AJSTR

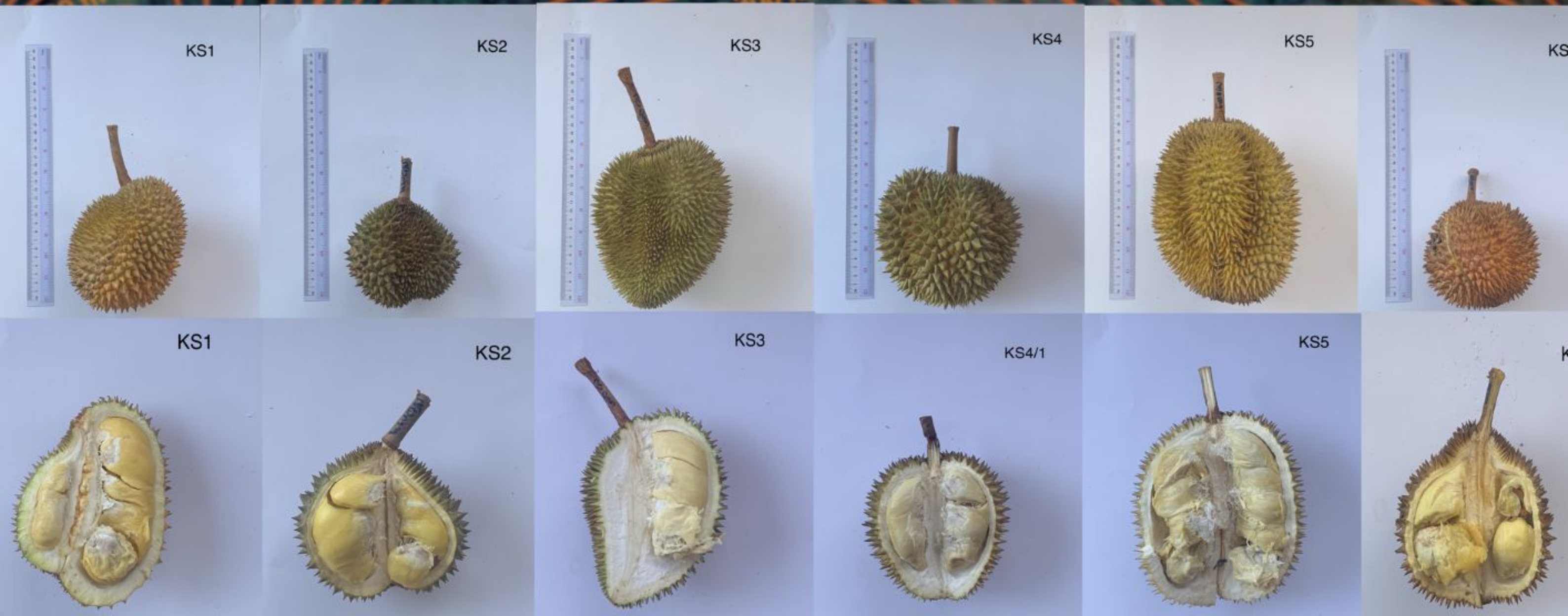
Vol.28 No.5
September - October 2025

ASEAN Journal
of Scientific and Technological Report (AJSTR)

Ecological Factors Governing the Persistence of Khlong-Saeng Durian: Implications for Conservation

Surachai Sang-Ngam et al.

ISSN 2773-8752 (online)



ASEAN
Journal of Scientific and Technological Reports

Online ISSN:2773-8752

ASEAN Journal of Scientific and Technological Reports (AJSTR)

Name ASEAN Journal of Scientific and Technological Reports (AJSTR)
Owner Thaksin University

Advisory Board Assoc. Prof. Dr. Nathapong Chitniratna
(President of Thaksin University, Thailand)
Assoc. Prof. Dr. Samak Kaewsuksaeng
(Vice President for Reserach and Innovation, Thaksin University, Thailand)
Assoc. Prof. Dr. Suttiporn Bunmak
(Vice President for Academic Affairs and Learning, Thaksin University, Thailand)
Assoc. Prof. Dr. Samak Kaewsuksaeng
(Acting Director of Reserach and Innovation, Thaksin University, Thailand)
Asst. Prof. Dr. Prasong Kessaratikoon
(Dean of the Graduate School, Thaksin University, Thailand)
Editor-in-Chief Assoc. Prof. Dr. Sompong O-Thong, Mahidol University, Thailand
Session Editors

1. Assoc. Prof. Dr. Jatuporn Kaew-On, Thaksin University, Thailand
2. Assoc. Prof. Dr. Samak Kaewsuksaeng, Thaksin University, Thailand
3. Assoc. Prof. Dr. Rattana Jariyaboon, Prince of Songkla University, Thailand
4. Asst. Prof. Dr. Noppamas Pukkhem, Thaksin University, Thailand
5. Asst. Prof. Dr. Komkrich Chokprasombat, Thaksin University, Thailand

Editorial Board Members

1. Prof. Dr. Hidenari Yasui, University of Kitakyushu, Japan
2. Prof. Dr. Jose Antonio Alvarez Bermejo, University of Almeria, Spain
3. Prof. Dr. Tjokorda Gde Tirta Nindhia, Udayana University in Bali, Indonesia
4. Prof. Dr. Tsuyoshi Imai, Yamaguchi University, Japan
5. Prof. Dr. Ullah Mazhar, The University of Agriculture, Peshawar, Pakistan
6. Prof. Dr. Win Win Myo, University of Information Technology, Myanmar
7. Prof. Dr. Yves Gagnon, University of Moncton, Canada
8. Assoc. Prof. Dr. Chen-Yeon Chu, Feng Chia University, Taiwan
9. Assoc. Prof. Dr. Gulam Murtaza, Government College University Lahore, Lahore, Pakistan
10. Assoc. Prof. Dr. Jompob Waewsak, Thaksin University, Thailand
11. Assoc. Prof. Dr. Khan Amir Sada, American University of Sharjah, Sarjah, United Arab Emirates.
12. Assoc. Prof. Dr. Sappasith Klomklao, Thaksin Univerrsity, Thailand
13. Asst. Prof. Dr. Dariusz Jakobczak, National University, Pakistan
14. Asst. Prof. Dr. Prawit Kongjan, Prince of Songkla University, Thailand
15. Asst. Prof. Dr. Shahrul Isamail, Universiti Malaysia Terengganu, Malaysia
16. Asst. Prof. Dr. Sureewan Sittijunda, Mahidol University, Thailand
17. Dr. Nasser Ahmed, Kyushu University, Fukuoka, Japan
18. Dr. Peer Mohamed Abdul, Universiti Kebangsaan Malaysia, Malaysia
19. Dr. Sriv Tharith, Royal University of Phnom Penh, Cambodia
20. Dr. Zairi Ismael Rizman, Universiti Teknologi MARA, Malaysia
21. Dr. Khwanchit Suwannoppharat, Thaksin University, Thailand

Staff: Journal Management Division

1. Miss Kanyanat Liadrak, Thaksin University, Thailand
2. Miss Ornkamon Kraiwong, Thaksin University, Thailand

Contact Us Institute of Research and Innovation, Thaksin University
222 M. 2 Ban-Prao sub-district, Pa-Pra-Yom district, Phatthalung province, Thailand
Tel. 0 7460 9600 # 7242 , E-mail: aseanjstr@tsu.ac.th

List of Contents

List of Contents

ASEAN
Journal of Scientific and Technological Reports
Online ISSN:2773-8752

ASEAN
Journal of Scientific and Technological Reports
Online ISSN:2773-8752

Research article

Exact Modeling and Solution of the Healthcare Facility Location-Allocation and Vehicle Routing Problem for Medication Delivery through Rural Primary Care Units

Simya Samohyusoh[1], Nikorn Sirivongpaisal[2], and Sirirat Suwatcharachaitiwong[3*]

[1] Faculty of Engineering, Prince of Songkla University, Songkhla, 90110, Thailand
[2] Faculty of Engineering, Prince of Songkla University, Songkhla, 90110, Thailand
[3] Faculty of Engineering, Prince of Songkla University, Songkhla, 90110, Thailand
* Correspondence: sirirat.su@psu.ac.th

Citation:
Samohyusoh, S.; Sirivongpaisal, N.; Suwatcharachaitiwong, S. Exact modeling and solution of the healthcare facility location-allocation and vehicle routing problem for medication delivery through rural primary care units. *ASEAN J. Sci. Tech. Report.* **2025**, *28*(5), e259125. https://doi.org/10.55164/ajstr.v28i5.e259125.

Article history:
Received: May 2, 2025
Revised: July 30, 2025
Accepted: August 16, 2025
Available online: August 30, 2025

Publisher's Note:
This article is published and distributed under the terms of Thaksin University.

Abstract: This study aims to analyze the optimal locations of Subdistrict Health Promoting Hospitals (SHPHs) for participating in the "Taking Medicine Nearby House" project by systematically allocating patients to appropriate SHPHs and formulating efficient delivery routes aimed at enhancing home based healthcare services for chronic disease patients particularly whom diagnosed with hypertension in Hat Yai District, Songkhla Province, with a focus on individuals facing mobility limitations. The research proposes a conceptual framework integrating spatial analysis and operations research techniques to improve community-level healthcare logistics. A two-stage solution was applied to solve the problem with the Location and Allocation Model and the Vehicle Routing Model. The first model selected the optimum location of SHPH and allocated patients to that location. The latter provided the medication distribution routing to each SHPS's patient. The exact method solved with LINGO satisfies the case study solution. The results indicate that, among nine candidate SHPHs, six were identified as optimal locations to participate in the service. Fifty-eight patients were effectively allocated to these facilities under capacity constraints, achieving a minimized total transportation and operational cost of 123,149.37 THB (1.000 THB = 0.02990 USD). The findings demonstrate that applying the Location Allocation Problem (LAP) and Vehicle Routing Problem (VRP) through exact solution methods can effectively support planning and logistics management in home-based healthcare systems and can provide potential for application in other regional contexts in the future.

Keywords: Location allocation problems; vehicle routing problems; exact method; home healthcare; medication delivery

1. Introduction

Currently, public hospitals in Thailand, particularly large- scale hospitals, are facing overcrowding due to the high volume of outpatient visits. In 2017, the public health service system under the Ministry of Public Health recorded up to 220 million outpatient visits annually. On average, regional hospitals served 3,152 outpatients per day, large general hospitals served 1,709 per day, and small general hospitals served 1,204 per day [1]. Although the number of outpatients in regional and general hospitals has shown a continuous downward trend between 2022 and 2024 for example, the average number of outpatients at regional hospitals decreased from 23,022 per day in 2022 to 15,490 in 2023, and it is projected to further decrease to 13,341 in 2024; while general hospitals saw a decline from 29,299 outpatients per day in 2022 to 20,303 in 2023, with an expected further drop to 18,709 in 2024 [2]. However, the volume of service users remains high, leading to persistent overcrowding in public hospitals. This overcrowding directly contributes to prolonged waiting times, delayed access to healthcare services, diminished quality of care, and increased patient dissatisfaction. The root causes of these issues primarily lie in the limitations of available resources and shortages of medical personnel. To address these challenges, strengthening the primary healthcare system, particularly through the family practice teams and community health service networks, has been recognized as a key strategy. This approach facilitates the redistribution of healthcare workloads away from hospitals, reduces the number of outpatient visits, and promotes a more continuous and accessible care system. It is especially beneficial for patients with chronic conditions, who can receive consistent and community-based care [3].

In 2017, the Ministry of Public Health announced a policy aimed at reducing hospital overcrowding by encouraging patients to receive medications outside of hospital settings through the "Taking Medicine Nearby House" project. This initiative allows patients to present prescriptions and receive medications at licensed community pharmacies, rather than having to wait for hospital-based medication delivery. The program primarily targets patients with chronic diseases who are free from complications and require continuous medication. This approach reflects the concept of healthcare supply chain management, which plays a crucial role in supporting home healthcare services, particularly for elderly patients or those unable to travel to obtain their medications. However, the implementation of this program has led to increased costs within the healthcare service system, particularly in logistics and pharmaceutical distribution. These costs encompass transportation, storage, and the delivery of medications, medical supplies, and equipment directly to patients' homes. Notably, transportation costs account for approximately 47.3% of total logistics expenditures in Thailand, representing a critical factor that must be effectively managed to control healthcare expenses and enhance the efficiency of the service system [4–6].

Hat Yai Hospital is a regional-level hospital providing healthcare services throughout Hat Yai District, Songkhla Province. According to data from the Ministry of Public Health in 2023, the hospital recorded a total of 1,117,720 service visits per year, with 70.45% originating from Hat Yai District, 14.93% from other districts within Songkhla Province, and 14.62% from other provinces. Although the hospital is officially equipped with approximately 700 beds, the actual utilization reached as high as 1,000 beds, resulting in a bed occupancy rate of 111.6%, which exceeds its service capacity [7]. Regarding outpatient services, the hospital recorded 235,280 outpatients in 2020, 223,200 in 2021, and 199,378 during the first half of 2022, averaging approximately 3,500 patients per day, significantly higher than its intended capacity of 2,500 patients per day. In 2020, a total of 1,805 patients participated in the "Taking Medicine Nearby House" project, ranking as the second-highest in the country. However, the implementation of the program in Hat Yai District continues to encounter challenges related to the geographical distribution of participating pharmacies. As there is no limitation on the number of pharmacies permitted to join the program, an imbalance in distribution has emerged. Urban areas such as Hat Yai, Kho Hong, Khlong Hae, and Khlong U-Taphao Subdistricts exhibit a high concentration of participating pharmacies. This oversaturation results in low patient volumes for some pharmacies, prompting their withdrawal from the program. Conversely, rural areas—including Khue Tao, Nam Noi, Tha Kham, Thung Yai, Chalung, Thung Tam Sao, Khuan Lang, Ban Phru, and Phatong—experience limited participation, compelling residents to travel considerable distances to access services. This situation contributes to increased transportation costs for patients.

In response to the challenges above, a medication delivery model involving Subdistrict Health Promoting Hospitals (SHPHs) has been proposed as part of the "Taking Medicine Nearby House" project. Under this model, SHPHs serve as central hubs for distributing medications to patients within their designated service areas. Village Health Volunteers (VHVs) are assigned to deliver medications, provide basic healthcare, and continuously monitor patients' health conditions to enhance the efficiency and coverage of healthcare services. Based on the findings from previous evaluations of the project, it has been observed that in urban areas, medications can be effectively distributed through community pharmacies due to their widespread availability and relatively high density. However, in peripheral or suburban areas, the number of participating pharmacies is considerably limited, thereby restricting patients' access to medication services in these regions, as illustrated in Figure 1. Consequently, this study focuses on exploring and proposing suitable location management strategies for primary healthcare units, namely Subdistrict Health Promoting Hospitals (SHPHs), to serve as distribution points in the "Taking Medicine Nearby House" project. The research particularly targets rural areas within Hat Yai District, Songkhla Province, encompassing nine subdistricts: Khue Tao, Nam Noi, Tha Kham, Thung Yai, Chalung, Thung Tam Sao, Khuan Lang, Ban Phru, and Phatong. A medication distribution model via subdistrict-level SHPHs is developed in parallel with a delivery route planning strategy to ensure that medications are delivered efficiently and comprehensively to patients, especially since most of the program's beneficiaries are elderly individuals who are unable to travel to receive their medications.

The primary objective of this study is to enhance healthcare accessibility by addressing LAP and VRP. This is achieved through the development and analysis of a mathematical model that employs exact methods to determine the most cost-effective solution, thereby resulting in an efficient medication distribution system, reduced operational costs, and improved quality of public health services at the community level.

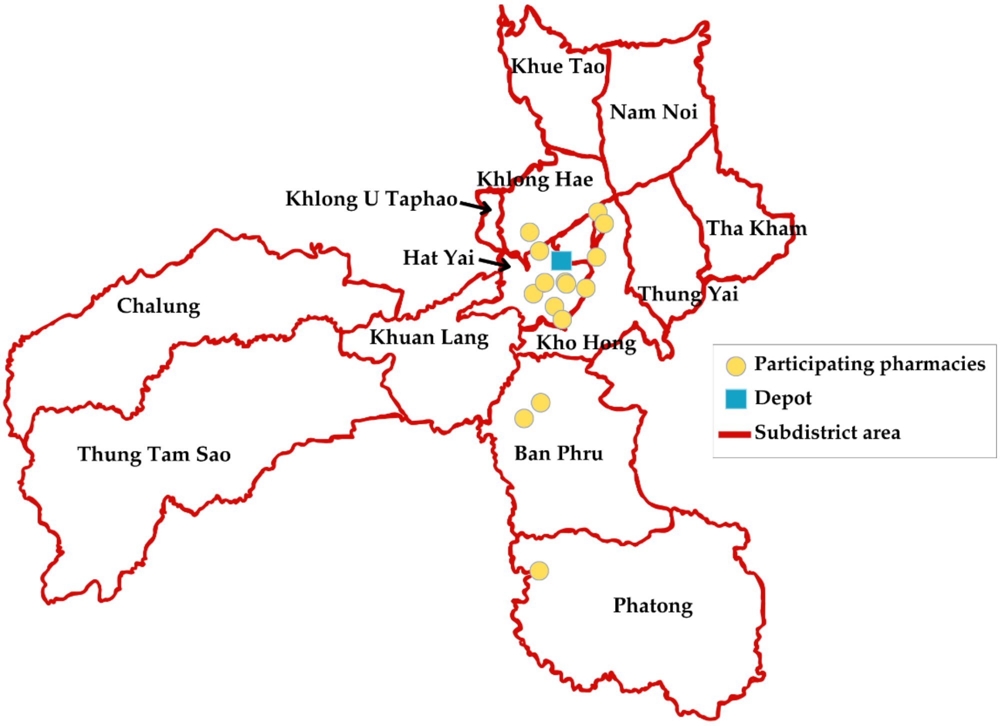


Figure 1. The Distribution of Pharmacies in the "Taking Medicine Nearby House" project, Hat Yai District, Songkhla Province.

This research proposes a novel medicine distribution model tailored to the Thai primary healthcare system. In this model, Subdistrict Health Promoting Hospitals (SHPHs) function as local distribution hubs, while Village Health Volunteers (VHVs)—who already provide home-based care—are engaged to deliver medicines directly to patients. This integration not only improves the efficiency and coverage of home healthcare services but also strengthens supply chain resilience by utilizing established community-based health infrastructures to maintain service continuity during disruptions. Furthermore, the study addresses LAP and VRP associated with medicine distribution through SHPHs. Mathematical modeling techniques are applied, and an exact method is employed to determine optimal facility locations and delivery routes, ensuring cost-effective and reliable healthcare logistics in a community context.

2. Materials and methods.

2.1 Integrated community and home-based healthcare systems

In recent years, home-based healthcare systems have garnered significant attention, especially for patients suffering from chronic illnesses and those with limited mobility. These systems are increasingly viewed as essential in enhancing healthcare accessibility, alleviating pressure on primary hospitals, and ultimately improving patients' quality of life. A critical enabler of this shift is the development of efficient healthcare logistics, particularly in terms of optimizing the delivery of medical services and medications directly to patients' homes. Several studies have explored various computational approaches to improve the efficiency and effectiveness of home healthcare logistics. For example, Phonin et al. [8] addressed VRP with Time Windows (VRPTW) in home healthcare, proposing a hybrid Tabu Search and Greedy Algorithm to minimize vehicle service time. Their method reduced the number of required vehicles by up to 59%, demonstrating strong potential for operational efficiency. Similarly, Atta et al. [9] explored the broader Home Health Care Routing and Scheduling Problem (HHCRSP), aiming to reduce costs, enhance service quality, and maintain care continuity. The evaluation of various algorithmic approaches highlighted the critical role of HHCRSP in delivering cost-effective, patient-centered home healthcare. Suwatcharachaitiwong et al. [4] studied systems that integrate both direct home delivery and patient self-collection from nearby pickup points, such as lockers or convenience stores. The study aimed to minimize total operational costs using a Genetic Algorithm. Simulation results indicated that the algorithm was effective in optimizing delivery logistics, supporting flexible models of medication distribution tailored to patients' needs and locations. In Kordi et al. [10], a multi-objective mixed-integer mathematical model was developed to optimize home healthcare service planning. For small-scale problems, they applied the epsilon-constraint method using CPLEX software. At the same time, for larger, real-world scenarios, they introduced a Multi-Objective Variable Neighborhood Search (MOVNS) algorithm to explore solution spaces effectively. The approach emphasized the need for adaptable and scalable algorithms capable of balancing multiple, often competing planning objectives. Addressing the challenges of uncertainty in home healthcare operations, Abdul Nasir and Kuo [11] proposed a chance-constrained optimization model that considered multi-depot and multi-period factors alongside precedence constraints. They developed a specialized three-stage solution methodology combined with stochastic simulation to enable robust and flexible planning under uncertain conditions. Expanding on technological innovations in healthcare logistics, Euchi et al. [12] explored the use of telemedicine and artificial intelligence (AI) to support home healthcare delivery, particularly for elderly patients. The research introduced AI-driven routing and scheduling techniques based on autonomous learning and search algorithms, enhancing decision-making in distributed healthcare environments and optimizing patient visit allocations.

The advancement of community-based healthcare models has played a vital role in strengthening healthcare systems by leveraging local resources and positioning sub-district health centers as central service hubs. Effective implementation requires coordinated planning of facility locations, patient distribution, and service routing. Addressing these needs, Ransikarbum et al. [13] proposed a dual-phase approach combining K-means clustering for hospital siting and a time-constrained CVRP model for pharmaceutical distribution. The method demonstrated practical effectiveness in both medium-term facility planning and short-term logistics. Salami et al. [14] developed a Multi-Period Capacitated Maximal-Covering Location-Allocation model that accounts for the dynamic nature of healthcare demand over time. By incorporating capacity constraints across multiple planning periods, their model ensures continuity of care and optimal allocation of limited healthcare resources. The approach utilizes Mixed-Integer Linear Programming (MILP), a method well-suited for solving complex, NP-hard problems in urban healthcare service design. In the realm of advanced delivery technologies, Shi et al. [15] introduced a bi-objective mixed-integer programming model for multi-trip drone routing with simultaneous pickup and delivery, demonstrating improved delivery speed, safety, and efficiency through a modified NSGA-II algorithm. Similarly, Zaid et al. [16] addressed the Home Healthcare VRP (HHCVRP) with a focus on social sustainability, employing a hybrid metaheuristic combining Ant Colony Optimization and Non-Dominated Sorting to enhance service quality and workload management in smart city contexts. The studies highlight the value of advanced optimization and emerging technologies in developing efficient, adaptable, and sustainable community-based healthcare systems.

2.2 Review of literature on location-allocation and vehicle routing problems in healthcare logistics

This section reviews key studies that apply location-allocation and vehicle routing models to optimize healthcare logistics, demonstrating their effectiveness in improving service coverage, resource allocation, and operational efficiency for example, the case study by Phutthaphooltrakun & Raothanachonlakul [17] applied the Maximal Covering Location Problem (MCLP) model to plan the locations of blood storage and distribution centers in Thailand. The model improved service coverage to 96.66%, demonstrating the potential of LAP in designing public health infrastructure. Tapabut et al. [18] studied the relocation of emergency parking stations by incorporating the elderly population as a key factor in the target area. Using LAP for analysis, the study aimed to allocate service points that better meet the specific needs of this demographic group. In the context of pharmacy management, Poomisirisawat et al. [19] proposed a Location-Inventory Problem (LIP) model that integrates the selection of pharmacy locations with optimal inventory level planning. An exact method was used to enhance the efficiency of medication distribution, reduce operational costs, and increase medication accessibility for patients in the area. Similarly, Pan et al. [20] applied the LAP model to analyze the coverage of tertiary hospitals in major cities across China, with a focus on equitable distribution and improving access to healthcare services. In the field of emergency management, Alghanmi et al. [21] examined the allocation of Points of Dispensing (PODs) for distributing medications and medical supplies during emergencies. Their approach prioritized population risk levels to ensure a rapid and effective response. Zhuo et al. [22] proposed a Multi-Objective LAP model for planning new community hospitals in Wuhou District, Chengdu, aiming to align hospital capacity with population needs and to enhance the quality and long-term efficiency of the healthcare system. Meanwhile, Murad et al. [23] analyzed the impact of location on healthcare accessibility by introducing a P-Median model for determining the locations of health service centers in Jeddah, Saudi Arabia. The model was designed to ensure that residents could access services within a 15-minute walking distance, underscoring the importance of spatial planning in promoting equity in access to basic healthcare services.

Shi et al. [24] investigated routing and scheduling for medicine delivery in home healthcare systems under demand uncertainty and developed a hybrid genetic algorithm to enhance routing decision efficiency. Similarly, Al Theeb et al. [25] proposed a multi-objective mixed-integer linear programming model that integrates a two-echelon VRP for vaccine supply chains in developing countries, aiming to reduce the number of undelivered vaccines. Euchi et al. [26] developed a VRP approach incorporating time windows and synchronized visits for home healthcare systems using artificial intelligence (AI) techniques to optimize scheduling and reduce operational costs. Durak et al. [27] introduced a mathematical model for routing and scheduling nurses' visits by considering ergonomic factors through a Fuzzy Inference System to evaluate workload suitability in the context of home healthcare services. Furthermore, HadjTaieb et al. [28] proposed a shortest-path routing approach for home healthcare providers using alternative energy vehicles under the Green VRP with Time Windows concept, which emphasizes reducing environmental impact and promoting the sustainability of out-of-hospital medical service systems.

Exact methods in healthcare logistics primarily utilize MILP and professional-grade solvers (such as CPLEX or Gurobi), employing techniques such as branch-and-bound, branch-price-and-cut, set-partitioning, or commodity-flow formulations to examine all feasible routes and determine the global optimum for VRP or HHCRSP problems. However, computational complexity increases rapidly as the number of patients or service networks expands, limiting optimal problem-solving to small-scale instances (typically not exceeding 50 patient nodes on average) due to the exponential growth in complexity. Research by van Montfort et al. [29] and recent work by Zhang & Zhang [30] indicate that optimal solutions can be obtained from deterministic models for small-scale instances, but for large-scale problems or those involving uncertainty, metaheuristics such as VNS provide superior results in terms of efficiency and scalability. Studies on exact solution methods, which can be applied to various logistics problems, have been extensively explored. Paradiso et al. [31] proposed an exact solution framework for the multi-trip VRP with time window constraints, focusing on urban logistics systems and last-mile delivery. Their model involves many variables and constraints. Salavati-

Khoshghalb et al. [32] applied an exact method to solve VRP under demand uncertainty. They developed an Integer L-shaped algorithm within a Branch-and-Cut framework and incorporated a return travel cost approximation technique to improve cost reduction efficiency. Zetina et al. [33] introduced an exact algorithm for the non-convex quadratic facility location problem, capable of handling instances with up to 1,000 nodes while accounting for the interaction costs between service centers and customers. Balti & Jemai [34] conducted a comprehensive study reviewing various approaches to the VRP and proposed a hybrid methodology that integrates exact optimization methods with Home Health Care (HHC) systems and Intelligent Transportation Systems (ITS) using Internet of Things (IoT) technologies. This approach enhances routing reliability for home healthcare service vehicles by accounting for traffic density and enabling real-time route adjustments to avoid congestion. Similarly, Zhang & Zhang [30] addressed the Vehicle Routing and Appointment Scheduling Problem (VRASP), aiming to optimize caregiver routing and appointment scheduling to reduce operational costs and improve service quality. The proposed solution, a customized Variable Neighborhood Search (VNS) algorithm combining regret-based insertion and Tabu Search, serves as a practical decision-support tool for HHC providers operating under uncertain conditions. In a related study, Linfati et al. [35] developed a two-phase heuristic algorithm for scheduling and routing in pharmaceutical delivery to highly dependent patients. The model, designed for daily visit scheduling, uses a flexible mathematical framework applicable to various scenarios. Hybrid metaheuristics—specifically, Simulated Annealing combined with Record-to-Record Travel—are employed to refine initial solutions, enhancing overall delivery efficiency. Van Montfort et al. [29] focused on the development of routing and scheduling plans for home-based caregiving services. The study introduced two MILP models—the Miller-Tucker-Zemlin (MTZ) and time-indexed formulations—demonstrating their effectiveness in reducing required staff, lowering service costs, and improving caregiver-task alignment. Notably, the research highlighted that the advantages of task division are not only dependent on the planning objective but also yield benefits when minimizing travel time.

The growing complexity of home healthcare logistics has driven research toward advanced optimization models that integrate routing and scheduling. These studies predominantly employ hybrid methodologies combining exact algorithms, heuristics, and real-time data from ITS and IoT technologies. Such approaches enhance service delivery by addressing dynamic factors including traffic, patient-specific needs, and caregiver capabilities. Moreover, these models extend beyond conventional routing problems by incorporating appointment scheduling, task allocation, and community-based frameworks. Collectively, they underscore the vital role of hybrid optimization and real-time data integration in improving the efficiency, flexibility, and quality of home healthcare logistics. Table 1 summarizes a comparative analysis of recent key studies on Location-Allocation and Vehicle Routing Problems and their extensions within home healthcare. It highlights variations in problem scope, target populations, planning levels, methodologies, and innovations, while indicating the use of real data and community health resources such as Subdistrict Health Promoting Hospitals (SHPHs) and Village Health Volunteers (VHVs).

Table 1. Comparative Overview of Studies on Location-Allocation and Vehicle Routing Problems in Home Healthcare

Researcher / Year	Problem Scope	Target Group	Planning Level	Problem Addressed	Methodology	Real Data	Uses SHPH /VHV	Key Innovation
van Montfort et al. (2024)	Integrating task-splitting into HHC routing & scheduling	Patients receiving HHC	Operational	VRPTW	MILP + Heuristics	✗	✗	Task splitting & managing temporal dependencies
Phonin et al. (2025)	Minimizing total completion time in HHC services	Elderly patients	Operational	VRPTW	Tabu Search with Greedy list	✗	✗	Fast metaheuristic for route efficiency
Atta et al. (2025)	Efficient routing & caregiver allocation	HHC patients	Operational	HHCRSP	Exact + Heuristics/ Metaheuristics	✗	✗	Integrating caregiver allocation in VRP
Kordi et al. (2023)	Multi-objective HHC optimization (cost, CO_2, workload, quality)	Patients needing HHC	Operational	Multi-objective VRP	ε-constraint + MOVNS	✗	✗	Balancing multiple real-world healthcare objectives
Abdul Nasir et al. (2024)	Mobile health facility placement & uncertain supply routing	Priority HHC patients	Strategic & Operational	Multi-depot VRP	Chance-constrained + Simulation	✓	✗	Uncertainty & disruption-focused HHC model
Zaid et al. (2024)	Real-time VRP for smart city HHC	Urban patients	Operational	HHCVRP	ACO + NSGA	✗	✗	Sensor data integration & social sustainability
Ransikarbum et al. (2024)	Hospital location & medicine distribution	Drug retailers & patients	Midstream & Last-mile	CVRP + Facility Location	K-means + CVRP + GIS	✓	✗	GIS-aided model with clustering & routing

Table 1. Comparative Overview of Studies on Location-Allocation and Vehicle Routing Problems in Home Healthcare (Continues)

Researcher / Year	Problem Scope	Target Group	Planning Level	Problem Addressed	Methodology	Real Data	Uses SHPH /VHV	Key Innovation
Shi et al. (2022)	Drone-based medicine delivery during emergencies	Communities during crises	Emergency Logistics	Drone Location Routing Problem	Bi-objective MIP + NSGA-II	✗	✗	Multi-trip drone model with pick-up & delivery
Euchi et al. (2020)	Routing & scheduling for vulnerable populations	Elderly, mobility-limited	Operational	HHC Routing & Scheduling	AI-based Learning & Search	✗	✗	Pure AI-driven solution without mathematical models
Balti & Jemai (2025)	VRP with real-time traffic adaptation	HHC patients	Operational	Traffic-aware VRP	Exact + ITS	✗	✗	Integration of Intelligent Transport Systems (ITS)
Schneider et al. (2019)	Overview of VRP research	General logistics sector	N/A	VRP overview	Editorial/Review	✗	✗	Heuristic and exact methods in the VRP landscape
Zhang & Zhang (2025)	Routing & appointment scheduling under uncertainty	Aging population	Operational	VRASP	Stochastic + VNS (Hybrid)	✗	✗	Combining stochastic modeling with regret-based heuristics
Linfati et al. (2018)	Scheduling & delivery routing for medication	Highly dependent patients	Operational	Medication Delivery VRP	2-Phase Heuristic (SA + RTR)	✓	✗	Cluster-based daily routing for medicine delivery
This Research	**Optimal SHPH location + routing for hypertension patients**	**Chronic disease patients with limited mobility**	**Strategic & Operational**	**LAP + VRP (medication delivery model)**	**Exact (LINGO) + GIS**	✓	✓	**Novel model integrating LAP + VRP with a real community-based health system**

* VHV (Village Health Volunteers)

* SHPH (Subdistrict Health Promoting Hospitals)

2.3 Application of healthcare logistics in the "taking medicine nearby house" project

This study presents an approach for applying the concepts of the Location Problem and the Routing Problem to enhance service efficiency under the "Taking Medicine Nearby House" project, with a primary focus on the actual residential locations of patients. The proposed approach aims to improve medication accessibility, reduce the burden of travel to main hospitals, and strengthen the role of Subdistrict Health Promoting Hospitals (SHPHs) as key health service distribution points within communities. A significant challenge in the program implementation lies in the SHPHs' site selection, where it is appropriate in both geographical location and service capacity to function effectively as service units under the program. Additionally, the optimal design of medication distribution routes, utilizing locally available resources such as VHVs, presents further complexity. Observations from the program's implementation in recent years indicate that the government has engaged private-sector pharmaceutical suppliers to be responsible for the procurement and distribution of medications to participating pharmacies, where patients can collect their prescribed medications directly. This service model is well-suited to urban contexts, as patients' residences are typically located near participating pharmacies. Conversely, in rural or suburban areas, the number of participating pharmacies is limited, resulting in patients having to travel long distances to access medication services. This has become a significant barrier to achieving equitable access to healthcare services.

To enhance access to healthcare services for people living in rural areas, the researcher proposes the concept of utilizing SHPHs as the primary local units for medication distribution, in place of pharmacies. Under this approach, suppliers would be responsible for distributing medications from central facilities to selected SHPHs. VHVs would then be assigned to deliver the medications directly to patients' homes, while also providing basic health check-up services during the same visit. This approach is expected to reduce the travel burden on patients, increase convenience in receiving services, and promote more comprehensive and sustainable access to healthcare in rural areas.

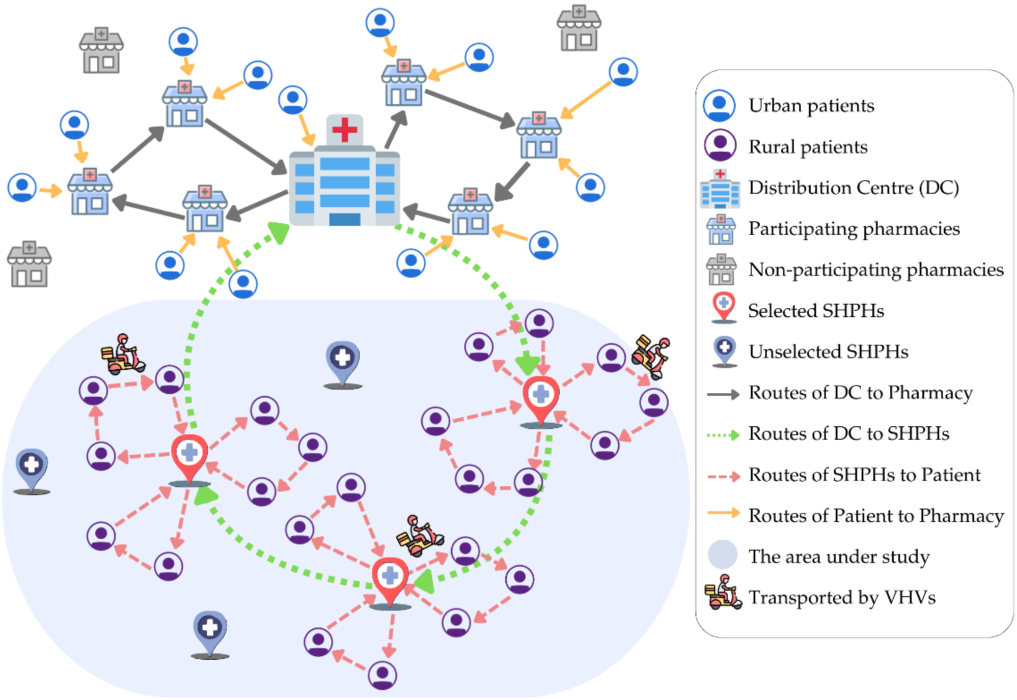


Figure 2. Conceptual Framework of the Research: Location-Allocation Problem and Vehicle Routing Problem.

This research problem is divided into two main levels: (1) the selection of appropriate SHPHs to participate in the program, along with the allocation of patient groups based on their respective catchment areas; and (2) the planning of medication delivery routes and individual patient monitoring using VHVs as the primary service providers. The solving process is conducted using an exact method, implemented through Lingo software to solve the mathematical model formulated as a logistics-based optimization problem.

The study area covers nine subdistricts located in the peri-urban zones of Hat Yai District, Songkhla Province: Khu Tao, Nam Noi, Tha Kham, Thung Yai, Chalung, Thung Tam Sao, Khuan Lang, Ban Phru, and Phatong. One SHPH was designated as a representative unit for each subdistrict. The target group comprises patients diagnosed with hypertension whose residential locations are fully identifiable (100%). These patients have a prevalence rate of 60–70% and exhibit mild symptoms, making them suitable for home-based monitoring by VHVs. Data from the nine subdistricts revealed a total of 58 patients. The primary objectives of the model are to identify appropriate SHPHs, allocate patients to the designated service units, and determine the most efficient medication delivery routes to ensure comprehensive home-based healthcare access while minimizing supply chain costs. The model operates under the following assumptions:

- **Equal Participation Cost:** Each primary healthcare unit (SHPH) incurs an equal fixed cost for participating in the program.
- **Uniform Patient Demand:** For simplification purposes, it is assumed that all patients have an equal level of medication demand, unless otherwise specified in extended scenarios. Although Equation (4) allows for varying service demand per patient, this assumption is adopted in the base case to reduce model complexity.
- **Daily Service Operation:** All delivery and healthcare service operations conducted by village health volunteers (VHVs) are assumed to be completed daily within a single planning period.
- **Motorcycle Accessibility:** Each VHV is assumed to have access to at least one motorcycle for use in delivering medication and providing home healthcare services.

2.4 Mathematical modeling of the location-allocation problem for healthcare facility planning

Strategic spatial planning plays a critical role in healthcare logistics, particularly in ensuring equitable and efficient access to services across diverse populations. Among the most established approaches for this purpose is LAP, a mathematical framework used to determine the optimal selection of service facility locations from a predefined set of alternatives. The objective is to maximize service coverage and operational efficiency while adhering to constraints such as distance, accessibility, and facility capacity. LAP models have been extensively applied in logistics, public health planning, and emergency response systems. To illustrate this concept, Figure 3 presents the fundamental structure of the Location-Allocation Problem, which underpins the spatial design of service networks to effectively and equitably meet population needs.

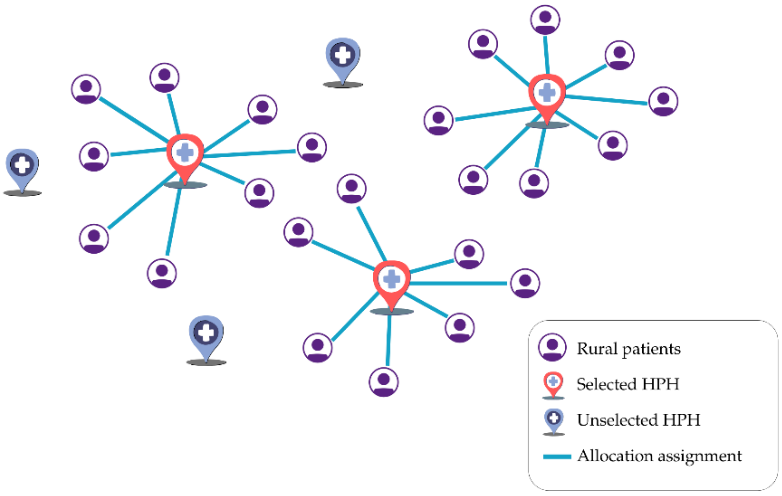


Figure 3. Location-Allocation Problem

The mathematical model representing the problem of selecting SHPHs to participate in the "Taking Medicine Nearby House" project was adapted from the model initially proposed by Perl & Daskin [37], as follows:

Indices

i	Set of potential locations of SHPHs eligible to participate in the "Taking Medicine Nearby House" project
j	Set of locations of patients with hypertension

Parameters

c_{ij}	Transportation cost from SHPH i to hypertensive patient j
f_i	Fixed cost or participation cost of SHPH i in the project
C_i	Capacity of SHPH i to serve patients
D_j	Demand for medication by hypertensive patient j

Decision Variables

$x_i \in \{0,1\}$	A binary variable indicating whether location i is selected as a participating SHPH (1 = selected, 0 = otherwise)
$y_{ij} \in \{0,1\}$	A binary variable indicating whether demand point j is assigned to SHPH i (1 = assigned, 0 = otherwise)

Objective Function

$$\text{Min} \sum_{i\in I} f_i x_i + \sum_{i\in I}\sum_{j\in J} c_{ij} y_{ij} \tag{1}$$

Constraints

$$\sum_{i\in I} y_{ij} = 1, \quad \forall j \in J \tag{2}$$

$$y_{ij} \leq x_i, \quad \forall i \in I, j \in J \tag{3}$$

$$\sum_{j\in J} D_j y_{ij} \leq C_i x_i, \quad \forall i \in I \tag{4}$$

$$x_i \in \{0,1\}, \; y_{ij} \in \{0,1\}, \quad \forall i \in I, j \in J \tag{5}$$

The objective function in Equation (1) aims to minimize the total system cost, which consists of the participation cost of the primary healthcare facilities (SHPHs) and the transportation cost for delivering medication to patients. Equation (2) ensures that each patient is assigned to only one SHPH. Equation (3) states that a patient can only be assigned to an SHPH that has been selected to participate in the project. Equation (4) represents the capacity constraint of each SHPH in terms of the total amount of healthcare service demand it can accommodate, where each patient may require a different level of service based on individual needs. Finally, Equation (5) defines the decision variables as binary, taking values of either 0 or 1.

2.5 Mathematical modeling of the vehicle routing problem for home-based healthcare services

Efficient route planning is a fundamental component of home-based healthcare logistics, particularly when distributing medical supplies to patients in decentralized communities. VRP provides a mathematical framework for optimizing delivery routes to minimize travel costs while meeting service requirements. In the context of home healthcare in Thailand, the distribution of medicines by SHPHs exemplifies this application. Figure 4 illustrates the core structure of the VRP, emphasizing its role in supporting cost-effective and reliable healthcare delivery beyond conventional clinical settings.

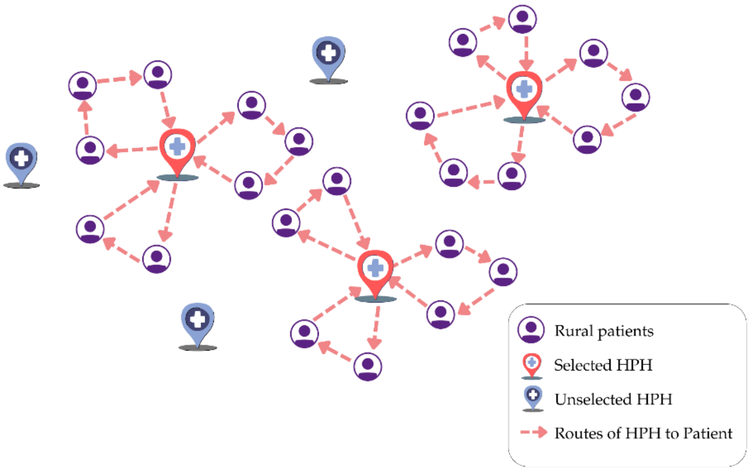


Figure 4. Vehicle Routing Problem

The development of a mathematical model focused on route planning for delivering medication to patients with hypertension under the "Taking Medicine Nearby House" project was carried out by adapting the mathematical model proposed by Kara, Laporte, & Bektas [38].

Indices

i	Set of locations SHPHs eligible to participate in the "Taking Medicine Nearby House" project
j	Set of locations of patients with hypertension

Parameters

c_{ij}	Transportation cost from SHPH *i* to hypertension patient *j*
Q	Vehicle capacity
q_j	Quantity of medicine required by the hypertension patient *j*

Decision Variables

$x_{ij} \in \{0,1\}$	A binary variable indicating whether there is a route from SHPH *i* to patient *j* (1 = route exists, 0 = otherwise)
u, m	Auxiliary variables

Objective Function

$$\text{Min} \sum_{i \neq j} c_{ij}\, x_{ij} \tag{6}$$

Constraints

$$\sum_{j=2}^{n} x_{1j} = m, \tag{7}$$

$$\sum_{i=2}^{n} x_{i1} = m, \tag{8}$$

$$\sum_{j=1,\, j \neq i}^{n} .\, x_{ij} = 1 \qquad (i=2,\ldots,n), \tag{9}$$

$$\sum_{i=1,\, i \neq j}^{n} x_{ij} = 1 \qquad (j=2,\ldots,n), \tag{10}$$

$$u_i - u_j + Q\,x_{ij} \leq Q - q_j \qquad (i,j=2,\ldots,n;\ i \neq j), \tag{11}$$

$$q_i \leq u_i \leq Q \qquad (i=2,\ldots,n), \tag{12}$$

$$x_{ij} = 0 \text{ or } 1 \qquad (i,j= 1,\ldots,n;\ i \neq j), \tag{13}$$

$$m \geq 1 \text{ and integer} \tag{14}$$

The objective function presented in Equation (6) aims to minimize the total transportation distance. Equations (7) and (8) ensure that each route must start and end at a primary healthcare facility (SHPH). Equations (9) and (10) impose the constraint that each patient must be visited exactly once. Equations (11) and (12) are used to control the total demand on each route, ensuring it does not exceed the vehicle capacity, while also preventing the occurrence of undesired sub-tours. Equation (13) defines the decision variables as binary, taking values of either 0 or 1. Lastly, Equation (14) specifies that the auxiliary variables must be strictly positive real numbers.

The integration of LAP and VRP using exact solution methods represents an appropriate approach for developing an efficient service model within the context of the "Taking Medicine Nearby House" project. This approach can effectively support policy-level decision-making in the equitable and sustainable allocation of healthcare resources to communities and patients.

3. Results and discussion

This research is based on a case study of the "Taking Medicine Nearby House" project conducted in Hat Yai District, Songkhla Province, with a primary focus on patients with chronic conditions and elderly individuals who are unable to travel to collect their medications independently, particularly those diagnosed with hypertension. The study encompasses several key components: the selection of primary healthcare units (SHPHs), the allocation of patients, the planning of medication delivery routes, and the provision of home healthcare services. The proposed model operates under specific constraints, namely that each SHPH may serve a maximum of ten patients, and the initial participation cost for each unit is fixed at 20,000 Baht. Each patient is assumed to require one package of medication. The principal objective of this research is to minimize the total cost of the healthcare service system, with a particular emphasis on transportation expenses. The analysis utilizes actual distance data between the residences of patients and the respective SHPHs, as well as the distances among these healthcare units within the rural areas of Hat Yai. Patient coordinates are derived from official house registration documents, and the distances are computed using the actual road network, processed through ArcGIS software. The results are presented in a distance matrix, as shown in Table 2.

Table 2. Distances between SHPHs and hypertensive patients (kilometers).

Coordinates	P1	P2	P3	P4	P5	P6	...	C57	C58
P1	0.00	30.53	10.32	11.38	12.61	5.64	...	12.42	15.80
P2	30.53	0.00	33.00	34.10	37.10	34.76	...	13.95	14.90
P3	10.32	33.00	0.00	0. 92	4.25	7.32	...	9.47	22.47
P4	11.38	34.10	0. 92	0.00	3.15	8.21	...	10.48	14.77
P5	12.61	37.10	4.25	3.15	0.00	7.76	...	13.72	18.97
P6	5.64	34.76	7.32	8.21	7.76	0.00	...	14.62	18.22
...	...	...	...	...	...	...	...	...	...
C57	12.42	13.95	9.47	10.48	13.72	14.62	...	0.00	5.72
C58	15.80	14.90	22.47	14.77	18.97	18.22	...	5.72	0.00

*P = Point SHPH, C = Hypertensive patients.

After collecting the input data, the mathematical model was transformed into a set of executable commands for processing with the LINGO software. This was done to determine the optimal locations of SHPHs, allocate patients appropriately, and identify delivery routes that comply with the specified constraints. Computational experiments were conducted using LINGO version 19 on a workstation equipped with an AMD Ryzen 5 processor and 32 GB of RAM.

3.1 Selection of SHPH locations participating in the project and patient allocation

To ensure the efficiency of Home Health Care services under the "Taking Medicine Nearby House" project, it is essential to consider both the locations of SHPHs and the spatial distribution of patients. The problem was solved to identify suitable SHPHs and allocate patients to each facility. The program completed the computation in just 2 seconds, and the results are presented in Table 3.

Table 3. Status and Patient Allocation.

Code	Location of SHPHs	Status	Allocated Patients
P1	Khu Tao	Participated	C22, C23, C25, C27, C28, C40, C41, C42, C43, C44
P2	Chalung	Participated	C29, C30, C31, C32, C37, C38, C39, C55,
P3	Thung Yai	Not Participated	-
P4	Thung Tam Sao	Participated	C10, C13, C24, C26, C33, C34, C35, C36, C45, C57
P5	Tha Kham	Not Participated	-
P6	Nam Noi	Not Participated	-
P7	Ban Phru	Participated	C11, C12, C14, C16, C17, C18, C19, C20, C21, C51
P8	Phatong	Participated	C1, C2, C3, C4, C5, C6, C7, C8, C9, C15
P9	Khuan Lang	Participated	C46, C47, C48, C49, C50, C52, C53, C54, C56, C58
			Total Cost 122,196.50 Bath

*P = Point SHPH, C = Hypertensive patients.

From testing the total number of SHPHs across 9 subdistricts and 58 hypertensive patients participating in the project, it was found that 6 subdistricts were suitable for participation. The allocation of patients in each subdistrict is as follows: Khu Tao Subdistrict had 10 patients, Chalung Subdistrict had 8 patients, Thung Tam Sao Subdistrict had 10 patients, Ban Phru Subdistrict had 10 patients, Phatong Subdistrict had 10 patients, and Khuan Lang Subdistrict had 10 patients. The total cost was 122,196.50 THB.

3.2 Transportation routing

Following the location selection of the SHPHs, transportation routes for medicine delivery by VHVs were systematically designed. Transportation costs were calculated based on actual travel distances, using a standardized rate of 36.29 THB per kilometer. This rate includes estimated vehicle-related expenses such as depreciation, labor, fuel, and maintenance costs. Specifically, the depreciation cost was based on a standard value of 60,000 THB for a 110cc motorcycle, assuming a useful life of five years. Using the geographic coordinates of the selected SHPHs and the assigned patient residences, described in Section 3.1, the optimization model was employed to generate cost-effective travel routes for each service area. These routes were designed to minimize the total transportation cost while ensuring complete delivery coverage. The results of route optimization are summarized in Table 4.

Table 4. Transportation Routing Results.

Code	Location of SHPHs	Transportation Order	Computation Time (min: sec)	Cost (THB)
P1	Khu Tao	P1-C28-C23-C22-C25-C27-C41-C42-C44-C43-C40-P1	00.02	115.28
P2	Chalung	P2-C55-C29-C39-C32-C31-C30-C38-C37-P2	00.05	310.03
P4	Thung Tam Sao	P4-C10-C13-C57-C45-C24-C26-C36-C35-C34-C33-P4	00.03	222.95
P7	Ban Phru	P7-C17-C20-C12-C21-C16-C18-C11-C19-C14-C51-P7	00.01	105.90
P8	Phatong	P8-C15-C9-C6-C1-C8-C5-C7-C3-C4-C2-P8	02.39	128.94
P9	Khuan Lang	P9-C50-C46-C48-C49-C53-C54-C58-C56-C47-C52-P9	00.01	69.76
			Total Cost	**952.87**

*P = Point SHPH, C = Hypertensive patients.

Table 4 illustrates the results of medicine distribution routing and transportation costs for the six SHPHs, namely in the subdistricts of Khu Tao, Chalung, Thung Tam Sao, Ban Phru, Phatong, and Kuan Lang, which provide home healthcare services, with a total transportation cost of 952.87 THB. Therefore, when including the costs of selecting the participating SHPH locations, patient allocation, and transportation route planning, the total cost amounts to 123,149.37 THB.

3.3 Sensitivity testing of model parameters

To evaluate the robustness of the model, a sensitivity analysis was conducted by varying key parameters, consisting of the fixed cost of facility participation (f_i), the service capacity of Sub-district Health Promoting Hospitals (SHPHs) (C_i), and patient demand (D_j). For the fixed cost (f_i), the base value was set at 20,000 THB for all facilities. Three scenarios were examined: (1) a uniform reduction to 10,000 THB, (2) a uniform increase to 30,000 THB, and (3) random values assigned to each facility within the range of 10,000 to 30,000 THB. This analysis aimed to assess whether increased investment in establishing service points would remain cost-effective and how such changes would influence the selection of SHPHs.

Regarding service capacity (C_i), the current setup allows each of the 9 facilities to accommodate a maximum of 10 patients, serving a total of 58 hypertensive patients. Three scenarios were evaluated: (1) reducing the maximum capacity to 7 patients per facility, (2) increasing it to 20 patients, and (3) randomly assigning capacities between 7 and 20 patients. These scenarios were analyzed to observe the impact of capacity variation on patient allocation and to determine whether such changes would affect the feasibility of serving all patients. For patient demand (D_j), while the baseline scenario assumes a constant demand of one unit of medication per person (i.e., one sachet of medication), two alternative cases are considered: (1) an increase in the number of units per person, and (2) a randomized demand ranging from 1 to 3 units, depending on the severity of each patient. Specifically, mild cases require 1 unit, chronic conditions require 2 units, and patients with multiple conditions require 3 units. These scenarios are intended to evaluate how variations in resource demand impact overall patient allocation and associated costs.

Notably, the transportation cost parameter (c_{ij}) is not modified in this analysis, as it is calculated based on the actual distance between each community health center (SHPH) and the patient's residence. The cost per kilometer is derived from empirical data specific to the study area, rendering this parameter a fixed input that reliably reflects the real-world context of healthcare service delivery. The case scenarios described above are presented in Table 5. Subsequently, the data were processed using the LINGO program. The results of the analysis for each scenario are summarized in Table 6, with Case 1 representing the baseline scenario.

Table 5. Case Scenarios for Sensitivity Analysis.

Case No.	Case Code	Fixed Cost (THB)	Capacity (Patients/SHPH)	Demand (Units/Person)	Remarks
1	Base	20,000	10	1	Baseline scenario
2	FC-1	10,000	10	1	Reduced fixed cost
3	FC-2	30,000	10	1	Increased fixed cost
4	FC-3	10,000–30,000	10	1	Randomized fixed cost
5	CAP-1	20,000	7	1	Reduced SHPH capacity
6	CAP-2	20,000	20	1	Increased SHPH capacity
7	CAP-3	20,000	7–20	1	Randomized SHPH capacity
8	DEM-1	20,000	20	3	Increased medication demand
9	DEM-2	20,000	20	1–3	Randomized medication demand
10	COMB-1	10,000–30,000	7–20	1–3	Combined randomized scenario

Table 6. Analysis Results of Case Scenarios for Sensitivity Analysis

Case No.	Case Code	Number of SHPHs Opened	Selected SHPHs	Number of Patients Allocated	Total SHPH Participation Cost (THB)	Runtime (seconds)
1	Base	6	P1, P2, P4, P7, P8, P9	58	122,196.50	0.38
2	FC-1	6	P1, P2, P3, P7, P8, P9	58	62,196.53	0.55
3	FC-2	6	P1, P2, P3, P7, P8, P9	58	182,196.50	0.41
4	FC-3	6	P1, P2, P3, P5, P7, P9	58	107,089.20	0.19
5	CAP-1	9	P1, P2, P3, P4, P5, P6, P7, P8, P9	58	182,579.90	0.16
6	CAP-2	3	P6, P7, P9	58	62,350.60	0.80
7	CAP-3	4	P3, P7, P8, P9	58	82,355.76	0.94
8	DEM-1	9	P1, P2, P3, P4, P5, P6, P7, P8, P9	58	188,102.00	0.17
9	DEM-2	6	P1, P2, P5, P7, P8, P9	58	124,388.80	0.61
10	COMB-1	9	P1, P2, P3, P4, P5, P6, P7, P8, P9	58	193,734.40	0.12

To assess the robustness of the proposed model under varying conditions, Table 6 presents how these factors impact total cost, facility utilization, and patient allocation. Adjusting the SHPH participation cost directly influenced total expenditure. Reducing the fixed cost to 10,000 THB (Case 2) lowered the total cost to 62,196.53 THB, while all 58 patients were successfully allocated. Increasing it to 30,000 THB (Case 3) raised costs significantly to 182,196.50 THB. A randomized cost between 10,000–30,000 THB (Case 4) resulted in a moderate cost of 107,089.20 THB. These results highlight that lowering fixed costs can improve budget efficiency without affecting service coverage. Changes in SHPH capacity also had notable effects, namely, limiting capacity to 7 patients (Case 5) requiring all 9 facilities, raising the total cost to 182,579.90 THB. Increasing capacity to 20 patients (Case 6) reduced the need to only 3 facilities and cut costs to 62,350.60 THB—the lowest among all cases. Random capacity between 7–20 patients (Case 7) used 4 SHPHs, with a moderate cost of 82,355.76 THB. Higher capacity boosts operational efficiency. Increasing demand to 3 units per patient (Case 8) led to the use of all SHPHs and a total cost of 188,102.00 THB. Randomized demand between 1–3 units

(Case 9) required 6 facilities and cost 124,388.80 THB. The model adapts well to varied demand levels, though higher or more variable demand increases the cost. Randomizing all three parameters simultaneously required all 9 facilities and yielded the highest cost of 193,734.40 THB. Despite the complexity, the runtime was just 0.12 seconds, indicating high computational efficiency. Across all scenarios, the model consistently allocated all 58 patients, demonstrating flexibility and robustness. However, fixed costs and service capacity were the most influential in determining total cost and facility requirements. These insights can help guide policymakers in optimizing resource allocation and budgeting under varying conditions.

3.4 Discussion

A review of the existing literature on HHC logistics reveals that most studies have primarily focused on improving operational efficiency through routing and scheduling optimization. Various mathematical and algorithmic techniques have been employed to address the VRP and its derivatives. For instance, van Montfort et al. (2024), Phonin et al. (2025), and Atta et al. (2025) emphasized optimizing task allocation, reducing service time, and aligning caregiver assignments with routing efficiency. However, these studies have not incorporated the structure of community-based health systems or considered the socio-spatial contexts of developing countries in their models. The present study introduces a novel approach by integrating both strategic and operational planning levels through the combination of LAP and VRP. The objective is to enhance healthcare access for patients with hypertension and limited mobility in Hat Yai District, Thailand. This differs from prior work such as that of Suwatcharachaitiwong et al. [4], which focused on hybrid delivery models (home delivery and self-collection) without embedding actual community-based primary healthcare infrastructure into the optimization framework. Importantly, this study is the first to explicitly incorporate Subdistrict Health Promoting Hospitals (SHPH) and Village Health Volunteers (VHV)—key components of Thailand’s primary healthcare system—into the model. In contrast, although works by Kordi et al. [10], Abdul Nasir and Kuo [11], and Shi et al. [24] addressed flexibility, uncertainty, and cost efficiency, they did not systematically reference local health delivery systems or adapt their models to community-specific infrastructures. In terms of methodology, this research utilizes exact optimization via LINGO in conjunction with Geographic Information Systems (GIS), enabling spatially grounded analysis suitable for real-world community-level planning. This contrasts with many studies that rely solely on mathematical algorithms such as MOVNS, genetic algorithms, or MILP, often tested on generated data, which may not be directly applicable in field settings. For example, while Paradiso et al. [31], Salavati-Khoshghalb et al. [32], and Balti & Jemai [34] proposed advanced techniques, their models are more aligned with urban or industrial logistics rather than healthcare delivery in rural or community contexts. Furthermore, the use of real patient data and authentic community health structures enhances the validity and applicability of the proposed model for local policy-making. This sets the study apart from much of the literature, which tends to rely on theoretical simulations. Consequently, the findings from this research offer a practical framework for proactive, community-based healthcare service planning in Thailand. They could be adapted to similar settings across Southeast Asia that share comparable public health infrastructures.

4. Conclusions

This study aims to design a logistics system that supports home healthcare services under the “Taking Medicine Nearby House Project” in the Hat Yai district, Songkhla province. The focus is on analyzing and planning various aspects, including the selection of suitable primary healthcare units (SHPHs), the allocation of chronic disease patients who are unable to travel for healthcare services, and the development of efficient medication delivery routes and home healthcare services. The analysis process utilizes a mathematical model combined with Geographic Information System (GIS) data to assess the actual distances on the road network, and the processing is conducted using the LINGO program. This approach facilitates the determination of optimal service point locations and the strategic distribution of resources in alignment with both service capacity constraints and the geographical characteristics of the target area. The analysis revealed that the application of the LAP model resulted in a total cost of 122,196.50 THB, while the VRP model produced an average service cost of 952.87 THB per delivery. When combined, the total cost of implementing both models amounted to 123,149.37 THB. Based on the model's output, six out of nine SHPHs were selected to serve as distribution centers. These included Khu Tao, Chalung, Thung Tam Sao, Ban Phru, Phatong, and Khuan Lang.

A total of 58 patients were successfully assigned to these SHPHs without exceeding the individual capacity limitations of any unit, thereby ensuring operational feasibility and balanced workload distribution. The resource allocation and route planning approach used in this research demonstrates the potential to reduce logistics costs in the system, especially transportation costs, which are a critical component of the Home Healthcare system. In conclusion, the application of logistics techniques combined with mathematical models and GIS systems has the potential to enhance the efficiency of primary healthcare service management in the Home Healthcare model. It can be applied to logistics planning for other public health projects in the future.

The implementation of the "Taking Medicine Nearby House" initiative, through the application of Location-Allocation and Vehicle Routing models, provides managerial insight into precise planning for selecting service points and optimizing medication distribution routes. This approach considers both the capacity constraints of individual healthcare units and the geographical characteristics of the target area. Such an integrated strategy enables a more balanced distribution of workloads and contributes to reducing overall system costs—particularly transportation expenses, which are a major component of home healthcare services. It offers a practical and actionable tool to support strategic decision-making, resource planning, and the design of more efficient community-level health service systems. This method proves especially valuable in delivering care to target populations with mobility limitations, such as chronically ill patients and the elderly. It is also adaptable to other public health initiatives, including vaccine distribution and the delivery of medical supplies to remote or underserved areas. In addition to enhancing operational efficiency, the model promotes equity in healthcare access and supports the development of sustainable health systems under resource and budget constraints. The approach aligns with Thailand's national health policy, which emphasizes the delivery of safe, high-quality healthcare services while reducing congestion, disparities, waiting times, and financial burdens [39]. This policy, developed by the National Health Security Office (NHSO) in collaboration with the Ministry of Public Health and a network of service providers—including VHVs, SHPHs, community hospitals, and general hospitals—offers a solid foundation for strengthening primary healthcare delivery nationwide. In this context, the model contributes not only technical solutions but also meaningful managerial insight for improving healthcare systems in a locally responsive and scalable manner.

The suggestion for improving the problem-solving model, based on the characteristics of issues related to both the selection of healthcare service locations and the planning of medication delivery and home services, highlights the necessity of developing a mathematical model that can appropriately integrate both aspects. This model is known as LRP, which presents a potential approach for designing healthcare systems that are both efficient and cost-effective. For smaller-scale case studies, such as those involving a limited number of service units and patients, the use of exact methods can still provide accurate results within an acceptable timeframe. However, when the problem size increases, such as in cases involving 9 service units and up to 157 patients, it was found that processing with the LINGO program required over 100 hours and could not provide a solution within the constrained time limit. Therefore, to enhance the efficiency of solving the problem under time constraints, it is recommended to consider the use of answer approximation techniques, such as Heuristics and Metaheuristics. These methods can significantly reduce computation time while providing practically acceptable results. Furthermore, the mathematical model should be improved to better reflect real-world situations by considering resource constraints, the service capacity of each healthcare unit, as well as geographic factors and travel characteristics specific to each area. This will ensure that the problem-solving approach is appropriate, can be practically applied, and supports strategic decision-making in planning the logistics system for Home Healthcare services effectively in the future.

5. Acknowledgements

Author Contributions: Conceptualization, Sirivongpaisal N. and Suwatcharachaitiwong S.; methodology, Sirivongpaisal N. and Suwatcharachaitiwong S.; software, Samohyusoh S.; validation, Sirivongpaisal N., Suwatcharachaitiwong S. and Samohyusoh S.; formal analysis, Sirivongpaisal N., Suwatcharachaitiwong S. and Samohyusoh S.; data curation, Samohyusoh S.; writing—original draft preparation, Samohyusoh S.; writing—review and editing, Sirivongpaisal N. and Suwatcharachaitiwong S.; visualization, Samohyusoh S.; supervision, Sirivongpaisal N. and Suwatcharachaitiwong S.; funding acquisition, Suwatcharachaitiwong S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Health Systems Research Institute (HSRI, grant number HSRI 65-031). The APC was funded by Health Systems Research Institute.

Conflicts of Interest: Declare conflicts of interest or state “

The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analysis, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

[1] Photha, T.; Singweeratham, N.; Techakehakij, W.; Wongphan, T. *Cost Assessment of the Pilot Program on Drug-Dispensing Services in Pharmacies to Reduce Hospital Overcrowding*; Health Systems Research Institute: Nonthaburi, Thailand, August 31, 2020.

[2] Passarat, B. Ministry of Public Health Demonstrates That Telemedicine Effectively Reduces Outpatient Numbers. *Post Today*. https://www.posttoday.com/smart-life/721315 (accessed Mar 21, 2025)

[3] Lohapaiboonkul, J.; Phannikun, T. Differential Evolution Method for Solving Location and Routing Problems: A Case Study of Palm Collection Points and Transportation in Narathiwat Province. *J. Ind. Tech. Ubon Ratchathani Rajabhat Univ.* **2019**, *9*(2), 1–12.

[4] Suwatcharachaitiwong, S.; Siriwongpaisan, N.; Phatchachansiri, R. *Medication and Medical Supplies Service System Design for Patients Receiving Medication at Type 1 Pharmacies in Hat Yai, Songkhla*; Health Systems Research Institute (HSRI), Targeted Research Program in Medicine and Public Health (Fiscal Year 2022): Nonthaburi, Thailand, November 2023.

[5] Health Insurance System Research Office (HISRO). Cost Analysis of Health Services: A Knowledge Base for the Development of Sustainable Health Service and Financing Systems. https://www.hsri.or.th/news /1/3607 (accessed Mar 21, 2025)

[6] Office of Logistics System Development Strategy. Thailand Logistics Report 2023. https://www.nesdc. go.th /download (accessed Mar 12, 2025)

[7] Prince of Songkla University. PSU Radio Station Policy, Hatyai Hospital and the Burden of Overwork; July 2, 2024. https://psub.psu.ac.th/?p=12272 (accessed March 30, 2025)

[8] Phonin, S.; Likasiri, C.; Pongvuthithum, R.; Chonsiripong, K. Managing Home Healthcare System Using Capacitated Vehicle Routing Problem with Time Windows: A Case Study in Chiang Mai, Thailand. *Multidisciplinary Digital Publishing Institute (Logistics)* **2025**, *9*(3), 85. https://doi.org/10.3390/ logistics9030085

[9] Atta, S.; Basto-Fernandes, V.; Emmerich, M. Concise Review of Home Health Care Routing and Scheduling Problem. *Operations Research Perspectives.* **2025**, *15*, 100347. https://doi.org/10.1016/j.orp.2025.100347

[10] Kordi, G.; Divsalar, A.; Emami, S. Multi-Objective Home Health Care Routing: A Variable Neighborhood Search Method. *Optimization Letters.* **2023**, *17*, 2257–2298. https://doi.org/10.1007/ s11590-023-01993-y

[11] Abdul Nasir, J.; Kuo, Y.-H. Stochastic Home Care Transportation with Dynamically Prioritized Patients: An Integrated Facility Location, Fleet Sizing, and Routing Approach. *Transp. Res. Part B Methodol.* **2024**, *184*, 102949. https://doi.org/10.1016/j.trb.2024.102949

[12] Euchi, J.; Zidi, S.; Laouamer, L. A Hybrid Approach to Solve the Vehicle Routing Problem with Time Windows and Synchronized Visits in Home Health Care. *Arab. J. Sci. Eng.* **2020**, *45*, 10637–10652. https://doi.org/10.1007/s13369-020-04828-5

[13] Ransikarbum, K.; Kritchanchai, D.; Chanpuypetch, W.; Niemsakul, J. Central Hospital Location and Distribution Planning Using Integrated K-Means and Vehicle Routing Algorithm in the Healthcare Chain. *Oper. Supply Chain Manag.* **2024**, *17*(2), 299–315. http://doi.org/10.31387/oscm0570429

[14] Salami, A.; Afshar-Nadjafi, B.; Amiri, M. A Two-Stage Optimization Approach for Healthcare Facility Location-Allocation Problems with Service Delivering Based on Genetic Algorithm. *Int. J. Public Health* **2023**, *68*, 1605015. https://doi.org/10.3389/ijph.2023.1605015

[15] Shi, Y.; Lin, Y.; Li, B.; Li, R. Y. M. A Bi-Objective Optimization Model for the Medical Supplies' Simultaneous Pickup and Delivery with Drones. *Comput. Ind. Eng.* **2022**, *171*, 108389. https://doi.org/ 10.1016/j.cie.2022.108389

[16] Zaid, A. A.; Asaad, A. R.; Othman, M.; Haj Mohammad, A. Multi-Objective Technology-Based Approach to Home Healthcare Routing Problem Considering Sustainability Aspects. *Multidisciplinary Digital Publishing Institute (Logistics)* **2024**, *8*(3), 75. https://doi.org/10.3390/logistics8030075

[17] Phutthaphooltrakun, P.; Raothanachonlakul, P. Location Selection for Blood Distribution Center: Case Study – Rayong Province. *Journal of Eastern Management and Technology.* **2020**, *17*(1). https://so06.tci-thaijo.org/index.php/umt-poly/article/view/241672

[18] Tapabut, K.; Chaikram, T.; Rattanopas Suwanlee, S.; Somart, J. Application of Geographic Information System to the Allocation of Emergency Medical Service in the Elderly in Chuen Chom District, Mahasarakham Province. *Journal of Science and Technology.* **2019**, *37* (3). https://li01.tci-thaijo.org/index.php/tstj/article/view/184099

[19] Poomisirisawat, H.; Suwatcharachaitiwong, S.; Sirivongpaisal, N. An Integrated Approach for Designing Healthcare Facilities with a Location-Inventory Model. *Science & Technology Asia.* **2024**, *29*(2), 53–62. https://ph02.tci-thaijo.org/index.php/SciTechAsia/article/view/254630

[20] Pan, J.; Deng, Y.; Yang, Y.; Zhang, Y. Location-Allocation Modelling for Rational Health Planning: Applying a Two-Step Optimization Approach to Evaluate the Spatial Accessibility Improvement of Newly Added Tertiary Hospitals in a Metropolitan City of China. *Social Science & Medicine.* **2023**, *338*, 116296. https://doi.org/10.1016/j.socscimed.2023.116296

[21] Alghanmi, N.; Alotaibi, R.; Alshammari, S.; Alhothali, A.; Bamasag, O.; Faisal, K. A Survey of Location-Allocation of Points of Dispensing During Public Health Emergencies. *Front. Public Health* **2022**, *10*, 811858. https://doi.org/10.3389/fpubh.2022.811858

[22] Zhuo, M.; Li, J.; Zhang, R. Redesigning Deployments of Community Hospitals with a Location-Allocation Model. *IFAC-Papers OnLine*. **2022**, *55*(10), 1272–1277. https://doi.org/10.1016/j.ifacol. 2022.09.565

[23] Murad, A.; Faruque, F.; Naji, A.; Tiwari, A. Using the Location-Allocation P-Median Model for Optimising Locations for Health Care Centres in the City of Jeddah. *Geospat. Health* **2021**, *16*(2), 1002. https://doi.org/10.4081/gh.2021.1002

[24] Shi, Y.; Boudouh, T.; Grunder, O. A Hybrid Genetic Algorithm for a Home Health Care Routing Problem with Time Window and Fuzzy Demand. *Expert Systems with Applications.* **2017**, *72*, 160–176. https://doi.org/10.1016/j.eswa.2016.12.013

[25] Al Theeb, N.; Abu-Aleqa, M.; Diabat, A. Multi-Objective Optimization of Two-Echelon Vehicle Routing Problem: Vaccines Distribution as a Case Study. *Computers & Industrial Engineering.* **2024**, *187*, 109590. https://doi.org/10.1016/j.cie.2023.109590 .

[26] Euchi, J.; Zidi, S.; Laouamer, L. A Hybrid Approach to Solve the Vehicle Routing Problem with Time Windows and Synchronized Visits in Home Health Care. *Computer Engineering and Computer Science.* **2020**, *45*, 10637–10652. https://doi.org/10.1007/s13369-020-04828-5

[27] Durak, Z.; Mutlu, O. Home Health Care Nurse Routing and Scheduling Problem Considering Ergonomic Risk Factors. *Heliyon* **2024**, *10*(1), e23896. https://doi.org/10.1016/j.heliyon.2023.e23896

[28] HadjTaieb, S.; Hani, Y.; Moalla Loukil, T.; El Mhamedi, A. Green VRP Applied to Home Health Care Problem. *IFAC-PapersOnLine*. **2022**, *55*(10), 3154–3159. https://doi.org/10.1016/j.ifacol.2022.10.214

[29] van Montfort, L.; Dullaert, W.; Leitner, M. Task-Splitting in Home Healthcare Routing and Scheduling. *arXiv* **2024**, arXiv:2406.19288 [math.OC]. https://doi.org/10.48550/arXiv.2406.19288

[30] Zhang, Y.; Zhang, Z. Joint Optimization of Service Routing and Scheduling in Home Health Care. *arXiv* **2025**, arXiv:2505.20474 [math.OC]. https://doi.org/10.48550/arXiv.2505.20474

[31] Paradiso, R.; Roberti, R.; Laganá, D.; Dullaert, W. An Exact Solution Framework for Multitrip Vehicle-Routing Problems with Time Windows. *Operations Research.* **2020**, *68*(1), 1–307. https://doi.org/10.1287/opre.2019.1874

[32] Salavati-Khoshghalb, M.; Gendreau, M.; Jabali, O.; Rei, W. An Exact Algorithm to Solve the Vehicle Routing Problem with Stochastic Demands under an Optimal Restocking Policy. *European Journal of Operational Research.* **2019**, *273*(1), 175–189. https://doi.org/10.1016/j.ejor.2018.07.039

[33] Zetina, C. A.; Contreras, I.; Jayaswal, S. An Exact Algorithm for Large-Scale Non-convex Quadratic Facility Location. *arXiv*. **2021**, arXiv:2107.09746. https://doi.org/10.48550/arXiv.2107.09746.

[34] Balti, M.; Jemai, A. Route Optimization for Home Healthcare Vehicles. *SAGE Journals.* **2025**, *3*(4). https://doi.org/10.1177/27723577241305712

[35] Linfati, R.; Escobar, J. W.; Escalona, J. A Two-Phase Heuristic Algorithm for the Problem of Scheduling and Vehicle Routing for Delivery of Medication to Patients. *Math. Probl. Eng.* **2018**, *2018*, 8901873. https://doi.org/10.1155/2018/8901873

[36] Schneider, M.; Gschwind, T.; Vigo, D. Advances in Vehicle Routing and Logistics Optimization: Exact Methods. *EURO J. Transp. Logist.* **2019**, *8*, 117–118. https://doi.org/10.1007/s13676-018-0139-6

[37] Perl, J.; Daskin, M. S. A Warehouse Location-Routing Problem. *Transportation Research Part B: Methodological.* **1985**, *19* (5), 381–396. https://doi.org/10.1016/0191-2615(85)90052-9

[38] Kara, I.; Laporte, G.; Bektas, T. A Note on the Lifted Miller–Tucker–Zemlin Subtour Elimination Constraints for the Capacitated Vehicle Routing Problem. *European Journal of Operational Research.* **2004**, *158*(3), 793–795. https://doi.org/10.1016/S0377-2217(03)00377-1

[39] Ministry of Public Health. Operational Plan for Fiscal Year 2021; Office of the Permanent Secretary, Ministry of Public Health: Nonthaburi, Thailand, 2020. [Online] Accessed July 17, 2025. Available from: http://bps.moph.go.th/new_bps/sites/default/files/plan_ops64.pdf

ASEAN
Journal of Scientific and Technological Reports
Online ISSN:2773-8752

Research article

Vegetative, Photosynthetic, and Anthocyanin Content of Turmeric (*Curcuma longa* L.) Applied with Organic Foliar Fertilizers

Pet Roey L. Pascual[1,2], Edivine A. Remedios[3], Nonna Fatima H. Abello[4*], Danny E. Carabio[5], Valerie U. Pascual[6], and Rosalyn P. Alburo[7]

[1] Crop Science, Cebu Technological – Barili Campus, Cagay, Barili 6036, Cebu, Philippines
[2] Center for Studies in Biotechnology, Cebu Technological University - Barili Campus, Cagay, Barili 6036, Cebu, Philippines
[3] Crop Science, Cebu Technological – Barili Campus, Cagay, Barili 6036, Cebu, Philippines
[4] Graduate School of Bioscience and Biotechnology, Chubu University, Kasugai, Aichi, 487-8501, Japan
[5] Agribusiness, Cebu Technological University - Barili Campus, Cagay, Barili 6036, Cebu, Philippines
[6] Development Communication, Cebu Technological University - Barili Campus, Cagay, Barili 6036, Cebu, Philippines
[7] Biodiversity, Environment and Natural Resources Research Center, Cebu Technological University, Argao, Cebu 6021, Philippines
* Correspondence: nonnafatima.abello@ctu.edu.ph

Citation:
Pascual, P.R.L; Remedios, E.A.; Abello, N.F.H.; Carabio, D.E.; Pascual, V.U.; Alburo, R.P. Vegetative, photosynthetic, and anthocyanin content of turmeric (*Curcuma longa* L.) applied with organic foliar fertilizers. *ASEAN J. Sci. Tech. Report.* **2025**, *28*(5), e257477. https://doi.org/10.55164/ajstr.28i5.257477

Article history:
Received: January 13, 2025
Revised: July 20, 2025
Accepted: August 16, 2025
Available online: September 14, 2025

Publisher's Note:
This article is published and distributed under the terms of Thaksin University.

Abstract: In the Philippines, turmeric is used as a preservative and as medicine. However, due to the limited knowledge of its fertilizer management, this has resulted in low growth and yield. Foliar fertilizers are now commonly used in organic farming. A study was done to evaluate the effect of different foliar fertilizers: seaweed, bamboo shoot, Japanese snail, and commercial liquid fertilizer, arranged in a randomized complete block design with three replications, based on vegetative, photosynthetic, and anthocyanin content using methanol extraction of turmeric. Photosynthetic parameters were measured using the Li-6800 Portable Photosynthesis System. Results demonstrated that fermented seaweed significantly enhanced vegetative growth, increasing plant height at 19.72 ± 2.00 SD cm (3rd week), 28.53 ± 1.68 SD cm (5th week), and 36.40 ± 1.31 SD cm (7th week), and leaf length by 2.57 cm relative to the control. In contrast, the application of fermented Japanese snail yielded the highest anthocyanin concentration at 5.18 ± 0.14 SE µg/g and markedly improved photosynthetic traits, including assimilation rate at 11.26 ± 0.31 mmol CO_2 m^{-2} s^{-1}, stomatal conductance to CO_2 at 393.71 ± 7.41 SD mmol m^{-2} s^{-1} and total conductance at 42 ± 1.18 SD Pa. Notably, anthocyanin content exhibited a strong positive correlation with assimilation rate (R=0.99) stomatal conductance (R=0.84) and total conductance (R=0.84), and a negative correlation with plant height (R=0.61). Thus, the application of fermented seaweed can enhance morphological parameters, while the fermented Japanese snail foliar application resulted in better photosynthetic efficiency and anthocyanin content.

Keywords: Japanese snail liquid fertilizer; organic foliar fertilizers; photosynthetic responses; seaweed liquid fertilizer; Anthocyanin

1. Introduction

Turmeric (*Curcuma longa* L.), locally referred to as luyang dilaw, dilaw, or dulao in the Philippines, is mainly used as a food additive, preservative, and coloring agent. Little known is that turmeric has many medicinal properties. Studies showed that curcumin (diferuloylmethane), the yellow bioactive compound in turmeric, has been shown to possess a wide range of biological effects, including anti-inflammatory, antioxidant, anticancer, antidiabetic, antibacterial, antifungal, antiviral, and many other health benefits [1]. Many people in the world use turmeric in various ways to improve their well-being.

The lack of farmers' knowledge of fertilizer management resulted in low growth, yield, and anthocyanin content. Moreover, turmeric plants are also prone to several diseases. Several authors, however, suggested that optimum growth, rhizome yield, net return, and cost-benefit ratio can be attained through the combined application of both organic and inorganic fertilizers [2-4]. The interest in organic farming in the Philippines is also rapidly gaining ground as the trend toward healthier lifestyles expands. With a focus on organic farming for health considerations and meeting stringent consumer standards, agricultural practices are increasingly being modified [5]. Using organic fertilizers, which play an important role in improving soil physical properties [6], can serve as an alternative practice to using N mineral fertilizers [7]. Since organic agriculture relies almost entirely on natural methods and avoids synthetic pesticides, the risk of synthetic pesticide pollution in ground and surface waters is minimal to nonexistent. Local resource utilization as fertilizers also contributes to rural development, resource sustainability, and reduced reliance on external input [8].

Foliar feeding generally favors the means of correcting nutrient deficiency, as these shortages include secondary nutrients and micronutrients. Through foliar fertilizer application, leaf activity increases, as observed in pepper [9], French bean [10], and sunflower [11]; thus, chlorophyll, photosynthesis, and anthocyanin increase in a relatively short time. Anthocyanin is widely studied for its medicinal properties as it prevents Type 2 Diabetes mellitus (T2DM), some types of cancer, and hypertension [12-14]. Furthermore, strong positive correlations among anthocyanin, flavonoid, and alkaloid contents in turmeric and their corresponding antioxidant activities (DPPH inhibition and reducing power) suggest that anthocyanins, as flavonoids, play a major additive role in turmeric's antioxidant properties [15]. Foliar application can target a particular stage of crop development to achieve specific objectives and is an excellent way to fine-tune a high-fertility program [16]. The key factors that decide optimum fertilization practices are crop nutrient absorption and crop yields. It is, therefore, vital to apply fertilizers effectively to reduce loss and increase the quality of nutrient use. Currently, Cebu Technological University – Barili Campus (CTU-Barili Campus) promotes organic farming management practices for improved production of indigenous vegetables and cereals. This research also emphasizes anthocyanin content in line with its goals to enhance secondary metabolites and other phytochemicals in plants. The study aims to evaluate the effect of different foliar fertilizers on the morphological, anthocyanin, and photosynthetic responses of turmeric. Foliar application can readily correct the nutrient deficiency, and it includes micronutrients that affect the secondary metabolite content of a test crop. Moreover, this may ascertain claims that the use of certain organic foliar fertilizers may improve the nutritional value of the crops. In this case, the anthocyanin content of turmeric.

2. Materials and Methods

2.1 Experimental Site and Description

The experiment was conducted inside one of the greenhouses of CTU-Barili Campus (10°7'53" N Latitude, 123°32'45" E longitude) with an area of 8.95 m x 6 m, with 5 pots per replication per treatment, at one plant per pot. The dimensions of each pot were 19 cm × 15.5 cm × 13.5 cm. The greenhouse has a light intensity of 40.79 μmol m^{-2} s^{-1} and a temperature of 25.92 °C.

2.2 Preparation of Seaweed Fermentation

18 kg of seaweed (Kappaphycus alvarezii) was bought from Barili Public Market and washed to remove dirt. For every ½ kg of seaweed, 500 mL of water was added and blended in an electric blender. This step was done repeatedly until all the seaweeds were blended. All liquids were removed by straining. The by-product was then added with 10 kg of molasses and stirred thoroughly. It was then stored inside a container covered with Manila paper and placed in a cool, dry area. After seven days, seaweeds were filtered through a strainer, and the liquid was dispensed into another container and stored in a cool and dry place.

2.3 Preparation of Japanese Snail Fermentation

Japanese snails weighing a total of 7 kg were collected at CTU-Barili Campus. The snails—including their shells—were then washed and smashed. It was then placed inside a container and added to 7 kg of molasses. The mixture was then stirred thoroughly. The container was covered with Manila paper and then stored in a cool and dry area. After seven days, the mixture was filtered through a strainer. The liquid was dispensed into another container and then stored in a cool, dry place.

2.4 Preparation of Bamboo Shoot Fermentation

Bamboo shoots (*Bambusa blumeana*) weighing a total of 8.5 kg were collected at CTU-Barili Campus. The outer covering of the bamboo shoots was peeled off, then chopped into pieces. The bamboo shoots were then placed inside a container and filled with 8.5 L of water. The mixture was stirred thoroughly. The container was covered with Manila paper and was stored in a cool, dry area. After 24 hours, the bamboo shoots were squeezed and filtered through a strainer. The liquid was dispensed into another container and stored in a cool, dry place.

2.5 Treatments and Nutrient Analysis

The treatments used in this study are T0, commercial natural liquid fertilizer, T, Fermented Seaweed, T2, fermented bamboo shoot, and T3, fermented Japanese snail. Following the recommended rate of a commercial seaweed fertilizer, 10 mL/L was used for all treatments and sprayed five times on the leaves once a week. One rhizome of turmeric was planted in each pot filled with 2.5 kg of soil. The nutrient content of the three fermented organic fertilizers, as shown in Table 1, was determined by the Regional Soils Laboratory, Department of Agriculture. Total Nitrogen, Total Phosphorus, and Total Potassium were obtained using the Kjeldahl Method [17], the Vanadomolybdate method [18], and the Flame atomic emission spectroscopy method [19], respectively.

Table 1. Total nutrient content of fermented organic foliar fertilizers

	Total Nitrogen (%)	Total Phosphorus (%)	P_20_5 (%)	Total Potassium (%)	K_2O (%)
Fermented seaweed	0.12	0.003	0.006	1.21	1.46
Bamboo shoot	0.11	0.012	0.028	0.23	0.28
Japanese snail	0.14	0.002	0.004	1.01	1.22

2.6 Plant Growth Metrics

2.6.1. Morphological and Physiological Assessment

Plant height (measured from the base to the tip of the plant held vertically), leaf length (from the base of the petiole to the tip of the leaf blade of the middle leaf), and the number of leaves were recorded from the 2nd up to the 8th week from planting. Transpiration rate (mmol H_2O m^{-2} s^{-1}), assimilation rate (mmol CO_2 m^{-2} s^{-1}), average stomatal conductance (mmol m^{-2} s^{-1}), total conductance (Pa), and pressure deficit at leaf temperature (kPa) were measured using only the Li-6800 Portable Photosynthesis System. Random sampling was done where five fully expanded leaves from all sample plants per treatment were measured from the top, bottom, and middle of the plant during mid-morning, starting at 10:00 AM.

2.6.2. Anthocyanin Determination

Fifty grams (50 g) of powdered turmeric was homogenized in 100 mL acidified methanol, then macerated for 1 hr. This was filtered and reextracted until a faint-colored extract was obtained following the protocol of [20]. The filtrates were transferred to a rotary evaporator, and the extract was dissolved in methanol. This was stored at 4°C. One milliliter (1 mL) of the filtered extract from the composite rhizomes of the ten sample plants from the four treatments in three replications was pipetted into two (2) test tubes. One milliliter (1 mL) of 0.01% HCI solution in 95% ethanol was added to each test tube. The first tube (A1) was added with 10mL of 2% HCl (pH=0.8), and the second tube (A2) with 10 mL of citric buffer (pH=3.5). Each of the separate solutions was carefully mixed using a vortex mixer, and the absorbances were measured at 520 nm against a 70% methanol blank. Three replicates were conducted for each of the samples. The anthocyanin content was expressed as cyanidin. The total anthocyanin content was calculated using the equation:

TAC = (A1-A2) x f

Where:

TAC = total anthocyanin content expressed as "9 cyanidin g

A1 = absorbance in 2% HCI (at pH 0.8)

f = MW x DF x CF1 x CF2 εχε

A2 = absorbance in citrate buffer (at pH 3.5)

MW = molecular weight of cyanidin-3-glucoside (449 g/mol)

DF = dilution factor (50mL/10g)

1 = path length (1 cm)

CF1 = conversion factor 1 (10 μg/g)

CF2 = conversion factor 2 (1 L/1000 mL)

/ε = molar extinction coefficient of cyanidin-3-glucoside (26,900 L/mol.cm)

2.7 Statistical Analysis

The effects of treatments were analyzed using Analysis of Variance (ANOVA), and a further test was done using Tukey's test at $p<0.05$ to test for variances between treatment means.

3. Results and Discussion

3.1 Morphological Responses

The application of organic foliar fertilizers has a notable impact on the growth of turmeric plants compared to the control. Throughout the experimental duration, the plants treated with fermented seaweed organic foliar fertilizer consistently exhibited higher growth rates (Fig. 1A). The application of fermented seaweed significantly produced the tallest plant at 36.40 ± 1.31 SD cm compared to other liquid fertilizers in the 8th week. Since week 2, turmeric plants treated with fermented seaweed consistently showed the most significant increase in plant height, indicating its strong promotive effect on shoot elongation throughout the growth period. By week 3, seaweed-treated plants had already outpaced those treated with bamboo shoot and Japanese snail, and this lead was maintained through week 8. Similar results were recorded in rice bean, which garnered the tallest plant at 128.53 ± 3.9 SD cm, and this also gives significant growth in *Lactuca sativa* (var. curly green), *Cucumis sativus, Lactuca sativa, Phaseolus aureus,* and *Cajanus cajan* when applied with seaweed liquid fertilizer [21-23].

Also, through foliar spraying, the seaweed-treated plants exhibited the longest leaves throughout the observation period. The difference became particularly notable from week 5 onward, with a leaf length increase at 19.35 ± 1.37 SD cm, where seaweed consistently outperformed bamboo shoot and Japanese snail treatments. This resulted in a 2.57 cm difference in length with the commercial liquid fertilizer. This corresponds with the result of Huda, M. N. et al. [24], where seaweed liquid fertilizer at 20% concentration increased leaf area. The extract contains micronutrients, auxins, cytokinins, and other growth-promoting substances in large amounts, which can effectively promote plant growth as well as improve leaf development [25-26].

Seaweed fertilizer provides crops with a full spectrum of high to medium micronutrients, including polysaccharides, amino acids, vitamins, and other active compounds. These are crucial for developing a strong root system, which enhances the absorption of soil nutrients, water, and gases [27]. These hormones and vitamins are very significant in enhancing cell size and cell division. When combined, they enhance each other, as cytokinins stimulate shoot formation and auxins aid in root development. At the same time, micronutrients improve soil health [28]. The findings of [22] express that both fermented seaweed liquid fertilizer and seaweed sludge give promising results on *Lactuca sativa* (var. curly green) plant growth. An increase in the number of leaves was observed with the application of fermented seaweed compared to the commercial liquid fertilizer. This confirmed the result of Jayasinghe, P. [29] on *Capsicum annum* with the treatment of seaweed liquid fertilizer, which enhances shoot length, dry weight, root, number of leaves, number of pods, and yield. There is an excellent improvement in leaf length as it was treated with fermented seaweed [30], as seaweed extracts significantly increase the physiological activities of the plants that are visible in the size of leaves, leaf length, and width of onion cultivars [31]. An anticipated high level of cytokinins, auxins, other growth hormones, and nutrients present in fermented seaweed showed an ameliorating role in turmeric.

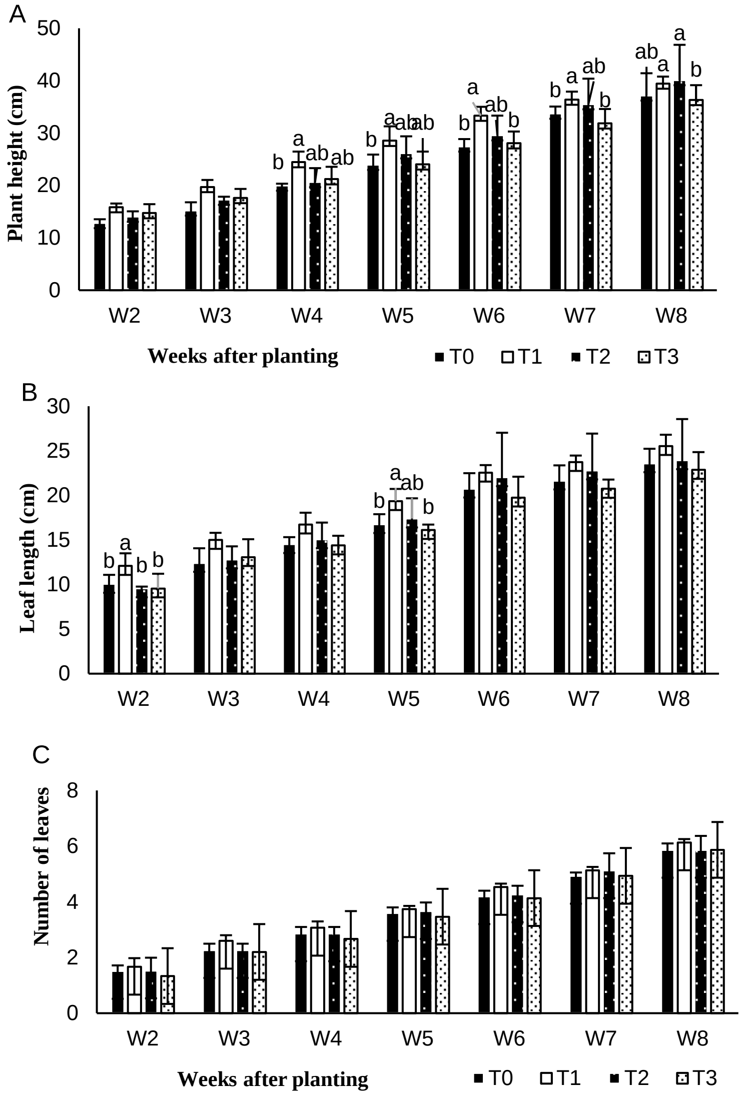


Figure 1. Weekly plant height (A), leaf length (B), and number of leaves (C) of Turmeric as affected by different organic foliar fertilizers. Different superscript letters indicate significant differences (Tukey HSD, $\alpha = 0.05$); Error bars represent mean ± SD (standard deviation

3.2 Physiological Responses

The experimental area experiences warm temperatures throughout the year, ranging from 74 °F to 91 °F and rarely dropping below 72 °F or exceeding 95 °F. These conditions typically promote active transpiration in plants. In this study, however, the application of different organic foliar fertilizers did not significantly affect the transpiration rate among treatments (Fig. 2B). This indicates that leaf water loss through transpiration remained consistent regardless of fertilizer type. The vapor pressure deficit at leaf temperature also showed no significant variation, suggesting a stable leaf microclimate across all treatments.

Despite uniform transpiration rates, the consistent water loss may be influenced more by environmental factors such as temperature and humidity than by the nutrient composition of the fertilizers applied. The lack of variation implies that the organic foliar fertilizers did not impose additional water stress or alter stomatal water regulation. Instead, plants maintained similar transpiration behavior under all treatment conditions, reflecting efficient stomatal control and stable internal water relations during the growth period.

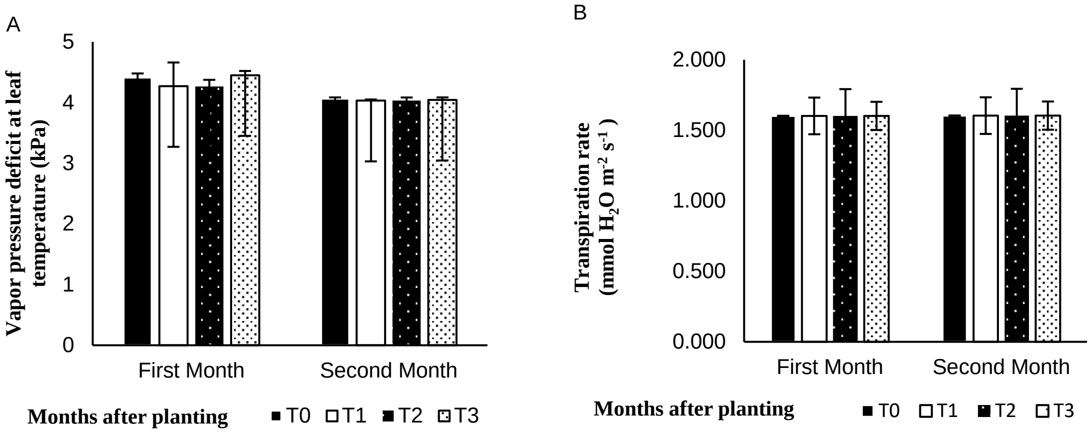


Figure 2. Vapor Pressure Deficit Leaf Temperature (A) and Transpiration Rate (B) of Turmeric as Affected by Different Organic Foliar Fertilizers. Different superscript letters indicate significant differences (Tukey HSD, $\alpha = 0.05$); Error bars represent mean ± SD (standard deviation).

Applying foliar fertilizer directly to the leaves typically boosts leaf activity and enhances photosynthesis. Due to this increased activity, the leaf increases its need for water. The increase in water uptake by the plant vascular system increases significantly over time. As shown in Fig. 3A, the net assimilation rate was increased at 11. 26 ± 0.31 mmol CO_2 m^{-2} s^{-1} after application of fermented Japanese snail surpassing the commercial foliar fertilizer and other foliar treatments. The assimilation of water and carbon dioxide results from photosynthesis, whereby the plant absorbs them and consequently transforms them directly into a multitude of organic molecules in the various cells of the plant, and is believed to be a growth predictor for some woody and herbaceous plants [32; 33]. Net assimilation rate increases in dry biomass per unit leaf area and is a dynamic physiological variable associated with breathing and photosynthetic rates. Carbon dioxide enters the leaf during photosynthesis due to a diffusion gradient between the atmosphere and the sites of carbon dioxide assimilation in the mesophyll. A decline in the assimilation rate meant that carbon dioxide and water were used less to produce essential biomolecules. As [34] stated, carbon dioxide and nitrate assimilation in plants are crucial for crop production. It appears that the nitrogen content in Japanese snails helps increase plant activity. In fact, [35] stressed that nitrogen would significantly affect the plant growth of lettuce as a effect of fermented bamboo shoots and Japanese snails.

Assimilation is closely linked to stomatal conductance, although this relationship is influenced by intrinsic water-use efficiency (iWUE = A/gs). The result, as seen in Figure 3B and 3C, indicated that it was fermented Japanese snail fertilizer that had the highest stomatal conductance and total conductance at 393.71 ± 7.41 mmol m^{-2} s^{-1} and 39.08 ± 1.18 mmol m^{-2} s^{-1}, respectively. Plants regulate stomatal conductance to optimize carbon uptake in relation to water loss [36-37]. Whereas, total conductance includes mesophyll, cuticular, and boundary layer conductance. Stomata were shown to react to a soil moisture deficit by producing abscisic acid from the roots, which helps regulate stomatal openings [38]. It was reported by [39] that organic fertilizer can increase the stomatal conductance of plants, such as green jujube leaves, increasing the photosynthetic rate. The trend in the net photosynthetic rate and stomatal conductance closely followed the changes in chlorophyll content [40; 41]. In addition, [42] stated that the improvements in leaf N content also significantly promote an increase in mesophyll and stomatal conductance. The addition of moderate levels of soil nitrogen significantly improved CO_2 diffusion conductance in the leaf anatomy and physiology of Manchurian ash and Mongolian oak. Furthermore, the two stages of CO_2 diffusion—the gas-phase and liquid-phase diffusional processes—were both greatly enhanced in the nitrogen-treated leaves. The effect of stomatal conductance on photosynthesis is mainly through carbon dioxide diffusion from the atmosphere to the substomatal cavities.

These enhancements in physiological characteristics in turmeric reflect improved CO_2 uptake and internal diffusion, resulting in elevated photosynthetic activity. A strong positive relationship was observed among assimilation rate, stomatal conductance, and total CO_2 conductance under T3, indicating a coordinated physiological response that optimized carbon assimilation in turmeric. The effect of T3 can be attributed to both its foliar action and its influence on soil fertility. [43] reported that golden apple snail residue improves soil pH, organic carbon, total and available nitrogen, and phosphorus—nutrients essential for photosynthesis and gas exchange. Nitrogen, in particular, supports chlorophyll production, Rubisco activity, and stomatal function, which are critical for C_3 photosynthesis [44]. The improved nitrogen availability under T3 likely contributed to the increased assimilation and conductance values. As highlighted by [45; 46], nitrogen-use efficiency plays a key role in plant responses to CO_2. Thus, the improvements under T3 likely result from both direct nutrient supply and enhanced soil conditions, creating an optimal environment for photosynthesis and CO_2 exchange in turmeric.

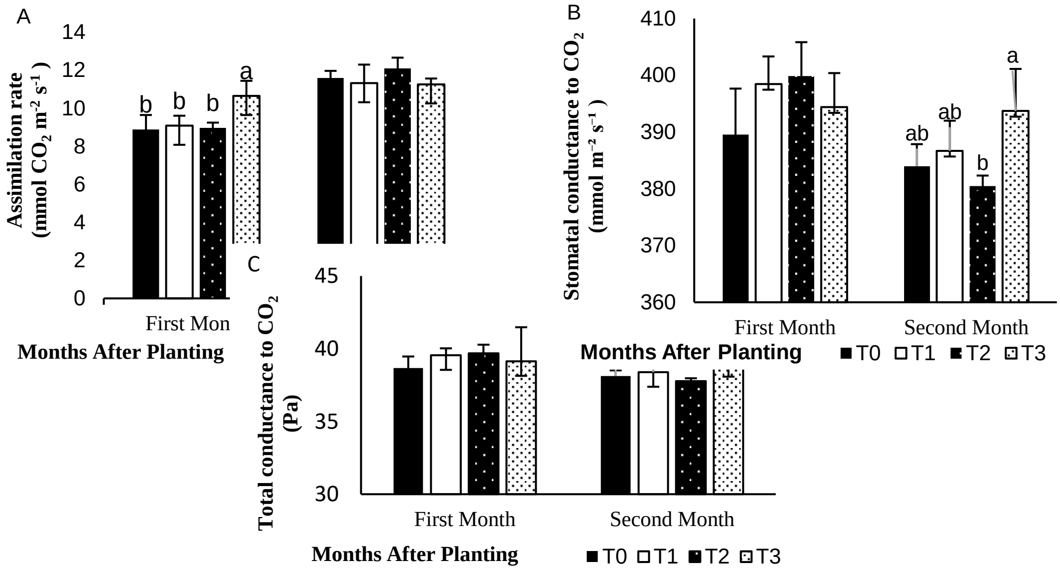


Figure 3. Assimilation rate (A), stomatal conductance to CO_2 (B), and total conductance to CO_2 (C) of Turmeric as affected by different organic foliar fertilizers. Different superscript letters indicate significant differences (Tukey HSD, $\alpha = 0.05$); Error bars represent mean ± SD (standard deviation).

Anthocyanins are a class of polyphenolic pigments that aid in both abiotic and biotic stress resistance as well as reproduction by luring pollinators and seed dispersers [47]. It is indicated in Figure 4 that anthocyanin content is significantly higher by 40% in Japanese snail fertilizer than in the other used organic fertilizer at 5.18±0.14 SE ug/g. As reported in the study of [48], total anthocyanin content significantly increased upon Calcium carbonate treatment, and Japanese snails are known to contain a high amount of CaCl2, which might be the reason why anthocyanin content is higher in T3. This was also supported by [49], where foliar spray of CaCl2 improved fruit total phenolics content, especially anthocyanins, by more than 20%. This was explained by [50] that Calcium acts as a second messenger in sugar signaling in grapes and wheat, where sugars promote light signaling-induced anthocyanin accumulation in seedlings through differential regulation. It has been suggested that changes in intracellular calcium levels regulate anthocyanin pigmentation in plants.

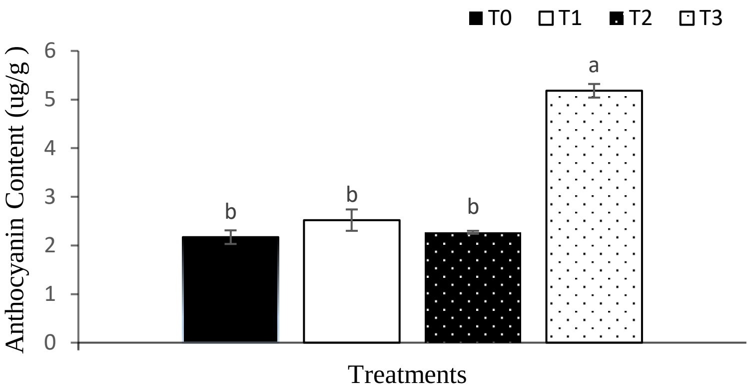


Figure 4. Anthocyanin content is affected by different organic foliar fertilizers. Different superscript letters indicate significant differences (Tukey HSD, α = 0.05); Error bars represent mean ± SEM (standard error of the mean).

The graph (Figure 5A) clearly shows that anthocyanin content has a strong negative relationship with the plant height of turmeric after application of Japanese snail fertilizer at R=-0.61. This also showed a significant positive strong relationship with the assimilation rate at R=0.99. As mentioned earlier, Japanese snail fertilizer contains carbon, which functions as a second messenger in sugar signaling within plants. In this process, sugars promote light signaling that triggers anthocyanin accumulation. When sucrose concentration was increased, anthocyanin levels in plants were significantly enhanced. However, this increase in sucrose concentration also resulted in a decrease in both plant height and leaf number [51]. It was reported by [52] that plants increase anthocyanins in response to abiotic stressors. However, it is well-known that abiotic stressors limit plant growth in environments where cell division and expansion are not ideal growth processes [53]. On the other hand, a high assimilation rate meant that CO_2 and water were more utilized for the production of essential biomolecules like anthocyanin. This was supported by the study of [54], where higher CO2 levels of 800 μmol·mol−1 significantly boosted anthocyanin production in ginger. The net assimilation rate and sustained growth rate showed a strong positive correlation with the maximum photosynthetic rate, which is also essential for anthocyanin accumulation [55]. In the study of [56], it was indicated that anthocyanins are highly associated with photosynthesis.

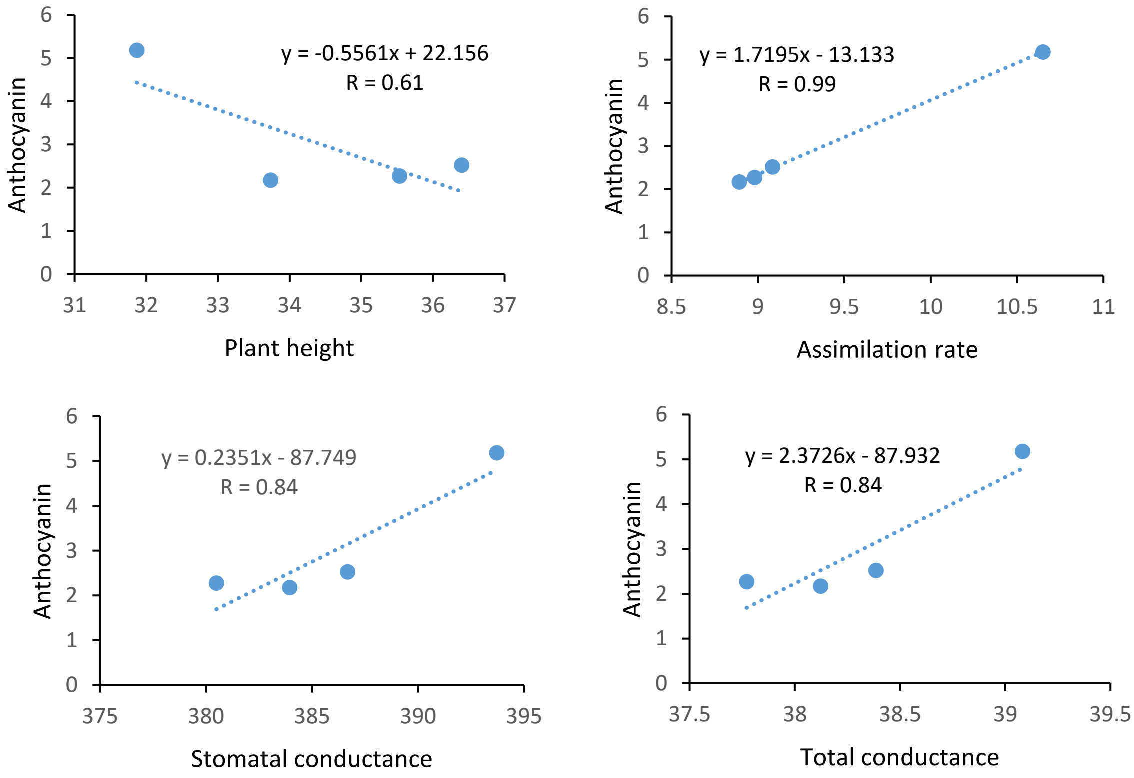


Figure 5. Relationship of anthocyanin content after application of Japanese snail fertilizer; (A) Plant height, (B) assimilation rate

4. Conclusions

Organic fermented seaweed fertilizer foliar application increased leaf length and plant height of turmeric. Concurrently, fermented Japanese snail fertilizer foliar application enhanced photosynthetic efficiency, as evidenced by stomatal conductance to CO_2 and total CO_2 conductance, as well as enhanced anthocyanin content. Future research may focus on the effects of different organic fertilizers on the enhancement of secondary metabolites in plants. These findings showcase the reactions of turmeric to different foliar fertilizers in terms of morphology, photosynthesis, and secondary metabolite content.

5. Acknowledgment

The authors wish to acknowledge all the support provided by Cebu Technological University for the realization of this research.

Author Contributions: Pascual PRL: Conceived idea, designed research methodology and analysed data; Remedios EA: Prepared materials, conducted study, collected data and wrote manuscript; Abello NFH: collected and analyzed data, wrote and edited the manuscript; Carabio DE: Literature Review, designed research methodology and data interpretation; Pascual VU: Manuscript critiquing, final reading and approval; Alburo RP: conducted and analyzed data on anthocyanin content

Funding: This study was financially supported by Cebu Technological University, Philippines under General Appropriations Act.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] Ishita, C.; Biswas, K.; Bandyopadhyay, U.; Banerjee, R. K. Turmeric and Curcumin: Biological Actions and Medicinal Applications. *Current Science*. **2004**, *87*(1), 44-53.

[2] Majeed, Y.; Ziaf, K.; Ghani, M. A.; Ahmad, I.; Ahmad, M. A.; Abbasi, K. Y.; Mujahid H. Cheema, K.L. Effect of different combinations of organic and synthetic sources of nutrients on growth, yield and quality parameters of turmeric under Faisalabad conditions. *Journal of Environmental and Agricultural Science*. **2020**, *22*, 1-7.

[3] Fernando, K. M. C.; Ekanayake, E. M. U. I. Plant growth, foliage senescence and rhizome yield of turmeric (Curcuma domestica L.) as affected by inorganic and organic fertilizers. *Tropical Agricultural Research and Extension*. **2022**, *25*(4), 344. https://doi.org/10.4038/tare.v25i4.5604

[4] Chandravanshi, O. K.; Meena, K. C.; Khan, K. A.; Soni, N.; Patidar, D. K. Responses of organic manures and inorganic fertilizers on growth, yield and economics of turmeric (*Curcuma longa* Linn.). *Journal of Medicinal Plants Studies*. **2021**, *9*(3), 243-247.

[5] Kramer, S.; Reganold, J.; Glover, J.; Bohannan, B.; Mooney, H. Reduced nitrate leaching and enhanced denitrifier activity and efficiency in organically fertilized soils. *Proceedings of the National Academy of Sciences of the United States of America*. **2006**, *103* (12), 4522-4527. https://doi.org/10.1073/pnas.0600359103

[6] Dauda, S.; Ajayi, F.; Ndor, E. Growth and yield of water melon *(Citrullus lanatus)* as affected by poultry manure application. *Journal of Agriculture and Social Sciences*. ***2008***, 4.

[7] Naeem, M.; Iqbal, J.; Bakhsh, M. A. A. Comparative Study of Inorganic Fertilizers and Organic Manures on yield and yield Components of Mungbean *(Vigna radiata L.)*. *Journal of Agriculture and Social Sci*ences. **2006**, 2, 227-229.

[8] Reganold, J.; Wachter, J. Organic agriculture in the twenty-first century. *Nature Plants*. **2016**, *2*. 15221. https://doi.org/10.1038/nplants.2015.221

[9] Mehdawe, A.; Mahadeen, A.; Al-Ramamneh, E. A. M. Foliar Application of Moringa Leaf Extracts Affects Growth, Yield and Mineral Composition of Pepper (Capsicum Annuum L.) Under Greenhouse Conditions. *Journal of Ecological Engineering*. **2023**, *24*(6), 329-337. https://doi.org/10.12911/22998993/163196

[10] Souri, M. K.; Aslani, M. Beneficial effects of foliar application of organic chelate fertilizers on French bean production under field conditions in a calcarous soil. *Advances in Horticultural Science*. **2018**, *32*(2), 265-272.

[11] Shindhe, V.; Dhanoji, M. M.; Meena, M. K.; Patil, R.P. Influence of foliar organic nutrition on growth, yield and yield components of sunflower. *J Pharmacogn Phytochem*. **2020**, *9*(1), 1267-1269.

[12] Marczylo, T. H.; Cooke, D.; Brown, K.; Steward, W. P.; Gescher, A. J. Pharmacokinetics and metabolism of the putative cancer chemopreventive agent cyanidin-3-glucoside in mice. *Cancer Chemotherapy Pharmacology*. **2009**, *64*, 1261-1268. https://doi.org/10.1007/s00280-009-0996-7

[13] Sandoval-Ramirez, B. A.; Catalan, U.; Fernandez-Castillejo, S.; Rubió, L.; Macià, A.; Solà, R. Anthocyanin tissue bioavailability in animals: possible implications for human health. A systematic review. *Journal of Agricultural and Food Chemistry*. **2018**, *66*, 11531-11543. https://doi.org/10.1021/acs.jafc.8b04014

[14] Lin, B. W.; Gong, C. C.; Song H. F.; Cui Y. Y. Effects of anthocyanins on the prevention and treatment of cancer. *British Journal of Pharmacology*. **2017**, *174*, 1226-1243. https://doi.org/10.1111/bph.13627

[15] Chinedum, E.; Kate, E.; Sonia, C.; Ironkwe, A.; Andrew, I. Polyphenolic Composition and Antioxidant Activities of 6 New Turmeric (Curcuma Longa L.) Accessions. *Recent Pat Food Nutr Agric*. **2015**, *7*(1), 22-27. https://doi.org/10.2174/2212798407666150401104716

[16] Bochalya, R.; Irungbam, P.; Bandyopadhyay, P.; Mallick, R.; Kant, K.; Gogoi, M. Performance of spring hybrid Sunflower (*Helianthus annuus L.* Var. GKSF-2002) under different foliar nutrients and growth regulators in West Bengal. **2018**.

[17] Kjeldahl, J. A. New Method for the Determination of Nitrogen in Organic Matter. *Zeitschrift für Analytische Chemie* **1883**, *22*, 366-382. https://doi.org/10.1007/BF01338151

[18] Baadenhuijsen, H.; Seuren-Jacobs, H. E.; Jansen, A. P. Continuous-flow determination of serum inorganic phosphate with a single reagent--the vanadomolybdate method re-evaluated., *Clinical Chemistry*. **1977,** *23*(7), 1275-1280. https://doi.org/10.1093/clinchem/23.7.1275

[19] Taylor, A. Encyclopedia of Spectroscopy and Spectrometry || Biomedical Applications of Atomic Spectroscopy. **1999**, 139-147. https://doi.org/10.1006/rwsp.2000.0022

[20] Wrolstad, R. E.; Acree, T. E.; Decker, E. A.; Penner, M. H.; Reid, D. S.; Schwartz, S. J.; Shoemaker, C. F.; Smith, D.; Sporns, P. Current Protocols in Food Analytical Chemistry || Extraction, Isolation, and Purification of Anthocyanins **2001**. https://doi.org/10.1002/0471142913.faf0101s00

[21] Pascual, P. R. L.; Carabio, D. E.; Abello, N. F. H.; Remedios, E. A.; Pascual, V. Enhanced assimilation rate due to seaweed biostimulant improves growth and yield of rice bean *(Vigna 10mbellate)*. *Agronomy Research*. **2021**.

[22] Pascual, P. R. L.; Carabio, D. E.; Rondina, M. E.; Abello, N. F. H.; Pascual, V. U. Fermented seaweed *(Kappaphycus alverezii)* by-product promotes growth and development of lettuce (*Lactuca sativa* var. Curly green). *Plant cell biotechnology and molecular biology*. **2020**, *21*(71-72), 208-214.

[23] Mohanty, D.; Adhikary, S. P.; Chatopadhyay, G. N.; Prof, P. C. Seaweed liquid fertilizer (slf) and its role in agriculture productivity. *Journal of Environmental Science*. **2013**, *3*, 23-26.

[24] Huda, M. N.; Mannan, M. A.; Bari, M. N.; Rafiquzzaman, S. M.; Higuchi, H. Red seaweed liquid fertilizer increases growth, chlorophyll and yield of mungbean (Vigna radiata). *Agronomy Research*. **2023**, *21*, https://doi.org/10.15159/ar.23.011

[25] Spinelli, F.; Fiori, G.; Noferini, M.; Sprocatti, M.; Costa, G. Perspectives on the use of a seaweed extract to moderate the negative effects of alternate bearing in apple and strawberry. *J. Horticut. Sci. Biotnol.* **2010**, 131-137. https://doi.org/10.1080/14620316.2009.11512610

[26] Tay, S. A. B.; Macleod, J. K.; Palni, L. M. S.; Letham, D. S. Detection of cytokinins in a seaweed extract *Phytochemistry*. **1985,** *24*, 2611-2614. https://doi.org/10.1016/S0031-9422(00)80679-2

[27] Mukherjee, A.; Patel, J. S. Seaweed extract: biostimulator of plant defense and plant productivity. *International Journal of Environmenatal Science and Technology*. **2019**, *17*, 553-558. https://doi.org/10.1007/s13762-019-02442-z

[28] Liu, Z.; Lijun, L. Effects of Plant growth regulators and saccharide on in vitro plant and tuberous root regeneration of Cassava. *Journal of Plant Growth Regulation*. **2011**, *30*(1), 11-19. https://doi.org/10.1007/s00344-010-9163-y

[29] Jayasinghe, P. Effect of Seaweed Liquid Fertilizer on Plant Growth of *Capsicum annum*. *Discovery An International Journal*. **2016**, *52*, 723-728.

[30] Munisamy, S.; Kosalaraman, K.; Devika, R. Seaweed plant nutrients to enhance the population and health of earthworm *(Eudrilus eugeniae)* along with crop betterment in chili *Capsicum annum* (var. Kkm-ch1) under pot culture. *Journal of Agriculture and Environment.* **2018,** *2*(6).

[31] Abbas, R. M.; Anwar, J.; Zafar-Ul-Hye, M.; Khan, R.; Rahi A.; Danish, S.; Datta, R. Effect of Seaweed Extract on Productivity and Quality Attributes of Four Onion Cultivars. *Horticulturae.* **2020,** *6*(2), 28. https://doi.org/10.3390/horticulturae6020028

[32] Shipley, B. Net assimilation rate, specific leaf area and mass ratio: which is most closely correlated with relative growth rate? A meta-analysis. *Funct Ecol.* **2006**, *20*, 565-574. https://doi.org/10.1111/j.1365-2435.2006.01135.x

[33] Poorter, H.; Nagel, O. W. The role of biomass allocation in the growth response of plants to different levels of light, CO2, nutrients and water: A quantitative review. *Functional Plant Biology.* **2000**, *27*. https://doi.org/10.1071/PP99173_CO

[34] Lawlor, D. W. Carbon and nitrogen assimilation in relation to yield: Mechanisms are the key to understanding production systems. *Journal of Experimental Botany.* **2002**, *53***,** 773-787. https://doi.org/10.1093/jexbot/53.370.773

[35] Carabio, D. E.; Pascual, V. U.; Abello, N. F. H.; Rondina, M. E.; Pascual, P. R. L. Combined application of fermented bamboo *(Bambusa spinosa)* and mollusk *(Achatina fulica)* liquid fertilizer can Improved lettuce (*Lactuca sativa* var. Curly green) production. *Plant cell biotechnology and molecular biology.* **2021**, *22*(3-4), 56-64.

[36] Mcausland, L.; Vialet-Chabrand, S.; Davey, P.; Baker, N. R.; Brendel, O.; Lawson, T. Effects of kinetics of light-induced stomatal responses on photosynthesis and water-use efficiency. *New Phytologist*. **2016**, *211*, 1209-1220. https://doi.org/10.1111/nph.14000

[37] Lawson, T.; Blatt, M. R. Stomatal size, speed, and responsiveness impact on photosynthesis and water use efficiency. *Plant Physiology*. **2014**, *164*, 1556-1570. https://doi.org/10.1104/pp.114.237107

[38] Liu, F.; Andersen, M.; Jacobsen, S.E.; Jensen, C. Stomatal control and water use efficiency of soybean (*Glycine max L.* Merr.) during progressive soil drying. *Environmental and Experimental Botany*. **2005**, *54*, 33-40. https://doi.org/10.1016/j.envexpbot.2004.05.002

[39] Wang, C. X.; Yue, X. W.; Shi, L. T.; Li, K.; Li, X. Y.; Fang, H. D.; Pan, Z. X. Responses of photosynthesis and yield of jujube to different proportions of organic fertilizer and biochar in Yuanmou dry-hot valley, *Chinese Journal of Tropical Crops*. **2022**, *43*, 28-136.

[40] Su, Y.; Huang, S. L. Effects of Bio-organic Fertilizer on Flue-cured Tobacco Photosynthetic Characteristics and Rhizosphere Soil Microorganism, *Journal of Agricultural Science and Technology*. **2022**, *24*, 164-171. https://doi.org/10.13304/j.nykjdb.2020.0731

[41] Chen, G. Y.; Chen, J.; Xu, D. Q. Thinking about the relationship between net photosynthetic rate and intercellular CO2 concentration, *Plant Physiol. Comm*. **2010**, *46*, 64-66

[42] Zhu, K.; Wang, A.; Wu, J.; Yuan, F.; Guan, D. X.; Jin, C.; Zhang, Y.; Gong, C. Effects of nitrogen additions on mesophyll and stomatal conductance in Manchurian ash and Mongolian oak. *Scientific Reports*. **2020**, *10*(1), 10038. https://doi.org/10.1038/s41598-020-66886-x

[43] Wang, J.; Lu, X.; Zhang, J.; Ouyang, Y.; Qin, Z.; Zhao, B. Using golden apple snail to mitigate its invasion and improve soil quality: a biocontrol approach. *Environ Sci Pollut Res Int*. **2020,** *27*(13), 14903-14914. https://doi.org/10.1007/s11356-020-07998-9

[44] Rho, T.; Doty, Sharon.; Kim, Soo-Hyung. Endophytes Alleviate the Elevated CO2-dependent Decrease in Photosynthesis in Rice, Particularly under Nitrogen Limitation. *Journal of experimental botany*. **2019**, *71*(2), 707-718. https://doi.org/10.1093/jxb/erz440

[45] Ainsworth, E. A.; Rogers, A. The response of photosynthesis and stomatal conductance to rising [CO2]: mechanisms and environmental interactions. *Plant, Cell & Environment*. **2007**, *30*, 258-270. https://doi.org/10.1111/j.1365-3040.2007.01641.x

[46] Zhu, X. G.; Long, S. P., Ort, D. R. Improving photosynthetic efficiency for greater yield. *Annual Review of Plant Biology*. **2010**, *61*, 235-261. https://doi.org/10.1146/annurev-arplant-042809-112206

[47] Liu, Y.; Tikunov, Y.; Schouten, R.; Marcelis, L. F. M.; Visser, R.; Bovy, A. Anthocyanin Biosynthesis and Degradation Mechanisms in Solanaceous Vegetables: A Review. *Frontiers in Chemistry*. **2018**, *6*, 52. https://doi.org/10.3389/fchem.2018.00052

[48] Zhang, X.; Wei, J.; Huang, Y.; Shen, W.; Chen, X.; Lu, C.; Su, N.; Cui, J. Increased Cytosolic Calcium Contributes to Hydrogen-Rich Water-Promoted Anthocyanin Biosynthesis Under UV-A Irradiation in Radish Sprouts Hypocotyls. *Frontiers in Plant Science.* **2018**, *9*. https://doi.org/10.3389/fpls.2018.01020

[49] Xu, W.; Peng, H.; Yang, T.; Whitaker, B.; Huang, H.; Sun, J.; Chen, P. Effect of calcium on strawberry fruit flavonoid pathway gene expression and anthocyanin accumulation. *Plant Physiology and Biochemistry.* **2014**, *82*, 289-298. https://doi.org/10.1016/j.plaphy.2014.06.015

[50] Shin, D. H.; Choi, M.; Lee, H. K.; Cho, M.; Choi, S. B.; Choi, G.; Park, Y. Calcium dependent sucrose uptake links sugar signaling to anthocyanin biosynthesis in Arabidopsis. *Biochemical and Biophysical Research Communications*. **2013,** *430*(2), 634-639. https://doi.org/10.1016/j.bbrc.2012.11.100

[51] Ai, T. N.; Naing, A. H.; Arun, M.; Lim, S. H.; Kim, C. K. Sucrose-induced anthocyanin accumulation in vegetative tissue of Petunia plants requires anthocyanin regulatory transcription factors. *Plant Science.* **2016**, *252*, 144-150. https://doi.org/10.1016/j.plantsci.2016.06.021

[52] Kovinich, N.; Kayanja, G.; Chanoca, A.; Otegui, M. S.; Grotewold, E. Abiotic stresses induce different localizations of anthocyanins in Arabidopsis. *Plant signaling & behavior*. **2015**, *10*(7). https://doi.org/10.1080/15592324.2015.1027850

[53] Zhang, H.; Zhao, Y.; Zhu, J. K. Thriving under Stress: How Plants Balance Growth and the Stress Response. *Developmental cell*. **2020,** *55*(5), 529-543. https://doi.org/10.1016/j.devcel.2020.10.012

[54] Ghasemzadeh, A.; Jaafar, H.; Karimi, E.; Ibrahim, M. H. Combined effect of CO2 enrichment and foliar application of salicylic acid on the production and antioxidant activities of Anthocyanin, flavonoids and isoflavonoids from ginger. *BMC Complement Altern Med*. **2012**, *12*, 229. https://doi.org/10.1186/1472-6882-12-229

[55] Li, X.; Schmid, B.; Wang, F.; Paine, C. E. Net assimilation rate determines the growth rates of 14 species of subtropical forest trees. *PloS one*. **2016,** *11*(3), e0150644. https://doi.org/10.1371/journal.pone.0150644

[56] Gould, K. S.; Markham, K. R.; Smith R. H.; Goris J. J. Functional role of anthocyanins in the leaves of Quintinia serrata A. Cunn., *Journal of Experimental Botany*. **2000,** *51*(347), 1107-1115, https://doi.org/10.1093/jexbot/51.347.1107

ASEAN
Journal of Scientific and Technological Reports
Online ISSN:2773-8752

Research article

Artificial Intelligence-Based Approach for Chronic Kidney Disease Detection

Gogulamudi Pradeep Reddy[1], Duppala Rohan[2], Kasaraneni Purna Prakash[3], Yellapragada Venkata Pavan Kumar[4*], and Maddikera Kalyan Chakravarthi[5]

[1] Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal 576104, Karnataka, India
[2] School of Computer Science and Engineering, VIT-AP University, Amaravati 522241, Andhra Pradesh, India
[3] Department of Computer Science and Engineering, Siddhartha Academy of Higher Education, Deemed to be University, Vijayawada 520007, Andhra Pradesh, India
[4] School of Electronics Engineering, VIT-AP University, Amaravati 522241, Andhra Pradesh, India
[5] Engineering Department, College of Engineering and Technology, University of Technology and Applied Sciences, Muscat, Oman
* Correspondence: pavankumar.yv@vitap.ac.in

Citation:
Reddy, P.G.; Rohan, D.; Prakash, P.K.; Kumar, P.V.Y.; Chakravarthi, K.M. Artificial intelligence-based approach for chronic kidney disease detection. *ASEAN J. Sci. Tech. Report.* **2025**, *28*(5), e258012. https://doi.org/10.55164./ajstr.v28i5.258012.

Article history:
Received: February 22, 2025
Revised: July 10, 2025
Accepted: August 27, 2025
Available online: September 14, 2025

Publisher's Note:
This article is published and distributed under the terms of Thaksin University.

Abstract: Chronic Kidney Disease (CKD) is a long-term medical condition in which the kidneys gradually lose their ability to function properly. Early detection of CKD is crucial for preventing severe complications and improving patient outcomes. Traditionally, CKD diagnosis has relied on manual analysis of clinical parameters and laboratory tests, which often lack scalability and precision. Artificial Intelligence (AI), through machine learning algorithms, has transformed healthcare by enabling automated and accurate disease detection. Datasets play a pivotal role in developing AI-based diagnostic systems, as the quality and balance of data significantly influence model performance. The majority of existing research on CKD detection has focused on balanced datasets, where data samples are evenly distributed across classes, to recommend the most effective classifiers for detection. However, in real-world scenarios, datasets are often imbalanced, with minority classes underrepresented, leading to biased models and poor detection of critical cases. Therefore, adopting suitable techniques to handle these imbalances is necessary. In this context, this paper addresses the issue by evaluating the performance of various classifiers on both slightly imbalanced and severely imbalanced CKD datasets. Through comprehensive experimentation, the research identifies that Gradient Boosting Machine (GBM) demonstrates robust performance across both slightly and severely imbalanced datasets by achieving 99.25% ± 0.68 and 92.26% ± 2.23 testing accuracy, 100% and 90.79% ± 3.9 AUROC, 0.01 ± 0.01 and 0.39 ± 0.15 LR-, 64.98% and 84.54% ± 3.41 H – measure. This work emphasizes the need for adaptable classifiers that reflect real-world data, improving the reliability of AI-based CKD diagnosis.

Keywords: Artificial intelligence in healthcare; Chronic Kidney Disease (CKD); Imbalanced Datasets; Machine Learning for Disease Detection; Gradient Boosting Machine (GBM)

1. Introduction

Chronic Kidney Disease (CKD) is a progressive medical condition characterized by the gradual loss of kidney function over time, preventing the kidneys from effectively filtering waste, excess fluids, and toxins from the blood. This condition affects millions of people globally and poses a significant burden on healthcare systems. CKD is often referred to as a "silent disease" because symptoms typically do not appear until the disease has progressed to advanced stages, making early detection critical to prevent severe complications such as cardiovascular diseases or kidney failure. The causes of CKD are diverse, with diabetes and hypertension being the most common reasons. Diabetes, characterized by elevated blood sugar levels, can damage the tiny blood vessels in the kidneys, impairing their ability to filter blood. Similarly, high blood pressure puts excessive stress on these delicate vessels, leading to functional decline. Other causes include glomerulonephritis, an inflammatory condition affecting the kidney's filtering units; polycystic kidney disease, a genetic disorder characterized by cyst formation; recurrent kidney infections; and the long-term use of nephrotoxic medications, such as certain painkillers. These factors, alone or in combination, contribute to the progression of CKD. While CKD progresses silently in its initial stages, a range of symptoms may appear as kidney function deteriorates. These include persistent fatigue, swelling in the extremities or face due to fluid retention, nausea, vomiting, changes in appetite, and cognitive difficulties such as trouble concentrating. Changes in urination patterns, including increased frequency, altered color, or the presence of foamy urine, may also occur. Advanced stages may lead to severe complications, including electrolyte imbalances, anemia, and shortness of breath, significantly impacting the patient’s quality of life. In addition to medical causes, lifestyle factors such as a high-sodium diet, lack of physical activity, obesity, and smoking significantly increase the risk of CKD. Environmental factors, including exposure to heavy metals, contaminated water, and air pollution, have also been linked to kidney dysfunction, particularly in regions with poor environmental regulations.

CKD is categorized into 5 stages as shown in Figure 1, based on the glomerular filtration rate (GFR), a measure of how well the kidneys filter blood. In Stage 1, kidney function is normal but with mild damage, and the GFR is 90 mL/min or higher. Stage 2 indicates a mild decline in kidney function, with GFR between 60 and 89 mL/min. Stage 3 reflects a moderate decline, with GFR ranging from 30 to 59 mL/min. By Stage 4, the disease is severe, with GFR between 15 and 29 mL/min, and Stage 5 signifies end-stage renal disease (ESRD), where the GFR drops below 15 mL/min, necessitating dialysis or transplantation [1]. Several diagnostic methods are employed to detect and monitor CKD [2]. Blood tests, including measurements of serum creatinine and blood urea nitrogen, are used to assess kidney function. Urine tests are used to detect abnormalities such as proteinuria or hematuria, which can indicate kidney damage. Imaging techniques, such as ultrasound and CT scans, help identify structural abnormalities or obstructions, while kidney biopsies provide a detailed analysis of kidney tissue to determine the extent of damage [3]. Estimates of GFR are critical for staging CKD and monitoring its progression. However, these diagnostic methods have several limitations. Late diagnosis is common due to the asymptomatic nature of early-stage CKD. Some methods, like biopsies, are invasive and carry risks. Cost and accessibility remain barriers, particularly in resource-constrained settings, while manual interpretation of test results introduces subjectivity and potential errors.

Artificial Intelligence (AI) offers promising solutions to these challenges [4]. By leveraging advanced machine learning algorithms, AI systems can analyze large and complex datasets with precision, enabling the early detection of subtle patterns in clinical data. Additionally, Explainable AI (XAI) is growing as a crucial component in healthcare applications, addressing the "black-box" nature of traditional AI models [5]. The Internet of Things (IoT) can play a transformative role by enabling real-time data collection from patients [6]. IoT devices, such as wearable sensors, smart health monitors, and connected medical equipment, can continuously track vital parameters like blood pressure, blood sugar levels, heart rate, and kidney-related biomarkers [7-8]. These devices can transmit the collected data to centralized systems for analysis, allowing for early identification of abnormalities and potential disease progression. The IoT can also facilitate remote monitoring, enabling healthcare providers to track patients' health without requiring frequent hospital visits, which is especially beneficial for individuals in remote or underserved areas. These systems enhance diagnostic accuracy, minimize human error, and facilitate scalable solutions that can handle large patient populations. By leveraging the data collected, AI supports personalized treatment by identifying patient-specific risk factors and predicting disease progression. Moreover, AI-powered tools can address real-world challenges, such as imbalanced datasets, which often result in biased models and poor detection of critical cases [9-10]. In this view, this paper evaluates various classifiers for both slightly imbalanced [11] and severely imbalanced datasets [12].

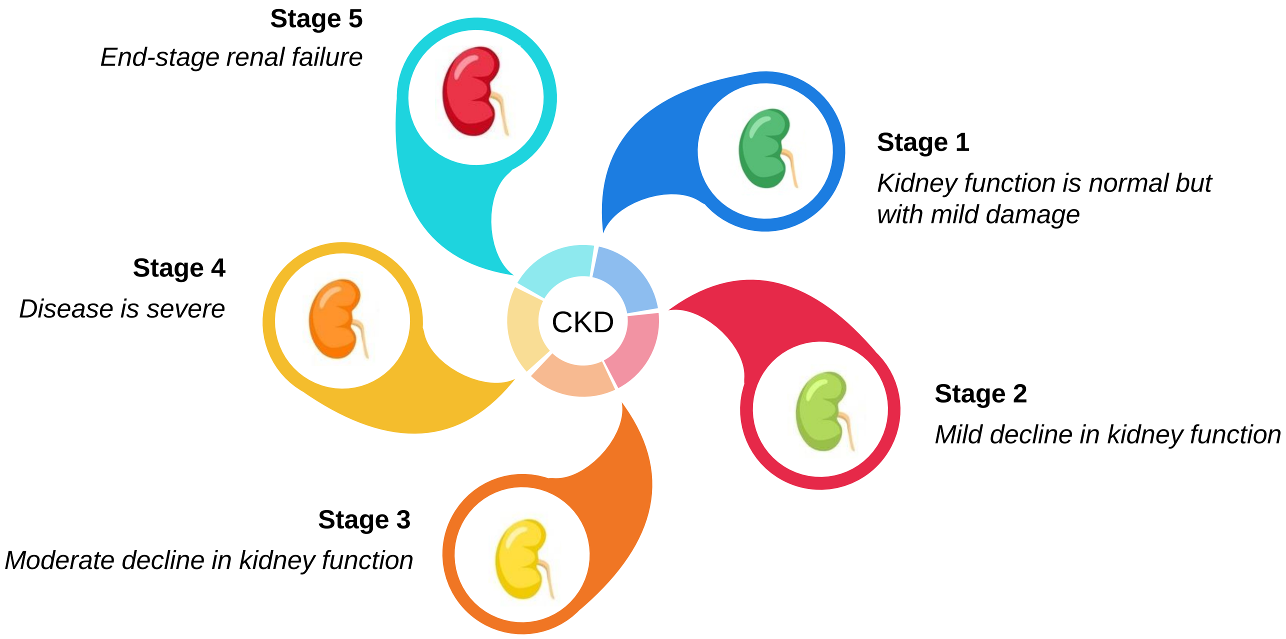


Figure 1. Various stages of CKD.

The rest of this paper is organized as follows: Section 2 provides an overview of related works, and Section 3 describes the proposed methodology in detail. Section 4 presents the results and discussion, analyzing the performance of various classifiers. Finally, Section 5 concludes the paper by summarizing the key findings and suggesting potential avenues for future research in the detection of CKD.

2. Related Work

This section provides a comprehensive review of the existing literature on the detection of CKD using various intelligence-based techniques. It highlights the progress made, key methods and approaches, and significant findings in this domain. A summary of key works in the literature is presented in Table 1. The prediction and early detection of CKD have gained significant attention in the research community due to CKD's high prevalence and associated complications. Various machine-learning models have been explored to enhance prediction accuracy and enable timely interventions [13]. Several comparative analyses have evaluated the performance of machine learning classifiers, such as Random Forest, SVM, XGBoost, and k-Nearest Neighbors [14]. Furthermore, deep learning frameworks, such as convolutional neural networks (CNNs), long short-term memory (LSTM) networks, and ensemble methods, have been proposed for CKD prediction [15-17]. The integration of advanced feature engineering techniques, including recursive feature elimination and optimization strategies like Bayesian tuning, has further enhanced the predictive capabilities of these models. Additionally, several studies have investigated the impact of data augmentation and resampling techniques. In addition to improving prediction accuracy, researchers have also focused on addressing challenges such as data privacy [18]. Privacy-preserving models using encrypted data through the Paillier homomorphic cryptosystem ensure secure predictive analysis without compromising patient confidentiality. Meta-heuristic feature selection techniques, including genetic and bat algorithms, have been used to optimize model performance and reduce computational complexity. Despite significant advances, most studies focus on balanced datasets, where each class is equally represented. However, real-world datasets often exhibit class imbalance, affecting the detection of minority cases. This imbalance affects the classifier's performance and restricts the identification of critical conditions. To address this limitation, this research evaluates classifiers on CKD datasets with both slight and severe class imbalance. The deployment of machine learning models into clinical practice has seen significant progress through frameworks like DEPLOYR [19], which enables seamless integration of real-time predictive models into electronic medical records, aligning model outputs with clinical workflows. Beyond technical integration, real-world performance monitoring is essential for finding valuable insights about model behavior post-deployment and identifying the need for ongoing surveillance [20]. The SHIELD-RT randomized study further addresses the practical challenges of deploying ML tools in a clinical setting, highlighting barriers such as clinician adoption, workflow alignment, and infrastructure readiness [21]. Multimodal AI frameworks have been proposed to support more generalizable and robust clinical applications by integrating diverse data sources, such as imaging, text, and structured data, into a unified system for decision support [22].

Table 1. Summary of various key works in CKD.

Ref.	Year	Objective	Methods/Approach	Relevance	Key Findings
[23]	2024	To classify and segment kidney MRI images for timely CKD detection and diagnosis.	Utilized Vision Transformers (EfficientNet_b1) for classification and ResNet18-Self-ONN-UNet++ for segmentation with CLAHE preprocessing and STAPLE post-processing.	Achieved 94.38% classification accuracy and superior segmentation performance, enhancing early CKD diagnosis using MRI data.	EfficientNet-b1 achieved 94.38% accuracy; the segmentation model achieved an IoU of 82.34% and a DS of 91.57%.
[24]	2024	To detect CKD using feature selection and hyperparameter tuning strategies for improved model performance.	Introduced Eurygaster Optimization Algorithm (EOA) with an ensemble of LSTM, BiGRU, and BiLSTM models. Applied SFLA for optimal hyperparameter selection.	Demonstrated superior CKD detection performance on benchmark datasets, emphasizing the importance of ensemble techniques.	EOAEDL-CKDD achieved optimal detection rates with improved accuracy and efficiency for early CKD detection.
[25]	2024	To explore non-invasive detection methods for medial vascular calcification (mVC) in CKD patients.	Analyzed peripheral pulse waveforms in the frequency domain, combined with traditional risk factors, to build predictive models using ensemble techniques.	Achieved 93% accuracy and an F-score of 0.96, highlighting the feasibility of non-invasive mVC detection in CKD management.	The ensemble model achieved an AUC ROC of 0.91, demonstrating the potential for non-invasive mVC detection in CKD patients.
[26]	2024	Introduces explainable AI for CKD diagnosis to address the limitations of "black-box" models and enhance interpretability.	Evaluated six machine learning algorithms, with extreme gradient boost (XGB) producing the best performance. SHAP and PDP were used to provide explainable predictions. Developed a graphical user interface (GUI) to assist healthcare professionals.	Pioneers the use of explainable AI in CKD diagnosis, aiding healthcare professionals in making informed decisions with interpretable insights.	XGB achieved the highest accuracy, and the SHAP and PDP methods provided transparent predictions, improving clinical utility and decision-making.

Table 1. Summary of various key works in CKD. (Continue)

Ref.	Year	Objective	Methods/Approach	Relevance	Key Findings
[27]	2023	Investigates the potential of various machine learning approaches for the early detection of CKD.	Predictive modeling was employed in conjunction with 12 machine learning classifiers, including XGBoost, to identify the most relevant features. The model evaluated 25 variables and identified the top 30% using predictive analytics.	Demonstrates the potential of predictive modeling and ML to enhance CKD diagnosis, reducing computational complexity and improving predictive accuracy.	XGBoost achieved the highest performance with 98.3% accuracy, precision, recall, and F1-score. Predictive modeling effectively identified crucial features for CKD diagnosis.
[28]	2023	To develop a novel hybrid deep learning network for early CKD detection and prediction.	Utilized DSCNN and Capsule Network with feature selection by Aquila Optimization. STOA optimized the DSCNN for classification.	Combines deep learning with IoT for enhanced CKD detection.	The proposed model outperformed state-of-the-art methods, demonstrating better classification accuracy with lower computational effort.
[29]	2021	To develop affordable machine learning models for early CKD diagnosis using selective clinical test attributes.	Evaluated models with k-fold cross-validation on optimized datasets derived from low-cost clinical test attributes.	Reduces diagnosis costs for CKD in resource-limited settings.	Optimized datasets and the random forest classifier achieved high accuracy, making it effective for CKD screening.
[30]	2020	To assess the applicability of ML techniques for CKD diagnosis in low-income, developing countries.	Conducted systematic reviews and experiments with ML techniques using k-fold cross-validation on CKD datasets.	Tailors CKD diagnostic approaches to developing countries.	The J48 decision tree achieved 95% accuracy, offering interpretable results suitable for low-resource settings.
[31]	2020	Automate CKD detection using salivary analysis and time-series data.	1-D Correlational Neural Network (CorrNN) and Bidirectional LSTM; tested with CKD sensing module.	Offers a novel salivary-based diagnostic tool for CKD.	Achieved an average accuracy rate of 98.08%, advancing CKD detection methodologies significantly.

3. Methodology

The experiment utilized two datasets: slightly imbalanced and severely imbalanced. The slightly imbalanced dataset was taken from the UCI Machine Learning Repository, and the severely imbalanced dataset was taken from Kaggle [11-12]. The slightly imbalanced dataset was collected over 2 months from a hospital, consisting of 24 features and a target variable. The target variable has two classes: 250 positive cases and 150 negative cases. The severely imbalanced dataset was collected from a Hospital in the United Arab Emirates, which consists of 21 features and a target variable with 56 positive and 435 negative cases. The dataset included the samples of 241 men and 250 women. The slightly imbalanced dataset consists of numerical and categorical columns. Imputation techniques were used to handle missing values. Missing numerical values were replaced with random samples from the same column. Missing categorical values were filled using the mode of the same column. Inconsistent entries in categorical columns were removed and replaced with clean values. At last, label encoding was used to convert categorical columns into numeric values, ensuring that each unique value in a column is mapped to a corresponding numeric label, and to prepare the dataset for machine learning and deep learning algorithms. SMOTE-Tomek was employed to handle the class imbalance in the datasets. SMOTE works by generating synthetic examples for the minority class based on feature space similarities between existing minority samples. This helps to reduce bias in the model's predictions by ensuring better representation of minority instances during training. Tomek links are pairs of instances from opposite classes that are each other's nearest neighbors. The presence of a Tomek link indicates ambiguity or noise at the class boundary. Removing such pairs helps sharpen the decision boundary and improve the classifier's ability to distinguish between classes. The 80% dataset was used for training the classifiers, and the 20% dataset was used for testing the classifiers' performance. Five-fold cross-validation was employed using the StratifiedKFold method from scikit-learn's model_selection module to ensure that each fold maintains approximately the same proportion of class labels. Each fold was indexed and tracked using a variable fold, initialized at 1 and incremented during each iteration. Additionally, a small constant epsilon (= 1e-10) was defined to avoid division by zero or numerical instability in metric calculations, particularly in cases where denominators could become very small.

A total of 16 machine learning and 5 deep learning classifiers were implemented. Machine learning classifiers include Gaussian Naïve Bayes (GNB), Bernoulli's Naïve Bayes (BNB), Complement Naïve Bayes (CNB), Multinomial Naïve Bayes (MNB), Decision Tree (DT), Extra Tree Classifier (ETC), K Nearest Neighbors (KNN), Support Vector Machine (SVM), Logistic Regression (LR), Stochastic Gradient Descent (SGD), Light Gradient Boosting Machine (LGBM), Histogram-based Gradient Boosting (HGB), Adaptive Boosting (ADB), Gradient Boosting Machine (GBM), Categorical Boosting (CB), and Extreme Gradient Boosting (XGB). Deep learning classifiers include Long Short-Term Memory (LSTM), Bidirectional LSTM (BiLSTM), Gated Recurrent Unit (GRU), Bidirectional GRU (BiGRU), and Recurrent Neural Network (RNN). LSTM, GRU, and RNN architectures share a similar sequential structure. Each classifier consisted of five layers with decreasing numbers of neurons – 128, 64, 32, 16, and 8 - utilizing the ReLU activation function. The five bidirectional layers in each model have decreasing numbers of units – 128, 64, 32, 16, and 8 - with the first layer specifying the input shape and utilizing the tanh activation, while subsequent layers use the ReLU activation. All five deep learning classifiers include a dropout layer with a dropout rate of 0.2 to prevent overfitting, as well as a final dense layer with a sigmoid activation function for predicting the disease. The proposed GBM model's hyperparameters are tabulated in Table 2.

Table 2. Parameters of GBM

S. No.	Hyper-Parameter	Value
1	loss	log_loss
2	learning_rate	0.1
3	n_estimators	200
4	subsample	0.8
5	criterion	friedman_mse
6	max_depth	7
7	min_samples_split	4
8	min_samples_leaf	2
9	max_features	log2
10	random_state	42
11	verbose	1

Thirteen metrics were considered for evaluating the classifiers. These metrics include testing accuracy, precision, sensitivity, F1 score, F2 score, specificity, balanced accuracy, Mathews Correlation Coefficient (MCC), AUROC, AUPRC, Likelihood Ratio Positive (LR+), Likelihood Ratio Negative (LR-), and H-measure. Testing accuracy can be used as a baseline metric, but it is often misleading in imbalanced datasets. So, other metrics must be considered for balanced evaluation. Precision and sensitivity were included to ensure the classifier does not over-predict the minority class and to measure how well the model captures the minority class. In imbalanced datasets, the F1 score is highly relevant, as it balances both sensitivity and precision. The F2 score gives more preference for sensitivity and is crucial when false negatives are more important than having false positives. While sensitivity focuses on the minority class, specificity evaluates how well the classifier identifies the majority class. Balanced accuracy measures the classifier's performance on both the minority and majority equally by averaging sensitivity and specificity. MCC is useful for imbalanced datasets as it evaluates the correlation between predictions and accurate labels by considering all true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). AUPRC is more useful than AUROC in imbalanced datasets. LR+ helps understand the strength of positive predictions and avoids over-reliance on the majority class bias. LR- ensures the classifier can eliminate cases as the majority class while correctly identifying most minority class cases. The H-measure is especially useful for imbalanced datasets because it allows for the weighting of false positives and negatives based on their respective importance in the specific problem domain. The H-measure was calculated using the trapezoidal rule and then normalized to the range of [0, 1]. For better interpretability, the results were expressed as percentages. It is considered an alternative to AUROC.

4. Results and Discussion

This section analyzes and compares the results based on the proposed methodology to provide key insights. It is divided into two subsections: the first presents the performance of classifiers on a slightly imbalanced CKD dataset, while the second examines their performance on a severely imbalanced dataset. A detailed discussion of these results is provided below.

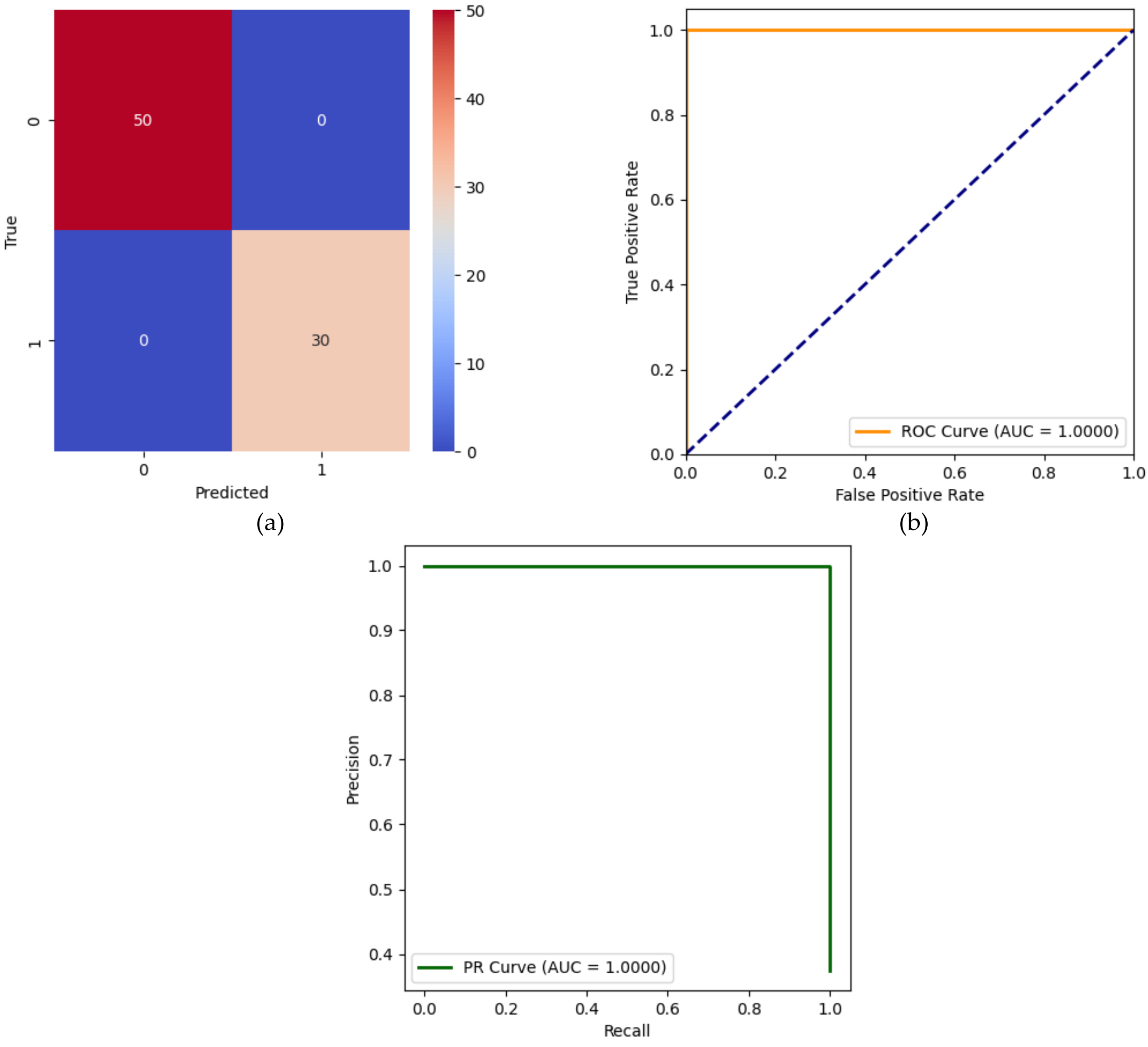
4.1. Results of classifiers on the slightly imbalanced dataset

GBM achieved the highest testing accuracy of 99.25% ± 0.68, surpassing all other models, including other high-performing classifiers such as HGB, CB, LSTM, and BiLSTM, each of which achieved accuracies of nearly 99%. GBM achieved a superior precision of 99.35% ± 1.44, outperforming models such as CB, LSTM, and BiLSTM, and is significantly higher than that of probabilistic models, including GNB, CNB, and MNB, which achieved precision scores below 88%, as shown in Table 3.

GBM achieved a sensitivity of 98.67% ± 1.83 and outperformed all models in specificity, achieving 99.6% ± 0.89%, which demonstrates its predictive power in classifying negative instances and minimizing false positives correctly. The F1 score of GBM was 98.99% ± 0.92, the highest among all classifiers, and achieved an F2 score of 98.79% ± 1.38. GBM also achieved the highest balanced accuracy of 99.13% ± 0.84, indicating its effectiveness across both positive and negative classes, especially in scenarios with class imbalance. An MCC of 98.42% ± 1.44 was also the highest observed, further reflecting the GBM's strong overall correlation between predicted and actual outcomes. GBM achieved an AUROC and AUPRC value of 100% ± 0.00, outperforming other models such as ETC, SVM, CB, LSTM, and BiLSTM, which reported AUROC and AUPRC scores above 99.8%. GBM achieved the highest H-measure of 64.98% ± 0.00, which was the highest among all models. GBM performed its best in the first fold, as shown in Figure 2.

Table 3. Results of classifiers on the slightly imbalanced dataset (Mean ± Variance).

S. No.	Classifier	Testing Accuracy (%)	Precision (%)	Sensitivity (%)	F1 Score (%)	F2 Score (%)	Specificity (%)	Balanced Accuracy (%)	MCC (%)	AUROC (%)	AUPRC (%)	LR+	LR-	H-measure (%)
1	GNB	94 ± 2.05	87.76 ± 4.81	98 ± 1.83	92.51±2.32	95.7 ± 1.17	91.6 ± 3.85	94.8 ± 1.54	88.03 ± 3.66	94.73 ± 1.61	86.61 ± 4.42	13.91 ± 6.63	0.02 ± 0.01	61.67 ± 0.91
2	BNB	97.75 ± 1.37	94.43 ± 3.21	100 ± 0.00	97.11±1.71	98.82 ± 0.71	96.40 ± 2.19	98.2 ± 1.10	95.41 ± 2.71	99.8 ± 0.38	99.66 ± 0.65	36.66 ± 18.25	0.00 ± 0.00	64.82 ± 0.18
3	CNB	92.5 ± 3.06	83.64 ± 5.67	100 ± 0.00	91.01±3.37	96.17 ± 1.51	88 ± 4.90	94 ± 2.45	85.79 ± 5.29	99.84 ± 0.15	99.76 ± 0.23	9.69 ± 4.41	0.00 ± 0.00	64.9 ± 0.07
4	MNB	92.5 ± 3.06	83.64 ± 5.67	100 ± 0.00	91.01±3.37	96.17 ± 1.51	88 ± 4.90	94 ± 2.45	85.79 ± 5.29	99.84 ± 0.15	99.76 ± 0.23	9.69 ± 4.41	0.00 ± 0.00	64.9 ± 0.07
5	DT	95.75 ± 2.59	93.04 ± 4.45	96 ± 4.35	94.43±3.44	95.35 ± 3.73	95.6 ± 2.97	95.8 ± 2.75	91.12 ± 5.48	96.49 ± 2.05	92.17 ± 4.69	$2.00\times10^9 \pm 4.47\times10^9$	0.04 ± 0.04	62.68 ± 1.47
6	ETC	98.25 ± 1.12	96.25 ± 3.42	99.33 ± 1.49	97.73±1.43	98.67 ± 0.95	97.6 ± 2.19	98.47 ± 0.87	96.39 ± 2.29	99.96 ± 0.04	99.94 ± 0.06	$3.93\times10^9 \pm 5.38\times10^9$	0.00 ± 0.01	64.96 ± 0.01
7	KNN	94.75 ± 2.40	88.39 ± 5.08	99.33 ± 1.49	93.48±2.91	96.89 ± 1.68	92 ± 3.74	95.67 ± 2.01	89.63 ± 4.62	98.39 ± 0.98	95.25 ± 2.95	18.58 ± 17.64	0.00 ± 0.01	63.82 ± 0.89
8	SVM	98.5 ± 1.05	96.79 ± 2.21	99.33 ± 1.49	98.03±1.37	98.81 ± 1.28	98 ± 1.41	98.67 ± 1.03	96.86 ± 2.19	99.93 ± 0.08	99.89 ± 0.14	$2.00\times10^9 \pm 4.47\times10^9$	0.00 ± 0.01	64.95 ± 0.04
9	LR	98 ± 1.43	95.62 ± 3.44	99.33 ± 1.49	97.41±1.82	98.55 ± 1.28	97.2 ± 2.28	98.27 ± 1.23	95.87 ± 2.90	99.93 ± 0.12	99.89 ± 0.18	$2.00\times10^9 \pm 4.47\times10^9$	0.00 ± 0.01	64.95 ± 0.05
10	SGD	98.25 ± 0.68	96.85 ± 3.13	98.67 ± 1.83	97.7±0.85	98.27 ± 0.91	98 ± 2.00	98.33 ± 0.41	96.36 ±1.37	99.91 ± 0.17	99.85 ± 0.29	$3.86\times10^9 \pm 5.29\times10^9$	0.01 ± 0.01	64.93 ± 0.08
11	LGBM	95.75 ± 2.59	93.9 ± 6.15	95.33 ± 5.58	94.41±3.45	94.91 ± 4.16	96 ± 4.24	95.67 ± 2.70	91.25 ± 5.39	99.76 ± 0.29	99.62 ± 0.48	$3.80\times10^9 \pm 5.20\times10^9$	0.04 ± 0.05	64.86 ± 0.14
12	HGB	99 ± 1.63	98.18 ± 4.07	99.33 ± 1.49	98.71±2.07	99.07 ± 1.36	98.8 ± 2.68	99.07 ± 1.36	97.96 ± 3.29	99.89 ± 0.20	99.83 ± 0.33	$7.93\times10^9 \pm 4.43\times10^9$	0.00 ± 0.01	64.93 ± 0.10
13	ADB	96.75 ± 2.59	94.8 ± 3.70	96.67 ± 3.33	95.72±3.42	96.28 ± 3.34	96.8 ± 2.28	96.73 ± 2.72	93.12 ± 5.50	96.73 ± 2.72	92.97 ± 5.36	$2.00\times10^9 \pm 4.47\times10^9$	0.03 ± 0.03	62.67 ± 1.82
14	GBM	99.25 ± 0.68	99.35 ± 1.44	98.67 ± 1.83	98.99±0.92	98.79 ± 1.38	99.6 ± 0.89	99.13 ± 0.84	98.42 ± 1.44	100 ± 0.00	100 ± 0.00	$7.86\times10^9 \pm 4.40\times10^9$	0.01 ± 0.01	64.98 ± 0.00
15	CB	99 ± 1.05	98.1 ± 2.81	99.33 ± 1.49	98.69±1.36	99.07 ± 1.12	98.8 ± 1.79	99.07 ± 0.92	97.92 ± 2.15	99.92 ± 0.18	99.87 ± 0.29	$5.93\times10^9 \pm 5.41\times10^9$	0.00 ± 0.01	64.94 ± 0.08
16	XGB	98 ± 1.43	95.62 ± 3.44	99.33 ± 1.49	97.41±1.82	98.55 ± 1.28	97.2 ± 2.28	98.27 ± 1.23	95.87 ± 2.90	99.87 ± 0.19	99.79 ± 0.30	$2.00\times10^9 \pm 4.47\times10^9$	0.00 ± 0.01	64.91 ± 0.09
17	LSTM	99 ± 1.05	98.1 ± 2.81	99.33 ± 1.49	98.69±1.36	99.07 ± 1.12	98.8 ± 1.79	99.07 ± 0.92	97.92 ± 2.15	99.93 ± 0.12	99.89 ± 0.19	$5.93\times10^9 \pm 5.41\times10^9$	0.00 ± 0.01	64.95 ± 0.05
18	BiLSTM	99 ± 1.05	98.1 ± 2.81	99.33 ± 1.49	98.69±1.36	99.07 ± 1.12	98.8 ± 1.79	99.07 ± 0.92	97.92 ± 2.15	99.93 ± 0.15	99.89 ± 0.24	$5.93\times10^9 \pm 5.41\times10^9$	0.00 ± 0.01	64.95 ± 0.07
19	GRU	98.75 ± 0.88	98.75 ± 2.80	98 ± 1.83	98.34±1.14	98.13 ± 1.20	99.2 ± 1.79	98.6 ± 0.80	97.38 ± 1.81	99.89 ± 0.14	99.83 ± 0.23	$7.80\times10^9 \pm 4.36\times10^9$	0.02 ± 0.01	64.93 ± 0.06
20	BiGRU	99 ± 1.05	98.1 ± 2.81	99.33 ± 1.49	98.69±1.36	99.07 ± 1.12	98.8 ± 1.79	99.07 ± 0.92	97.92 ± 2.15	99.91 ± 0.17	99.85 ± 0.28	$5.93\times10^9 \pm 5.41\times10^9$	0.00 ± 0.01	64.93 ± 0.08
21	RNN	98.5 ± 1.05	98.08 ± 2.82	98 ± 1.83	98.01±1.37	98 ± 1.34	98.8 ± 1.79	98.4 ± 0.98	96.85 ± 2.19	99.91 ± 0.10	99.85 ± 0.16	$5.86\times10^9 \pm 5.35\times10^9$	0.02 ± 0.01	64.93 ± 0.05



(a) (b)

(c)

Figure 2. (a) Confusion Matrix, (b) AUROC, (c) AUPRC of GBM on the slightly imbalanced dataset.

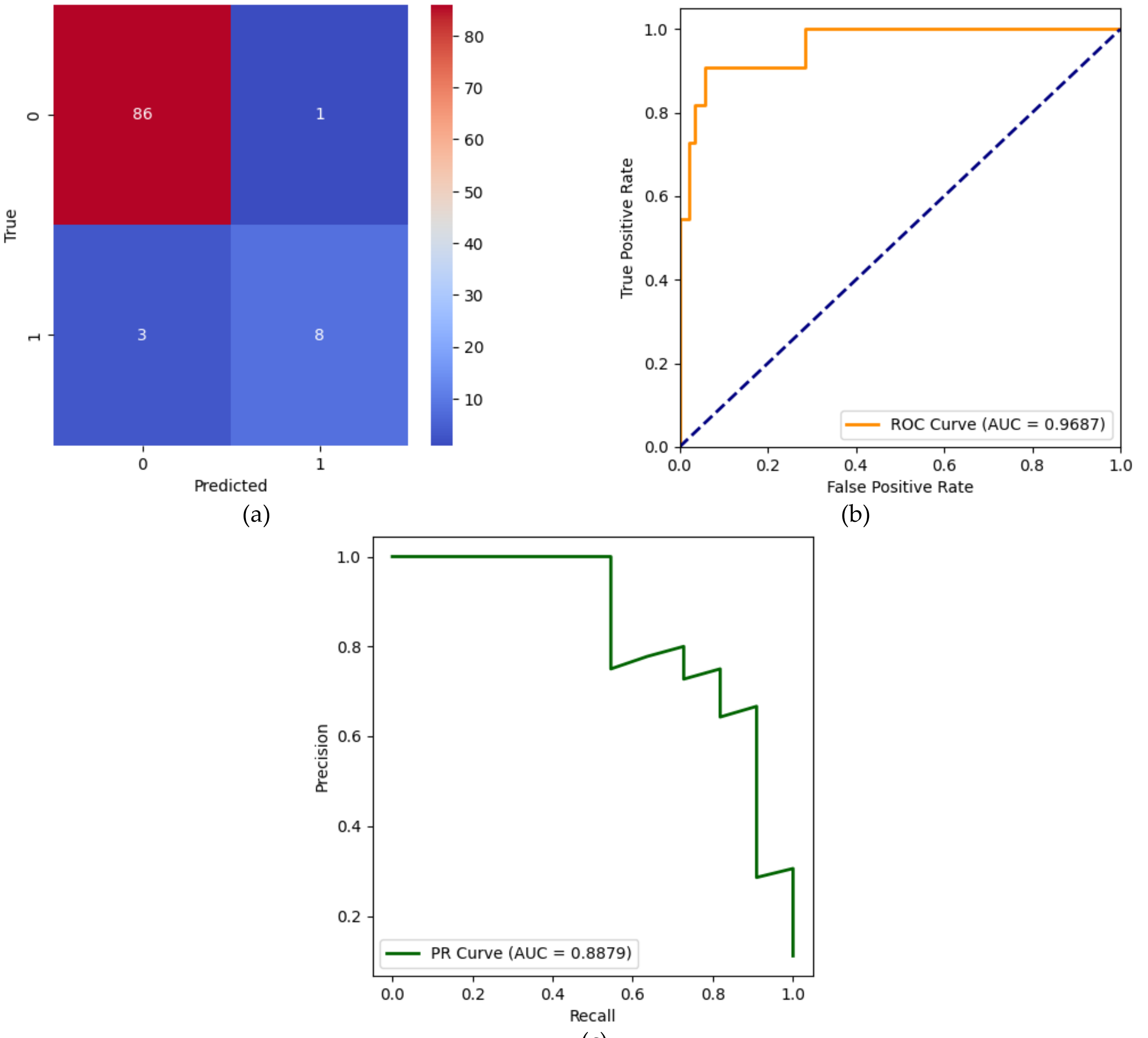
GBM demonstrated superior performance across nearly all evaluation metrics. It achieved the highest testing accuracy, precision, F1 score, specificity, balanced accuracy, MCC, AUROC, and AUPRC. GBM also recorded a high LR+ value, indicating its stronger ability to identify true positives with fewer false positives. Additionally, GBM achieved a lower LR- (0.01) compared to traditional models, demonstrating its ability to reduce false negatives effectively. It also achieved the highest H-measure of 64.98%, making it statistically strong.

4.2 Results of classifiers on a severely imbalanced dataset

GBM achieved the highest testing accuracy of 92.26% ± 2.23 among all models, outperforming ensemble methods like ADB, CB, and GRU, as well as deep learning architectures such as BiLSTM and RNN, reflecting its strong generalization capability on unseen data. GBM outperformed other models, including ADB and CB, in terms of precision, particularly in its ability to avoid false positives. GBM outperformed other models in terms of F1 score, followed closely by CB and GRU, highlighting GBM's ability to handle class imbalance. GBM stands out with a specificity of 96.09% ± 2.65, demonstrating its ability to correctly identify negative cases, which is crucial in domains where false positives can be costly. GBM achieved a balanced accuracy of 79.18% ±7.21 and an MCC of 61.09% ± 11.10, suggesting a decent correlation between predicted and actual labels, even in the presence of class imbalance. In terms of AUROC, GBM scored 90.79% ± 3.90 and achieved the highest AUPRC of 69.11% ± 11.41. GBM achieved an LR+ of 26.2 ± 22.45, the highest among all classifiers, indicating a powerful increase in the odds of disease presence when the test is positive, and an LR- of 0.39, reflecting a decreased probability of disease when the test is negative, as shown in Table 4. Figure 3 depicts the second fold results of GBM, which are the highest among the five folds. Table 3 shows that several models, including CB, LSTM, BiLSTM, and GRU, performed competitively on the slightly imbalanced dataset, achieving near-perfect metrics across accuracy, precision, and AUROC. However, GBM outperformed all classifiers, achieving high AUROC and AUPRC values of 100%, the highest precision of 99.35% ± 1.44, and the highest balanced accuracy of 99.13% ± 0.84. On the severely imbalanced dataset, shown in Table 4, GBM again emerged as the top performer, with the highest testing accuracy of 92.26% ± 2.23, specificity of 96.09% ± 2.65, and AUPRC of 69.11% ± 11.41, along with the highest LR+ of 26.2 ± 22.45. It also maintained strong sensitivity and F1 scores, indicating a reliable balance between correctly detecting both positive and negative cases. From these results, it is evident that GBM performed robustly across different imbalance scenarios and consistently outperformed other classifiers, making it the most suitable model to be proposed.

Table 4. Results of classifiers on severely imbalanced dataset (Mean ± Variance).

S. No.	Classifier	Testing Accuracy (%)	Precision (%)	Sensitivity (%)	F1 Score (%)	F2 Score (%)	Specificity (%)	Balanced Accuracy (%)	MCC (%)	AUROC (%)	AUPRC (%)	LR+	LR-	H-measure (%)
1	GNB	82.08 ± 1.89	34.92 ± 3.23	67.73 ± 16.73	45.81 ± 6.57	56.7 ± 11.11	83.91 ± 3.04	75.82 ± 7.35	39.49 ± 9.04	85.18 ± 4.95	58.13 ± 9.60	4.2 ± 0.68	0.38 ± 0.18	79.66 ± 4.21
2	BNB	82.89 ± 3.28	37.48 ± 5.22	76.67 ± 8.41	50.19 ± 5.73	63.21 ± 6.49	83.22 ± 3.69	79.94 ± 4.37	45.1 ± 7.01	89.73 ± 2.92	63.83 ± 6.34	4.74 ± 1.11	0.28 ± 0.10	83.62 ± 2.46
3	CNB	75.16 ± 4.28	28.56 ± 5.12	76.97 ± 13.14	41.55 ± 6.92	57.28 ± 9.18	74.94 ± 4.33	75.96 ± 7.15	35.82 ± 9.85	84.6 ± 3.00	54.72 ± 6.74	3.16 ± 0.84	0.3 ± 0.18	79.15 ± 2.52
4	MNB	75.16 ± 4.28	28.56 ± 5.12	76.97 ± 13.14	41.55 ± 6.92	57.28 ± 9.18	74.94 ± 4.33	75.96 ± 7.15	35.82 ± 9.85	84.6 ± 3.00	54.72 ± 6.74	3.16 ± 0.84	0.3 ± 0.18	79.15 ± 2.52
5	DT	83.92 ± 5.60	39.16 ± 13.35	62.42 ± 16.40	47.72 ± 13.98	55.4 ± 14.93	86.67 ± 5.11	74.55 ± 9.57	40.66 ± 16.81	80.37 ± 8.86	38.72 ± 13.03	5.53 ± 2.99	0.43 ± 0.20	74.47 ± 6.31
6	ETC	90.02 ± 2.20	55.38 ± 9.49	71.21 ± 17.73	61.47 ± 9.96	66.71 ± 13.70	92.41 ± 2.65	81.81 ± 8.28	57.03 ± 11.61	90.66 ± 3.38	63.74 ± 10.05	10.46 ± 4.79	0.30 ± 0.18	84.44 ±2.98
7	KNN	82.49 ± 2.89	34.14 ± 7.14	56.97 ± 12.38	42.58 ± 8.66	50.13 ± 10.38	85.75 ± 2.65	71.36 ± 6.49	34.62 ± 10.43	79.6 ± 3.69	37.02 ± 8.77	4.13 ± 1.21	0.5 ± 0.14	73.42 ± 4.62
8	SVM	84.94 ± 2.94	39.81 ± 6.68	58.64 ± 22.29	45.84 ± 9.63	52.21 ± 15.78	88.28 ± 5.29	73.46 ± 9.03	39.67 ± 10.89	86.21 ± 4.90	54.31 ± 15.76	5.29 ± 1.65	0.45 ± 0.22	80.58 ± 4.24
9	LR	85.75 ± 3.84	41.68 ± 12.37	62.27 ± 23.81	49.1 ± 14.94	55.92 ± 19.01	88.74 ± 4.25	75.5 ± 11.37	42.99 ± 17.45	87.02 ± 5.37	56.57 ± 17.81	6.12 ± 3.31	0.42 ± 0.26	81.29 ± 4.63
10	SGD	86.15 ± 2.09	42.29 ± 7.41	65.61 ± 21.93	50.97 ± 11.60	58.64 ± 16.50	88.74 ± 2.74	77.17 ± 10.15	45.15 ± 14.08	87.48 ± 6.11	57.47 ± 15.17	5.86 ± 1.67	0.38 ± 0.24	81.67 ± 5.36
11	LGBM	86.56 ± 2.28	44.78 ± 6.13	73.18 ± 20.35	54.78 ± 8.49	64.17 ± 13.85	88.28 ± 3.49	80.73 ± 9.07	50.05 ± 11.04	89.38 ± 5.54	63.69 ± 16.08	6.47 ± 1.79	0.29 ± 0.22	83.32 ± 4.82
12	HGB	89.62 ± 2.19	55.28 ± 9.53	62.27 ± 20.05	56.75 ± 10.85	59.58 ± 16.06	93.1 ± 3.15	77.69 ± 9.08	52.2 ± 11.41	87.96 ± 4.31	62.43 ± 15.49	10.36 ± 4.32	0.40 ± 0.20	82.09 ± 3.73
13	ADB	91.45 ± 1.13	66.16 ± 9.11	56.97 ± 12.38	59.8 ± 6.48	57.87 ± 10.27	95.86 ± 2.09	76.42 ± 5.36	56.08 ± 5.01	91.06 ± 3.78	63.24 ± 9.49	17.09 ± 8.75	0.44 ± 0.11	84.79 ± 3.38
14	GBM	92.26 ± 2.23	70.13 ± 14.06	62.27 ± 15.38	64.24 ± 11.31	62.78 ± 13.62	96.09 ± 2.65	79.18 ± 7.21	61.09 ± 11.10	90.79 ± 3.9	69.11 ± 11.41	26.2 ± 22.45	0.39 ± 0.15	84.54 ± 3.41
15	CB	91.04 ± 2.65	59.89 ± 9.52	65.91 ± 17.75	62.05 ± 12.10	64.15 ± 15.26	94.25 ± 2.15	80.08 ± 8.87	57.54 ± 13.44	90.91 ± 4.47	68.65 ± 14.15	12.48 ± 4.84	0.36 ± 0.18	84.65 ± 3.90
16	XGB	84.94 ± 3.56	39.4 ± 9.40	60.45 ± 24.03	46.75 ± 12.87	53.72 ± 18.04	88.05 ± 4.78	74.25 ± 10.87	40.41 ± 15.29	86.89 ± 5.49	56.72 ± 18.60	5.32 ± 2.13	0.44 ± 0.25	81.17 ± 4.74
17	LSTM	88.38 ± 3.53	49.06 ± 15.58	61.97 ± 25.45	53.72 ± 17.29	58.07 ± 21.26	91.72 ± 2.98	76.85 ± 12.54	48.35 ± 19.94	88.09 ± 7.28	60.43 ± 14.56	8.64 ± 5.02	0.41 ± 0.27	82.19 ± 6.42
18	BiLSTM	89.2 ± 2.78	52.08 ± 10.52	63.94 ± 17.42	56.97 ± 12.27	60.83 ± 14.91	92.41 ± 2.09	78.18 ± 8.85	51.53 ± 14.14	89.14 ± 4.73	65.35 ± 14.03	9 ± 3.32	0.39 ± 0.18	83.11 ± 4.20
19	GRU	91.24 ± 2.35	61.35 ± 11.26	63.94 ± 17.42	61.88 ± 11.86	62.93 ± 14.83	94.71 ± 1.92	79.33 ± 8.60	57.43 ± 13.35	89.33 ± 4.64	61.31 ± 13.94	13.88 ± 6.83	0.38 ± 0.18	83.28 ± 4.06
20	BiGRU	87.39 ± 4.17	47.77 ± 9.53	62.27 ± 15.38	53.18 ± 9.57	57.93 ± 11.79	90.57 ± 5.17	76.42 ± 6.97	47.3 ± 10.96	90.41 ± 4.03	63.11 ± 11.45	7.59 ± 3.08	0.41 ± 0.15	84.21 ± 3.44
21	RNN	90.63 ± 3.35	60.48 ± 19.75	54.85 ± 19.82	56.58 ± 16.92	55.33 ± 18.20	95.17 ± 2.49	75.01 ± 10.18	52.02 ± 19.13	88.23 ± 5.65	60.04 ± 11.88	17.92 ± 17.20	0.47 ± 0.21	82.33 ± 5.01



(a) (b)

(c)

Figure 3. (a) Confusion Matrix, (b) AUROC, (c) AUPRC of GBM on a severely imbalanced dataset.

5. Conclusions

Chronic Kidney Disease (CKD) is a pervasive health issue that requires timely and accurate detection to mitigate severe health complications. While Artificial Intelligence (AI) has significantly advanced CKD diagnostic capabilities, addressing data imbalance remains essential for achieving accurate results. Although many AI techniques perform well on balanced datasets, a thorough investigation of methods that excel in handling imbalanced datasets has not been conducted extensively. This research highlights the importance of addressing data imbalance to improve the performance of AI models in real-world scenarios, where instances of the minority class are often underrepresented. Through comprehensive experimentation, this research evaluates the performance of multiple classifiers on both slightly imbalanced and severely imbalanced CKD datasets. The findings highlight that GBM exhibits superior performance. With specificities of 99.6% ± 0.89 and 96.09% ± 2.65 for slightly and severely imbalanced datasets, respectively, the GBM model demonstrated significant performance in predicting true negatives. Furthermore, metrics such as AUROC (100% and 90.79% ± 3.9), LR+ (7.86 ×10^9 ± 4.40 ×10^9 and 26.2 ± 22.45), LR- (0.01 ± 0.01 and 0.39 ± 0.15), and H-measure (64.98% and 84.54% ± 3.41) highlight the model's efficacy in effectively distinguishing between classes.

5.1 Limitations

Although the proposed model has shown decent performance across all evaluation metrics, including a testing accuracy of 99.25%, an AUROC of 100%, and a balanced accuracy of 99.13%, some limitations remain. The datasets utilized for the experiment may not represent the demographic characteristics of patients across the globe. Model's generalizability may be affected by variations in diagnosis practices and demographic characteristics. There is a possibility of a decline in performance when the model is faced with missing or noisy data. The proposed model was trained using the past data, which might not have taken into consideration changes in clinical patterns over time.

5.2 Clinical and Healthcare Implications

The performance of GBM shows its potential in clinical practice for the early detection of CKD. It can be used as a support system for nephrologists and physicians for identifying high-risk patients at the early stages to reduce the likelihood of progression of the disease. GBM-based decision support tools can be used in electronic healthcare record (EHR) systems to identify abnormal patients during check-ups. The predictive power of the GBM makes it a promising system for reducing diagnostic delays and improving long-term patient outcomes. By continuously analyzing patient data, GBM can help medical practitioners identify patients who are suffering from CKD, leading to more effective disease management, which is beneficial in primary care settings.

5.3 Feasibility of Real-World Implementation

Due to the efficiency and relatively low computational demands at inference time, the practical implementation of GBM in real-world healthcare settings is promising and feasible. It can be integrated into clinical software systems, even in resource-limited environments such as rural or under-funded hospitals. It can also be operated on cloud-based platforms that connect to local health centers. Integration with IoT-based monitoring systems to enable continuous monitoring of CKD-relevant patterns through wearable or home-based devices can support long-term patient monitoring. In addition, GBM can be deployed using scalable architectures that support modular updates, making it easier to improve and adapt the system over time as more patient data becomes available to ensure that the model remains accurate and relevant across diverse patient populations and clinical scenarios.

5.4 Future Scope

Future research should include validating the proposed model using larger datasets and focus on integrating explainable AI (XAI) techniques that can help clinicians better understand the predictions, improving healthcare trust. Additionally, federated learning can be employed to train models across healthcare institutions without compromising patient privacy. The incorporation of attention mechanisms in sequence-based deep learning models like LSTM or Transformer architectures can allow the system to focus on clinically significant temporal patterns, such as gradual changes. Exploring hybrid models that combine machine learning techniques with deep learning architectures may yield both accuracy and interpretability, balancing computational efficiency with robustness.

Author Contributions: Conceptualization, G.P.R. and D.R.; methodology, K.P.P. and G.P.R.; software, K.P.P. and M.K.C.; validation, Y.V.P.K. and M.K.C.; formal analysis, Y.V.P.K.; investigation, D.R.; resources, G.P.R.; data curation, Y.V.P.K.; writing—original draft preparation, K.P.P.; writing—review and editing, D.R.; visualization, K.P.P.; supervision, Y.V.P.K.; project administration, M.K.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] Takkavatakarn, K.; Oh, W.; Cheng, E.; Nadkarni, G. N.; Chan, L. Machine Learning Models to Predict End-Stage Kidney Disease in Chronic Kidney Disease Stage 4. *BMC Nephrol* **2023**, *24*(1), 376. https://doi.org/10.1186/s12882-023-03424-7

[2] Nishanth, A.; Thiruvaran, T. Identifying Important Attributes for Early Detection of Chronic Kidney Disease. *IEEE Rev. Biomed. Eng.* **2018**, *11*, 208–216. https://doi.org/10.1109/RBME.2017.2787480

[3] Hodneland, E.; Keilegavlen, E.; Hanson, E. A.; Andersen, E.; Monssen, J. A.; Rorvik, J.; Leh, S.; Marti, H.-P.; Lundervold, A.; Svarstad, E.; Nordbotten, J. M. In Vivo Detection of Chronic Kidney Disease Using Tissue Deformation Fields From Dynamic MR Imaging. *IEEE Trans. Biomed. Eng.* **2019**, *66*(6), 1779-1790. https://doi.org/10.1109/TBME.2018.2879362

[4] Pradeep Reddy, G.; Rohan, D.; Kumar, Y. V. P.; Prakash, K. P.; Srikanth, M. Artificial Intelligence-Based Effective Detection of Parkinson's Disease Using Voice Measurements. *Engineering Proceedings* **2024,** *82*(1), 28. https://doi.org/10.3390/ecsa-11-20481

[5] Reddy, G. P.; Kumar, Y. V. P. Explainable AI (XAI): Explained. In 2023 IEEE Open Conference of Electrical, Electronic and Information Sciences (eStream); *IEEE*: Vilnius, Lithuania, **2023**; pp 1-6. https://doi.org/10.1109/eStream59056.2023.10134984

[6] Ma, F.; Sun, T.; Liu, L.; Jing, H. Detection and Diagnosis of Chronic Kidney Disease Using Deep Learning-Based Heterogeneous Modified Artificial Neural Network. *Future Generation Computer Systems* **2020**, *111*, 17-26. https://doi.org/10.1016/j.future.2020.04.036

[7] Ali, M. M.; Islam, M. S.; Uddin, M. N.; Uddin, Md. A. A Conceptual IoT Framework Based on Anova-F Feature Selection for Chronic Kidney Disease Detection Using Deep Learning Approach. *Intelligence-Based Medicine* **2024**, *10*, 100170. https://doi.org/10.1016/j.ibmed.2024.100170

[8] Chen, G.; Ding, C.; Li, Y.; Hu, X.; Li, X.; Ren, L.; Ding, X.; Tian, P.; Xue, W. Prediction of Chronic Kidney Disease Using Adaptive Hybridized Deep Convolutional Neural Network on the Internet of Medical Things Platform. *IEEE Access* **2020**, *8*, 100497-100508. https://doi.org/10.1109/ACCESS.2020.2995310

[9] Ramu, K.; Patthi, S.; Prajapati, Y. N.; Ramesh, J. V. N.; Banerjee, S.; Rao, K. B. V. B.; Alzahrani, S. I.; Ayyasamy, R. Hybrid CNN-SVM Model for Enhanced Early Detection of Chronic Kidney Disease. *Biomedical Signal Processing and Control* **2025**, *100*, 107084. https://doi.org/10.1016/j.bspc.2024.107084

[10] Bhaskar, N.; Manikandan, S. A Deep-Learning-Based System for Automated Sensing of Chronic Kidney Disease. *IEEE Sens. Lett.* **2019**, *3(10)*, 1–4. https://doi.org/10.1109/LSENS.2019.2942145

[11] Rubini, L.; Soundarapandian, P.; Eswaran, P. *Chronic Kidney Disease*. **2015**. https://archive.ics.uci.edu /dataset/336/chronic+kidney+disease (accessed 2025 18/02/2025)

[12] AI-Shamsi, S.; Regmi, D.; Govender, R. D.; Chicco, D. *Chronic kidney disease EHRs Abu Dhabi*. **2021**. https://www.kaggle.com/datasets/davidechicco/chronic-kidney-disease-ehrs-abu-dhabi (accessed 2025 18/02/2025)

[13] Norouzi, M.; Kahriman, E. A. A Machine Learning-Based Early Diagnosis Model for Chronic Kidney Disease Using SPegasos. *Netw Model Anal Health Inform Bioinforma* **2024**, *13*(1), 20. https://doi.org/10.1007 /s13721-024-00457-2

[14] Debal, D. A.; Sitote, T. M. Chronic Kidney Disease Prediction Using Machine Learning Techniques. *J Big Data* **2022**, *9*(1), 109. https://doi.org/10.1186/s40537-022-00657-5

[15] Saif, D.; Sarhan, A. M.; Elshennawy, N. M. Deep-Kidney: An Effective Deep Learning Framework for Chronic Kidney Disease Prediction. *Health Inf Sci Syst* **2023**, *12*(1), 3. https://doi.org/10.1007/s13755-023-00261-8

[16] Saif, D.; Sarhan, A. M.; Elshennawy, N. M. Early Prediction of Chronic Kidney Disease Based on Ensemble of Deep Learning Models and Optimizers. *Journal of Electrical Systems and Inf Technol* **2024**, *11*(1), 17. https://doi.org/10.1186/s43067-024-00142-4

[17] Al-Jamimi, H. A. Synergistic Feature Engineering and Ensemble Learning for Early Chronic Disease Prediction. *IEEE Access* **2024**, *12*, 62215–62233. https://doi.org/10.1109/ACCESS.2024.3395512

[18] Gogoi, P.; Valan, J. A. Privacy-Preserving Predictive Modeling for Early Detection of Chronic Kidney Disease. *Netw Model Anal Health Inform Bioinforma* **2024**, *13*(1), 16. https://doi.org/10.1007/s13721-024-00452-7

[19] Corbin, C.K.; Maclay, R.; Acharya, A.; Mony, S.; Punnathanam, S.; Thapa, R.; Kotecha, N.; Shah, N.H.; Chen, J.H. DEPLOYR: A Technical Framework for Deploying Custom Real-Time Machine Learning Models into the Electronic Medical Record. *Journal of the American Medical Informatics Association* **2023**, *30*, 1532-1542, doi: https://doi.org/10.1093/jamia/ocad114

[20] Thomas, L.; Hyde, C.; Mullarkey, D.; Greenhalgh, J.; Kalsi, D.; Ko, J. Real-World Post-Deployment Performance of a Novel Machine Learning-Based Digital Health Technology for Skin Lesion Assessment and Suggestions for Post-Market Surveillance. *Front. Med.* **2023**, *10*, 1264846, doi: https://doi.org/10.3389/fmed.2023.1264846

[21] Hong, J.C.; Eclov, N.C.W.; Stephens, S.J.; Mowery, Y.M.; Palta, M. Implementation of Machine Learning in the Clinic: Challenges and Lessons in Prospective Deployment from the System for High Intensity EvaLuation During Radiation Therapy (SHIELD-RT) Randomized Controlled Study. *BMC Bioinformatics* **2022**, *23*, 408, doi: https://doi.org/10.1186/s12859-022-04940-3

[22] Soenksen, L.R.; Ma, Y.; Zeng, C.; Boussioux, L.; Villalobos Carballo, K.; Na, L.; Wiberg, H.M.; Li, M.L.; Fuentes, I.; Bertsimas, D. Integrated Multimodal Artificial Intelligence Framework for Healthcare Applications. *npj Digit. Med.* **2022**, *5*, 149, doi: https://doi.org/10.1038/s41746-022-00689-4

[23] Islam, Md. S. B.; Sumon, Md. S. I.; Sarmun, R.; Bhuiyan, E. H.; Chowdhury, M. E. H. Classification and Segmentation of Kidney MRI Images for Chronic Kidney Disease Detection. *Computers and Electrical Engineering* **2024**, *119*, 109613. https://doi.org/10.1016/j.compeleceng.2024.109613

[24] Awad Yousif, S. M.; Halawani, H. T.; Amoudi, G.; Osman Birkea, F. M.; Almunajam, A. M. R.; Elhag, A. A. Early Detection of Chronic Kidney Disease Using Eurygasters Optimization Algorithm with Ensemble Deep Learning Approach. *Alexandria Engineering Journal* **2024**, *100*, 220-231. https://doi.org/10.1016/j.aej.2024.05.011

[25] Bialonczyk, U.; Debowska, M.; Dai, L.; Qureshi, A. R.; Söderberg, M.; Lindholm, B.; Stenvinkel, P.; Poleszczuk, J. Detection of Medial Vascular Calcification in Chronic Kidney Disease Based on Pulse Wave Analysis in the Frequency Domain. *Biomedical Signal Processing and Control* **2024**, *94*, 106250. https://doi.org/10.1016/j.bspc.2024.106250

[26] Dharmarathne, G.; Bogahawaththa, M.; McAfee, M.; Rathnayake, U.; Meddage, D. P. P. On the Diagnosis of Chronic Kidney Disease Using a Machine Learning-Based Interface with Explainable Artificial Intelligence. *Intelligent Systems with Applications* **2024**, *22*, 200397. https://doi.org/10.1016/j.iswa.2024.200397

[27] Islam, Md. A.; Majumder, Md. Z. H.; Hussein, Md. A. Chronic Kidney Disease Prediction Based on Machine Learning Algorithms. *Journal of Pathology Informatics* **2023**, *14*, 100189. https://doi.org/10.1016/j.jpi.2023.100189

[28] Venkatrao, K.; Kareemulla, S. HDLNET: A Hybrid Deep Learning Network Model With Intelligent IOT for Detection and Classification of Chronic Kidney Disease. *IEEE Access* **2023**, *11*, 99638-99652. https://doi.org/10.1109/ACCESS.2023.3312183

[29] Rashed-Al-Mahfuz, Md.; Haque, A.; Azad, A.; Alyami, S. A.; Quinn, J. M. W.; Moni, M. A. Clinically Applicable Machine Learning Approaches to Identify Attributes of Chronic Kidney Disease (CKD) for Use in Low-Cost Diagnostic Screening. *IEEE J. Transl. Eng. Health Med.* **2021**, *9*, 1-11. https://doi.org/10.1109/JTEHM.2021.3073629

[30] Sobrinho, A.; Queiroz, A. C. M. D. S.; Dias Da Silva, L.; De Barros Costa, E.; Eliete Pinheiro, M.; Perkusich, A. Computer-Aided Diagnosis of Chronic Kidney Disease in Developing Countries: A Comparative Analysis of Machine Learning Techniques. *IEEE Access* **2020**, *8*, 25407-25419. https://doi.org/10.1109/ACCESS.2020.2971208

[31] Bhaskar, N.; Suchetha, M.; Philip, N. Y. Time Series Classification-Based Correlational Neural Network With Bidirectional LSTM for Automated Detection of Kidney Disease. *IEEE Sensors J.* **2021**, *21*(4), 4811-4818. https://doi.org/10.1109/JSEN.2020.3028738

ASEAN
Journal of Scientific and Technological Reports
Online ISSN:2773-8752

Research article

Bridging the Yield Gap in Soybean Farming: Technical Efficiency and Key Determinants from a Stochastic Frontier Study in Chiang Mai Province, Thailand

Ploiphailin Tantiwit[1, 2], Pornsiri Suebpongsang[3], Buncha Somboonsuke[4], Prathanthip Kramol[5], Jittima Singvejsakul[6], Ayut Nissapa[7], and Rungrat Saeyang[8*]

[1] Faculty of Natural Resources, Prince of Songkla University, Songkhla, 90110, Thailand
[2] Faculty of Agriculture, Chiang Mai University, Chiang Mai, 50200, Thailand
[3] Faculty of Agriculture, Chiang Mai University, Chiang Mai, 50200, Thailand
[4] Faculty of Natural Resources, Prince of Songkla University, Songkhla, 90110, Thailand
[5] Faculty of Agriculture, Chiang Mai University, Chiang Mai, 50200, Thailand
[6] Faculty of Agriculture, Chiang Mai University, Chiang Mai, 50200, Thailand
[7] Faculty of Natural Resources, Prince of Songkla University, Songkhla, 90110, Thailand
[8] Faculty of Natural Resources, Prince of Songkla University, Songkhla, 90110, Thailand
[*] Correspondence: rungrat.s@psu.ac.th

Citation:
Tantiwit, P.; Suebpongsang, P.; Somboonsuke, B.; Kramol, P.; Singvejsakul, J.; Nissapa, A.; Saeyang, R. Technical Efficiency of Soybean Production in ChiangMai Province *ASEAN J. Sci. Tech. Report.* **2025**, *28*(5), e258940. https://doi.org/10.55164/ajstr.v28i5.258940

Article history:
Received: April 23, 2025
Revised: August 14, 2025
Accepted: August 27, 2025
Available online: September 14, 2025

Abstract: This study aimed to analyze the technical efficiency of soybean production and identify factors influencing technical inefficiency among soybean farmers in Chiang Mai Province. Cross-sectional data from 100 farmers across Chiang Dao, Hang Dong, Phrao, and Mae Rim districts were analyzed using stochastic frontier analysis (SFA) with a Cobb–Douglas production function integrated with a copula model, allowing for flexible modeling of the dependency between error components. The average technical efficiency score was 0.77, indicating moderate efficiency, with a 23% potential yield improvement through optimized resource use. Labor hours, fertilizer expenditure, and seed quantity significantly increased yields, while land ownership and farming experience reduced inefficiency; larger farm size was linked to higher inefficiency. Land rental emerged as a key factor for improving efficiency, suggesting policy support for agreements with rice farmers who leave land fallow after harvest. Given soybeans short growth cycle and low water demand, targeted interventions—such as farmer-to-farmer knowledge exchange, structured training on optimal input use, and scale-specific strategies from low-cost improvements for smallholders to precision agriculture for commercial farms could enhance productivity, resource efficiency, and profitability in northern Thailand.

Keywords: Copula function; Soybean production; Stochastic frontier analysis; Technical efficiency

1. Introduction

Soybean is a crucial food crop for Thailand's economy, serving as both a dietary staple and a primary raw material for the oil extraction, food processing, and animal feed industries. Thailand produces approximately 50,000–60,000 tons of soybeans annually [1], while domestic demand reaches nearly 3.4 million tons. Consequently, domestic production satisfies only around

1–2% of total demand, and the country relies heavily on imports, which account for over 98% of domestic consumption [2].

Thailand's food and animal feed industries are expanding and growing significantly while domestic soybean production continues to decline. As a result, the government and businesses relying on soybeans as a key raw material have greatly emphasized encouraging farmers to cultivate more leguminous crops. They have also supported initiatives to enhance the efficiency of legume crop production. Consequently, soybeans have been classified as a substitute crop for imports to increase domestic production, meet rising demand, and mitigate trade deficits resulting from the impact of free trade agreements (FTAs). Additionally, increasing domestic soybean production helps mitigate the risks associated with price volatility in other agricultural products, such as rainy-season feed corn and dry-season rice. This strategy aims to ensure food security and sustain local livelihoods, laying the foundation for sustainable agricultural production.

The uncertainty in global trade, particularly due to the ongoing conflict between Russia and Ukraine, has increased transportation costs driven by rising fuel prices and higher prices for oilseed crops [2–3]. Despite ongoing efforts in both the public and private sectors to promote soybean production through improved efficiency, cost reduction, and better land management, challenges persist. Soybeans are short-season crops with lower returns than other crops, requiring significant labor, especially during harvest. Additionally, the high cost of production input discourages farmers from growing soybeans. In years when water availability is sufficient, many farmers opt to grow alternative crops that require less labor and yield higher returns, such as off-season rice and feed corn [4]. This shift in cultivation patterns has led to a continuous decline in Thailand's soybean production, from 22,802 tons in 2020 to 20,802 tons in 2022. To address the shortage of soybean raw materials, Thailand has significantly increased its imports of soybeans and soybean products, with 3.20 million tons of soybeans imported in 2023 to meet domestic demand, particularly for food processing [5–6]. According to the Food and Agriculture Organization of the United Nations (FAO), in 2021, Thailand's average soybean yield was 1.65 tons per hectare, which was higher than the average yields in several Southeast Asian countries, such as Vietnam, Cambodia, Indonesia, the Philippines, and Myanmar, where yields ranged from 1.37 to 1.60 tons per hectare. However, Thailand's yield was still lower than that of Laos, which achieved an average of 1.80 tons per hectare. Compared to global soybean production, which has an average yield of 2.87 tons per hectare, Thailand's production remains relatively low in terms of yield per unit area and total output [7].

The upper northern region of Thailand is a major soybean cultivation area and a significant contributor to the country's soybean production. This region consists of six provinces: Mae Hong Son, Nan, Chiang Rai, Phrae, Chiang Mai, and Lampang, covering a total soybean cultivation area of 6,785.60 hectares, which accounts for 49.30% of the total national soybean planting area of approximately 13,761.76 hectares. In 2022, Chiang Mai province had a soybean cultivation area of over 663.2 hectares, representing 12.6% of the soybean growing area in the upper northern region or 5.3% of the total national soybean cultivation area [8]. Chiang Mai ranked third in soybean cultivation within the upper northern region, following Mae Hong Son and Nan. Additionally, Chiang Mai has been actively promoted as a province for soybean cultivation using the "Chiang Mai 60" variety, which was developed and improved by the Chiang Mai Field Crops Research Center. This variety is well-suited to the northern region's conditions and has been trial-planted in Chiang Mai to optimize production. The province was also selected as a model province for a project aimed at increasing soybean yields and reducing production costs [9]. During the 2019/20 planting season, Chiang Mai cultivated soybeans across nine districts, covering 796.32 hectares and producing 1,243 tons. The total production quantity decreased from the 2018/19 planting season, when the cultivated area was 1,063.2 hectares, and the total production was 1,744 tons. Mae Taeng district had the largest soybean growing area, followed by Mae Rim and San Pa Tong districts.

An analysis of Chiang Mai's average soybean yield per rai in the 2019/20 planting season revealed an average yield of 1.56 tons per hectare, which is lower than the national average of 1.68 tons per hectare. Chiang Mai's yield was also lower than Mae Hong Son’s (1.86 tons/ha) and Nan’s (1.74 tons/ha). Furthermore, Chiang Dao, Hang Dong, Phrao, and Mae Rim districts recorded relatively low average yields, ranging from 1.40 to 1.60 tons per hectare, which was lower than the yields of five other districts, where production ranged between

1.71 and 1.88 tons per hectare [9–10]. This suggests that narrowing the yield gap offers significant potential to enhance soybean production in Chiang Mai, particularly among smallholder farmers, and can be realized through improvements in technical efficiency. Several studies have quantified technical efficiency in soybean production using stochastic frontier analysis (SFA). For instance, Sharma et al. [11] explored soybean farmers in Madhya Pradesh, India, and reported a mean efficiency score of 0.72, indicating a 28% potential output gain through the adoption of best-practice technologies. They also identified labor and machine usage as underutilized inputs, while better contact with extension services significantly reduced inefficiencies. In northern Ghana, Etwire et al. [12] found an average technical efficiency of 0.53 (i.e., a 47 % output gap), with key determinants including farm location, farmers' age, and participation in an Agricultural Value Chain Mentorship Project. Musaba et al. [13] examined small-scale soybean producers in Zambia using a Cobb–Douglas SFA model. They found an average efficiency of 0.503, corresponding to a 49.7% potential output increase, with factors such as education, household size, market distance, extension access, and herbicide use affecting inefficiency levels. These findings demonstrate that while substantial inefficiencies persist, targeted institutional support and extension services can meaningfully improve productivity. Beyond India, Ghana, and Zambia, efficiency assessments have surfaced in other contexts. Alabi et al. [14] examined soybean producers in Northwest Nigeria using a stochastic frontier production model. They found an average efficiency of 0.5377, corresponding to a 46.23% potential output increase, with factors such as household size, age, education level, experience, number of extension contacts, and membership in cooperatives influencing efficiency levels. Based on research conducted by Otitoju and Arene [15], the translog stochastic frontier model is applied to examine the level and determinants of technical efficiency of soybean farmers in Benue State, Nigeria. The average technical efficiency score was 0.73, indicating that there is an opportunity to improve soybean productivity by 27%. The factors significantly impacting efficiency were gender, age, and experience. Besides that, a study in Indonesia's Tabanan Regency found that the average technical efficiency of soybean production was 0.77. The study also identified farming experience as a key driver of inefficiency, as noted by Rinaldi et al. [16].

In Thailand, numerous researchers have applied the stochastic frontier analysis framework to assess technical efficiency and its determinants across various agricultural commodities at both national and regional levels. For rice production, Puphoung [17] investigated farm-level efficiency and reported relatively high performance, with paddy field size, education, experience, participation in a farmer group, and debt identified as key positive determinants. Similarly, Somcom and Wana [18] found that irrigation management, farming experience, training participation, affiliation with agricultural organizations, and diversification of farming activities significantly enhanced the efficiency of rice farmers, while Rahman et al. [19] highlighted farm size as an influential factor affecting the level of technical inefficiency. For cassava, Sriwichai [20] identified land preparation practices and access to credit as critical determinants of efficiency differences among farmers. In chili production, Krasachat [21] employed a stochastic frontier production function approach to compare farms adopting Good Agricultural Practices (GAP) and those without GAP certification, concluding that GAP adoption significantly increased technical efficiency; Sornin and Athipanyakul [22] emphasized that the Cobb–Douglas SFA model is well suited for estimating the technical productivity of chili pepper production, with cultivated area, labor force, and input value identified as significant factors influencing output. Regarding soybeans, Phummuong [23] analyzed rainfed soybean production using stochastic frontier analysis (SFA) and found an average technical efficiency of 0.66. Key factors positively associated with efficiency included labor and herbicide use. Conversely, increased seed use, farm flooding, and excessive application of liquid fertilizer had adverse effects. The findings emphasize the importance of enhancing farmers' access to information and promoting crop rotation prior to planting soybeans to improve efficiency under rainfed conditions. Further, Lesak [24] employed the Stochastic Nonparametric Envelopment of Data (StoNED) approach to analyze soybean farmers operating in irrigated areas of Chiang Mai Province, reporting an average technical efficiency score of 0.620. The study further identified that higher educational attainment of the household head and female labor participation were positively and significantly associated with improvements in technical efficiency.

Previous research in Thailand and internationally has demonstrated the utility of Stochastic Frontier Analysis (SFA) in identifying technical efficiency and productivity gaps, as well as their determinants, across a range of crops. As a parametric approach, SFA can be employed to estimate a range of production functions, such as Cobb–Douglas, Translog, and others [25]. While existing work contributes valuable insights, notable methodological and contextual gaps persist. More recently, some studies have advanced this framework by integrating SFM with copula analysis, which enables the modeling of potential dependency between the two error components in the SFM structure, thereby offering greater flexibility and precision in efficiency estimation [26]. However, this methodological innovation has been applied primarily to rice production [25, 27], with limited attention to soybean production in Thailand, despite the crop's economic and nutritional importance, particularly in regions such as Chiang Mai Province, where farming systems are highly diverse. Building on these methodological advantages, this study applied a copula-integrated SFA framework to examine technical efficiency in soybean production. This approach allowed for potential dependence between inefficiency and random shocks, thereby enhancing the robustness and precision of efficiency estimates. Applied to Chiang Mai Province, it also addresses a notable gap in the literature by providing region-specific evidence on both the level and determinants of technical inefficiency in soybean farming. This study aimed to estimate the technical efficiency of soybean production in Chiang Mai Province using a copula-integrated stochastic frontier model, identify the socioeconomic, farm management, and resource-use factors influencing technical inefficiency, and provide targeted recommendations to enhance production efficiency. The results will inform farmers on how to improve productivity and input use, while offering policymakers evidence to design strategies for strengthening soybean production efficiency at both regional and national levels.

2. Materials and Methods

This study utilizes cross-sectional data from a survey on soybean cultivation in Chiang Mai during the 2020/21 production year. The research focuses on the production efficiency of soybean farms in four districts: Chiang Dao, Hang Dong, Phrao, and Mae Rim. These districts account for 50% of Chiang Mai's total soybean cultivation area and have an average yield per rai lower than the province's average.

The sample size was determined using the Taro Yamane formula [28] at a 90% confidence level. The formula used is:

$$n = \frac{N}{1 + N(e^2)}$$

Where:

- n = required sample size
- N = total population (965 soybean farmers in the selected districts)
- e = acceptable margin of error (0.1)

The calculation resulted in a sample size of 90.61. However, to enhance the reliability of the analysis, the study rounded the sample size to 100 farmers. A stratified random sampling method was employed, ensuring the proportion of farmers from each district was represented. Random selection was applied within each stratum, as soybean farmers in the selected districts share similar socioeconomic characteristics. Field data were collected between February and July 2022 using a structured questionnaire consisting of both open-ended and closed-ended questions. Subject matter experts reviewed the questionnaire to ensure content validity and subsequently pilot-tested it with 40 farmers in Mae Taeng District to assess clarity, appropriateness, and reliability. Feedback from the pilot was used to refine the wording and sequencing of the questions. Prior to the survey, four research assistants participated in a one-day training program covering interview techniques, ethical considerations, and the use of the questionnaire. The primary data collection was conducted mainly by the lead researcher (approximately 80% of interviews), with support from the trained assistants. Face-to-face interviews were conducted after obtaining informed consent from each respondent, resulting in a 100% response rate. To ensure data quality, daily checks were performed to verify the completeness and consistency of responses, and any discrepancies were clarified with respondents on the same day. The study employed the Stochastic Frontier Analysis (SFA) method to evaluate technical efficiency in soybean production. SFA is particularly suitable for agricultural contexts because production efficiency among farmers often exhibits high variance, and this method can separate the impacts of uncontrollable external random factors from those arising from inefficiencies in production practices [29]. In the SFA framework, the composite error terms consist of two components: v, representing external noise or randomness beyond the control of farmers, and u, representing inefficiency caused by production-related factors. Under the traditional distributional assumption, these two error components are assumed to be statistically independent [25–26, 29].

The first step of the analysis aimed to determine the most suitable production function specification for modeling soybean production in Chiang Mai Province. Two commonly used functional forms in efficiency studies — the Cobb-Douglas production function and the Translog production function —were considered. The Cobb–Douglas and Translog production functions are both widely applied within the stochastic frontier analysis (SFA) framework due to their complementary strengths in modeling production technology [30]. The Cobb–Douglas form is valued for its simplicity, ease of estimation, and interpretability, making it particularly suitable for farm-level studies with limited data or potential multicollinearity issues, although it assumes constant elasticities of substitution between inputs. In contrast, the Translog function offers greater flexibility by allowing elasticities of substitution to vary and by incorporating interaction terms between inputs, enabling it to capture more complex and non-linear production relationships. Employing both functional forms within SFA enables a comparative assessment, ensuring that the chosen specification best reflects the underlying production technology while striking a balance between model interpretability and flexibility [31].

Building on the model selection process, the study further incorporated the copula-based SFA approach as an extension to both the Cobb–Douglas and Translog specifications. The copula method describes joint multivariate distributions and integrates the two error components, allowing for greater flexibility in the stochastic frontier model. Unlike traditional assumptions that require v and u to be independent, the copula approach enables them to be dependent [29]. By explicitly modeling this dependence structure, the copula-integrated SFA framework can yield more robust and precise efficiency estimates, providing a richer understanding of the determinants of technical inefficiency in soybean production. Four models were compared: the Cobb-Douglas production function and the Translog production function, both with and without incorporating the copula approach. The model selection was based on the Bayesian Information Criterion (BIC), where the best model is the one with the lowest BIC value [32]. The best model has the lowest BIC value. BIC balances model fit and complexity by penalizing models with a large number of parameters, thus reducing the risk of overfitting. Compared to the Akaike Information Criterion (AIC), BIC applies a more substantial penalty for complexity, making it more likely to select simpler models, especially as the sample size increases [33–34]. Maximum Likelihood Estimation (MLE) was employed to estimate the coefficients of the variables [35], with all computations performed using RStudio.

The production factors used in the models were derived from literature reviews, including works by Etwire et al. [12], Lesak [24], Otitoju and Arene [15], Phummuong [23], Si and Wang [36], and Sriwichai [20]. The factors used in the stochastic frontier production function, along with their definitions and measurement units, were presented in Table 1. All continuous input and output variables were log-transformed using the natural logarithm to facilitate the interpretation of estimated coefficients as elasticities, thereby enabling the evaluation of proportional changes in output relative to proportional changes in each input. For variables containing zero observations (e.g., no pesticide application during the production cycle), a constant value of one was added prior to transformation to prevent undefined logarithmic values. This adjustment ensured the retention of all observations in the dataset while maintaining the integrity and consistency of the analysis.

Table 1. Variables used in the stochastic frontier production function

Variables	Definitions	Measurement Units
Seeds ($x1$)	Quantity of soybean seeds used for planting	kilograms per rai
Machinery ($x2$):	Use of agricultural machinery and equipment in soybean production, such as tractors, seeders, and threshers	working hours per rai
Human labor ($x3$):	Total amount of manual labor input, including family and hired labor, involved in various stages of soybean production	working hours per rai
Fertilizer ($x4$):	Total expenditure incurred for using chemical and/or organic fertilizers during soybean production per rai	Thai Baht (THB) per rai
Pesticides ($x5$):	Total expenditure incurred for using chemical and/or biological pesticides used to control pests, diseases, and weeds in soybean production	Thai Baht (THB) per rai

Remark: 1 rai = 0.16 hectare or 1 hectare = 6.25 rais

The Cobb-Douglas production function with the copula approach (1) and the Translog production function with the copula approach (2) can be represented as follows:

$$\text{Iny}_i = \beta_0 + \beta_1 \text{Inx}_{1i} + \beta_2 \text{Inx}_{2i} + \beta_3 \text{Inx}_{3i} + \beta_4 \text{Inx}_{4i} + \beta_5 \text{Inx}_{5i} + v_i - u_i \qquad (1)$$

$$\text{Iny}_i = \beta_0 + \beta_1 \text{In}\, x_{1i} + \beta_2 \text{In}\, x_{2i} + \beta_3 \text{In}\, x_{3i} + \beta_4 \text{In}\, x_{4i} + \beta_5 \text{In}\, x_{5i} + \frac{1}{2}\beta_{11}(\text{In}\, x_{1i})^2 + \frac{1}{2}\beta_{22}(\text{In}\, x_{2i})^2 + \frac{1}{2}\beta_{33}(\text{In}\, x_{3i})^2 + \frac{1}{2}\beta_{44}(\text{In}\, x_{4i})^2 + \frac{1}{2}\beta_{55}(\text{In}\, x_{5i})^2 + \beta_{12}\, \text{In}\, x_{1i}x_{2i} + \beta_{13}\, \text{In}\, x_{1i}x_{3i} + \beta_{14}\, \text{In}\, x_{1i}x_{4i} + \beta_{15}\, \text{In}\, x_{1i}x_{5i} + \beta_{23}\, \text{In}\, x_{2i}x_{3i} + \beta_{24}\, \text{In}\, x_{2i}x_{4i} + \beta_{25}\, \text{In}\, x_{2i}x_{5i} + \beta_{34}\, \text{In}\, x_{3i}x_{4i} + \beta_{35}\, \text{In}\, x_{3i}x_{5i} + \beta_{45}\, \text{In}\, x_{4i}x_{5i} + v_i - u_i \qquad (2)$$

Where:

y_i	represents the average soybean yield of farmers (kg per rai/ 1 rai = 0.16 hectare)
$x_{1i}, \ldots, x_{5i}$	are the production input variables
β_0	is the constant term
$\beta_1, \beta_2, \ldots, \beta_5$	are the parameters to be estimated
In	denotes the natural logarithm (logarithm base e)
v_i	represents errors due to uncontrollable external factors
u_i	represents errors due to internal factors that can be controlled
i	represents individual farmers cultivating soybeans, with values ranging from 1 to 100

The technical efficiency of soybean production among farmers in Chiang Mai can be estimated using a production function, where the efficiency score ranges from 0 to 1. If soybean production's technical efficiency (TE) equals 1.0, the producer has achieved the highest possible level of technical efficiency. Conversely, if the TE score is less than 1.0, it implies the presence of inefficiency [37]. According to Adinya et al. [38], the estimated technical efficiency levels can be classified into five categories:

- Very low (< 0.60)
- Low (0.61 - 0.70)
- Moderate (0.71 - 0.80)
- High (0.81 - 0.90)
- Very high (> 0.90)

This study employs a linear model estimated using multiple regression analysis with the Ordinary Least Squares (OLS) method to analyze factors affecting technical inefficiency in soybean production. This approach is chosen because multiple independent variables may influence the dependent variable, allowing for the simultaneous analysis of multiple factors [18]. The selection of variables for the analysis follows the stepwise regression method [39]. Six independent variables expected to influence technical inefficiency in soybean production are included, while each farmer's technical inefficiency (TI) is the dependent variable. The model is expressed in equation (3)

$$TI = \delta_0 + \delta_1 DOWN + \delta_2 DIRR + \delta_3 DLAB + \delta_4 DSUPP + \delta_5 EXP + \delta_6 SIZ + \varepsilon \quad (3)$$

Where:

TI = Technical inefficiency

δ_0 = Constant term

$\delta_1, \delta_2, \delta_3, \ldots, \delta_6$ = Parameters to be estimated

$DOWN$ = Dummy variable representing land ownership (1 = rented land, 0 = owned land)

$DIRR$ = Dummy variable indicating whether the land is in an irrigated area (1 = irrigated, 0 = non-irrigated)

$DLAB$ = Dummy variable representing labor type (1 = mixed labor, including household labor, 0 = hired labor only)

$DSUPP$ = Dummy variable indicating support from relevant agencies (1 = received support, 0 = did not receive support)

EXP = Experience in soybean farming (years)

SIZ = Total soybean cultivation area (rai)

$\boldsymbol{\varepsilon}$ = Error term

In equation (3), if an independent variable contributes to reducing technical inefficiency, its corresponding parameter will have a negative sign (-). Conversely, if an independent variable increases inefficiency, its parameter sign will be positive (+). The variables included in the analysis are based on a literature review of studies by Chubtong [40], Phummuong [23], and Si and Wang [36].

3. Results and Discussion

3.1 Technical Efficiency Analysis of Soybean Production

The analysis of basic data for the variables used in the stochastic frontier production function of soybean production in Chiang Mai Province, based on a sample of 100 farmers, revealed the following important statistics (Table 2): The average soybean yield per farmer was 310.13 kilograms per rai (approximately 1.94 tons per hectare). The average amount of soybean seed used was 16.68 kilograms per rai (0.10 tons per hectare). The average machine hours used for production were 3.12 hours per rai (19.5 hours per hectare). The average labor hours used for production were 12.39 hours per rai (77.45 hours per hectare). The average expenditure on fertilizer was 125 THB per rai (781.25 THB per hectare), and the average expenditure on pesticides was 99 THB per rai 618.75 THB per hectare. The improved soybean yields during the study period may be attributed to the favorable climatic conditions in the year of data collection, which were conducive to soybean cultivation. Additionally, rising domestic soybean prices are likely to encourage farmers to intensify their production efforts, resulting in increased productivity.

Table 2. Descriptive statistics of variables for the stochastic frontier production function

Variables	Average	S.D.	Minimum	Maximum
Production of soybean (Y) (kg/rai)	310.13	72.84	150.00	446.667
Amount of soybean seeds (X_1) (kg/rai)	16.68	3.40	12.00	30.00
Farm machinery used (X_2) (hours of work/rai)	3.12	3.49	1.00	22.00
Labour used (X_3) (hours of work/rai)	12.39	12.39	1.50	72.33
Cost of fertilizers (X_4) (THB/rai)	125.00	13.52	98.00	152.00
Cost of pesticides (X_5) (THB/rai)	99.00	15.25	82.00	143.00

Source: author’s calculation

Remark: 1 rai = 0.16 hectare or 1 hectare = 6.25 rais

Testing the functional form of the stochastic frontier production function for soybean production in Chiang Mai Province involved four alternative specifications: the Cobb-Douglas model, the Cobb-Douglas model integrated with the copula approach, the Translog model, and the Translog model integrated with the copula approach. As shown in Table 3, the Cobb-Douglas specification with the copula approach achieved the lowest BIC value (-210.9685), indicating the most favorable balance between model fit and parsimony. This outcome confirmed that the Cobb-Douglas form, when combined with the copula framework, is the most appropriate for capturing the production structure in the study area. Consequently, this specification was adopted for the subsequent estimation of production function coefficients and technical efficiency scores for soybean farmers.

Table 3. Bayesian Information Criterion (BIC) for model selection

	Cobb-Douglas	**Cobb-Douglas +Copula**	**Translog**	**Translog +Copula**
Bayesian Information Criteria (BIC)	-206.8829	-210.9685*	-190.6069	-184.1074

Source: author's calculation.
Note: * the lowest value of BIC

The estimation results of the coefficients for the stochastic frontier production function of soybean production by farmers in Chiang Mai province, using the Cobb-Douglas function combined with the copula model (Table 4), show that the variables for the number of labor hours used in production and the expenditure on fertilizer have a statistically significant impact on soybean yields at the 0.01 level. Meanwhile, the variable for the amount of soybean seed used in production has a statistically significant impact on soybean yields at the 0.05 level. All three factors have positive coefficients, indicating a positive relationship with soybean production. The estimation results of the coefficients for the stochastic frontier production function of soybean production by farmers in Chiang Mai province, using the Cobb-Douglas function combined with the copula model (Table 4), show that the variables for the number of labor hours used in production and the expenditure on fertilizer have a statistically significant impact on soybean yields at the 0.01 level. Meanwhile, the variable for the amount of soybean seed used in production has a statistically significant impact on soybean yields at the 0.05 level. All three factors have positive coefficients, indicating a positive relationship with soybean production. Table 4 presents the estimated parameters from the Cobb–Douglas stochastic production frontier with the copula specification. Holding other factors constant, the elasticity estimates indicate that a 1% increase in soybean seed use leads to a 0.286% increase in output, while a 1% increase in labor hours results in a 0.350% increase in output. Fertilizer expenditure also has a positive and statistically significant effect, where a 1% increase raises yield by 0.024%. In contrast, the number of farm machinery hours and pesticide expenditure do not significantly affect soybean yield in this sample, suggesting that there are limited productivity gains from additional spending on these inputs under current production conditions. The results highlight labor and seed use as the most influential inputs in soybean production, with labor showing the most significant elasticity. This suggests that interventions targeting labor productivity, such as training programs or labor-saving technologies, could yield substantial improvements in output. Likewise, optimizing seed use through improved seed quality or better planting techniques may further enhance productivity. Additionally, the sum of these elasticities is 0.752, indicating decreasing returns to scale in soybean production.

Table 4. Estimated parameters for the stochastic production frontier of Cobb-Douglas with Copula

Variables	**Coefficient**	**Standard error**
constant	1.684***	0.192
Amount of soybean seeds (X_1) (kg/rai)	0.286**	0.127
Farm machinery used (X_2) (hours of work/rai)	0.085	0.057
Labor used (X_3) (hours of work/rai)	0.350***	0.060
Cost of fertilizers (X_4) (THB/rai)	0.024***	0.007
Cost of pesticides (X_5) (THB/rai)	0.007	0.008

Source: author's calculation. Note: *** and ** represent statistical significance at 0.01 and 0.05 levels (or confidence at 99 and 95 percent levels), respectively.
Remark: 1 rai = 0.16 hectare or 1 hectare = 6.25 rais

The estimation results of the technical efficiency of soybean production by farmers in Chiang Mai province, as shown in Table 5, reveal that 13% of the farmers have a high level of technical efficiency (0.81-0.90). Meanwhile, 3% of farmers have a low level of technical efficiency (0.61-0.70), while the majority, over 84%, have their technical efficiency levels clustered in the range of 0.77-0.79. The value of technical efficiency indicates that most farmers have a medium level of technical efficiency (0.71-0.80). None of the sample farmers has achieved the highest technical efficiency (greater than 0.90). This distribution reflects a moderate level of efficiency overall, indicating scope for improvement in both production practices and technology adoption. The average technical efficiency of the farmers is 0.77, or 77.0%, indicating that there is still room for improvement in soybean production under the current technology available during the research period by 23.0%. The observed mean technical efficiency of 0.77 aligns with production theory under the stochastic frontier framework, where deviations from the frontier reflect suboptimal input allocation, managerial inefficiencies, or environmental constraints. The significant positive elasticities for labor, seed quantity, and fertilizer expenditure indicate that these inputs contribute directly to marginal output gains, consistent with neoclassical production theory, which posits that diminishing returns are gradual within the optimal input range. The lack of significance for machinery hours and pesticide expenditure suggests either overutilization or a mismatch between their application and yield response.

Table 5. Distribution of technical efficiency of soybean production in Chiang Mai province

Technical efficiency level		Sample (n=100)	Percentage (%)
Very low	(<0.60)	0	0
Low	(0.61-0.70)	3	3
Medium	(0.71-0.80)	84	84
High	(0.81-0.90)	13	13
Very High	(>0.90)	0	0
Total		100	100
Average = 0.77 Median = 0.77 Maximum = 0.81 Minimum = 0.66 S.D. 0.02			

Source: author's calculation.

Compared with previous soybean studies, Chiang Mai's TE is higher than findings from several African contexts (0.53) in northern Ghana Etwire et al. [12], 0.503 in Zambia Musaba et al. [13], and 0.5377 in North West Nigeria Alabi et al. [14], and also exceeds the 0.72 reported for Madhya Pradesh, India, Sharma et al. [11]. It is similar to Indonesia's Tabanan Regency (0.77) [16] and somewhat higher than Benue State, Nigeria (0.73) [15]. Within Thailand, Chiang Mai's TE surpasses earlier soybean estimates under rainfed conditions (0.66) Phummuong [23] and a StoNED-based study in irrigated Chiang Mai (0.620) Lesak [24]. These differences are plausible given the context and method. First, agro-ecological and institutional conditions vary: studies with lower TE often report constraints related to extension access, market distance, input quality, or household characteristics (e.g., education, age, cooperative membership), while programs that strengthen extension and input use tend to reduce inefficiency [11-14]. Second, methodological choices matter: Chiang Mai's estimates come from a Cobb–Douglas, copula-integrated SFA, whereas some comparators use conventional SFA (assuming independent errors) or StoNED Lesak [24]. Relaxing the independence assumption can yield more precise efficiency measures. Third, temporal and regional differences in seed varieties, input price environments, and recent extension initiatives can lead to an upward shift in TE. Consistent with our results, labor and seed have the largest elasticities, while machine hours are not statistically significant. This pattern aligns with small, fragmented plots, where careful labor/seed management dominates gains, and mechanization advantages remain limited.

3.2 Analysis of Factors Affecting Technical Inefficiency in Soybean Production

The results of the multiple regression analysis examining the factors influencing technical inefficiency in soybean production indicated that the model is appropriate, as evidenced by a coefficient of determination (R^2) of 0.470. This value suggested that all independent variables in the model, including the land ownership status, labor hiring status, irrigation zone, irrigation method, farming experience, total soybean cultivation area, and other related factors, collectively explain 47.0% of the variation in technical inefficiency. The remaining 53.0% of the variation was likely due to other factors not captured by the model, such as physical, biological, economic, or farm management variables associated with soybean production. The analysis revealed an F-statistic of 13.75 with a p-value of 0.0000, less than or equal to 0.05, indicating that the independent variables significantly explain the variation in the dependent variable at the 5% significance level. Furthermore, the Durbin-Watson statistic of 1.86, which is close to the benchmark value of 2, suggests that autocorrelation is absent among the residuals. Although the model has a relatively low R^2 value (Table 6), it is still considered suitable for explaining the relationship. At the 99% confidence level, the factors that significantly influence technical inefficiency in soybean production are land ownership status (1 = rented land, 0 = owned land), farming experience (in years), and the total area of soybean cultivation (in rai) (Table 6).

Table 6. Estimation of a Technical Inefficiency Model for Soybean Production in Chiang Mai Province.

Variable		**Coefficient**	**Standard Error**	**T-stat**
Constant		0.242***	0.006	43.575
Dummy of land ownership (DOWN)		-0.011***	0.004	-2.683
Dummy of Irrigation area (DIRR)		-0.004	0.005	-0.833
Dummy of labor involvement (family and hired) (DLAB)		-0.001	0.004	-0.329
Dummy of support from institution (DSUPP)		-0.006	0.005	-1.153
Experience of farmers in soybean production (EXP)		-0.001***	0.000	-5.485
Soybean land area (SIZ)		0.002***	0.000	4.210
R	0.686		RMSE	0.020
R-Squared	0.470		Coef. Var	9.007
Adj R-Squared	0.436		MSE	0.000
F-Statistics	13.755		MAE	0.015
Durbin-Watson	1.865			

Notes: *** represents statistically significant at 0.01 level (or confidence at 99 percent level), F-Statistic = F [6, 100] = 13.755 (p-value = 0.0000)

The regression analysis revealed that land ownership has a statistically significant effect on technical inefficiency in soybean production, with a coefficient of -0.011. Specifically, a 1% increase in the proportion of land rented for soybean cultivation is associated with a 0.011% reduction in technical inefficiency, indicating an improvement in technical efficiency. One possible explanation is that farmers who rent land face higher monetary costs, which may incentivize them to plan their production more efficiently and optimize the use of available inputs in order to maximize income through increased marketable output. The higher average yield per rai among farmers who rent land (324.80 kg/rai or 2.03 tons/ha) compared to those who own land (294.92 kg/rai or 1.84 tons/ha) further supports this observation. This result is consistent with the findings of Chubtong [40] and Donkor and Owusu [41], who argue that renting land tends to reduce technical inefficiency. Specifically, renting land may incentivize farmers to utilize resources more effectively, thereby increasing agricultural productivity. The analysis showed that experience in soybean cultivation significantly reduces technical inefficiency, with a coefficient of -0.001. Specifically, a 1% increase in a farmer’s experience in soybean production is associated with a 0.001% reduction in technical inefficiency. This finding is consistent with the research of Omar and Fatah [42], who demonstrated that farmers' experience in crop cultivation positively influenced their decision-making in managing coconut production in Johor, Malaysia. Similarly, studies by Alabi et al. [14], Puphoung [17], and Somcom and Wana [18] found that increased experience enables farmers to address challenges and obstacles during production more effectively. Therefore, farmers with more years of experience in soybean cultivation are typically more skilled and better able to understand production conditions. Their experience enables them to select appropriate technologies and gather knowledge from various sources, ultimately improving production efficiency. In addition, the size of the cultivated area significantly contributes to an increase in technical inefficiency in soybean production, with a coefficient of 0.002. Specifically, a 1% increase in the area dedicated to soybean cultivation is associated with a 0.002% increase in technical inefficiency. This relationship may arise from farmers' difficulties in effectively managing larger areas, leading to reduced management quality and production efficiency. These findings align with the research of Ajibefun et al. [43] and Puphoung [17], who observed that larger agricultural areas are associated with higher levels of technical inefficiency in production. Therefore, farmers should focus on comprehensive production management across the cultivated area by planning and allocating time to thoroughly monitor all areas. This approach will help increase the technical efficiency of soybean production.

4. Conclusions

This study employed the stochastic frontier analysis (SFA) framework, selecting the Cobb-Douglas production function integrated with a copula model as the most appropriate specification for analyzing the technical efficiency of soybean production by farmers in Chiang Mai Province. The methodological choice enabled a more flexible treatment of the potential dependency between inefficiency effects and random shocks, thereby improving the robustness of the efficiency estimates. The results revealed an average technical efficiency score of 0.77, implying a 23 % potential yield gain if farmers optimize input use under the existing technology set. The findings align with production economics principles, where deviations from the frontier reflect suboptimal input allocation and managerial constraints. Positive and significant elasticities for labor, seed use, and fertilizer expenditure confirm that these inputs remain within the increasing returns portion of the production function, indicating scope for productivity gains without breaching diminishing returns. The insignificance of machinery and pesticide expenditure suggests possible overuse or inefficient application, supporting the resource misallocation hypothesis. When benchmarked against international soybean studies, such as those by Otitoju and Arene [15] in Nigeria, Sharma et al. [11] in India, and Rinaldi et al. [16] in Indonesia, Chiang Mai's TE level is moderate and broadly comparable. However, contextual differences in determinants emerge: while African and South Asian contexts often emphasize education, access to extension services, and credit availability, the results highlight the roles of land ownership, farming experience, and cultivated area size. Land rental emerged as a key factor reducing inefficiency, suggesting policy support for agreements with rice farmers who leave land fallow after harvest could expand soybean cultivation. Given soybeans' short growth cycle and low water demand, such strategies align with their role as a secondary crop in Thai farming systems. Further, farmer experience was also critical, highlighting the need for structured knowledge exchange, targeted training, and improved access to agronomic information. Policies should facilitate land rental markets, strengthen farmer-to-farmer mentoring, and expand site-specific training on seed density, labor management, and fertilizer application. Interaction effects suggest that bundled interventions, such as combining improved seed use with labor training, can yield greater efficiency gains than isolated measures. Additionally, practical implications vary by scale: smallholders benefit from low-cost improvements supported by group-based extension; medium-scale farmers from semi-mechanization and input calibration; and commercial-scale farmers from integrating precision agriculture and digital monitoring. By addressing these areas, stakeholders can bridge existing efficiency gaps, enhance resource use, and improve the resilience and profitability of soybean production systems in northern Thailand.

5. Acknowledgements

I want to express my sincere gratitude for the financial support and research funding provided by the Graduate School of Prince of Songkla University, the Graduate School of Chiang Mai University, and the Master's Degree Joint Program under the Participatory and Integrative Support for Agricultural Initiative (PISAI), supported by the ERASMUS+ Capacity Building in Higher Education Programme.

Author Contributions: Initiatives and Hypotheses: Ploiphailin Tantiwit, Rungrat Saeyang, Pornsiri Suebpongsang. Research Implementation: Involvement in the design of experiments, testing of measurement tools, data collection methods, and criteria: Ploiphailin Tantiwit, Rungrat Saeyang, Prathantip Kramol, Pornsiri

Suebpongsang. Data collection, data analysis, interpretation: Ploiphailin Tantiwit, Rungrat Saeyang, Ayut Nissapa, Jittima Singvejsakul. Critique of results, presentation, comparison with conclusions or existing knowledge or theories: Ploiphailin Tantiwit, Rungrat Saeyang, Ayut Nissapa. Contribution to manuscript writing: Ploiphailin Tantiwit, Rungrat Saeyang, Ayut Nissapa, Pornsiri Suebpongsang.

Funding: Support for research funding: Graduate School, Prince of Songkla University, Graduate School of Chiang Mai University, Master's degree joint program under the Participatory and Integrative Support for Agricultural Initiative PISAI under the support of the ERASMUS+ Capacity Building in Higher Education Programme.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] USDA. *Oilseeds and products annual: Thailand*; Report No. TH2025-0009. Foreign Agricultural Service, U.S. Department of Agriculture, **2025**. https://apps.fas.usda.gov/newgainapi/api/Report/DownloadReportByFileName?fileName=Oilseeds%20and%20Products%20Annual_Bangkok_Thailand_TH2025-0009 (accessed Aug 16, 2022).

[2] Office of Agricultural Economics. *Situation of important agricultural products and trends in 2023*; Ministry of Agriculture and Cooperatives, **2022**. https://www.oae.go.th/assets/portals/1/files/jounal/2565/231225652566.pdf (accessed Jan 21, 2023). (in Thai).

[3] Department of Agriculture. *Research and development on soybean*; Ministry of Agriculture and Cooperatives, **2025**. https://www.doa.go.th/research/attachment.php?aid=2246 (accessed Nov 19, 2022). (in Thai).

[4] Office of Agricultural Economics. *Agricultural economic information by product in 2021*; Ministry of Agriculture and Cooperatives, **2022**. https://www.oae.go.th/assets/portals/1/files/jounal/2565/commodity2564.pdf (accessed Dec 14, 2023). (in Thai).

[5] Office of Agricultural Economics. *Agricultural economic information by product for the year 2022*; Ministry of Agriculture and Cooperatives, **2023**. https://www.oae.go.th/assets/portals/1/files/jounal/2566/commodity2565.pdf (accessed Mar 14, 2024). (in Thai).

[6] Office of Agricultural Economics. *Situation of important agricultural products and trends in 2024*; Ministry of Agriculture and Cooperatives, **2023**. https://www.oae.go.th/assets/portals/1/files/jounal/2566/trend2567.pdf (accessed Feb 3, 2024). (in Thai).

[7] FAOSTAT. Crop and livestock products. Food and Agriculture Organization of the United Nations. https://www.fao.org/faostat/en/#data/QCL (accessed Mar 21, 2025). (in Thai).

[8] Office of Agricultural Economics. *Agricultural statistics of Thailand 2022*; Ministry of Agriculture and Cooperatives, **2023**. https://www.oae.go.th/assets/portals/1/files/jounal/2566/yearbook2565.pdf (accessed Mar 14, 2024). (in Thai).

[9] Department of Agricultural Extension. *Crop production data report*; Ministry of Agriculture and Cooperatives, **2025**. https://production.doae.go.th/service/data-state-location/index (accessed Jan 26, 2025). (in Thai).

[10] Office of Agricultural Economics. *Agricultural economic information by product in 2021*; Ministry of Agriculture and Cooperatives, **2021**. https://www.oae.go.th/assets/portals/1/files/jounal/2565/commodity2564.pdf (accessed Jan 29, 2022). (in Thai).

[11] Sharma, P. U.; Dupare, B. U.; Patel, R. M. Technical efficiency of soybean production in Madhya Pradesh: a stochastic Frontier Approach. *Soybean Research* **2016**, *14* (1), 68–77.

[12] Etwire, P. M.; Martey, E.; Dogbe, W. Technical efficiency of soybean farms and its determinants in Saboba and Chereponi districts of northern Ghana: A stochastic frontier approach. *Sustainable Agriculture Research* **2013**, *2*(4), 106–116. https://doi.org/10.5539/sar.v2n4p106.

[13] Musaba, E. C.; Banda, B. Analysis of technical efficiency of small-scale soybean farmers in Mpongwe District, Zambia: A Stochastic Frontier Analysis. *Journal of Agriculture and Veterinary Science* **2020**, *13*(12), 49–56. https://doi.org/10.9790/2380-1312024956.

[14] Alabi, O. O.; Maharazu, I.; Aluwong, J. S.; Atteh, P. A.; Obida, A. Socioeconomic evaluation and technical efficiency of soybean (*Glycine max*) production in northwest Nigeria: A parametric approach. *SVU-International Journal of Agricultural Sciences* **2024**, *6*(3), 36–46. https://doi.org/10.21608/svuijas.2024.287999.1368.

[15] Otitoju, M. A.; Arene, C. J. Constraints and determinants of technical efficiency in medium-scale soybean production in Benue State, Nigeria. *Journal of Agricultural Research* **2010**, *5* (17), 2276–2280.

[16] Rinaldi, J.; Arya, N. N.; Mahaputra, I. K.; Elisabeth, D. A. A.; Resiani, N. M. D.; Arsana, I. G. K. D.; Silitonga, T. F. Production factors, technical, and economic efficiency of soybean (*Glycine max* L. Merr.) farming in Indonesia. *Open Agriculture* **2023**, *8* (1), 20220194. https://doi.org/10.1515/opag-2022-0194.

[17] Puphoung, S. Technical efficiency of farmers' rice production in Chiang Rai province analysis. *Prawarun Agricultural Journal* **2024**, *21*(1), 106–117. (in Thai).

[18] Somcom, S.; Wana, C. Technical efficiency analysis of rice production of farmers in Chai-Nat. *Journal of Chandrakasemsarn* **2020**, *26*(2), 179–196. (in Thai).

[19] Rahman, S.; Wiboonpongse, A.; Sriboonchitta, S.; Chaovanapoonphol, Y. Production efficiency of Jasmine rice producers in Northern and North-Eastern Thailand. *Journal of Agricultural Economics* **2009**, *60*(2), 419–435. https://doi.org/10.1111/j.1477-9552.2008.00198.x.

[20] Sriwichai, S. Technical efficiency of farmers' cassava production in Phayao Province. M.S. Thesis, Chiang Mai University, Chiang Mai, Thailand, **2014**. (in Thai).

[21] Krasachat, W. Technical inefficiency of chili farms in Thailand: A parametric approach. *Test Engineering & Management* **2020**, *83*, 30415–30421.

[22] Sornin, K.; Athipanyakul, T. Production Management and Production Costs of Conventional Chili and Safer Chili of Farmers in Chaiyaphum Province. In *Proceedings of the 15th Graduate Research Conference*; Khon Kaen University, Thailand, **2014**; pp 415–422. (in Thai).

[23] Phummuong, J. Technical efficiency of soybean production in the rainfed area of the lower northern region, Thailand. M.S. Thesis, Chiang Mai University, Chiang Mai, Thailand, **2002**. (in Thai).

[24] Lesak, P. Technical efficiency of soybean production in irrigated zone in Chiang Mai Province by StoNED method. M.S. Thesis, Chiang Mai University, Chiang Mai, Thailand, **2013**. (in Thai).

[25] Chanaim, S.; Srichikul, W.; Mensaklo, E. A. Copula-Based Stochastic Frontier Analysis of Thai Jasmine Rice Production in the Northeast of Thailand. *TEM Journal* **2023**, *12*(4), 2312–2318. https://doi.org/10.18421/TEM124-42.

[26] Arsad, R.; Isa, Z.; Abidin, N. H. Z.; Kamarudin, N. Measuring stock performance using stochastic frontier analysis model with dependent error approach. *International Journal of Advanced and Applied Sciences* **2022**, *9*(12), 1–10. https://doi.org/10.21833/ijaas.2022.12.001.

[27] Chaovanapoonphol, Y.; Singvejsakul, J.; Sriboonchitta, S. Technical efficiency of Rice production in the upper north of Thailand: clustering copula-based stochastic frontier analysis. *Agriculture* **2022**, *12*(10), 1585. https://doi.org/10.3390/agriculture12101585.

[28] Yamane, T. *Statistics: An Introductory Analysis*, 3rd ed.; Harper and Row: New York, **1973**.

[29] Poonkham, K. Analysis of efficiency of convention hotel in Thailand by using stochastic frontier with copula. M.S. Thesis, Chiang Mai University, Chiang Mai, Thailand, **2013**. (in Thai).

[30] Cohen, J. P. Production Functions for Medical Services. *Encyclopedia of Health Economics* **2014**, 180–183. https://doi.org/10.1016/B978-0-12-375678-7.01010-5.

[31] European Commission. *Directorate-General for Agriculture and Rural Development – Unit A.3.: Assessment of CAP Contribution to Sustainable Productivity*; European Union, **2025**. https://eu-cap-network.ec.europa.eu/sites/default/files/publications/2025-04/eu-cap-network-guidelines-assessment-of-cap-contribution-annexes.pdf (accessed Aug 12, 2025).

[32] Wiboonpongse, A.; Liu, J.; Sriboonchitta, S.; Denoeux, T. Modeling dependence between error components of the stochastic frontier model using copula: Application to intercrop coffee production in Northern Thailand. *International Journal of Approximate Reasoning* **2015**, *65*, 34–44. https://doi.org/10.1016/j.ijar.2015.04.001.

[33] Schwarz, G. Estimating the dimension of a model. *The annals of statistics* **1978**, *6*(2), 461–464. https://doi.org/10.1214/aos/1176344136.

[34] Burnham, K.; Anderson, D. *Model selection and multi-model inference: A Practical Information-Theoretic Approach*, 2nd ed.; Springer, **2004**.

[35] Aigner, D.; Lovell, C. A. K.; Schmidt, P. Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics* **1977**, *6*, 21–37. https://doi.org/10.1016/0304-4076(77)90052-5.

[36] Si, W.; Wang, X. Productivity growth, technical efficiency, and technical change in China's soybean production. *African Journal of Agricultural Research* **2011**, *6*(25), 5606–5613. https://doi.org/10.5897/AJAR11.1080.

[37] Thipbharos, P. Technical efficiency approach of agricultural producers by stochastic frontier analysis. *Journal of Economics Chiang Mai University* **2016**, *20*(2), 93–124. (in Thai).

[38] Adinya, I. B.; Offem, B. O.; Ikpi, G. U. Application of a stochastic frontier production function for measurement and comparison of technical efficiency of mandarin fish and clown fish production in lowland reservoirs, ponds, and dams of Cross River State, Nigeria. *The Journal of Animal & Plant Sciences* **2011**, *21*(3), 595–600.

[39] Butsaenkom, P. Selection of predictive variables into the multiple regression equation. *Journal of Educational Measurement Mahasarakham University* **2012**, *17*(1), 43–60. (in Thai).

[40] Chubtong, K. Technical efficiency of pineapple production in Prachuap Khiri Khan and Chiang Rai provinces. M.S. Thesis, Chiang Mai University, Chiang Mai, Thailand, **2014**. (in Thai).

[41] Donkor, E.; Owusu, V. Effects of land tenure systems on resource-use productivity and efficiency in Ghana's rice industry. *African Journal of Agricultural and Resource Economics* **2014**, *9*(4), 286–299. https://doi.org/10.22004/ag.econ.197016.

[42] Omar, Z.; Fatah, F. A. Determinants of technical efficiency among coconut smallholder production in Johor, Malaysia: A Cobb-Douglas stochastic frontier production approach. *IOP Conf. Series: Earth and Environmental Science* **2021**, *757*, 012013. DOI: 10.1088/1755-1315/757/1/012013.

[43] Ajibefun, I.; Battese, G.; Daramola, A. Determinants of technical efficiency in smallholder food crop farming: Application of stochastic frontier production function. *Journal of International Agriculture* **2002**, *41*(3), 225–240.

[44] Sritongtae, C.; Monkham, T.; Sanitchon, J.; Lodthong, S.; Srisawangwong, S.; Chankaew, S. Identification of Superior Soybean Cultivars through the Indication of Specific Adaptabilities within Duo-Environments for Year-Round Soybean Production in Northeast Thailand. *Agronomy* **2021**, *11*(3), 585. https://doi.org/10.3390/agronomy11030585.

ASEAN
Journal of Scientific and Technological Reports
Online ISSN:2773-8752

Research article

Vulnerability Assessment of Agricultural Produce to Flooding in Libacao, Aklan, Philippines

Nenia Bohulano[1*]

[1] College of Agriculture, Forestry, and Environmental Sciences, Aklan State University, Banga, Aklan, Philippines, 5601
* Correspondence: nenia.bohulano@asu.edu.ph

Citation:
Bohulano, N. Vulnerability assessment of agricultural produce to flooding in libacao, aklan, Philippines. *ASEAN J. Sci. Tech. Report.* **2025**, *28*(5), e258574. https://doi.org/10.55164/ajstr.v28i5.258574

Article history:
Received: April 1, 2025
Revised: July 18, 2025
Accepted: August 16, 2025
Available online: September 14, 2025

Abstract: Climate change affects agricultural productivity, food security, and rural livelihoods globally. Changes in temperature, rainfall patterns, and extreme weather events like droughts, floods, and stronger typhoons cause substantial damage to crops, livestock, and infrastructure and harm the agricultural sector as a whole. The Philippines is highly vulnerable to natural disasters due to its geographical and environmental setting. Aklan Province in Western Visayas, Philippines, boasts of its high geographic diversity, having five major river systems. Aklan River is the largest, longest, and third-largest drainage basin on the island. A conducted geohazard mapping and assessment identified Libacao to be vulnerable to the threat of flooding. This study focused on assessing the vulnerability of the agricultural produce in the Municipality of Libacao, Aklan, which is one of the communities along the Aklan River. The barangays of Calacabian, Calamcam, Casit-an, Dalagsaan, Guadalupe, Janlud, Julita, Loctuga, Magugba, Manika, Ortega, Oyang, Pampango, Pinonoy, and Rivera were assessed as they were situated along the Aklan River. The assessment identified rice, abaca, coconuts, bamboo, and banana as the top agricultural produce in Libacao. Vulnerability assessment showed that rice production in Calacabian, Calamcam, Casit-an, Loctuga, Pampango, Pinonoy, and Rivera was highly vulnerable. Also, Casit-an, Loctuga, Pampango, and Rivera were assessed to be very highly vulnerable to coconuts. For bananas, the barangays of Calacabian, Casit-an, Pampango, and Pinonoy were identified to have very high vulnerability. With the vulnerable areas identified, the Local Government Unit (LGU) of Libacao may consider the results in their planning and strategies.

Keywords: Agricultural vulnerability; flood hazard; adaptive capacity

1. Introduction

The significant impacts of climate change, especially on the agricultural sector, have been concerning. The Food and Agriculture Organization (FAO) of the United Nations reported that climate change is affecting agricultural productivity, food security, and rural livelihoods across the globe [1]. Some impacts of climate change on agriculture include changes in temperature, rainfall patterns, and extreme weather events like droughts, floods, and stronger typhoons. Extreme weather events cause substantial damage to crops, livestock, and infrastructure, which in turn leads to a considerable economic loss for farmers and the agricultural sector as a whole. The Philippines, due to its geographical and environmental setting, is highly vulnerable to natural disasters as an archipelago situated in the Pacific Ring of Fire, with more than 7,000 islands and 36,000 kilometers of coastline [2]. The Philippines has endured enormous losses from one disaster after another and has experienced 317 extreme weather events during the last decade. It incurred damages worth at least PHP 515.51 billion due to disasters from 2010 to 2020, of which 98% were climate-related [3]. Aklan Province in Western Visayas, Philippines, boasts of its high geographic diversity, ranging from rivers to white sandy beaches, mangroves, and mountainous landscapes. The province has five major river systems. Aklan River is the largest, longest, and third-largest drainage basin on the Island of Panay. The Aklan River system is part of the vital Aklan River Forest Reserve, which serves several municipalities along its banks.

The Philippines is among the top ten countries in Asia affected by flooding due to annual monsoonal rains and numerous typhoons [4]. Flooding is a type of natural disaster caused by extreme rainfall and has direct local effects. In such a way that it transforms the morphology of the riverbed, damages materials, or destroys infrastructures. In addition, it causes the loss of human life and has a significant socioeconomic impact over time. As stated by dela Torre et al. [4], flood damages are unavoidable, but by identifying which areas are susceptible to severe effects of this hazard due to underlying biophysical and socioeconomic factors, mitigating measures can be adopted by agricultural planners and farm managers. As such, the Aklan River causes disruptions when there is prolonged heavy rain. The increase in the average water level of the river leaves the residents living along the riverbanks or low-lying areas to take immediate action, such as evacuation. Likewise, damaged farmlands were prevalent.

A report by the Mines and Geosciences Bureau [5] identified towns in the Province of Aklan that were vulnerable to the threat of flooding. Results showed that Libacao, along with Madalag, Banga, Malinao, Lezo, Numancia, and the capital town of Kalibo, were the most vulnerable to the threat of flooding. These towns were situated along the Aklan River and experienced the impacts and consequences of river floods in recent years. The unceasing heavy rainfall has caused the Aklan River to overflow into the areas along the river. In the June 2008 Panay Island flooding resulting from Typhoon "Frank", there were reports of widespread landslides in the mountain range of Panay Island, also including the Aklan River headwaters. It suggested the possibility of damming tributaries that may have contributed to the flooding [6]. In 2014, Typhoon "Seniang" hit and threatened the municipality of Libacao along with neighboring towns with heavy downpours and uninterrupted rainfall that led to flooding, which affected the residents living along the Aklan River, especially farm households, and damaged the infrastructure. Tropical Depression "Agaton" in 2018 also flooded the areas along the Aklan River in the towns of Libacao, Kalibo, Madalag, Numancia, Banga, and Malinao. Severe Tropical Storm "Paeng" has also recently devastated the province of Aklan. The province reported a total of eight deaths in Aklan, of which five deaths were in Libacao, and the agriculture sector incurred the most significant amount of damage with around PHP 119.30 million [7].

The staggering effects of climate change, particularly the effects of flooding on crops, have become a grave concern worldwide [8]. As the Aklan River continues to threaten the natural resources, physical resources, and the inhabitants along the river, it is imperative to assess the vulnerability of these communities to hydrologic hazards, specifically flooding. The study focused on the assessment of vulnerability of the agricultural produce in one of the communities along the Aklan River, the Municipality of Libacao. The study aimed to determine the profile of the communities within the municipality that are exposed to hydrologic hazards, specifically flooding, and to assess their vulnerability in terms of their agricultural resources. Results can provide an objective and scientific basis for risk management decisions, priority setting, and resource allocation in environmental disaster management, and provide a rational basis for an enhanced implementation and policy development of the Local Government Unit (LGU).

2. Materials and Methods

2.1 Study area

The Municipality of Libacao in the Province of Aklan is located in Panay Island of the Western Visayas (Region VI), Philippines. It is a landlocked third-class municipality with a determined population of 28,272 as of the 2020 Census [9]. The municipality has a land area of 254.98 square kilometers or 98.45 square miles, which makes up around 14.49% of Aklan's total area. It is divided into 24 barangays. Libacao is situated along the Aklan River, which is the largest and longest river, with a primary waterway of 1,761.6 kilometers downstream [10]. This is the major river basin that traverses several municipalities in the province and is also within the Proclaimed Aklan Watershed and Forest Reserve of the province, with an area of 2,064 km^2 located in the municipalities of Libacao and Madalag. The map of the study area is shown in Figure 1.

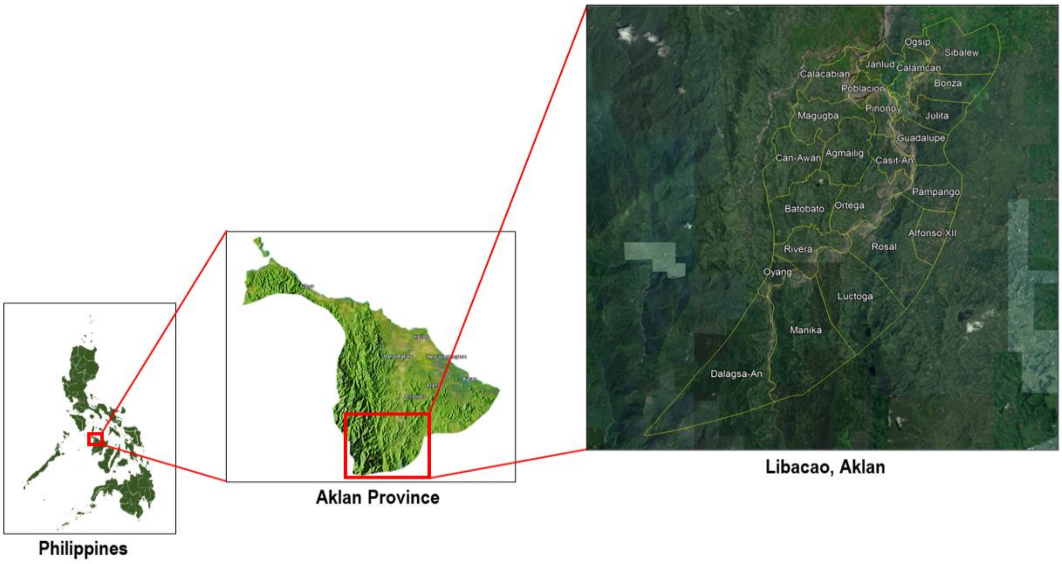


Figure 1. Map of the study area of Libacao, Aklan [11]. Yellow lines represent political boundaries

2.2. Vulnerability assessment

Vulnerability assessment of agricultural produce to flooding in Libacao was based on indicators clustered into respective components, such as the exposure, sensitivity, and adaptive capacity. Selected 15 barangays along the Aklan River were the following: Dalagsaan, Manika, Oyang, Rivera, Loctuga, Ortega, Pampango, Casit-an, Guadalupe, Pinonoy, Julita, Calamcam, Janlud, Magugba, and Calacabian. Participatory assessments such as Participatory Rural Appraisal (PRA) were also done with the participation of the community residents to assess and draw out information on the status, problems, and issues of the community. Stakeholders' participation in Disaster and Climate Risk Assessments (CDRA), focus group discussions (FGDs), and key informant interviews (KIIs) was also performed for the assessment. Field assessment and secondary data analysis were also done for the validation of the community [12].

2.3 Data collection and analysis

A collection of primary datasets was done through farmer household (HH) surveys on the 55 households located along the Aklan River. In addition, a Key Informant Interview (KII) was conducted with 5 representatives of the barangays, which include the Barangay Captain, Barangay Health Worker, Barangay Kagawad in-charge of the Enviroment and Agriculture, Barangay Secretary, and Barangay Risk Reduction Management Officer and focus group discussion with the Municipal Environment and Natural Resources Office (MENRO), Municipal Disaster Risk Reduction and Management Office, and Municipal Agriculturist Officer. Hence, data validation was conducted online with the presence of the Municipal Mayor and MENRO. A copy of the vulnerability result was presented to the municipal level. Survey materials consisted of questions investigating the status and characteristics of the farmlands and the socioeconomic and agroecological attributes of the farmer households. A descriptive research method was used in the data collection. The situation or given state of affairs in terms of specified aspects or factors was described in this. Described were the characteristics of individuals or groups of households in the community or physical environments and the conditions of the communities along the Aklan River basin, specifically the agricultural practices or the farming systems in the community. Descriptive statistics were used to analyze the qualitative and quantitative data gathered from the community, and situational analysis was done to validate further the information generated. Significant differences between means were evaluated by One-way Analysis of Variance (ANOVA) at $p<0.05$ using IBM SPSS Statistics 26. Frequencies and means were used in the study. The vulnerability scoring (LGU Guidebook on the Formulation of Local Climate Change Action Plan [13] using the formula and matrix below:

$$\text{RELATIVE VULNERABILITY} = \frac{\text{THREAT LEVEL (based on the exposure and sensitivity analysis)}}{\text{ADAPTIVE CAPACITY}}$$

MATRIX:

HAZARD	**THREAT LEVEL**	**ADAPTIVE CAPACITY**	**RELATIVE VULNERABILITY**
EXAMPLE: DROUGHT	**4 (Medium High)**	**3 (Medium)**	**4/3 = 1.33 (Low vulnerability)**

LEGEND:

THREAT LEVEL	**ADAPTIVE CAPACITY SCORE**					**RELATIVE VULNERABILITY**
	High (5)	**Medium High (4)**	**Medium (3)**	**Medium Low (2)**	**Low (1)**	
High (5)	1	1.25	1.66	2.5	5	High (4-5)
Medium High (4)	0.8	1	1.33	2	4	Medium High (2.1-3.9)
Medium (3)	0.6	0.75	1	1.5	3	Medium (1.5-2)
Medium Low (2)	0.4	0.5	0.66	1	2	Medium Low (1-1.49)
Low (1)	0.2	0.25	0.33	0.5	1	Low (>1)

3. Results and Discussion

3.1 Profile of the area and the respondents

Shown in Table 1 is the profile of the selected barangays situated along the Aklan River in terms of their population, number of households, number of farmers, agricultural land use, and total land area. The population in each barangay varies from 344 to 2,090. The barangay of Casit-an was the least populated, and Manika was the most populated. Moreover, the number of households ranged from 90 households in Casit-an to 631 households in Manika. It was observed that Janlud had 115 farmers tilling 44.0 ha of land, while Manika had the highest number of farmers at 975. Furthermore, the amount of land used for agricultural activities ranged from 44 to 591 hectares.

Table 1. Profile of selected barangays situated along the Aklan River.

Barangay	**Population**	**No. of Households**	**No. of Farmers**	**Agricultural Land Use (ha)**
Calacabian	943	235	410	54.50
Calamcam	1,012	166	180	378.97
Casit-an	344	90	255	153.66
Dalagsaan	1,908	393	450	390.50
Guadalupe	1,784	453	551	446.00
Janlud	1,306	319	115	44.00
Julita	2,077	483	690	271.50
Loctuga	1,832	346	290	578.00
Magugba	904	243	305	177.00
Manika	2,090	631	975	63.00
Ortega	1,295	310	640	368.50
Oyang	935	268	720	68.00
Pampango	1,504	396	896	515.80
Pinonoy	828	183	262	591.00
Rivera	648	160	157	207.00

Aklan is mainly an agricultural province since almost 110,452.0 hectares or 60.44% of its total land area is vast plains or rolling hills devoted to rice and other crops [10]. In Libacao, the top agricultural produce from the selected barangays was identified as shown in Table 2. The most common produce that was cultivated includes Rice, Abaca, Banana, Coconuts, Bamboo, and other produce.

Table 2. Agricultural Produce, Area Planted, and Crop Exposure/Vulnerability Rating of selected barangays situated along the Aklan River.

Barangay	Agricultural Produce	Total Area (ha)	Crop exposure/Vulnerability Rating
Calacabian	Abaca	2	Very low
	Bamboo	15	Very High
	Banana	2	Very High
	Coconuts	10	High
	Corn	2	Very High
	Rice	20	Very High
	Root crops (Sweet potato, Peanut)	2	Very High
	Vegetable	1.5	Very High
Calamcam	Abaca	13.25	Medium
	Copra	138	Medium
	Rice	57.75	Very High
	Root crops	174.97	Medium
Casit-an	Abaca	15.54	Very low
	Bamboo	45.94	Very low
	Banana	15	Very High
	Coconuts	16	Very High
	Corn	10	Very High
	Rice	41.18	Very High
	Root crops	10	Very High
Dalagsaan	Abaca	200	Very low
	Bamboo	20	Very low
	Banana	5	Very low
	Rice	165.5	Very low
Guadalupe	Abaca	100	Very Low
	Bamboo	5	Very Low
	Banana	100	Very Low
	Copra	100	Very Low
	Corn	25	Low
	Fruit Trees	66	Very Low
	Rice	25	Low
	Root crops/ Vegetables	25	Medium
Janlud	Abaca	7	Very Low
	Coconuts	10	Very Low
	Rice	25	Low
	Root crops	2	Very Low

Table 2. Agricultural Produce, Area Planted, and Crop Exposure/Vulnerability Rating of selected barangays situated along the Aklan River. (Continues)

Barangay	Agricultural Produce	Total Area (ha)	Crop exposure/Vulnerability Rating
Julita	Abaca	170.25	Very Low
	Bamboo	170.25	Very Low
	Coconuts	170.25	Medium
	Fruit Trees	170.25	Very Low
	Root crops	217.50	Medium
	Rice	217.50	Medium
	Vegetables	217.50	Medium
Loctuga	Abaca	500	Very Low
	Coconuts	10	Very High
	Fruit Trees	5	Very Low
	Rice	50	Very High
	Root crops	10	Very Low
	Vegetables	3	Very Low
Magugba	Abaca	10	Very Low
	Banana	15	Very Low
	Bamboo	67	Low
	Cassava	15	Very Low
	Coconuts	5	Very Low
	Corn	15	Medium
	Rice	50	Low
Manika	Abaca	50	Very Low
	Bamboo	10	Very Low
	Coconuts	1	Very Low
	Corn	1	Very High
	Rice	1	High
Ortega	Abaca	139	Medium
	Coconuts	118	Medium
	Corn	45	Medium
	Rice	66.50	Low
Oyang	Abaca	35	Very Low
	Bamboo	3	Very Low
	Banana	2	Very Low
	Coconuts	2	Very Low
	Rice	25	High
	Root crops (Taro)	1	Very Low
Pampango	Abaca	350	Medium
	Bamboo	50	Medium
	Banana	30	Very High
	Copra	20	Very High
	Fruit Trees	8	Very High
	Ginger	5	Very High

Table 2. Agricultural Produce, Area Planted, and Crop Exposure/Vulnerability Rating of selected barangays situated along the Aklan River. (Continues)

Barangay	Agricultural Produce	Total Area (ha)	Crop exposure/Vulnerability Rating
	Rice	52.8	Very High
Pinonoy	Abaca	20	Very Low
	Bamboo	4	Very Low
	Banana	2	Very High
	Coconuts	509	High
	Fruit Trees	15	Very Low
	Peanut	1	Very High
	Rice	40	Very High
Rivera	Abaca	150	Very Low
	Bamboo	15	Low
	Coconuts	5	Very High
	Rice	37	Very High

3.2 Vulnerability assessment

The Intergovernmental Panel on Climate Change [14] stated that vulnerability encompasses a variety of concepts and elements, including sensitivity or susceptibility to harm and lack of capacity to cope and adapt. It involves a combination of factors that determine the degree to which someone's life, livelihood, property, and other assets are at risk from an event or cascade of events in nature [15]. Vulnerability can be a function of the system's sensitivity to climate change, exposure to climatic hazards, and its adaptive capacity to these changes. Vulnerability assessment is a process of identifying and evaluating the potential risks and impacts of a hazard or stressor on a system or population. These vulnerability assessments identify the factors that contribute to vulnerability and inform the development of strategies to reduce or manage those risks. Exposure is one of the dimensions of vulnerability assessment. It is defined as the degree to which a system is exposed to climatic variations, considering both the frequency and extent of its contact with a hazard [16]. Shown in Figure 2 is the exposure of the top crops, such as rice, abaca, coconuts, bamboo, and banana, to flooding. In terms of rice (Figure 2a), six barangays were evaluated to have high exposure (>40%), namely: Calacabian, Pampango, Calamcam, Ortega, Rivera, and Loctuga. This resulted in the economic value of crop loss that ranged from Php1,620,000.00 to Php103,685,400.00 with an exposed area of 20 to 50 ha (Table 3). Calacabian was assessed to have the highest exposure among them. Pinonoy, Manika, and Casit-an were assessed to have medium exposure (<40%) in the area exposed from 0.3 to 15 ha. The crop economic value lost ranged from Php21,048.00 to Php1,800,000.00 (Table 3). The barangays of Guadalupe, Janlud, Magugba, Oyang, and Julita were evaluated to have low exposure to flooding hazards. On the other hand, Dalagsaan has no areas exposed to flooding hazards. For abaca, Barangay Ortega was highly exposed, and Calamcam and Pampango had low exposure (≤20%), while the rest were not exposed, as shown in Figure 2b. For coconuts (Figure 2c), four barangays were highly exposed areas (Pampango, Loctuga, Ortega, and Rivera), and Casit-an had medium exposure. The highly exposed areas ranged from 2 to 57 ha with an economic value lost of 75,000.00 to 1,425,000.00, while the Casit-an area exposed was 5 ha with an economic loss of 150,000.00 (Table 3). Calacabian, Calamcam, Pinonoy, and Julita were areas with low exposure. In terms of areas with bamboo (Fig. 2d), Pampango, Rivera, and Magugba were found to have low exposure. The rest of the areas, namely: Calacabian, Casit-an, Dalagsaan, Guadalupe, Julita, Manika, Oyang, and Pinonoy, were not exposed to the identified hazard. Shown in Figure 2e are the areas of Libacao that have high exposure when it comes to bananas. It was assessed that Pampango, Pinonoy, and Calacabian were highly exposed areas. The economic value lost from these barangays ranges from Php5,000 to Php1,020,000.00 from 1 to 25 ha of bananas. Also, Casit-an was evaluated to have medium exposure (5 ha) with an economic value of Php50,000.00 for bananas. Areas in Dalagsaan, Guadalupe, Magugba, and Oyang were not exposed. In the rice areas with high and medium exposure, trainings were conducted on rice production using varieties suited to flooded conditions. While in the areas of coconuts and bananas with high and medium exposure, off-farm livelihood training (fruits and food processing) was conducted.

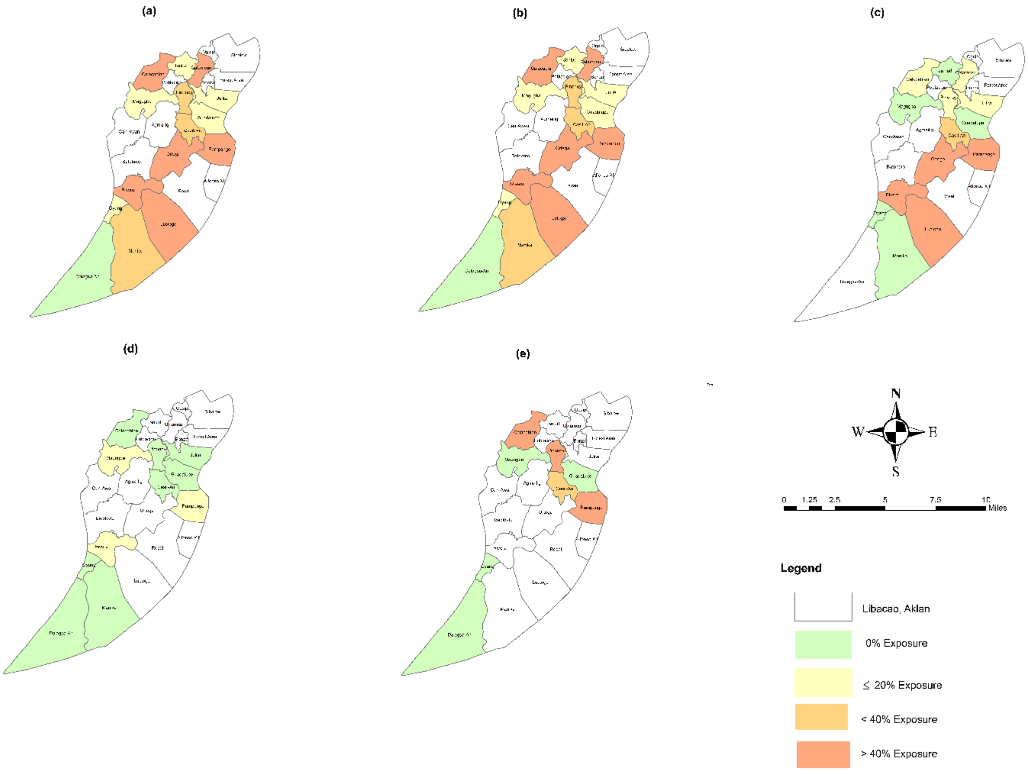


Figure 2. Exposure to flooding hazard of the top agricultural produce: (a) rice, (b) abaca, (c) coconuts, (d) bamboo, and (e) banana in Libacao, Aklan, Philippines.

Another dimension considered in vulnerability assessments is sensitivity. It refers to the degree to which a system is affected, either adversely or beneficially, by a hazard [17]. Some crops may be more sensitive to changes in rainfall or water availability, which makes them more vulnerable to the impacts of flooding. Sensitivity indicators utilized in this study include the ratio of farmers with access to climate information, the ratio of farmers employing sustainable production techniques, and the ratio of farmers with access to irrigation. Based on the conducted surveys (Table 4), it was revealed that more than half (59.00%) of abaca farmers had access to climate information. However, less than half (46.67%) of them reported not employing sustainable production techniques. Furthermore, a significant ($p<0.05$) number of them do not have access to irrigation. Abaca is considered a less water-intensive crop. However, extended dry spells could harm production and farmer livelihoods, reducing their ability to adjust. For bamboos, more than half have access to climate information (57.73%) and were employing sustainable production techniques (50.91%). Conversely, none of them have access to irrigation. The lack of an irrigation system during critical growth periods of bamboo will limit their ability to manage water stress, hence affecting their overall ability to respond to climate variability. About 46.25% of banana farmers have access to climate information for their production, and about 51.88% of the banana farmers employ sustainable production techniques. Also, none of them have access to irrigation for their crops. Their adaptive capacity is naturally diminished by this reliance on natural rainfall, which could result in significant crop losses during dry seasons. As for coconuts, a high number of them (60.36%) have access to climate information.

On the other hand, only around 40.71% of coconut farmers were employing sustainable production techniques, and none of them had access to irrigation. Despite its resilient, prolonged exposure to dry spells, it constrains their overall adaptive capacity. Abaca, bamboo, banana, and coconut farmers do not have access to irrigation since these crops do not necessarily require irrigation. Around 65.47% of rice farmers have access to climate information. Fewer rice farmers were reported to be employing sustainable production practices (27.00%) and have access to irrigation (20.67%). The limited use of sustainable methods and insufficient irrigation infrastructure resulted in rice farmers' limited ability to adapt and overall vulnerability. A significant number ($p<0.05$) of farmers for abaca, bamboo, banana, coconuts, and rice were dependent on rainfall and other means of water for their crops since they do not have access to irrigation. These will later result in crop yield losses, food insecurity (rice, bananas, and coconuts), financial strain, soil degradation, limited early warning and response effectiveness, and psychological and social impact. It is recommended to do better extension activities through translating climate data, demonstrating sustainable practices, farmer field schools, and digital extension. In addition, capacity building for extension workers, strengthening research-extension-farmer linkages, promoting flood-resilient crop varieties, and rainwater harvesting are suggested for adaptive capacity.

Table 3. Different crops, barangays, annual average output per hectare, exposed area, and exposed value

Crops	Barangay	Annual Average Output per hectare (Philippine peso (Php) per ha)	Exposed Area (ha)	Exposed Value (Php)
Rice (High Exposure)	Calacabian	100,000.00	20	2,000,000.00
	Pampamgo	33,600	50	1,680,000.00
	Calamcam	2,184,000.00	47.48	103,685,400.00
	Ortega	36,000	45	1,620,000.00
	Rivera	107,500	22.2	2,386,500.00
	Loctuga	65,000	25	1,625,000.00
Rice (Medium Exposure)	Pinonoy	120,000	15	1,800,000.00
	Manika	70,160	0.3	21,048.00
	Casit-an	85,680	11	942,480.00
Coconuts (High Exposure)	Pampango	17,280	10	172,800.00
	Loctuga	15,000	5	75,000.00
	Ortega	25,000	57	1,425,000.00
	Rivera	93,296	2	186,592.00
Coconuts (Medium Exposure)	Casit-an	30,000	5	150,000.00
Bananas (High Exposure)	Pampango	40,800	25	1,020,000.00
	Pinonoy	5,000	1	5,000
	Calacabian	15,000	1	15,000
Bananas (Medium Exposure)	Casit-an	10,000	5	50,000.00

Table 4. Sensitivity indicators for the top agricultural produce of Libacao, Aklan, Philippines.

Parameter*	Abaca	Bamboo	Banana	Coconut	Rice
Access to climate information (%)					
With	59.00 ± 27.66^{a}	57.73 ± 34.74^{a}	46.25 ± 28.75^{a}	60.36 ± 28.32^{a}	64.47 ± 24.97^{a}
Without	41.00 ± 27.66^{a}	42.27 ± 34.74^{a}	53.75 ± 28.75^{a}	39.64 ± 28.32^{a}	35.53 ± 28.74^{a}
Employing sustainable production techniques (%)					
Yes	46.67 ± 41.52^{a}	50.91 ± 43.92^{a}	51.88 ± 45.19^{a}	40.71 ± 38.92^{a}	27.00 ± 28.90^{a}
No	53.33 ± 41.52^{a}	49.09 ± 43.92^{a}	48.13 ± 45.19^{a}	59.29 ± 38.92^{a}	73.00 ± 28.90^{b}
Access to irrigation (%)					
With	0.00 ± 0.00^{a}	0.00 ± 0.00^{a}	0.00 ± 0.00^{a}	0.00 ± 0.00^{a}	20.67 ± 36.93^{a}
Without	100.00 ± 0.00^{b}	100.00 ± 0.00^{b}	100.00 ± 0.00^{b}	100.00 ± 0.00^{b}	79.33 ± 36.93^{b}

* means with different letters in each section per crop are significantly different from each other ($p<0.05$)

Adaptive capacity is another important dimension considered in vulnerability assessments. Adaptive capacity refers to the ability of a being or a system to adjust to the impacts of a hazard [17] or evolve [18] to accommodate environmental hazards and neutralize potential damages, or to take advantage of opportunities of planning to expand its range of variability for coping [16]. Access to financing, such as crop insurance, availability of alternative livelihoods, and access to government extension programs, were chosen as adaptive capacity indicators for this study. Access to crop insurance provided by the Philippine Crop Insurance Corporation (PCIC) program, the Philippine Coconut Authority (PCA), the Department of Agriculture (DA), agricultural cooperatives, and others was based on the conducted HH, FGDs, and the data provided by PCIC. Shown in Table 4 are the mentioned adaptive capacity indicators for the study. It was reported that crop insurance was available through PCIC, DA, PCA, and the Libacao Cooperative. Also, alternative livelihoods mentioned included backyard gardening, poultry, livestock, cash for work, and others. Furthermore, there were several government extension programs available provided by DA, Department of Labor and Employment (DOLE), Department of Social Welfare and Development (DSWD), Department of Trade and Industry (DTI), and Philippine Fiber Industry Development Authority (PhilFIDA) which include programs such as Abaca Trainings, Farm Schools, Food Processing Trainings, Livestock Trainings, Vegetable Farming Trainings, and Sustainable Livelihood Programs.

Table 4. Adaptive capacity indicators for the top agricultural produce of Libacao, Aklan, Philippines.

Crop	Access to financing	Alternative livelihood	Government extension programs
Abaca	Crop Insurance (PCIC, DA, Libacao Cooperative)	Backyard Gardening and Poultry, Cash for work, Livestock, SAP, Sustainable Livelihood by DOLE and DSWD, and Vegetable seedlings	Abaca Trainings, Farm Schools, Food Processing Trainings, Livestock Trainings, Vegetable Farming Trainings, Sustainable Livelihood Program by DSWD
Bamboo	Crop Insurance (PCIC and DA)	Backyard Gardening and Poultry, Cash for work, Livestock, SAP, Sustainable Livelihood by DOLE and DSWD, and Vegetable seedlings	Farm Schools, Food Processing Trainings, Livestock Trainings, Vegetable Farming Trainings, Sustainable Livelihood Program by DSWD
Banana	Crop Insurance (PCIC and DA)	Backyard Gardening and Poultry, Cash for work, Livestock, Sustainable Livelihood by DOLE and DSWD, and Vegetable seedlings	Food Processing Trainings, Livestock Trainings, Vegetable Farming Trainings, Sustainable Livelihood Program by DSWD
Coconut	Crop Insurance (PCIC, PCA, DA, Libacao Cooperative)	Backyard Gardening and Poultry, Cash for work, Livestock, SAP, Sustainable Livelihood by DOLE and DSWD, and Vegetable seedlings	Farm Schools, Food Processing Trainings, Livestock Trainings, Vegetable Farming Trainings, Sustainable Livelihood Program by DSWD
Rice	Crop Insurance (PCIC, DA, Libacao Cooperative)	Backyard Gardening and Poultry, Cash for work, Livestock, SAP, Sustainable Livelihood by DOLE and DSWD, and Vegetable seedlings	Abaca Trainings, Farm Schools, Food Processing Trainings, Livestock Trainings, Vegetable Farming Trainings, Sustainable Livelihood Program by DSWD

Vulnerability assessment is a valuable tool to identify systems that are at risk of a hazard. In terms of agriculture, it can help in identifying the areas and crops that are at risk of a hazard, such as flooding in the case of this study. By factoring in the different dimensions of exposure, sensitivity, and adaptive capacity into the consideration, these assessments help give a better understanding of the risk and vulnerabilities posed to these agricultural systems or crops.

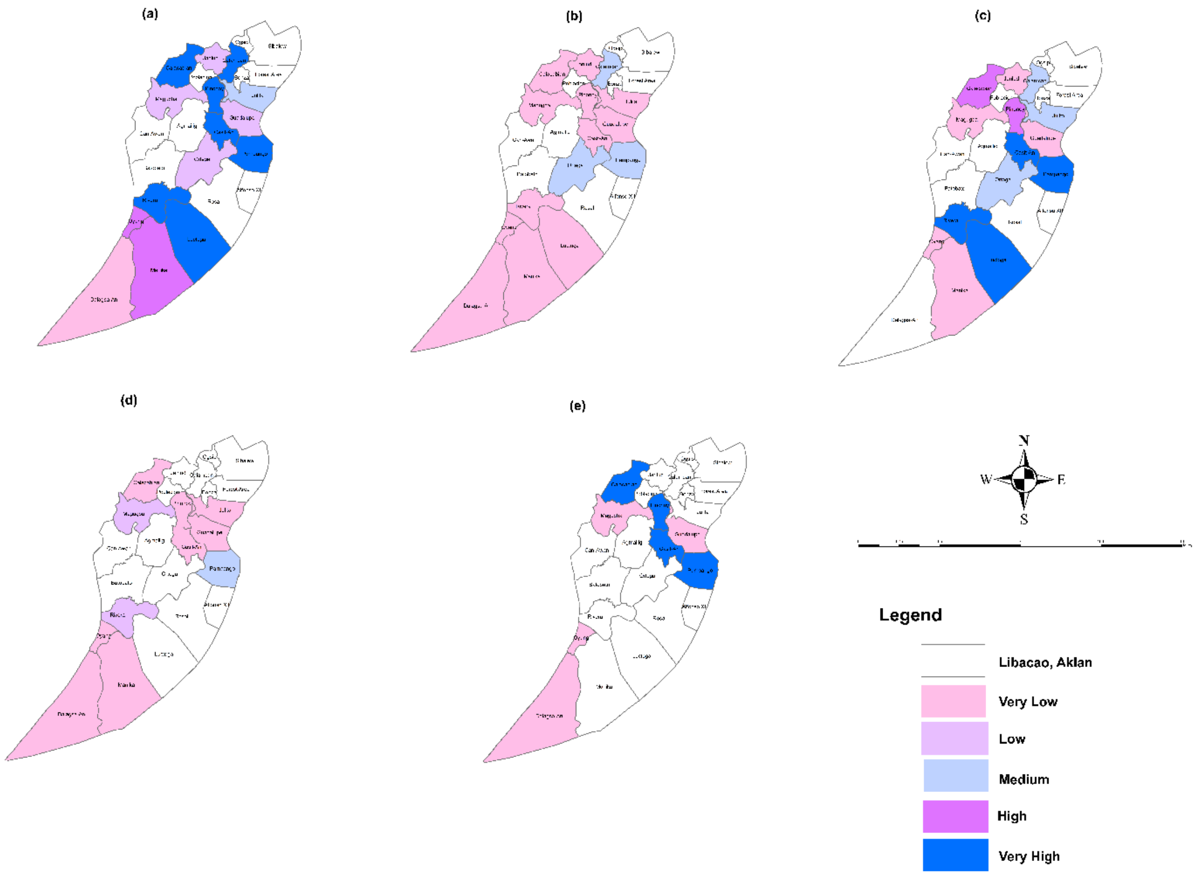


Figure 3. Vulnerability to flooding of the top agricultural produce: (a) rice, (b) abaca, (c) coconuts, (d) bamboo, and (e) banana in Libacao, Aklan, Philippines.

For rice, seven barangays were identified to have a very high vulnerability, namely: Calacabian, Calamcam, Casit-an, Loctuga, Pampango, Pinonoy, and Rivera (Fig. 3a). Likewise, the barangays of Manika and Oyang were identified to be highly vulnerable, while Julita was evaluated to have a medium vulnerability. On the other hand, the four barangays of Guadalupe, Janlud, Magugba, and Ortega were identified to have low vulnerability, while Dalagsaan had a very low vulnerability. When it comes to abaca (Fig. 3b), the three barangays of Calamcam, Ortega, and Pampango were identified as having medium vulnerability. The other barangays planting abaca had very low vulnerabilities, as follows: Calacabian Casit-an, Dalagsaan, Guadalupe, Janlud, Julita, Loctuga, Magugba, Manika, Oyang, Pinonoy, and Rivera. Additionally, Casit-an, LeLoctuga, Pampango, and Rivera were assessed to be very highly vulnerable for coconuts (Fig. 3c). Also, Calacabian and Pinonoy were highly vulnerable. The barangays of Calamcam, Julita, and Ortega were assessed to have a medium vulnerability. Moreover, the five barangays of Guadalupe, Janlud, Magugba, Manika, and Oyang were identified to have low vulnerabilities. Shown in Figure 3d are the different levels of vulnerability of bamboo in terms of the flooding hazard. It was identified that only Pampango had a medium vulnerability. Magugba and Rivera had low vulnerability. Furthermore, the remaining barangays identified had very low vulnerability, namely: Calacabian, Casit-an, Dalagsaan, Guadalupe, Julita, Manika, Oyang, and

Pinonoy. Lastly, for bananas (Fig. 3e), four barangays were identified to have very high vulnerability. These were the barangays of Calacabian, Casit-an, Pampango, and Pinonoy. On the other hand, the remaining barangays of Dalagsaan, Guadalupe, Magugba, and Oyang had very low vulnerability.

4. Conclusions

Rice production in the areas of Calacabian, Calamcam, Casit-an, Loctuga, Pampango, Pinonoy, and Rivera was highly vulnerable. Also, for coconuts, Casit-an, Loctuga, Pampango, and Rivera were assessed to be very highly vulnerable. For bananas, the barangays of Calacabian, Casit-an, Pampango, and Pinonoy were identified to have very high vulnerability. Identifying vulnerable areas for specific crops will be beneficial for Local Government Units (LGUs), policymakers, and planners in preparing and adapting to measures for flooding events. Henceforth, initiatives to increase the coping capacity of communities to such events may be included in the local government's Local Disaster Risk Reduction and Management Plan (LDRRMP) and Local Climate Change Action Plan (LCCAP).

5. Acknowledgements

The study acknowledges everyone who has contributed to this research, particularly the consulted local experts, the Local Government Unit of Libacao, Aklan, the Barangay Officials of selected barangays, and the farmer household survey respondents.

Funding: This research was funded by the COMMISSION ON HIGHER EDUCATION (CHED) and the support of the Aklan State University.

Conflicts of Interest: The author declares no conflict of interest.

References

[1] Climate Change and Food Security: Risks and Responses. Food and Agriculture Organization of the United Nations, **2015**. https://www.fao.org/3/i5188e/I5188E.pdf (accessed Mar 11, 2023).

[2] Examining the Philippines' Disaster Risk Reduction and Management System. Senate Economic Planning Office (SEPO), **2017**. https://legacy.senate.gov.ph/publications/SEPO/PB_Examining%20PH%20DRRM%20System_Revised_27June2017.pdf (accessed Mar 11, 2023).

[3] Climate Finance At A Glance. Senate Economic Planning Office (SEPO), **2022**. https://legacy.senate.gov.ph/publications/SEPO/AAG_Climate%20Finance_Final_27June%202022.pdf (accessed Mar 11, 2023).

[4] Dela Torre, D. M. G.; Dela Cruz, P. K. A.; Jose, R. P. et al. Geospatial Assessment of Vulnerabilities of Croplands to Flooding Risks: A Case Study of Philippine River Basins. *Int Arch Photogramm Remote Sens Spat Inf Sci - ISPRS Arch.* **2019.** https://doi.org/10.5194/isprs-archives-XLII-4-W19-173-2019

[5] Mines and Geosciences Bureau Region 6. Results of the MGB Geohazard Assessment of Barangays in Balete, Buruanga, Ibajay, Kalibo, Libacao, Madalag, Malay and Nabas in Aklan Province. http://www.mgb6.org/wp-content/uploads/2015/01/aklan-geohazard-summary-report-1-page.doc (accessed Mar 11, 2023).

[6] Catane, S. G.; Abon, C. C.; Saturay, R. M. et al. Landslide-amplified flash floods—The June 2008 Panay Island flooding, Philippines. *Geomorphology* **2012**, *169-170*, 55-63. https://doi.org/10.1016/j.geomorph.2012.04.008

[7] Aklan records over P288M 'Paeng' damage. Philippine Information Agency (PIA), **2022**. https://pia.gov.ph/news/2022/11/04/aklan-records-over-p288m-paeng-damage (accessed Mar 20, 2023).

[8] Mallari, A. E. Climate Change Vulnerability Assessment in the Agriculture Sector: Typhoon Santi Experience. *Procedia – Social and Behavioral Sciences* **2016**, *216*, 440-451. https://doi.org/10.1016/j.sbspro.2015.12.058

[9] Cities and Municipalities Competitive Index. https://cmci.dti.gov.ph/lgu-profile.php?lgu=Libacao (accessed Mar 10, 2023).

[10] Google Earth. Libacao, Aklan. https://earth.google.com (accessed Mar 10, 2023). (accessed Jul 12, 2025).

[11] Philippine Statistics Authority Provincial Statistical Office [PSA]. *The Countryside in Figures Province of Aklan. 2020 Edition*; PSA: Aklan, Philippines, **2020**.

[12] Housing and Land Use Regulatory Board [HLURB]. *Supplemental Guidelines on Mainstreaming Climate Change and Disaster Risks in the Comprehensive Land Use Plan*; Housing and Land Use Regulatory Board, Climate Change Commission, United Nations Development Programme, Australian Government, **2015**.

[13] Maralli Jr, A.; Megia, P. J.; Dimaano, E.; Aparicio, J.; Mendoza, K. M. *LGU Guidebook on the Formulation of Local Climate Change Action Plan (LCCAP) Book 1*; Department of the Interior and Local Government (DILG). https://region5.dilg.gov.ph/camarinessur/wp-content/uploads/2018/05/BOOK-1-LGU-Guidebook-in-LCCAP-Formulation-Process.pdf

[14] Intergovernmental Panel on Climate Change [IPCC]. Summary for policymakers. In *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*; Cambridge University Press, **2014**; pp 1-32. https://doi.org/10.1017/CBO9781107415324.004

[15] Wisner, B.; Blaikie, P. M.; Blaikie, P.; Cannon, T.; Davis, I., Eds. *At risk: natural hazards, people's vulnerability and disasters*; Routledge, **2004**. https://doi.org/10.4324/9780203974575

[16] Jose, R.; Apura, R. J.; Dela Torre, D. M. et al. Assessing the vulnerability of agricultural crops to riverine floods in Kalibo, Philippines using composite index method. *GISTAM 2017 - Proc 3rd Int Conf Geogr Inf Syst Theory, Appl Manag* **2017**, 184-194. https://doi.org/10.5220/0006284601840194

[17] Locatelli, B.; Herawati, H.; Brockhaus, M. et al. *Methods and Tools for Assessing the Vulnerability of Forests and People to Climate Change*; Center for International Forestry Research, **2008**.

[18] Doch, S.; Diepart, J. C.; Heng, C. A multi-scale flood vulnerability assessment of agricultural production in the context of environmental change: The case of Sangkae River watershed, Battambang Province. In Diepart, J. C., Ed. *Learning for resilience: Insights from Cambodia's rural communities*; The Learning Institute: Phnom Penh.

[19] Doch, S.; Diepart, J. C.; Heng, C. A multi-scale flood vulnerability assessment of agricultural production in the context of environmental change: The case of Sangkae River watershed, Battambang Province. In *Learning for resilience: Insights from Cambodia's rural communities*; Diepart, J. C., Ed.; The Learning Institute: Phnom Penh, **2015**; pp 19–49.

ASEAN
Journal of Scientific and Technological Reports
Online ISSN:2773-8752

Research article

Effects of Different Levels of Vermicompost on Black Rice (*Oryza sativa* L.) Cultivation

Harry Casimero Luces

[1] Faculty of College of Agriculture, Forestry and Food Science, University of Antique-Hamtic Campus, 5715, Philippines
* Correspondence: harry.luces@antiquespride.edu.ph

Citation:
Luces, H. Effects of different levels of vermicompost to black rice (*Oryza sativa* L.) cultivation. *ASEAN J. Sci. Tech. Report.* **2025**, *28*(5), e257803. https://doi.org/10.55164/ajstr.v28i5.257803.

Article history:
Received: February 8, 2025
Revised: August 17, 2025
Accepted: September 1, 2025
Available online: September 14, 2025

Abstract: Given the growing need to address the challenges of black rice cultivation and the environmental impacts of synthetic fertilizers, this study seeks to evaluate the response of a black rice variety to different levels of vermicompost application. The development of organic farming would bring a healthier environment, safe for humans, and result in increased food production that addresses Sustainable Development Goals (SDGs) such as Life on Land (SDG 15) and Zero Hunger (SDG 2). The experiment was conducted using a Randomized Complete Block Design (RCBD) with four treatments, each replicated three times. The treatments were as follows: A–20 tons of vermicompost per hectare, B – 25 tons of vermicompost per hectare, C–30 tons of vermicompost per hectare, and D – inorganic fertilizer applied at a rate of 90-30-30 kg NPK per hectare. The results revealed no significant differences among the treatment means at all stages of plant growth, including growth indicators such as the number of tillers and plant height, as well as yield components, including panicle length, fresh and dry grain weight, and the number of filled and unfilled grains. These findings suggest that applying 20, 25, or 30 tons of vermicompost per hectare produces similar effects to using synthetic fertilizer. The study highlights the potential of vermicompost as a sustainable and environmentally friendly alternative to synthetic fertilizers for black rice cultivation, thereby reducing reliance on synthetic fertilizers that may have long-term adverse effects on soil health.

Keywords: Fertilizer management; nutrient management; organic fertilizer; vermiculture; sustainable farming

1. Introduction

Black rice (*Oryza sativa* L.), often referred to as "forbidden rice" due to its historical association with royalty in ancient China [1], has gained significance as a crop due to its exceptional nutritional and health benefits. Black rice is rich in anthocyanins, which possess a potent antioxidant property and are linked to a reduced risk of chronic illnesses, such as diabetes, heart disease, and cancer [2]. Its nutrient-dense profile, which includes essential amino acids, dietary fiber, vitamins, and minerals [3], makes it a popular choice among health-conscious consumers and a versatile ingredient in various food products. However, despite its increasing demand, black rice faces unique agronomic challenges that hinder its cultivation and productivity.

Black rice typically exhibits a tall phenotype and a long vegetative stage, which increases the risk of lodging and reduces productivity compared to white rice [4]. Furthermore, limited agronomic research has left farmers without optimal cultivation practices, further restricting their adoption and profitability [5]. The development of sustainable and effective practices to enhance black rice cultivation is imperative, not only to address these challenges but also to conserve its genetic and cultural heritage, which is at risk of erosion due to the dominance of high-yielding commercial varieties [6]. Current agricultural practices heavily rely on synthetic fertilizers to boost crop yields. While synthetic fertilizers may be effective in the short term, they contribute to numerous environmental and health issues, including soil degradation [7], water pollution [8], greenhouse gas emissions [9], and loss of biodiversity [10]. Additionally, these fertilizers often lead to soil acidification [11] and nutrient imbalances [12], diminishing long-term soil fertility. In the Philippines, where rice is a staple crop, the overuse of synthetic inputs poses serious risks to agricultural sustainability and food security [13].

In contrast, vermicompost offers a promising and sustainable alternative. Produced through the decomposition of organic matter by earthworms and microorganisms, vermicompost is rich in essential macro- and micronutrients, plant growth hormones, and beneficial microbes [14]. It enhances soil texture, aeration, and water retention while promoting plant growth and suppressing pests and diseases [15]. Studies in the Philippines have demonstrated the cost-effectiveness and productivity benefits of vermicompost [16], particularly in reducing reliance on synthetic fertilizers [16]and improving soil health [17]. Given the growing need to address the challenges of black rice cultivation and the environmental impacts of synthetic fertilizers, this study seeks to evaluate the response of a black rice variety to different levels of vermicompost application. The findings aim to provide evidence-based recommendations for sustainable cultivation practices that enhance black rice productivity while promoting environmental stewardship. This research contributes to the development of sustainable agricultural systems and supports the broader goals of food security, environmental conservation, and the preservation of traditional rice varieties.

This study was conducted to investigate the effects of varying vermicompost levels on black rice (Oryza sativa L.) cultivation, specifically in terms of growth and yield. Specifically, it aimed to determine the growth in terms of the number of tillers and height, as well as yield in terms of the length of the panicle, fresh and dry weight, and the number of filled and unfilled grains of black rice applied with different levels of vermicompost.

2. Materials and Methods

2.1 Vermicompost Production

For the preparation of vermicompost, a 30:1 ratio of carbon to nitrogenous sources was followed [24]. Carbon sources included rice straw, paper, and dried banana leaves, while nitrogenous sources comprised dried goat and cattle manure and leaves from nitrogen-fixing plants. These materials were shredded and allowed to decompose for 30 days. After this initial composting period, the material was fed to African Night Crawlers. Vermicompost was harvested after an additional 30 days of processing by the worms.

2.2 Black Rice Seedling Production and Research Design

Rice seedlings were produced using the "Dapog Method". The 12-day-old seedlings were then transplanted into the rice paddies at a spacing of 20 cm x 20 cm [19]. The experimental area, measuring 330 square meters (27.5 meters long and 12 meters wide), was laid out using a Randomized Complete Block Design (RCBD) with four treatments replicated three times. Each plot had dimensions of 2 meters x 5 meters with a 0.5-meter distance between plots. The treatments were as follows:

Treatment A: 20 tons of vermicompost per hectare
Treatment B: 25 tons of vermicompost per hectare
Treatment C: 30 tons of vermicompost per hectare
Treatment D: Commercial fertilizer (urea and T-14 at 90–30–30 kg NPK per hectare)

2.3 Data collection

Data were collected from ten representative samples marked 20 days after transplanting, except for the fresh and dry weight of grains, which were measured from all plants within the plot. *Growth Indicators*: Plant height was measured at 30, 60, and 110 days after transplanting, from the base of the plant to the tip of its tallest leaf [20]. The number of tillers was also counted at 30, 60, and 110 days after transplanting, based on the ten representative samples [21]. *Yield Components:* The length of the panicle was determined by measuring from its base to the tip and was recorded from ten representative samples [22]. Filled and unfilled grains were counted separately per panicle, using the same representative samples [23].

2.4 Statistical Analysis

All gathered data were recorded, tabulated, and analyzed using Analysis of Variance (ANOVA) for Randomized Complete Block Design through the Statistical Tool for Agricultural Research (STAR) version 2.0.1 [18].

3. Results and Discussion

3.1 Height of black rice at different stages of growth applied with different levels of vermicompost

The average height of black rice at 30, 60, and 110 days after transplanting is presented in Figure 1. Treatment A recorded mean heights of 46.98 cm, 100.80 cm, and 113.73 cm, respectively. Treatment B showed mean heights of 49.50 cm, 96.48 cm, and 112.43 cm. Treatment C had mean heights of 47.05 cm, 100.43 cm, and 113.43 cm. Finally, Treatment D achieved mean heights of 47.95 cm, 107.91 cm, and 122.27 cm. The height of black rice increased as the number of days lapsed, and it varied with the application of different rates of vermicompost and inorganic fertilizer. There was a rapid increase in plant growth from 30 days to 60 days after transplanting at all levels of vermicompost; the increase was more pronounced, as indicated by the overlapping of the lines. The range of increase in the height of the plants applied with inorganic fertilizer was wide, from 47.95 cm to 107.91 cm, from 30 to 60 days after transplanting, respectively, as evidenced by the wider gap between Treatment D (yellow line) and the rest of the treatments A, B, and C (blue, orange, and gray). From 60 to 110 days after transplanting, the rate of plant growth was already slow, as observed in all treatments. However, statistical analysis using analysis of variance for the Randomized Complete Block Design revealed that the differences among treatment means in terms of height at all stages of plant growth did not vary significantly at the 5% level of significance for black rice height. This means that the effects of 20, 25, and 30 tons of vermicompost are comparable to those of synthetic fertilizer at a rate of 90-30-30 kg NPK per hectare. Similar findings were observed, showing that vermicompost significantly enhanced rice seedling biomass, photosynthesis, and chlorophyll content, thereby promoting vigorous early-stage growth [25]. Likewise, vermicompost improves nutrient uptake and microbial activity, which are critical for vegetative development and plant height [26], [27]. Studies on combined applications further showed that vermicompost used with reduced chemical fertilizers can sustain strong vegetative growth while reducing chemical input [28], [29]. Evidence from other crops also aligns with these results. For instance, vermicompost has been shown to improve chlorophyll content and photosynthetic efficiency in spinach [30], while a similar study reported enhanced vegetative biomass in beans, even under water stress [31]. Additionally, vermicompost is beneficial in improving soil structure and root development, which supports plant height gains [32].

3.2 Number of Tillers of Black Rice at Different Levels of Vermicompost

The results showed that at 30 days after transplanting, the highest average number of tillers (4.00) was observed in black rice applied with 90-30-30 kg NPK per hectare. This was followed by those treated with 25 tons of vermicompost per hectare (3.47), 20 tons of vermicompost per hectare (3.37), and 30 tons of vermicompost per hectare (3.27). At 60 days after transplanting, black rice treated with 20 tons of vermicompost per hectare produced the highest average number of tillers (7.70), followed by 90-30-30 kg NPK per hectare (7.37), 30 tons of vermicompost per hectare (7.23), and 25 tons of vermicompost per hectare (6.60). 110 days after transplanting, the highest average number of productive tillers was recorded in black rice treated with 20 tons vermicompost per hectare (8.10), followed by 90-30-30 kg NPK per hectare (7.73), 30 tons vermicompost per hectare (7.43), and 25 tons vermicompost per hectare (6.90) (**Table 1**). Statistical analysis using Analysis of Variance (ANOVA) revealed no significant differences among treatment means at the 0.05 level of significance in the three-observation period. This means that the effects of 20, 25, and 30 tons of vermicompost are comparable to those of synthetic fertilizer at a rate of 90-30-30 kg NPK per hectare. These findings are consistent with the reports of other studies that observed vermicompost application in rice significantly enhanced the number of tillers per hill, particularly when used with reduced chemical fertilizer input [29], [33]. A similar study highlighted the role of vermicompost in promoting microbial activity and nutrient availability, resulting in improved shoot development and tiller production [26]. Additionally, improvements in chlorophyll content and photosynthesis under vermicompost treatment contribute to higher tillering capacity [25]. Moreover, the improvement in tiller number may also be attributed to better soil physical and chemical properties brought about by vermicompost [27], [32]. The increased availability of NPK and enhanced soil structure favor root branching and subsequent tiller initiation. This suggests that vermicompost can provide a sustainable alternative to inorganic fertilizers for maintaining tiller production in rice cultivation.

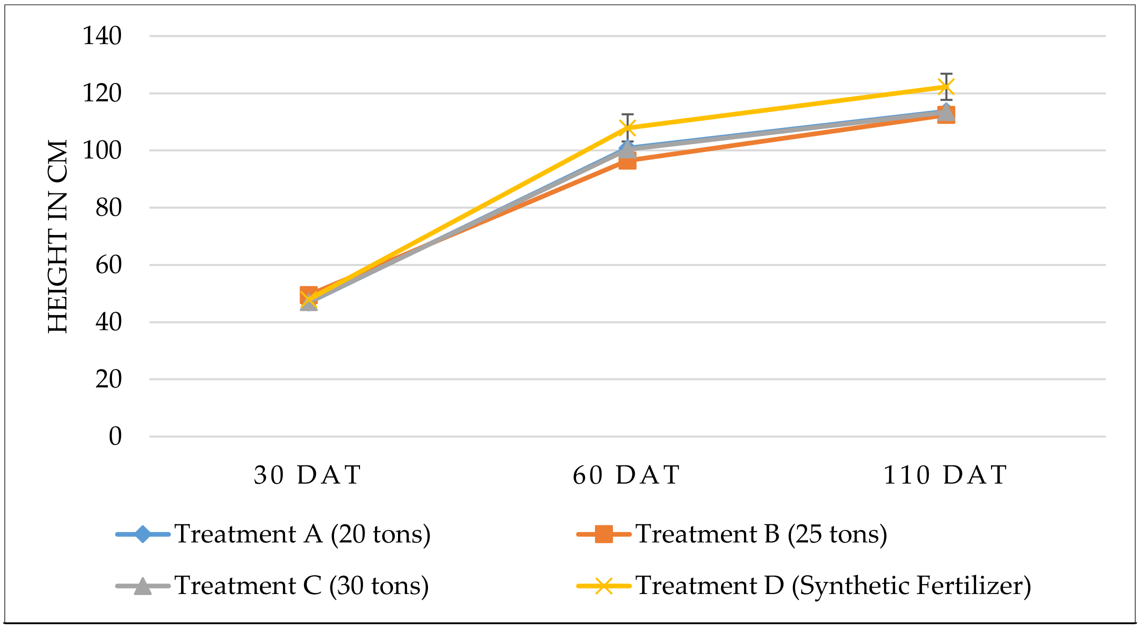


Figure 1. Height of the black rice at different stages of growth applied with different levels of vermicompost

Table 1. Number of tillers of black rice 30, 60, and 110 days after transplanting applied with different levels of vermicompost

Treatments	30 DAT	60 DAT	110 DAT
A – 20 t vermicompost/ha.	3.37 ± 0.41^{ns}	7.70 ± 0.85^{ns}	8.10 ± 1.71^{ns}
B – 25 t vermicompost/ha.	3.47 ± 0.25	6.60 ± 1.60	6.90 ± 1.99
C – 30 t vermicompost/ha.	3.27 ± 1.23	7.23 ± 0.81	7.43 ± 2.02
D – 90-30-30 kg NPK/ha.	4.00 ± 1.17	7.37 ± 1.76	7.73 ± 0.75

Values represent the mean ± standard deviation

^{ns}No Significant Difference ($p > 0.05$)

3.3 Yield Components of Black Rice at Different Levels of Vermicompost

The results showed that the average panicle length of black rice was highest in the 90-30-30 kg NPK per hectare treatment (22.17 cm), followed by applications of 20 tons vermicompost per hectare (21.43 cm), 30 tons vermicompost per hectare (21.31 cm), and 25 tons vermicompost per hectare (21.02 cm). The average number of filled grains per panicle was most significant in the treatment with 25 tons of vermicompost per hectare (72.75 grains), followed by 20 tons of vermicompost per hectare (70.28 grains), 30 tons of vermicompost per hectare (68.86 grains), and 90-30-30 kg NPK per hectare (61.55 grains). The average number of unfilled grains per panicle was highest in the 90-30-30 kg NPK per hectare treatment (57.75 grains), followed by 25 tons vermicompost per hectare (46.36 grains), 30 tons vermicompost per hectare (45.60 grains), and 20 tons vermicompost per hectare (44.98 grains). The fresh weight yield of black rice was highest in the 90-30-30 kg

NPK per hectare treatment (7.21 kg), followed by 20 tons vermicompost per hectare (6.92 kg), 30 tons vermicompost per hectare (6.67 kg), and 25 tons vermicompost per hectare (6.66 kg). The dry weight yield of black rice was highest in the treatment with 20 tons vermicompost per hectare (6.12 kg), followed by 90-30-30 kg NPK per hectare (5.71 kg), 30 tons vermicompost per hectare (5.59 kg), and 25 tons vermicompost per hectare (5.59 kg) (**Table 2**). Statistical analysis using the Analysis of Variance (ANOVA) for Randomized Complete Block Design (RCBD) revealed no significant differences among treatment means at the 0.05 level for all measured components. This indicates that the application of vermicompost at any tested level is comparable to the 90-30-30 kg NPK per hectare treatment in terms of panicle length, filled and unfilled grain count, and both fresh and dry weight yields.

Similar studies reported that vermicompost application significantly improved panicle traits and yield in rice [29], [33]. Similarly, vermicompost enriched with beneficial microbes enhances nutrient uptake and grain filling, contributing to improved productivity [26]. Likewise, the integration of vermicompost with chemical fertilizers resulted in longer panicles and higher grain yield than chemical fertilizers alone [28]. Additionally, improved photosynthetic capacity under vermicompost application supports higher grain filling [25], while organic amendments improve soil conditions and nutrient availability, thereby favoring yield components [27], [32]. Similar yield-enhancing effects of vermicompost have also been reported in non-rice crops. A significant increase in pod number, seed number, and seed biomass was observed in beans treated with vermicompost [31], while potatoes and beets achieved higher yields with vermicompost compared to conventional compost [34]. In vegetables, the application of vermicompost produced substantial increases in both the fresh and dry yields of pepper [35] and resulted in higher tomato yields [36]. Likewise, the combined application of vermicompost enhanced yields in spinach, peas, and radish under reduced chemical fertilizer treatments [37].

4. Conclusions

Based on the results of the study, it is concluded that any level of vermicompost from 20 to 30 tons showed comparable results to black rice in terms of plant growth, both in the growth indicators, such as number of tillers and height and in yield components such as length of panicle, both fresh and dry weight of grains, and the number of filled and unfilled grains with the use of inorganic fertilizer at the rate of 90-30-30 kg NPK per hectare.

Table 2. Yield components of black rice applied with different levels of vermicompost

Treatments	Panicle Length (cm)	Number of Filled Grains per Panicle	Number of Unfilled Grains per Panicle	Fresh Weight (kg)	Dry Weight (kg)
A – 20 t vermicompost/ha.	21.43 ± 0.61^{ns}	70.28 ± 14.57^{ns}	44.98 ± 9.17^{ns}	6.92 ± 1.38^{ns}	6.12 ± 1.33^{ns}
B – 25 t vermicompost/ha.	21.02 ± 0.27	72.75 ± 10.86	46.36 ± 5.79	6.66 ± 0.76	5.5 ± 0.71
C – 30 t vermicompost/ha.	21.31 ± 0.11	68.86 ± 10.66	45.60 ± 3.76	6.67 ± 1.01	5.54 ± 0.3
D – 90-30-30 kg NPK/ha.	22.17 ± 1.04	61.55 ± 4.53	57.75 ± 7.52	7.21 ± 0.76	5.71 ± 1.17

Values represent the mean ± standard deviation

^{ns}No Significant Difference ($p > 0.05$)

Acknowledgements The author sincerely thanks Professor Januaria Buncag and Professor Maria Olga Blanco for their invaluable guidance and support throughout the study.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] Kushwaha, U. K. S. *Black Rice: Research, History and Development*; Springer: 2016. https://doi.org/10.1007/978-3-319-30153-2

[2] Das, M.; Dash, U.; Mahanand, S. S.; Nayak, P. K.; Kesavan, R. K. Black rice: A comprehensive review on its bioactive compounds, potential health benefits, and food applications. *Food Chem. Adv.* **2023**, *3*, 100462. https://doi.org/10.1016/j.focha.2023.100462

[3] Panda, D. K.; Jyotirmayee, B.; Mahalik, G. Black rice: A review from its history to chemical makeup to health advantages, nutritional properties, and dietary uses. *Plant Sci. Today* **2022**, *9*, 1–15. https://doi.org/10.14719/pst.1817

[4] Dewi, K.; Agustina, R. Z.; Nurmalika, F. Effects of blue light and paclobutrazol on seed germination, vegetative growth, and yield of black rice (Oryza sativa L. 'Cempo Ireng'). *Biotropia* **2016**, *23*(2), 85–96. https://doi.org/10.11598/btb.2016.23.2.478

[5] Bhandari, R.; Ansari, M. J.; Alharbi, S. A.; Kushwaha, U. S.; Ghimire, P. Productivity and profitability of black rice as affected by transplanting methods and crop geometry. *Heliyon* **2024**, *10*(14), e34741. https://doi.org/10.1016/j.heliyon.2024.e34741

[6] Hijam, L.; Mandal, R.; Chakraborty, M.; Maying, B. Utilization of rice genetic resources for nutritional and medicinal benefits. In *Bioprospecting of Ethnomedicinal Plant Resources: Sustainable Utilization and Restoration*; Scopus, **2024**; pp 351–374.

[7] Tripathi, S.; Srivastava, P.; Devi, R. S.; Bhadouria, R. Influence of synthetic fertilizers and pesticides on soil health and soil microbiology. In *Agrochemicals Detection, Treatment and Remediation: Pesticides and Chemical Fertilizers*; Elsevier, 2020; pp 25–54. https://doi.org/10.1016/B978-0-08-103017-2.00002-7

[8] Guo, Y.; Wang, J. Spatiotemporal changes of chemical fertilizer application and its environmental risks in China from 2000 to 2019. *Int. J. Environ. Res. Public Health* **2021**, *18*(22), 11911. https://doi.org/10.3390/ijerph182211911

[9] Şener, S.; Türemiş, N. F.; Tanır, F. Agrochemical usage for sustainable fruit production and human health. In *Agrochemicals Detection, Treatment and Remediation: Pesticides and Chemical Fertilizers*; Elsevier, **2020**; pp 291–305. https://doi.org/10.1016/B978-0-08-103017-2.00012-X

[10] Tyagi, J.; Ahmad, S.; Malik, M. Nitrogenous fertilizers: Impact on environment sustainability, mitigation strategies, and challenges. *Int. J. Environ. Sci. Technol.* **2022**, *19*(11), 11649–11672. https://doi.org/10.1007/s13762-022-04027-9

[11] Anani, O. A.; Adetunji, C. O.; Osarenotor, O.; Inamuddin. Biofertilizer utilization in the agricultural sector. In *Biofertilizers: Study and Impact*; Wiley, **2021**; pp 293–307. https://doi.org/10.1002/9781119724995.ch9

[12] He, H.; Peng, M.; Hou, Z.; Li, J. Organic substitution contrasting direct fertilizer reduction increases wheat productivity, soil quality, microbial diversity and network complexity. *Environ. Technol. Innovations* **2024**, *36*, 103784. https://doi.org/10.1016/j.eti.2024.103784

[13] Banerjee, S.; Mitra, S.; Velhal, M.; Desmukh, V.; Ghosh, B. Impact of agrochemicals on the environment and human health: The concerns and remedies. *Int. J. Exp. Res. Rev.* **2021**, *26*, 125–140. https://doi.org/10.52756/ijerr.2021.v26.010

[14] Cruz, G. S. J.; Hermes, P. H.; Miriam, S. V.; López, A. M. Benefits of vermicompost in agriculture and factors affecting its nutrient content. *J. Soil Sci. Plant Nutr.* **2024**, *24*(3), 4898–4917. https://doi.org/10.1007/s42729-024-01880-0

[15] Hussain, N.; Abbasi, S. A. Efficacy of the vermicomposts of different organic wastes as "clean" fertilizers: State-of-the-art. *Sustainability* **2018**, *10* (4), 1205. https://doi.org/10.3390/su10041205

[16] Guerrero, R. D.; Rafael, D. Vermicomposting improves farm economics. *BioCycle* **2008**, *49*(8), 58–60.

[17] Zurbano, L. Y.; Cabanela, J. B.; Orijuela, N. P.; Villanueva, J. B. Response of string beans (*Vigna unguiculata* subsp. *sesquipedalis* L.) on saline soil amended with vermicompost. *Univ. J. Agric. Res.* **2022**, *10* (5), 563–568. https://doi.org/10.13189/ujar.2022.100511

[18] International Rice Research Institute. *Statistical Tool for Agricultural Research* [Computer software]. http://bbi.irri.org (accessed Sep 13, 2025).

[19] Mondal, M. M. A.; Puteh, A. B.; Razi Ismail, M.; Yusop Rafii, M. Optimizing plant spacing for modern rice varieties. *Int. J. Agric. Biol.* **2013**, *15*(1), 175–178.

[20] Zhang, J.; Xie, T.; Yang, W.; Zhou, G. Research status and prospect on height estimation of field crop using near-field remote sensing technology. *Smart Agric.* **2021**, *3*(1). https://doi.org/10.12133/j.smartag.2021.3.1.202102-SA033

[21] Kumar, C. P. S.; Naik, B. M.; Suji, D. B. Assessment of association analysis for grain yield and its attributing traits in rice (*Oryza sativa* L.). *Plant Arch.* **2020**, *20*, 1733–1737.

[22] Sun, J.; Ren, Z.; Cui, J.; Tang, C.; Luo, T.; Yang, W.; Song, P. A high-throughput method for accurate extraction of intact rice panicle traits. *Plant Phenomics* **2024**, *6*, 0213. https://doi.org/10.34133/plantphenomics.0213

[23] Samant, T. K. Impact of front line demonstration on yield and economics of hybrid rice (Rajalaxmi). *Indian J. Agric. Res.* **2015**, *49*(1), 88–91. https://doi.org/10.5958/0976-058X.2015.00013.X

[24] Xie, Y.; Zhou, L.; Dai, J.; Chen, J.; Yang, X.; Wang, X.; Wang, Z.; Feng, L. Effects of the C/N ratio on the microbial community and lignocellulose degradation during branch waste composting. *Bioprocess Biosyst. Eng.* **2022**, *45* (7), 1163–1174. https://doi.org/10.1007/s00449-022-02732-w

[25] Ruan, S.; Luo, H.; Zeng, X.; Wen, R.; Wu, F.; Tang, X. Effects of different vermicompost rates on growth, 2-acetyl-1-pyrroline, photosynthesis and antioxidant responses of fragrant rice (*Oryza sativa* L.) seedlings. *Phyton* **2021**, *90*(4), 1273–1283. https://doi.org/10.32604/phyton.2021.015610

[26] Mahanta, K.; Jha, D. K.; Rajkhowa, D. J.; Manoj-Kumar. Microbial enrichment of vermicompost prepared from different plant biomasses and their effect on rice (*Oryza sativa* L.) growth and soil fertility. *Biol. Agric. Hortic.* **2012**, *28*(4), 241–250. https://doi.org/10.1080/01448765.2012.738556

[27] Ramazanoglu, E. Effects of vermicompost application on plant growth and soil enzyme activity in wheat (*Triticum aestivum* L.) monitored by thermal imaging. *Cogent Food Agric.* **2024**, *10*(1), 2373872. https://doi.org/10.1080/23311932.2024.2373872

[28] Sungthongwises, K.; Tanpan, T. Effect of vermicompost-chemical mixed fertiliser on the growth and macronutrient use efficiency of upland rice cv. Sakonnakhon. *Songklanakarin J. Sci. Technol.* **2018**, *40*(5), 1039–1042.

[29] Rao, G. B. S.; Baradhan, G.; Sureshkumar, S. M.; Immanuel, R. R.; Ramesh, S. Influence of integrated nutrient management practices on growth and yield of rice–greengram cropping system. *Plant Arch.* **2019**, *19*, 441–443.

[30] Xu, C.; Mou, B. Vermicompost affects soil properties and spinach growth, physiology, and nutritional value. *HortScience* **2016**, *51*(7), 847–855. https://doi.org/10.21273/hortsci.51.7.847

[31] Aguilar-Benítez, G.; Peña-Valdivia, C. B.; Castro-Rivera, R.; Lara-Ávila, J. P.; Cruz-Crespo, E.; Rojas-Velázquez, A. N. Vermicompost and water stress effects on dry beans (*Phaseolus vulgaris* L.): Productive parameters and water relations. *Phyton-Int. J. Exp. Bot.* **2017**, *86*, 28–39.

[32] Tammam, A. A.; Shehata, M. R. A. M.; Pessarakli, M.; El-Aggan, W. H. Vermicompost and its role in alleviation of salt stress in plants–I. Impact of vermicompost on growth and nutrient uptake of salt-stressed plants. *J. Plant Nutr.* **2023**, *46*(7), 1446–1457. https://doi.org/10.1080/01904167.2022.2072741

[33] Manandhar, S.; Tuladhar, R.; Prajapati, K.; Singh, A.; Varma, A. Effect of *Azotobacter chroococcum* and *Piriformospora indica* on *Oryza sativa* in presence of vermicompost. In *Mycorrhiza – Nutrient Uptake, Biocontrol, Ecorestoration*, 4th ed.; Varma, A., Prasad, R., Tuteja, N., Eds.; Springer, **2018**; pp 327–339. https://doi.org/10.1007/978-3-319-68867-1_18

[34] Akbassova, A. D.; Sainova, G. A.; Aimbetova, I. O.; Sunakbaeva, D. K.; Akeshova, M. M. Impact of polyfunctional vermicompost on the productivity of vegetable root crops. *Fresenius Environ. Bull.* **2016**, *25*(9), 3755–3759.

[35] Zhang, M.; Liu, Y.; Wei, Q.; Liu, L.; Gu, X.; Gou, J.; Wang, M. Ameliorative effects of vermicompost application on yield, fertilizer utilization, and economic benefits of continuous cropping pepper in Karst areas of Southwest China. *Agronomy* **2023**, *13*(6), 1591. https://doi.org/10.3390/agronomy13061591

[36] Liu, X. C.; Chen, L.; Li, S. Q.; Shi, Q. H.; Wang, X. Y. Effects of vermicompost fertilization on soil, tomato yield and quality in greenhouse. *Chin. J. Appl. Ecol.* **2021**, *32*(2), 549–556. https://doi.org/10.13287/j.1001-9332.202102.022.

[37] Syed, S.; Saman, R. U.; Manzoor, A.; Arsalan, M.; Latif, A.; Khan, M.; Ullah, R.; Bilal, M.; Latif, R.; Ahmad, E.; Amjad, M.; Aslam, M.; Sarwar, S.; Shah, S. H.; Masood, A. Maximizing vegetable growth potential: Evaluating the synergistic impact of vermicompost and NPK fertilizer amended soil. *Pak. J. Agric. Res.* **2024**, *37*(3), 207–216. https://doi.org/10.17582/JOURNAL.PJAR/2024/37.3.207.216

ASEAN
Journal of Scientific and Technological Reports
Online ISSN:2773-8752

Research article

Light Curve and Period Change Behavior of a Binary Star V392 Orionis

Wiraporn Maithong[1]*, Torik Hengpiya[2], and Chaloemchon Wannathong[3]

[1] Faculty of Science and Technology, Chiang Mai Rajabhat University, Chiang Mai, 50300, Thailand
[2] National Astronomical Research Institute of Thailand (Public Organization), Songkhla, 90000, Thailand
[3] Faculty of Science Technology and Agriculture, Yala Rajabhat University, Yala, 95000, Thailand
* Correspondence: wiraporn_mai@cmru.ac.th

Citation:
Maithong , W.; Hengpiya , T.; Wannathong , C. Light curve and period change behavior of a binary star V392 orionis . *ASEAN J. Sci. Tech. Report.* **2025**, *28*(5), e259078. https://doi.org/10.55164/ajstr.v28i5.259078.

Article history:
Received: April 30, 2025
Revised: August 18, 2025
Accepted: September 1, 2025
Available online: September 14, 2025

Abstract: This study investigates the period change of the eclipsing binary system V392 Orionis (V392 Ori), which is classified as an Algol-type binary star. The research utilized data from the TESS Eclipsing Binary Catalog obtained via the Transiting Exoplanet Survey Satellite (TESS). Light curves were generated through photometric analysis to determine the times of minimum light. The historical and current observational data were used to construct the O–C diagram. The O–C diagram derived from PE and CCD observations exhibits a downward parabolic trend, indicating a decreasing orbital period at a rate of 5.76×10^{-3} seconds per year.

Keywords: V392 Ori; Period Change; O-C Diagram

1. Introduction

Binary systems are stellar systems composed of two stars orbiting a common center of mass under mutual gravitational attraction [1]. The study of binary star systems allows astronomers to determine the physical properties of stars and understand stellar evolution [2]. The binary star systems can be classified into two major types: visual binary systems, where individual stars can be distinguished through a telescope, and close binary systems, where the components are so close that they appear as a single star and can only be identified as binaries through other analytical methods. The type of a binary system is typically classified by constructing a light curve, which plots the variation in brightness over time. A dip in the light curve indicates an eclipse event where one star blocks the other along the observer's line of sight. In a whole orbital period of a binary system, two eclipses typically occur, corresponding to the primary and secondary eclipses. The time it takes for the system to complete one full orbit is referred to as the orbital period.

V392 Orionis (V392 Ori) is an Algol-type close eclipsing binary star with a spectral type of A5V. Its celestial coordinates are RA 06h 11m 25.17s and Dec +18° 32′ 59.60″. [3] The orbital period is 0.659284 days [4] and decreasing. [5] This study aims to analyze the period change of V392 Ori through its O–C diagram, contributing to a better understanding of the system's long-term behavior and evolution.

2. Materials and Methods

The study of the light variation of the binary star V392 Ori involved data collected from the TESS Eclipsing Binary Catalog, which contains observations from the Transiting Exoplanet Survey Satellite (TESS). The TESS, operated by NASA, is accessible via the website http://tessebs.villanova.edu [6]. The TESS conducts observations of exoplanets, variable stars, and binary star systems, collecting data within the wavelength range of 600–1,000 nanometers, which corresponds to portions of the R to I bands in the Johnson UBV photometric system, as illustrated in Figure 1. [7]

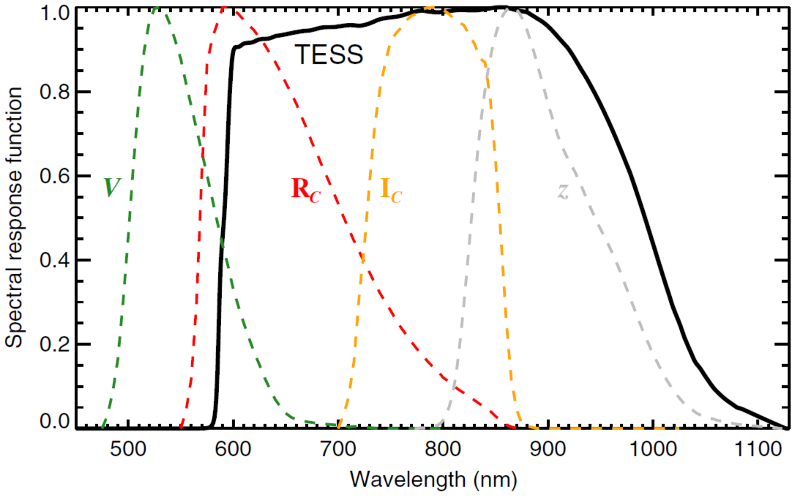


Figure 1. The black line is the wavelength bandwidth of the TESS observation.

This work uses Python code to download the data of V392 Ori, including the BJD time of observation and flux. The data at the latest time were used to plot the light curve as illustrated in Figure 2.

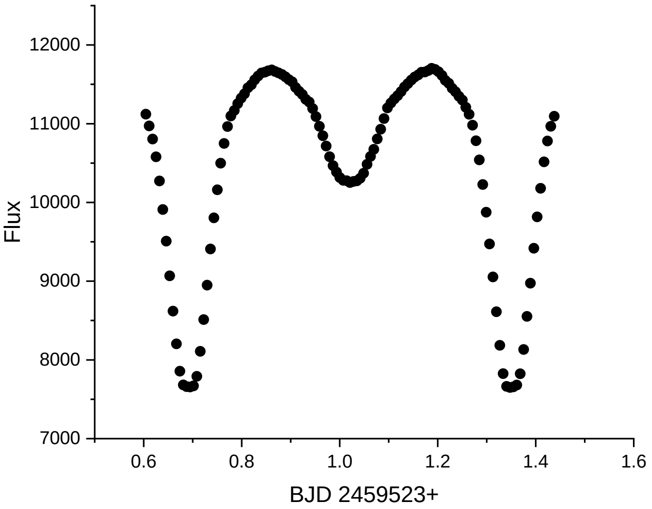


Figure 2. The light curve of V392 Ori, plotted from TESS observational data.

3. Results and Discussion

In Figure 2, the vertical axis represents the flux, and the horizontal axis represents time in Barycentric Julian Date (BJD), a time standard that accounts for the motion of light and other relevant factors by referencing the barycenter of the solar system. From this light curve, the times of minimum light were determined by fitting parabolic curves to the regions around the minima. The minimum time obtained in this study is BJD 2459523.691. The period change analysis was conducted following the procedure described by

Bob Nelson [8], as referenced by the American Association of Variable Star Observers (AAVSO), where the Heliocentric Julian Date (HJD) is converted to Barycentric Julian Date (BJD). The linear ephemeris used in this study is:

$$\text{Min. I} = \text{BJD } 2425506.620 + 0.659284\ E \quad (1)$$

Where: Min. I = BJD time of primary minimum
E = Epoch (number of orbital cycles since the initial reference time)

The O–C values of the eclipsing binary system V392 Ori, based on the times of minimum compiled by Bob Nelson [8], were computed from equation (1) as shown in Table 1.

Table 1. O–C values for the eclipsing binary system V392 Ori

Time of Minimum BJD	Epoch	O-C	Photograph	Source
2425506.62	0	0	pg	GCVS 4
2425506.624	0	0.004	pg	MVS 3.199
2426224.583	1089	0.002724	pg	MVS 3.199
2426743.425	1876	-0.011784	pg	MVS 3.199
2428453.629	4470	0.00952	pg	VSS 2.56
2428494.467	4532	-0.028088	pg	VSS 2.56
2428547.276	4612	0.038192	pg	VSS 2.56
2428809.597	5010	-0.03584	pg	VSS 2.56
2429167.598	5553	-0.026052	pg	VSS 2.56
2429229.567	5647	-0.029748	pg	VSS 2.56
2429231.59	5650	0.0154	pg	VSS 2.56
2429251.35	5680	-0.00312	pg	VSS 2.56
2429340.354	5815	-0.00246	pg	VSS 2.56
2429630.469	6255	0.02758	pg	VSS 2.56
2429641.317	6271.5	-0.002606	pg	VSS 2.56
2429729.328	6405	-0.00602	pg	VSS 2.56
2429943.596	6730	-0.00532	pg	VSS 2.56
2429957.471	6751	0.024716	pg	VSS 2.56
2430326.609	7311	-0.036324	pg	VSS 2.56
2430375.422	7385	-0.01034	pg	VSS 2.56
2430778.27	7996	0.015136	pg	MVS 3.199
2432118.601	10029	0.021764	pg	MVS 3.199
2432908.4	11227	-0.001468	pg	MVS 3.199
2433186.588	11649	-0.031316	pg	MVS 3.199
2433357.358	11908	-0.015872	pg	VSS 2.56
2433361.332	11914	0.002424	pg	MVS 3.199
2433378.15	11939.5	0.008682	pg	AC 151.26
2433597.32	12272	-0.033248	pg	AC 151.26
2433630.3	12322	-0.017448	pg	AC 151.26
2433659.29	12366	-0.035944	pg	AC 151.26
2434088.15	13016.5	-0.040186	pg	AC 151.26
2434420.12	13520	-0.01968	pg	AC 151.26
2434653.544	13874	0.017784	pg	MVS 3.199
2435135.41	14605	-0.05282	pg	MVS 3.199
2435164.446	14649	-0.025316	pg	MVS 3.199
2435892.339	15753	0.018148	pg	MVS 3.199

2435921.352	15797	0.022652	pg	MVS 3.199
2436604.322	16833	-0.025572	pg	MVS 3.199
2436608.28	16839	-0.023276	pg	MVS 3.199
2436896.412	17276	0.001616	pg	HABZ 98
2437312.418	17907	-0.000588	pg	MVS 644
2437345.369	17957	-0.013788	pg	HABZ 98
2437349.36	17963	0.021508	pg	HABZ 98
2437917.624	18825	-0.0173	pg	MVS 3.199
2438088.397	19084	0.001144	pg	HABZ 98
2438372.526	19515	-0.02126	pg	MVS 3.199
2438411.458	19574	0.012984	pg	MVS 3.199
2438440.458	19618	0.004488	pg	HABZ 98
2438670.58	19967	0.036372	pg	MVS 3.199
2439024.547	20504	-0.032136	pg	MVS 3.199
2439055.556	20551	-0.009484	pg	MVS 3.199
2439057.581	20554	0.037664	pg	MVS 3.199
2439059.517	20557	-0.004188	pg	HABZ 98
2439088.545	20601	0.015316	pg	MVS 3.199
2439528.276	21268	0.003888	pg	HABZ 98
2440981.332	23472	-0.002048	pg	HABZ 98
2441329.444	24000	0.007	pg	HABZ 98
2441331.416	24003	0.001148	pg	HABZ 98
2441333.394	24006	0.001296	pg	HABZ 98
2442404.725	25631	-0.004204	pg	BBS 19
2442417.28	25650	0.0244	pg	BBS 20
2442475.28	25738	0.007408	pg	HABZ 98
2443212.368	26856	0.015896	pg	BBS 33
2445052.421	29647	0.007252	pg	VSSC 60.22
2445054.398	29650	0.0064	pg	VSSC 60.22
2447231.361	32952	0.013632	vis	BBS 88
2447233.331	32955	0.00578	vis	BBS 88
2447554.39	33442	-0.006528	vis	BBS 91
2447564.282	33457	-0.003788	vis	BBS 91
2447566.28	33460	0.01636	vis	BAVM 52
2447587.367	33492	0.006272	vis	BBS 91
2447914.367	33988	0.001408	vis	BBS 94
2447922.308	34000	0.031	vis	BBS 94
2447941.423	34029	0.026764	vis	BBS 94
2447943.419	34032	0.044912	vis	BBS 94
2448272.361	34531	0.004296	vis	IBVS 4097
2448272.364	34531	0.006996	vis	IBVS 4097
2448307.308	34584	0.009144	vis	BBS 97
2448332.352	34622	0.000352	vis	BBS 97
2448688.365	35162	-8E-06	vis	BBS 101
2450390.627	37744	-0.009296	vis	BBS 116
2450483.597	37885	0.00126	PE	BBS 114
2451925.451	40072	0.001652	PE	IBVS 5296
2452647.368	41167	0.002772	PE	IBVS 5643
2452654.622	41178	0.004348	CCD	IBVS 5502
2452655.278	41179	0.000964	CCD	IBVS 5438

2452715.274	41270	0.00182	CCD	IBVS 5676
2452998.766	41700	0.002	CCD	IBVS 5493
2453354.121	42239	0.002724	CCD	VSB 43
2453450.376	42385	0.00286	CCD	IBVS 5741
2453671.897	42721	0.003936	CCD	IBVS 5677
2453674.535	42725	0.0048	CCD	IBVS 5741
2453696.291	42758	0.004228	CCD	VSB 44
2453700.246	42764	0.003924	CCD	VSB 44
2453715.41	42787	0.004592	CCD	IBVS 5741
2453758.261	42852	0.001632	PE	IBVS 5802
2454068.788	43323	0.006268	CCD	IBVS 5760
2454091.534	43357.5	0.00667	PE	IBVS 5761
2454476.553	43941.5	0.003614	PE	IBVS 5874
2454521.382	44009.5	0.001652	CCD	OEJV 0094
2454809.161	44446	0.003436	CCD	VSB 48
2454829.599	44477	0.003232	PE	IBVS 5918
2454881.684	44556	0.004996	CCD	IBVS 5894
2455135.177	44940.5	0.003798	CCD	VSB 50
2455136.165	44942	0.002172	CCD	VSB 50
2455158.913	44976.5	0.004974	CCD	IBVS 5920
2455163.854	44984	0.001544	CCD	IBVS 5929
2455244.287	45106	0.001796	PE	IBVS 5959
2455591.73	45633	0.002128	CCD	IBVS 5992
2455602.278	45649	0.001684	PE	IBVS 6070
2455838.629	46007.5	-0.00063	CCD	OEJV 0142
2456256.945	46642	-0.000828	CCD	IBVS 6042
2457042.487	47833.5	0.004586	PE	IBVS 6152
2457296.632	48219	-0.004196	CCD	IBVS 6230
2457414.308	48397.5	-0.01029	CCD	IBVS 6196
2457702.091	48834	-0.004556	CCD	VSB-063
2458058.105	49374	-0.004416	CCD	VSB-64
2458219.63	49619	-0.004178	CCD	JAVSO 49, 106
2458785.954	50478	-0.005088528	CCD	RHN 2019
2459523.691	51597	-0.0052129	CCD	This Study
2459938.063	52225.5	0.005658	CCD	VSB, 108

Note:

pg (photographic): Refers to observations derived from astronomical photographic plates or films, which were commonly used in early stellar photometry.

vis (visual): Refers to visually estimated observations, obtained either with the unaided eye or through a telescope. The timing of events is subject to the observer's judgment and experience.

PE (photoelectric): Refers to measurements acquired using a photoelectric photometer, a method prevalent before the development of CCD technology. This technique provided improved accuracy over visual methods.

CCD (Charge-Coupled Device): Refers to observations obtained using CCD sensors, a modern and highly precise technique in photometric measurements that has become the standard in contemporary astronomical research. The O–C values derived from the observed times of minimum light (O) and the predicted times (C) were plotted against the corresponding epochs to produce the O–C diagram. This study considers the period change in three solutions: the first is computed from all data, the second calculates from vis to present, and the last one analyzes especially PE and CCD. The first solution is the O-C diagram of V392 Ori, as shown in Figure 3.

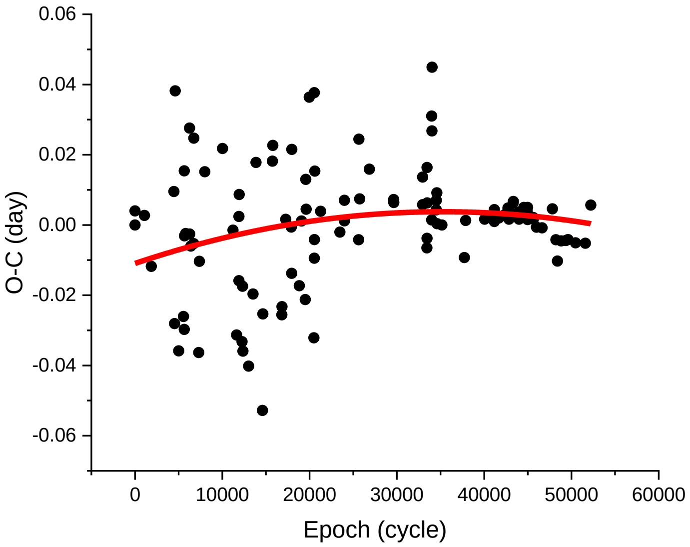


Figure 3. The O-C diagram for V392 Ori, based on all available data.

The parabolic nature of the O–C diagram indicates a trend of increasing orbital period. A second-order polynomial fit was applied using the Quadratic Polynomial Fitting Method, resulting in the following equation:

$$O\text{-}C = -0.01094 + 8.35174 \times 10^{-7}\ \text{Epoch} - 1.18474 \times 10^{-11}\ \text{Epoch}^2 \qquad (2)$$

From this, the rate of orbital period change (dP/dE) can be calculated as:

$$\begin{aligned} dP/dE &= 2(-1.18474 \times 10^{-11}) \\ &= -2.3694 \times 10^{-11}\ \text{day/cycle} \\ &= -1.134 \times 10^{-3}\ \text{sec/year} \end{aligned}$$

This negative value of dP/dE indicates that the orbital period of the binary system V392 Ori is decreasing. This suggests that the distance between the two stellar components is decreasing, consistent with the Angular Momentum Loss (AML) theory [9]. The loss of angular momentum in a binary star system drives orbital evolution, which may ultimately result in the merger of the two components into a single, rapidly rotating star. The second solution, the O-C diagram of V392 Ori, was computed from data except for pg, as shown in Figure 4.

The O-C diagram from Figure 4 shows that there is a periodic pattern (blue line) on the second-order polynomial line (red line). The second-order polynomial and sinusoidal fit to the O–C data yielded the following equation:

$$\begin{aligned} O\text{-}C = {} & 0.05561 - 1.7805 \times 10^{-6}\ \text{Epoch} + 1.2656 \times 10^{-11}\ \text{Epoch}^2 \qquad (3) \\ & + -0.00145 + 0.00349 \sin(\pi((\text{Epoch}- 1447.9893)/ 4972.98766)) \end{aligned}$$

So,

$$\begin{aligned} dP/dE &= 2 \times 1.2656\text{x}10^{-11} \\ &= 2.5312\ \text{x}10^{-11}\quad \text{day/cycle} \\ &= 1.21\ \text{x}10^{-3}\quad \text{sec/year} \end{aligned}$$

The positive rate of change of the period (1.21 × 10^{-3} seconds/year) suggests that the orbital separation between the two stars is increasing. This supports the Thermal Relaxation Oscillation (TRO) theory [10], which suggests that variations in the internal structure and energy balance of the components, combined with mass transfer between them, can drive cyclic changes between contact and semi-detached configurations, thereby influencing the orbital evolution of the system. This periodic variation indicates the likely presence of a third body in the system. The light-time effect gives an orbital radius of approximately 0.67 AU for the third body, with an orbital period of about 17.95 years. The third solution, the O-C diagram of the V392 Ori, was calculated from PE and CCD data as shown in Figure 5.

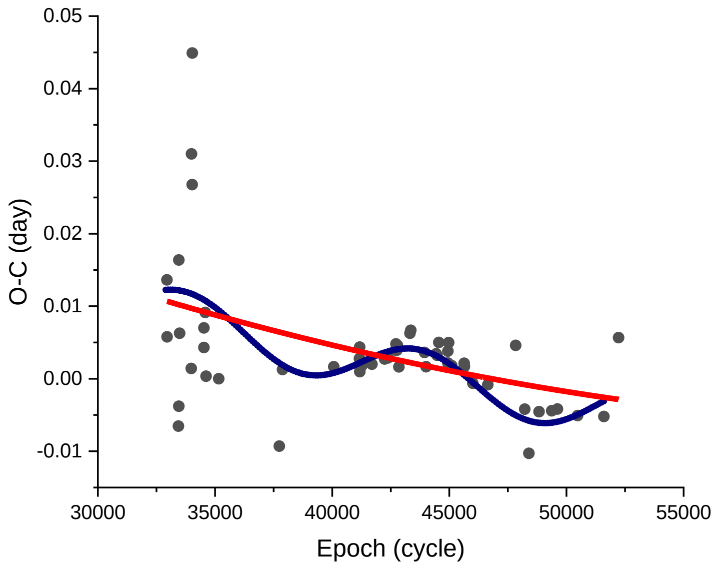


Figure 4. The O-C diagram for the V392 Ori from vis to present data.

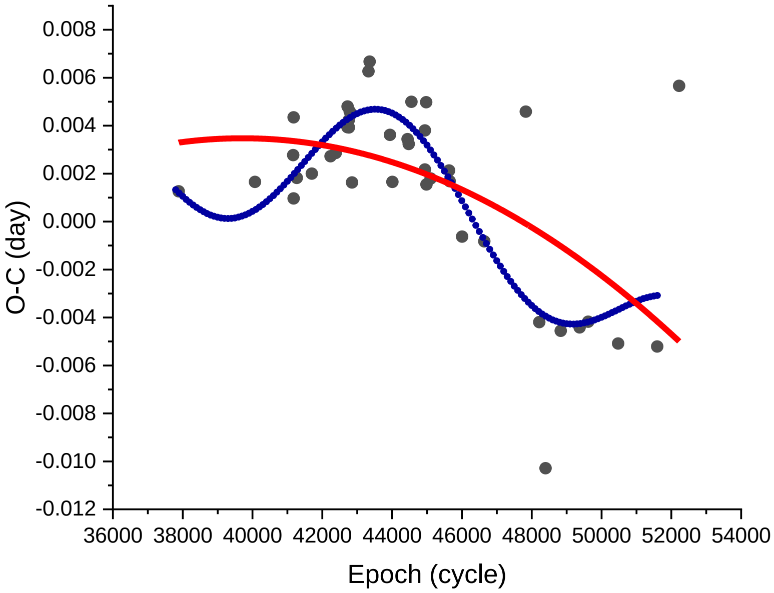


Figure 5. The O-C diagram for V392 Ori, based on PE and CCD data.

The O-C diagram from Figure 5 exhibits a similar pattern to that in Figure 3, which is a periodic pattern (blue line) on the second-order polynomial line (red line). Therefore, the O–C data yielded the following equation:

$$\begin{aligned}\text{O-C} &= -0.08178 + 4.29419 \times 10^{-6}\,\text{Epoch} - 5.40715 \times 10^{-11}\,\text{Epoch}^2 \\ &\quad -0.000642856 + 0.0027\sin(\pi((\text{Epoch}- 23541.55743)/ 4510.50998))\end{aligned} \qquad (4)$$

So,

$$dP/dE = -5.176 \times 10^{-3} \quad \text{sec/year}$$

The negative rate of change of the period (-5.176×10^{-3} seconds/year) is consistent with the evolution of AML theory. The sinusoidal variation may indicate the presence of a third body in this system, which is positioned approximately 0.47 AU from the primary body and has an approximate orbital period of 16.28 years.

4. Conclusions

This study of the eclipsing binary system V392 Ori used data from the TESS Eclipsing Binary Catalog, obtained via the TESS satellite. A light curve was constructed and analyzed to determine the times of minimum light, which were then used to build an O–C diagram. The analysis using the Quadratic Polynomial Fitting Method revealed three solutions for the orbital period of V392 Ori. The second condition was calculated from the previous data to the present data. The orbital period was found to be increasing at a rate of 1.21×10^{-3} seconds/year, suggesting ongoing mass transfer and cyclical structural changes in the binary system, as predicted by the TRO theory. Furthermore, periodic variations in the O–C diagram indicate the potential presence of a third body with an orbital radius of about 0.67 AU and a period of approximately 17.95 years. On the other hand, the first condition was computed from all data, while the last solution was computed from data excluding pg and vis. The period change is decreasing at rates of 1.134×10^{-3} and 5.176×10^{-3} seconds/year, respectively. The evolution corresponds to the AML theory. Furthermore, the last solution presented for the third body's behavior is at a distance of 0.47 AU, with an orbital period of around 16.28 years. When considering the observations in modern astronomy, which includes modern instrumental photography, the third condition may show a trustworthy solution. However, this binary should be observed further to study its evolution.

5. Acknowledgements

The author would like to thank the National Aeronautics and Space Administration for the TESS Eclipsing Binary Catalog.

Author Contributions: Conceptualization, methodology, investigation, W.M., formal analysis, writing—original draft preparation, C.W. and W.M.; Observation, T.H., writing—review and editing, and project administration, W.M.

Funding: This research was funded by Thailand Science Research and Innovation (TSRI) through Chiang Mai Rajabhat University, Fiscal Year 2025..

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] Maithong, W.; Tajumpa, **A.** Period Change and Third Body Study of a Binary System V781 Tauri. *Udon Thani Rajabhat University Journal of Science and Technology* **2019**, *7*(2), 179–187. (in Thai)

[2] Soonthornthum, B. *Astrophysics*.. Chiang Mai University. **2007**. (In Thai)

[3] SIMBAD Astronomical Database - CDS (Strasbourg). (Accessed 2024, September 22). http://simbad.u-strasbg.fr/simbad/

[4] Liu, X.; Tan, H.-S.; Leung, K. UBV observations of short-period close binary V392 Orionis. *Astrophysics and Space Science* **1988**, *141*(1), 21–25. https://doi.org/10.1007/BF00641912

[5] Hong, K.; Lee, J. W.; Koo, J.; Park, J.; Kim, S.; Rittipruk, P.; Kanjanasakul, C.; Han, C. Time-series Spectroscopy of the Oscillating Eclipsing Algol System V392 Orionis. *The Astronomical Journal* **2019**, *157*, 1–8. https://doi.org/10.3847/1538-3881/aaf39f

[6] TESS Eclipsing Binary Catalog. (Accessed 2024, September 22). http://tessebs.villanova.edu/

[7] Ricker, G. R. et al. Transiting Exoplanet Survey Satellite. *Journal of Astronomical Telescopes, Instruments, and Systems* **2015**, *1*(1), 014003-1–10.

[8] Bob Nelson’s Database of Eclipsing Binary O-C Files, AAVSO. (Accessed 2024, October 22). https://www.aavso.org/bob-nelsons-o-c-files

[9] Qian, S. Orbital Period Changes of Contact Binary Systems: Direct Evidence for Thermal Relaxation Oscillation Theory. *Mon. Not. R. Astron. Soc.* **2001**, *328*, 914–924. https://doi.org/10.1046/j.1365-8711.2001.04921.x

[10] Lucy, L. B.; Wilson, R. E. Observational Tests of Theories of Contact Binaries. *The Astrophysical Journal* **1979**, *231*, 502–513. https://doi.org/10.1086/157212

ASEAN
Journal of Scientific and Technological Reports
Online ISSN:2773-8752

Academic article

Morphological Profiling of Lung Cancer Through Explainable Machine Learning

Karuna Tanthanawarakun[1], and Prompong Sugunnasil[2*]

[1] Faculty of Engineering, Chiang Mai University Chiang Mai, Chiang Mai, 50200, Thailand
[2] College of Arts, Media, and Technology, Chiang Mai University, Chiang Mai, 50200, Thailand
[*] Correspondence: prompong.sugunnasil@cmu.ac.th

Citation:
Tanthanawarakun, K.; Sugunnasil, P. Morphological profiling of lung cancer through explainable machine learning. *ASEAN J. Sci. Tech. Report.* **2025**, *28*(5), e257888. https://doi.org/10.55164/ajstr.v28i5.257888.

Article history:
Received: February 14, 2025
Revised: August 19, 2025
Accepted: September 1, 2025
Available online: September 14, 2025

Abstract: Lung cancer continues to be the predominant cause of cancer-related mortality globally, representing 11.6% of all cancer cases as reported by WHO, with approximately 2.2 million new diagnoses and 1.8 million fatalities recorded in 2020. For lung cancer diagnosis, Computed Tomography (CT) imaging serves as a critical tool in identifying both solid and subsolid ("ground glass") nodules. Although CT image segmentation has demonstrated significant clinical value, healthcare practitioners require a comprehensive understanding of the underlying algorithmic mechanisms to ensure diagnostic precision. This research investigates the application of interpretable machine learning methods for feature extraction from lung CT imaging data. We conduct a comparative analysis between transparent and opaque classification algorithms utilizing a comprehensive dataset comprising 1,229 normal and 1,010 abnormal pulmonary CT scans. The processed data undergoes evaluation using both interpretable models (including Logistic Regression, Decision Trees, and K-Nearest Neighbor) and black-box models (such as Multi-Layer Perceptrons, Convolutional Neural Networks, and Support Vector Machines). Our findings indicate that interpretable algorithms consistently outperform their black-box counterparts across multiple metrics. The evaluation framework incorporates accuracy, F1 score, precision, recall, computational efficiency, and resource utilization measurements. Results demonstrate exceptional classification accuracy for pulmonary malignancy detection while preserving explanatory capability, thereby providing clinicians with both practical and transparent diagnostic assistance. This investigation contributes to the development of accountable artificial intelligence systems for deployment in mission-critical healthcare environments.

Keywords: CT Scan image; lung cancer image; explainable AI; encoding; machine learning

1. Introduction

Lung cancer persists as one of the leading causes of cancer-related death in both men and women globally [1]. Two main categories of lung cancer are identifiable: non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC) Transformation from NSCLC to SCLC has been reported in clinical studies [2]. NSCLC represents approximately 85% of all lung cancer cases [3], encompassing various histological subtypes including adenocarcinoma, squamous cell carcinoma, and large cell carcinoma. Alarmingly, 70% of patients receive diagnoses at advanced stages, and 15% of cases remain undetected until reaching late stages [4]. The mortality incidence (MI), representing the ratio between fatalities and total diagnosed patients, reaches a staggering 80%.Consequently, developing precise detection techniques for lung cancer has become critically important.

Computed tomography scanning (CT scan) functions as a diagnostic tool that synthesizes multiple computerized views to create detailed internal visualizations of bodily structures. The scan yields valuable information regarding various pulmonary conditions, including pneumonia, cancer, vascular thrombosis, smoking-induced damage, and COVID-19 infection. With increasing reliance on CT scans, radiologists face mounting pressure to detect, analyze, and quantify these images in real-time despite resource constraints. Computerized techniques offer potential solutions to reduce resource demands, particularly in terms of time and personnel requirements, in lung cancer screening protocols. Various approaches, including Machine Learning [5], the Synthetic Minority Over-sampling Technique (SMOTE) [4], and Deep Learning [6], have been proposed to enhance image analysis processes.

Given that analytical outcomes directly impact human lives, simplistic binary classifications (such as benign or malignant) prove insufficient for justifying the utilization of information[7-8]. Despite achieving accurate results, the opaque nature of many machine learning algorithms remains problematic, with decision-making processes poorly understood. Consequently, medical applications of machine learning necessitate additional supporting evidence to validate methodological approaches. The interpretability and explainability of machine learning algorithms for CT scan image analysis, therefore, emerge as crucial drivers in medical image processing [9]. The benefits of explainable machine learning extend beyond mere accountability to include model refinement and improvement.

In this research, we propose a comprehensive comparative study between conventional machine learning approaches and explainable machine learning methodologies for feature extraction from lung CT images. Our primary objective is to classify images, determining whether specimens exhibit benign or malignant characteristics. The dataset comprises paired associations between lung CT images and corresponding feature vectors, with each vector element representing quantitative image information, such as the quantities of nodules of specific sizes in designated lung regions. Features undergo exhaustive generation based on predefined criteria, including nodule dimensions and anatomical location [10]. SCLC displays distinctive characteristics from NSCLC, notably exhibiting more aggressive growth patterns and a greater propensity for widespread metastasis [11]. CT-based lung cancer detection methodologies categorize findings according to cellular structure and anatomical location, providing detailed cross-sectional visualizations [12]. While CT scanning offers valuable diagnostic capabilities, including screening program support, three-dimensional imaging, and biopsy guidance, limitations persist regarding ionizing radiation exposure, false-positive results, and dependence on expert radiological interpretation [13].

Our research introduces an innovative data engineering technique designed to enhance classification efficiency for lung cancer diagnosis using CT scan images [14]. We employ explainable machine learning methods to capture distinctive imaging features and conduct rigorous comparative analysis of data encoding using interpretable classification algorithms [15]. Our findings consistently demonstrate that explainable approaches outperform non-interpretable alternatives, yielding superior diagnostic accuracy for pulmonary malignancies [16].

2. Materials and Methods

2.1 Interpretability and Explainable Machine Learning

The field of explainable machine learning addresses the critical need for algorithmic transparency, particularly in healthcare applications where understanding model decisions directly impacts patient care. While conventional machine learning often prioritizes performance metrics, explainability focuses on providing human-comprehensible justifications for model predictions [17]. Interpretability encompasses multiple dimensions, including algorithm transparency, feature relevance, and decision logic visualization. As noted by Doshi-Velez and Kim, interpretability in machine learning can be defined as "the ability to explain or present in understandable terms to a human" [9]. This definition emphasizes the human-centric nature of explainability, recognizing that different stakeholders—such as clinicians, patients, and researchers—may require different types of explanations. For medical imaging applications, model interpretability facilitates trust establishment among healthcare professionals by revealing the reasoning behind diagnostic suggestions. This transparency becomes particularly crucial when algorithms influence treatment decisions with potentially life-altering consequences [8]. Our research employs interpretability methods that strike a balance between explanatory clarity and diagnostic accuracy, addressing the unique challenges posed by pulmonary nodule classification.

2.2 CT Image Feature Analysis

Computed tomography feature analysis involves the systematic extraction and quantification of imaging characteristics that correlate with pathological findings. In lung cancer assessment, these features capture crucial information about nodule properties, including size, shape, margin characteristics, and internal density patterns [10]. Pulmonary nodules in CT images typically present as approximately spherical opacities with diameters of less than 30mm. These can be classified based on their density profiles as solid, part-solid, or ground-glass opacities, each carrying different clinical implications regarding malignancy potential [11]. Feature extraction algorithms quantify these visual characteristics into numerical descriptors, making them suitable for algorithmic processing. Our methodology employs a comprehensive feature extraction approach that captures first-order statistics (intensity distribution measures), shape metrics (morphological descriptors), texture features (spatial pixel relationship patterns), and wavelet-based characteristics [4]. By transforming visual data into quantitative measurements, these techniques enable precise comparison between cases and facilitate reproducible assessment of subtle imaging findings that may elude visual detection.

2.3 Detection of Lung Cancer Disease Based on CT Image

Lung cancer detection from CT images presents a complex diagnostic challenge, requiring the precise identification of potentially malignant pulmonary nodules against a background of typical anatomical structures. The process begins with image preprocessing and segmentation to isolate lung parenchyma from surrounding tissues [12]. Computer-aided detection systems have evolved significantly, progressing from basic rule-based algorithms to sophisticated deep learning architectures. Early detection methods demonstrated limited sensitivity, while contemporary approaches achieve substantially improved performance. As demonstrated by Li et al., automated systems utilizing deep features extracted from autoencoders can achieve accuracy levels of 75.01% with a sensitivity of 83.35% [18]. Research by Faisal and colleagues explored various machine learning approaches for nodule classification, finding that ensemble methods achieved impressive accuracy ratings exceeding 99% under controlled testing conditions [19]. Additionally, multilayer architectures employing convolutional networks with fusion-based classification have demonstrated detection rates above 99% with minimal false positives per image [20]. Our approach incorporates explainable feature engineering techniques that enhance diagnostic transparency while maintaining high classification performance. This methodology addresses the fundamental challenge of providing interpretable results that can meaningfully inform clinical decision-making beyond simple binary classification outputs.

2.4 Algorithmic Transparency in Clinical Decision Support

Machine learning algorithms in medical image analysis can be categorized along a spectrum of interpretability, from inherently transparent "white box" models to complex "black box" systems [21]. White box approaches—including linear regression, decision trees, and logistic regression—provide clear relationships between input features and predictions. Conversely, black box models, such as deep neural networks and complex ensembles, often achieve superior accuracy but operate through mechanisms that are resistant to straightforward human understanding. This tradeoff between transparency and performance presents significant challenges in clinical implementation. As demonstrated by studies utilizing the IQ-OTH/NCCD dataset, support vector machines with polynomial kernels achieved 89.89% classification accuracy, substantially outperforming linear alternatives (82.02%) [22]. This performance differential highlights the complexity of nodule classification tasks while emphasizing the need for sophisticated modeling approaches. The emerging field of Explainable AI (XAI) addresses transparency limitations through techniques that illuminate decision processes, including feature importance ranking, partial dependence analysis, and model-agnostic interpretation methods [23]. Notably, comparative studies have shown that decision tree algorithms can achieve remarkable performance (93.24% accuracy) even without extensive preprocessing, highlighting the potential value of inherently interpretable models in clinical contexts [16]. Our methodology synthesizes these insights by implementing a balanced approach that enhances model transparency through feature engineering while preserving classification performance. This strategy aligns with the requirements for medical AI systems that provide not only accurate results but also explainable outputs that integrate meaningfully into clinical workflows.

2.5 Methods

This section outlines the specifics of the proposed methodology, detailing the processes and methods used in conducting the research. This encompasses the study's design.

Overview of the proposed architecture

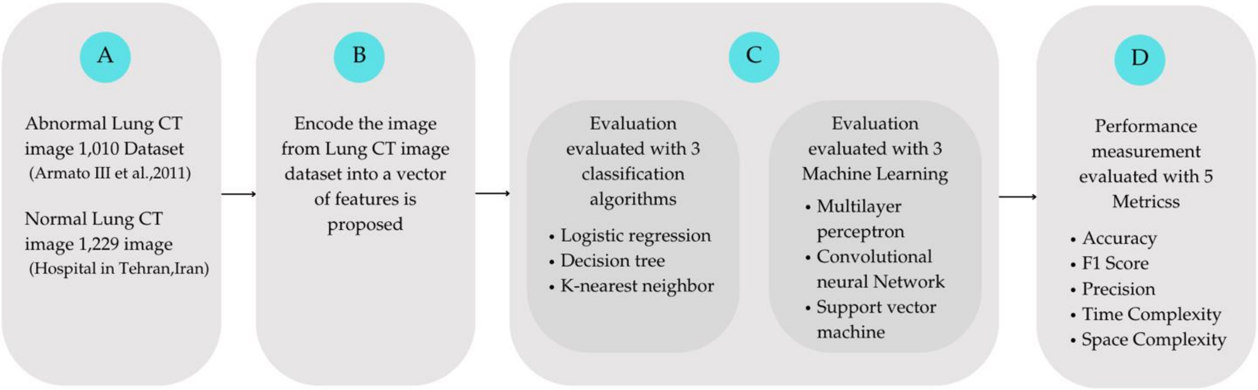


Figure 1. The overall architecture of the experiment.

Feature engineering plays a crucial role in analyzing lung cancer using CT scans. In this intricate process, we transform the raw data of the images into a structured format by converting salient features into pixel values that a computational model can interpret and understand. The first step involves identifying and extracting nodules, which are small, rounded opacities within the pulmonary tissue. Additionally, we isolate solid masses, which are denser, irregular growths that could indicate more advanced pathology. By distilling these complex visual cues into a set of features, we enable the algorithm to identify patterns and facilitate the accurate diagnosis of lung cancer. (Figure 1) presents an overview of the proposed architecture. (a) The input consists of a lung CT image dataset that includes normal and non-normal lung cancer CT images. We performed an image pre-processing to reduce biased information and prevent overfitting. The input consists of two types of images with the same format, which have been reshaped, resized, and reformatted. (b) This study proposes a process for encoding images from the lung CT image dataset into a vector of features. In the process of feature engineering for lung cancer detection using CT scans, we convert image-based features into quantifiable pixel values. This involves the meticulous extraction of both nodules and solid masses, transforming them into discernible features for further analysis (c) In addition, we compared the proposed image classification process with other baseline models using two types of evaluations: Explainable machine learning (i.e Logistic regression, Decision tree and K-nearest neighbor) and unexplainable algorithm (i.e. Multilayer perceptron, Convolutional Neural Network and Support vector machine). (d) The final step is performance measurement using five classification metrics: Accuracy, F1 Score, Precision, Time Complexity, and Space Complexity. We propose that the size and location of the nodule affect the outcome.

2.5.1 Collection of datasets

We collected the Lung Image Database Consortium Image Collection (LIDC-IDRI) [24] Figure 2) ,which consists of 244,527 diagnostic and lung cancer screening thoracic computed tomography (CT) scans fro 1,010 patients. Additionally, we collected normal CT images from large public Covid-19 (SARS-CoV-2) lung CT scan datasets, which include 1,229 negative cases (i.e., both normal and non-COVID images). The data is available as 512x512px PNG images and has been collected from real patients in the radiology centers of teaching hospitals in Tehran, Iran [25]. In conclusion, we have 1,224 cases for normal lung CT scans and 1,010 cases for lung with cancer.

2.5.2 Pre-Processing Data and Feature Extraction

This experiment requires data pre-processing because raw CT scan files are not directly applicable to the classification process. Since the images originate from different sources, we standardized them through reformatting, recoloring, and resizing to reduce potential bias in the dataset. This process helps ensure consistency across input images during the feature extraction phase. For the first dataset, which comprised non-normal images, we used DICOM files—an established standard for medical imaging that healthcare providers and equipment manufacturers have widely adopted. In this study, we focused exclusively on the lung region by removing unnecessary parts, such as the heart and background structures.

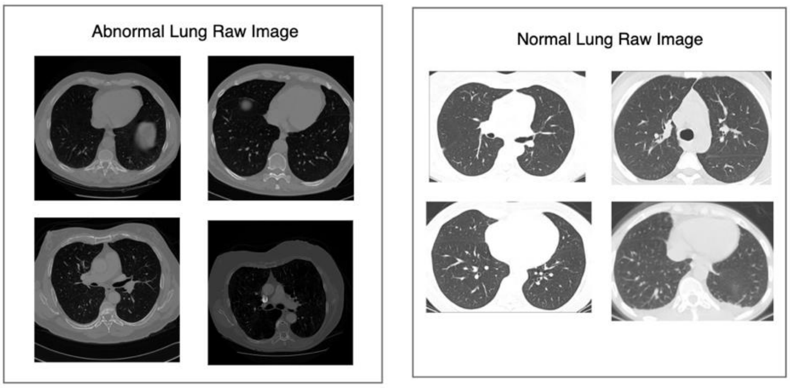


Figure 2. The example of the Abnormal Lung and Normal Lung Raw Dataset.

The CT scan lung images in this dataset exhibit multi-dimensional. Our approach involves initially converting these images from a 3D format to a 2D representation and subsequently performing a precise cropping operation to extract the lung region. During the lung region extraction process, we calculated the area encompassing most of the lung region, ensuring that it encapsulates the most relevant portion of the complete image. Subsequently, we reformatted the entire image into a PNG file as part of the preprocessing procedure. The second dataset comprises normal lung images, consisting of 1,229 negative cases. These normal lung images are stored as 2D PNG files with dimensions of 415x345 pixels, which differ from the size of the non-normal lung images. To maintain uniformity with the non-normal lung images, we conducted resizing and zoom-out operations. Following this, we utilized the OpenCV (cv2) library for preprocessing and executed cropping operations to isolate the lung region accurately.

2.5.2.1 Data Pre-processing

To ensure consistency, interpretability, and computational efficiency, two distinct pre-processing pipelines were employed in this study. These pipelines were tailored to support the different experiments. The aim was to transform raw lung CT scan images, both normal and abnormal, into structured feature sets suitable for training and evaluating machine learning models. In the first experimental approach, CT scan images were transformed into raw pixel-level vectors, preserving the grayscale intensity values directly from the image arrays. Non-Normal Dataset: DICOM files were subjected to lung segmentation using masking techniques to isolate lung regions see (Figure 3). Each image slice was then binarized, converted to grayscale, and flattened into one-dimensional pixel vectors(Figure 4). These vectors were exported as CSV files for model input.

(A) Original (B) Mask of Lungs (C) Segmented Lungs

Figure 3. The Example of Abnormal Lung and Normal Lung Masking Process.

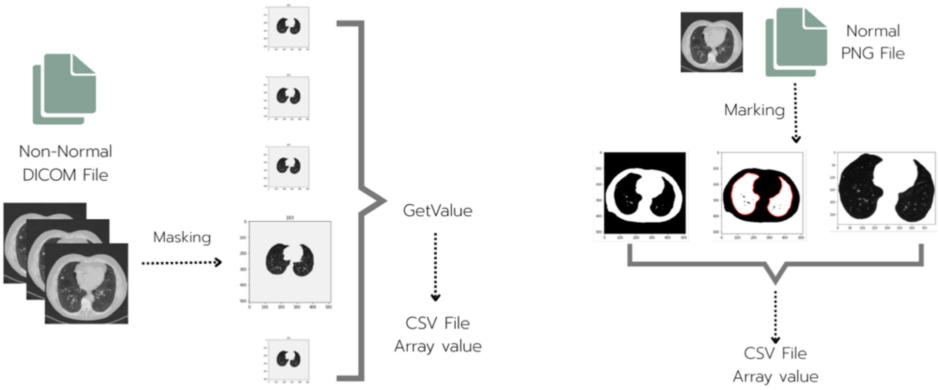


Figure 4. AbNormal Lung and Normal Feature Annotation and Vector Encoding Process.

The normal dataset is composed of PNG image files underwent manual region-of-interest (ROI) annotation to isolate lung regions. The annotated images were processed in the same manner as the non-normal dataset, ensuring consistency in vector representation (Figure 4).

The resulting feature vectors served as direct input for both explainable models (Logistic Regression, Decision Tree, KNN) and traditional machine learning models (CNN, MLP, SVM). The second dataset design employed an analytical approach that incorporated advanced morphological feature engineering techniques to extract measurable attributes from pulmonary radiographs. The methodology involved a series of systematic procedures executed in sequential order. Employing the OpenCV (cv2) computational library, comprehensive contour analysis was conducted to identify and differentiate various pulmonary structures. This analytical procedure facilitated the detection of potential malignant and benign nodules, parenchymal tissue characteristics, vascular elements, and nodule candidates. The implementation of advanced computer vision algorithms enabled the precise delineation of these structures based on their radiographic density, morphology, and spatial relationships.

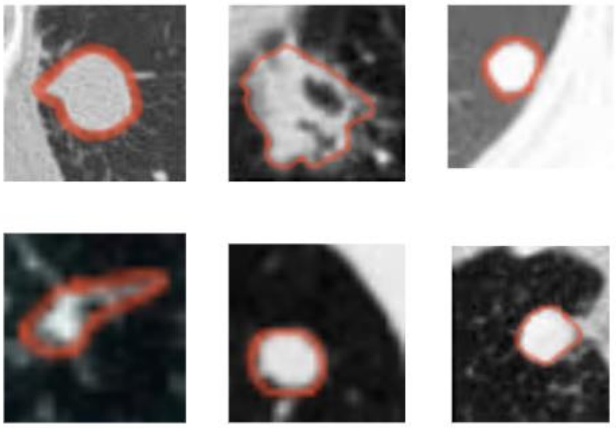
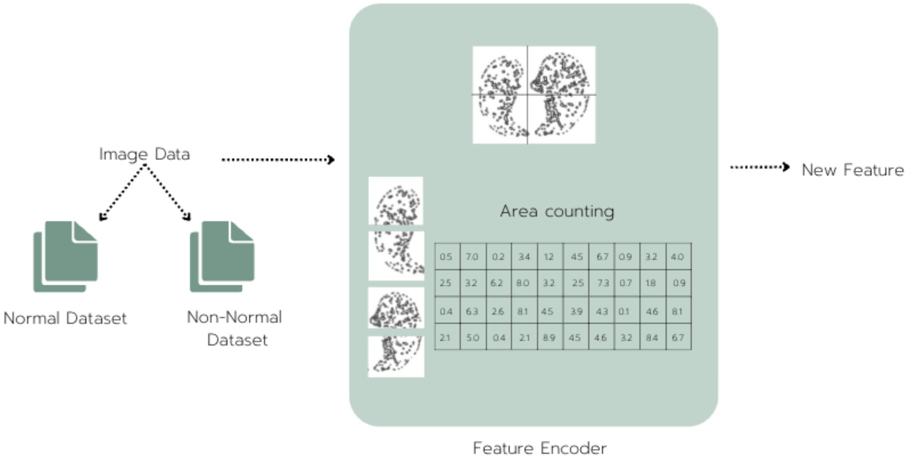


Figure 5.The Example of lung nodules.

Pulmonary radiographs were partitioned into four anatomically distinct quadrants, as illustrated in (Fig5). This segmentation strategy permitted localized analysis of morphological characteristics, enhancing the spatial specificity of feature extraction.



0.5	7.0	0.2	3.4	1.2	4.5	6.7	0.9	3.2	4.0
2.5	3.2	6.2	8.0	3.2	2.5	7.3	0.7	1.8	0.9
0.4	6.3	2.6	8.1	4.5	3.9	4.3	0.1	4.6	8.1
2.1	5.0	0.4	2.1	8.9	4.5	4.6	3.2	8.4	6.7

Figure 6. Shows the contributing zone of the CT scan image.

Following the visualization and statistical analysis of contour area distributions (Figure 7), optimal threshold parameters were established within the range of 0 to 70,000 spatial units. These thresholds were implemented with systematic increments of 5 units across the entire range, establishing discrete categorical boundaries for subsequent feature quantification.

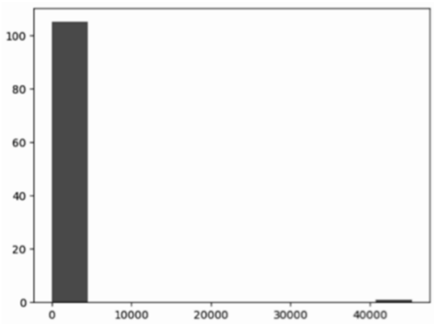


Figure 7. Graphical representation of contour result data distribution across the analyzed pulmonary specimens.

A comprehensive feature mapping framework was developed wherein contour areas were enumerated within each predefined threshold range. This methodical approach facilitated the transformation of complex morphological characteristics into a structured, high-dimensional feature space, as illustrated in (Figure 6). The features extracted from both experimental pipelines—the raw vector dataset (Experiment 1) and the morphologically encoded dataset (Experiment 2)—were utilized in the classification phase to evaluate model performance. Specifically, both feature sets were applied to two categories of machine learning models. Explainable Models, including Logistic Regression, Decision Tree, and K-Nearest Neighbor (KNN), which provide transparent decision-making processes suitable for clinical interpretation. Traditional or Non-Interpretable Models, such as Multi-layer Perceptron (MLP), Support Vector Machine (SVM), and Convolutional Neural Network (CNN), which leverage more complex computational architectures for pattern recognition.

Explainable Machine Learning (XAI): Explainable Artificial Intelligence (XAI) refers to transparent models that have an intrinsic architecture that satisfies at least one of the three transparency dimensions discussed in a previous section.

1.1) Decision tree

Decision trees belong to the family of supervised learning algorithms that organize decision sequences into intuitive tree-like structures (Figure 8). The architecture begins with the root node, positioned at the apex of the tree, which encompasses the entire training dataset and serves as the starting point for analysis [15]. As the algorithm progresses, it evaluates internal nodes where specific attributes are selected based on their ability to effectively separate data into more uniform subgroups according to predetermined criteria. The process then advances to data partitioning, where the algorithm divides information based on specific thresholds or categories of the selected feature. This division creates distinct pathways representing different decision outcomes. The algorithm continues this pattern recursively, generating new decision points and branches as it navigates deeper into the tree hierarchy, until it reaches termination conditions such as maximum depth or minimum improvement thresholds. At the tree's endpoints, leaf nodes represent final classification outcomes where predictions are assigned to new data instances. In classification applications, these terminal nodes contain specific category labels that are determined by the unique sequence of decisions encountered while traversing from the root to the respective leaf position.

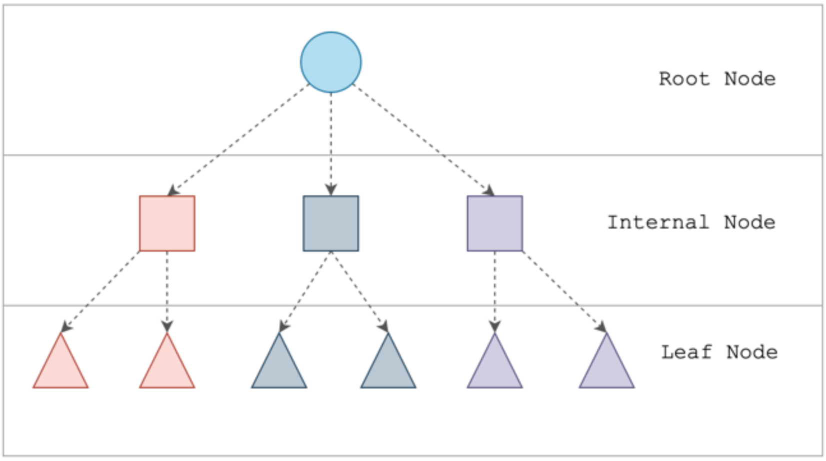


Figure 8. Structure of the Decision tree machine learning model.

Recent research has focused on evaluating various machine learning approaches for early detection of pulmonary malignancies [15]. A comparative analysis examined multiple classification algorithms, including probabilistic methods (Naïve Bayes), boundary-based approaches (Support Vector Machine), connectionist systems (Neural Networks and Multi-Layer Perceptron), hierarchical models (Decision Tree), and gradient-enhanced variants (Gradient-boosted Decision Tree). Additionally, ensemble methods combining multiple algorithms through Random Forest and Majority Voting techniques were assessed, with particular attention to combinations that integrate the capabilities of MLP, GBT, and SVM. Among all methodologies evaluated, the Gradient Boosted Tree demonstrated superior performance characteristics, achieving a classification accuracy of 90% across testing scenarios. This finding highlights the potential value of boosted decision tree architectures in clinical diagnostic applications, particularly for the early identification of lung cancer, where accurate classification directly impacts treatment planning and patient outcomes.

1.2) K-nearest Neighbors

The K-nearest Neighbors (KNN) algorithm functions as a supervised learning classifier that utilizes proximity relationships to determine the classification of data points. While capable of addressing both regression and classification problems, KNN primarily serves as a classification method, operating on the fundamental principle that similar data instances naturally cluster together in feature space [26] (Figure 9). In their research, the investigators implement an enhanced k-Nearest-Neighbors approach supplemented with genetic algorithm optimization for feature selection. This strategic combination effectively reduces dimensionality while simultaneously improving classifier performance. The methodology targets an explicitly accurate diagnosis of disease progression stages in lung cancer patients. To maximize the effectiveness of their proposed system, the research team conducts systematic experimentation to identify the optimal value for the k parameter—the number of neighboring points considered during classification. Through rigorous testing on a comprehensive lung cancer dataset, their approach demonstrates superior accuracy in disease stage determination compared to conventional methods. The genetic algorithm component plays a crucial role in identifying the most diagnostically relevant features while eliminating redundant or noisy variables that could potentially degrade classification performance. This feature selection process not only improves accuracy but also enhances computational efficiency by reducing the dimensionality of the feature space, allowing for more rapid diagnosis without sacrificing precision. The methodology's success in accurately classifying different disease stages highlights its potential clinical value, particularly in settings where timely and accurate staging directly impacts treatment planning and patient management strategies.

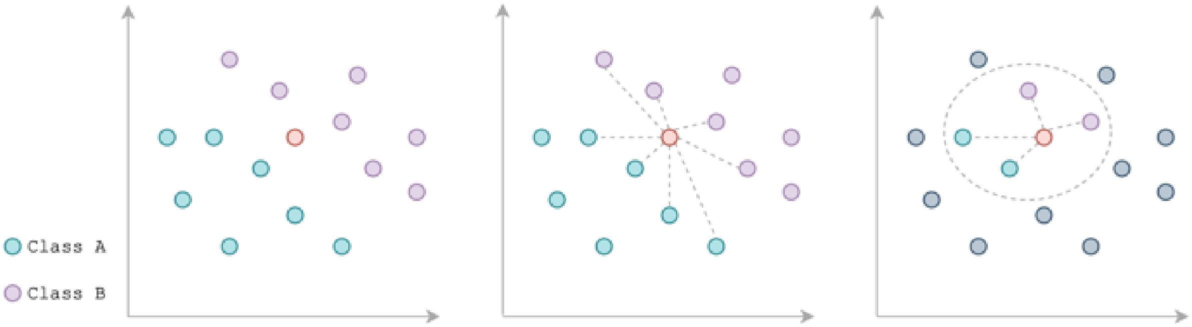


Figure 9. Structure of K-nearest Neighbors (KNN) machine learning model.

1.3) Logistic Regression

Logistic Regression represents a versatile statistical methodology widely applied in classification scenarios and predictive modeling. This approach effectively estimates occurrence probabilities for specific events within datasets containing multiple independent variables. When applied to binary classification problems—such as differentiating between non-COVID and COVID cases—the model generates probability outputs constrained between 0 and 1. A fundamental aspect of this technique involves the logit transformation, which converts odds ratios (the success probability relative to the failure probability) into a more mathematically tractable form. Research by Chen et al. [27] examined various classification methods for lung cancer prediction, with Logistic Regression selected as one of the key comparative algorithms. Their investigation utilized comprehensive datasets from the UCI Machine Learning Repository, incorporating a diverse set of predictive attributes. These features encompassed demographic factors (Age, Gender), environmental exposures (Air Pollution, Dust Allergy, Occupational Hazards), lifestyle variables (Alcohol Use, Balanced Diet, Obesity, Smoking, Passive Smoking), genetic predisposition (Genetic Risk, Chronic Lung Disease), and symptom indicators (Chest Pain, Hemoptysis, Fatigue, Weight Loss, Dyspnea, Wheezing, Dysphagia, Digital Clubbing, Respiratory Infections, Persistent Cough, and Sleep Disturbances). Within their analytical framework, outcome classifications employed a three-tier labeling system: malignant tumors received designation '2', benign tumors were labeled '1', and tumor-free healthy individuals were assigned '0'. Performance evaluation revealed that Logistic Regression achieved 66.7% accuracy in correctly classifying cases across these categories. Notably, while Logistic Regression demonstrated reasonable predictive capability, Support Vector Machine (SVM) algorithms consistently outperformed other classification methodologies within the comparative analysis. This performance differential highlights the complexity of lung cancer prediction and suggests that while Logistic Regression offers interpretability advantages, more sophisticated algorithms may provide enhanced predictive power for complex medical classification tasks.

2) Traditional Machine Learning

Unexplainable machine learning refers to a model that lacks transparency or interpretability. This kind of machine learning model is complex and challenging to understand.

2.1) Multi-layer Perceptron

The Multilayer Perceptron (MLP), as described by Tamarindo et al. in 2018, emerges as a crucial component within the realm of feed-forward neural networks. Comprising three distinct layers – an input layer, a hidden layer, and an output layer – the MLP's architecture is eloquently depicted in (Figure 10). The primary role of the input layer is to receive and process the incoming signals. On the other hand, the output layer assumes responsibility for pivotal functions such as classification and prediction. Notably, MLP's innovative prowess lies in its capacity to accommodate an infinite sequence of hidden layers, discretely interposed between the input and output layers. This network structure, characterized by a unidirectional flow of data from the input to the output layer, closely aligns with the feed-forward network paradigm. The training regimen used to equip the nodes within the MLP with the requisite knowledge is underpinned by the backpropagation learning technique. Through this method, MLPs refine their model parameters, thereby enhancing their capability to generalize and make accurate predictions. Notably, MLPs exhibit a remarkable capacity to address problems that defy linear separability, rendering them exceptionally adept at approximating a wide spectrum of continuous functions.

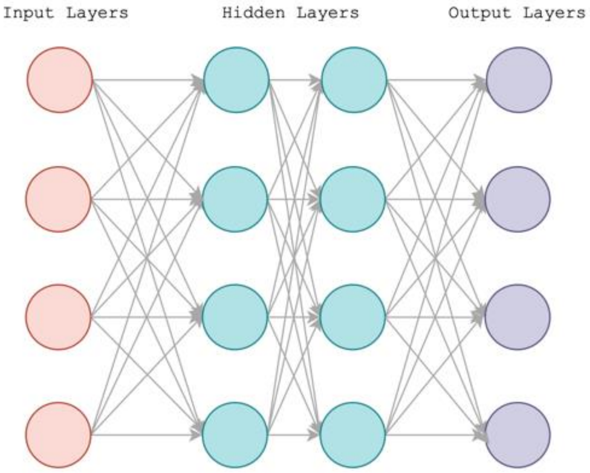


Figure 10. Structure of a Multi-layer perceptron (MLP) machine learning model

This study presented a practical methodology for the detection and classification of lung cancer in CT scan images [5]. The approach involved extracting texture features, including statistical features, which were then fed into various classification algorithms. These algorithms included k-nearest neighbors, support vector machines, decision trees, multinomial naive Bayes, stochastic gradient descent, random forests, and multi-layer perceptrons (MLPs). The analysis outcomes revealed that the MLP classifier achieved the highest accuracy, reaching an impressive rate of 88.55%, surpassing the performance of the other classifiers examined in the study.

2.2) Convolutional Neural Network

Convolutional Neural Networks distinguish themselves from traditional neural architectures through their exceptional ability to process structured grid data, particularly images and audio signals (Figure 11). Their specialized structure comprises three principal layer types. The convolutional layer serves as the network's foundation, utilizing sliding filter matrices to detect local patterns through the generation of feature maps. Pooling layers perform downsampling operations, reducing spatial dimensions while preserving essential information and providing translation invariance. Fully connected layers serve as the final processing stage, integrating the detected features to inform classification decisions. This hierarchical organization creates a progressive feature extraction pipeline. The initial convolutional layers identify fundamental visual elements—

detecting basic attributes such as color boundaries, edges, and simple textures. As information propagates through intermediate layers, the network begins recognizing more complex patterns and structural components of increasing abstraction. The deeper layers gradually assemble these components into recognizable object parts, with each successive layer capturing increasingly sophisticated representations. This architectural progression enables the network to transform raw pixel data into meaningful semantic concepts, ultimately converging at the fully connected layer where comprehensive image understanding occurs and final classification decisions are made. This hierarchical feature extraction mirrors how human visual processing combines elemental visual information into coherent object recognition, but accomplishes this through learned mathematical operations rather than biological processes.

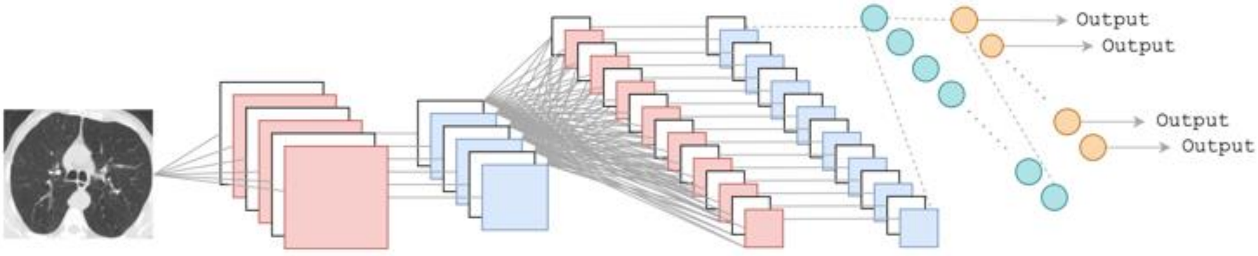


Figure 11. Convolutional Neural Network (CNN) machine learning model.

The research by Cai et al. [28] addresses the critical challenge of developing effective computerized detection models for COVID-19 diagnosis using CT imaging, with particular focus on overcoming the limitations of small dataset sizes. The investigators propose an innovative architectural approach—the stacked autoencoder detector model—which demonstrates significant performance improvements across multiple evaluation metrics. Their methodology employs a two-phase implementation strategy. Initially, they construct four specialized autoencoder modules that form the foundation layers of their detection system. These autoencoders are specifically engineered to extract high-dimensional, information-rich features from thoracic CT images that might otherwise remain undetected through conventional approaches. In the second phase, these individual autoencoder components are integrated through a cascading arrangement, with their outputs connected to a fully-connected dense layer and ultimately to a softmax classification layer. This creates a comprehensive end-to-end detection pipeline capable of distinguishing COVID-19 manifestations from normal pulmonary imaging. Performance evaluation reveals remarkable effectiveness, particularly noteworthy given the limited amount of available training data. The system achieves exceptional results across all primary assessment metrics: 94.7% overall accuracy, 96.54% precision (indicating minimal false positives), 94.1% recall (demonstrating high sensitivity to positive cases), and a balanced F1-score of 94.8%.

2.3) Support vector machine

The Support Vector Machine (SVM) represents a sophisticated classification and regression methodology, distinguished by its ability to optimize decision boundaries while controlling model complexity. This technique particularly excels when processing high-dimensional feature spaces containing thousands of predictor variables, making it exceptionally versatile across diverse application domains. The wide-ranging applicability of SVM extends to numerous fields, including customer analytics, computer vision applications, biological data processing, natural language understanding, network security, protein analysis, and audio recognition systems, demonstrating its remarkable flexibility and effectiveness. At its core, SVM operates by transforming input data into higher-dimensional feature spaces where separation between classes becomes mathematically tractable. The algorithm identifies optimal hyperplanes that maximize the margin between different categories, creating robust decision boundaries that generalize effectively to new data. Through kernel transformations, SVMs can represent complex nonlinear relationships while maintaining computational efficiency, enabling powerful classification capabilities even with challenging datasets. For example, consider the following (Figure 12), in which the data points are categorized into two distinct groups.

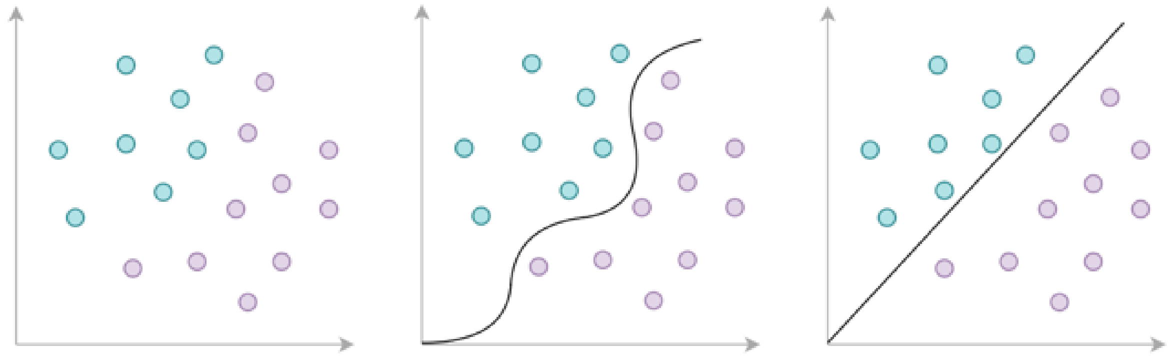


Figure 12. Support vector machine (SVM) machine learning model.

Research by Rahman et al. [14] demonstrates the effectiveness of SVM in medical imaging through an innovative approach for pulmonary nodule analysis in CT scans. Their methodology employs a multi-stage process, starting with the precise segmentation of lungs using active contour models. This initial delineation is followed by specialized masking techniques that transform connected nodules into isolated structures, making them suitable for classification. The critical detection phase employs an SVM classifier leveraging a combination of two-dimensional stochastic features and three-dimensional anatomical characteristics. This hybrid feature approach enables the system to accurately identify subtle nodule patterns while maintaining contextual awareness of surrounding anatomical structures. Following detection, the system applies active contour modeling to extract precise nodule boundaries, facilitating volumetric and morphological analysis.

This classification framework successfully differentiates between nodules connected to pulmonary structures (such as vessel attachments or bronchial wall nodules) and isolated parenchymal nodules. The complete system achieves 89% detection accuracy while maintaining minimal false positive rates, demonstrating SVM's capability to handle complex medical image classification tasks with high precision.

Performance evaluation metrics

Performance evaluation metrics were applied to measure the effectiveness of a machine learning algorithm. Five classification metrics were utilized, including accuracy, F1 score, precision, time complexity, and space complexity. These metrics are crucial for evaluating the accuracy, efficiency, and efficacy of a model or system, and for identifying areas that require improvement. **Accuracy**, which measures the proportion of correct predictions made by the model or algorithm, is one of the most used performance evaluation metrics in machine learning. However, it may not be applicable for multiclass classification problems where the output has more than two categories. The F1 score, which balances precision and recall, is a flexible metric for dealing with imbalanced datasets.

$$\frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{FP}+\mathrm{TN}+\mathrm{FN}}$$

F1 score is a metric commonly used in machine learning and statistics to evaluate the performance of a classification model, particularly when dealing with imbalanced datasets. It is a single numerical value that combines precision and recall, providing a balanced measure of a model's accuracy.

$$\frac{2\mathrm{TP}}{2\mathrm{TP}+\mathrm{FP}+\mathrm{FN}}$$

Precision and recall are two other important performance evaluation metrics that measure the proportion of true positive predictions among all positive predictions and the proportion of true positive predictions among all actual positive instances, respectively. These metrics are particularly Fundamental in applications where false positives or false negatives can have serious consequences, such as medical diagnosis.

$$\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}}$$

Time complexity and space complexity are measures of the computational resources an algorithm requires during execution. Time complexity quantifies how the runtime increases relative to the input size, typically expressed using Big O notation. This mathematical representation establishes an upper bound on growth rates, enabling performance comparisons between different algorithms. Space complexity analyzes memory utilization patterns, assessing both primary algorithm storage needs and auxiliary memory requirements. For data-intensive applications such as medical image processing, space complexity becomes particularly critical when handling large datasets like high-resolution CT scans. When evaluating machine learning models for clinical applications, these metrics take on special significance. While complex models may achieve superior classification accuracy, their computational demands might render them impractical for real-time diagnostic settings. Algorithms demonstrating favorable time and space complexity characteristics offer advantages in resource-constrained environments where processing efficiency directly impacts clinical workflow. The balance between algorithmic complexity and computational efficiency represents a key consideration when developing diagnostic systems, particularly those intended for deployment in settings with limited computational resources or where rapid results are essential for patient care.

3. Results and Discussion

3.1 Experimental Setup

In this study, we utilized the Lung Image Database Consortium Image Collection (LIDC-IDRI), as reported by Armato, III. et al. [24], which comprises 244,527 thoracic computed tomography (CT) scans for diagnostic and lung cancer screening purposes from 1,010 patients. We also incorporated normal CT scans from a publicly accessible COVID-19 (SARS-CoV-2) lung CT dataset, which includes 1,229 negative cases, encompassing both normal and non-COVID-19 images. These datasets provide images in a 512x512px PNG format, sourced from actual patients at the radiology departments of academic hospitals in Tehran, as documented by Aria, M. et al. [25]. The final dataset includes 1,224 normal lung CT scans and 1,010 lung cancer cases.

3.2 Experiment Result

Table 1. Performance results of traditional machine learning models with normal and non-normal lung image datasets.

Model	Accuracy	Precision	Recall	F1 Score	Time Complexity (min)	Time Complexity (MB)
MLP	0.55	0.27	0.50	0.35	206.63	4811.45
CNN	0.96	0.97	0.96	0.96	3.40	5628.08
SVM	0.95	0.96	0.95	0.95	108.20	9767.71

In this section, we present two distinct solutions. The first solution pertains to the image classification results, where we report the outcomes of our efforts in distinguishing individuals as having either normal lung conditions or lung cancer, based on image analysis in (Table 1). The second solution focuses on the results of our feature engineering endeavors, which also aim to identify individuals in the same manner, differentiating between normal and cancerous cases, specifically those related to the lungs. We implemented on Google Colaboratory notebooks, which allocated an A100 GPU. In this experiment, various performance metrics were employed to assess the effectiveness of the machine learning model. A total of two experimental runs were conducted, each consisting of three iterations, with the results subsequently averaged for inclusion in this report. Initially, the model was trained and tested using both normal and non-normal lung image datasets. The results have been presented in (Table 2). The experiment yielded compelling results. The Convolutional Neural Network (CNN) classification model emerged as the standout performer, achieving an impressive accuracy rate of 0.95. This results in effectiveness in accurately classifying lung images as either normal or those with lung cancer. On the other hand, when it comes to computational efficiency, in terms of both runtime complexity and space complexity, the Logistic Regression classification model outperforms the others. This model demonstrated an advantageous balance between accuracy and resource utilization, making it a strong candidate for applications where efficiency is a critical consideration. In essence, our findings emphasize the importance of selecting the right tool for the task at hand. While the CNN excelled in accuracy, the Logistic Regression model showcased superior efficiency, highlighting the need to strike a balance between precision and computational resources in practical applications.

Table 2. Performance comparison of different machine learning models with normal and non-normal lung image datasets.

Model	Accuracy	Precision	Recall	F1 Score	Time Complexity (min)	Space Complexity (MB)
Logistic Regression	0.91	0.92	0.90	0.91	0.71	3859.11
Decision Tree	0.94	0.95	0.94	0.94	22.6	5268.47
K-Nearest Neighbor	0.94	0.95	0.94	0.94	160.57	7793.39
MLP	0.55	0.27	0.50	0.35	206.63	4811.45
CNN	0.96	0.97	0.96	0.96	3.40	5628.08
SVM	0.95	0.96	0.95	0.95	108.20	9767.71

In the second experimental phase, we utilized a morphological feature engineering dataset for training and testing, employing the same machine learning model as in the first phase. The outcomes of these experiments are summarized in (Table 3). Notably, the highest accuracy rate achieved was 0.88, and this was attained by both the Multilayer Perceptron and Support Vector Machine classification models. These results underscore the effectiveness of these models in handling morphological feature engineering data for lung cancer classification. Similar to the first experimental phase, when considering time complexity and space complexity, the Logistic Regression classification model demonstrated superior efficiency. This finding reaffirms the Logistic Regression model as a compelling choice for applications where computational resources need to be optimized while maintaining high classification accuracy.

Table 3. Performance comparison of different machine learning models with the Feature Engineering Encoder dataset.

Model	Accuracy	Precision	Recall	F1 Score	Time Complexity (min)	Space Complexity (MB)
Logistic Regression	0.86	0.87	0.87	0.86	0.04	176.80
Decision Tree	0.87	0.88	0.88	0.87	0.48	241.45
K-Nearest Neighbor	0.87	0.88	0.88	0.87	3.21	494.99
MLP	0.88	0.89	0.89	0.88	21.99	220.83
CNN	0.82	0.83	0.83	0.82	0.16	257.79
SVM	0.88	0.89	0.87	0.87	7.15	447.40

Our analysis revealed two distinct dataset types. Models trained on standardized non-normal and normal images achieved higher accuracy compared to those trained on feature-encoded datasets. However, while differences in accuracy were apparent, precision and F1-score values showed only marginal variation.

Significantly, the use of standardized images substantially reduced computational demands, with both time and memory complexity significantly lowered. This finding highlights the potential for considerable efficiency gains in lung cancer classification tasks. When comparing explainable machine learning models with traditional non-interpretable approaches, we observed broadly comparable performance. The choice of algorithm—whether Decision Tree, Logistic Regression, K-Nearest Neighbor, CNN, or SVM—did not substantially alter the overall diagnostic effectiveness when applied to the respective datasets. To assess whether performance differences were statistically significant, paired t-tests were conducted across cross-validation folds. CNN achieved significantly higher accuracy compared to Logistic Regression and Decision Tree ($p < 0.05$). In contrast, no significant difference was observed between CNN and SVM ($p = 0.18$). Explainable models (Decision Tree, Logistic Regression, and KNN) demonstrated statistically comparable performance to non-interpretable models on feature-engineered datasets, as evidenced by overlapping 95% confidence intervals. These results suggest that small variations in performance do not undermine the interpretability advantage of an explainable app. Overall, our findings underscore the importance of considering both accuracy and resource efficiency in medical image classification. While certain datasets may achieve marginally higher accuracy, the performance gap is often not statistically significant. Prioritizing computationally efficient and interpretable methods is therefore essential for real-world clinical deployment, where resource constraints and the need for transparent decision-making are critical factors.

3.3 Discussion

Unlike previous studies that primarily prioritized accuracy, our approach introduces a novel morphological feature engineering pipeline that reduces computational burden while encoding radiologically interpretable features. This dual benefit bridges the gap between AI accuracy and clinical usability. Our comparative analysis highlights the trade-offs between accuracy, efficiency, and interpretability. While CNN achieved the highest accuracy (0.96), explainable models such as Decision Trees and K-Nearest Neighbors produced comparable results (0.94), with only a marginal difference of ~2%. Significantly, morphological feature engineering reduced processing time and memory usage by up to 95% while maintaining clinically acceptable accuracy (up to 0.88). By framing predictions in terms of radiological features such as nodule size, opacity, and distribution, explainable models not only preserved diagnostic reliability but also enhanced transparency and clinical trust. Overall, these findings suggest that explainable models, combined with feature engineering, offer a practical and scalable solution for real-world deployment, supporting radiologists in their diagnostic workflows rather than replacing them.

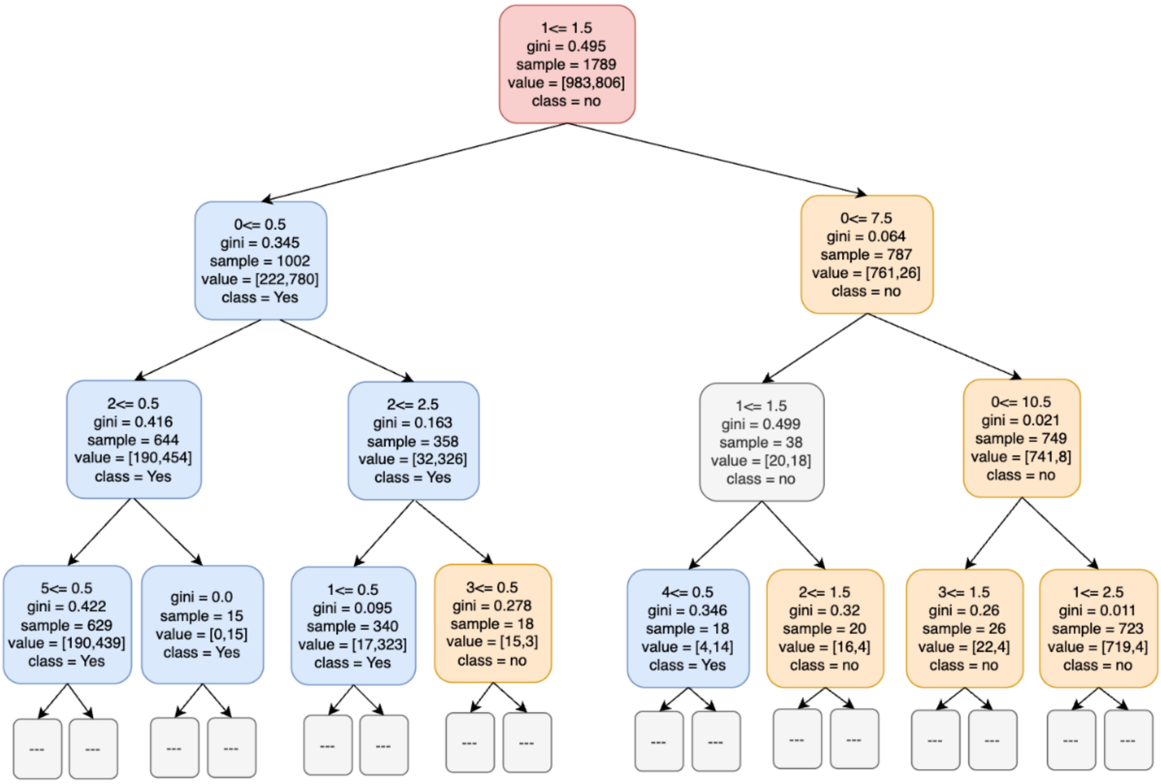


Figure 13. Decision Tree classifier for lung cancer detection.

The tree structure illustrates how radiological features (e.g., nodule size, opacity) are used in decision rules, with internal nodes representing feature thresholds and leaf nodes indicating class predictions (“cancer” or “normal”) shown in (Figure 13). The gini impurity metric quantifies class purity at each node, where lower values denote stronger separation. For example, one path shows that if feature_1 ≤ 0.5 and feature_3 > 1.0, the model predicts “lung cancer” with high confidence. This visualization clarifies the reasoning process, enabling clinicians to trace decision paths and verify which features most influenced each prediction. Such transparency is critical for integrating AI into clinical workflows, where interpretability and accountability are essential.

4. Conclusions

This research introduced a novel application of explainable machine learning methodologies for feature encoding from lung cancer CT imaging data. By systematically comparing explainable algorithms with black-box approaches, our framework demonstrated reliable effectiveness in identifying malignant pulmonary nodules. Our results show that explainable models can achieve diagnostic performance comparable to conventional methods, while offering the added advantage of interpretability. A key contribution of this study lies in the morphological feature engineering pipeline, which substantially reduced computational demands in terms of both processing time and memory usage. This efficiency gain represents a meaningful step toward practical deployment in clinical environments. The broader implications extend to real-world healthcare, where rapid, accurate, and interpretable decision support is essential. By aligning with radiologists’ diagnostic reasoning, explainable models can enhance trust and usability in clinical workflows, thereby supporting early detection of lung malignancies and contributing to improved patient outcomes. Looking ahead, this methodology holds promise for broader medical imaging tasks, including oncology, dermatology, neurology, and other radiological applications. Future research should refine feature extraction techniques, validate performance in diverse patient populations, and explore hybrid approaches that combine feature engineering with deep learning. Explainable machine learning models integrated with morphological feature engineering offer a compelling alternative to black-box approaches. Although CNNs deliver slightly higher accuracy, explainable models provide transparency, accountability, and clinical trust—attributes essential for the responsible integration of AI in healthcare. By balancing accuracy, efficiency, and interpretability, explainable AI can serve as a practical decision-support tool, ultimately laying the foundation for trustworthy and sustainable clinical adoption.

5. Acknowledgements

The authors would like to thank the Department of Computer Science at Chiang Mai University for providing computational resources and the radiology departments of the teaching hospitals in Tehran for their assistance with the CT scan datasets.

Author Contributions: Conceptualization, K.T. and P.S.; methodology, K.T.; software, K.T.; validation, K.T. and P.S.; formal analysis, K.T.; investigation, K.T.; resources, P.S.; data curation, K.T.; writing—original draft preparation, K.T.; writing—review and editing, P.S.; visualization, K.T.; supervision, P.S.; project administration, P.S. All authors have read and agreed to the published version of the manuscript.

References

[1] Takenaka, M.; Hanagiri, T.; Shinohara, S.; Kuwata, T.; Chikaishi, Y.; Oka, S.; Tanaka, F. The prognostic significance of HER2 overexpression in non-small cell lung cancer. *Abbreviated Journal Name* **2011**.

[2] Oser, M. G.; Niederst, M. J.; Sequist, L. V.; Engelman, J. A. Molecular drivers and cells of origin in the transformation from non-small-cell lung cancer to small-cell lung cancer. *Lancet Oncol.* **2015**, *16*(4), E165-E172. https://doi.org/10.1016/S1470-2045(14)71180-5

[3] de Castro, J.; Rodríguez, M. C.; Martínez-Zorzano, V. S.; Sánchez-Rodríguez, P.; Sánchez-Yagüe, J. Erythrocyte fatty acids as potential biomarkers in the diagnosis of advanced lung adenocarcinoma, lung squamous cell carcinoma, and small cell lung cancer. *Abbreviated Journal Name* **2014**, *142*(1), 111-120. https://doi.org/10.1309/AJCP1QUQQLLT8BLI

[4] Yu, L.; Tao, G.; Zhu, L.; Wang, G.; Li, Z.; Ye, J.; Chen, Q. Prediction of pathologic stage in non-small cell lung cancer using machine learning algorithm based on CT image feature analysis. *Abbreviated Journal Name* **2019**. https://doi.org/10.1186/s12885-019-5646-9

[5] Singh, G. A. P.; Gupta, P. K. Performance analysis of various machine learning-based approaches for detection and classification of lung cancer in humans. *Abbreviated Journal Name* **2019**.

[6] Kadir, T.; Gleeson, F. Lung cancer prediction using machine learning and advanced imaging techniques. *Abbreviated Journal Name* **2018**, *7*(3). https://doi.org/10.21037/tlcr.2018.05.15

[7] Elish, M. C. The stakes of uncertainty: developing and integrating machine learning in clinical care. *Abbreviated Journal Name* **2018**. https://doi.org/10.1111/1559-8918.2018.01213

[8] Tjoa, E.; Guan, C. A survey on explainable artificial intelligence (XAI): Toward medical XAI. *IEEE Transactions on Neural Networks and Learning Systems* **2020**, *32*(11). https://doi.org/10.1109/TNNLS.2020.3027314

[9] Doshi-Velez, F.; Kim, B. Towards a rigorous science of interpretable machine learning. *Abbreviated Journal Name* **2017**.

[10] Balagurunathan, Y.; Kumar, V.; Gu, Y.; Kim, J.; Wang, H.; Liu, Y.; Goldgof, D. B.; Hall, L. O.; Korn, R.; Zhao, B.; Schwartz, L. H.; Basu, S.; Eschrich, S.; Gatenby, R. A.; Gillies, R. J. Test–Retest Reproducibility Analysis of Lung CT Image Features. *Med. Phys.* **2014**, *41*(5), 2405–2427.

[11] Paing, M. P.; Choomchuay, S. Ground glass opacity (GGO) nodules detection from lung CT scans. *Abbreviated Journal Name* **2017**. https://doi.org/10.1109/ISESD.2017.8253338

[12] Ramalho, G. L. B.; Rebouças Filho, P. P.; Medeiros, F. N. S. D.; Cortez, P. C. Lung disease detection using feature extraction and extreme learning machines. *Abbreviated Journal Name* **2014**. https://doi.org/10.1590/rbeb.2014.019

[13] Abdillah, B.; Bustamam, A.; Sarwinda, D. Image processing based detection of lung cancer on CT scan images. *Abbreviated Journal Name* **2017**. https://doi.org/10.1088/1742-6596/893/1/012063

[14] Keshani, M.; Azimifar, Z.; Tajeripour, F.; Boostani, R. Lung nodule segmentation and recognition using SVM classifier and active contour modeling: A complete intelligent system. *Abbreviated Journal Name* **2013**, *43*(4), 287-300. https://doi.org/10.1016/j.compbiomed.2012.12.004

[15] Faisal, M. I.; Bashir, S.; Khan, Z. S.; Khan, F. H. An evaluation of machine learning classifiers and ensembles for early stage prediction of lung cancer. *Abbreviated Journal Name* **2018**. https://doi.org/10.1109/ICEEST.2018.8643311

[16] Günaydin, Ö.; Günay, M.; Şengel, Ö. Comparison of Lung Cancer Detection Algorithms. In *Proceedings of the 2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT)*, April 2019; IEEE, **2019**; pp 1–4. https://doi.org/10.1109/EBBT.2019.8741826

[17] Tonekaboni, S.; Joshi, S.; McCradden, M. D.; Goldenberg, A. What clinicians want: contextualizing explainable machine learning for clinical end use. *Abbreviated Journal Name* **2019**.

[18] Kumar, D.; Wong, A.; Clausi, D. A. Lung nodule classification using deep features in CT images. *Abbreviated Journal Name* **2015**. https://doi.org/10.1109/CRV.2015.25

[19] Thallam, C.; Peruboyina, A.; Raju, S. S. T.; Sampath, N. Early stage lung cancer prediction using various machine learning techniques. In *2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*; IEEE: City, Country, **2020**; pp 1285–1292. https://doi.org/10.1109/ICECA49313.2020.9297576

[20] Li, X.; Shen, L.; Xie, X.; Huang, S.; Xie, Z.; Hong, X.; Yu, J. Multi-resolution convolutional networks for chest X-ray radiograph-based lung nodule detection. *Artificial Intelligence in Medicine* **2020**, *103*, 101744. https://doi.org/10.1016/j.artmed.2019.101744

[21] Daisy, T. White Box vs. Black Box Algorithms in Machine Learning. ActiveState, 2023. https://www.activestate.com/ (accessed 2023-07-09).

[22] Kareem, H. F.; AL-Husieny, M. S.; Mohsen, F. Y.; Khalil, E. A.; Hassan, Z. S. Evaluation of SVM Performance in the Detection of Lung Cancer in Marked CT Scan Dataset. *Indonesian J. Electr. Eng. Comput. Sci.* **2021**, *21*(3), 1731. https://doi.org/10.11591/ijeecs.v21.i3.pp1731-1738

[23] Guidotti, R.; Monreale, A.; Ruggieri, S.; Turini, F.; Giannotti, F.; Pedreschi, D. A Survey of Methods for Explaining Black Box Models. *ACM Comput. Surv.* **2018**, *51*(5), 1–42. https://doi.org/10.1145/3236009

[24] Armato III, S. G.; McLennan, G.; Bidaut, L.; McNitt-Gray, M. F.; Meyer, C. R.; Reeves, A. P.; Clarke, L. P. The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans. *Abbreviated Journal Name* **2011**.

[25] Aria, M.; Ghaderzadeh, M.; Asadi, F.; Jafari, R. Lung CT Scan Dataset from Tehran Teaching Hospitals. Mendeley Data, 2021. https://doi.org/10.17632/hn6vr7r5cm.1 (accessed 2023-07-09).

[26] Maleki, N.; Zeinali, Y.; Niaki, S. T. A. A k-NN method for lung cancer prognosis with the use of a genetic algorithm for feature selection. *Abbreviated Journal Name* **2021**, *164*, 113981. https://doi.org/10.1016/j.eswa.2020.113981

[27] Radhika, P. R.; Nair, R. A.; Veena, G. A comparative study of lung cancer detection using machine learning algorithms. *Abbreviated Journal Name* **2019**.

[28] Li, D.; Fu, Z.; Xu, J. Stacked-autoencoder-based model for COVID-19 diagnosis on CT images. *Abbreviated Journal Name* **2021**, *51*, 2805-2817. https://doi.org/10.1007/s10489-020-02002-w

ASEAN
Journal of Scientific and Technological Reports
Online ISSN:2773-8752

Research article

Forecasting Rainfall in the Ping, Wang, Yom, and Nan River Basins of Thailand using Decomposition and Holt-Winters Methods Enhanced by GRG Nonlinear Optimization

Pradthana Minsan[1*], and Watha Minsan[2]

[1] Faculty of Science and Technology, Chiang Mai Rajabhat University, Chiang Mai, 50300, Thailand
[2] Faculty of Science, Chiang Mai University, Chiang Mai, 50200, Thailand
* Correspondence: pradthana_min@g.cmru.ac.th

Citation:
Minsan, P.; Minsan, W.; Forecasting rainfall in the Ping, Wang, Yom, and Nan River Basins of Thailand using decomposition and holt-winters methods enhanced by GRG nonlinear optimization. *ASEAN J. Sci. Tech. Report.* **2025**, *28*(5), e259297. https://doi.org/10.55164/ajstr.v28i5.259297.

Article history:
Received: May 13, 2025
Revised: August 20, 2025
Accepted: September 1, 2025
Available online: September 14, 2025

Abstract: The Ping, Wang, Yom, and Nan rivers in Northern Thailand are vital for national sustainability, serving as key sources of energy and public utilities. This study utilizes secondary data from monitoring stations operated by the Upper Northern Region Irrigation Hydrology Center under the Royal Irrigation Department (RID), Thailand. The primary objective is to develop optimal parameter estimation methods for rainfall forecasting in the upper Ping, Wang, Yom, and Nan river basins. Time series forecasting techniques examined include decomposition with the Whale Optimization Algorithm (WOA-D), Holt-Winters smoothing enhanced by WOA (WOA-HW), and Generalized Reduced Gradient (GRG) nonlinear optimization methods (GRG-D and GRG-HW). These are compared with traditional decomposition and Holt-Winters models developed using Minitab (Minitab-D) and Excel (ForecastSheet-HW). The secondary objective is to forecast rainfall over a 24-month horizon to support trend analysis and improve water resource management. Seasonal ARIMA (SARIMA) is also employed for comparative analysis. Results show that traditional tools, such as Minitab-D and ForecastSheet-HW, are accessible and effective in minimizing forecasting errors. Metaheuristic models such as WOA offer improved accuracy but require programming expertise. GRG solvers provide a practical balance, offering near-comparable accuracy without the need for coding. GRG-D and GRG-HW produced forecasts closely matching actual rainfall across all basins. GRG-HW and WOA-HW achieved the lowest sMAPE values for the Yom and Nan Rivers, GRG-D outperformed Minitab-D for the Wang River, and ForecastSheet-HW remained most effective for the Ping River.

Keywords: Generalized reduced gradient; whale optimization algorithm; rorecasting; time series

1. Introduction

Accurate rainfall forecasting in Thailand's major river basins is crucial for effective water resource management, informed agricultural planning, flood prevention, and the development of infrastructure related to water management. Thailand comprises several major river basins, among which the Ping, Wang, Yom, and Nan rivers form significant tributaries contributing to the Chao Phraya River, the country's primary waterway. Analyzing rainfall patterns in these basins is therefore essential for developing precise forecasting models.

Over the past decade, Thailand has faced significant challenges due to climate change, which has considerably affected rainfall patterns [1]. Notable examples include the severe flooding of 2011, which caused extensive economic damage, and the recent unprecedented floods in northern Thailand in 2024, primarily driven by the La Niña phenomenon and intensified by typhoons Yagi and Soulik. These conditions caused heavy and persistent rainfall, leading to substantial water accumulation, flash floods, and river overflow, severely impacting over 43,000 households across 19 provinces, resulting in at least 45 fatalities. Chiang Mai, situated on the Ping River, experienced two significant flood events within a year, underscoring the severe limitations of existing flood-warning systems. These events underscore the critical need for improved water management and enhanced early warning systems to mitigate future disasters. Conversely, periods of severe drought have also impacted agricultural productivity and water availability, as evidenced during the 2015–2016 period. Consequently, the increased variability in rainfall makes accurate forecasting particularly vital for disaster preparedness and sustainable water resource management.

The application of statistical and data science methodologies for rainfall forecasting has garnered significant attention globally [2]. Mishra and Desai applied the Box-Jenkins method to forecast droughts using the standardized precipitation index (SPI) series in India's Kansabati river basin, achieving forecasts consistent with observed data 1–2 months ahead, with accuracy decreasing for longer lead times. Wang et al. employed an integrated Ensemble Empirical Mode Decomposition-Autoregressive Integrated Moving Average (EEMD-ARIMA) approach to forecast annual runoff data for Biuliuhe, Dahuofang, and Mopanshan reservoirs in China, demonstrating effective forecasting capabilities [3]. Within the context of Thailand, Jiroge et al. developed forecasting models for inflow to reservoirs managed by the Electricity Generating Authority of Thailand (EGAT), employing decomposition, Holt-Winters, Box-Jenkins methods, and combined forecasting methods, utilizing Minitab and Excel Office 365 for data analysis [4].

Recent years have seen growing interest in combining metaheuristic algorithms with traditional time series forecasting methods. An early contribution in this direction was made by Jiang et al., who improved the Holt–Winters smoothing method by incorporating the Fruit Fly Optimization Algorithm (FOA) to forecast monthly electricity consumption [5]. Inspired by this approach, more recent work in 2024 has focused on applying similar techniques to hydrological forecasting. For instance, Minsan and Minsan integrated the Whale Optimization Algorithm (WOA) with both decomposition and Holt–Winters methods to predict reservoir inflows in southern Thailand [6]. They also applied Cuckoo Search (CS) in combination with these methods for forecasting monthly inflows into reservoirs in eastern Thailand [7]. Beyond hydrology, these integrated metaheuristic approaches have proven effective in other domains, including PM2.5 air quality forecasting and government revenue prediction [8–10].

Therefore, this research is structured around two primary objectives. First, to develop optimal parameter estimation models using Root Mean Square Error (RMSE) as a performance metric, employing time series analysis methods to forecast rainfall in the upper Ping, Wang, Yom, and Nan river basins. The approaches include decomposition combined with WOA (WOA-D), Holt-Winters smoothing enhanced by WOA (WOA-HW), and methods integrating Generalized Reduced Gradient (GRG) nonlinear optimization, namely GRG-D and GRG-HW. Additionally, these methods are compared with decomposition models created by Minitab and Holt-Winters models developed using Excel during the training data phase. The second objective is to forecast rainfall for a 24-month horizon, utilizing these models for practical rainfall trend analysis and prediction, to enhance efficient water management in northern Thailand. Seasonal ARIMA (SARIMA) methodology is also employed for comprehensive analysis. Model performance comparisons and selections are made based on the Symmetric Mean Absolute Percentage Error (sMAPE) criterion, serving as the primary performance indicator for selecting the most suitable forecasting model for each of the four river basins.

2. Materials and Methods

2.1 Data Preparation

This research aims to identify suitable forecasting methods for rainfall in the upper northern region of Thailand. Monthly rainfall data were obtained from the Upper Northern Region Irrigation Hydrology Center, Royal Irrigation Department (RID), Thailand, covering four major river basins: Ping, Wang, Yom, and

Nan. The secondary data [11] were recorded monthly in millimeters at monitoring stations managed by RID. Specific data sources for each river basin included:

Ping River: Stations 071511 (Ban Huai Luek, Chiang Dao district) and 07132 (Chiang Dao district), Chiang Mai Province.

Wang River: Stations W.16A (Ban Hai, Chae Hom district) and Kiew Lom Dam (Mueang district), Lampang Province.

Yom River: Stations Y.20 (Ban Huai Sak, Song district) and Y.1C (Ban Nam Khong, Mueang district), Phrae Province.

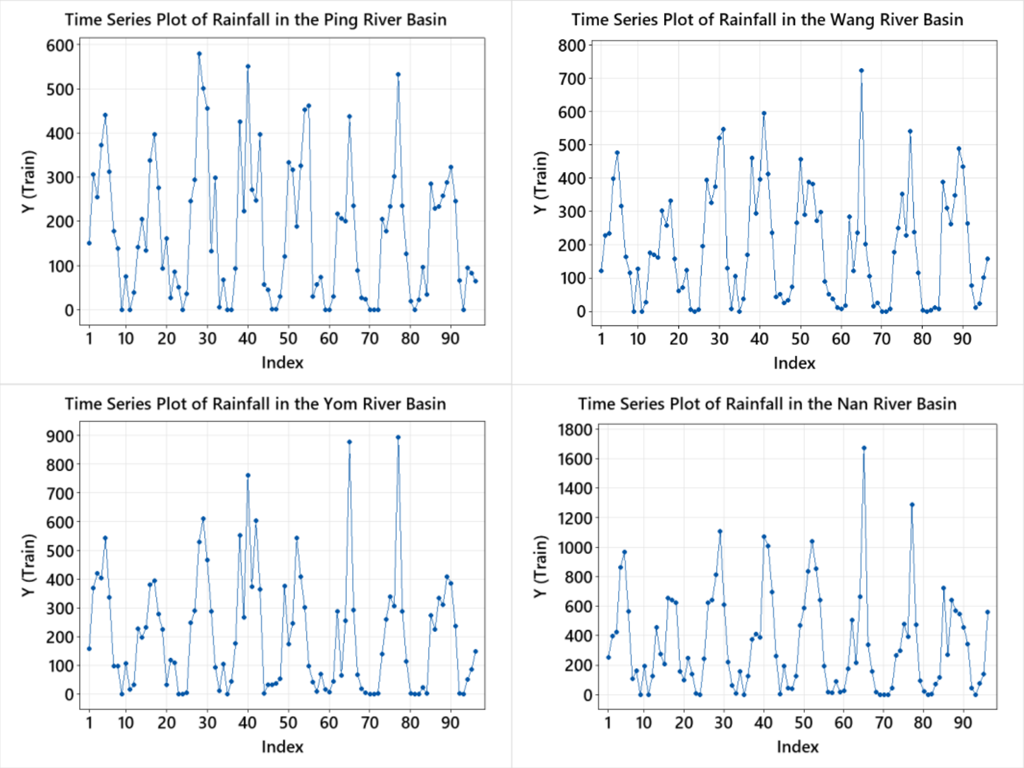
Nan River: Stations 28053 (Thung Chang district), 28073 (Tha Wang Pha district), and Mueang district, Nan Province.

The dataset encompasses 120 months from April 2014 to March 2024. The data were partitioned into two subsets:

Training Dataset (80% of the data, April 2014 to March 2022, totaling 96 months) for developing forecasting models.

Testing Dataset (20% of the data, April 2022 to March 2024, totaling 24 months) for evaluating model performance.

Initial examination of Figure 1 indicated that the time series exhibited both seasonal and trend components. Deseasonalized data analyzed through the Autocorrelation Function (ACF) (Figure 2) showed significant correlations at lag 1, which gradually decreased with increasing lag, indicating a trend component. Detrended data (Figure 3) revealed prominent negative autocorrelations at lags 6 and 18, and positive autocorrelations at lags 12 and 24, suggesting the presence of seasonal fluctuations. Thus, the data contain both seasonal and trend characteristics. Variance stability was assessed using Levene's Test to evaluate homoscedasticity. Prior to the test, the time series data were transformed to reduce autocorrelation, ensuring independence among observations. The results (Table 1) showed p-values greater than the significance level of 0.05 for all river basins, indicating constant variance across all river basins. Consequently, an additive model was deemed appropriate for subsequent analyses and forecasting.


Time Series Plot of Rainfall in the Ping River Basin

Time Series Plot of Rainfall in the Wang River Basin

Time Series Plot of Rainfall in the Yom River Basin

Time Series Plot of Rainfall in the Nan River Basin

Figure 1. Monthly precipitation in the Ping, Wang, Yom, and Nan Basins.

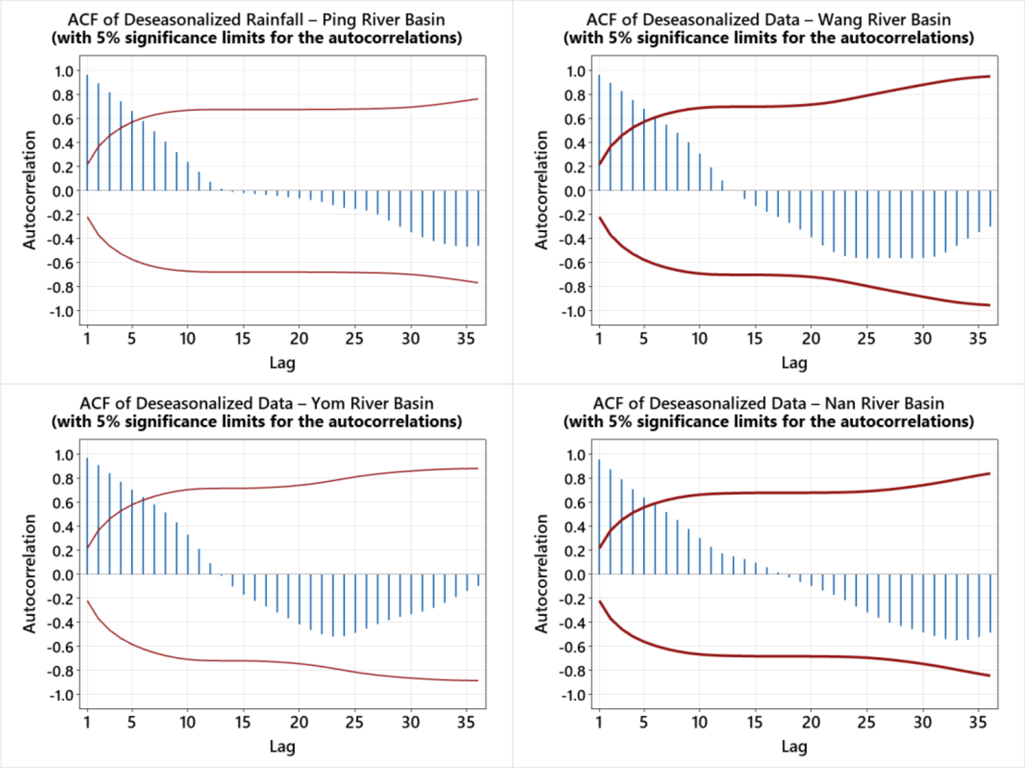


Figure 2. ACF of the deseasonalized time series data of the four river basins.

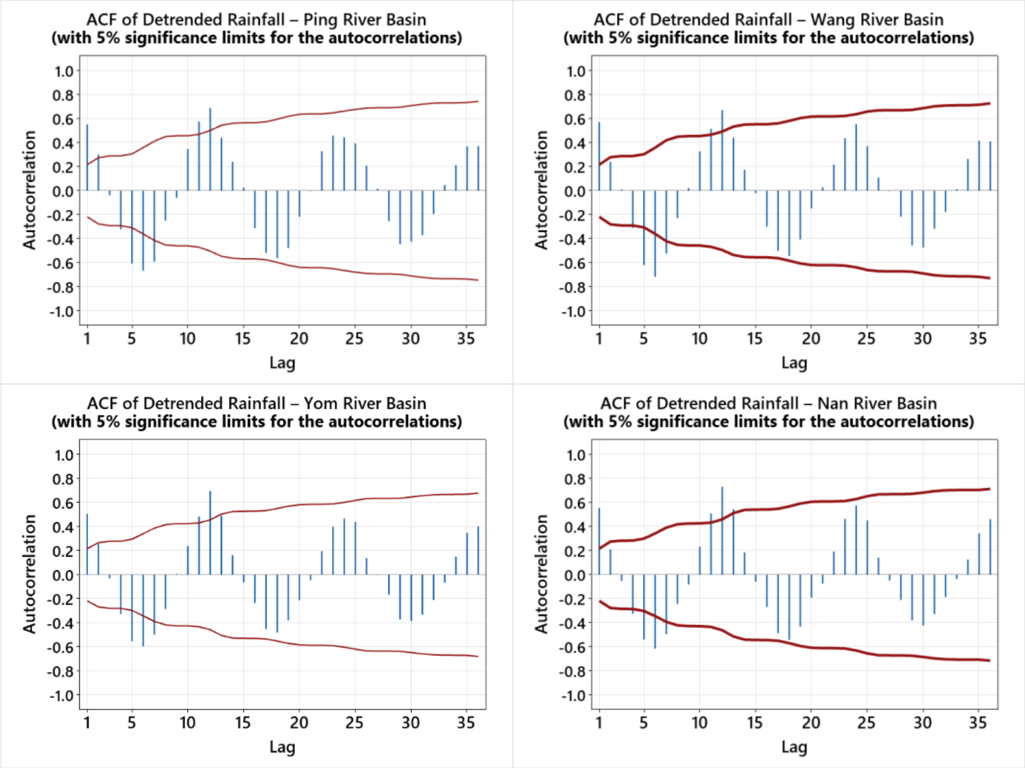


Figure 3. ACF of the detrended time series data of the four river basins.

Table 1. Results of Levene's test for homogeneity of variances across the four river basins.

Levene's Test	**River Basin**			
	Ping	**Wang**	**Yom**	**Nan**
Test Statistic	2.31	0.22	0.10	0.69
P-Value	0.132	0.638	0.749	0.408

2.2 Holt-Winters Method

The Holt-Winters method is widely recognized for its ability to accommodate changes in level, trend, and seasonality within time series data. It can effectively accommodate both additive and multiplicative forms, depending upon the underlying characteristics of the time series data. Notably, the Holt-Winters method remains robust even with limited training data. For the additive model, the forecasting process follows equations (1) through (4) sequentially, as detailed below:

Additive Forecasting:

$$\hat{Y}_{t+p} = \hat{T}_t + p\hat{\beta}_t + \hat{S}_{t-s+1+((p-1)\,\mathrm{mod}\,s)} \text{ for } p = 1, 2, ..., \tag{1}$$

$$\hat{T}_t = \alpha(Y_t - \hat{S}_{t-s}) + (1-\alpha)(\hat{T}_{t-1} + \hat{\beta}_{t-1}), \tag{2}$$

$$\hat{\beta}_t = \gamma(\hat{T}_t - \hat{T}_{t-1}) + (1-\gamma)\hat{\beta}_{t-1}, \tag{3}$$

$$\hat{S}_t = \delta(Y_t - \hat{T}_t) + (1-\delta)\hat{S}_{t-s}, \tag{4}$$

In these equations, $\hat{Y}_{t+p}$ denotes the forecasted data at a future time $t+p$, and p denotes the number of periods ahead to forecast. The parameter s is set to 12 months in this study. Additionally, $\hat{T}_t$ denotes the estimated level at time t, $\hat{\beta}_t$ denotes the trend component, and $\hat{S}_t$ denotes the seasonal component.

2.2.1 Integration of Holt-Winters with Whale Optimization Algorithm (WOA-HW)

Following the approach of Minsan and Minsan [6, 8], the Holt-Winters method requires the estimation of three key smoothing parameters: level (α), trend (γ), and seasonality (δ). Each parameter lies within the range [0, 1] and plays a crucial role in adjusting the model components to best reflect the underlying time series behavior, including long-term trends and seasonal fluctuations. To determine the optimal values for these parameters, this study employs the Whale Optimization Algorithm (WOA), a nature-inspired metaheuristic algorithm. The goal is to minimize forecasting error by tuning the parameters specifically for the additive Holt-Winters model. The optimization process, including a detailed explanation of its implementation, is illustrated in the pseudo-code provided in Figure 4. The performance of the WOA-HW model is evaluated using the RMSE, which serves as the objective function for optimization. The objective is to minimize the RMSE, as defined in Equation (5), calculated over the training dataset.

```
The number of whales: $N = 100$, the number of parameters: $m = 3$, maximum iterations: $T_{\max} = 300$, time limit: $MaxTime = 30$ sec., the fitness value fails to improve after a specified number of iterations $T_{improve} = 50$, the bound of search area: Range Initialize $X_i = (x_i^j, x_i^2, x_i^3), X^*$
While (t < $T_{\max}$ or time < $MaxTime$ or the fitness value fails to improve after a specified $T_{improve}$)
For i = 1 to $N$
    Check if any search agent goes beyond the search space and amend it
    For $j$ = 1 to $m$
        $p = rand[0,1]$
        Update $a, r, A, C, D, D', b, l, X_{rand}$
        If $p \geq 0.5$ then
            Update $x_i^j$ the position of the current search agent #Exploitation Phase
        Elseif $p < 0.5$ and $|A| < 1$
            Update $x_i^j$ the position of the current search agent #Encircling Prey
        Elseif $p < 0.5$ and $|A| \geq 1$
            Update $x_i^j$ the position of the current search agent #Exploration Phase
        Endif
    End for
End for
Calculate $fitness(X_i)$ using WOA-HW by Equation (5)
Update $X^*$ if there is a better solution
$t = t + 1$
End while
Return $X^* = (x^{1*}, x^{2*}, x^{3*})$ # Objective Minimize RMSE($\alpha^*, \gamma^*, \delta^*$) where $\alpha^* = x^{1*}, \gamma^* = x^{2*}, \delta^* = x^{3*}$
```

Figure 4. Pseudo-code of the WOA-HW [6, 12].

$$\text{Objective Minimize RMSE}(\alpha, \gamma, \delta),$$

$$\text{Variable range } \begin{cases} 0 \leq \alpha \leq 1 \\ 0 \leq \gamma \leq 1, \\ 0 \leq \delta \leq 1 \end{cases}$$

$$\text{RMSE} = \frac{1}{n_1} \sum_{t=1}^{n_1} \left(Y_t - \hat{Y}_t \right)^2, \tag{5}$$

where n_1 denotes the number of observations in the training dataset (96 months in this study),

Y_t denotes the observed rainfall, and

$\hat{Y}_t$ denotes the corresponding forecasted value from the WOA-HW model.

2.2.2 Integration of Holt-Winters with GRG Nonlinear Optimization (GRG-HW)

This method builds upon the WOA-HW model by incorporating GRG Nonlinear optimization, available through Microsoft Excel's Solver add-in. Based on the findings of Minsan [13], GRG Nonlinear, when used in Excel 2019, is capable of efficiently identifying optimal parameters for forecasting models. Therefore, in this study, GRG Nonlinear was employed as an alternative optimization technique to WOA, to determine the optimal values of the three smoothing parameters (α, γ, and δ) for the Holt-Winters model.

The GRG-HW model follows the same forecasting equations (1) to (4) as in WOA-HW, with GRG Nonlinear used to replace the WOA metaheuristic for parameter tuning. The optimization setup was configured in Excel Solver with detailed parameter settings to ensure consistency and replicability. These settings are summarized as follows:

Constraint precision: 0.000001
Solving limits: Maximum time = 30 seconds
Use automatic scaling: True
Convergence: 0.0001
Population size: 100
Random seed: 0
Require bounds on variables: True
Derivatives: Forward; and use multistart: True

This configuration enables Solver to efficiently explore the solution space and determine optimal values for the Holt-Winters parameters. The objective function minimized in the GRG-HW model is the same RMSE as defined in Equation (5). The final forecast values, $\hat{Y}_t$, are generated using the parameter set obtained from the GRG-HW model, evaluated on the same training dataset of 96 monthly rainfall observations used in WOA-HW. To facilitate replication, the Excel file containing the GRG-HW Solver setup, parameter constraints, and example data has been provided as supplementary material [14].

2.2.3 Forecasting in Excel using Forecast Sheet (ForecastSheet-HW)

Microsoft Excel 365 includes a built-in tool called Forecast Sheet, which enables users to create time series forecasting models through a guided interface. This tool employs the exponential smoothing method and enables the automatic or manual configuration of the three smoothing parameters (level, trend, and seasonality) used in Holt-Winters forecasting. In this study, the ForecastSheet-HW approach is incorporated as a baseline method and compared with the GRG-HW and WOA-HW models discussed in earlier sections. The comparison aims to evaluate whether the Forecast Sheet can estimate the three smoothing parameters and generate forecasts with comparable accuracy and model adaptability. For consistency, the following settings were manually specified when using the Forecast Sheet in Excel:

Seasonality: Set manually to 12
Include forecast statistics: True

These configurations ensure a controlled environment for evaluating the forecasting performance of ForecastSheet-HW in relation to other optimization-based Holt-Winters implementations.

2.3 Decomposition

This study explores three decomposition-based forecasting approaches. The first approach involves traditional decomposition using Minitab software, referred to in this study as Minitab-D. The other two approaches, described in Sections 2.3.1 and 2.3.2, incorporate optimization algorithms.

2.3.1 Integration of Decomposition with Whale Optimization Algorithm (WOA-D)

Following the approach proposed by Minsan and Minsan [6, 8], the decomposition method is applied to forecast rainfall time series by estimating 14 parameters: $\hat{\beta}_0, \hat{\beta}_1, \hat{S}_1, \hat{S}_2, \hat{S}_3, \hat{S}_4, \hat{S}_5, \hat{S}_6, \hat{S}_7, \hat{S}_8, \hat{S}_9, \hat{S}_{10}, \hat{S}_{11}, \hat{S}_{12}$. These parameters are optimized within the range [0, 1] and are rescaled based on a scaling method proposed in Minsan and Minsan [6, 8]. The effectiveness of the WOA-D model is evaluated by minimizing RMSE, as defined by the objective function below:

$$\text{Objective Minimize RMSE}(\hat{\beta}_0, \hat{\beta}_1, \hat{S}_1, \hat{S}_2, \ldots, \hat{S}_{12})$$

$$\text{Variable range} \begin{cases} \hat{\beta}_0' - 0.2\left|\hat{\beta}_0'\right| \le \hat{\beta}_0 \le \hat{\beta}_0' + 0.2\left|\hat{\beta}_0'\right| \\ \hat{\beta}_1' - 0.2\left|\hat{\beta}_1'\right| \le \hat{\beta}_1 \le \hat{\beta}_1' + 0.2\left|\hat{\beta}_1'\right| \\ \mathrm{LB}_i \le \hat{S}_i \le \mathrm{UB}_i) \text{ for } i = 1, 2, \ldots, 12 \end{cases}$$

$$\mathrm{RMSE} = \frac{1}{n_1}\sum_{t=1}^{n_1}\left(Y_t - \hat{Y}_t\right)^2, \tag{6}$$

where $\hat{Y}_t$ denotes the corresponding forecasted value from the WOA-D model.

The seasonal component bounds for the additive model UB_i LB_i are derived as follows: $\mathrm{UB}_i = \frac{|\Delta Y_t|}{2}$, $\mathrm{LB}_i = -\frac{|\Delta Y_t|}{2}$ where $\Delta Y_t = Y_t - Y_{t-1}$. This range constraint is adapted from [6, 8] to enhance the efficiency of the optimization search space, particularly when implemented using spreadsheet-based solvers such as Excel. The trend component is represented by a simple linear equation and estimated using the ordinary least squares (OLS) method as follows: $\hat{Y}_t' = \hat{\beta}_0' + \hat{\beta}_1' t$. The optimization process, including a detailed explanation of its implementation, is illustrated in the pseudo-code provided in Figure 5.

```
The number of whales: $N = 100$, the number of parameters: $m = 14$, maximum iterations: $T_{\max} = 300$, time limit: $MaxTime = 30$ sec., the fitness value fails to improve after a specified: $T_{improve} = 50$, the bound of search area: Range Initialize $X_i = (x_i^j, x_i^2, ..., x_i^{14}), X^*$
While (t < $T_{\max}$ or time < $MaxTime$ or the fitness value fails to improve after a specified $T_{improve}$)
For i = 1 to $N$
        Check if any search agent goes beyond the search space, and amend it
    For $j = 1$ to $m$
            $p = rand[0,1]$
            Update $a, r, A, C, D, D', b, l, X_{rand}$
            If $p \geq 0.5$ then
               Update $x_i^j$ the position of the current search agent #Exploitation Phase
            Elseif $p < 0.5$ and $|A| < 1$
               Update $x_i^j$ the position of the current search agent #Encircling Prey
            Elseif $p < 0.5$ and $|A| \geq 1$
               Update $x_i^j$ the position of the current search agent #Exploration Phase
            Endif
    End for
End for
Scaling Parameters
Calculate $fitness(X_i)$ using WOA-D by Equation (6)
Update $X^*$ if there is a better solution
$t = t + 1$
End while
Return $X^* = (x^{1*}, x^{2*}, ..., x^{14*})$
# Objective Minimize RMSE($\hat{\beta}_0^*, \hat{\beta}_1^*, \hat{S}_1^*, \hat{S}_2^*, \hat{S}_3^*, \hat{S}_4^*, \hat{S}_5^*, \hat{S}_6^*, \hat{S}_7^*, \hat{S}_8^*, \hat{S}_9^*, \hat{S}_{10}^*, \hat{S}_{11}^*, \hat{S}_{12}^*$) where $\hat{\beta}_0^* = x^{1*}, \hat{\beta}_1^* = x^{2*}, ..., \hat{S}_{12}^* = x^{14*}$
```

Figure 5. Pseudo-code of the WOA-D [6, 12].

2.3.2 Integration of Decomposition with GRG Nonlinear Optimization (GRG-D)

This method builds upon the WOA-D approach described in Section 2.3.1 by substituting the metaheuristic algorithm with the GRG Nonlinear optimization available in Microsoft Excel. The GRG Nonlinear algorithm was applied to determine 14 parameters for the decomposition model, including the trend coefficients ($\hat{\beta}_0, \hat{\beta}_1$) and 12 seasonal components ($\hat{S}_1, \hat{S}_2, \hat{S}_3, \hat{S}_4, \hat{S}_5, \hat{S}_6, \hat{S}_7, \hat{S}_8, \hat{S}_9, \hat{S}_{10}, \hat{S}_{11}, \hat{S}_{12}$). The structure of the objective function and forecasting equations remains consistent with those used in the WOA-D model. Following the procedure of Minsan and Minsan [6, 8], the GRG Solver in Excel was configured with the setting "Use automatic scaling" enabled, as it improves the effectiveness of parameter estimation, particularly when the data scale is small. This adjustment aligns with the scaling approach applied in WOA-D. To facilitate replication, the Excel file containing the GRG-D Solver setup, parameter constraints, and example data has been provided as supplementary material [14]. Users can test the model directly using the default Solver configuration. Model performance is evaluated using RMSE, as defined in Equation (6). Forecast values $\hat{Y}_t$ are calculated using the GRG-D model on the training dataset, with the default Solver tolerance parameters retained, similar to the configuration used in GRG-HW.

2.4 Box-Jenkins Method

The SARIMA model, or Seasonal Autoregressive Integrated Moving Average, extends the classical ARIMA model by incorporating seasonal components. It is commonly expressed as SARIMA(p,d,q)(P, D, Q)$_s$, where s denotes the seasonal period (e.g., 12 for monthly data). This model accounts for both non-seasonal and seasonal autoregressive (AR), moving average (MA), and differencing components. Equation (7) provides the general form of the SARIMA model:

$$\phi_p(B)\Phi_P(B^L)(1-B)^d(1-B^L)^D Y_t = \delta + \theta_q(B)\Theta_Q(B^L)\varepsilon_t\,. \tag{7}$$

Where $\delta = \phi_p(B)\Phi_P(B^L)\mu$, with μ being the mean of the stationary series

B denotes the backward shift operator, such that $B^p y_t = y_{t-p}$

d denotes the degree of non-seasonal differencing

D denotes the degree of seasonal differencing

L denotes the length of the seasonal cycle (e.g., $L=12$ for monthly data)

$\phi_p(B) = (1-\phi_1 B-\phi_2 B^2 - \ldots - \phi_p B^p)$ denotes a non-seasonal autoregressive operator of order p

$\theta_q(B) = (1-\theta_1 B-\theta_2 B^2 - \ldots - \theta_q B^q)$ denotes a non-seasonal moving average operator of order q

$\Phi_P(B^L) = (1-\Phi_1 B^L-\Phi_2 B^{2L} - \ldots - \Phi_P B^{PL})$ denotes a seasonal autoregressive operator of order P

$\Theta_Q(B^L) = (1-\Theta_1 B^L-\Theta_2 B^{2L} - \ldots - \Theta_Q B^{QL})$ denotes the seasonal moving average operator of order Q

The SARIMA model combines both non-seasonal and seasonal components into a unified framework, allowing it to effectively capture both short-term fluctuations and recurring seasonal patterns within time series data.

The Box-Jenkins approach involves a systematic sequence of steps to identify, estimate, validate, and forecast using SARIMA models. The key steps are as follows:

2.4.1 Stationarity Check

The time series is analyzed using the autocorrelation function (ACF) and partial autocorrelation function (PACF) to determine whether the series is stationary. If the series is found to be non-stationary, differencing must be applied to achieve stationarity.

2.4.2 Regular Differencing

If the ACF and PACF suggest a non-stationary trend component, a non-seasonal differencing of order $d=1$ or $d=2$ may be applied to stabilize the mean of the series.

2.4.3 Seasonal Differencing

If the ACF and PACF exhibit a clear seasonal pattern, seasonal differencing must be applied using an appropriate seasonal order D, typically 1 or 2, depending on the periodicity of the seasonal structure.

2.4.4 Data Transformation

If the variance of the series is non-constant, a variance-stabilizing transformation, such as the Box-Cox transformation, may be applied. This transformation helps stabilize the variance across time and supports the use of modeling techniques that assume homoscedasticity, thereby improving the reliability of parameter estimation and residual diagnostics.

2.4.5 Model Identification

Once a stationary series is obtained, the initial model structure is identified by analyzing the ACF and PACF plots to determine the appropriate values for the non-seasonal and seasonal orders (p, d, q) and (P, D, Q), respectively.

2.4.6 Parameter Estimation

Model parameters are estimated using an appropriate method, typically the Ordinary Least Squares (OLS) method or the Maximum Likelihood Estimation (MLE) method.

2.4.7 Model Diagnostics

Model adequacy is assessed through the following diagnostic procedures:

1. Test the significance of each model parameter using the t-test.
2. Examine residual autocorrelation using the Ljung-Box Q-test at appropriate lags (e.g., lag 12) to check for randomness of residuals.
3. Assess residual normality using the Kolmogorov–Smirnov (KS) test.
4. Evaluate mean equality of residuals over time using the t-test to determine temporal stability.
5. Assess the variance equality of residuals using Levene's test to verify homoscedasticity across forecast horizons.
6. Use the Autoregressive Conditional Heteroscedasticity–Lagrange Multiplier (ARCH-LM) test to detect conditional heteroscedasticity, indicating the presence of time-varying volatility in residuals.

If any diagnostic test fails, steps 2.4.5 through 2.4.7 should be revisited to refine the model until all assumptions are reasonably satisfied.

2.4.8 Forecasting

Once a suitable model passes all diagnostics, it is used to generate forecasts for the desired period. The resulting forecasts are expected to provide reliable estimates for future values in the time series.

2.5 Evaluation Criteria

The evaluation in this study comprises two main aspects. The first focuses on in-sample accuracy, which evaluates the parameter estimation process using the training dataset. The objective is to minimize the RMSE, calculated as:

$$\text{RMSE} = \frac{1}{n_1}\sum_{t=1}^{n_1}\left(Y_t - \hat{Y}_t\right)^2$$

Where n_1 is the number of observations in the training dataset (96 months), and Y_t and $\hat{Y}_t$ denote the actual and forecasted rainfall values, respectively.

The RMSE values were computed from six forecasting approaches:

1. Minitab-D: Parameter estimation using Minitab.
2. WOA-D: Parameter estimation using Python and Whale Optimization Algorithm.
3. GRG-D Solver: Parameter estimation in Excel using GRG Nonlinear Solver.
4. ForecastSheet-HW: Holt-Winters parameters estimated using Excel's Forecast Sheet.
5. WOA-HW: Holt-Winters parameters estimated using Whale Optimization Algorithm.
6. GRG-HW: Parameters estimated using the GRG Solver in Excel for the Holt-Winters method.

The RMSE results from methods 1 through 3 (Minitab-D, WOA-D, GRG-D) and 4 through 6 (ForecastSheet-HW, WOA-HW, GRG-HW) are compared within their respective groups to identify the most efficient parameter estimation strategy.

The second evaluation aspect focuses on forecasting accuracy, assessed using the testing dataset. This stage aims to evaluate each model's ability to forecast rainfall 24 months ahead (i.e., two years into the future). The performance metric employed is the Symmetric Mean Absolute Percentage Error (sMAPE). The adjusted MAPE was first introduced by Flores in 1986 [15]. It has been widely adopted in hydrological forecasting studies and has also been implemented in the HydroErr open-source library by Elise et al. [16], which supports both Python and MATLAB environments.

One of the main advantages of sMAPE is that the forecast error is expressed as a percentage bounded between 0% and 100%, making it easier to interpret and compare across different models and time periods. The formula is defined as:

$$\text{sMAPE} = \frac{100}{24}\sum_{t=97}^{n}\frac{\left|Y_t - \hat{Y}_t\right|}{\left|Y_t\right| + \left|\hat{Y}_t\right|}$$

Where Y_t and $\hat{Y}_t$ denote the actual and forecasted rainfall values for the testing period. Here, $n = 120$ refers to the total number of months in the full dataset, with testing data covering months 97 through 120. Although $n = 120$ the sMAPE is computed using the last 24 months (i.e., $t = 97$ to 120)

Seven forecasting approaches were compared based on their sMAPE values: Minitab-D, WOA-D, GRG-D, ForecastSheet-HW, WOA-HW, GRG-HW, and Box-Jenkins. The approach with the lowest sMAPE is considered the most effective in terms of long-term forecasting accuracy.

3. Results and Discussion

3.1 Analysis of Training Dataset Performance

This study compares the performance of parameter estimation for decomposition-based forecasting using three methods: Minitab-D, WOA-D, and GRG-D. Additionally, it evaluates the performance of parameter estimation for Holt-Winters models using three methods: ForecastSheet-HW, WOA-HW, and GRG-HW. All models were applied to monthly rainfall data from the upper northern region of Thailand, covering the Ping, Wang, Yom, and Nan river basins. The rainfall data were obtained from hydrological monitoring stations under the Upper Northern Region Irrigation Hydrology Center, Royal Irrigation Department, and span 10 years. The RMSE values used to evaluate the forecasting performance on the training dataset are presented in Table 2.

Table 2. RMSE of the training dataset for each river basin.

Model		River basin			
		Ping	Wang	Yom	Nan
Decomposition	Minitab-D	84.2	89.5	110.0	165.1
	WOA-D	81.1	87.7*	104.1*	162.9*
	GRG-D	80.9*	87.8	104.7	163.9
Holt-Winters	ForecastSheet-HW	121.6	125.3	138.2	218.7
	WOA-HW	91.2*	102.2*	124.8*	187.5*
	GRG-HW	91.2*	102.2*	124.8*	187.5*

Note: The lowest RMSE value in group decomposition, and group Holt-Winters is *

For the WOA-HW and WOA-D methods, the number of iterations was set to 300, with execution performed using Python via Google Colab. Each iteration required approximately 20 seconds to converge. GRG-D and GRG-HW models were implemented in Excel using the GRG Nonlinear Solver. These optimization approaches aimed to enhance forecast accuracy by effectively tuning parameters. Compared to manual parameter estimation using Minitab-D and Excel's ForecastSheet-HW, the optimization-based approaches generally yielded better performance. The results presented in Table 2 show the following key findings. WOA-D and GRG-D both produced lower RMSE values than Minitab-D across all river basins. Among them, WOA-D achieved the lowest RMSE values in three out of four basins (Wang, Yom, and Nan), while GRG-D slightly outperformed WOA-D only in the Ping River basin. For Holt-Winters models, both WOA-HW and GRG-HW yielded the same RMSE values and clearly outperformed ForecastSheet-HW across all river basins. The best-performing models in each group (decomposition and Holt-Winters) are marked with an asterisk (*) in Table 2.

3.2 Analysis of Forecasting Performance on the Testing Dataset

3.2.1 Box-Jenkins

In the testing phase, the Box-Jenkins method was employed to construct seasonal ARIMA (SARIMA) models, following the methodology proposed by Minsan and Minsan [7]. They emphasized that forecasting future observations often involves higher uncertainty due to the accumulation of model and estimation errors. The testing dataset (i.e., the last 24 months) was used to assess how well the SARIMA models could predict future rainfall values across the four river basins. The goal was to ensure that the Box-Jenkins models not only fit the training data but also perform well in out-of-sample forecasting. The best SARIMA model for each basin was selected based on its ability to minimize residual autocorrelation and pass all relevant diagnostic tests, including the t-test for coefficient significance, Levene's test for homoscedasticity, the ARCH–LM test for conditional heteroscedasticity, the Kolmogorov–Smirnov (KS) test for residual normality, and the Ljung–Box test for independence. These models, along with their equations and transformation techniques (if applied), are summarized below. Ping River Basin: The selected model was SARIMA(0,1,1)(1,0,0)$_{12}$, with no transformation applied. The forecasting equation is

$$Y_t = Y_{t-1} + 0.8185Y_{t-12} - 0.8185Y_{t-13} - 0.9760e_{t-1}.$$

Wang River Basin: The model SARIMA(0,1,2)(0,1,1)$_{12}$ was used, with a Box-Cox transformation applied to the data $Z = Y^{\frac{1}{1.1}}$. The forecasting equation is

$$Z_t = Z_{t-1} + Z_{t-12} - Z_{t-13} - 0.7870e_{t-1} - 0.2041e_{t-2} - 0.8608e_{t-12} + 0.6774e_{t-13} + 0.1757e_{t-14}.$$

Yom River Basin: The SARIMA(0,1,2)(0,1,3)$_{12}$ model was selected, along with a Box-Cox transformation with $Z = Y^{\frac{1}{1.1}}$. The equation is

$$\begin{aligned} Z_t = {} & Z_{t-1} + Z_{t-12} - Z_{t-13} - 1.129e_{t-1} + 0.219e_{t-2} - 0.681e_{t-12} + 0.7689e_{t-13} \\ & - 0.1490e_{t-14} - 0.549e_{t-24} + 0.6197e_{t-25} - 0.1203e_{t-26} + 0.522e_{t-36} - 0.5891e_{t-37} + 0.1143e_{t-38}. \end{aligned}$$

Nan River Basin: The chosen model was SARIMA(0,1,1)(1,1,0)$_{12}$, using a Box-Cox transformation with $Z = Y^{\frac{1}{1.448}}$. The equation is

$$Z_t = Z_{t-1} + 0.446Z_{t-12} - 0.446Z_{t-13} + 0.554Z_{t-24} - 0.554Z_{t-25} - 0.9192e_{t-1}.$$

Table 3 summarizes the selected Box–Jenkins SARIMA models for the Ping, Wang, Yom, and Nan river basins, together with adequacy diagnostics including the t-test, Levene's test, ARCH-LM test, Kolmogorov–Smirnov (KS) test, and Ljung–Box (LB) test at lag 12. All models satisfied the diagnostic criteria with no evidence of residual autocorrelation, variance heterogeneity, non-normality, or conditional heteroskedasticity, confirming their suitability for forecasting rainfall in each river basin.

Table 3. Box-Jenkins models and model adequacy diagnostics by river basin.

River basin	Box-Jenkins SARIMA(p,d,q)(P,D,Q)$_{12}$	t-test (P-value)	Levene (P-value)	Arch-LM (P-value)	KS (P-value)	LB Lag 12 (P-value)
Ping	SARIMA(0,1,1)(1,0,0)$_{12}$	0.05 (0.963)	0.00 (0.963)	11.372 (0.251)	0.090 (0.054)	14.15 (0.166)
Wang	SARIMA(0,1,2)(0,1,1)$_{12}$	0.15 (0.878)	0.33 (0.565)	4.672 (0.862)	0.085 (0.145)	13.28 (0.151)
Yom	SARIMA(0,1,2)(0,1,3)$_{12}$	-0.71 (0.480)	0.05 (0.832)	10.411 (0.318)	0.074 (>0.150)	5.23 (0.632)
Nan	SARIMA(0,1,1)(1,1,0)$_{12}$	0.11 (0.916)	3.2 (0.078)	14.312 (0.112)	0.095 (0.066)	6.59 (0.763)

3.2.2 Summary of Forecasting Performance Based on sMAPE

To identify the most effective forecasting models for monthly rainfall across each river basin, the symmetric Mean Absolute Percentage Error (sMAPE) values from the testing dataset were compared, as shown in Table 4. The model with the lowest sMAPE value in each basin was selected as the most appropriate for 24-month-ahead forecasting. The results can be summarized as follows. For the Ping River Basin, the ForecastSheet-HW method achieved the lowest sMAPE of 37.0, outperforming all other models. Thus, ForecastSheet-HW is identified as the most suitable model for forecasting rainfall in this basin. In the Wang River Basin, the GRG-D method yielded the lowest sMAPE value of 32.1, indicating its superiority over other decomposition-based and Holt-Winters models. Therefore, GRG-D was selected as the best model for this basin. For the Yom and Nan River Basins, both GRG-HW and WOA-HW shared the lowest sMAPE values of 40.6 and 32.5, respectively. This suggests that the optimization-enhanced Holt-Winters models provided the most accurate forecasts for these two basins. Overall, when evaluating all models across the four basins, GRG-D or GRG-HW optimization methods produced the lowest sMAPE values in three out of four river basins. In contrast, the ForecastSheet-HW method was the most suitable model only for the Ping River Basin.

Table 4. sMAPE of the testing dataset for monthly rainfall in each river basin.

Model		**River basin**			
		Ping	**Wang**	**Yom**	**Nan**
Decomposition	Minitab-D	40.8	34.8	42.0	33.0
	WOA-D	46.3	34.4	50.0	34.9
	GRG-D	45.3	32.1*	45.1	33.5
Holt-Winters	ForecastSheet-HW	37.0*	47.0	43.7	57.3
	WOA-HW	42.2	39.0	40.6*	32.5*
	GRG-HW	42.2	39.0	40.6*	32.5*
Box-Jenkins		40.6	35.5	43.1	35.5

Note: * indicates the lowest sMAPE for each river basin.

Table 5 and Figure 6 present the 24-month-ahead rainfall forecasts for each river basin from April 2024 to March 2026. The results indicate that the highest inflow volumes are expected during the rainy season, typically between August and September for the Ping, Wang, and Yom Rivers, and July to August for the Nan River. This seasonal trend aligns closely with Thailand's climatological patterns, where monsoon rainfall peaks during these months. The forecasted rainfall volumes clearly demonstrate strong seasonal variation, with particularly high rainfall predicted between July and September, reflecting the annual rainy season. These peaks suggest a significant increase in water inflows to the river systems during this period. Such forecasts are highly valuable for water resource planning and management. They support proactive strategies in maintaining river levels, preparing for flood risks during the rainy season, and allocating water for downstream use during the dry season. Overall, the 24-month forecasts provide essential insights for sustainable water management across the upper northern region of Thailand.

Table 5. Forecasted Monthly Rainfall for Each River Basin (in millimeters), April 2024 – March 2026.

Month/Year	River basin			
	Ping	Wang	Yom	Nan
Apr-24	123	160	148	296
May-24	326	289	315	409
Jun-24	209	258	272	458
Jul-24	305	322	401	799*
Aug-24	411*	445*	524*	983*
Sep-24	397*	336*	434*	586
Oct-24	255	230	185	176
Nov-24	19	62	36	104
Dec-24	15	22	9	24
Jan-25	11	50	48	124
Feb-25	37	15	19	18
Mar-25	15	34	38	128
Apr-25	126	158	148	296
May-25	329	288	315	409
Jun-25	212	257	272	458
Jul-25	308	321	401	799*
Aug-25	414*	444*	524*	983*
Sep-25	400*	335*	434*	586
Oct-25	258	229	185	176
Nov-25	22	60	36	104
Dec-25	18	20	9	24
Jan-26	14	49	48	124
Feb-26	40	14	19	18
Mar-26	18	33	38	128

Note: * indicates the first and second highest rainfall volumes within each 12 months.

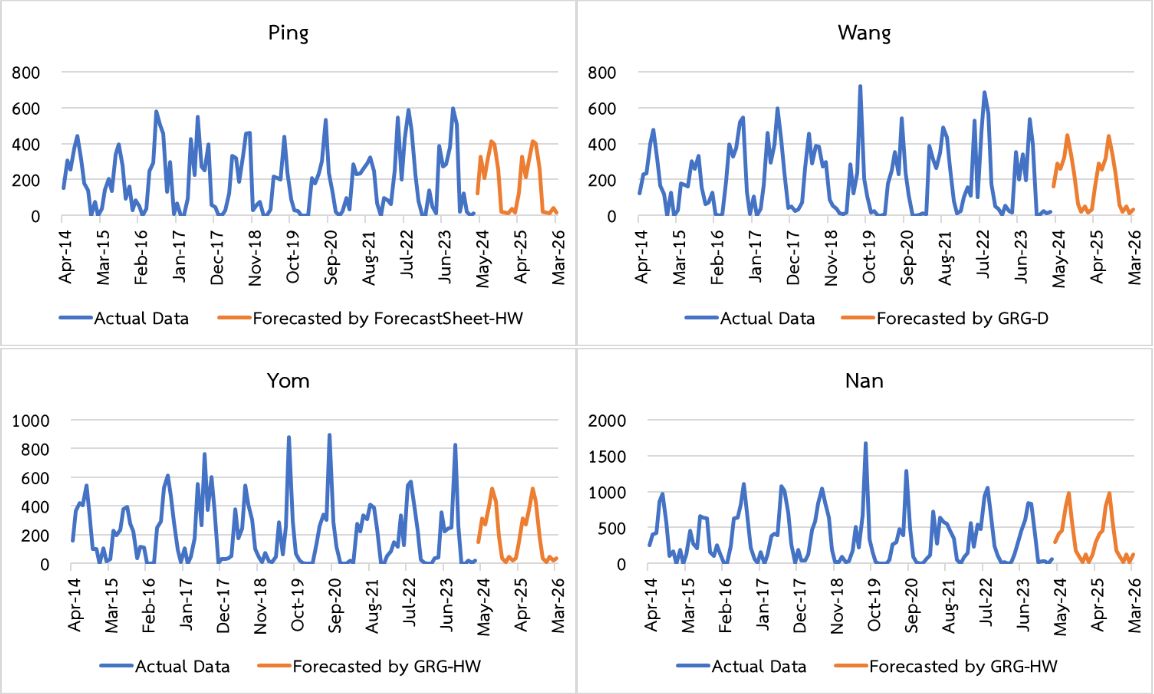


Figure 6. Actual and Forecasted Monthly Rainfall for Each River Basin Using the Most Appropriate Forecasting Methods.

4. Conclusions

This study examined the use of the GRG Nonlinear Solver, a built-in optimization tool in Microsoft Excel, to estimate parameters for time series forecasting models with trend and seasonality, specifically the decomposition (GRG-D) and Holt-Winters (GRG-HW) models. Based on the 24-month ahead forecasts, the Holt-Winters models proved most effective for the Ping, Yom, and Nan Rivers, reflecting rainfall patterns where recent observations should be given greater importance and older data progressively less weight through exponential smoothing. In contrast, the Wang River forecasts were best captured by the decomposition model, which treats all historical data points equally in the prediction process. These results indicate that the suitability of each method is closely tied to the way it weights past data, aligning with the inherent temporal characteristics of rainfall in each basin. While traditional estimation tools such as Minitab-D and ForecastSheet-HW are user-friendly and widely accessible, they often fall short in minimizing forecasting error as defined by the RMSE criterion. In particular, it is known to be inefficient in exploring the entire solution space and often fails to reach the global optimum. In contrast, metaheuristic algorithms such as the Whale Optimization Algorithm (WOA) can significantly improve forecasting accuracy, but they require programming expertise, which limits their practical use in non-technical settings. The GRG Nonlinear Solver offers a balanced alternative. It eliminates the need for coding while achieving RMSE values that are competitive with WOA, with differences typically within one decimal place. In this study, GRG-D and GRG-HW produced forecasts closely aligned with actual rainfall across the four basins. During the 24-month testing period, GRG-HW and WOA-HW delivered the lowest sMAPE values in the Yom and Nan Rivers, while GRG-D consistently provided the most accurate forecasts for the Wang River and showed robust performance overall. ForecastSheet-HW, however, remained the best-performing model for the Ping River. These findings suggest that the GRG Nonlinear Solver is a practical and effective alternative for rainfall forecasting, particularly when ease of use is a key consideration. To further enhance its utility, future research may explore the tuning of GRG parameters, such as solver bounds and convergence settings. Additionally, integrating the forecasting models with physical and climatic variables, such as reservoir inflows, watershed characteristics, and regional weather patterns, may support more robust decision-making in long-term water resource planning.

5. Acknowledgement

We would like to thank the Department of Mathematics and Statistics, Faculty of Science and Technology, Chiang Mai Rajabhat University, for providing us with research space and equipment. We also gratefully acknowledge the Data Science Research Center, Department of Statistics, Faculty of Science, Chiang Mai University, for providing access to Minitab software, which was used in the data analysis.

Author Contributions: For research articles with multiple authors, a brief paragraph outlining their individual contributions must be included. The following statements should be used "Conceptualization, Pradthana Minsan; methodology, P.M.; software and coding, P.M. and Watha Minsan.; analysis and visualization, P.M.; validation, P.M., and W.M.; writing—original draft preparation, P.M.; writing—review and editing, P.M., and W.M. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] Limsakul, A.; Singhruck, P. Long-term trends and variability of total and extreme precipitation in Thailand. *Atmospheric Research* **2016**, *169*, 301–317. https://doi.org/10.1016/j.atmosres.2015.10.015

[2] Mishra, A. K.; Desai, V. R. Drought forecasting using stochastic models. *Stoch Environ Res Risk Assess* **2005**, *19*, 326–339. https://doi.org/10.1007/s00477-005-0238-4

[3] Wang, W. C.; Chau, K. W.; Xu, D. M.; Chen, X. Y. Improving forecasting accuracy of annual runoff time series using ARIMA based on EEMD decomposition. *Water Resour Manage* **2015**, *29*(8), 2655–2675. https://doi.org/10.1007/s11269-015-0962-6

[4] Saeying, J.; Minsan, W.; Taninpong, P. Forecasting model for the amount of water flowing into the reservoirs of the electricity generating authority of Thailand (EGAT). *Recent Science and Technology* **2023**, *15*(2), 494–510. (In Thai)

[5] Jiang, W.; Wu, X.; Gong, Y.; Yu, W.; Zhong, X. Holt–Winters smoothing enhanced by fruit fly optimization algorithm to forecast monthly electricity consumption. *Energy* **2020**, *193*, 116779. https://doi.org/10.1016/j.energy.2019.116779

[6] Minsan, W.; Minsan, P. Decomposition and Holt-Winters enhanced by the whale optimization algorithm for forecasting the amount of water inflow into the large dam reservoirs in southern Thailand. *Journal of Current Science and Technology* **2024**, *14*(2), article 38.

[7] Minsan, P.; Minsan, W. Monthly volumes of water inflow into the large dam reservoirs in eastern Thailand forecasting by the cuckoo search optimization enhanced decomposition and Holt-Winters techniques. *Thai Journal of Operations Research* **2024a**, *12*(2), 69–89. (In Thai)

[8] Minsan, P.; Minsan, W. Decomposition and Holt-Winters techniques enhanced by whale optimization algorithm: Case study of pm2.5 forecasting in 8 northern provinces of Thailand. *Thai Science and Technology Journal* **2024b**, *32*(6), 12–34. (In Thai)

[9] Minsan, W.; Minsan, P. Incorporating decomposition and the Holt-Winters method into the whale optimization algorithm for forecasting monthly government revenue in Thailand. *Science & Technology Asia* **2023**, *28*(4), 38–53.

[10] Minsan, W.; Minsan, P.; Panichkitkosolkul, W. Enhancing decomposition and Holt-Winters weekly forecasting of pm2.5 concentrations in Thailand's eight northern provinces using the cuckoo search algorithm. *Thailand Statistician* **2024**, *22*(4), 963–985.

[11] The upper northern region irrigation hydrology center, the Royal Irrigation Department (RID), Thailand. https://www.hydro-1.net/ (accessed 2025-01-16).

[12] Mirjalili, S.; Lewis, A. The whale optimization algorithm. *Advances in Engineering Software* **2016**, *95*, 51–67. https://doi.org/10.1016/j.advengsoft.2016.01.008

[13] Minsan, P. Comparing methods of optimization in solver of Microsoft excel 2007 and 2019: A case study of statistical models. *The Journal of King Mongkut's University of Technology North Bangkok* **2021**, *31*(3), 496–511. https://doi.org/10.14416/j.kmutnb.2021.05.013

[14] Minsan, W.; Minsan, P. GRG-HW and GRG-D optimization models for rainfall forecasting in the Yom river basin. *Figshare* **2025**. https://doi.org/10.6084/m9.figshare.28953389

[15] Flores, B. E. A pragmatic view of accuracy measurement in forecasting. *Omega* **1986**, *14*(2), 93–98.

[16] Jackson, E. K.; Roberts, W.; Nelsen, B.; Williams, G. P.; Nelson, E. J.; Ames, D. P. Introductory overview: Error metrics for hydrologic modelling – A review of common practices and an open source library to facilitate use and adoption. *Environmental Modelling & Software* **2019**, *119*, 32–48. https://doi.org/10.1016/j.envsoft.2019.05.001.

ASEAN
Journal of Scientific and Technological Reports
Online ISSN:2773-8752

Research article

Evaluation of the Milling Performance on Wood-Plastic Composite

Wikanet Phetsuwan[1], Kunlapat Thongkaew[2*] and Chainarong Srivabut[3]

[1] Faculty of Engineering, Rajamangala University of Technology Srivijaya, Songkhla, 90000, Thailand
[2] Faculty of Engineering, Prince of Songkla University, Songkhla, 90112, Thailand
[3] Faculty of Engineering, Rajamangala University of Technology Srivijaya, Songkhla, 90000, Thailand
* Correspondence: kunlapat.t@psu.ac.th

Citation:
Phetsuwan, W.; Thongkaew, K.; Srivabut, C. Evaluation of the milling performance on wood-plastic composite. *ASEAN J. Sci. Tech. Report.* **2025**, *28*(5), e259643. https://doi.org/10.55164/ajstr.v28i5.259643.

Article history:
Received: June 5, 2025
Revised: August 27, 2025
Accepted: September 1, 2025
Available online: September 14, 2025

Abstract: This research aims to study the milling factors that affect the quality of milled surfaces, including surface roughness and burr formation. A composite material composed of recycled polypropylene plastic and rubber wood powder was selected. The side milling parameters evaluated comprise spindle speeds of 330, 610, and 850 rpm; feed rates of 43, 120, and 200 mm/min; and cutting depths of 1, 3, and 5 mm. Surface roughness decreased with increased spindle speeds, but reduced feed rates and depths of cut. Burr formation diminished with increased spindle speed and feed rate, but required a lower depth of cut. Regression models for surface roughness and burr development were established, demonstrating prediction errors of approximately 1.54% and 2.29%. The optimal milling parameters to minimize surface roughness were determined to be a spindle speed of 850 rpm, a feed rate of 43 mm/min, and a depth of cut of 1 mm, resulting in a surface roughness of 5.184 µm. The suggested conditions to reduce burr formation were a spindle speed of 610 rpm, a feed rate of 200 mm/min, and a depth of 1 mm, which resulted in a burr height of 0.251 mm. The optimal parameters for minimizing both burrs and roughness were a spindle speed of 610 rpm, a feed rate of 43 mm/min, and a depth of cut of 1 mm. This yielded 6.958 µm of roughness and 0.255 mm of burr creation. These conditions eliminate the necessity for post-milling finishing and correspond to ISO 21920 surface texture standards.

Keywords: Milling operation; wood-plastic composite; surface roughness; burr formation

1. Introduction

Rubber wood is a popular material for furniture production due to its high strength, beautiful texture, and corrosion resistance [1]. Currently, the average annual production of rubber wood furniture and construction materials is 2,700,000 cubic meters, valued at 34 billion baht per year [2]. The processing of rubber wood generates wood waste, which is then utilized to manufacture wood-plastic composites (WPCs), consequently enhancing the value of this waste [3]. WPCs serve as a substitute for natural wood, offering better properties, such as high strength, good impact resistance, ease of molding, resistance to mold growth, and environmental friendliness [4]. As a result, the value of WPCs is expected to increase to 13.244 billion USD by 2030 [5]. The production of WPCs involves mixing wood sawdust and plastics, such as polyethylene, in appropriate proportions and then extruding them to create composite materials. These materials can be formed using molds in various methods, such as hot compression, extrusion, or injection molding [6]. However, the formed materials still have limitations in terms of geometry and size due to mold constraints. Consequently, to manufacture products with specific dimensions and patterns, additional machining processes, such as milling, drilling, cutting, or turning, are necessary [7]. Milling is a popular technique in post-machining due to its high cutting rate and the ability to machine diverse geometries.

There are some studies related to milling for processing WPCs [8-12]. Thanate et al. [8] investigated the effects of milling parameters on the surface roughness of WPCs and found that the depth of cut had the most significant impact on surface roughness, followed by feed rate and spindle speed, respectively. A comparison of milling performance for different plastics used in WPCs, such as polyethylene (PE), polypropylene (PP), and polyvinyl chloride (PVC), showed that WPCs made from PP plastic had the lowest surface roughness due to its high density and high melting temperature [9]. Chainarong et al. [10] employed Response Surface Methodology (RSM) to determine the optimal milling conditions for WPCs, with a spindle speed of 1,000 rpm, a feed rate of 315 mm/min, and a cutting depth of 1 mm, resulting in a surface roughness of approximately 2.865 μm. They also studied the impact of milling parameters on the surface roughness of WPCs made from high-density polyethylene (HDPE) and rubberwood sawdust (RWS) [11], finding that surface roughness decreased as spindle speed increased, while both feed rate and cutting depth decreased. Wikanet et al. [12] investigated the milling factors affecting the qualities of WPCs and found that surface roughness decreased with increased spindle speed. The burr formation decreased with an increase in spindle speed and feed rate, but a decrease in cutting depth. In addition, up-milling resulted in higher burr formation than down-milling because the cutting-edge rotation was opposite to the direction of workpiece feeding.

The analysis of previous research revealed that while milling factors, including spindle speed, feed rate, and depth of cut, have been examined for their impact on the surface roughness of WPCs, the majority of studies have focused on surface milling. There has been little comprehensive research on side milling and slot milling, which influence burr formation as well as surface roughness. This research focuses on investigating the milling parameters and patterns that affect the surface quality of milled surfaces and burr formation in WPCs. The criteria evaluated comprise spindle speed, feed rate, and depth of cut in both shoulder milling and slot milling conditions. The determination of the suitable milling condition resulting in good machined quality was also considered. The results of this study will enhance WPC processing in the manufacturing of furniture and construction materials in many forms, thereby improving the market price of these products.

2. Materials and Methods

2.1 Material and Equipment

2.1.1 Wood-Plastic Composite

The WPC developed by Chainarong et al. [13] was employed in this research. The components and properties of this WPC are detailed in Table 1. The material size is 55 × 55 × 18 mm, as shown in Figure 1.

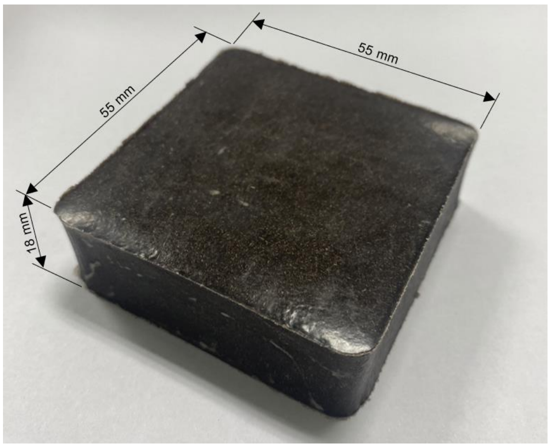


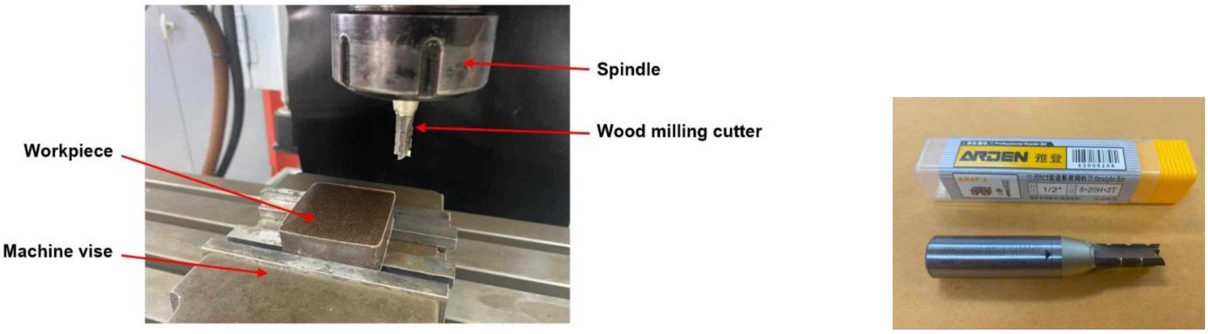
Figure 1. Wood-Plastic Composite

Table 1. Components and properties of WPC [13]

Components of WPC (weight percentage)	Properties of WPC
- Recycled Polypropylene: 51.8 % - Rubberwood Powder: 35.9 % - Calcium Carbonate: 7.2 % - Other: 5.1%	- Tensile strength and modulus: 24.82 MPa and 1.20 GPa - Compressive strength and modulus: 22.88 MPa and 1.32 GPa

2.1.2 Milling Machine and Equipment

The experiment utilized a Full Mark VBM-3V vertical milling machine (Figure 2(a), capable of achieving a maximum spindle speed of 1,250 rpm and a maximum feed rate of 875 mm/min. A 3-flute tungsten carbide cutter (ARDEN, model AK47-2) with an 8 mm diameter was employed. The cutting edges are sharpened and polished with optimal rake and relief angles to minimize cutting forces, reduce surface roughness, and prevent burr development. The tool design optimizes chip evacuation and heat dissipation, making it appropriate for both hardwood and composite wood materials, while maintaining a balance between material removal rate and surface quality, as shown in Figure 2(b).



(a) (b)

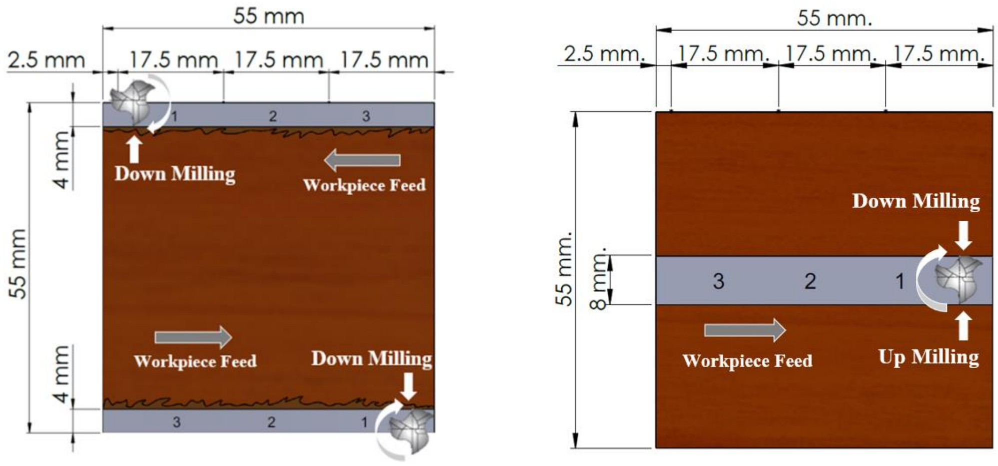
Figure 2. (a) Experimental setup and (b) Wood milling cutter

2.2 Experimental Design

A full factorial design was employed in this study to investigate the main effects and interaction effects of the milling parameters for all conditions. Three experimental factors were established, each with three levels, resulting in a total of 27 experimental conditions, as shown in Table 2. The experiments were conducted three times, resulting in a total of 81 trials. The milling operations of side milling and slot milling (Figure 3) entailed the cutting of the workpiece in the direction same as the workpiece feed direction over a distance of 55 mm. The side milling results will be compared with those of slot milling (Figure 3(b)), which was completed by Wikanet et al. [12].

Table 2. Milling parameters

Parameter	Unit	Level		
		Low	Medium	High
Speed (*S*)	rpm.	330	610	850
Feed Rate (*F*)	mm/min	43	120	200
Depth of Cut (*D*)	mm	1	3	5



(a) (b)

Figure 3. Milling processes: (a) side milling and (b) slot milling [12]

The average surface roughness (Ra) was evaluated using a Mitutoyo surface roughness tester (model SJ-210) equipped with a stylus tip radius of 2 µm, a resolution of 0.35 µm, and a measurement speed of 0.5 mm/s (Figure 4(a)). The burr width was subsequently evaluated using a CARL ZEISS AXIO IMAGER A1m microscope with magnifications ranging from 1.25x to 100x and 2 mm/revolution (Figure 4(b)). The burr formation (WB) was assessed by collecting samples over a measurement length from the mill edge to the large burr area. For both parameters, three repeated measurements were performed at different locations, with a measurement length of 17.5 mm, to ensure accuracy and reliability (Figure 3).

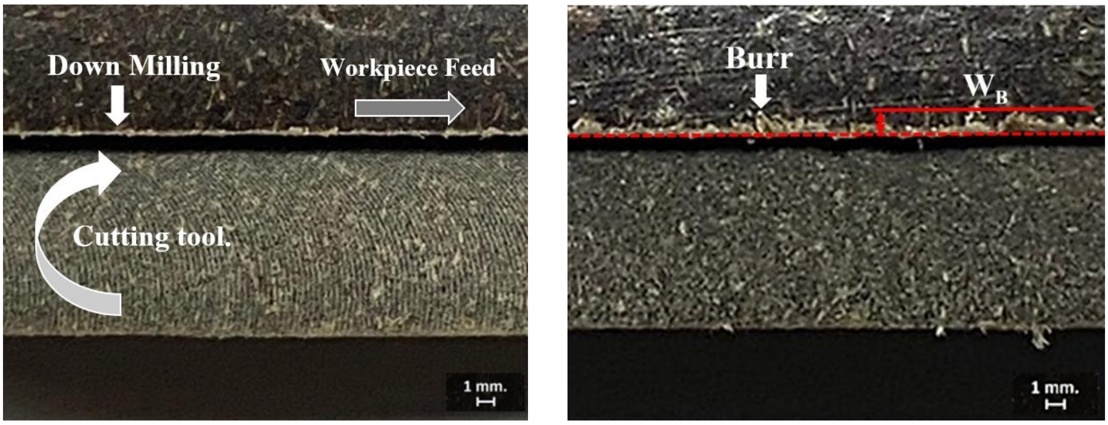
(a) (b)

Figure 4. (a) Mitutoyo SJ-210 surface roughness tester and (b) CARL ZEISS AXIO IMAGER A1m microscope.

3. Results and Discussion

3.1 Milling Performance on Wood-Plastic Composites

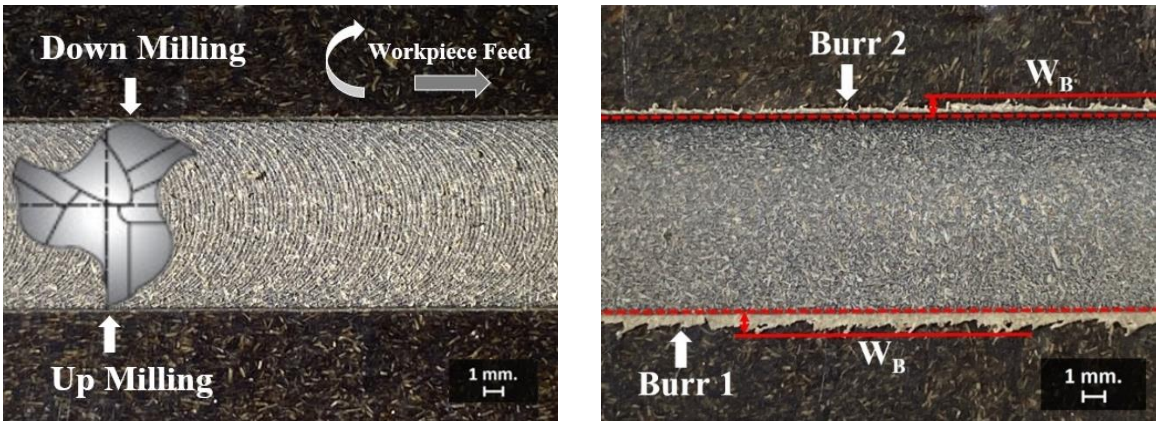
The experimental results for side milling showed that the average surface roughness was 9.803 µm with a standard deviation of 2.529 µm, while the average burr formation was 0.280 mm with a standard deviation of 0.020 mm. Negligible burr formation (0.245 mm) was observed at a spindle speed of 610 rpm, feed rate of 200 mm/min, and depth of cut of 1 mm, as illustrated in Figure 5(a). However, under these conditions, the milled surface was relatively rough, with an average roughness of 16.213 µm. In contrast, at a spindle speed of 850 rpm, a feed rate of 43 mm/min, and a depth of cut of 1 mm, a larger burr formation of 0.273 mm was observed, while the surface roughness decreased significantly to 5.184 µm (Figure 5(b)). The burr formation was significantly induced by high spindle speed, low feed rate, and feed depth as softening material.



(a) (b)

Figure 5. Characteristics of side milled workpiece: (a) Negligible burr formation and (b) burr formation.

In comparison with slot milling (Figure 6) from the previous study [12], side milling at 610 rpm spindle speed, 200 mm/min feed rate, and 1 mm depth of cut produced a surface roughness of 6.575 µm, with burr formation of 0.770 mm for up-milling and 0.257 mm for down-milling (Figure 6a). Under the same depth of cut but with 850 rpm spindle speed and 43 mm/min feed rate, slot milling produced a similar surface roughness (6.575 µm), while burr formation increased to 0.806 mm for up-milling and 0.269 mm for down-milling (Figure 6b). These results indicate that side and slot milling show no significant difference in overall machining performance. However, slot milling tends to generate burrs on both sides of the cut edge, with up-milling producing larger burrs because the cutting edge rotates in the opposite direction to the feed direction [14].



(a) (b)

Figure 6. Characteristics of slot milled workpiece: (a) Negligible burr formation and (b) burr formation [12].

3.2 The Effects of Side Milling Factors on Surface Roughness

The Analysis of Variance (ANOVA) was employed to analyze the results of the side milling using Design-Expert. The data was initially examined for normality, and it was determined to be both normal and independently distributed. Thus, the data can be used to investigate the effect of milling factors on surface roughness (R_a). The main effects and interaction effects of spindle speed (*S*), feed rate (*F*), and feed depth (*D*) were analyzed at a 95% confidence level. The results in Table 3 show that all primary parameters had a substantial impact on surface roughness, with spindle speed having the most significant influence. The surface roughness was considerably influenced by the interaction between spindle speed and feed rate (*SF*), as well as between feed rate and feed depth (*FD*). However, the interaction between spindle speed and feed depth (*SD*), and the three-way interaction (*SFD*) were not significant (p-value > 0.05). The insignificance can result from the mixed positive and negative correlations of the cutting parameters.

Table 3. ANOVA analysis on surface roughness

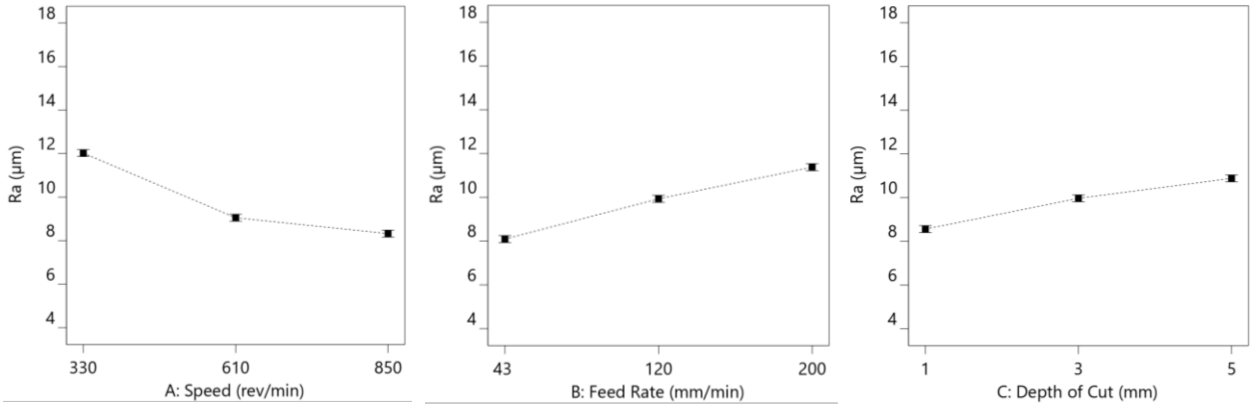
Source	Sum of Squares	df	Mean Square	F-Value	P-Value
Model	488.51	26	18.79	60.94	< 0.0001*
S-Speed	207.68	2	103.84	336.81	< 0.0001*
F-Feed Rate	146.50	2	73.25	237.59	< 0.0001*
D-Depth of Cut	73.77	2	36.88	119.64	< 0.0001*
SF	49.41	4	12.35	40.06	< 0.0001*
SD	1.57	4	0.3924	1.27	0.2920
FD	5.14	4	1.29	4.17	0.0051*
SFD	4.44	8	0.5554	1.80	0.0970
Pure Error	16.65	54	0.3083		
Total	505.16	80			
R^2	96.70				
Adjusted R^2	95.12				

Note: * Parameters have a significant effect at the 0.05 significance level.

The regression model to predict surface roughness (*Ra*) was developed by excluding these variables and reconsidering them, as illustrated in Equation (1). The coefficient of determination (R^2) was 96.70%, and the adjusted R^2 was 95.12%, indicating that the equation can be used to predict results accurately. From Equation (1), the subscripts 1 and 2 represent the levels of each machining parameter, corresponding to the low and high levels, respectively. For example, S_1 represents the low-speed level of 330 rpm, while S_2 represents the high-speed level of 850 rpm.

$$Ra = 9.80 + 2.22S_1 - 0.7472S_2 - 1.71F_1 + 0.1308F_2 - 1.24D_1 + 0.1648D_2 - 0.6321S_1F_1 + 0.8549S_2F_1 - 0.7965S_1F_2 + 0.9881S_2F_2 + 0.0337F_1D_1 - 0.1056F_2D_1 \quad (1)$$

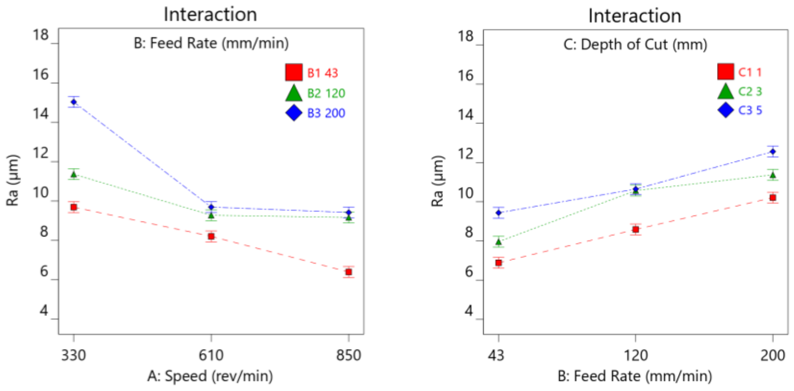
The relationship between the main milling factors that influence surface roughness is illustrated in Figure 7. The roughness of the surface decreases as the spindle speed increases (Figure 7(a)). Nevertheless, the surface roughness increases as the feed rate and depth of cut are increased (Figures 7(b) and (c)). According to Merchant's orthogonal cutting model [15], the cutting force is directly proportional to the feed rate and depth of cut, while being inversely proportional to the cutting speed, which is influenced by the spindle speed. Increased spindle speeds enhance surface smoothness by diminishing shear friction during milling, whereas higher feed rates and larger cuts intensify these forces. An increase in cutting force raises the cutting temperature, resulting in greater surface roughness and burr formation, which is particularly relevant to the ductile properties of WPCs. This conclusion coincides with the research presented in [10,16]



(a) (b) (c)

Figure 7. The effects of main milling parameters on surface roughness: (a) spindle speed, (b) feed rate, and (c) feed depth

The effects of combined milling parameters on surface roughness are shown in Figure 8, where B1–B3 correspond to feed rates of 43, 120, and 200 mm/min, and C1–C3 correspond to depths of cut of 1, 3, and 5 mm, respectively. Analysis of interaction effects indicates that a high spindle speed is essential for achieving a low surface roughness value; concurrently, a low feed rate must be maintained (Figure 8(a)). The feed rate and depth of cut exhibit a comparable pattern, as lower feed rates and depths of cut lead to reduced surface roughness (Figure 8(b)).



(a) (b)

Figure 8. The effects of combined milling parameters on surface roughness: (a) spindle speed vs. feed rate and (b) feed rate vs. feed depth

The optimal side milling condition for minimizing surface roughness was achieved at a spindle speed of 850 rpm, a feed rate of 43 mm/min, and a depth of cut of 1 mm, yielding a predicted surface roughness of approximately 5.184 µm (Table 4). This value meets the roughness requirements specified in ISO 21920:2021 surface texture [17] without the necessity of post-machining surface finishing. The predictive equation was validated by conducting a total of 10 milling conditions (Table 4), and the results showed that the average difference between the predicted surface roughness values and the actual surface roughness was 1.544%, which is small and acceptable for implementation.

Table 4. Evaluation of the side-milled surface roughness regression model

No.	**Speed (rpm)**	**Feed Rate (mm/min)**	**Depth of Cut (mm)**	**Cal. R_a (F_t)**	**Actual R_a (d_t)**	**Error $e_t = d_t - F_t$**	$\frac{\lvert e_t\rvert}{d_t} \times 100$
1	850	43	1	5.184	5.256	0.072	1.370
2	610	120	1	7.925	8.332	0.407	4.885
3	610	43	3	8.068	8.058	-0.010	0.124
4	850	43	3	6.260	6.253	-0.007	0.112
5	610	43	5	9.541	9.568	0.027	0.282
6	330	120	5	12.073	12.110	0.037	0.306
7	850	120	1	7.815	8.062	0.247	3.064
8	330	43	1	8.477	8.612	0.135	1.568
9	610	200	5	10.869	11.111	0.242	2.178
10	610	200	1	8.519	8.653	0.134	1.549
						MAPE	1.544

3.3 The Effects of Side Milling Factors on Burr Formation

ANOVA was utilized to examine the impact of milling conditions on burr development. Table 5 indicates that all primary and interaction milling parameters strongly influence burr development, with feed depth exerting the most significant effect. Similar to surface roughness analysis, the combined parameters of the three-way interaction (*SFD*) do not significantly affect burr formation. This result can also be attributed to the different correlations among cutting parameters.

Table 5. ANOVA analysis on burr formation

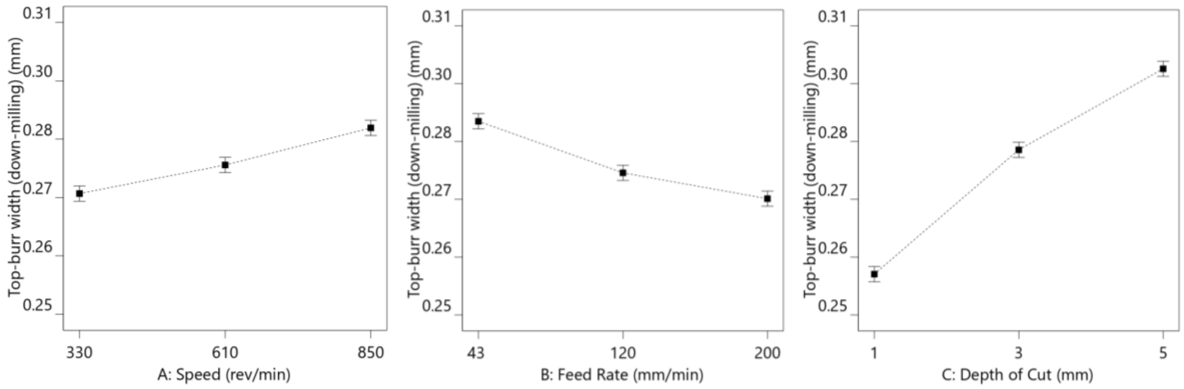
Source	Sum of Squares	df	Mean Square	F-Value	P-Value
Model	0.0312	18	0.0017	59.60	< 0.0001*
S-Speed	0.0022	2	0.0011	37.02	< 0.0001*
F-Feed Rate	0.0031	2	0.0016	53.88	< 0.0001*
D-Depth of Cut	0.0216	2	0.0108	371.38	< 0.0001*
SF	0.0013	4	0.0003	11.41	< 0.0001*
SD	0.0010	4	0.0002	8.49	< 0.0001*
FD	0.0020	4	0.0005	17.14	< 0.0001*
SFD	0.0003	8	0.0000	1.49	0.1812
Pure Error	0.0015	54	0.0000		
Total	0.0330	80			
R^2	95.53				
Adjusted R^2	93.37				

Note: * Parameters have a significant effect at the 0.05 significance level.

A regression equation for predicting burr formation (W_B) was established in Equation (2). The coefficient of determination (R^2) is 95.53%, while the adjusted R^2 is 93.37%, indicating that the equation is suitable for predicting results. Similarly, as in Eq. (1), subscripts 1 and 2 denote the low and high settings of each machining parameter, respectively. For instance, S_1 corresponds to the low-speed level of 330 rpm, whereas S_2 corresponds to the high-speed level of 850 rpm.

$$W_B = 0.2802 - 0.0060S_1 - 0.0005S_2 - 0.0083F_1 - 0.0017F_2 - 0.0212D_1 + 0.0028D_2 - 0.0027S_1F_1 - 0.0049S_2F_1 - 0.0004S_1F_2 + 0.0044S_2F_2 + 0.0017S_1D_1 - 0.0017S_2D_1 - 0.0066S_1D_2 + 0.0032S_2D_2 - 0.0047F_1D_1 + 0.0044F_2D_1 + 0.0096F_1D_2 - 0.0054F_2D_2 \quad (2)$$

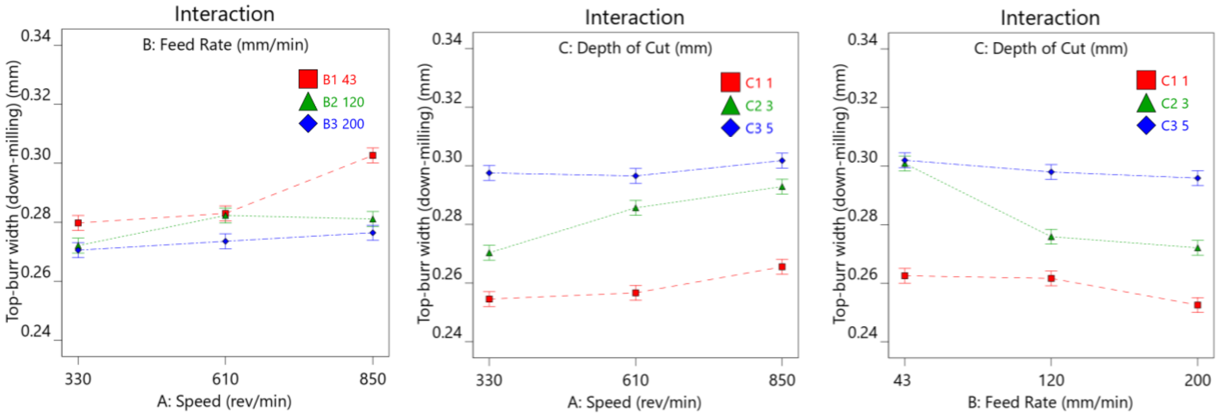
It was demonstrated that increasing spindle speed results in enhanced burr formation (Figure 9(a)). Nonetheless, increasing the feed rate and decreasing the depth of cut leads to a reduction in burr formation (Figures 9(b) and (c)). A greater depth of cut implies increased material removal, hence increasing burr development.



(a) (b) (c)

Figure 9. The effects of main milling parameters on burr formation: (a) spindle speed, (b) feed rate, and (c) feed depth

Figure 10 illustrates the impact of various milling parameters on burr development, with B1–B3 representing feed rates of 43, 120, and 200 mm/min, and C1–C3 denoting depths of cut of 1, 3, and 5 mm, respectively. Figure 10 illustrates that reducing burr formation requires a combination of a lower spindle speed and a small depth of cut, along with a relatively high feed rate. According to Merchant's orthogonal cutting model, this parameter setting reduces the overall cutting force and heat generation in the workpiece, thereby minimizing burr formation on the machined surface [15]. This result agrees with the findings of [18]. The optimal milling condition for the lowest burr formation was determined through an analysis. The optimal cutting conditions for minimizing burr formation were a spindle speed of 610 rpm, a feed rate of 200 mm/min, and a depth of cut of 1 mm, resulting in a predicted burr formation of 0.251 mm (Table 6), and eliminating the need for post-machining.



(a) (b) (c)

Figure 10. The effects of combined milling parameters on burr formation: (a) spindle speed vs. feed rate, (b) spindle speed vs. feed depth, and (c) feed rate vs. feed depth

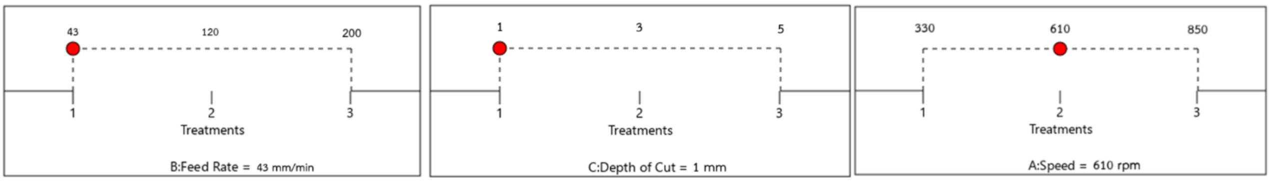
The regression model was validated by conducting 10 milling conditions as given in Table 6. The result was similar to those used for surface roughness analysis. The average difference between the predicted burr formation value and the actual burr formation was 2.294%, which is considered acceptable in comparison.

Table 6. Evaluation of the burr formation regression model

No.	Speed (rpm)	Feed Rate (mm/min)	Depth of Cut (mm)	Cal. Burr (F_t)	Actual Burr (d_t)	Error $e_t = d_t - F_t$	$\frac{\lvert e_t \rvert}{d_t} \times 100$
1	850	43	1	0.277	0.272	-0.005	1.838
2	610	120	1	0.264	0.262	-0.002	0.763
3	610	43	3	0.299	0.288	-0.011	3.819
4	850	43	3	0.318	0.301	-0.017	5.648
5	610	43	5	0.295	0.291	-0.004	1.375
6	330	120	5	0.296	0.302	0.006	1.987
7	850	120	1	0.264	0.260	-0.004	1.538
8	330	43	1	0.256	0.264	0.008	3.030
9	610	200	5	0.294	0.289	-0.005	1.730
10	610	200	1	0.251	0.248	-0.003	1.210
						MAPE	2.294

3.4 Optimal Milling Parameters for Minimizing Surface Roughness and Burr Formation

An analysis of optimal side milling parameters was conducted to minimize surface roughness and burr formation. An I-optimal design algorithm was employed to identify the best parameter combinations, providing practical guidance for wood machining manufacturers (Figure 11). The results revealed that the optimal conditions consisted of a low feed rate of 43 mm/min, a low depth of cut of 1 mm, and a moderate spindle speed of 610 rpm, resulting in a surface roughness of 6.958 µm and a burr formation of 0.255 mm, with a desirability assessment of 84.8%. These milling conditions result in low surface roughness and minimal burr formation, in compliance with ISO 21920 requirements. It is recommended that manufacturers eliminate the need for post-finishing processes in WPC milling, especially for furniture applications that require proper assembly.



(a) (b) (c)

Figure 11. Optimal side milling parameters for minimizing both surface roughness and burr formation: (a) feed rate, (b) depth of cut, and (c) spindle speed

4. Conclusions

The evaluation of the side milling performance of WPCs revealed that the cutting parameters, including spindle speed, feed rate, and feed depth, have a significant effect on surface roughness and the formation of burrs. The most significant factors influencing surface roughness were spindle speed, feed rate, and depth of cut. On the other hand, feed rate, spindle speed, and depth of cut had the most significant effects on burr formation. The results of regression models for surface roughness and burr formation revealed that the predicted and experimental values varied by about 1.54% and 2.29%, respectively. The optimal milling parameters for minimizing roughness include a spindle speed of 850 rpm, a feed rate of 43 mm/min, and a cutting depth of 1 mm, resulting in a surface roughness of 5.184 µm. The optimal milling parameters for minimizing burr formation included a spindle speed of 610 rpm, a feed rate of 200 mm/min, and a cutting depth of 1 mm, resulting in a burr formation of 0.251 mm. To minimize both surface roughness and burr, the recommended milling parameters were 610 rpm spindle speed, 43 mm/min feed rate, and 1 mm cutting depth. This resulted in a surface roughness of 6.958 µm and a burr formation of 0.255 mm, respectively. This corresponds to ISO 21920's surface texture specification and does not require any post-milling finishing. This demonstrates that the predicted regression models can be employed to estimate the quality of milling. The result is advantageous for industries involved in the machining of composite plastics and wood products. To further enhance milling performance, such as increasing speed while maintaining cut quality, considerations may include the type and dimensions of the cutter, as well as the utilization of various composite plastic and wood materials.

5. Acknowledgements

This research would like to express its gratitude to the Department of Industrial Engineering, Faculty of Engineering, Rajamangala University of Technology Srivijaya, and the Department of Industrial and Manufacturing Engineering, Faculty of Engineering, Prince of Songkla University, for their generous support in providing the equipment, machinery, and facilities used in the research and experimentation.

Author Contributions: Conceptualization, W.P. and K.T.; methodology, W.P.; validation, K.T. and C.S.; formal analysis, K.T.; data curation, W.P.; writing—original draft preparation, W.P.; writing—review and editing, K.T. and C.S. Funding acquisition was equally supported by all authors, ensuring the successful completion of the research. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] Department of International Trade Promotion. Opportunities for Thai Rubberwood in China. https://www.ditp.go.th/post/172905 (accessed 2024-11-01).

[2] Office of Agricultural Research and Development (Public Organization). Rubber, latex, and natural rubber are among the economically significant crops. https://www.thansettakij.com/business/trade-agriculture/591319 (accessed 2024-11-01).

[3] Srivabut, C.; Ratanawilai, T.; Hiziroglie, S. Effect of nanoclay, talcum, and calcium carbonate as filler on properties of composites manufactured from recycled polypropylene and rubberwood fiber. *Constr. Build. Mater.* **2018**, *162*, 450–458. https://doi.org/10.1016/j.conbuildmat.2017.12.048

[4] Ratanawilai, T.; Taneerat, T. Alternative polymeric matrices for wood-plastic composites: Effects on mechanical properties and resistance to natural weathering. *Constr. Build. Mater.* **2018**, *172*, 349–357. https://doi.org/10.1016/j.conbuildmat.2018.03.266

[5] Acumen Research and Consulting. Wood Plastic Composites Market Analysis - Global Industry Size, Share, Trends and Forecast 2022-2030. http://www.acumenresearchandconsulting.com/wood-plastic-composites-market (accessed 2024-11-01).

[6] Cheewawuttipong, W.; Homkhiew, C.; Rawangwong, S. A comparative study on the effect of oil palm fiber contents and types on properties of rubberwood sawdust polypropylene composites. *RMUTSV Res. J.* **2022**, *14* (22), 31–46. (in Thai)

[7] Hamamoto, N. *Manufacturing Processes*; CED Technology: Bangkok, 2016.

[8] Ratnabali, T.; Pitsuwan, P.; Chirasampathaa, P.; Homkiew, C. The Influence of Cutting Parameters on the Surface Finish of Wood-Plastic Composites. *Ladkrabang Eng. J.* **2015**, *32*(2), 43–48. (in Thai)

[9] Ratanawilai, T.; Jeenapong, S.; Srivabut, S. Suitable Condition for Face Milling of Wood-Plastic Composite Products on Surface Roughness. *Ladkrabang Eng. J.* **2023**, *40*(1), 153–163. (in Thai)

[10] Srivabut, C.; Rawangwong, S.; Homkhiew, S.; Rodjananugoon, J. Optimal Condition on Surface Roughness in Side Milling of High Density Polyethylene and Rubberwood Flour Composites using Response Surface Methodology. *Ladkrabang Eng. J.* **2022**, *39*(1), 23–34. (in Thai)

[11] Srivabut, C.; Rawangwong, S.; Homkhiew, C. Influence of Milling Parameters on Surface Roughness of Wood-Plastic Composites Applying I-Optimal Experimental Design. *Ladkrabang Eng. J.* **2022**, *39*(4), 22–35. (in Thai)

[12] Phetsuwan, W.; Thongkaew, K. A Study of Slot Milling Factors Affect the Machined Surface Quality of Wood-Plastic Composite. *Recent Sci. Technol.* **2025**, *17*(2). (in Thai)

[13] Srivabut, C.; Ratanawilai, T.; Hiziroglie, S. Response surface optimization and statistical analysis of composites made from calcium carbonate filler-added recycled polypropylene and rubberwood fiber. *J. Thermoplast. Compos. Mater.* **2019**, *35*(3), 1–25. https://doi.org/10.1177/0892705719889988

[14] Yabo, Z.; Qingshun, B.; Yangyang, S.; Donghai, L. Burr formation mechanism and machining parameter effect in slot micro-milling titanium alloy Ti6Al4V. *Int. J. Adv. Manuf. Technol.* **2022**, *123*(5), 2073–2086. https://doi.org/10.1007/s00170-022-10298-w

[15] Groover, M. P. Theory of Metal Machining. In *Fundamentals of Modern Manufacturing: Materials, Processes, and Systems*, 4th ed.; Wiley & Sons, Inc: United States of America, **2010**; pp 483–507

[16] Zhu, Y.; Buck, D.; Guan, J.; Song, M.; Tang, Q.; Guo, X.; Zhu, Z. Effects of Milling Methods on Cutting Performance of Wood–Plastic Composites Based on Principal Component Analysis. *Forests* **2024**, *15*(1516), 1–12. https://doi.org/10.3390/f15091516

[17] ISO. ISO 21920-1:2021, Geometrical product specifications (GPS) -Surface texture: Profile-Part 1: Indication of surface texture. https://cdn.standards.iteh.ai/samples/72196/5012ce84af1a4629a79764a48401a1db/ISO-21920-1-2021.pdf (accessed 2025-08-12).

[18] Akkoyun, F.; Cevik, Z. A.; Ozsoy, K.; Ercetin, A.; Arpaci, I. Image Processing Approach to Investigate the Correlation between Machining Parameters and Burr Formation in Micro-Milling Processes of Selective-Laser-Melted AISI 316L. *Micromachines* **2023**, *14*(1376), 1–16. https://doi.org/10.3390/mi14071376

ASEAN
Journal of Scientific and Technological Reports
Online ISSN:2773-8752

Research article

Root Length and Diameter at Flag Leaf Stage Correlate with Important Yield Parameters in Corn (*Zea mays* L.) Grown in Alkaline Soil Under Drought Conditions

Noriel Jay A. Magsayo[1*], Elvira D. Jamio[2], Melissa I. Canunayon[3], Daniel B. Tangpos[4], Julius D. Caritan[5], and Pet Roey L. Pascual[6]

[1] Graduate School, Cebu Technological University – Barili Campus, Barili, 6306 Cebu, Philippines
[2] Graduate School, Cebu Technological University – Barili Campus, Barili, 6306 Cebu, Philippines
[3] Graduate School, Cebu Technological University – Barili Campus, Barili, 6306 Cebu, Philippines
[4] Graduate School, Cebu Technological University – Barili Campus, Barili, 6306 Cebu, Philippines
[5] Graduate School, Cebu Technological University – Barili Campus, Barili, 6306 Cebu, Philippines
[6] Food Science, Agribusiness, and Development Communication, Cebu Technological University – Barili Campus, Barili, 6306 Cebu, Philippines
* Correspondence: norieljaymagsayo@gmail.com

Citation:
Magsayo, N.J.; Jamio, E.; Canunayon, M.; Tangpos, D.; Caritan, J.; Pascual, P.R. Root length and diameter at flag leaf stage correlate with important yield parameters in corn (*Zea mays* L.) grown in alkanline soil under drought conditions. *ASEAN J. Sci. Tech. Report.* **2025**, *28*(5), e259419. https://doi.org/10.55164/ajstr.v28i5.259643.

Article history:
Received: May 19, 2025
Revised: August 26, 2025
Accepted: September 1, 2025
Available online: September 14, 2025

Abstract: Root growth and development are critical in determining corn yield, with nutrient and energy demands at a higher level during flag leaf stage. To simulate common abiotic stresses, an experiment was conducted under alkaline and drought conditions, which often limit root and shoot growth, nutrient uptake, and productivity. A randomized complete block design was used to evaluate the relationship between root length, root diameter, shoot growth, and yield components at the flag leaf stage. The results showed that grain yield had no relationship with shoot development but showed a strong positive correlation with nodal root length and diameter. Ear height (EH), ear weight (EW), unshelled weight (UW), shelled weight (SW), ear length (ER), and computed yield (CY) are strongly correlated with root length at $r = 0.741$, $r = 0.578$, $r = 0.591$, $r = 0.869$, and $r = 0.874$, respectively. Moreover, ear height, ear weight, unshelled weight, shelled weight, ear length, and computed yield are strongly correlated positively with the root length at $r = 0.725$, $r = 0.831$, and $r = 0.822$, $r = 0.625$, $r = 0.408$, and $r = 0.622$, respectively. These results indicate that nodal root length and thickness at the flag leaf stage are major determinants of eventual corn yield. The study emphasizes the importance of improved management practices to enhance root and shoot development during the flag leaf stage. It further suggests that breeding programs should focus on nodal root length and diameter, particularly under alkaline soil and drought conditions.

Keywords: Root length; root diameter; stress; correlation; nodal root

1. Introduction

Corn is one of the most widely grown crops worldwide. Drought stress is one of the environmental factors that highly affects its productivity [1]. Similarly, for drought-tolerant corn breeding efforts in areas with limited water on alkaline soils, the link between root features and yield-related traits is also crucial. Several research articles have investigated the effects of drought on a range of plant physiological processes, such as root development [2]. The experiment was conducted in Barili, Cebu, where the soil is naturally alkaline and recent climatic conditions have resulted in prolonged drought periods. Measuring corn under these combined stress factors was necessary to simulate the actual field environment that local farmers face. Improving corn yields is a formidable challenge facing local farmers in Cebu, Philippines, which is characterized by alkaline soils and frequent droughts. Although geneticists and breeders have made significant advances in creating root traits that improve productivity in drought conditions, a deeper understanding of the correlation between whole-plant traits and root functional traits is necessary to enhance crop performance in these environments further [3]. Some root traits have been shown to enhance the potential for productivity among plants under drought stress, such as those that involve the formation of highly dense, densely branched root structures that effectively capture moisture supplies deep in soils [2,3]. Notably, studies have shown that improved drought tolerance is often correlated with increased root length, characterized by more branching with smaller diameters, higher specific root lengths, and higher root densities in various crop species, particularly in the deeper parts of the soil profile [3,4]. It is the most important factor allowing root access to subsurface water reserves that may persist into the critical reproductive and grain-filling stages. It has been identified as a key factor for plants to maintain optimal metabolic processes under conditions characterized by late-season water shortages.

Interestingly, previous studies suggest that water stress may not significantly impact the development of corn roots. In some instances, drought stress can even promote better root growth and extension [2, 5]. Physiologically, drought-tolerant genotypes often exhibit enhanced resilience by accumulating osmoregulatory molecules, such as proline and amino acids, which reduce cellular osmotic potential and maintain water absorption and cell turgor pressure under stress [6]. Improving water-use efficiency is critical in arid and semi-arid regions, where water scarcity limits crop yields. Therefore, drought tolerance should be a target for improvement under certain changing environmental conditions. This study aims to investigate the relationship between root length, root diameter, and shoot development at the flag leaf stage and important yield parameters in corn grown in alkaline soil under drought conditions, highlighting the key root traits that can be used to improve production and enhance corn's ability to withstand these stress conditions.

2. Materials and Methods

2.1 Preparation of Experimental Area

The research was conducted at the experimental and production farm of Cebu Technological University - Barili Campus (10°7'53" N, 123°32'45" E), covering a study area of 100 square meters. In terms of chemical characteristics, the soils have an alkaline pH of more than 7.5 [7]. The experimental area was plowed twice to improve soil aeration and drainage. The study utilized five top-performing varieties in alkaline soil, each replicated four times, resulting in a total of 20 experimental blocks. The plots were divided into furrows spaced 75 cm apart. Each treatment plot measured 5 meters in length, with a one-meter space allocated between blocks to facilitate data collection and regular inspection.

2.2 Climatic Conditions

The research was conducted from February to June 2024. Cebu, Philippines, was warm and humid during the period. Temperatures in March ranged from 25°C to 31°C, accompanied by moderate rainfall of approximately 103 mm over 8 days. In April, temperatures were slightly higher at 32°C, but rainfall was lower at approximately 81 mm over 7 days. May recorded the highest temperature at 33°C, but had more rainfall, with intervals of 127-147 mm within 8 to 9 days [8]. This study, however, coincides with a dry season due to the El Niño phenomenon [9]. Below-average precipitation has been documented, especially in April, when drought conditions prevailed [10].

2.3 Seed Preparation, Field Planting, and Experimental Design

Corn seeds were soaked in coconut water for twenty-four (24) hours to improve the germination rate and sown directly with two seeds per hill. The spacing between rows was 75 cm for all treatments, with 25 cm spacing between planting hills. This planting arrangement ensured proper spacing for each crop, allowing for accurate comparison across treatments. The field experiment was laid out in a

Randomized Complete Block Design (RCBD) having five treatments, which were hybrid and conventional corn varieties, with ten (10) samples per treatment, replicated four times with a spacing of 0.25m between hills and 0.75m between rows. The following were the treatments: T1 = WPS327, T2 = TCT1868, T3 = SAPT84HY, T4 = M22B-01, and T5 = GSI130YR.

2.4 Fertilization

To ensure soil productivity, 7 bags of organic fertilizer were applied as a basal application across the entire field, resulting in a fertilization rate of 17.5 tons/ha. The fertilization application rate was also given. There was a split application of NPK fertilizers, in particular, nitrogen (46-0-0), phosphorus (16-20-0), and potassium (0-0-60). The first application of nitrogen was made at 15 days after planting, followed by the second application of the mixture of nitrogen and potassium at 30 days after planting, while the third application of the combination of the three fertilizers was carried out at 45 days after planting. These applications of fertilizers helped provide steady nutrients for the crop throughout its sensitive growth periods, resulting in higher corn yields.

2.5 Thinning and Weeding

To ensure adequate spacing of the corn plants, thinning to one seedling per hill was done 14 days after sowing for all corn varieties. Followed by uniform hand weeding across all rows in each replication to manage weed pressure. Additionally, hill up was carried out 14 days after crop emergence to increase drainage, proper root aeration, and suppress weed growth. These crop management practices optimized root growth development by providing better soil stability, allowing each corn plant to absorb necessary nutrients and sunlight.

2.6 Water and Pest Management

Considering the lack of rainfall due to drought conditions during the vegetative stage, watering was done three times a week to maintain the soil moisture needed by the corn plants. Fall armyworms (FAW) and Asiatic corn borers (ACB) were observed, and synthetic insecticides were used during severe infestation.

2.7 Data Gathered

2.7.1 Agronomic Characteristics of Corn

Plant height (cm) was measured from the base of the plant to the base of the tassel using a measuring tape. Days to silking were determined by counting the number of days from planting to silk emergence. Root length (cm) was measured by measuring the length of the nodal root of the corn plant at the flag leaf stage using a ruler. Root diameter (mm) was determined by measuring the width of the nodal root at the flag leaf stage using a caliper. Ear height (cm) was measured at the maturity stage, from the ground to the base of the ear. Average ear weight (kg) was recorded by dividing the total weight of sample ears with husks by the number of samples using a digital weighing scale. Average ear length (cm) was measured by dividing the total length of the sample ears without husks by the number of samples using a tape measure. Shelled weight (g) was determined by shelling and weighing sample ears by the number of samples per treatment using a digital weighing scale. The average weight of unshelled ears (g) was determined by weighing all sample ears without husks and dividing by the total sample ears per treatment. Computed yield (kg). The total yield with 15% moisture content was computed using this formula:

$$Grain\ Yield\ (t/ha) = \frac{SW(g)\ x\ 55{,}000\ x\ (100-MC)}{(100-18) x\ (1{,}000{,}000)}$$

Where:

SW = shelled weight (g)
55,000 = assumed plant population per hectare (constant)
MC = actual harvest moisture
15 = standard moisture basis used
1,000,000 = conversion factor from grams to metric tons

2.7.2 Chemical Composition of Corn

Total Soluble Solid (TSS). Twenty grams of kernels were blended and homogenized with 100 mL of distilled water to make corn juice. TSS in Brix was measured using a handheld refractometer (model HI-96801, Hanna Instruments Ltd., USA), calibrated with distilled water by placing 1-3 drops of juice on the instrument's prism and taking the reading. Titratable Acidity (%). Five milliliters (5 mL) of corn juice was taken and diluted with 45 mL of distilled water. The 50 mL extract (aliquot) was measured and transferred into an Erlenmeyer flask, to which two drops of 1% phenolphthalein indicator were added. This was then titrated with 0.1% NaOH until a pink color was achieved, using a Cordial 1642TF Glass burette, and the volume of NaOH was recorded. Titratable acidity was expressed as percent lactic acid equivalent, since lactic acid is the predominant organic acid in corn. TA was calculated using this formula.

$$TA\ (\%\ lactic\ acid) = \frac{V\ x\ N\ x\ 0.09}{Sample\ weight\ (g)}\ x\ 100$$

Where:

- V = volume of NaOH used in titration (mL)
- N = normality of NaOH (0.1 N)
- Milliequivalent weight of lactic acid = 0.09 (g/meq.)
- Sample weight (g) = weight/volume of extracted juice used in titration.

2.8 Data Analysis

This study performed a linear correlation analysis to evaluate the relationship between corn's root, shoot, and yield parameters. The Pearson correlation coefficient (r) was calculated to determine the strength and direction of the linear relationship between these variables using SPSS.

Table 1. Pearson correlation coefficient interpretation.

r	**Interpretation**
> 0.70	Very strong positive correlation
0.40 to 0.69	Strong positive correlation
0.30 to 0.39	Moderate positive correlation
0.20 to 0.29	Weak positive correlation
0.01 to 0.19	No relationship
-0.01 to -0.19	No relationship
-0.20 to -0.29	Weak negative correlation
-0.30 to -0.39	Moderate negative correlation
-0.40 to -0.69	Strong negative correlation
<-0.70	Very strong negative correlation

3. Results and Discussion

3.1 Relationship between Nodal root length at flag leaf stage and Yield components of corn

The relationship between the nodal root length during the flag leaf stage and various yield components of corn was assessed to understand the influence of root architecture on corn performance under drought and alkaline conditions. The primary functional roots during the flag leaf stage are nodal roots, which are known to be essential for nutrition uptake, hydraulic support, and soil exploration [11]. The analysis revealed varying degrees of correlation between nodal root length, days to silking, ear height, ear weight, unshelled and shelled weight, ear length, total soluble solids (TSS), Titratable acidity (TA), and computed yield per hectare. Figure 1b illustrates a strong negative correlation between nodal root length and days to silking ($r = -0.505$). This suggests that an increase in nodal root length may be associated with marginally earlier reproductive development, potentially due to enhanced early-stage water absorption. This result is consistent with Chaudhary [12], who reported that corn genotypes with deeper rooting systems exhibited shorter flowering times under saline stress. Similarly, Ju [13] demonstrated genetic linkages between root traits and phenology, indicating that enhanced root length contributes to reduced anthesis–silking interval (ASI), a well-known indicator of stress tolerance. Conversely, Figure 1g illustrates a moderate positive correlation between nodal root length and total soluble solids (TSS) with a correlation coefficient of $r = 0.392$. This suggests that corn with longer roots tends to have a higher soluble sugar content in its kernels. Longer root systems likely enhance nutrient and water uptake, which sustains photosynthetic activity and carbohydrate allocation to developing ears. This explains the positive but moderate relationship observed here. Similar findings were reported by Li [14], who showed that root length density correlated positively with sugar accumulation and yield under mulching conditions. However, compared with stronger correlations between nodal root length and physical traits such as ear height and grain weight, the relationship with TSS is weaker. This is consistent with the notion that sugar metabolism is influenced by multiple physiological pathways, not solely by root architecture. Wang [15] found that while crown root traits in corn influenced carbohydrate allocation, kernel sugar content was strongly affected by genotypic differences in starch biosynthesis pathways.

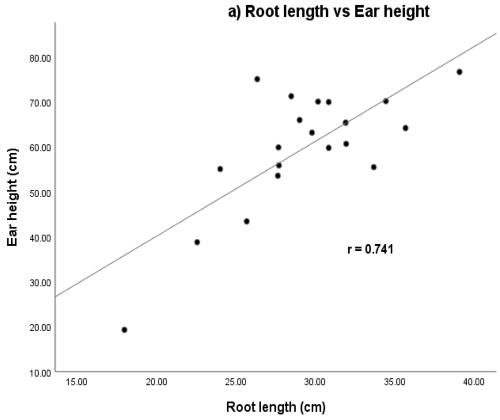
On the other hand, Figure 1h shows a weak positive correlation between nodal root length and titratable acidity (TA) with a correlation coefficient of $r = 0.203$. Suggesting that the relationship between root system development and kernel acidity is minor. Unlike yield-related traits, which strongly depend on root capacity for water and nutrient uptake, organic acid levels are more tightly regulated by genotype-specific biochemical processes and environmental factors such as light intensity, temperature, and post-harvest respiration [16]. A similar conclusion was drawn by Silva [17], who found that root traits and biochemical grain quality traits were weak compared to those with yield attributes. Notably, figures 1c, 1d, and 1f show a strong positive association. Ear weight, unshelled weight, and ear length are strongly correlated with the nodal root length at $r = 0.578$, $r = 0.591$, and $r = 0.598$, respectively. These relationships indicate that as the root system becomes longer, corn plants tend to produce larger and heavier ears. Longer root systems improved nutrient uptake and enhanced water availability, all of which contribute to kernel initiation and filling. A longer ear indicates improved sink capacity, while higher ear and unshelled weights suggest greater biomass partitioning into reproductive structures. These findings are consistent with the work of Silva and Nzuve [17,18] who showed that root vigor contributed significantly to ear development and grain filling under field and low-input conditions. Additionally, Figures 1a, 1e, and 1i revealed a very strong positive correlation between nodal root length and three key agronomic traits of corn: ear height (r = 0.741), shelled weight (r = 0.869), and computed yield (r = 0.874). These results indicate a robust association between underground and aboveground growth, where deeper root systems strongly support higher ear development and eventual yield performance. Higher ear placement tends to exhibit greater vigor and biomass accumulation, longer roots provide greater access to deep soil nitrogen and water, which are critical for maximizing kernel set and grain weight. Root system architecture plays a significant role in drought avoidance, nutrient acquisition, and photosynthate partitioning, all of which translate into higher yields [12, 17]. Ju [13] demonstrated that QTLs for root length often co-localized with yield-related QTLs, providing genomic evidence for these strong relationships. However, as ear height is also strongly correlated with nodal root length, care must be taken to avoid excessively tall plants that may increase the risk of lodging under high-density planting conditions.

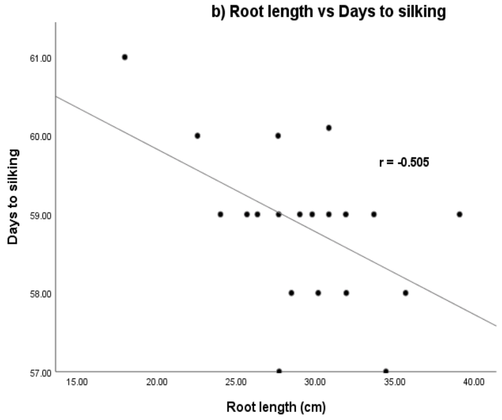
3.2 Relationship between Nodal root diameter at flag leaf stage and yield components of corn

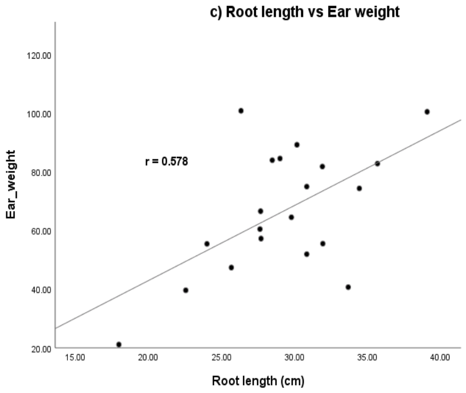
The correlations observed between nodal root diameter and key phenological and yield performance of corn reveal the significance of root morphology in regulating corn performance in alkaline and drought stress conditions. Figure 2g shows no relationship between nodal root diameter and total soluble solids. Nodal root diameter is primarily a structural trait that provides anchorage and lodging resistance, but it does not significantly improve nutrient acquisition [19]. Additionally, a strong negative correlation was observed in Figure 2 B between nodal root length and days to silking, with a correlation coefficient of $r = -0.493$. This indicates that corn plants with thicker nodal roots tend to silk earlier than those with thinner nodal roots. Thicker nodal roots improve a plant’s hydraulic conductivity, enabling more efficient uptake and transport of water and nutrients [20]. Early silking is particularly advantageous under drought, as it reduces the anthesis-silking interval (ASI) and improves the likelihood of successful fertilization [21]. On the other hand, a moderate positive correlation was found between nodal root diameter and titratable acidity at Figure 2h with $r = 0.363$. This indicates that corn plants with thicker nodal roots tend to have kernels with slightly higher organic acid content. Thicker nodal roots with higher nutrient transport capacity, particularly for nitrogen and minerals such as potassium and magnesium [20]. These nutrients are closely involved in organic acid metabolism within kernels, including the citric acid and lactic acid pathways that determine TA [16]. By contrast, correlation of titratable acidity (TA) with nodal root diameter is generally weaker than that with yield-related parameters, as TA is more strongly determined by genotype-specific biochemical pathways and post-harvest physiological processes [17]. The moderate relationship observed suggests that nodal root diameter plays only an indirect role in modulating grain acidity. Additionally, figures 2e, 2f, and 2i showed a strong positive correlation between shelled weight, ear length, and computed yield with nodal root diameter at $r = 0.625$, $r = 0.408$, and $r = 0.622$, respectively. Suggesting that corn plants with thicker nodal roots tend to develop heavier shelled grain, longer ears, and eventual yield. This is due to the improved water and nutrient transport efficiency in thicker nodal roots, which supports ear elongation and kernel set [13, 20]. These strong correlations highlight nodal root diameter as a valuable indirect selection of traits for yield improvement. Moreover, figures 2a, 2c, and 2d showed a very strong positive relationship between nodal root diameter and ear height ($r = 0.725$), ear weight ($r = 0.831$), and unshelled weight ($r = 0.822$). These results indicate that plants with thicker nodal roots consistently produced taller shoots and larger, heavier ears. Similar findings were reported by Silva[17], who found that root structural traits made strong contributions to ear development and yield. The strong correlations confirm that nodal root diameter is a valuable indirect selection trait for yield improvement. In breeding programs, selecting for thicker nodal roots could simultaneously enhance plant vigor, ear size, and yield stability, particularly under stress conditions where efficient resource use and strong anchorage are essential.

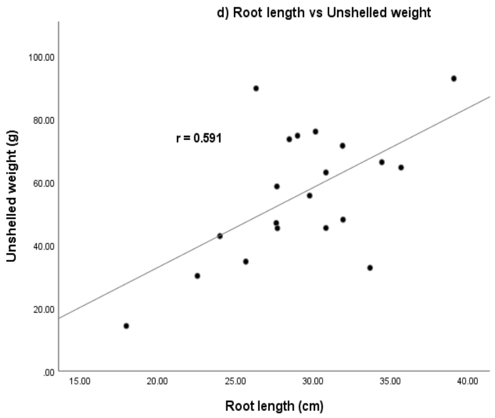
3.3 Relationship between Plant height at flag leaf stage and yield components of corn

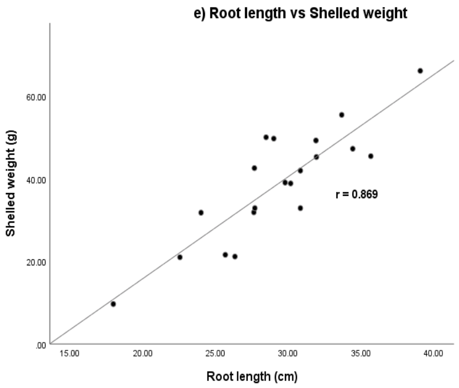
Plant height is a crucial agronomic trait in corn, often associated with vigor and yield potential. Taller plants generally capture more light and produce larger ears, although excessive height can increase the risk of lodging. Figures 3b and 3g showed weak negative correlations between plant height and days to silking (r =-0.298) and total soluble solids (TSS) (r =- 0.258), respectively. Indicating that taller corn plants tend to flower slightly earlier and may allocate more assimilates to vegetative biomass rather than soluble sugar accumulation in kernels. This is consistent with earlier reports that kernel quality traits are less strongly associated with corn plant stature than yield-related parameters [17]. Early silking in taller plants may reflect more efficient growth and faster transition to reproductive development [21]. Figures 3e, 3f, and 3i showed that plant height has no relationship with shelled weight, ear length, and computed yield. This suggests that, unlike root traits, plant height is not a reliable predictor of ear size or grain yield, as kernel development and yield are more strongly controlled by root efficiency, source-sink balance, and genetic regulation of reproductive traits [17,18]. Additionally, a weak positive correlation was observed in Figure 3h between plant height and titratable acidity (TA). This relationship suggests that vegetative vigor may have a minor influence on organic acid accumulation, potentially due to increased photosynthetic activity and metabolic flux in taller plants. Similar findings have been reported in tomato by Kumar [22], where plant height was weakly associated with fruit quality traits, including titratable acidity. On the other hand, strong positive correlations were observed in Figures 3a, 3c, and 3d between plant height and key yield-related traits in corn, including ear height ($r = 0.586$), ear weight ($r = 0.692$), and unshelled weight ($r = 0.687$). These findings suggest that taller plants tend to exhibit greater ear development and higher biomass accumulation, supporting their positive contribution to overall yield performance. Taller plants generally exhibit greater vegetative vigor, with increased leaf area and photosynthetic capacity that enhance the supply of assimilates to reproductive organs. This results in higher ear placement and a larger ear mass. Similar findings were reported by Samskriti [23], who found that plant height was strongly associated with ear height and grain yield in promising hybrids. Mallhi [24] likewise observed that plant height enhanced ear height and ear size. Under drought stress, Tasnim [25] confirmed that plant height positively correlated with ear length, indicating its stability as a vigor trait across environments. Nutrient management further reinforces this relationship. Liu [26] demonstrated that plant and ear height were strongly correlated with yield under organic and inorganic fertilizer regimes. Similarly, Ayad [27] reported significant positive associations between plant height and ear weight across environments, showing that shoot vigor and reproductive sink capacity are closely linked. These results confirm that plant height can serve as an indirect indicator of ear development, particularly in genotypes where increased vegetative vigor translates into larger ears and greater ear size and weight.

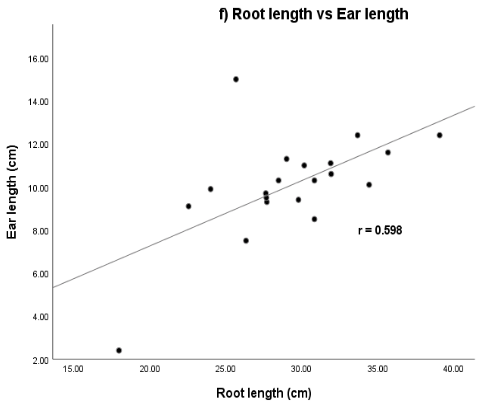

a) Root length vs Ear height

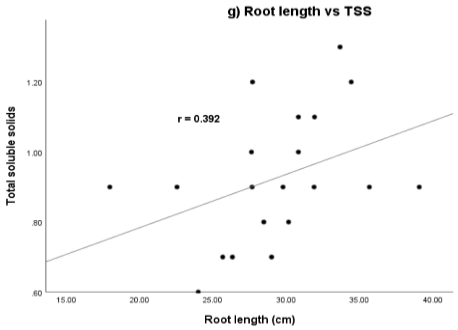

b) Root length vs Days to silking

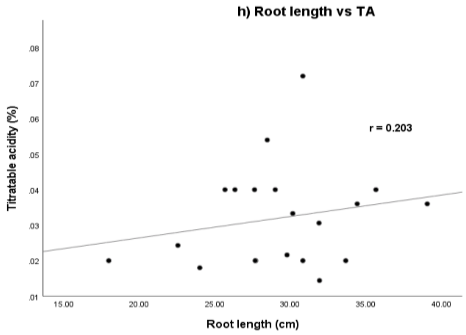

c) Root length vs Ear weight


d) Root length vs Unshelled weight


e) Root length vs Shelled weight


f) Root length vs Ear length


g) Root length vs TSS


h) Root length vs TA

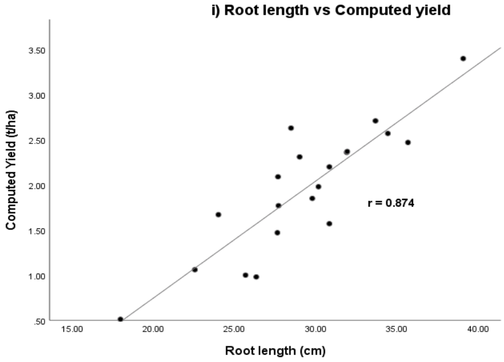
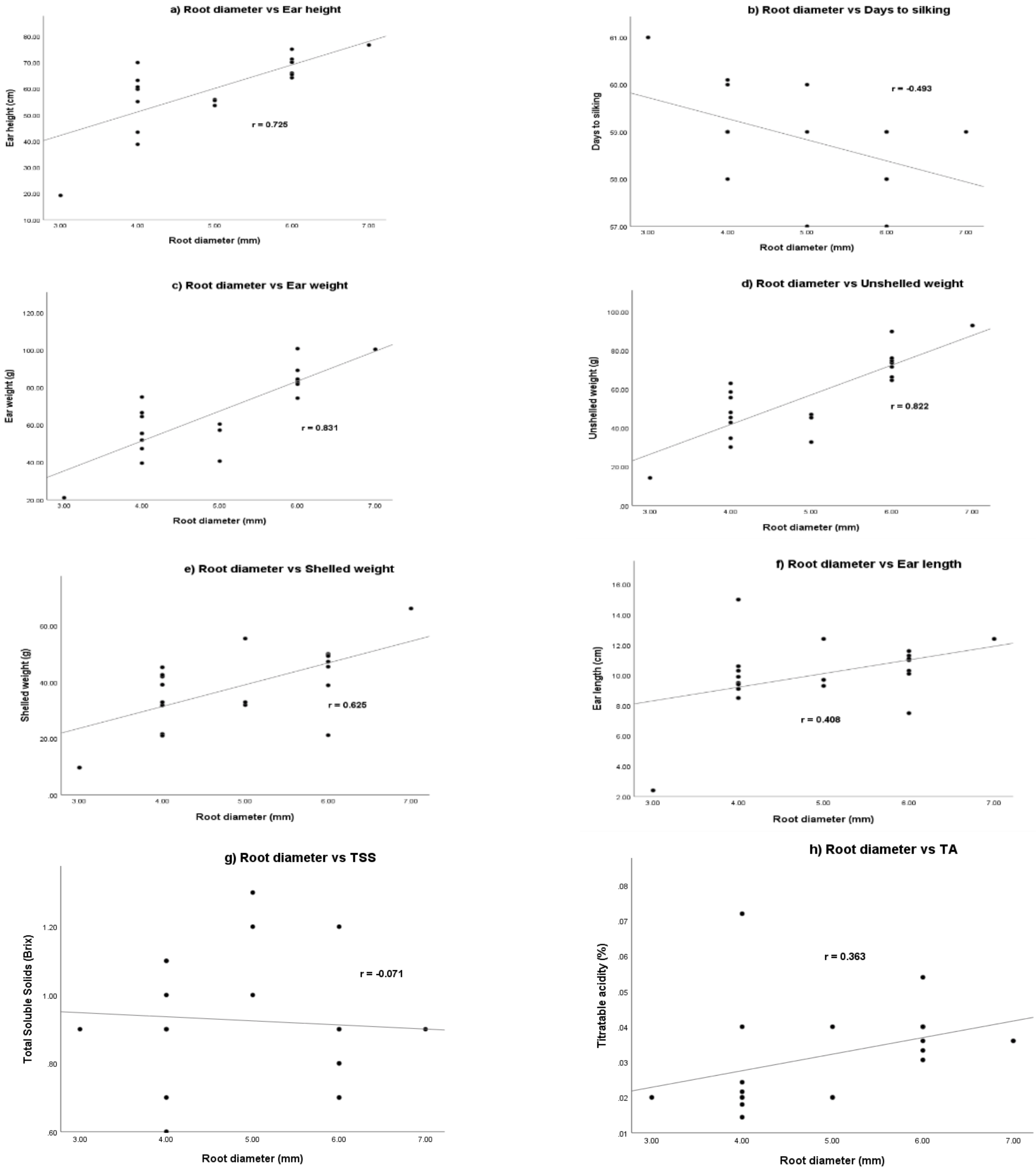


Figure 1. Relationship between Nodal root length and Yield parameters of corn



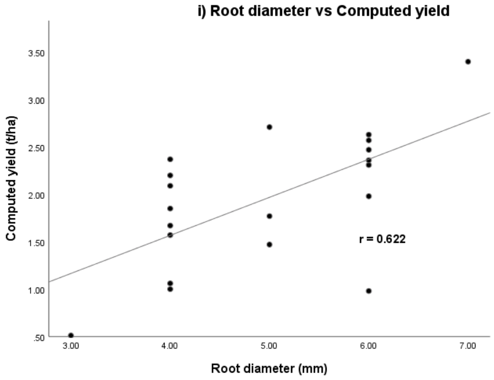
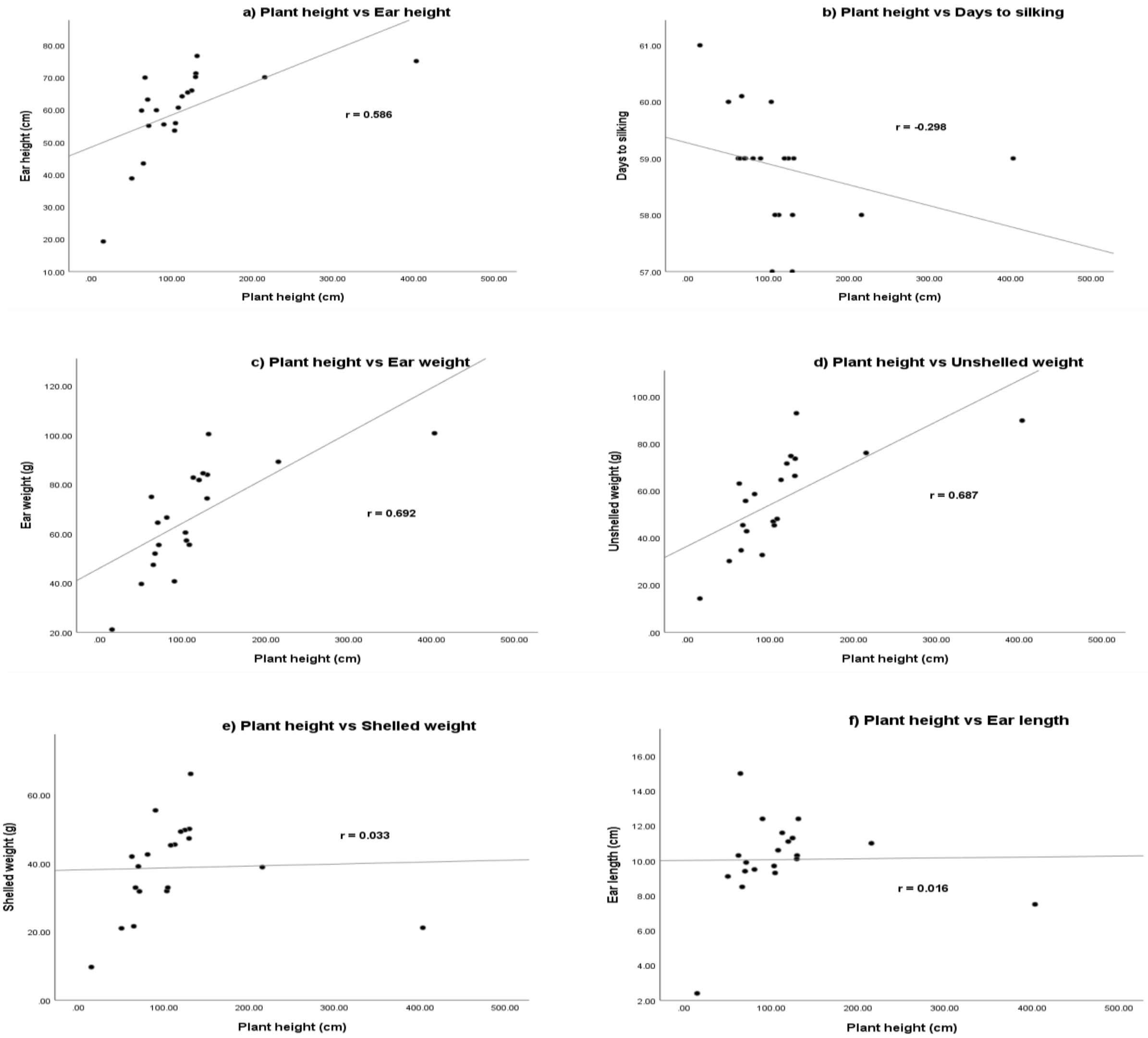


Figure 2. Relationship between Nodal root diameter and Yield parameters of corn



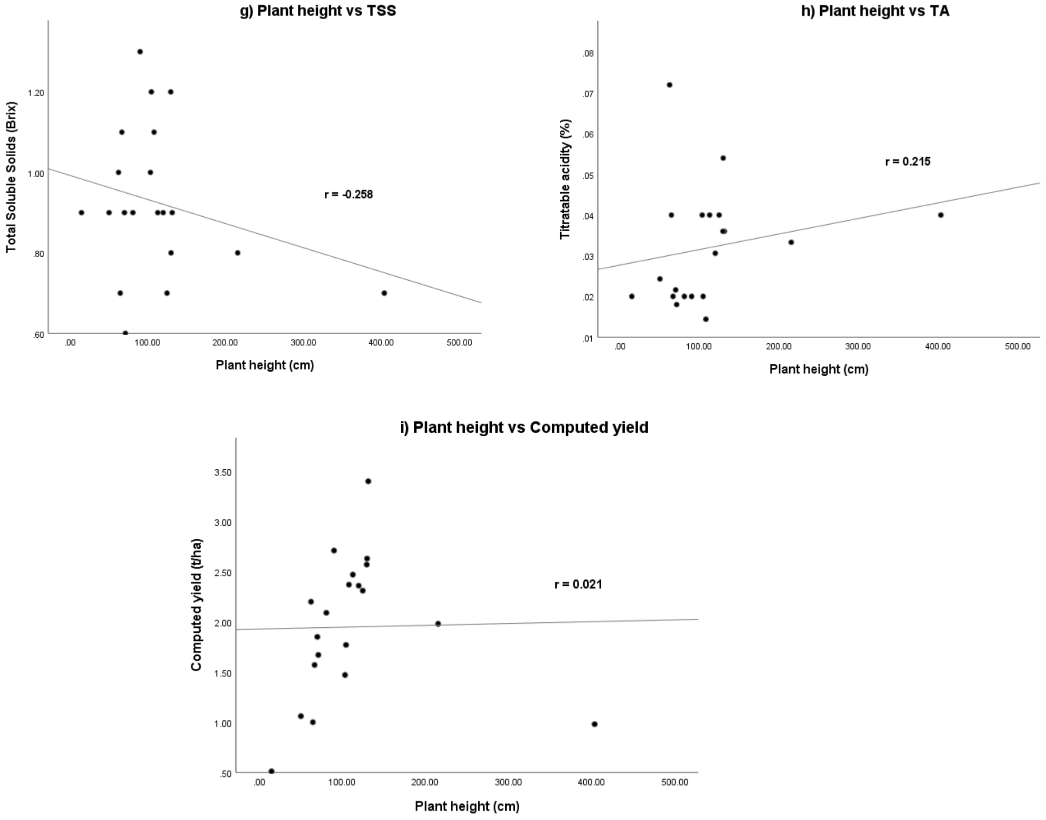


Figure 3. Relationship between plant height and yield and yield components

4. Conclusions

These collective findings highlight the crucial role of root traits and shoot development, specifically nodal root length and diameter at the flag leaf stage, in influencing various yield-related parameters in corn plants grown under alkaline and drought conditions. The correlations observed in the study, ranging from moderate to strong, highlight the complex interplay between root morphology, plant development, and yield potential. However, while beneficial for productivity, excessive height may increase susceptibility to lodging, suggesting that breeders should aim for balanced plant structure, combining moderate height with strong stalks and robust root systems

5. Acknowledgements

The authors would like to acknowledge Cebu Technological University, Barili Campus, and the DOST-Strand N for their support in making this study a success.

Author Contributions: Conceptualization, P.R.P.; methodology, N.J.M., D.T., and J.C.; software, N.J.M.; validation, E.J., M.C.; formal analysis, N.J.M.; investigation, N.J.M., E.J., M.C., D.T., and J.C.; resources, N.J.M., E.J., M.C., D.T., and J.C..; data curation, N.J.M.; writing—original draft preparation, N.J.M.; writing—review and editing, N.J.M., P.R.P; visualization, N.J.M.; supervision, P.R.P.; project administration, N.J.M., E.J., M.C., D.T., and J.C.; funding acquisition, N.J.M., E.J., M.C., D.T., and, J.C. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: This research was made possible through the support provided by the Department of Science and Technology (DOST) STRAND N program and Cebu Technological University - Barili Campus. The authors declare no conflicts of interest, as the funding agency did not influence the study's design, methodology, data analysis, or conclusions.

References

[1] Abraha, T.; Nyende, A. B.; Githiri, S.; Kasili, R.; Araia, W.; Shaibu, A.; Rabiu, I.; Adnan, A. A. Variability of Root and Physiological Traits of Different Maturity Groups of Maize (Zea mays L.). *J. Plant Breed. Crop Sci.* **2015**, *7*(9), 232–238. https://doi.org/10.5897/JPBCS2015.0517

[2] Wahab, A.; Abdi, G.; Saleem, M. H.; Ali, B.; Ullah, S.; Shah, W.; Mumtaz, S.; Yasin, G.; Muresan, C. C.; Marc, R. A. Plants' Physio-Biochemical and Phyto-Hormonal Responses to Alleviate the Adverse Effects of Drought Stress: A comprehensive Review. *Plants* **2022**, *11*, 1620. https://doi.org/10.3390/plants11131620

[3] Comas, L. H.; Becker, S. R.; Cruz, V. M. V.; Byrne, P. F.; Dierig, D. A. Root traits contributing to plant productivity under drought. *Front Plant Sci.* **2013**, *4*, 442. https://doi.org/10.3389/fpls.2013.00442

[4] Wajhat-Un-Nisa; Sandhu, S.; Ranjan, R.; *et al.* Root plasticity: an effective selection technique for identification of drought tolerant maize (Zea mays L.) inbred lines. *Sci. Rep.* **2023**, *13*, 5501. https://doi.org/10.1038/s41598-023-31523-w

[5] Adee, E.; Roozeboom, K.; Balboa, G. R.; Schlegel, A.; Ciampitti, I. A. Drought-Tolerant Corn Hybrids Yield More in Drought-Stressed Environments with No Penalty in Non-Stressed Environments. *Front. Plant Sci.* **2016**, *7*, 1534. https://doi.org/10.3389/fpls.2016.01534

[6] Rosa, V.; Silva, A.; Brito, D.; Júnior, J.; Silva, C.; Costa, M.; Oliveira, J.; Ribeiro, C. Drought Stress during the Reproductive Stage of Two Soybean Lines. *Pesqui. Agropecu. Bras.* **2020**, *55*, e01736. https://doi.org/10.1590/S1678-3921.pab2020.v55.01736

[7] Enojada, G. R.; Asio, V. B. Morphological, physical, and chemical characteristics of soils derived from limestone rocks in Barili, CEBU. *J. Trop. Technol.* **2018**, *18*(2).

[8] The Weather Channel. https://weather.com/weather/tenday/l/25414ec6b3178100e232564a4347cadead55c8e0758dce9d74ea69f0b1228ba (accessed 2025-09-13).

[9] Philippine Atmospheric, Geophysical and Astronomical Services Administration. https://www.pagasa.dost.gov.ph/weather/weather-outlook-selected-philippine-cities (accessed 2025-09-13).

[10] Eram, M. M. Cebu virtually had no rains in April. *Cebu Daily News*, May 1, 2024. https://cebudailynews.inquirer.net/571007/cebu-virtually-had-no-rains-in-april (accessed 2025-09-13).

[11] Gao, Y.; Lynch, J. P. Reduced Crown Root Number Improves Water Acquisition under Water Deficit Stress in Maize (Zea mays L.). *J. Exp. Bot.* **2016**, *67*(15), 4545–4557. https://doi.org/10.1093/jxb/erw243

[12] Chaudhary, W.; Ali, M.; Bajwa, K.; Iqbal, A.; Khan, M.; Shahid, A.; Aslam, M.; Kamran, S.; Bajwa, A.; Iqbal, M.; Azmat; Khan, A.; Shahid, M. Correlation Analysis of Maize Genotypes under Saline Stress and Its Impact on Morphological Characteristics. *Life Sci. J.* **2017**, *14*(7), 84–89. https://doi.org/10.7537/marslsj140717.15

[13] Ju, C.; Zhang, W.; Liu, Y.; Li, J.; Wang, C.; Yang, X.; Yan, J. Genetic Analysis of Seedling Root Traits Reveals the Association of Root Trait with Other Agronomic Traits in Maize. *BMC Plant Biol.* **2018**, *18*, 171. https://doi.org/10.1186/s12870-018-1383-5

[14] Li, Y.; Yang, J.; Shi, Z.; Pan, W.; Liao, Y.; Li, T.; Qin, X. Response of Root Traits to Plastic Film Mulch and Its Effects on Yield. *Soil Tillage Res.* **2021**, *209*, 104930. https://doi.org/10.1016/j.still.2020.104930

[15] Wang, H.; Tang, X.; Yang, X.; Xu, C.; Kang, L.; Tong, H.; Zhang, Q.; Xiao, Y. Exploiting Natural Variation in Crown Root Traits via Genome-Wide Association Studies in Maize. *BMC Plant Biol.* **2021**, *21*, 346. https://doi.org/10.1186/s12870-021-03127-x

[16] Li, R.; Li, M.; Ashraf, U.; Liu, S.; Zhang, J. Exploring the Relationships Between Yield and Yield-Related Traits for Rice Varieties Released in China From 1978 to 2017. *Front. Plant Sci.* **2019**, *10*, 543. https://doi.org/10.3389/fpls.2019.00543

[17] Silva, T. N.; Moro, G. V.; Moro, F. V.; Santos, D. M. M. Correlation and Path Analysis of Agronomic and Morphological Traits in Maize. *Rev. Ciênc. Agron.* **2016**, *47*(2), 351–357. https://doi.org/10.5935/1806-6690.20160041

[18] Nzuve, F.; Githiri, S.; Mukunya, D. M.; Gethi, J. Genetic Variability and Correlation Studies of Grain Yield and Related Agronomic Traits in Maize. *J. Agric. Sci.* **2014**, *6*(9), 166. https://doi.org/10.5539/jas.v6n9p166

ASEAN
Journal of Scientific and Technological Reports
Online ISSN:2773-8752

Research article

Ecological Factors Governing the Persistence of Khlong-Saeng Durian: Implications for Conservation

Surachai Sang-Ngam[1], Wirot Chaowiset[2], Sudarat Theerapisit[3, 5], Suppawan Promprao[4], and Siriluk Sintupachee[1, 5*]

[1] Faculty of Science and Technology, Nakhon Si Thammarat Rajabhat University, Nakhon Si Thammarat, 80280, Thailand
[2] Faculty of Science and Technology, Suratthani Rajabhat University, Suratthani, 84100, Thailand
[3] Faculty of Education, Nakhon Si Thammarat Rajabhat University, Nakhon Si Thammarat, 80280, Thailand
[4] School of Science, Walailak University, Thaiburi, Thasala District, Nakhon Si Thammarat, 80160, Thailand
[5] Specialized Research Unit for Insects and Herbs, Nakhon Si Thammarat Rajabhat University, Nakhon Si Thammarat, 80280, Thailand
* Corresponding author: siriluk_sint@nstru.ac.th

Citation:
Sang-Niam, S.; Chaowiset, w.; Theerapisit, S.; Promprao, S.; Sintupachee, S. Ecological factors governing the persistence of khlong-saeng durina: implications for conservation. *ASEAN J. Sci. Tech. Report.* **2025** *28*(5), e259593. https://doi.org/10.55164/ajstr.v28i5.259593.

Article history:
Received: May 11, 2028
Revised: August 13, 2025
Accepted: September 1, 2025
Available online: September 14, 2025

Abstract: The Khlong-Saeng durian (*Durio* spp.) of Surat Thani is undergoing a decline, leading to an investigation of environmental factors affecting its longevity. This study examined the relationship between soil properties and the age of native durian trees in Ban Khao Thep Phithak. Using geographic indices, correlation analysis, and preliminary phytochemical screening of peel waste, we identified factors influencing tree lifespan. Durian trees occurred in areas with temperatures of 26–31°C and relative humidity between 58.87% and 87.37%. A strong positive correlation was found between tree age and the dependent variable ($r = 0.75$, $p < 0.05$). All age groups were associated with silty soils, with significant variation in texture ($p < 0.001$, $F(10, 22) = 42.96$). Soil chemistry also differed significantly with tree age. Key parameters included soil pH ($p < 0.001$), organic matter ($p < 0.001$), nitrogen ($p < 0.001$), potassium ($p < 0.001$), phosphorus ($p < 0.001$), calcium ($p < 0.001$), and sodium ($p < 0.001$). Phytochemical screening showed antioxidant activity in the peels of trees over 50 years old, suggesting an age-related accumulation of bioactive compounds. Together, these results provide the first evidence linking soil nutrients and phytochemical markers to the longevity of Khlong-Saeng durian. The findings highlight the importance of maintaining optimal soil conditions and conserving high-elevation habitats to safeguard this unique genetic resource. Such knowledge is essential for sustainable cultivation strategies and long-term conservation of native durian in Southern Thailand.

Keywords: *Durio* spp.; environmental factors; native durian conservation; phytochemical screening; sustainable utilization

1. Introduction

Southern Thailand, a region characterized by its significant biodiversity and distinct tropical ecosystems, serves as a vital reservoir for unique plant genetic resources. Within this ecologically important area, Surat Thani Province is notable for being the native habitat of several endemic fruit varieties, among which the Khlong-Saeng durian (*Durio* sp.) holds a special place [1]. This specific local landrace is highly regarded by local communities for its exceptionally fragrant aroma and uniquely delicious flavor profile, representing a valuable component of the region's agricultural heritage and biodiversity [1]. Its distinctiveness underscores the importance of understanding and preserving such locally adapted plant varieties [2]. However, the continued existence of the Khlong-Saeng durian is facing considerable pressure. Historical land management practices, including extensive timber harvesting, have altered the landscape. More recently, the expansion of commercial monocultures, particularly oil palm and rubber plantations, has led to significant habitat fragmentation and loss for native species. Compounded by potentially insufficient public awareness and promotion ("poor public relations"), these pressures have resulted in a noticeable decline in the Khlong-Saeng durian population. This decline signifies not only the potential loss of a unique fruit but also contributes to genetic erosion, diminishing the overall biodiversity and resilience of the local ecosystem [3].

Conserving locally adapted landraces, such as the Khlong-Saeng durian, is of paramount importance. These varieties often possess unique genetic traits conferring resilience to local pests, diseases, and environmental conditions, representing an invaluable resource for future crop improvement and adaptation strategies, especially in the face of climate change [4]. Furthermore, safeguarding such varieties contributes to maintaining the rich tapestry of global biodiversity and preserves elements of local cultural identity intrinsically linked to native flora. The *in situ* conservation (protection within their natural habitat) of these genetic resources is therefore a critical environmental and societal goal [5]. Understanding the factors that enable long-term survival, or longevity, is fundamental to effective plant conservation. Generally, plant longevity is influenced by a complex interplay of genetic predispositions and environmental factors. Critical environmental variables often include soil characteristics (such as nutrient availability, like potassium, and pH), water regimes, elevation, topography, and prevailing climatic conditions. Recent studies have demonstrated that optimal durian persistence is often associated with moderate elevations (200–400 m), well-drained soils with balanced macronutrient profiles, and relative humidity levels that support consistent flowering and fruiting cycles [6]. Research in tropical fruit tree ecology has further emphasized that soil organic matter content and seasonal moisture stability can significantly extend tree lifespan by promoting root health and nutrient cycling [7]. Studies in diverse ecosystems, like the environmentally determined flora distributions in the Canary Islands or the observed links between elevation, persistence, and longevity in Mediterranean conifers, illustrate how specific environmental niches support long-lived plant populations. However, a significant knowledge gap remains concerning the specific drivers of longevity for tropical tree species, particularly those like the Khlong-Saeng durian, which are not widespread commercial cultivars. Factors governing the persistence of these unique tropical trees are largely undocumented. The primary aim is to identify and analyze the key environmental variables systematically – encompassing both physical site attributes (e.g., elevation, slope, aspect) and soil chemical properties – that exert a significant influence on the longevity and persistence of native Khlong-Saeng durian trees within their natural habitat in the Ban Khao Thep Phithak area of Surat Thani. Therefore, a secondary objective involves the preliminary phytochemical screening of Khlong-Saeng durian peel samples. This information is crucial for exploring potential avenues for the sustainable utilization of this peel waste, thereby fostering awareness of its potential value and local applications. Ultimately, this investigation seeks to establish a much-needed foundational knowledge base. Highlights of this research include (1) identifying the environmental and soil chemical factors most strongly correlated with the persistence of native durian trees, (2) providing the first phytochemical profile of Khlong-Saeng durian peel, and (3) linking ecological conservation strategies with potential economic benefits from agricultural byproduct utilization. The findings are expected to provide direct, evidence-based insights to guide targeted conservation strategies for the threatened Khlong-Saeng durian, focusing on preserving the environmental conditions conducive to its long-term survival. Simultaneously, by characterizing the peel's phytochemistry, the study aims to contribute to more holistic and sustainable management practices within the region, promoting the responsible use of resources and potentially adding value to agricultural byproducts. This research endeavors to bridge critical gaps in tropical plant ecology, conservation science, and sustainable resource management.

2. Materials and Methods

2.1 Study area

The research was conducted in Ban Khao Thep Phithak, located within the Khao Phang sub-district of Ban Ta Khun district, Surat Thani Province. This area is characterized by a valley-like formation, marked as No.1, 2, and 3, and intersects with the reservoir of Rajjaprabha Dam. It features a drainage system connecting the surrounding regions. Access to this area is limited to a single entrance and exit, with coordinates ranging from latitude 8.940 to 8.954 and longitude from 98.812 to 98.826 (Figure 1).

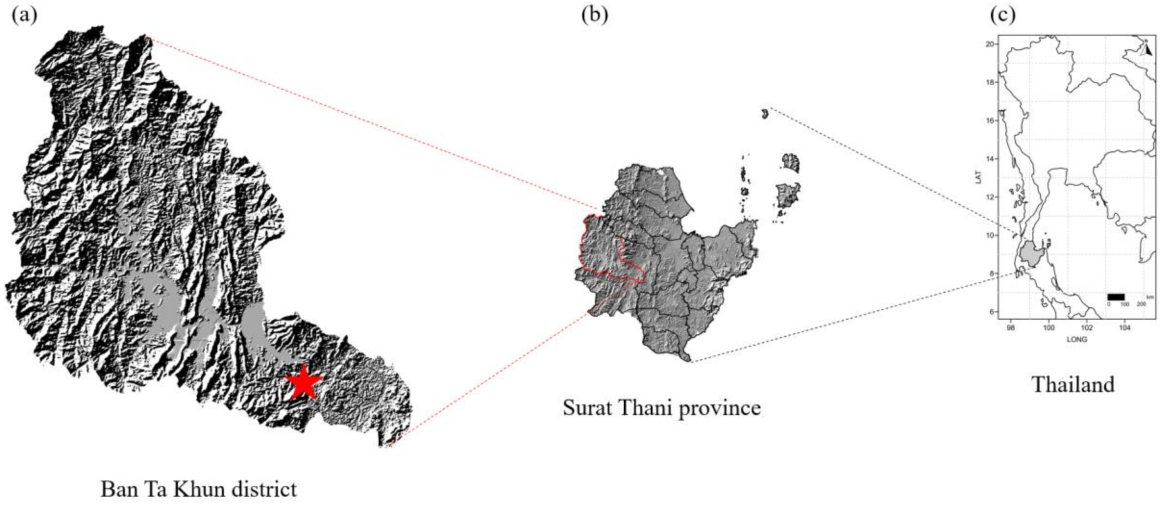


Figure 1. The study area, the red star represented the working area, the Ban Khao Thep Phithak in the Khao Phang sub-district, Ban Ta Khun district (a), in Surat Thani province (b), in the southern part of Thailand (c).

2.2 Field sampling

Data collection took place between July and September 2024. The sample included durian trees situated in Khlong Saeng (KS) and encompassed essential information for academic investigation. Measurements and records were obtained for various parameters, including temperature, soil pH, soil moisture, air humidity, elevation, tree circumference, and canopy dimensions. Additionally, fruit samples were gathered in August 2024, coinciding with the fruiting season.

2.3 Durian peel sampling

To obtain crude extracts from Durian peels, the initial step involved chopping the samples, followed by drying them at 50°C for 24 hours. Subsequently, the dried samples were subjected to methanol extraction at a ratio of 10 g of the dry sample per 100 mL of methanol, utilizing reflux extraction over a 2-hour duration. The resultant solution was then filtered through filter paper to isolate the extract. The solvent was subsequently removed under vacuum conditions using a rotary evaporator (KNF RC600, Germany). A 1 mL portion of the sample was obtained by centrifugation at 1,000 rpm for 5 minutes, allowing for the analysis of essential phytochemical constituents.

2.4 Environmental Conditions

Environmental parameter

Environmental parameters were gathered in triplicate at each durian tree on the day of collection. Four key environmental factors were assessed: temperature, relative humidity, soil moisture, and elevation above sea level. Temperature and relative humidity were measured using an automatic hygrometer (TSI® Model 9545/9545-A, VelociCalc®, United States of America) by recording these values within the radius of the canopy beneath the durian tree. Elevation and height above sea level for each tree were determined using a GIS meter (GPSMAP 66S, Garmin, Taiwan). The procedure involved aiming at the tree's base, waiting for the data to stabilize, and then recording the readings.

Soil Physiology Analysis

The particle size distribution method was used to examine soil texture in dried soil samples that were sieved through a 2 mm screen. The proportions of sand, silt, and clay were determined using the pipette method in accordance with USDA rules. The USDA soil texture triangle was used to determine soil texture [8]. Soil moisture measurements were gathered using the oven-drying method, which entailed removing branches and extraneous debris from the topsoil surrounding the tree prior to examination.

Soil Chemical Property Analysis

A pH meter was used to assess soil pH and electrical conductivity (EC) at a 1:2 ratio of soil to distilled water. Soil samples were sieved through a 0.5 mm mesh and evaluated for organic matter content using the Walkley-Black method [9]. Total nitrogen content was determined using the Kjeldahl method [10], while accessible phosphorus was determined using the Bray II method [11]. To analyze potassium, calcium, and sodium, soil samples were extracted using 1 M ammonium acetate (NH_4OAc) at pH 7. The concentrations of these elements were determined using an Atomic Absorption Spectrophotometer (AAS, Thermo iCE 3000 Series, USA). Model 410 Flame Photometer (Sherwood Scientific, UK) was used. To examine potassium, atomic absorption spectroscopy was used to assess calcium and sodium [12].

2.5 The morphology

2.5.1 Clinometer

The height of a large tree was estimated using the clinometer method, a method that involves angling [13]. The height of the entire tree is calculated as H1 = tan a(AC)+H2 where H is a height of the tree (Figure 2) where tan a is the angle formed by the tangent of the angle ABC, H_1 is the height of the tree from ground level to the top of the tree, and H_2 is the height of the tree from the ground to the eye level of the observer.

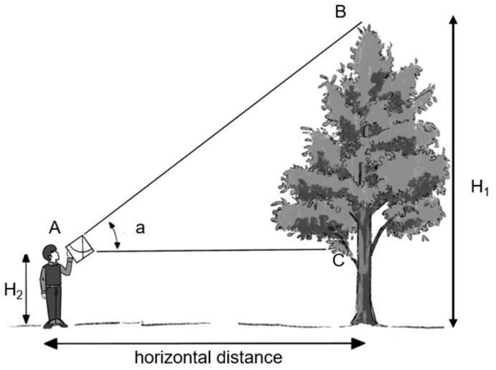


Figure 2. An instrument and method for measuring tree height by the clinometer method.

2.5.2 Diameter Distribution

Calculate the diameter of the stem. Take a tape measure and measure the circumference of the trunk at the height of the chest DBH (Diameter at Breast Height), then use the formula to calculate the trunk diameter as follows: $D = x / \pi$; where D = stem diameter, and x = circumference.

2.5.3 Crown cover

The surveyor's position appears like a sturdy pillar at the edge of the canopy due to the canopy's dimensions. To determine the distance to the canopy's edge from the opposite side in both the north-south and east-west directions, employ a tape measure and select the greater of the two values when obstructions are present. It may not be feasible to measure both sides due to limited access. To calculate the radius, measure the distance between two preliminary points using a tape measure. The diameter is subsequently found by doubling this length.

2.6 Geographical index analysis

Create a geographical map depicting the prevalence of KS durian trees within the regions of Ban Khao Thep Pitak Villages 1, 2, and 3, situated in Khao Phang sub-district, Ban Ta Khun District. This map is generated by integrating collected coordinates and environmental data into the mapping system. Tree coordinates are depicted using a three-dimensional framework, with the X-axis signifying latitude, the Y-axis signifying longitude, and the Z-axis representing various environmental variables, including altitude above sea level, temperature, soil pH, soil moisture content, and air humidity, for each data point.

2.7 Phytochemical screening

The unrefined extract was applied to TLC silica gel 60 F254 aluminium plates (10 cm x 20 cm, Merck, Switzerland), each with a 0.8 cm strip, using a Linomat-5 (CAMAG, Switzerland) and a Nitrogen aspirator. Two microliters of crude extract, at a concentration of 1 mg/mL, were spotted on the TLC plate. Subsequently, the TLC plates were developed in an ADC-2 Automatic Developing Chamber (CAMAG, Switzerland) with a mobile phase composed of toluene, acetonitrile, and ethyl acetate in a 35:5:15 ratio. The phytochemical profile was documented under white light, at 254 nm, and at 366 nm using the CAMAG TLC Visualizer2. Phytochemical groups were identified by applying specific reagents to the TLC plates bearing the phytochemical fingerprints and observing the reaction of the chemical groups. Alkaloids were treated with the Dragendorff reagent, and the plate was heated at 100°C for 5 minutes. The reaction was visualized as a red-orange color under 366 nm. Terpenoids and steroids were analyzed using the vanillin reagent, which was sprayed onto the plate, and then the plate was heated at 100°C for 5 minutes. The reaction appeared as a brown-purple color under 366 nm. Coumarins and anthraquinones were treated with the potassium hydroxide reagent, and the plate was heated at 100°C for 5 minutes. The reaction exhibited a blue-red color under 366 nm. Phenolics, phenols, terpenes, sugars, and steroids were identified using the p-Anisaldehyde reagent, which was sprayed onto the plate, and then the plate was heated at 100°C for 5 minutes. The reaction manifested as blue, red, gray, or green colors under 366 nm, respectively. Flavonoids were visualized under a 366 nm wavelength after the aluminum chloride reagent was sprayed and the plate was heated at 100°C for 5 minutes. To assess the phytochemical response on the TLC plate to the antioxidant test, a separate TLC plate was sprayed with a 0.5% 2,2-diphenyl-1-picrylhydrazyl (DPPH) solution in methanol. This action resulted in the substance turning yellow after a 5-minute incubation at room temperature in the dark.

2.8 Statistical analysis

To investigate the relationship between environmental factors and the longevity of a local durian tree using the program R (Ver 4.3.1 for Windows). The data were analyzed using a one-way analysis of variance (ANOVA). Significant differences were identified, and multiple comparisons were performed using Tukey's test at a 95% confidence level.

3. Results and Discussion

3.1 The evaluation of the geographical index

Geographic mapping was utilized to visualize the spatial distribution of native durian trees in the study area, integrating both coordinate data and environmental factors. Durian trees were categorized by age into five classes: 200-250 years, 150-200 years, 100– 150 years, 50-100 years, and less than 50 years. These categories were represented by different color-shaded symbols on the map (Figure 3). The durian trees were primarily distributed within a narrow geographic range, spanning latitudes 8.940–8.954 and longitudes 98.812–98.826. The surveyed area was characterized by a single access point and coordinates ranging from latitude 8.938 to 8.948 and longitude 98.818 to 98.824. A total of 83 durian trees were recorded, located in Villages 1, 2, and 3. Age classification, based on interviews with landowners, revealed that 12 trees were over 100 years old, 37 were between 50 and 99 years old, and 34 were younger than 50 years. Most trees were concentrated in Village 1, with no trees found in Village 2. In Village 3, four trees were recorded, including one over 100 years old and three aged between 55 and 99 years (Figure 3). The age distribution and spatial clustering patterns observed in this study are consistent with findings from similar research on durian populations in Southern Thailand and Peninsular Malaysia, where older trees are often concentrated in areas with long-term traditional cultivation practices and minimal land-use change [14]. Studies on other perennial tropical fruit trees, such as mangosteen and rambutan, have also reported that historical planting patterns, coupled with access to water sources and suitable microclimates, strongly influence the persistence of older trees [15]. The absence of trees in certain villages may reflect past land conversion, soil condition differences, or socioeconomic factors that limit durian cultivation, as noted in previous agroforestry studies [16]. This suggests that conservation of the oldest durian trees may require targeted protection measures in specific geographic zones, along with community-based initiatives to maintain genetic diversity and local heritage value.

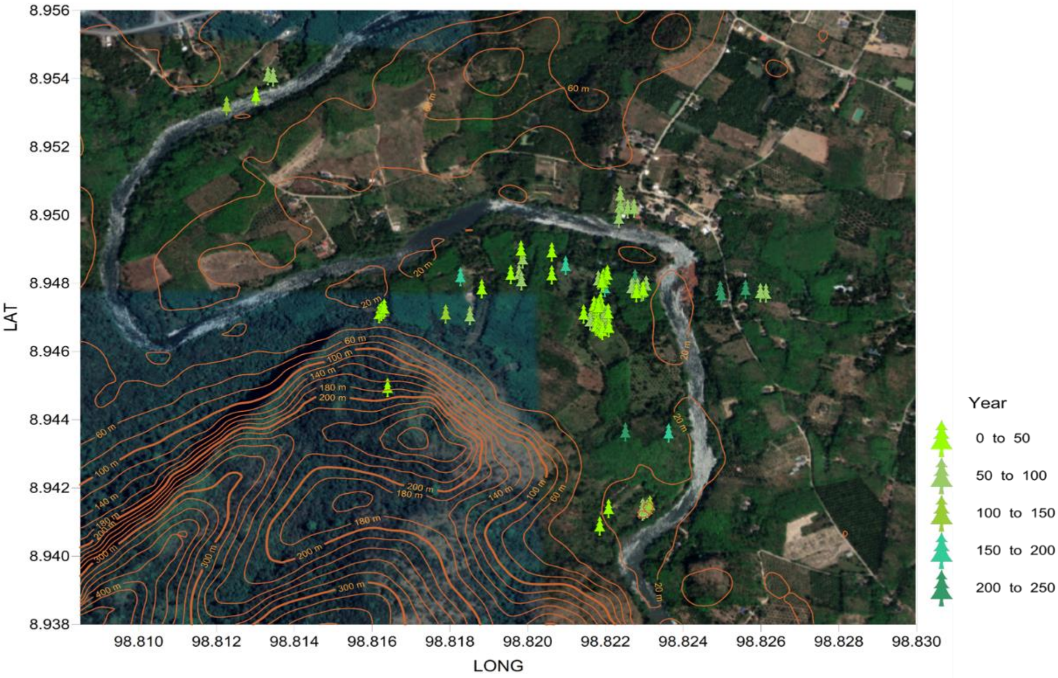


Figure 3. The distribution of the durian tree according to its age.

3.2 Environmental Characters

3.2.1 Elevation, Temperature, Soil Moisture, and Humidity

The study assessed environmental parameters, including elevation above sea level, ambient temperature, soil moisture, and relative humidity. Elevation ranged from 29 to 62 meters, temperature ranged from 26°C to 31°C, soil moisture content ranged from 9.5% to 14%, and air humidity ranged from 58.87% to 87.37%. A Chi-square analysis revealed a statistically significant correlation between tree age and temperature ($p < 0.05$). Durian cultivation is typically viable from 0 to 800 meters above sea level and within temperature ranges of 25–32°C. In comparison to traditional durian-growing regions in Indonesia, the study area showed slightly lower average temperatures (27–31°C), which may influence the distribution and persistence of native durians on Ternate Island [17]. Similar studies in Indonesia have reported durian growth at elevations ranging from 13 to 427 meters, with notable findings such as 58 indigenous durians aged over 50 years thriving at 400 meters above sea level in Ngawi, Indonesia [18]. Additionally, local durians from Batuah Village, aged 24–27 years, are known for their distinctive fragrance, which is attributed to specific humidity conditions (56–65%) and the presence of volatile sulfur and ester compounds [19]. In contrast, most pine trees can endure for up to 100 years within temperature ranges of 2.2 to 3.4°C [20], in comparison to other areas where traditional durian is also grown. Indonesia, in particular, discovered that the average temperature of the study area is lower than the air temperature, which ranges from 27-31°C during surveys, affecting the distribution and persistence of native durian on Ternate Island, Indonesia [17]. The results of the environmental factors assessment for durian (*Durio zibethinus*) planting in Pangkatan District, Labuhanbatu Regency, revealed that the temperature ranged between 25°C and 32°C [21]. The age of the durian tree, however, is not mentioned in the studies. Durians are grown at an average annual temperature of 27.8°C, providing volatile compounds and fragrances characteristic of the durian variety Durio zibethinus, which produces sulfur and esters, as well as aromatic compounds not typically found elsewhere. The durian was between 24 and 27 years old when it was discovered in Batuah Village, Loa Janan District, Kutai Kartanegara, East Kalimantan, Indonesia. This is possible because local durian has a stronger odor than other durian varieties [19]. The variety of the durian. The durian plantation in Gua Musang, Kelantan, was established at 26.4°C on a slope with the highest point 439 m above sea level and the lowest point 410 m above sea level.

The findings of this study indicate that the elevation range and microclimatic conditions of the surveyed area fall within the generally accepted thresholds for durian cultivation (0–800 m above sea level, 25–32°C). The slightly lower average temperatures recorded in the study area (27–31°C) compared to some traditional durian-growing regions in Indonesia may contribute to variations in fruiting patterns, flowering cycles, and overall tree longevity. Previous research in Ngawi, Indonesia, reported the successful cultivation of indigenous durians aged over 50 years at an elevation of 400 m above sea level, suggesting that moderate elevation, coupled with stable thermal conditions, can support long-term tree survival and maintain high fruit quality [22]. Elevation also influences soil moisture retention, as higher altitudes often exhibit cooler temperatures and reduced evapotranspiration rates, thereby affecting root water availability. Studies in East Kalimantan have shown that durians grown in areas with relative humidity levels between 56% and 65% develop distinctive aromatic profiles due to the enhanced biosynthesis of volatile sulfur compounds and esters, which are influenced by moisture-related metabolic pathways [19]. In contrast, excessive fluctuations in soil moisture can lead to nutrient leaching or waterlogging, both of which can stress the trees and alter fruit chemistry.

3.3 Physical Properties of Soil by Tree Age

The analysis of soil texture revealed that all durian trees grew in soils with a silty texture. A one-way ANOVA demonstrated statistically significant differences in soil characteristics by age group at the 99% confidence level ($p < 0.001$, F(10, 22) = 42.96). The silt content ranged from 81.93 ± 0.42% to 95.51 ± 0.56%, with the highest sand proportion in the 0–10-year group and the lowest in the 100–200-year group. Soil moisture content also varied significantly ($p < 0.001$, F(10, 22) = 7.218), ranging from 11.73 ± 0.49% to 15.93 ± 0.64%. The highest moisture was recorded in the 100–200-year group and the lowest in the 40–50-year group (Table 1).

3.4 Chemical Properties of Soil by Tree Age

The soil chemical composition varied significantly with the age of the durian trees. Soil pH ranged from 4.43 ± 0.02 (90–100 years) to 5.93 ± 0.04 (100–200 years), with significant differences detected ($p < 0.001$, F(10, 22) = 1.469). Electrical conductivity (EC) values ranged from 44.23 ± 0.25 to 173.70 ± 0.15 mS/cm ($p < 0.001$, F(10, 22) = 3.365). Organic matter (OM) content also showed significant variation ($p < 0.001$, F(10, 22) = 689.8), ranging from 1.23 ± 0.03% (70-80 year group) to 2.89 ± 0.02% (40-50 year group), with the highest OM content in the 40–50-year group. Soil nitrogen levels ranged from 667.10 ± 1.96 to 2,365.00 ± 31.27 ppm ($p < 0.001$, F(10,22)=6.202), with the 50–60-year group exhibiting the highest content. Phosphorus content varied from 1.55 ± 0.22 to 32.72±0.17 ppm ($p < 0.001$, F(10,22)=16.113), and potassium levels ranged from 34.67±0.06 to 72.47 ± 0.06 ppm ($p < 0.001$, F(10,22)=1.264), peaking in the 80–90-year group. The calcium content significantly differed across age groups ($p < 0.001$, F(10, 22) = 10.740), ranging from 25.67 ± 0.59 to 487.80 ± 0.55 ppm. Sodium also showed variation ($p < 0.001$, F(10, 22) = 1.324), with values from 10.34 ± 0.17 to 98.60±0.46 ppm. These results suggest that the age of the durian tree plays a significant role in influencing the chemical characteristics of the surrounding soil (Table 2).

Table 1. Physical properties of soil according to the age range.

Age range (years)	**Physical properties of soil**				**Soil type**
	soil particles			**Soil moisture content (%)**	
	Sand (%)	**Silt (%)**	**Clay (%)**		
0-10	1.44 ± 0.032	95.51 ± 0.563	3.05 ± 0.538	13.53 ± 1.147	Silt
10-20	1.59 ± 0.053	91.76 ± 0.208	6.35 ± 0.256	13.03 ± 0.954	Silt
20-30	1.39 ± 0.022	87.70 ± 0.332	10.91 ± 0.310	14.00 ± 1.041	Silt
30-40	1.47 ± 0.035	94.15 ± 0.048	4.38 ± 0.070	14.23 ± 0.452	Silt
40-50	1.43 ± 0.041	94.42 ± 0.337	4.14 ± 0.323	11.73 ± 0.493	Silt
50-60	1.44 ± 0.046	95.42 ± 0.141	3.14 ± 0.168	15.33 ± 0.746	Silt
60-70	1.75 ± 0.048	94.79 ± 0.242	3.45 ± 0.262	15.57 ± 0.572	Silt
70-80	1.52 ± 0.012	95.47 ± 0.258	3.00 ± 0.245	14.50 ± 0.980	Silt
80-90	1.46 ± 0.044	92.87 ± 0.236	5.66 ± 0.218	14.73 ± 0.666	Silt
90-100	1.72 ± 0.045	92.71 ± 0.235	5.56 ± 0.275	14.20 ± 0.361	Silt
100-200	1.40 ± 0.010	81.93 ± 0.420	16.67 ± 0.426	15.93 ± 0.642	Silt

Note: all pairs of the age of the tree show a statistically significant result at $p < 0.05$

Soil chemical properties, including organic matter, pH, and concentrations of macro- and micronutrients, were observed to vary with tree age in this study. Older trees (over 100 years) were generally associated with soils richer in organic matter and available potassium, possibly due to prolonged litter fall and root turnover that contribute to nutrient cycling. This finding is consistent with those from durian agroforestry systems in Malaysia, where older stands exhibited higher cation exchange capacity (CEC) and greater microbial biomass, indicating healthier rhizosphere activity [23]. Younger trees, in contrast, were often found in soils with lower organic carbon and reduced nutrient retention, which could be attributed to more recent land conversion or less developed root systems. Humidity and soil moisture levels further interact with soil chemistry by influencing microbial decomposition rates and the mineralization of nutrients. High humidity can enhance the breakdown of organic matter, increasing the availability of nitrogen and phosphorus—nutrients essential for flowering and fruit set [24]. Conversely, in poorly drained soils, high moisture levels may promote anaerobic conditions, leading to reduced nutrient uptake efficiency and potential accumulation of phytotoxic compounds.

3.4 Morphological character

The morphology of the Khlong Saeng (KS) native durian trees was evaluated using three key parameters, which consisted of fruiting cycle length, tree height, and crown cover—all of which showed significant correlations with tree age. Notably, the fruiting cycle length demonstrated statistically significant relationships with age (Figure 4a), tree height (Figure 4b), and crown cover (Figure 4c).

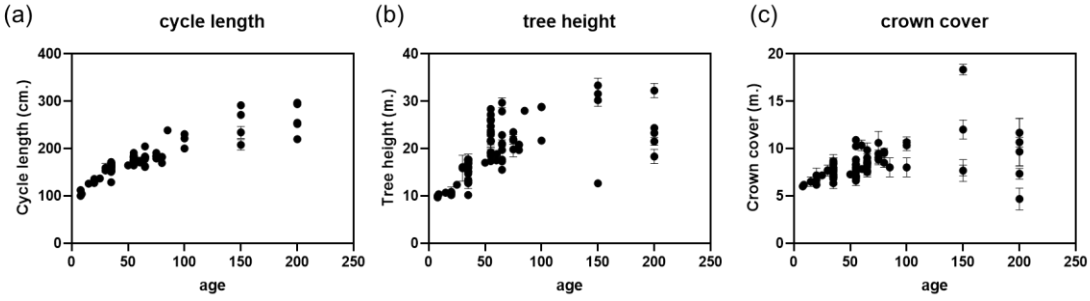


Figure 4. Correlation between tree age and morphological characteristics: (a) cycle length, (b) tree height, and (c) crown cover of native durian trees.

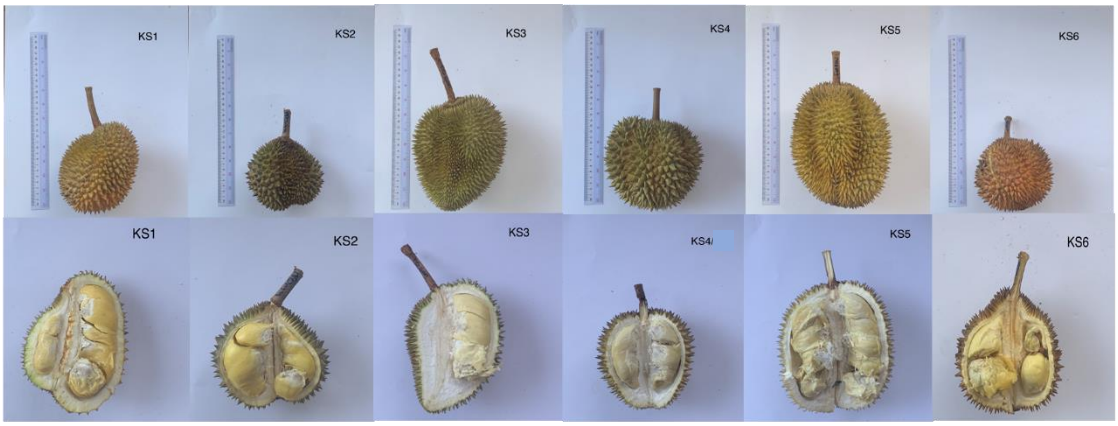
Table 2. Chemical properties of soil according to the age range.

Age range (years)	Chemical properties of soil							
	Soil pH	**EC mS/cm**	**OM (%)**	**N (ppm)**	**P (ppm)**	**K (ppm)**	**Ca (ppm)**	**Na (ppm)**
0-10	5.32 ± 0.018	173.70 ± 0.152	2.32 ± 0.023	885.20 ± 4.790	9.27 ± 0.070	63.43 ± 0.212	467.20 ± 1.689	33.50 ±0.302
10-20	5.27 ± 0.031	106.10 ± 0.148	1.50 ± 0.027	1326.00 ± 6.202	2.07 ± 0.074	53.77 ± 0.056	327.40 ± 0.381	34.60 ± 0.200
20-30	5.42 ± 0.024	69.30 ± 0.202	1.52 ± 0.031	962.80 ± 0.868	8.567 ± 0.088	51.43 ± 0.119	355.70 ± 1.613	32.50 ± 0.345
30-40	4.82 ± 0.020	75.20 ± 0.168	2.15 ± 0.060	1465.00 ± 9.741	1.55 ± 0.217	56.17 ± 0.122	234.10 ± 0.954	33.50 ± 0.608
40-50	4.71 ± 0.017	77.60 ± 0.103	2.89 ± 0.020	1129.00 ± 2.013	7.69 ± 0.183	55.80 ± 0.100	326.00 ± 1.012	26.20 ± 0.561
50-60	5.92 ± 0.022	72.37 ± 0.355	1.75 ± 0.014	2365.00 ± 31.270	32.72 ± 0.165	54.87 ± 0.064	487.80 ± 0.550	10.34 ± 0.173
60-70	4.82 ± 0.024	44.23 ± 0.251	1.61 ± 0.018	838.30 ± 5.227	7.83 ± 0.112	34.67 ± 0.059	110.90 ± 0.352	32.40 ± 0.000
70-80	5.42 ± 0.026	52.80 ± 0.200	1.23 ± 0.025	667.10 ± 1.956	5.60 ± 0.055	56.53 ± 1.672	207.10 ± 0.664	10.80 ± 0.357
80-90	5.72 ± 0.018	90.80 ± 0.204	2.53 ± 0.035	1713.00 ± 12.115	19.66 ± 0.131	72.47 ± 0.063	142.80 ± 0.403	11.17 ± 0.254
90-100	4.43 ± 0.015	65.97 ± 0.227	1.52 ± 0.022	1280.00 ± 4.056	8.25 ± 0.016	50.77 ± 0.057	25.67 ± 0.585	98.60 ± 0.458
100-200	5.93 ± 0.035	79.03 ± 0.849	2.16 ± 0.067	723.60 ± 2.560	2.573 ± 0.134	37.27 ± 0.048	385.80 ± 0.061	25.27 ± 0.671

Note: all pairs of the age of the tree show a statistically significant result at $p < 0.05$

For trees aged 200 years, the average fruiting cycle length was 263.20 ± 4.59 cm, while trees aged over 100 years had a cycle length of 247.64 ± 5.08 cm. Trees aged 50–99 years exhibited an average of 179.02 ± 2.62 cm, and those younger than 50 years had a cycle length of 149.59 ± 3.07 cm. In terms of tree height, trees aged over 100 years averaged 25.60 ± 0.99 m, slightly higher than the 200-year-old group (24.94 ± 1.02 m). Trees aged 50–99 years had an average height of 21.83 ± 0.78 m, while those under 50 years measured 14.12 ± 0.72 m on average. For crown cover, 100-year-old trees had the largest average (9.89 ± 1.06 m^2), followed by 200-year-old trees (8.93 ± 1.27 m^2), 50–99-year-old trees (8.40 ± 0.51 m^2), and the youngest group under 50 years (7.32 ± 0.42 m^2). These findings indicate a positive correlation between tree age and fruiting cycle length. Older trees generally exhibit longer fruit development periods, which could influence resource allocation, hormonal regulation, and physiological processes as the tree matures [25, 26]. While environmental conditions were not directly compared across age groups in this study, it is possible that long-term exposure to local conditions also contributes to these trends. The crown cover, reflecting the canopy spread, typically increases with age as trees expand their branching to maximize sunlight capture [27]. The observed peak in crown cover among trees over 100 years, followed by a slight decline in the 200-year group, may be attributed to natural senescence, branch dieback, or historical pruning practices. These morphological traits—cycle length, height, and crown spread—are valuable indicators of developmental stages in KS native durian trees and could be used to estimate tree age in the absence of planting records. Given the correlation between morphological traits and fruit yield, understanding these age-related patterns can support better orchard management. Tailoring interventions such as pruning or fertilization based on tree age and morphological characteristics could enhance productivity and sustainability.

During fieldwork, durian samples were collected from trees aged 30 to 70 years, with a focus on fruit-bearing specimens. Trees were selected from multiple gardens across the region to assess morphological diversity among KS local varieties. In a related study in North Aceh, Indonesia, durian trees ranged in age from 20 to 200 years, with heights between 20 and 83 meters, and exhibited variations in trunk and leaf color associated with age and site-specific water availability. Notably, optimal yields of 200 fruits per tree were observed at elevations between 0 and 910 meters, with temperatures ranging from 22 to 46°C, and a relative humidity of around 30%, for trees aged 80–150 years. In Malaysia, durian trees aged 25 to 150 years grew under 50–80% humidity and at temperatures of 25–32°C, reaching heights of 18.5–38 meters. Our KS samples were labeled KS1–KS6, with respective tree heights: KS1 (24.67 m), KS2 (18.83 m), KS3 (13.33 m), KS4 (19.17 m), KS5 (15.87 m), and KS6 (19.67 m). Their fruiting cycle lengths were: KS1 (181.67 cm), KS2 (170.67 cm), KS3 (168.67 cm), KS4 (176.00 cm), KS5 (159.00 cm), and KS6 (177.67 cm). Crown cover measurements were as follows: KS1 (6.60 m), KS2 (7.83 m), KS3 (7.50 m), KS4 (8.83 m), KS5 (7.67 m), and KS6 (8.50 m).



lemon yellow with smoot texture soft yellow with smoot texture soft yellow with smoot texture soft yellow with smoot texture soft yellow with smoot texture lemon yellow with smoot texture

Figure 5. The local durian in Surat Thani Province. Local names Khlong-Saeng Durian; KS1 = Mang Khut, KS2 = Mum Fai, KS3 = Khiao Thawai, KS4 = Thang Lot, KS5 = Gan Petch, and KS6 = Kor Trong.

The analysis of fruit shape among these samples identified four major categories: globose (KS5, KS6), oblate (KS2, KS4), elliptic (KS1), and obovoid (KS3). Flesh color was classified into dark yellow and soft yellow types. Native durians in Thailand typically exhibit an oval fruit shape, weigh between 2.20 and 3.42 kg, and have bright yellow flesh [28]. In contrast, local durians in Ternate, Indonesia, exhibit oval-spherical thorns, white skin, and yellowish-brown, oval-shaped seeds [29]. Environmental conditions significantly impact the longevity and productivity of durian trees. Previous studies have shown that optimal durian growth occurs at 20–30°C [30], under suitable conditions of air humidity, soil pH, and moisture. Excessive moisture, however, can lead to rot and fungal diseases. In Malaysia, soil pH values of 6.88–7.60 and soil moisture levels of 57.37–67.07% were found to be optimal [31]. In contrast, this study observed a slightly more acidic soil pH, ranging from 5.67 to 6.33. A related Indonesian study found that durian trees at elevations of 200–250 meters produced both rectangular and oblate fruits, with textures ranging from white to pale-yellow to lemon yellow [32]. Furthermore, the highest durian pollen germination rate (63%) was reported in Thailand at 30°C, highlighting temperature as a key factor for reproductive success [33]. The local durian varieties in Surat Thani province exhibit diverse morphological and organoleptic traits, reflecting both genetic diversity and adaptation to the province's unique environmental conditions. In this study, four distinct fruit shape categories were identified—globose, oblate, elliptic, and obovoid—suggesting that multiple landraces or locally adapted lines are present within the region. Such variation is consistent with previous reports on native Thai durians, which often exhibit an oval or oblong fruit shape, medium to large fruit size (2.20–3.42 kg), and bright yellow flesh coloration [28]. The variation in shape may influence both consumer preference and market value, as shape uniformity is often associated with commercial grading standards, while unique forms can attract niche markets. Flesh color variation, observed here as dark yellow and soft yellow, may be linked to differences in carotenoid and flavonoid concentrations, which not only influence appearance but may also contribute to antioxidant properties [34]. The color spectrum observed in Surat Thani durians aligns with reports from other Southeast Asian regions, such as Indonesia's Ternate durians, where seed color and skin tone vary markedly depending on environmental conditions [17].

Environmental factors in Surat Thani appear to play a central role in shaping both tree longevity and fruit quality. The province's climate, with average temperatures often ranging between 27 °C and 31°C, falls close to the optimal range of 20 °C to 30°C for durian growth and reproduction. However, the slightly more acidic soil pH recorded in this study (5.67–6.33) differs from the neutral to slightly alkaline conditions (pH 6.88–7.60) reported as optimal in Malaysia [31]. This suggests that Surat Thani durians may possess a degree of tolerance to moderately acidic soils, a trait potentially advantageous in tropical zones with high rainfall and leaching rates. Soil moisture levels, although not explicitly reported in the result segment, are critical in durian cultivation. Excessive moisture can predispose trees to root rot and fungal infections [35]. The slightly acidic soils of Surat Thani may also affect nutrient availability, particularly of calcium and magnesium, which are essential for fruit development and structural integrity. The durability of durian trees in such conditions may indicate long-term ecological adaptation, supporting the survival of certain landraces despite less-than-optimal pH levels. From a reproductive biology perspective, the success of pollen germination is highly dependent on temperature. The highest germination rate reported in Thailand (63% at 30°C) [33] is consistent with the prevailing temperature range in Surat Thani, potentially explaining the sustained productivity of older trees in the region. When compared to Indonesian durians grown at higher elevations (200–250 m) with diverse fruit shapes, Surat Thani's local varieties may prioritize flavor and flesh color consistency over morphological extremes, possibly due to market preferences and historical selection by farmers. Overall, the morphological diversity, flesh color variation, and environmental resilience of local durians in Surat Thani suggest a well-adapted population with significant potential for selective breeding and premium branding. These characteristics, combined with the province's established reputation as a major durian-producing area in southern Thailand, provide strong opportunities for value addition through cultivar registration, geographical indication (GI) labeling, and agritourism initiatives. Future studies should integrate genetic analysis with detailed sensory profiling to better understand the linkage between genotype, phenotype, and environmental adaptation in these local varieties.

3.5 Phytochemicals and antioxidant effects of DPPH on TLC plate

The phytochemical analysis of crude durian peel extract, using the derivatized reagent method, revealed the presence of five distinct groups of substances: terpenoids, steroids, phenolics, phenols, and sugar esters. These findings are summarized in Table 3.

Table 3. Phytochemicals of crude durian peel extract.

Phytochemicals	**reagents**	**crude durian peel extract**						
		KS1	KS2	KS3	KS4	KS5	KS6	KS7
Terpenoid	Vanillin	-	-	-	+	+	+	+
Steroids	Vanillin	-	-	-	+	+	+	+
Phenolic	Aluminium chloride	+	+	+	+	-	-	+
Phenol	p-Anisaldehyde	+	+	-	-	-	-	-
Sugar steroid	p-Anisaldehyde	+	+	-	-	-	-	-

Note: + means positive result; - means negative result. Local names Khlong-Saeng Durian; KS1 = Mang Khut, KS2 = Mum Fai, KS3 = Khiao Thawai, KS4 = Thang Lot, KS5 = Gan Petch, KS6 = Kor Trong and KS7 = Sho Nui

Terpenoid compounds were evident in extracts KS4, KS5, KS6, and KS7, with steroid compounds also being apparent in these same extracts. Phenolic substances were found in KS1, KS2, KS3, KS4, and KS7, and distinct phenolic compounds were detected in KS1 and KS2 extracts. Furthermore, steroid substances were identified explicitly in extracts KS1 and KS2, as outlined in Table 2. The DPPH assay demonstrated the antioxidant activity of the crude durian peel extract, with results shown on the TLC plate in Figure 6. This test visualized the antioxidant properties of various compounds in the extract.

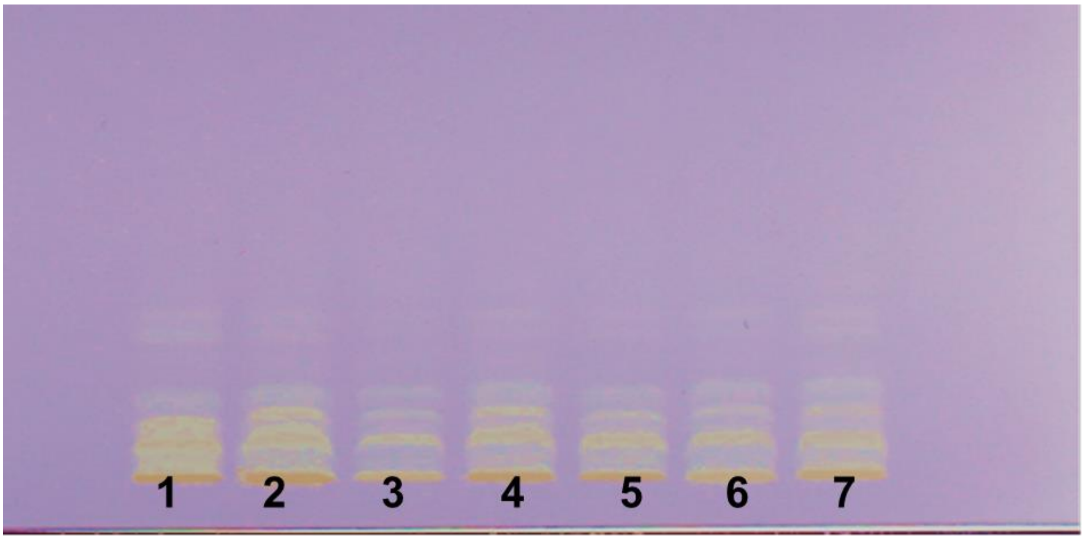


Figure 6. Antioxidant activity of the durian peel extract, assessed via DPPH reaction on a TLC plate under white light illumination at room temperature.

Durian fruit peels are often seen as environmental waste due to the high volume of byproducts, particularly during the blooming season in plantation areas. The growing emphasis on zero waste in agriculture has led to increased interest in utilizing durian peels [36-37]. In this study, the crude peel extract was tested for antioxidant activity using TLC screening. Previous studies have also investigated durian peels as potential raw materials with pharmacological properties that target issues such as diabetes, dental caries, and root canal infections. These investigations explored the antioxidant activity in durian peels, aiming to uncover bioactive compounds that could serve as nutraceuticals with antioxidant, anti-inflammatory, and health-promoting benefits [38, 39]. The TLC plate clearly showed that the crude durian peel extract contains multiple compounds exhibiting antioxidant activity. The presence of yellow spots (1-7) on the TLC plate suggests that several phytochemicals contribute to this effect. This aligns with the screening results (Table 3), where terpenoids, steroids, and phenolic compounds were identified, all of which likely separate and contribute to the antioxidant activity observed on the TLC plate (Figure 6). This study underscores the potential of durian peels, often discarded as agricultural waste, as a valuable source of natural antioxidants. This aligns with the growing movement to valorize food waste and identify sustainable sources of bioactive compounds for various applications in pharmaceuticals, nutraceuticals, and food preservation [40]. Given the high volume of durian peels generated, the study highlights their untapped potential for high-value applications due to their rich phytochemical content. Moreover, durian peels have been traditionally explored for their pharmacological properties, particularly in managing diabetes, dental caries, and root canal infections, further emphasizing the importance of studying their bioactive components.

4. Conclusions

This research on the Khlong Saeng native durian has provided significant insights into the morphological characteristics, environmental factors, and phytochemical properties of these trees, contributing to a better understanding of their development, ecological significance, and potential agricultural value. Morphological Character: The study found a significant correlation between tree age and morphological traits, including fruit cycle length, tree height, and crown cover, with older trees showing longer fruit cycles, greater height, and larger crown covers. Environmental Factors: The growth and age of the durian trees were influenced by local environmental conditions, including temperature, humidity, and soil properties, which varied across different regions. Phytochemicals and Antioxidant Effects: Phytochemical analysis of durian peel extracts revealed the presence of terpenoids, steroids, and phenolic compounds, with antioxidant activity demonstrated through DPPH reaction on a TLC plate, indicating the potential for bioactive applications. This highlights the untapped value of durian peel waste in the pharmaceutical and nutraceutical industries. Future research could focus on isolating specific bioactive compounds from durian peels and evaluating their therapeutic potential for human health. In addition, the findings could be utilized to develop sustainable agroforestry practices that protect native durian biodiversity while enhancing the income of local farmers. Establishing guidelines for selective breeding, conservation zoning, and value-added product development could ensure the long-term preservation and economic viability of Khlong Saeng native durians.

5. Acknowledgements

Thank you to Mr. Rittirong Ritkul, Chairman of Khlong-Saeng Community Enterprise, for providing Khlong-Saeng durian information and the sampling allowance. Thank to Mr. Thanupon Chantakul and Dr. Nuttakarn Naepimai for their assistance in collecting information on physical environmental factors. Thank you to Asst. Prof. Kasemsak Saetang for assisting with data analysis and mapping training. Their contributions have greatly enhanced the comprehensiveness and accuracy of this study.

Author Contributions: Conceptualization, S.S.; methodology, S.S., W.C. and S.S.; data collection, S.S. and S.S.; statistical analysis, S.S. and S.S.; resources, S.S., S.S.; writing original draft preparation, S.S.; writing review and editing, S.S., S.T. and S.S.; visualization, S.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] Lilavanichakul, A.; Pathak, T. B. Thai Farmers' Perceptions on Climate Change: Evidence on Durian Farms in Surat Thani Province. *Clim. Serv.* **2024**, *34*, 100475. https://doi.org/10.1016/j.cliser.2024.100475

[2] Srisunthon, P.; Chawchai, S. Land-Use Changes and the Effects of Oil Palm Expansion on a Peatland in Southern Thailand. *Front. Earth Sci.* **2020**, *8*, 559868. https://doi.org/10.3389/feart.2020.559868

[3] Tali, B. A.; Khuroo, A. A.; Nawchoo, I. A.; Ganie, A. H. Prioritizing Conservation of Medicinal Flora in the Himalayan Biodiversity Hotspot: An Integrated Ecological and Socioeconomic Approach. *Environ. Conserv.* **2019**, *46*(2), 147–154. https://doi.org/10.1017/S0376892918000425

[4] Nakorn, S. N.; Chalumpak, C. Effect of Tree Age and Fruit Age on Fruit Development and Fruit Quality of Pummelo Var. Tabtimsiam. *J. Agric. Technol.* **2016**, 637-645.

[5] Aregay, N.; Belew, D.; Zenebe, A.; Haile, M.; Gebresamuel, G.; Girma, A. Tree Age and Harvesting Season Affected Physicochemical and Bioactive Compounds of Elite Type of Gunda Gundo Orange (*Citrus* spp.) in Northern Ethiopia. *Int. J. Fruit Sci.* **2021**, *21*(1), 26–39. https://doi.org/10.1080/15538362.2020.1852150

[6] Amran, A.; Ariffin, M. R.; Isa, I. M.; Ahmed, O. H.; Herman, G.; Muhamad, S. H.; Nor, N. A. M.; Khairuddin, M. N. Physicochemical Properties of Soil Cultivated with *Durio zibethinus* Murr. in Gua Musang, Kelantan. *AGRIVITA J. Agric. Sci.* **2023**, *45*(2), 278–287. http://doi.org/10.17503/agrivita.v45i2.3601

[7] Omokaro, G. O.; Osarhiemen, I. O.; Idama, V.; Airueghian, E. O.; West, S. T.; Igbigbi, F. E.; Nnake, D. C.; Obolokor, E.; Ahmed, A.; Omoshie, V. O. The Role of Organic Amendments and Their Impact on Soil Restoration: A Review. *Asian J. Environ. Ecol.* **2024**, *23*(11), 41–52. https://doi.org/10.9734/ajee/2024/v23i11620

[8] Soil Survey Staff. *Soil Taxonomy: A Basic System of Soil Classification for Making and Interpreting Soil Surveys*, 2nd ed.; U.S. Department of Agriculture, Natural Resources Conservation Service: Washington, DC, **1999**.

[9] Walkley, A. J.; Black, I. A. Estimation of Soil Organic Carbon by the Chromic Acid Titration Method. *Soil Sci.* **1934**, *37*, 29–38.

[10] Bremner, J. M.; Keeney, D. R. Determination and Isotope-Ratio Analysis of Different Forms of Nitrogen in Soil: Exchangeable Ammonium, Nitrate and Nitrite by Extraction Distillation Methods. *Soil Sci. Soc. Am. J.* **1966**, *30*, 577–583. https://doi.org/10.2136/sssaj1966.03615995003000050015x

[11] Wuenscher, R.; Pignatelli, V.; Zorn, M.; Glaser, B. A Comparison of 14 Soil Phosphorus Extraction Methods Applied to 50 Agricultural Soils from Central Europe. *Plant Soil Environ.* **2015**, *61*(2), 86–96. https://doi.org/10.17221/932/2014-PSE

[12] David, D. J. The Determination of Exchangeable Sodium, Potassium, Calcium and Magnesium in Soils by Atomic-Absorption Spectrophotometry. *Analyst* **1960**, *85*(1012), 495–503. https://doi.org/10.1039/AN9608500495

[13] GLOBE. About GLOBE Global Learning and Observations to Benefit the Environment. *Available at*: www.globe.gov/about/overview, accessed August 2022.

[14] Lindenmayer, D. B.; Laurance, W. F. The Ecology, Distribution, Conservation and Management of Large Old Trees. *Biol. Rev.* **2017**, *92*(3), 1434–1458. https://doi.org/10.1111/brv.12290

[15] Nuraniputri, U.; Daryanto, H. K. S.; Kuntjoro, K. Produksi Manggis pada Beberapa Kelompok Umur Tanaman dan Faktor-Faktor yang Mempengaruhi Produksi Manggis di Kabupaten Sukabumi, Jawa Barat. *J. Agribisnis Indones. (J. Indones. Agribus.)* **2016**, *4*(1), 67–78. https://doi.org/10.29244/jai.2016.4.1.67-78

[16] Wong Pui Yan, P. Shaping Pathways of the Musang King and Black Thorn in the Penang Island Durian Industry. *Continuum* **2025**, *39*(2), 218–231. https://doi.org/10.1080/10304312.2024.2338477

[17] Sundari, S.; Mas'ud, A.; Arumingtyas, E. L.; Hakim, L.; Azrianingsih, R.; Wahyudi, D. Taxonomical Status of Local Durian (*Durio* spp.) From Ternate Island, North Maluku Based on Morphological Character and Geographical Factor. *Int. J. Conserv. Sci.* **2019**, *10*(4), 711–720.

[18] Boonyopakorn, P.; Bualeard, P. Applying Neuro-Fuzzy System to Analyze Durian Minerals within Soil for Precision Agriculture. *Proceedings of the 17th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, IEEE, **2020**; pp 135–138.

[19] Belgis, M.; Wijaya, C. H.; Apriyantono, A.; Kusbiantoro, B.; Yuliana, N. D. Volatiles and Aroma Characterization of Several *Lai* (*Durio kutejensis*) and Durian (*Durio zibethinus*) Cultivars Grown in Indonesia. *Sci. Hortic.* **2017**, *220*, 291–298. https://doi.org/10.1016/j.scienta.2017.03.041

[20] Leštianska, A.; Fleischer, P.; Merganičová, K.; Fleischer, P.; Nalevanková, P.; Střelcová, K. Effect of Provenance and Environmental Factors on Tree Growth and Tree Water Status of Norway Spruce. *Forests* **2023**, *14*(1), 156. https://doi.org/10.3390/f14010156

[21] Elizabeth, M. N.; Syawal, H. F. Evaluation of Land Suitability for Durian (*Durio zibethinus*) Plants in Pangkatan District, Labuhanbatu Regency. *J. Pertan. Tropik* **2021**, *8*(2), 117–121. https://doi.org/10.32734/jpt.v8i2.6502

[22] Fischer, G.; Parra-Coronado, A.; Balaguera-López, H. E. Altitude as a Determinant of Fruit Quality with Emphasis on the Andean Tropics of Colombia. A Review. *Agron. Colomb.* **2022**, *40*(2), 212–227. https://doi.org/10.15446/agron.colomb.v40n2.101854

[23] Neha; Bhople, B. S.; Sharma, S. Seasonal Variation of Rhizospheric Soil Properties under Different Land Use Systems at Lower Shivalik Foothills of Punjab, India. *Agrofor. Syst.* **2020**, *94*(5), 1959–1976. https://doi.org/10.1007/s10457-020-00512-7

[24] Videla, X.; Parada, A. M.; Nario, A.; Pino, I.; Hood, R. Efecto del Contenido de Agua en la Mineralización Bruta e Inmovilización de Nitrógeno. *Agric. Téc.* **2005**, *65*(1), 74–78. http://dx.doi.org/10.4067/S0365-28072005000100008

[25] Rashied, T.; Aboagye, L. M.; Osafo, E. A.; Darko, R.; Dassah, A.; Obirih-Opareh, J. Effect of Tree Age on Fruit Characteristics, Seed Emergence and Seedling Growth in Rambutan (*Nephelium lappaceum* L.). *J. Hortic. Sci.* **2022**, *17*(1), 245–248. https://doi.org/10.25081/jhs.2022.v17.i1.991

[26] Ali-Dinar, H.; Mohammed, M.; Munir, M. Effects of Pollination Interventions, Plant Age and Source on Hormonal Patterns and Fruit Set of Date Palm (*Phoenix dactylifera* L.). *Horticulturae* **2021**, *7*(11), 427. https://doi.org/10.3390/horticulturae7110427

[27] Rajandran, T.; Tan, X. Y.; Tan, S. S.; Jamal, F. N. B. The Influence of Weather Patterns on Durian Yield in Malaysia. *Indian J. Appl. Res.* **2024**, *14*(5). https://doi.org/10.36106/ijar

[28] Somsri, S. *Durian in Thailand: A Success Story*; Asia-Pacific Association of Agricultural Research Institutions: Bangkok, Thailand, **2018**; pp 5–37.

[29] Kurniadinata, O. F.; Wenpei, S. Z.; Zaini, A. Six Potential Superior Durian Plants Resulted by Cross Breeding of *D. zibethinus* and *D. kutejensis* from East Kalimantan, Indonesia: Initial Identification. *J. Trop. Hortic.* **2019**, *2* (2), 45–49. https://doi.org/10.33089/jthort.v2i2.24

[30] Ketsa, S.; Wisutiamonkul, A.; Palapol, Y.; Paull, R. E. The Durian: Botany, Horticulture, and Utilization. *Hortic. Rev.* **2020**, *47*, 125–211. https://doi.org/10.1002/9781119625407.ch4

[31] Isa, I. Z. M.; Mustafa, N. A. A.; Shahril, M. Z.; Abdul, A.; Hamid, N. H. Z. A.; Dek, M. S. P. Proximate Composition, Physicochemical, and Antioxidant Properties of Three Durian (*Durio zibethinus*) Cultivars. *J. Trop. Plant Physiol.* **2019**, *11*(2), 22–30.

[32] Sritongchuay, T.; Kremen, C.; Bumrungsri, S. Effects of Forest and Cave Proximity on Fruit Set of Tree Crops in Tropical Orchards in Southern Thailand. *J. Trop. Ecol.* **2016**, *32*(4), 269–279. https://doi.org/10.1017/S0266467416000353

[33] Nguyen, T. V.; Mitlohner, R.; Bich, N. V.; Do, T. V. Environmental Factors Affecting the Abundance and Presence of Tree Species in a Tropical Lowland Limestone and Non-Limestone Forest in Ben En National Park, Vietnam. *J. For. Environ. Sci.* **2015**, *31*(3), 177–191. https://doi.org/10.7747/JFES.2015.31.3.177

[34] Bahadori, S.; Torabi Giglou, M.; Esmaielpour, B.; Dehdar, B.; Estaji, A.; Hano, C.; Gohari, G.; Vergine, M.; Vita, F. Antioxidant Compounds of Potato Breeding Genotypes and Commercial Cultivars with Yellow, Light Yellow, and White Flesh in Iran. *Plants* **2023**, *12*(8), 1707. https://doi.org/10.3390/plants12081707

[35] Ramli, M. S. A.; Zainal Abidin, M. S.; Md Reba, M. N.; Pui Boon Hean; Abd Rahman, M. A.; Lim Way Foong; Keshinro Kazeem Kolawole; Rizqi Andry Ardiansyah. Improving Water-Efficient Irrigation in Terrain *Durio zibethinus* Farming Using Hybrid Ant Colony Optimization-Based Soil

Moisture Prediction Model. *ELEKTRIKA-J. Electr. Eng.* **2024**, *23*(2), 32–43. https://doi.org/10.11113/elektrika.v23n2.537

[36] Suciati, A. Anti-Diabetes Activity and Phytochemical Analysis of Durian Fruit Methanol Fraction. *Int. J. Islam. Complement. Med.* **2023**, *4* (1), 37–44. https://doi.org/10.55116/IJICM.V4I1.59

[37] Rizky, A. M.; Anastasia, D.; Merdekawati, L. E. The Potential Antibacterial Power of Ethanol Extract of Durian Peel (*Durio zibethinus* Murr.) Against *Enterococcus faecalis*. *Sriwijaya J. Dent.* **2020**, *1*(1), 1–12.

[38] Octiara, E.; Meliala, C. P.; Sikumbang, L. Antibacterial Activity of Durian Peel Ethanol Extract (*Durio zibethinus* Murr.) Against *Streptococcus mutans* and *Enterococcus faecalis*. *Biomed. Pharmacol. J.* **2023**, *16*(2). https://dx.doi.org/10.13005/bpj/2670

[39] Charoenphun, N.; Klangbud, W. K. Antioxidant and Anti-Inflammatory Activities of Durian (*Durio zibethinus* Murr.) Pulp, Seed, and Peel Flour. *PeerJ* **2022**, *10*, e12933. https://doi.org/10.7717/peerj.12933

[40] Sang-Ngam, S.; Promden, W.; Sangphech, N.; Chukaew, A.; Sintupachee, S. Anti-Inflammatory and Antioxidant Properties of Crude Methanol Extracts from the Flowers and Peels of Khlong Seng Local Durian (*Durio zibethinus* Murr.). *Trends Sci.* **2024**, *21*(7), 7793–7793. https://doi.org/10.48048/tis.2024.7793

ASEAN
Journal of Scientific and Technological Reports
Online ISSN:2773-8752

Research article

Plant-Based Feed and Forage Research for Livestock: A Bibliometric Review

Chona P. Napalinga[1*]

[1] Faculty of College of Agriculture, Forestry, and Food Science, University of Antique-Hamtic Campus, 5715, Philippines
* Correspondence: chona.napalinga@antiquespride.edu.ph

Citation:
Napalinga, C. Plant-based feed and forage research for livestock: a bibliometric review. *ASEAN J. Sci. Tech. Report.* **2025**, *28*(5), e259394. https://doi.org/10.55164/ajstr.v28i5.259394

Article history:
Received: May 17, 2025
Revised: August 25, 2025
Accepted: September 1, 2025
Available online: September 14, 2025

Abstract: This bibliometric review research analyzes trends related to "Plant-Based Feed and Forage for Livestock," utilizing research data from 2013 to 2024. Eighty-two (82) English language journal articles indexed in Scopus were utilized as data. This study focuses on publications in agricultural and environmental sciences. The study aimed to identify all keywords, author keywords, authors, and countries related to research on plant-based feed and forage. It was calculated using VOSviewer to determine density, network, and overlay. This highlights the growing importance of sustainable alternative feed. This bibliometric review highlights the role of plant-based feed and forage in contributing to a sustainable livestock production system, supporting Sustainable Development Goal 2 (Zero Hunger) by enhancing food security through resilient and efficient resource feeding strategies. Shifting to resource-intensive conventional feeds that are environmentally friendly alternatives contributes to SDG 12 (Responsible Consumption and Production). Additionally, reducing greenhouse gas emissions through adaptation of plant-based feed for livestock production, aligned with SGD 13 (Climate Action), and promoting land use practices that reduce deforestation and preserve biodiversity, supports SDG 15 (Life on Land). Mapping the co-citation network and overlaying this research topic provides insights into the development and dissemination of new knowledge in agriculture. Analysis identifies emerging trends and gaps in current research, suggesting areas for future investigation. Overall, this bibliometric review provides a comprehensive overview of the diverse research landscape on plant-based feed and forage for livestock production, specifically for ruminant animals, in temperate and tropical regions. It emphasizes alignment between this research and sustainable goals.

Keywords: Plant-based feed; forage crops; bibliometric review; VOSviewer; sustainable livestock feed

1. Introduction

Sustainable agricultural practices have increasingly been a focus of the global livestock industry today, as it addresses environmental concerns and meets the growing demand for food, produced ethically. The shift prompted significant interest in using plant-based feed and forage for livestock production, which could be a promising solution to reduce the environmental impact of the traditional livestock feed system [1]. Conventional feed production relies heavily on grain and soy, which contribute to deforestation, biodiversity loss, and significant greenhouse gas emissions. In Brazil, mechanized soy farming in

Rondônia and Mato Grosso has led to deforestation, resulting in the release of up to 0.77 tonnes of CO_2 equivalents for every tonne of soy exported. Between 2010 and 2015, this totalled 223.46 million tonnes [2, 3]. Cereal grains such as wheat and barley contribute to soil disturbance and nitrous oxide emissions from fertilizers [4]. In South Asia, increasing CO_2 emissions from cereal farming endanger long-term productivity [3]. This result encouraged the researchers, policymakers, and industry stakeholders to explore plant-based alternatives to mitigate negative environmental impacts while maintaining animal productivity [1].

The potential of plant-based feed and forage research is increasingly recognized as a means to enhance sustainable livestock production. Alfalfa and clover are leguminous plants that contain a lot of crude protein (8.22%–22.19%), which is essential for animal growth and reproduction [5, 6]. Grasses provide a significant amount of crude fiber (32.06%–32.92%), which aids in digestion [7]. Mixed cropping increases neutral and acid detergent fiber, which is good for the health of the rumen [8]. These forages also provide macro minerals, such as Ca, Mg, K, and P, as well as micronutrients like Zn, Cu, Fe, and Mn [6, 9]. Legumes also contain condensed tannins that help animals utilize protein more efficiently and eliminate less nitrogen, which is beneficial for their health and the environment [10-11]. These essential nutrients improve livestock nutrition, enhance soil fertility, and promote biodiversity in the area. These leguminous plants are known for their nitrogen-fixing properties, which help enrich the soil's organic matter and reduce the need for synthetic fertilizers that can degrade soil health. Plant-based forage is a type of plant that serves as an efficient source of forage feed for livestock, requiring only a small amount of water to grow. In contrast, traditional grain-based feeds require more water, which can contribute to a more sustainable agricultural practice [12].

It has been shown that incorporating plant-based feed alternatives can improve the digestion process and reduce the incidence of diseases in livestock. Phytogenic compounds enhance mitochondrial function and energy biosynthesis, facilitating effective nutrient absorption [13]. Dietary fibers derived from plant by-products also affect protein bioavailability and hydrolysis, which are crucial for nutrient absorption [14]. Furthermore, the bioactive compounds in these feeds exhibit antimicrobial, antioxidant, and immunoregulatory characteristics that enhance immune responses and reduce oxidative stress [15, 16]. These substances help reduce detrimental intestinal microbes and promote gut health [16]. Importantly, phytogenic additives can mimic the effects of antibiotic growth promoters—now prohibited in numerous areas—without fostering antibiotic resistance [17, 18]. This study can serve as a basis for future research on exploring different forage species in terms of their nutrients, adaptability, and suitability for various climatic conditions, thereby informing sustainable farming practices. It was also highlighted that forage crops play a crucial role in reducing soil erosion, enhancing water retention, and promoting carbon sequestration, which is essential for combating climate change and fostering a resilient agricultural system. He further explained that plant-based feeds are being explored for their potential in reducing methane gas emission from ruminants and to further develop them as more environmentally friendly feed alternatives for livestock animals compared to traditional grain feed production in animal production systems [19]. Moreover, it has been observed that land planted with various forages for livestock has expanded significantly in recent decades. This may be because of the growing interest in sustainable livestock farming. Certain plant species are being studied for their potential adaptability and contribution to soil health conditions [20].

Plant-based feed and forage research is closely aligned with the objectives of organic farming and sustainable livestock production. Plant-based feeds and forage are produced without the use of synthetic fertilizers and pesticides to control pests or fertilize the soil, thereby reducing the ecological footprint in livestock production. Animal feed produced organically promotes sustainability by offering quantifiable financial and environmental advantages. Organic systems prevent pollution-causing synthetic inputs and lessen eutrophication and acidification per land unit [21, 22]. Forage legumes improve soil health by enhancing nitrogen fixation and carbon sequestration [23]. Livestock production accounts for only 18% of calories and 25% of protein, yet it consumes 70–80% of the world's agricultural land. Organic feed helps maximize land use by lowering dependency on feed that competes with food [24, 25]. Furthermore, organic methods enhance the management of water footprints [26]. These practices support a healthier ecosystem and meet the demand for sustainability in animal product production, contributing to advancements in achieving sustainable agricultural production globally [12]. Furthermore, plant-based feed and forage offer various environmental benefits, including sustainable intensification, improved food security, and reduced greenhouse gas emissions [27].

Improved forage corps integration in mixed production systems can contribute to the restoration of degraded land areas, enhance climate resilience, and reduce greenhouse gas emissions per unit of livestock products [28]. Forage production systems can accumulate soil carbon, and leguminous forages have a positive effect on soil nitrogen retention compared to cereal grain crops [29]. Nutritionally, forage crops are key determinants of livestock growth performance, reproduction, behavior patterns, productivity, and end-product quality [30]. Improved forage germplasm has shown significantly higher herbage yield productivity, and feeding regimes with improved leguminous forages have resulted in increased milk yield and dry matter intake in ruminant animals [31]. Specialized forage growing and livestock feeding activities can enhance smallholder welfare and farm income, contributing to the economic importance of livestock production [32]. The cool-season forage crops, such as fodder radish, offer a promising alternative for autumn and winter forages [33]. The Livestock Plus approach in livestock feed production, which includes sowing improved forages, can lead to the sustainable intensification of mixed systems, generating multiple benefits [28]. However, challenges that hinder livestock producers include a scarcity of quality feed, land tenure issues, limited access to resources, weak institutional frameworks, poor infrastructure, and environmental degradation [34]. Opportunities lie in the potential of improved forages to play a central role in sustainable intensification, which requires multidisciplinary approaches to quantify synergies and trade-offs [31]. Gene banking to conserve forage germplasm is facing challenges today; these challenges need to be addressed for immediate action to prevent loss and improve livestock systems [35]. Furthermore, research and development focused on maximizing the potential of plant-based feed and forage for sustainable livestock production.

This bibliometric review aims to analyze research trends related to plant-based feed and forage, utilizing VOSviewer as the tool for calculation. This is a quantitative research method used to evaluate the structure and development in a thematic area by examining Scopus publication data, citation patterns per author, and the network per author. The Scopus database is used to identify the most influential authors, institutions, and research topics related to the field of forage and pasture crop production. Additionally, this review will provide insights into how research on plant-based feed and forage aligns with broader sustainable agricultural goals in livestock production. This bibliometric review will contribute to the understanding of the scientific landscape surrounding plant-based feed and forage research, and this will offer a valuable insight into the progression of knowledge in this critical area of sustainable agriculture.

2. Materials and Methods

A bibliometric database of "Plant-Based Feed and Forage Research for Livestock" was constructed using the Scopus bibliographic database. A keyword search was performed to identify "Plant-Based Feed and Forage Research for Livestock" related studies published between 2013 and 2024, using the following flexible retrieval conditions: "plant-based feed" OR "plant protein" OR "alternative feed sources" OR "forage crops" OR "non-traditional feed" OR "fodder plants" AND "livestock nutrition" OR "animal nutrition" OR "ruminant feed" OR "herbivorous livestock" OR "livestock feeding systems" OR "feed efficiency" AND "sustainable feed" OR "agricultural sustainability" OR "livestock feed innovation" to locate publications that contained these words in their titles, abstracts, or keywords list. The results of the search in Scopus were filtered, yielding 82 records. Following the conventions used in other bibliometric studies, further analysis was restricted to the year range of 2013 to 2024, specifically the subject areas of Agricultural and Biological Sciences and Environmental Science. Furthermore, the document type was focused only on published articles, and the language used was English. The articles were analyzed and exported into CSV format, then imported into the bibliometric analysis tool VOSviewer, and several computations were performed. These computations involved identifying countries, authorship patterns, all keywords, and author keyword networks. VOSviewer version 1.6.20, released in October 2023 and designed by Ness Jan van Eck and Ludo Waltman, was used for co-occurrence maps and keyword analysis.

3. Results and Discussion

3.1 Results Based on VOSviewer Computations for All Keywords: Density, Network, and Overlay

VOSviewer computation for all keywords. The density shown in Figure 1 encompasses all keywords related to "forage" and "forage quality", their evolution, and interconnections, highlighting the key subjects in forage research. The density plot highlights the dominant role of forage and forage quality, and shows fitting clusters in crop productivity, forage types, nutrient value, environmental factors, and geographical relevance, particularly in China. In addition, the network map, presented in Figure 2, illustrates intersectional connections across research fields, demonstrating that forage studies encompass disciplines such as agriculture, environmental sciences, and animal nutrition. The specific clusters identified revolve around yield (e.g., intercropping, maize), nutritional quality (e.g., nitrogen, protein, legumes), and environmental challenges (e.g., climate change). In addition, as reflected in Figure 3, the overlay map illustrates longitudinal trends. Early publications (2019) focused on physiology and botanical aspects, while more recent articles (2020-2021) have addressed sustainability and agronomic innovations to meet global challenges. Targeting sustainability and agronomic innovations to address global challenges. Emerging interests include adapting forage systems to semi-arid regions, integrating sustainable practices, and enhancing resilience to climate variability. Throughout, the research remains interconnected across disciplines, bridging agronomy, ecology, and livestock nutrition to optimize forage yield and quality while tackling modern agricultural and environmental issues.

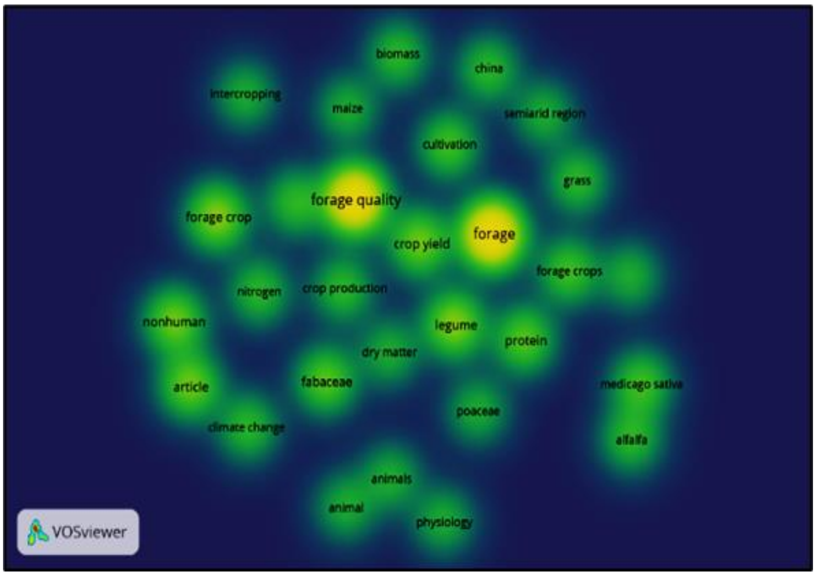


Figure 1. VOSviewer Computations for All Keywords Density.

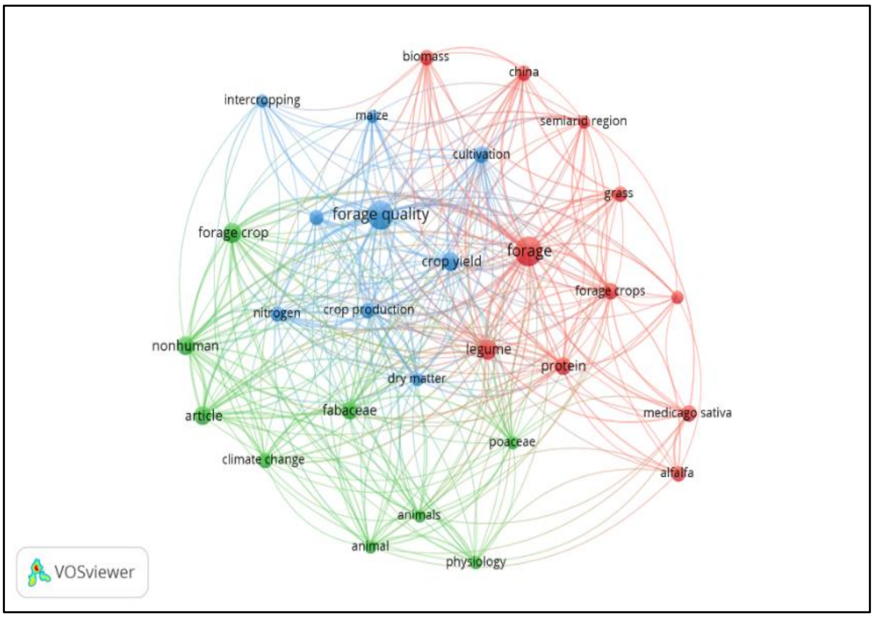


Figure 2. VOSviewer Computations for All Keywords Network.

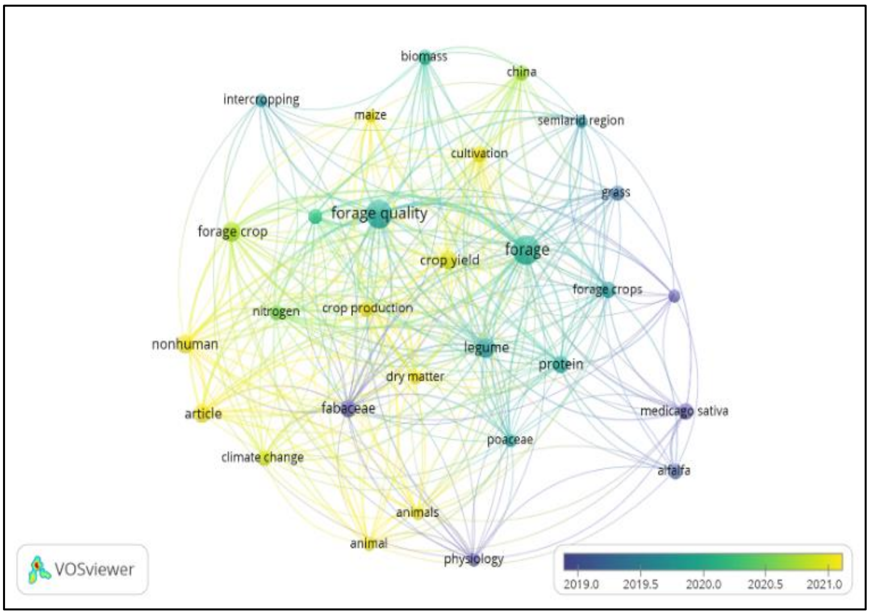


Figure 3. VOSviewer Computations for All Keywords Overlay.

3.2 Results Based on VOSviewer Computations for Authors Density, Network, and Overlay

The VOSviewer author visualizations highlight the density, network, and overlay of contributions and collaborations among researchers, focusing on figures like Devendra Singh Ginwal, R.K. Meena, Rakesh Kumar, Uttam Kumar, and Hardev Ram. As shown in Figure 4, the density map reveals high citation or activity frequency, with yellow gradients marking central authors and green to blue gradients representing decreasing influence. Figure 5 shows the network map, which features strong co-authorship connections, indicated by close positioning and thick connecting lines, suggesting frequent collaborations within small, focused groups. Authors like Ginwal and Meena appear central, reflecting their significant collaborative roles. Moreover, Figure 6 displays an overlay visualization that highlights recent contributions using a color gradient from blue (older activity) to yellow (recent publications), indicating these authors' ongoing engagement in their field. These maps collectively underline the interconnectedness of the authors' work, clustering them into research subgroups, while showcasing their influence through citation density and collaboration patterns. Such visualizations are valuable for identifying key contributors, understanding collaborative dynamics, and exploring potential partnerships in the research community. The study by Ginwal [36] aligns with VOSviewer bibliometric analyses, which provide empirical evidence to optimize intercropping ratios to enhance forage quality and nutrient yields. Frequent collaboration and citation in the author density map show influence in sustainable forage research. A tightly connected cluster among five authors was shown in the co-authorship network visualization, indicating an interdisciplinary synergy between them. Topical relevance was highlighted in overlay visualizations with the keyword "forage quality," which is trending in sustainable agriculture. These VOSviewer results strengthen the significance of the study, showing that the strategic collaboration between authors and the thematic focus of this research contribute to advancing livestock nutrition and resilient cropping systems. The research findings of these five authors also align with the Sustainable Development Goals. SDG 2 (Zero Hunger) promotes food security and improved nutrition, while SDG 12 (Responsible Consumption and Production) encourages resource-efficient feeding strategies. Additionally, it supports SDG 13 (Climate Action) through the use of a climate-resilient, organic, and legume-based feed production system, which reduces greenhouse gas emissions and enhances soil carbon sequestration. Lastly, through the protection of the ecosystem and biodiversity, it supports SDG 15 (Life on Land). The sustainable plant-based feed and forage production systems help reduce land degradation and promote biodiversity through diversified cropping and reduced chemical use.

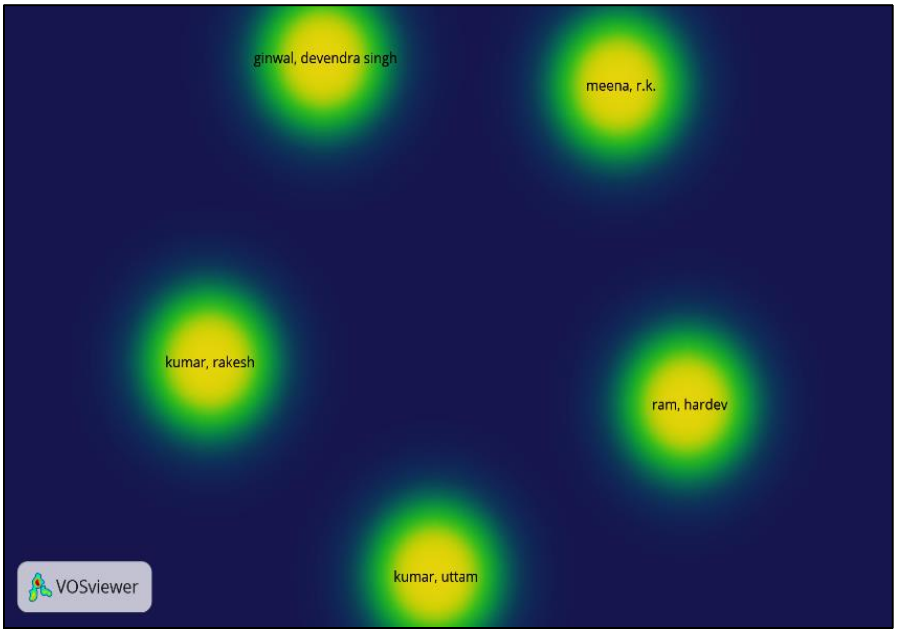


Figure 4. VOSviewer Computations for Authors Density.

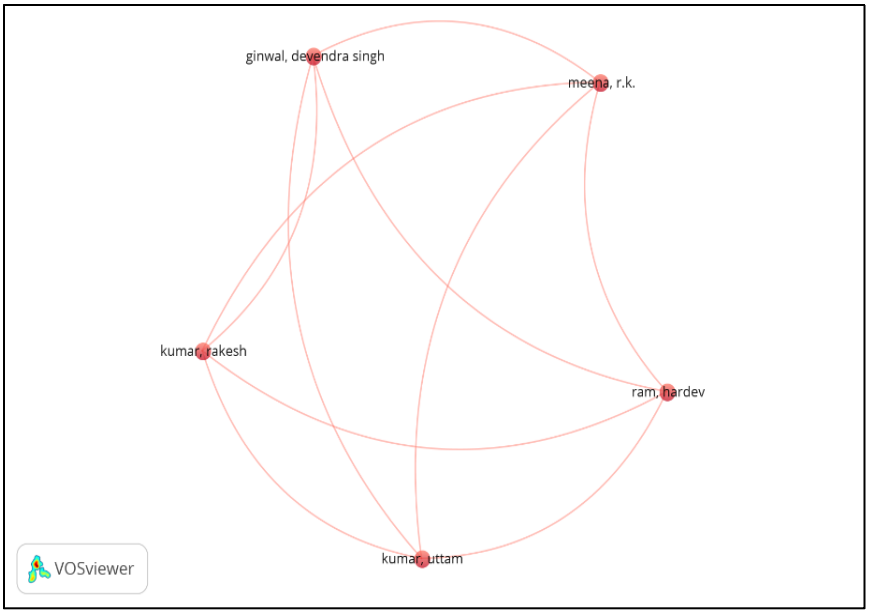


Figure 5. VOSviewer Computations for Authors Network.

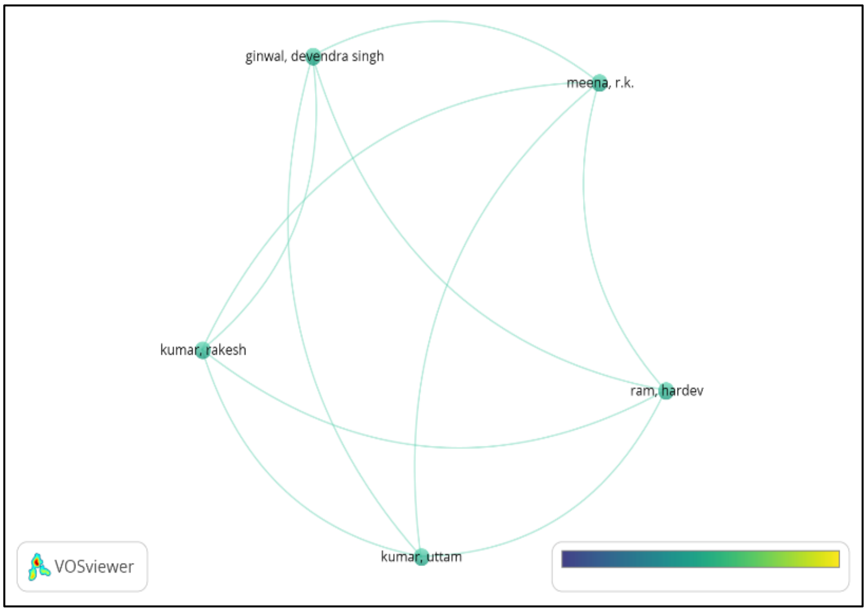


Figure 6. VOSviewer Computations for Authors Overlay.

3.3 Results Based on VOSviewer Computations for Authors' Keywords Density, Network, and Overlay

The VOSviewer visualizations for the author's keyword density, network, and overlay maps highlight the central themes, collaborations, and trends in forage research. Presented in **Figure 7** are the Author's visualizations, which reveal key contributors such as Devendra Singh Ginwal and R.K. Meena, whose prominence reflects frequent citations and collaborations, with interconnected clusters indicating strong co-authorship within focused research groups. Keyword density maps highlight terms such as "forage crop," "forage crops," and "forage quality," with "forage crop" exhibiting the highest prominence, indicating its central role in the research dataset. Furthermore, **Figure 8** shows the keyword network maps, which reveal limited connectivity between these terms, suggesting underexplored links between forage crop types and their quality. The overlay maps presented in Figure 9 showed that they trace the temporal evolution of research, showing that while "forage crop" is an established topic, "forage quality" has gained prominence more recently, highlighting a shift toward evaluating nutritional aspects. Weak links between terms suggest research gaps, offering opportunities to study the integration of crop productivity and forage quality more explicitly. Overall, the visualizations underscore trends in forage research while identifying key contributors, core terms, and potential areas for interdisciplinary or integrated studies. This analysis can guide future research, particularly in connecting forage crop studies with quality improvements for sustainable agriculture and animal nutrition.

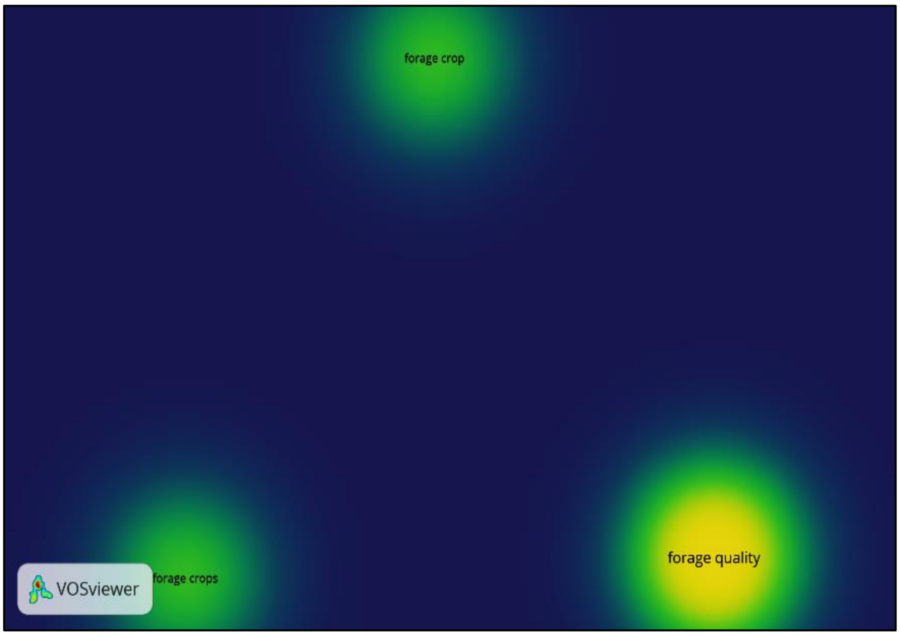


Figure 7. VOSviewer Computations for Authors' Keywords Density.

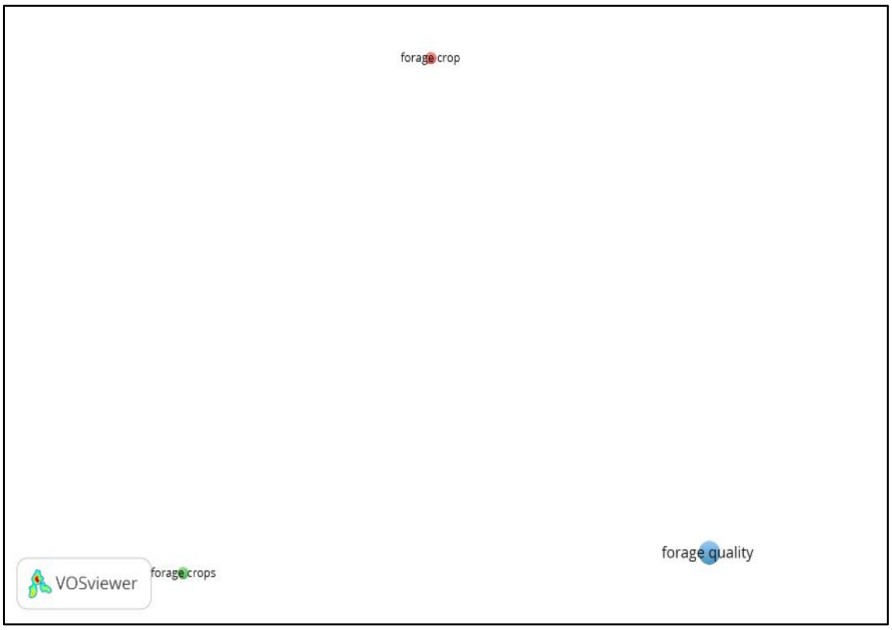


Figure 8. VOSviewer Computations for Authors' Keywords Network.

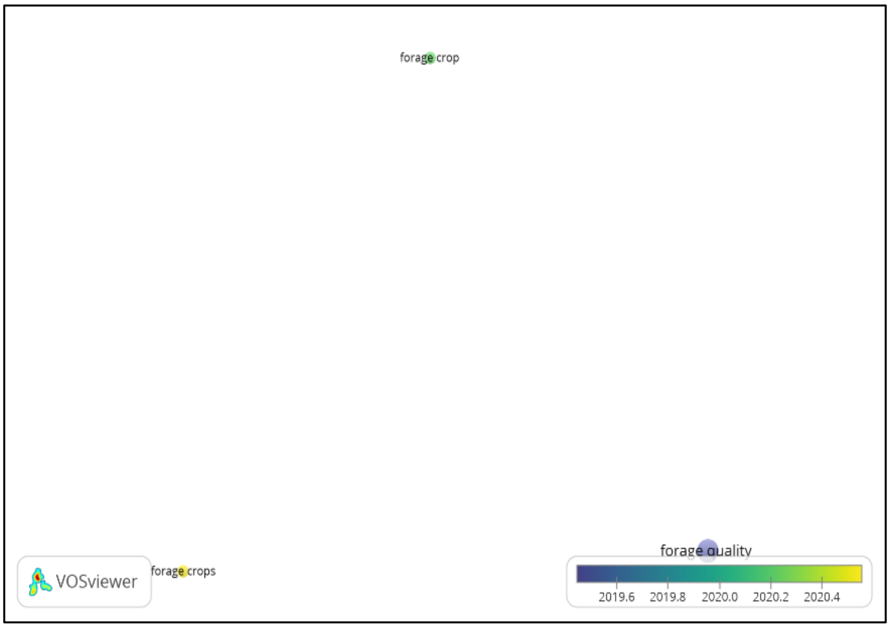


Figure 9. VOSviewer Computations for Authors Keywords Overlay.

Keywords and their number of occurrences are shown in Table 1. It shows that the most frequent keyword term was “forage quality” with 17 occurrences, followed by the terms "forage crop" and “forage crops” with 5 occurrences each.

Table 1. The ranking of the most frequently occurring keywords for the terms “Plant-based feed” and “Forage”.

Keyword	Occurrences
Forage crop	5
Forage crops	5
Forage quality	17

3.4 Results Based on VOSviewer Computations for Countries Density, Network, and Overlay

The VOSviewer visualization highlights the prominence of the United States and China as leading contributors to the analyzed research dataset, with their influence depicted through node size and intensity in a density, network, and overlay map. As shown in **Figures 10, 11, and 12, the United States and China were countries that were both distinctly visible, indicating their significant involvement; yet,** the spatial separation between their nodes suggests limited direct collaboration or co-authorship. Color gradients reflect temporal trends, with activity spanning from 2019 to 2021, where the United States and China dominate individually but lack a strong interconnection. This separation highlights regional silos, potentially indicating missed opportunities for synergy between these global research leaders. The heatmap effect, transitioning from yellow (high intensity) to cooler colors (blue), further reinforces their substantial, yet isolated, contributions. Such a visualization underscores the need for fostering international cooperation, as bridging the collaboration gap between the United States and China could amplify innovation and accelerate progress in the studied field. This analysis provides strategic insights for policymakers and researchers, encouraging efforts to enhance global research networks and capitalize on the strengths of these two major players.

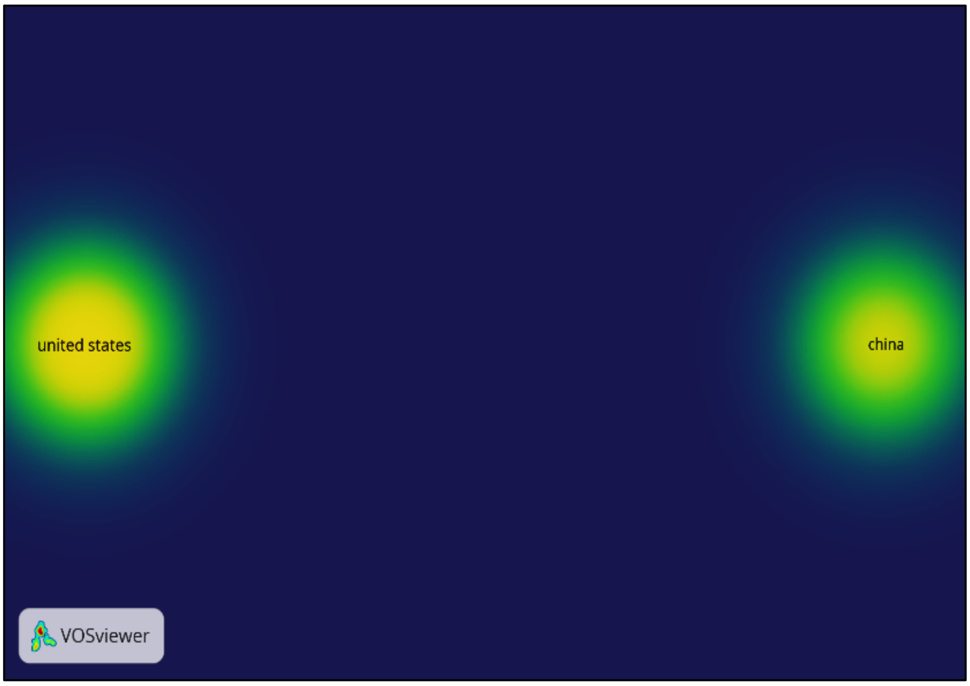


Figure 10. VOSviewer Computations for Countries Density.

Figure 11. VOSviewer Computations for Countries Network.

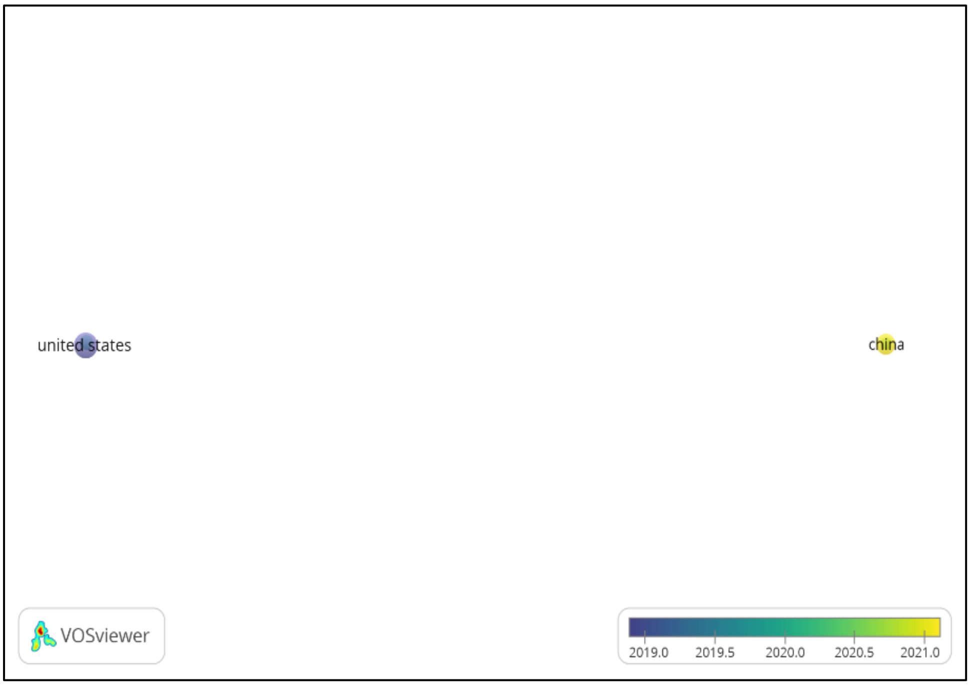


Figure 12. VOSviewer Computations for Countries Overlay.

Data on the countries, along with their corresponding documents and citations, is shown in Table 2. Among the four countries, the United States had the highest number of documents submitted (21 documents) with 387 citations, followed by China with 13 documents and 189 citations. Furthermore, Korea has 7 documents with 100 citations, while Pakistan has 5 documents with 41 citations, respectively.

Table 2. Countries with corresponding documents and citations

Country	Documents	Citations
United States	21	387
China	13	189
South Korea	7	100
Pakistan	5	41

3.5. Ratio of articles per year, by authors, by source, by affiliations, by country, by type, and by subject area.

The bibliometric analysis reveals several key trends in research output across various dimensions. The temporal analysis of documents by year reveals fluctuating research activity between 2013 and 2024, as illustrated in Figure 13, with significant peaks in 2014 and 2022, followed by a slight decline in 2023 and 2024. The peak in 2022 may reflect increased research funding or interest in specific topics, whereas the decline in 2024 might be attributed to incomplete data or shifts in focus. Figure 14 presents the analysis of documents per year by source, which shows that journals such as Crop and Pasture Science and Plants dominate specific years, indicating topical priorities in agricultural and environmental sciences. Understanding these journal-specific trends can help researchers target high-impact publications and identify thematic shifts within the field. Moreover, Figure 15 shows the documents authored by the author, revealing that a core group, led by Shen, Y., is responsible for a significant portion of the research output. The document-by-author ratio was based on output volumes, highlighting those who publish the most. Meanwhile, the data on VOSviewer covers computations on author density, network, and overly discussed influence, collaboration, and relevance, identifying who is central, cited, and active in the field. That is why Shen, Y. ranked highest in terms of document ratio, but in the VOSviewer maps, Ginwal, Meena, Kumar, and Ram appeared more central due to their greater influence and stronger networks, even though they are less prolific compared to Shen Y. This suggests the presence of collaborative networks that drive consistent contributions. Institutional analysis, as shown in Figure 16, revealed that leading affiliations, such as Lanzhou University and the USDA Agricultural Research Service, are major contributors. These institutions are complemented by others from South Korea, Saudi Arabia, and Australia, reflecting global collaboration and cooperation. The United States leads in research output, followed by China, underscoring its leadership in scientific contributions, likely supported by robust funding and infrastructure, as shown in Figure 17. A clear trend toward publishing research findings as peer-reviewed journal articles, as presented in Figure 18, reveals that 100% of the documents classified as "Articles" demonstrate a scholarly emphasis on maintaining academic rigor and disseminating findings to a broader audience. Collectively, Figure 19 presents the key highlights of the analysis, which focus on agricultural and environmental research, significant international collaboration, and the dominance of high-quality journal publications in advancing knowledge in this field.

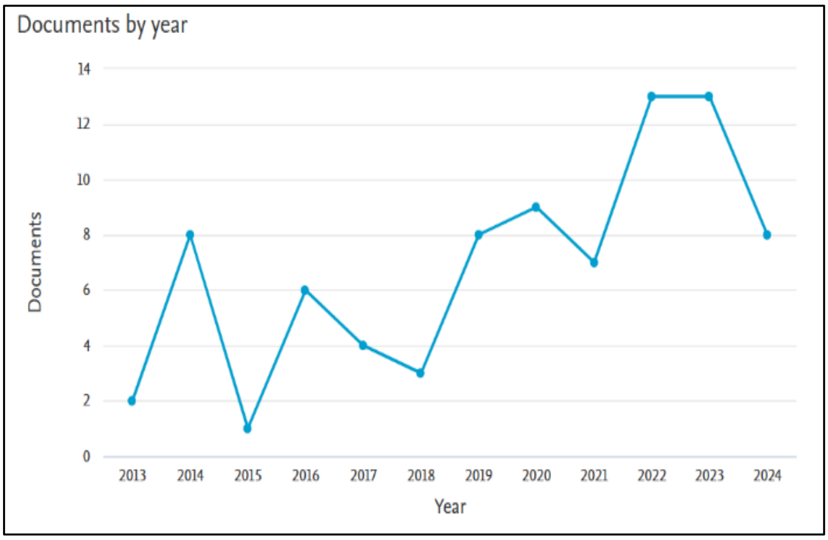


Figure 13. Ratio of articles per year.

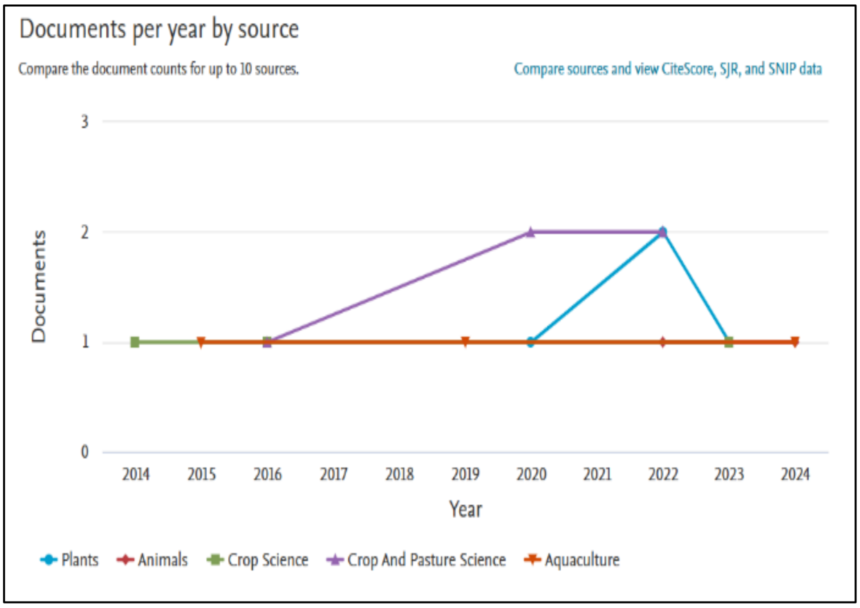


Figure 14. Ratio of articles by source.

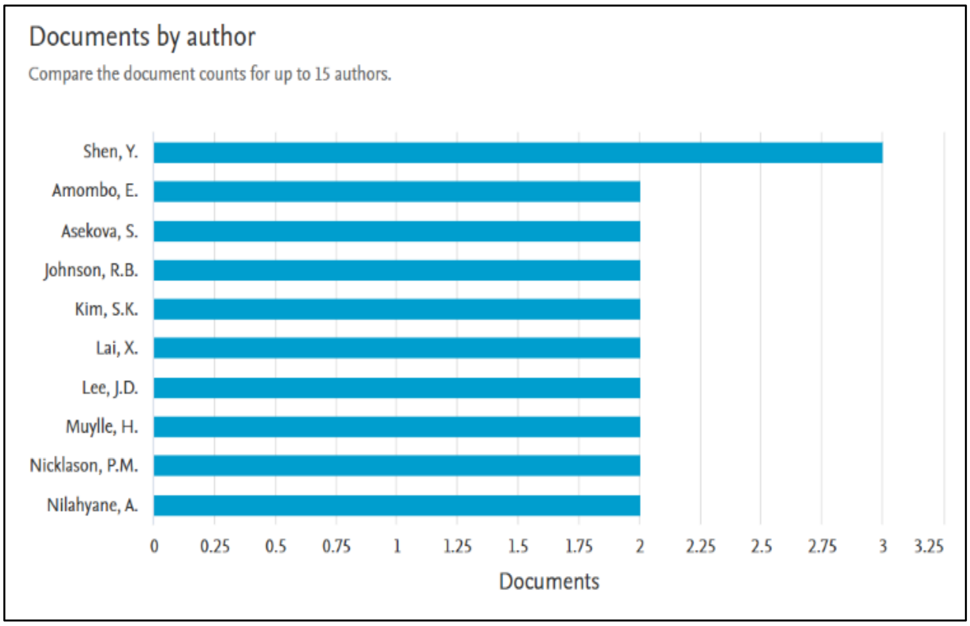


Figure 15. Ratio of articles by authors.

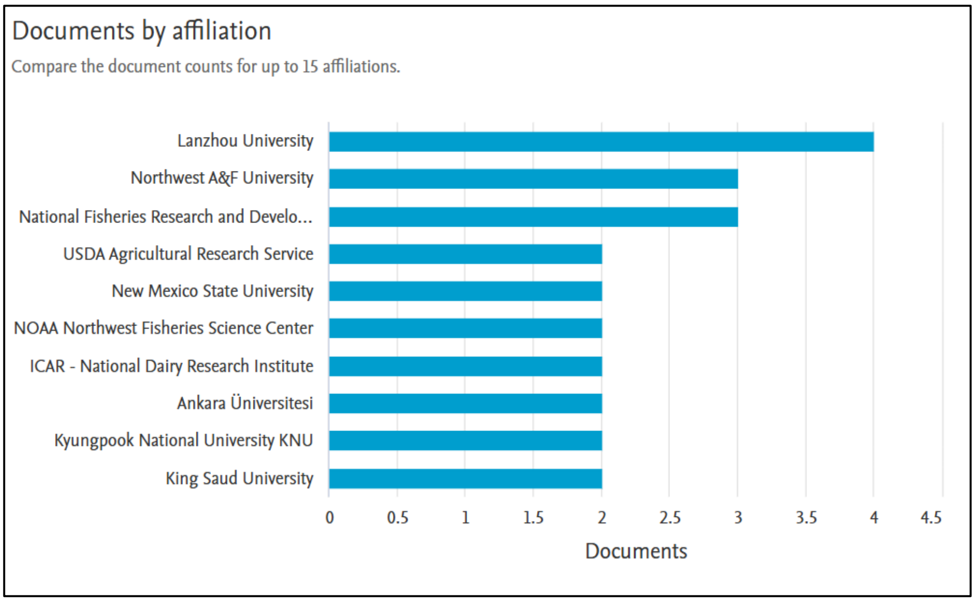


Figure 16. Ratio of articles by affiliations.

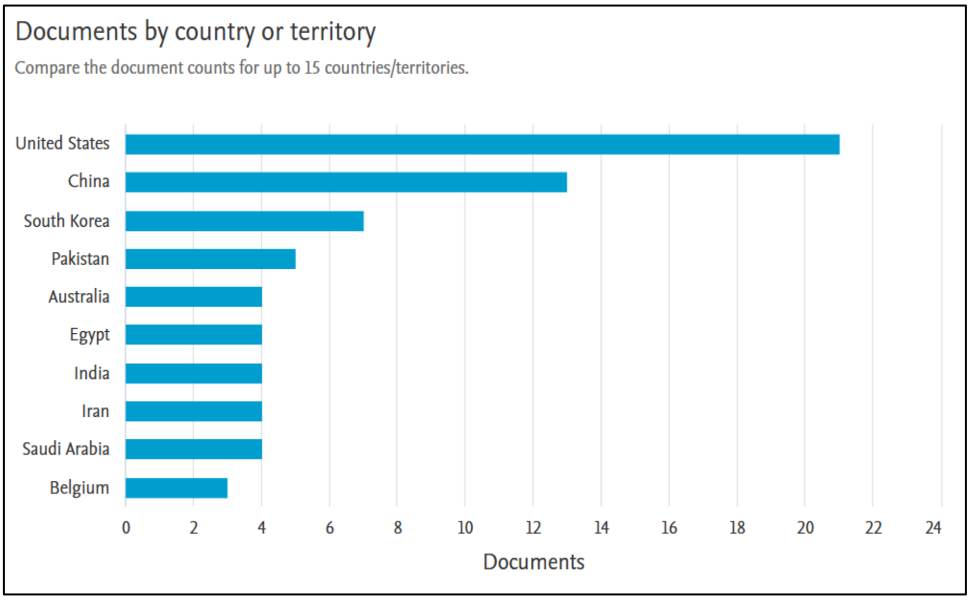


Figure 17. Ratio of articles by country.

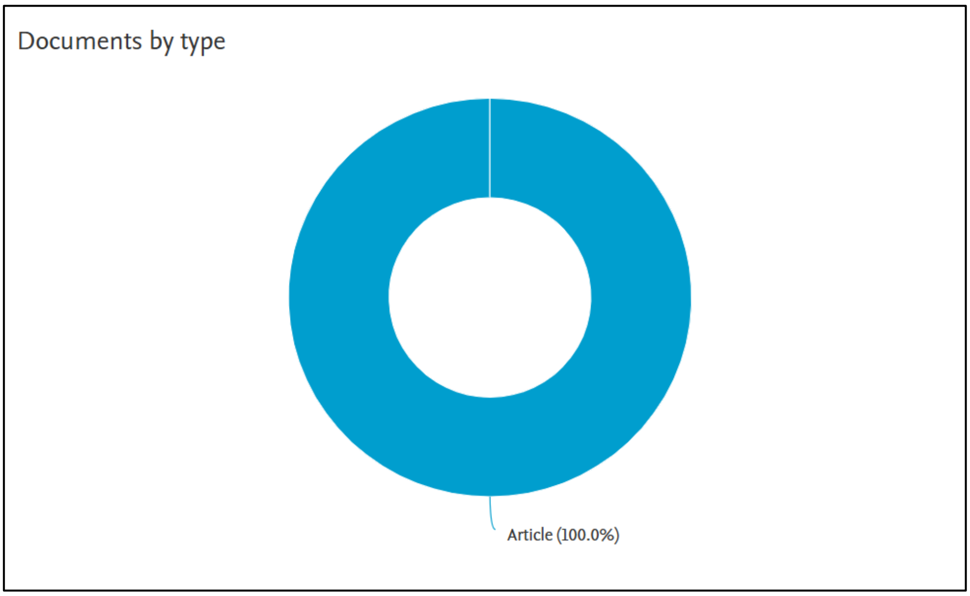


Figure 18. Ratio of articles by type.

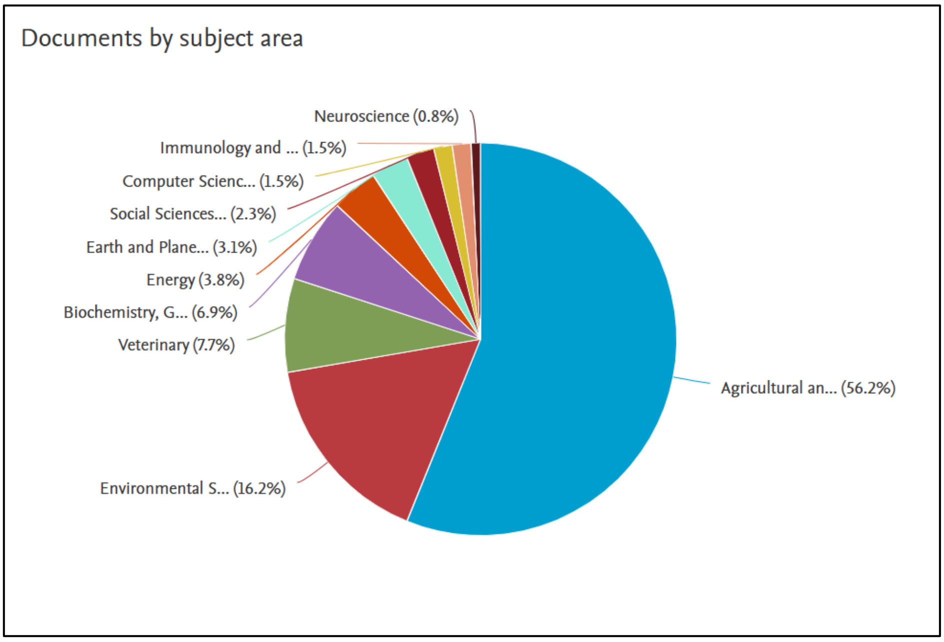


Figure 19. Ratio of articles by type.

4. Conclusions

The bibliometric review of plant-based feeds and forage research for livestock, gathered from 2013 to 2024, provides valuable insights into key trends, contributors, and thematic research development. This intensifies significant research endeavors that focus on the quality of forage, productivity, environmental sustainability, and other related aspects. VOSviewer visualizations reveal the robust interdisciplinary connections between agriculture, environmental science, and the field of animal production, addressing agricultural challenges such as climate change, crop resilience, and sustainability. Collaborative research by prominent authors such as Devendra Singh Ginwal and R.K. Meena was one of the key findings of this research. Forage crops and forage quality are the main highlights of the keyword analysis, which means there is an emerging interest in nutritional and sustainable practices in raising livestock. The findings suggest an opportunity for the deeper integration of crop production and forage quality in future research studies. The United States and China are the countries with the largest number of research studies related to plant-based feed and forage, highlighting the need for direct collaboration between these countries. Bridging the gap in collaboration between these two countries could foster innovation and thereby accelerate progress. Lanzhou University and the USDA Agricultural Research Service are the institutional leading contributors, demonstrating that robust infrastructure and international collaboration in advancing this field are very important. The analysis of research output over time shows that peaks occur in 2014 and 2022, corresponding to an increasing interest in the specific thematic area. Crop and Pasture Science journals dominate the publication outlets for research that emphasizes the importance of targeting platforms with significant impact. Overall, the results identify emerging trends, key contributors, and thematic priorities, which offer a future roadmap for research. To address the gap in collaboration between countries, it is essential to integrate the underexplored research area to enhance the global network further, optimizing forage research to support sustainable agricultural goals and livestock nutrition and production in a growing economy.

5. Acknowledgements

The author sincerely thank Professor Greta G. Gabinete for her invaluable guidance and support throughout the study.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] Dey, B.; Notenbaert, A.; Makkar, H.; Mwendia, S.; Rao, I. Realizing Economic and Environmental Gains from Forage Alternatives. *CAB Rev. Perspect. Agric. Vet. Sci. Nutr. Nat. Resour.* **2022**, *10*.

[2] Milne, E.; Cerri, C. E. P.; Carvalho, J. L. N. Agricultural Expansion in the Brazilian State of Mato Grosso; Implications for C Stocks and Greenhouse Gas Emissions. In *Environmental Science and Engineering*; Springer: Berlin, Heidelberg, **2010**; pp 447–460. https://doi.org/10.1007/978-3-642-00493-3_21

[3] Kibria, M. G.; Aspy, N. N.; Ullah, E.; Dewan, M. F.; Hasan, M. A.; Hossain, M. A.; Haseeb, M.; Hossain, M. E. Quantifying the Effect of Agricultural Greenhouse Gas Emissions, Food Production Index, and Land Use on Cereal Production in South Asia. *J. Clean. Prod.* **2023**, *432*, 139764. https://doi.org/10.1016/j.jclepro.2023.139764

[4] Hillier, J.; Walter, C.; Malin, D.; Garcia-Suarez, T.; Mila-i-Canals, L.; Smith, P. A Farm-Focused Calculator for Emissions from Crop and Livestock Production. *Environ. Model. Softw.* **2011**, *26*(9), 1070–1078. https://doi.org/10.1016/j.envsoft.2011.03.014

[5] Bo, P. T.; Bai, Y.; Dong, Y.; Shi, H.; Soe Htet, M. N.; Samoon, H. A.; Zhang, R.; Tanveer, S. K.; Hai, J. Influence of Different Harvesting Stages and Cereals–Legume Mixture on Forage Biomass Yield, Nutritional Compositions, and Quality under Loess Plateau Region. *Plants* **2022**, *11*(20), 2801. https://doi.org/10.3390/plants11202801

[6] Lebeloane, M. M.; Famuyide, I. M.; Elgorashi, E. E.; McGaw, L. J.; Kgosana, K. G. Evaluation of Minerals, Trace Elements, and Antinutritional Factors in Selected Legume Fodder Species (Fabaceae) with the Potential to Improve Cattle Nutrition and Gastrointestinal Health. *S. Afr. J. Bot.* **2024**, *171*, 120–128. https://doi.org/10.1016/j.sajb.2024.05.051

[7] Bo, P. T.; Dong, Y.; Zhang, R.; Htet, M. N. S.; Hai, J. Optimization of Alfalfa-Based Mixed Cropping with Winter Wheat and Ryegrass in Terms of Forage Yield and Quality Traits. *Plants* **2022**, *11*(13), 1752. https://doi.org/10.3390/plants11131752

[8] Kunelius, H. T.; Dürr, G. H.; McRae, K. B.; Fillmore, S. A. E. Performance of Timothy-Based Grass/Legume Mixtures in Cold Winter Region. *J. Agron. Crop Sci.* **2006**, *192*(3), 159–167. https://doi.org/10.1111/j.1439-037X.2006.00195.x

[9] Lindström, B. E. M.; Frankow-Lindberg, B. E.; Dahlin, A. S.; Wivstad, M.; Watson, C. A. Micronutrient Concentrations in Common and Novel Forage Species and Varieties Grown on Two Contrasting Soils. *Grass Forage Sci.* **2013**, *68*(3), 427–436. https://doi.org/10.1111/gfs.12006

[10] Jonker, A.; Yu, P. The Occurrence, Biosynthesis, and Molecular Structure of Proanthocyanidins and Their Effects on Legume Forage Protein Precipitation, Digestion and Absorption in the Ruminant Digestive Tract. *Int. J. Mol. Sci.* **2017**, *18*(5), 1105. https://doi.org/10.3390/ijms18051105

[11] Lagrange, S. P.; Macadam, J. W.; Villalba, J. J. The Use of Temperate Tannin Containing Forage Legumes to Improve Sustainability in Forage–Livestock Production. *Agronomy* **2021**, *11*(11), 2264. https://doi.org/10.3390/agronomy11112264

[12] Acar, R.; Koç, N.; Sumiahadi, A. Investigation of Yield, Yield Components, and Nitrogen-Fixing Ability of Wild Rocket (Diplotaxis tenuifolia) as a Forage Crop in Turkey. *Arab. J. Geosci.* **2019**, *12*(23), 740. https://doi.org/10.1007/s12517-019-4959-y

[13] Attia, Y. A.; El-Hack, M. E. A.; Alagawany, M. M.; Elnaggar, A. S. *Phytogenic and Phytochemical as Alternative Feed Additives for Animal Production*; Bentham Science Publishers: Sharjah, UAE, **2025**; pp 1–198. https://doi.org/10.2174/97898153227671250101

[14] Grundy, M. M. L.; Tang, J.; van Milgen, J.; Renaudeau, D. Cell Wall of Feeds and Their Impact on Protein Digestibility: An In Vitro Method Applied for Pig Nutrition. *Anim. Feed Sci. Technol.* **2022**, *293*, 115467. https://doi.org/10.1016/j.anifeedsci.2022.115467

[15] Georganas, A.; Giamouri, E.; Christodoulou, C.; Mavrommatis, A.; Zoidis, E.; Papadomichelakis, G.; Simitzis, P. E.; Tsiplakou, E.; Pappas, A. C.; Fegeros, K. Impact of Bioactive Compounds on Animals and the Environment. In *Bioactive Compounds and Their Importance*; Bentham Science Publishers: Sharjah, UAE, **2021**; pp 61–103.

[16] Guil-Guerrero, J. L.; Ramos, L.; Moreno, C.; Zúñiga-Paredes, J. C.; Carlosama-Yépez, M.; Ruales, P. Plant-Food By-Products to Improve Farm-Animal Health. *Anim. Feed Sci. Technol.* **2016**, *220*, 121–135. https://doi.org/10.1016/j.anifeedsci.2016.07.016

[17] Wallace, R. J.; Oleszek, W.; Franz, C.; Hahn, I.; Baser, K. H. C.; Mathe, A.; Teichmann, K. Dietary Plant Bioactives for Poultry Health and Productivity. *Br. Poult. Sci.* **2010**, *51*(4), 461–487. https://doi.org/10.1080/00071668.2010.506908

[18] Attia, Y. A.; Addeo, N. F.; Bovera, F.; Al-Hack, M. E. A.; AlBanoby, M. A.; Alhotan, R. A.; Khafaga, A. F.; Hafez, H. M.; de Oliveira, M. C. Phytogenic Substances as Novel Feed Supplements and Their Application in Livestock Nutrition. In *Phytogenic and Phytochemical as Alternative Feed Additives for Animal Production*; Bentham Science Publishers: Sharjah, UAE, **2025**; pp 4–18. https://doi.org/10.2174/9789815322767125010005

[19] Jiang, C.; You, Y.; Lai, X.; Zhang, Z.; Gao, W.; Ma, L. Maximizing Food Equivalent Unit Yield for Forage Crops. *Ind. Crops Prod.* **2024**, *218*, 118942.

[20] Mahmah, S.; Mebarkia, A.; Rekik, F. A Comparative Study on Narbon Vetch and Common Vetch. *J. Agric. Sci. (Belgrade)* **2023**, *68*(3), 263–270.

[21] Van Wagenberg, C. P. A.; De Haas, Y.; Hogeveen, H.; Van Krimpen, M. M.; Meuwissen, M. P. M.; Van Middelaar, C. E.; Rodenburg, T. B. Animal Board Invited Review: Comparing Conventional and Organic Livestock Production Systems on Different Aspects of Sustainability. *Animal* **2017**, *11*(10), 1839–1851.

[22] Espinoza-Villavicencio, J. L.; Palacios-Espinosa, A.; Ávila-Serrano, N.; Guillén-Trujillo, A.; De Luna-De La Peña, R.; Ortega-Pérez, R.; Murillo-Amador, B. Organic Livestock, an Alternative of Cattle Development for Some Regions of Mexico: A Review. *Interciencia* **2007**, *32*(6). https://www.scopus.com/inward/record.uri?eid=2-s2.0-35648991439

[23] Kumar, R.; Yadav, M. R.; Arif, M.; Mahala, D. M.; Kumar, D.; Ghasal, P. C.; Yadav, K. C.; Verma, R. K. Multiple Agroecosystem Services of Forage Legumes towards Agriculture Sustainability: An Overview. *Indian J. Agric. Sci.* **2020**, *90*(8), 1367–1377. https://doi.org/10.56093/ijas.v90i8.105882

[24] Scialabba, N. E.-H. Livestock and Future Food Supply Scenarios. In *Managing Healthy Livestock Production and Consumption*; Academic Press: Cambridge, MA, **2021**; pp 107–121. https://doi.org/10.1016/B978-0-12-823019-0.00011-8

[25] van Huis, A.; Gasco, L. Insects as Feed for Livestock Production: Insect Farming for Livestock Feed Has the Potential to Replace Conventional Feed. *Science* **2023**, *379*(6628), 138–139. https://doi.org/10.1126/science.adc9165

[26] Govoni, C.; Chiarelli, D. D.; Rulli, M. C. A Global Dataset of the National Green and Blue Water Footprint of Livestock Feeds. *Sci. Data* **2024**, *11*(1), 1419. https://doi.org/10.1038/s41597-024-04264-2

[27] Duan, C.; Yu, C.; Shi, P.; Huangqing, D.; Zhang, X.; Dai, E. Assessing Trade-Offs Among Productive, Economic, and Environmental Indicators of Forage Systems in Southern Tibetan Crop-Livestock Integration. **2023**.

[28] Chand, S.; Indu; Singhal, R. K.; Govindasamy, P. Agronomical and Breeding Approaches to Improve the Nutritional Status of Forage Crops for Better Livestock Productivity. **2022**. https://doi.org/10.1111/gfs.12557

[29] MacLeod, N.; Waldron, S.; Wen, S.-L. A Comprehensive Approach for Assessing the Economic Contribution of Forage and Livestock Improvement Options to Smallholder Farming Enterprises. **2015**. https://doi.org/10.1016/S2095-3119(15)61091-7

[30] Mbambalala, L.; Rani, Z. T.; Mpanza, T. D. E.; Mthana, M. S.; Ncisana, L.; Mkhize, N. R. Fodder Radish as a Potential Alternative Feed Source for Livestock in South Africa. **2023**. https://doi.org/10.3390/agriculture13081625

[31] Ates, S.; Cicek, H.; Bell, L. W.; Norman, H. C.; Mayberry, D. E.; Kassam, S.; Hannaway, D. B.; Louhaichi, M. Sustainable Development of Smallholder Crop-Livestock Farming in Developing Countries. **2018**. https://doi.org/10.1088/1755-1315/142/1/012076

[32] Rao, I.; Peters, M.; Castro, A.; *et al.* LivestockPlus – The Sustainable Intensification of Forage-Based Agricultural Systems to Improve Livelihoods and Ecosystem Services in the Tropics. **2015**. https://doi.org/10.17138/TGFT(3)59-82

[33] Maass, B. L.; Pengelly, B. C. Tropical and Subtropical Forage Germplasm Conservation and Science on Their Deathbed! **2019**. https://doi.org/10.1177/0030727019867961

[34] Paul, B. K.; Groot, J. C. J.; Maass, B. L.; Notenbaert, A. M. O.; Herrero, M.; Tittonell, P. A. Improved Feeding and Forages at a Crossroads: Farming Systems Approaches for Sustainable Livestock Development in East Africa. **2020**. https://doi.org/10.1177/0030727020906170

[35] Paul, B. K.; Koge, J.; Maass, B. L.; Notenbaert, A.; Peters, M.; Groot, J. C. J.; Tittonell, P. Tropical Forage Technologies Can Deliver Multiple Benefits in Sub-Saharan Africa. A Meta-Analysis. **2020**. https://doi.org/10.1007/s13593-020-00626-3

[36] Ginwal, D. S.; Kumar, R.; Ram, H.; Meena, R. K.; Kumar, U. Quality Characteristics and Nutrient Yields of Maize and Legume Forages under Changing Intercropping Row Ratios. *Indian J. Anim. Sci.* **2019**, *89*(3), 281–286. https://doi.org/10.56093/IJANS.V89I3.88079

ASEAN
Journal of Scientific and Technological Reports
Online ISSN:2773-8752

Research article

Profiling Festival-Period Traffic Accidents in Thailand: Clustering and Risk Factors

Witchaya Rattanametawee[1], Sriamporn Rebankoh[2], Khwansiri Sirimangkhala[3], and Naowarat Manitcharoen[4]*

[1] Faculty of Science, Mahasarakham University, Mahasarakham, 44150, Thailand
[2] College of Industrial Technology, King Mongkut's University of Technology North Bangkok, Bangkok, 10800, Thailand
[3] Big Data Institute (Public Organization), Ministry of Digital Economy and Society, Bangkok, 10900, Thailand
[4] College of Industrial Technology, King Mongkut's University of Technology North Bangkok, Bangkok, 10800, Thailand
* Correspondence: naowarat.n@cit.kmutnb.ac.th

Citation:
Rattanametawee, W.; Rebankoh, S; Sirimangkhala, K.; Manitcharoen, N. Profiling festival-period traffic accidents in Thailand: clustering and risk factors. *ASEAN J. Sci. Tech. Report.* **2025**, *28*(5), e259830. https://doi.org/10.55164/ajstr.v28i5.259830.

Article history:
Received: June 17, 2025
Revised: August 30, 2025
Accepted: September 6, 2025
Available online: August 14, 2025

Publisher's Note:
This article is published and distributed under the terms of the Thaksin University.

Abstract: The analysis of road accident data currently focuses on collecting and examining information from multiple perspectives, including factors such as the date and time of occurrence, details of the injured parties, alcohol consumption, and safety measures such as seatbelt and helmet usage. This study employs clustering techniques to group similar accident events. In addition to clustering, this study integrates logistic regression and decision tree techniques to enhance predictive capabilities. Logistic regression is used to estimate the probability of accident severity based on contributing factors, while decision tree modeling helps identify key decision rules that influence accident outcomes. To analyze the severity of road accidents, logistic regression can be employed to model the probability of severe outcomes based on contributing factors. A case study in Thailand is conducted to explore accident trends, which helps develop effective safety measures and policies. The findings emphasize the need to refine procedures for transferring emergency patients and implement stricter safety protocols to enhance care efficiency and mitigate adverse consequences.

Keywords: Clustering; graffic accident; logistic regression.

1. Introduction

Road accidents during the Thai New Year and Songkran festivals are a significant public safety concern, with Thailand experiencing the highest road traffic injury (RTI) mortality rate globally. The festive periods, particularly from April 11 to 17, see a marked increase in accidents, attributed to factors such as increased travel, alcohol consumption, and inadequate road safety awareness. Thailand ranks first in road accident fatalities, with a rate of 36.2 per 100,000 people [1]. During Songkran, the number of accidents significantly exceeds normal days, with a notable increase in severe injuries and fatalities [2, 3]. Key risk factors include alcohol consumption, age, and driving behavior [4]. EMS utilization during these holidays is linked to higher mortality rates among severe RTI patients, indicating potential inefficiencies in emergency response [3]. Despite efforts to improve road safety during these festive periods, the persistent high rates of accidents suggest a need for enhanced public awareness and preventive measures. Addressing the underlying causes of these accidents remains crucial for reducing fatalities and injuries. These pose a significant concern for public health and safety. The high frequency of accidents during these periods necessitates an in-depth analysis to identify underlying patterns and risk factors.

This study aims to address this gap by conducting a clustering analysis of traffic accident data, with a specific focus on the New Year and Songkran festivals in Thailand. This research holds substantial importance for several reasons: Identification of Risk Factors. By employing clustering techniques, the study will identify distinct groups of accidents based on shared characteristics. The results will enable the pinpointing of specific risk factors. Evaluation of Existing Measures: The findings can be used to evaluate the effectiveness of existing safety measures and identify areas where improvements are needed. Policy Recommendations: The results of this research can be used to inform the development of evidence-based policies and programs aimed at reducing road traffic accidents during festive periods. The specific objectives of this study are to analyze the patterns and characteristics of road traffic accidents that occur during the New Year and Songkran festivals in Thailand, identify the key risk factors associated with these accidents, and develop clusters of accidents based on shared characteristics. The study will cover a specific timeframe, from 2008 to 2014, encompassing both the New Year and Songkran festivals. The analysis will focus on national-level data collected by the National Institute for Emergency Medicine.

Logistic regression is a widely used statistical method for classification tasks, particularly in predicting binary outcomes. It estimates the probability of a dependent variable belonging to a specific category (e.g., success/failure) based on one or more independent variables. By applying the logistic function, the model ensures that predictions remain within the range of 0 to 1, making it well-suited for binary classification problems [5]. Logistic regression parameters are estimated using the maximum likelihood method, which iteratively determines the model that best fits the data. The model's coefficients represent the effect of independent variables on the log-odds of the predicted outcome, making their interpretation essential for understanding the influence of different factors [6]. Decision trees are a versatile and interpretable machine learning tool used across various fields for classification and regression tasks. They mimic human decision-making by creating a tree-like model that splits data based on the most informative features. This method is particularly valued for its ability to handle both categorical and numerical data, as well as its ease of understanding for human users [7]. While decision trees offer significant advantages in terms of interpretability and ease of use, they may not always perform as well as more complex models in certain scenarios, particularly when dealing with high-dimensional data or intricate relationships [8]. Clustering methodologies are of paramount importance in the realms of data mining and machine learning, facilitating the aggregation of data points based on their inherent similarities without the need for pre-established labels. A multitude of approaches are available, each characterized by distinctive algorithms and practical applications, rendering clustering a highly adaptable instrument across a diverse array of disciplines. K-means is the most widely used method, where data points are assigned to clusters based on their proximity to centroids [9]. It excels in small to medium datasets but struggles with larger datasets [10]. A comparative analysis reveals that, although K-means enjoys widespread recognition, Nonnegative Matrix Factorization (NMF) frequently achieves superior accuracy, particularly within lower-dimensional datasets [11]. Clustering methodologies are integral to big data analytics, facilitating the categorization of extensive datasets to derive enhanced insights. While clustering methodologies possess substantial analytical power, they are not immune to challenges, including susceptibility to noise and the need for parameter calibration, which can complicate their implementation in practical scenarios.

The application of tourism data in Thailand to provide suitable management recommendations. It utilizes clustering techniques in conjunction with logistic regression to categorize provinces. The analysis, which utilizes 14 variables, found that the optimal number of groups is three: primary provinces (18 provinces), secondary provinces (29 provinces), and tertiary provinces (30 provinces) [12]. It also highlights that the key variables influencing these provincial groupings are the number of overnight stays by foreign tourists, the number of domestic tourists, and their average expenditure. This analysis aims to develop a system for classifying the severity of road accidents in Thailand. It uses several machine learning algorithms, including Logistic Regression (LR), Random Forest (RF), and K-Nearest Neighbor (KNN). Additionally, different feature selection techniques were applied to create a total of nine predictive models. The findings support targeted road safety policies and highlight the importance of integrating data-driven approaches to reduce accident severity [13]. This study focuses on combining clustering with logistic regression and decision trees, analyzing beyond simple grouping to create a robust predictive model. First, clustering organizes a large dataset into meaningful groups based on similarities. Then, logistic regression can be applied to these clusters to predict the probability of a new data point belonging to a specific group. Simultaneously, a decision tree can be used to visually explain the key rules or criteria that define each cluster, providing clear and interpretable insights into what makes each group unique. By integrating these three techniques, the study gains the ability not only to group data but also to predict and explain future outcomes accurately.

2. Materials and Methods

2.1 Data

This article analyzes traffic accident data for the period 2008-2014, specifically during the New Year and Songkran festivals in Thailand. These data were obtained from the National Institute for Emergency Medicine in Thailand, a government agency responsible for overseeing and improving emergency medical services (EMS) nationwide [14]. It was established to enhance the quality, accessibility, and efficiency of emergency medical care, ensuring that people in need receive timely and adequate medical attention during emergencies. Descriptions of the dataset are as follows.

Table 1. Data description.

Column name	Description	Possible values
Date	Date of incident	13 Apr 2010, …
Time	Time of incident	20:01-21:00, …
Festival	The festival where the incident occurred	New year, Songkran
Gender	Victim's gender	Male, Female
Age	Victim's age	1, … , 99
RoadType	Road classification	Highway, Rural road, Urban road, Unknown
Status	Status of the victim	Driver, Passenger, Pedestrian, Unknown
Vehicle	Vehicle involved in the accident	Bicycle, Bus, Car, Car/Taxi, Motorcycle, Motorized tricycle, Pickup truck, Tricycle, Truck, Van, Other, Unknown
AnotherVehicle	The other driver's vehicle	Bicycle, Bus, Car, Car/Taxi, Motorcycle, Motorized tricycle, Pickup truck, Tricycle, Truck, Van, Other, Unknown
Measure	Preventive measures	Helmet, Seat belt, Unknown, No measure
Drinking	Drinking alcohol	Yes, No, Unknown
Transfer	The person who transported the patient	ALS, BLS, Died instantly, Foundation/Volunteer, FR, ILS, Not Transfer, Police officer, Victim/Relatives
ReferAdmit	Has he/she been admitted to the hospital?	Yes, No
Solution	Treatment outcome	Recovery, Died instantly, Died subsequently
Duration	Treatment duration	6 (Days), 15 (Days), …
Province	Province Name	Bangkok, Kalasin, …

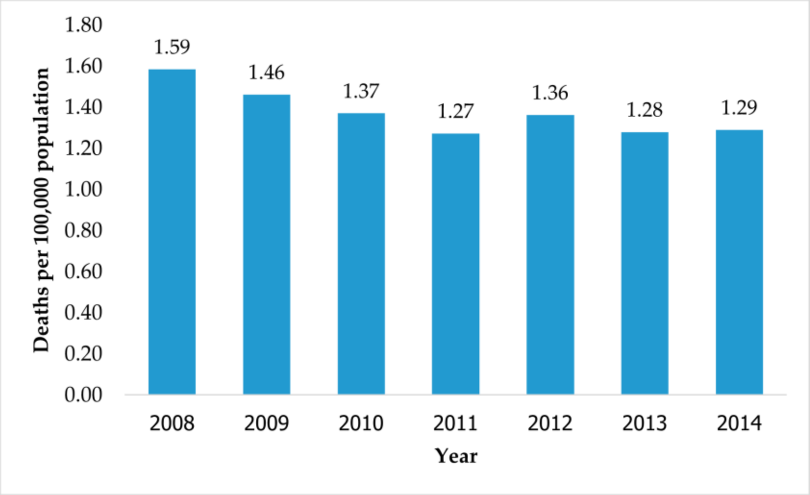


Figure 1. Deaths per 100,000 population by road accident during the New Year and Songkran festival.

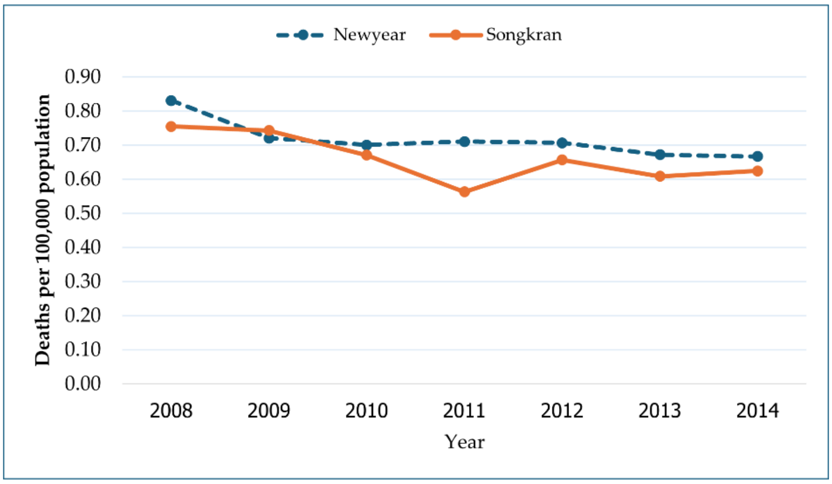


Figure 2. Comparison of deaths per 100,000 population during the New Year and Songkran festivals.

Figure 1 illustrates the annual deaths per 100,000 population from 2008 to 2014. Deaths peaked in 2008, then generally declined with some fluctuation through 2014, though rates remained relatively high throughout the period. Figure 2 compares deaths during the New Year and Songkran festivals from 2008 to 2014. The New Year consistently had higher fatalities than Songkran across all years. Both festivals showed declining death tolls over time, with a notable dip during the New Year 2011. Despite safety improvements, the New Year continues to pose a greater risk than Songkran

2.2 Methodology

Analyzing road accident data is crucial for understanding the factors contributing to these incidents and developing effective strategies for prevention. This process involves several key steps, each contributing to a comprehensive understanding of the data and its implications. Here is a detailed breakdown of the data analysis process for road accidents.

2.2.1. Importing Libraries

The first step involves importing the necessary libraries into the Python environment. Libraries such as Pandas are essential for data manipulation and cleaning, while NumPy is used for numerical operations. Matplotlib and Seaborn are essential for data visualization, enabling the creation of informative charts and graphs. Scikit-learn (sklearn) is indispensable for machine learning tasks, including feature selection and model building. Other libraries, such as imblearn and statsmodels, may be necessary for upsampling and statistical analysis.

2.2.2 Importing Data

The subsequent step entails importing road accident data from government databases. This data is typically loaded into a Pandas DataFrame, offering a structured and efficient format for handling tabular data.

2.2.3 Removing Duplicated Rows

Duplicate entries can distort analytical outcomes and lead to erroneous conclusions. Therefore, it is essential to identify and eliminate duplicate rows from the dataset, ensuring that each record is unique and contributes meaningfully to the analysis.

2.2.4 Detecting and Removing Outliers

Outliers, defined as data points that deviate significantly from the norm, can have a substantial impact on statistical analyses and the performance of machine learning models. Techniques such as box plots and z-scores are employed to detect these anomalies, which are then either removed or appropriately transformed based on the study's context.

2.2.5 Data Preparation

Data preparation encompasses the cleaning and transformation of raw data to render it suitable for analysis. This process includes handling missing values, standardizing or normalizing numerical variables, and encoding categorical variables. The integrity of the subsequent analysis is heavily dependent on the quality of data preparation.

2.2.6 Assessing Multicollinearity via VIF

Multicollinearity, the phenomenon where independent variables exhibit high correlations, can result in unstable and unreliable regression coefficients. The Variance Inflation Factor (VIF) is computed to quantify multicollinearity, enabling the identification and remediation of these issues within the dataset.

2.2.7 Upsampling

Road accident datasets often suffer from class imbalance; for example, severe accidents may be underrepresented compared to minor accidents. Upsampling techniques, such as SMOTE, are implemented to balance the class distribution, thereby enhancing the performance and robustness of machine learning models.

2.2.8 Encoding Categorical Variables

Since machine learning models predominantly require numerical input, categorical variables must be transformed into numerical formats. One-hot encoding and label encoding are standard methods used for this purpose, where one-hot encoding creates binary variables for each category, and label encoding assigns a unique integer to each category.

2.2.9 Dataset Partitioning

For evaluating model performance, the dataset is partitioned into three subsets: training, validation, and testing sets. The training set is used to build the model, the validation set to fine-tune hyperparameters, and the testing set to evaluate the model's final performance. In this setup, the dataset is divided into training and testing subsets using an 80:20 ratio. Specifically, 80% of the data is allocated for model training, while the remaining 20% is reserved for testing.

2.2.10 Feature Selection Using Logistic Regression and Decision Trees

Feature selection involves identifying a subset of relevant variables that most effectively predict the target outcome. Techniques such as logistic regression and decision trees are employed to evaluate the importance of each feature, thereby enhancing model performance, mitigating overfitting, and improving interpretability.

2.2.11 Visualization of Feature Coefficients

The visual representation of the coefficients derived from the logistic regression model provides insight into the relationships between the predictors and the outcome variable. Such visualizations facilitate the identification of key factors contributing to road accidents.

2.2.12 Clustering

Clustering is applied to the accident data to identify distinct clusters based on shared characteristics. This clustering approach is particularly advantageous for targeting interventions and developing prevention strategies.

2.2.13 Conclusions

The final step involves synthesizing the findings from the data analysis. This includes identifying the principal factors that contribute to road accidents, evaluating the performance of the machine learning models employed, and formulating evidence-based recommendations for prevention strategies. Additionally, the analysis of clusters derived from clustering provides nuanced insights into the heterogeneous characteristics of different accident groups, thereby informing more targeted interventions and safety measures.

3. Results and Discussion

3.1 Results

This section presents the primary findings from logistic regression and clustering analyses. We first discuss the most influential factors contributing to road accidents, as identified through logistic regression coefficients. We then examine the clustering analysis results, which categorize the data into distinct groups based on road type, vehicle type, and transfer method.

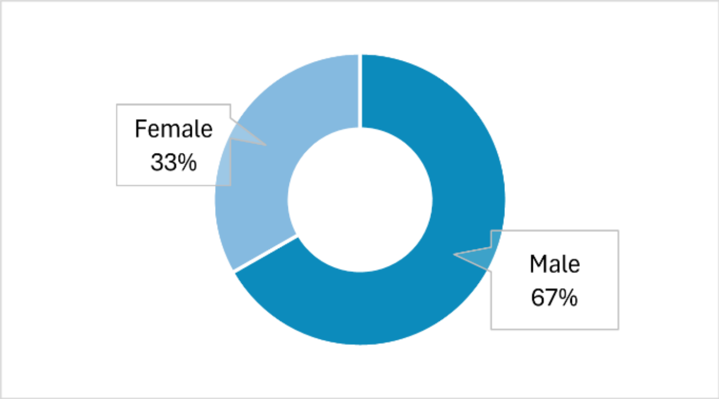


Figure 3. Accident Victim Statistics by Gender.

Figure 3 shows men account for 262,273 individuals (67%) compared to 130,858 women (33%), representing a 2:1 ratio. Similarly, Figure 4 illustrates a significantly larger gender disparity, with men comprising 4,876 individuals (79.36%) and women comprising 1,286 (20.64%).

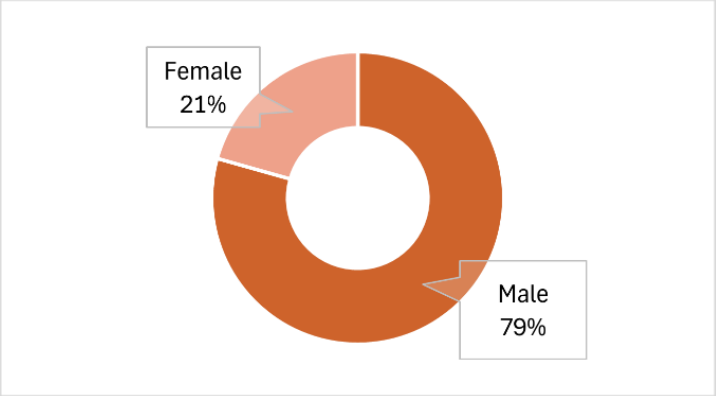


Figure 4. Mortality statistics by gender.

According to Figure 4, men comprise 4,876 individuals (79.36%), while women comprise 1,286 individuals (20.64%). This notable difference indicates a higher proportion of male cases compared to female cases, suggesting the potential influence of gender-related factors on the observed outcomes.

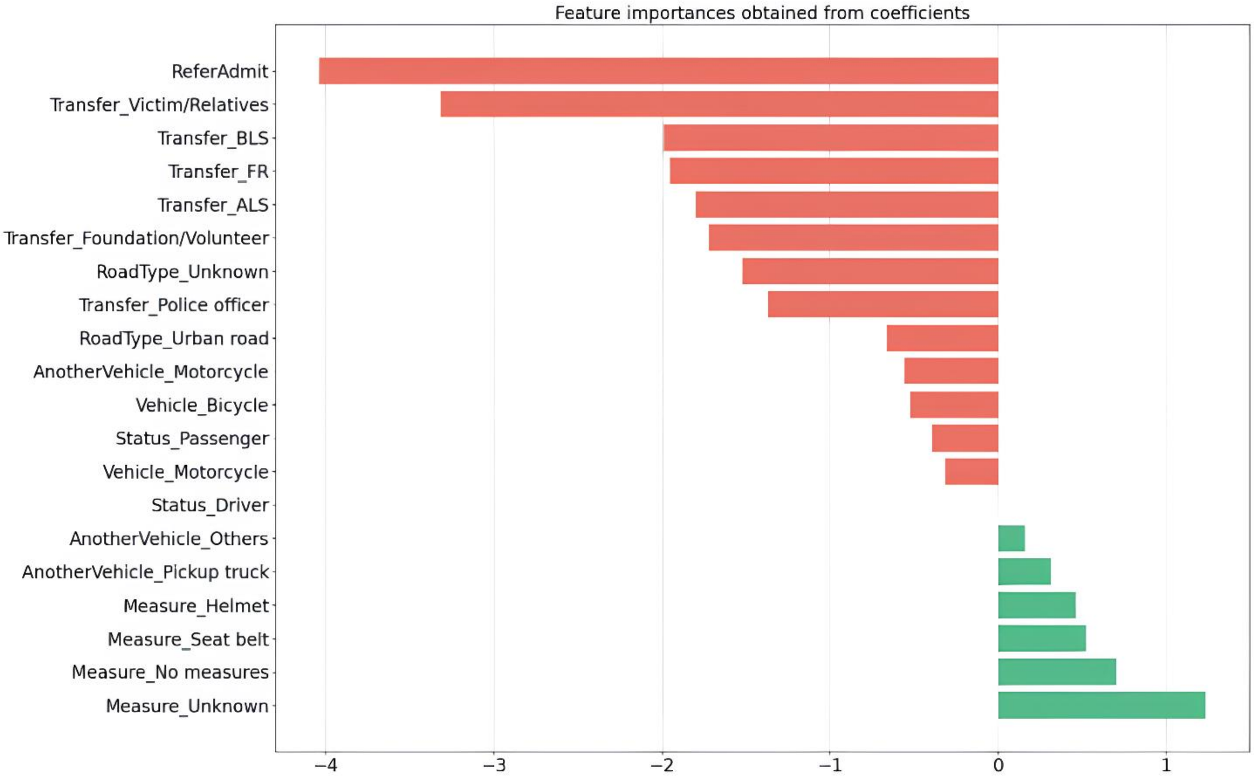
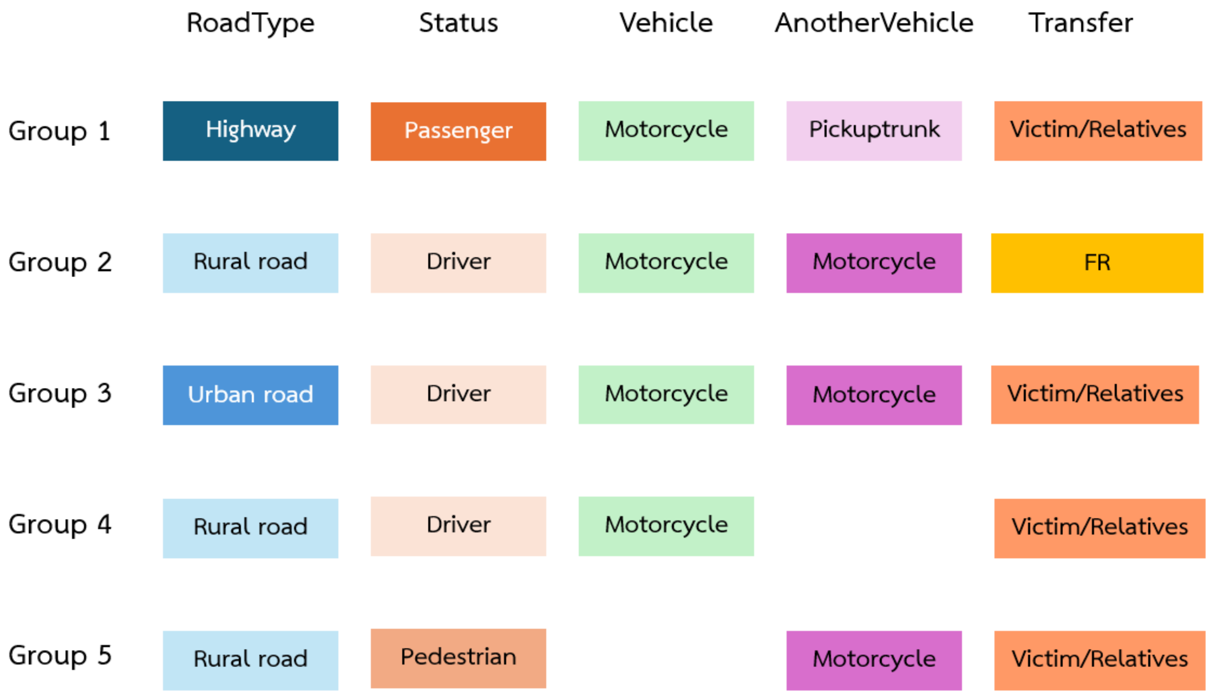


Figure 5. The features of importance derived from logistic regression coefficients.

Figure 5 displays the feature importances derived from the logistic regression coefficients. Red bars indicate features that reduce accident probability (ReferAdmit, Transfer_Victim/Relatives, Transfer_BLS), while green bars show features that increase it (Measure_Unknown, Measure_No measures, Measure_Seat belt). The analysis reveals that transfer methods and safety measures significantly impact outcomes, suggesting that improved transfer systems and safety protocols could enhance care efficiency."

The analysis utilized logistic regression to predict the likelihood of road accidents. Prior to modeling, the data underwent a thorough cleaning process to ensure accuracy and reliability. Multicollinearity was assessed using the Variance Inflation Factor (VIF) to identify and address any potential issues among predictor variables. The final model achieved a high accuracy rate of 98.67 %, demonstrating its strong predictive capability.



	RoadType	Status	Vehicle	AnotherVehicle	Transfer
Group 1	Highway	Passenger	Motorcycle	Pickuptrunk	Victim/Relatives
Group 2	Rural road	Driver	Motorcycle	Motorcycle	FR
Group 3	Urban road	Driver	Motorcycle	Motorcycle	Victim/Relatives
Group 4	Rural road	Driver	Motorcycle		Victim/Relatives
Group 5	Rural road	Pedestrian		Motorcycle	Victim/Relatives

Figure 6. Scenarios concerning road incidents and traffic patterns.

This visual, Figure 6, represents the results of a clustering analysis, categorizing road accident data into five distinct groups based on key features: Road Type, Status, Vehicle, Another Vehicle, and Transfer. Each group is defined by specific combinations of these features, offering insights into common accident scenarios. Detailed Analysis is described as follows.

Group 1: Highway accidents involving motorcycle passengers colliding with pickup trucks, with victim/relative transfers

Group 2: Rural road accidents between motorcycle drivers and other motorcycles, using FR transfer method

Group 3: Urban motorcycle-to-motorcycle collisions involving drivers, with victim/relative transfers

Group 4: Single-vehicle motorcycle accidents on rural roads with victim/relative transfers

Group 5: Pedestrian-motorcycle accidents on rural roads with victim/relative transfers

3.2 Discussion

The clustering analysis reveals five distinct accident patterns characterized by road type, user role, vehicle involvement, and transfer methods. Motorcycles emerge as the dominant factor, appearing in every cluster either as the primary or secondary vehicle involved. The analysis identifies critical accident scenarios: motorcycle-pickup truck collisions on highways, motorcycle-to-motorcycle crashes in both rural and urban settings, single-vehicle motorcycle accidents, and motorcycle-pedestrian incidents in rural areas. These patterns highlight environment-specific risks and reveal important gaps in emergency response systems. Notably, rural accidents often rely on transfers of victims or their relatives rather than professional emergency services, which can potentially impact care outcomes. The findings underscore the need for targeted interventions: enhanced motorcycle safety protocols for highways, improved collision prevention strategies for urban and rural motorcycle interactions, and strengthened emergency transfer systems, particularly in rural areas where professional emergency response may be limited."

4. Conclusions

This study reveals significant patterns in road accident data during festive periods. A notable gender disparity emerged, with males being twice as likely to experience accidents and four times more likely to suffer fatal outcomes compared to females. The analysis demonstrates that patient transfer methods have a critical influence on injury severity, highlighting the urgent need to strengthen emergency response systems. The clustering analysis identified motorcycles as the dominant factor across all accident scenarios, spanning highways, urban roads, and rural settings, and involving drivers, passengers, and pedestrians. Importantly, the study confirms the protective value of safety measures, with the use of helmets and seatbelts showing positive impacts on outcomes, reinforcing the importance of sustained prevention campaigns. These findings underscore the need for targeted interventions, including gender-specific safety education, motorcycle-focused safety protocols tailored to different road environments, enhanced emergency transfer systems (particularly in rural areas), and continued promotion of protective equipment use. Such targeted policies and resource allocation could significantly improve road safety and reduce accident burden during high-risk festive periods. Future analyses should incorporate environmental and infrastructure variables to enhance predictive accuracy. Key additions include road surface conditions (dry/wet), road geometry (straight/curve/intersection), weather conditions (clear/fog/rain), and lighting conditions (daylight/nighttime with/without lighting). Integrating these factors would provide deeper insights into contextual influences on accident severity and improve the model's ability to capture complex environmental relationships.

5. Acknowledgements

The authors gratefully acknowledge the Open Government Data of Thailand for providing access to valuable datasets that facilitated this research.

Funding: This research was funded by the College of Industrial Technology, King Mongkut's University of Technology North Bangkok (Grant No. Res-CIT0629/2023).

Author Contributions: All the authors have contributed by reading and revising the current study, including numerical simulation, writing, and review

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] Simmachan, T.; Wongsai, N.; Wongsai, S.; Lerdsuwansri, R. Modeling road accident fatalities with underdispersion and zero-inflated counts. *PLoS ONE* **2022**, *17*(11), e0269022. https://doi.org/10.1371/journal.pone.0269022

[2] Palwisut, P. A Risk Prediction Model of Road Accidents During Long Holiday in Thailand Using Ensemble Learning with Decision Tree Approach. *Suan Sunandha Sci. Technol. J.* **2023**, *10*(2), 213–221. https://doi.org/10.53848/ssstj.v10i2.499

[3] Riyapan, S.; Thitichai, P.; Chaisirin, W.; Nakornchai, T.; Chakorn, T. Outcomes of Emergency Medical Service Usage in Severe Road Traffic Injury during Thai Holidays. *West J. Emerg. Med.* **2018**, *2*, 266–275. https://doi.org/10.5811/westjem.2017.11.35169

[4] Boonserm, E.; Wiwatwattana, N. Using Machine Learning to Predict Injury Severity of Road Traffic Accidents During New Year Festivals From Thailand's Open Government Data. In *Proceedings of the 2021 9th International Electrical Engineering Congress (iEECON)*, Pattaya, Thailand, **2021**; pp 464–467. https://doi.org/10.1109/iEECON51072.2021.9440287

[5] Dong, Q.; Chen, X.; Huang, B. *Logistic regression*; Elsevier BV, **2024**; pp 141–152.

[6] Dey, D.; Haque, M. S.; Islam, M. M. *et al.* The proper application of logistic regression model in complex survey data: a systematic review. *BMC Med. Res. Methodol.* **2025**, *25*, 15. https://doi.org/10.1186/s12874-024-02454-5

[7] Saranceva, S. Applying the Decision Tree Method in the Field of Management Activities. *Ergodizajn* **2024**, 241–246. https://doi.org/10.30987/2658-4026-2024-2-241-246

[8] Fleissner, M.; Zarvandi, M.; Ghoshdastidar, D. Decision Trees for Interpretable clusters in mixture modelsand deep representations. **2024**.

[9] Sonia, Y. Study of Existing Methods & Techniques Of K-Means Clustering. **2024**. https://doi.org/10.53555/kuey.v30i4.1755

[10] Velunachiyar, S.; Sivakumar, K. Some Clustering Methods, Algorithms and their Applications. *Int. J. Recent Innovation Trends Comput. Commun.* **2023**. https://doi.org/10.17762/ijritcc.v11i6s.6946

[11] Basiri, F.; Amer, A.; Ranjbar Naserabadi, M. J.; Moghimi, M. M. A Comparative Study of K-means and NMF Clustering Algorithms. In *2023 2nd International Engineering Conference on Electrical, Energy, and Artificial Intelligence (EICEEAI)*, Zarqa, Jordan, **2023**; pp 1–4. https://doi.org/10.1109/EICEEAI60672.2023.10590510

[12] Boonkrong, P.; Simmachan, T.; Sittimongkol, R.; Lerdsuwansri, R. Data-Driven Approach in Provincial Clustering for Sustainable Tourism Management in Thailand. *Thail. Stat.* **2025**, *23*, 481–500.

[13] Simmachan, T.; Wongsai, S.; Lerdsuwansri, R.; Boonkrong, P. Impact of COVID-19 Pandemic on Road Traffic Accident Severity in Thailand: An Application of K-Nearest Neighbor Algorithm with Feature Selection Techniques. *Thail. Stat.* **2024**, *23*, 129–143. https://doi.org/10.1371/journal.pone.0309234

ASEAN
Journal of Scientific and Technological Reports
Online ISSN:2773-8752

Research article

Enhanced Electrical and Optical Properties of Cu-Doped ZnO Nanorods Synthesized via Co-Precipitation Method

Onanong Detchaiyaphum[1], **and Buppachat Toboonsung**[1, 2*]

[1] Graduate School, Nakhon Ratchasima Rajabhat University, Nakhon Ratchasima, 30000, Thailand
[2] Faculty of Science and Technology, Nakhon Ratchasima Rajabhat University, Nakhon Ratchasima, 30000, Thailand
* Correspondence: buppachat.t@nrru.ac.th

Citation:
Detchaiyaphum, O.; Toboonsung, B. Enhanced electrical and optical properties of Cu-doped ZnO nanorods synthesized via co-precipitation method. *ASEAN J. Sci. Tech. Report.* **2025**, *28*(5), e259802. https://doi.org/10.55164/ajstr.v28i5.259802.

Article history:
Received: June 14, 2025
Revised: September 9, 2025
Accepted: September 11, 2025
Available online: September 14, 2025

Abstract: Copper-doped zinc oxide (CZO) nanomaterials were prepared by the co-precipitation method, with a 0.5 M $ZnCl_2$ solution as the starting material, doped with 0-5 wt.% $CuCl_2$. The 100 wt.% Cu sample, synthesized without Zn, acts as a pure CuO reference. A 1 M NaOH solution was used as the precipitating agent to adjust the pH to 10. The final product was calcined at 500 °C for 3 h. Morphological analysis using SEM revealed that the CZO samples with 0-5 wt.% Cu exhibited a rod-shaped morphology, whereas the 100 wt.% Cu sample displayed a sheet-like structure with mixed nanoparticles. XRD confirmed the hexagonal wurtzite crystal structure in CZO with no detectable secondary phases, indicating successful incorporation of Cu^{2+} ions into the ZnO lattice. Elemental composition analysis using EDS supported this finding, showing a progressive increase in Cu content from 0.83 wt.% at 1 wt.% doping to 6.03 wt.% at 5 wt.%, accompanied by a corresponding decrease in Zn content. These results suggest that Cu^{2+} ions were effectively substituted for Zn^{2+} within the crystal lattice without forming impurity phases. Optical properties and energy band gap analysis, conducted using fluorescence and ultraviolet-visible spectrophotometry, indicated optimal conditions at 1 wt.% Cu doping. This level corresponded to the lowest band gap energy and the highest electrical conductivity value of 2.71×10^{-3} $(\Omega\cdot cm)^{-1}$, demonstrating a strong correlation between optical absorption and electrical performance. This study presents a systematic investigation into the effect of low-level Cu doping on the structural, optical, and electrical properties of ZnO nanomaterials.

Keywords: Copper-doping; Zinc oxide; Fluorescence spectroscopy; Optical properties; Electrical conductivity

1. Introduction

Nanotechnology currently plays a significant role in advancements across various fields, including medicine, food, cosmetics, electronics, and electrical applications. These developments often enhance the properties of materials, such as electrical and optical performance, to achieve greater efficiency. The synthesis of nanomaterials can be carried out in both thin-film and powder forms, utilizing methods such as spark discharge [1-2], sputtering [3], spray pyrolysis [4], sol-gel processing [5], co-precipitation [6], chemical vapor deposition (CVD) [7], and hydrothermal synthesis [8], among others. Various types of nanomaterials have garnered significant attention, such as titanium dioxide [2], zinc oxide, iron oxide [9], nickel oxide [6], and copper oxide [1], among others. Among these, ZnO is widely studied due to its exceptional optical and electrical properties. As a wide-bandgap semiconductor (3.37 eV), ZnO finds applications in diverse fields, including optoelectronic devices [10], solar cells [11-12], gas sensors [13-14], and antibacterial treatments [15, 16]. Doping with different metals can further enhance ZnO's electrical and optical properties. For instance, doping ZnO with copper (Cu) has been shown to reduce the band gap energy and improve electrical conductivity. Previous studies have indicated that Cu doping reduces the size of nanoparticles and enhances charge carrier mobility, thereby contributing to improved conductivity [17]. However, excessive Cu content may lead to the formation of secondary phases or localized defects, which can hinder electron transport and degrade material performance. Research on Co-doped ZnO nanoparticles prepared via co-precipitation revealed that ZnO retains its hexagonal wurtzite structure even after doping. Doping with 3%, 5%, and 7% Co was found to reduce particle size, while also decreasing the band gap energy, thereby improving electrical conductivity [18]. Similarly, studies on ZnO/CuO composites have demonstrated that increasing CuO concentration progressively reduces grain size, leading to changes in surface morphology and roughness [19]. These findings highlight the importance of precise control over the doping levels to optimize material properties.

While several studies in the literature have described Cu-doped ZnO nanostructures, with an emphasis on optical or structural properties, a comprehensive study of the relationship between low Cu dopant levels, morphology, and electrical properties as a function of the synthetic conditions has not been explored. Focusing on the correlative doping capability of morphology and electrical performance, a detailed investigation of the co-precipitated Cu-doped ZnO (CZO) NRs is presented in this work. The novelty of this work is the systematic analysis of low Cu doping concentrations (0-5 wt.%) under controlled conditions, and a 100 wt.% Cu sample as the CuO pure reference. The 100 wt. % Cu sample, prepared in the absence of the zinc precursor, acts as a pure CuO reference for comparison. Additionally, the ability to produce well-shaped nanorods without the use of surfactants presents a facile and potentially scalable synthetic method. The strong correspondence between the narrowed band gap and enhanced conductivity makes these CZO nanorods promising for optoelectronic devices, such as photodetectors, solar cells, and transparent conducting oxides. Morphological analysis was performed using scanning electron microscopy (SEM), and crystal structure was examined through X-ray diffraction (XRD). Additionally, the optical properties and energy band gap were analyzed using fluorescence spectrophotometry and ultraviolet-visible (UV-Vis) spectrophotometry, respectively. The electrical conductivity of the samples was evaluated using a two-point probe resistance measurement to assess the potential of CZO nanoparticles for electrical applications.

2. Materials and Methods

2.1 Synthesis Process

The copper-doped zinc oxide (CZO) nanomaterials were synthesized using zinc chloride ($ZnCl_2$, Ajax Finechem) as the precursor, with a solution concentration of 0.5 M. Copper chloride ($CuCl_2$, Ajax Finechem) was used as the dopant at concentrations of 0, 1, 3, and 5 wt.% relative to zinc chloride. Additionally, a separate control sample was prepared using only $CuCl_2$ (without $ZnCl_2$) and is referred to as 100 wt.% Cu to denote the absence of Zn precursor. This sample was intended to represent pure CuO for comparative analysis, rather than a doped ZnO system. A 1 M sodium hydroxide (NaOH, Labscan) solution was prepared as the precipitating agent for the co-precipitation process. The $ZnCl_2$ and $CuCl_2$ solutions were prepared using a magnetic stirrer for 15 minutes at 350 revolutions per minute (rpm). Subsequently, the NaOH solution was added dropwise to the mixture while continuously monitoring the pH until it reached 10. The resulting suspension was further stirred for an additional 45 minutes to ensure homogeneity. After stirring, the mixture was left to settle, allowing the precipitate to form. The residue was then washed repeatedly with deionized water until the pH became neutral. The washed precipitate was calcined in a furnace at a heating rate of 5°C/min from room temperature to 500°C, followed by isothermal treatment at 500°C for 3 hours. After calcination, the resulting material was ground into a fine powder using an agate mortar for further characterization.

2.2 Characterization studies

Morphological analysis of the CZO nanomaterials was performed using scanning electron microscopy (SEM, JEOL JSM-6010), which reveals the shape and size of the synthesized particles. Crystal structure analysis was conducted using X-ray diffraction (XRD, Rigaku SmartLab, 40 kV), which enables examination of the crystalline structure. The average crystallite size (D) was calculated using the Scherrer equation (1):

$$D = 0.9\lambda / \beta\cos\theta \tag{1}$$

where β is the full width at half maximum (FWHM) of the XRD peak (in radians), θ is the diffraction angle (in radians), and λ is the wavelength of the Cu-Kα (1.54 Å).

The optical properties were analyzed using fluorescence spectroscopy (Jasco, FP–8300), which measures the luminescence of the samples. This data was further used to calculate the band gap energy using Planck's equation, as shown in Equation 2. For light absorption testing, a UV-Vis spectrophotometer (Thermo Fisher, Genesys 10S) was utilized. The results were analyzed to determine the band gap energy using the Tauc relation. The band gap energy (Eg) [20] was calculated using Equation 3.

$$Eg = hc / \lambda \tag{2}$$

$$(\alpha h\upsilon)^2 = A(h\upsilon - Eg) \tag{3}$$

where Eg is the energy band gap (in eV), h is the Planck constant (6.626×10^{-34} J/s), υ is the frequency (Hz), c is the light velocity (3×10^8 m/s), λ is the wavelength (nm), A is the energy constant, and α is the absorbance coefficient.

The electrical conductivity analysis was conducted by measuring the electrical resistance of the synthesized material using a two-point probe method [21]. A multimeter was connected to two copper electrode plates, with a 100-micrometre-thick plastic sheet placed between them. The sample chamber had a dimension of 10×10 mm^2 [21,22]. A 50-milligram sample of the synthesized material was weighed and placed on the plastic sheet in the chamber on the first copper plate. The second copper plate was then placed on top, and the setup was pressed with a hydraulic press at 500 psi. The measured electrical resistance (UNI-T, UT139) was used to calculate the electrical resistivity of the material using Equation 4. Subsequently, the electrical conductivity was determined using Equation 5 [21, 22]. Each measurement was repeated five times to ensure accuracy.

$$\rho = RA / L \tag{4}$$

$$\sigma = 1 / \rho \tag{5}$$

where ρ is the electrical resistivity ($\Omega\cdot$m), R is the resistance (Ω), A is the surface area (1×1 cm^2), L is the thickness (1 μm), and σ is the electrical conductivity (($\Omega\cdot$m)$^{-1}$), as shown in Figure 1.

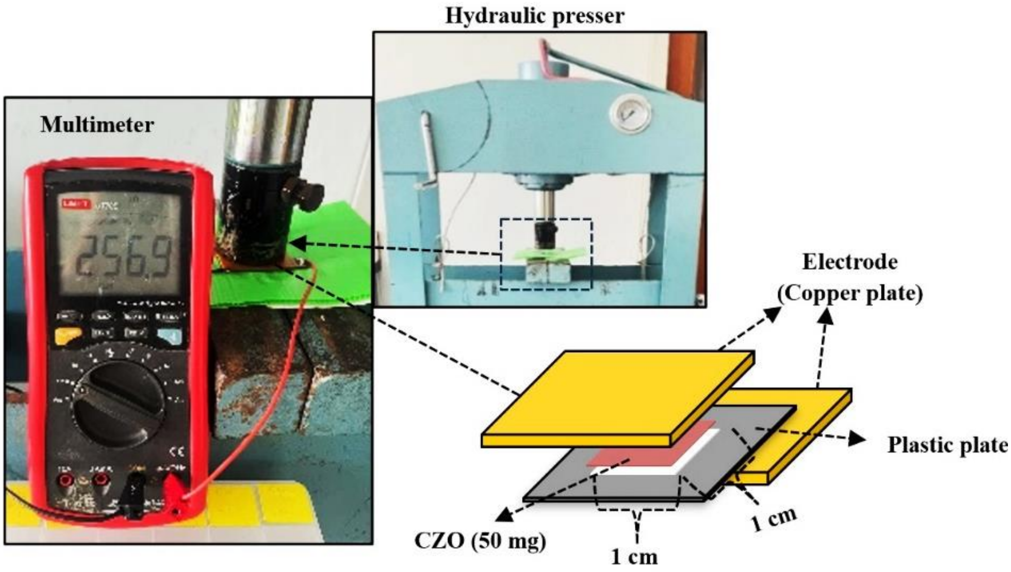


Figure 1. Schematic diagram and experimental setup for measuring the electrical resistance of copper-doped ZnO (CZO) powder.

3. Results and Discussion

3.1 Morphology Analysis

The morphological analysis using scanning electron microscopy (SEM) revealed that the CZO samples with 0-5 wt.% Cu exhibited a nanorod-like structure with irregular sizes, as illustrated in Figure 2. It was considered that the formation of a rod-like structure in the CZO samples may be due to the anisotropic growth inherent in the hexagonal wurtzite structure of the ZnO. In this crystal system, the [001] direction (c-axis) has faster growth kinetics owing to the greater surface energy of the polar (001) plane than the non-polar {100} facets, resulting in the one-dimensional elongation and formation of nanorods [24, 26]. Such growth will be even more favored under the current synthesis conditions: a high pH environment (pH 10), produced by NaOH, not only allows for fast nucleation but also stabilizes the (001) polar surface, leading to preferential c-axis growth. In this study, the co-precipitation method used in the experiment, which is combined with the heat treatment at 500 °C, can afford ample thermal energy for Ostwald ripening (i.e., smaller particles dissolve and redeposit onto larger ones), which is accompanied by the increase of the crystallinity of the nanoparticles and the preferential growth. The lack of organic surfactants or complexing reagents suggests that the morphology is predominantly controlled by the thermodynamic stability of crystal faces in an alkaline medium, rather than by capping action. In comparison to reports by various researchers who observed spherical to irregular-shaped ZnOnanoparticles using the sol-gel route [5], hydrothermal synthesis at lower temperatures [8], and capped agents that limited anisotropic growth. For example, Sangeetha et al. prepared spherical ZnO nanoparticles via the sol-gel method due to the isotropic growth of these particles in alcoholic media [5]. At the same time, Al-Khezraji et al. detected patchy clusters from ZnO/CuO composites studied by precipitation without in situ pH control [19]. The contributions presented in this work provide new insights into the key role of synthesis parameters such as pH and thermal treatment on the final particle shape. The diameters of the nanorods ranged from approximately 290 to 450 nm, while their lengths varied between 1200 and 1520 nm. The aspect ratio (length-to-diameter ratio) was calculated as 4.11, 4.14, 3.66, and 2.87 for doping levels of 0%, 1%, 3%, and 5% wt.%, respectively. The highest aspect ratio was observed at 1 wt.% Cu doping. In contrast, the sample containing 100 wt.% Cu showed a sheet-like structure with mixed nanoparticles. This sample was synthesized without any $ZnCl_2$, and thus, in the absence of Zn^{2+} ions, copper ions (Cu^{2+}) precipitate as $Cu(OH)_2$ under alkaline conditions, which readily decomposes to monoclinic CuO upon calcination at 500 °C. This explains the distinct sheet-like morphology and phase formation, consistent with the XRD and EDS results. This suggests that the presence of excess Cu results in a distinct growth mechanism and morphology. Despite variations in copper content, the overall rod-shaped morphology was preserved in the doped samples, suggesting that low-level Cu doping does not significantly alter the structural evolution of ZnO during synthesis.

3.2 Crystal Structure Analysis

X-ray diffraction (XRD) patterns confirmed that CZO samples with 0-5 wt.% Cu exhibited the hexagonal wurtzite structure of ZnO, consistent with JCPDS standard card number 36-1451 (Figure 3). The diffraction peaks corresponding to the (100), (002), (101), (102), (110), (103), (200), (112), and (201) planes were observed, indicating good crystallinity. For the 100 wt.% Cu sample, the diffraction pattern matched the monoclinic structure of CuO (JCPDS No. 48-1548), with characteristic peaks at (110), (11-1), (111), (200), (20-2), (020), (202), (11-3), (113), and (220). No additional peaks related to metallic Cu or secondary phases were detected in the CZO samples with low Cu concentrations. The absence of detectable Cu-related peaks suggests that Cu^{2+} ions have successfully substituted for Zn2+ ions within the ZnO lattice without forming secondary phases, which is consistent with previous studies [12, 17]. Using the Scherrer equation, the average crystallite size was estimated to be approximately 25 nm for CZO samples with 0-5 wt.% Cu, whereas the 100 wt.% Cu sample had a smaller crystallite size of about 16 nm. It should be noted that the 100 wt.% Cu sample was synthesized in the absence of any zinc precursor. As a result, the formation of a monoclinic CuO phase is expected, since Cu^{2+} ions first hydrolyze to $Cu(OH)_2$, which subsequently decomposes into CuO during calcination. This phase is structurally and compositionally distinct from the Cu-doped ZnO series, thereby serving as a pure CuO reference sample.

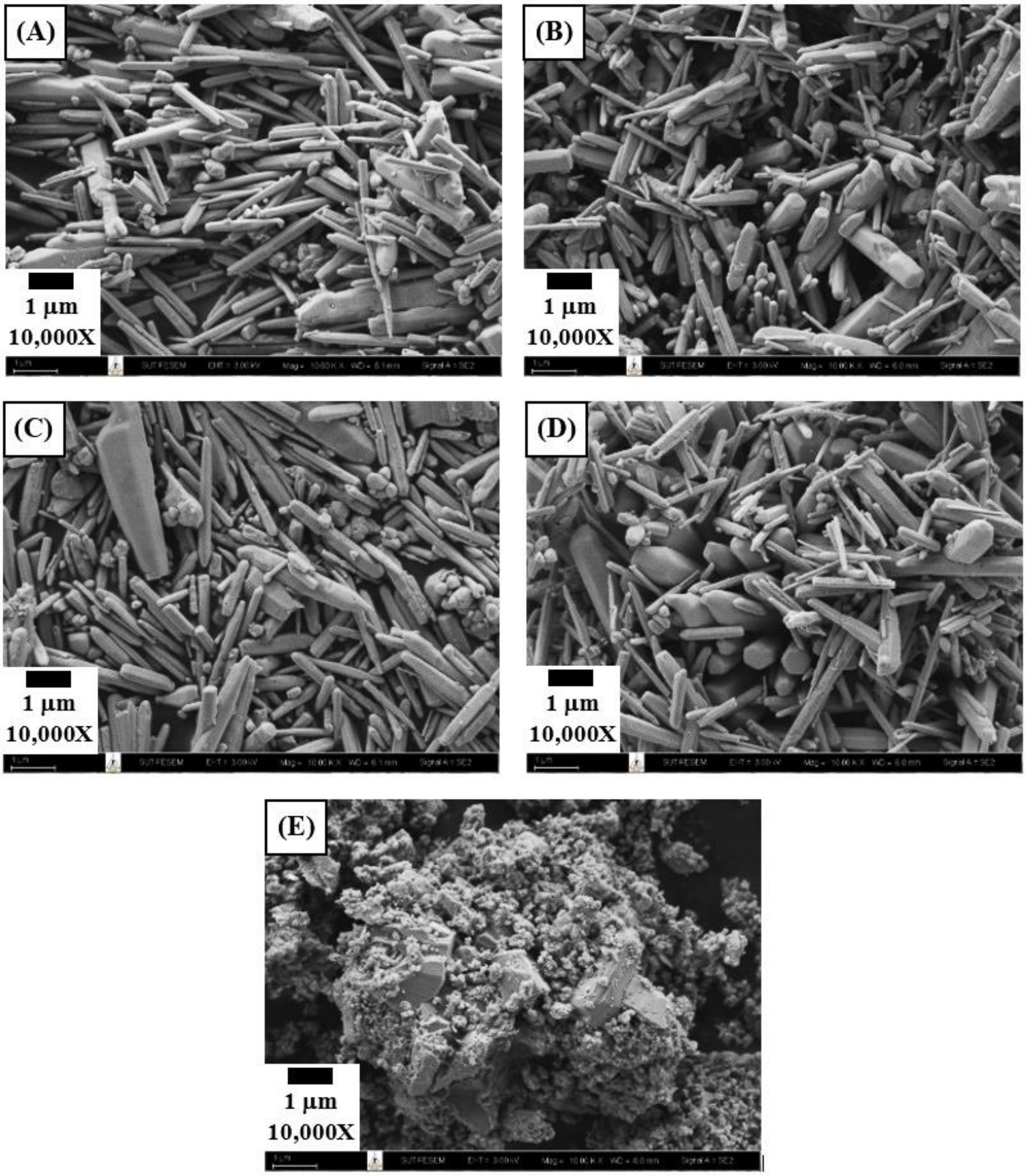


Figure 2. SEM images of CZO for (A) 0 wt.%, (B) 1 wt.%, (C) 3 wt.%, (D) 5 wt.%, and (E) 100 wt.% (control sample).

3.3 EDS Analysis and Elemental Composition

To confirm the successful incorporation of Cu into the ZnO lattice at low doping concentrations (1-5 wt.%), energy-dispersive X-ray spectroscopy (EDS) was conducted to determine the elemental composition of each sample. The results, summarized in Table 1, reveal a progressive increase in Cu content from 0.83 wt.% at 1 wt.% doping to 6.03 wt.% at 5 wt.% doping, accompanied by a corresponding decrease in Zn content. These findings support the hypothesis that Cu^{2+} ions are substitutionally incorporated into the ZnO lattice without forming secondary phases detectable by X-ray diffraction (XRD). In contrast, the 100 wt.% Cu sample exhibits a drastic increase in Cu content (71.81 wt.%) and a sharp decrease in Zn (1.51 wt.%), indicating the formation of a distinct CuO phase. The EDS data provide evidence for the successful incorporation of Cu into the ZnO lattice, with increasing Cu content as the doping level increases, at 1 wt.% Cu doping, the amount of

Cu increased slightly without significantly altering the Zn/O ratio, suggesting a uniform distribution of the dopant within the lattice. However, at higher doping levels (3-5 wt.%), the increase in Cu content corresponded to a decrease in Zn, which is indicative of Cu^{2+} substituting Zn^{2+} ions within the crystal structure, for the 100 wt. The negligible Zn content (1.51 wt.%) in the Cu sample confirms that the material consists predominantly of pure CuO nanostructures. This compositional evidence aligns with the XRD results, which show a monoclinic CuO phase, and the distinct sheet-like morphology observed in the SEM image. These findings strongly support the hypothesis that, at low doping levels (1-5 wt.%), Cu^{2+} ions successfully substitute Zn^{2+} within the ZnO lattice without forming secondary phases.

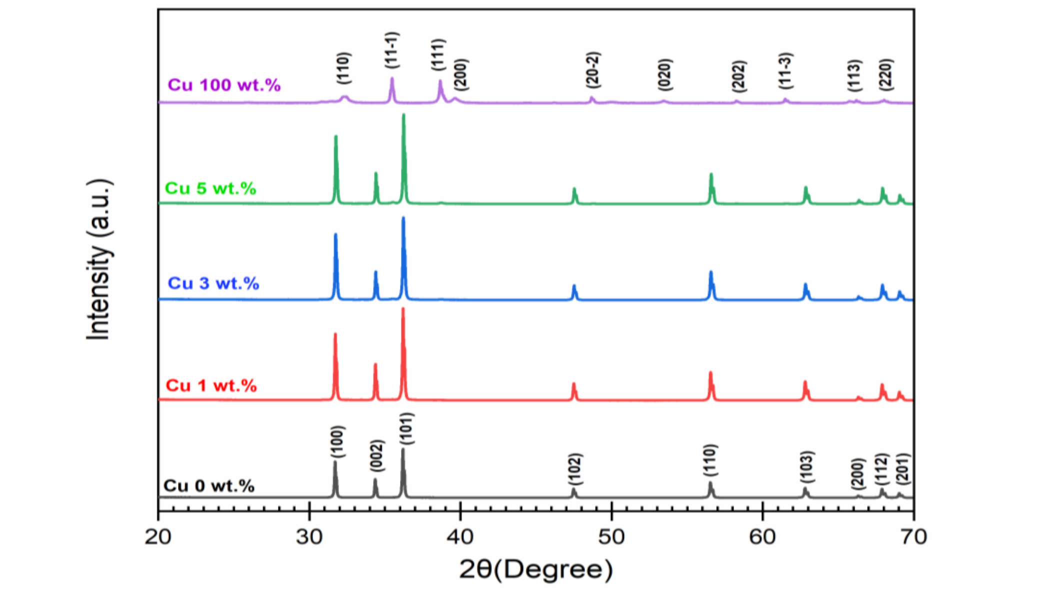


Figure 3. XRD pattern of 0-100 wt.% CZO.

Table 1. Elemental composition of CZO samples obtained from EDS analysis.

Cu doping (wt.%)	Zn (wt.%)	O (wt.%)	Cu (wt.%)
0	72.38	26.85	0.77
1	76.86	22.31	0.83
3	72.97	24.77	2.26
5	71.74	22.23	6.03
100 (control sample)	1.51	26.68	71.81

3.4 Fluorescence Spectroscopy Analysis

The luminescence of the samples was analyzed using fluorescence spectroscopy, and the peaks were modeled under the spectrum data using a Lorentzian function, resulting in three distinct peaks as shown in Figure 4. The CZO samples at concentrations of 0, 1, 3, 5, and 100 wt.% exhibited luminescence in the violet region (323-398 nm), corresponding to electron excitation from the ground state to a higher energy level. This excitation process is designated as Peak 1. Peak 2, observed in the blue region (422-444 nm), corresponds to fluorescence emission as electrons returned to the ground state. Peak 3, located in the green area (548-563 nm), is attributed to defect-related recombination, often associated with oxygen vacancies or zinc interstitials, which produce deep-level emission. These wavelengths were used to calculate the band gap energy using Planck's equation, where the band gap energy for excitation and fluorescence was found to be lower for Cu doping at 1, 3, and 5 wt.% compared to undoped ZnO with values of 3.11 eV and 2.78 eV, respectively. The reduction in band gap energy at 1 wt.% Cu correlates with enhanced charge carrier generation, which contributes to the improvement in electrical conductivity. However, at higher doping levels (3–5 wt.%), the slight increase in band gap energy and the presence of defect-related emission peaks suggest that excess Cu introduces localized states within the band gap, which may act as recombination or scattering centers. Furthermore, EDS analysis provided quantitative evidence of Cu incorporation at 1 wt.% Cu doping, the Zn/O ratio remained relatively unchanged, suggesting uniform distribution of Cu within the lattice; however, at 100 wt.% Cu, Zn content dropped to only 1.51 wt.%, confirming that it corresponded to pure CuO rather than a doped ZnO system. These findings strongly support the conclusion that low-level Cu doping effectively modifies the electronic structure of ZnO, reducing the band gap and enhancing light absorption. In contrast, excessive Cu content introduces localized defect states or impurity phases that act as scattering centers, diminishing carrier mobility and overall performance.

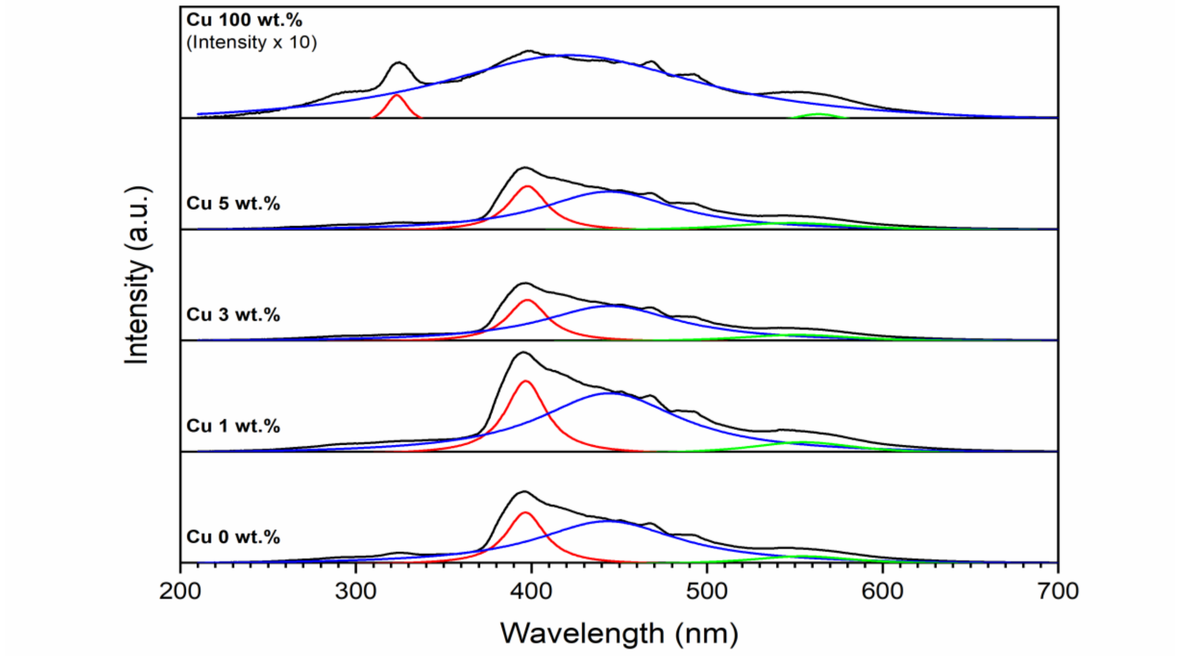


Figure 4. Fluorescence spectra of 0-100 wt.% CZO.

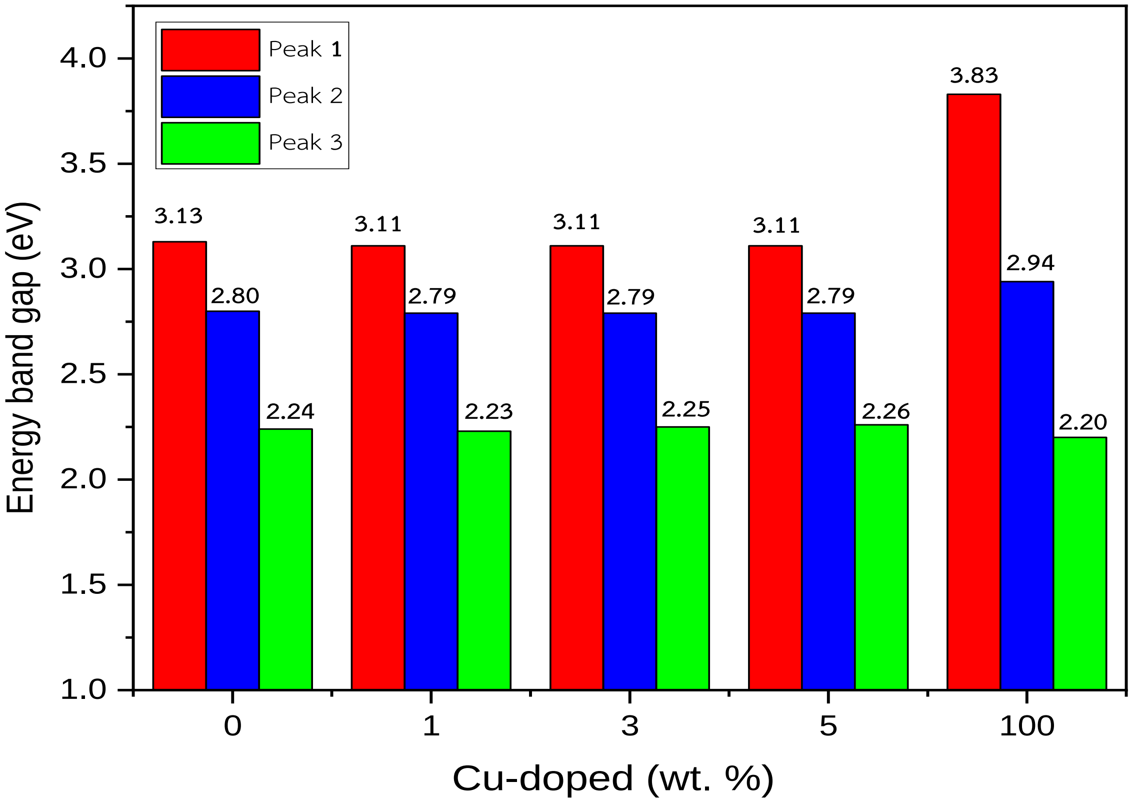


Figure 5. Three peaks of the energy band gap from fluorescence analysis in Figure 4.

3.5 UV-Vis Spectroscopy Analysis

The optical absorption properties of ZnO doped with Cu at concentrations of 0, 1, 3, 5, and 100 wt.% were analyzed using UV-Vis spectroscopy, as shown in Figure 6. The Tauc method was applied to calculate the band gap energy using Equation (3). The absorption occurred in the visible range (violet), with maximum wavelengths at 402, 428, 430, 449, and 350 nm. The corresponding band gap energies were calculated as 2.90, 2.72, 2.82, 2.83, and 2.16 eV, respectively. The lowest band gap energy was observed at 2.72 eV for Cu doping at 1 wt.%. It is noted that the band gap of nanomaterials is not only sensitive to the dopant concentrations, but also to the particle size and morphology. In this work, the diameter of the nanorod was found to decrease with an increase in Cu content, which generally results in a blue shift (an increase in band gap) due to quantum confinement. However, the measured red shift (from 2.90 eV to 2.72 eV) indicates that the effect of Cu doping prevails, primarily due to the hybridization of Cu 3d states with O 2p orbitals at the valence band edge, which narrows the band gap [23]. Such trends are opposite at high doped levels (3-5 wt.%), and the higher band gap of the Cu-doped ZnO might be attributed to the generation of defect states such as oxygen vacancies or the presence of Cu interstitials, which form interstitial levels in the mid-gap and counteract the narrowing effect. This observation contrasts with the study by Sajjad et al. [17], in which the band gap decreased continuously with an increase in copper content. This deviation can be attributed to the differences in doping units (wt.% vs. mol%) and the emergence of defect-related effects at higher doping levels in this study. These results reveal the existence of an optimal doping range for minimizing the band gap. This behavior has been previously reported by Cu-doped ZnO systems, in which the introduction of Cu^{2+} ions produces an impurity level close to the valence band toward the band gap energy [17,23]. The observed decrease in band gap from 2.90 eV (0 wt.%) to 2.72 eV (1 wt.% Cu) supports the successful substitution of Cu^{2+} into the ZnO lattice. The optical absorption analysis, performed using UV-Vis spectroscopy, showed consistent trends with the fluorescence results. CZO samples doped with 1 wt.% Cu exhibited the lowest band gap energy of 2.72 eV, confirming the effectiveness of low-level Cu doping in modifying the electronic structure of ZnO. At higher Cu concentrations (3 and 5 wt.%), the band gap energy increased slightly to 2.82 eV and 2.83 eV, respectively, aligning with the observed increase in defect-related emission from fluorescence measurements. This correlation between fluorescence and UV-Vis spectroscopy supports the hypothesis that Cu^{2+} ions are substitutionally incorporated into the ZnO lattice, leading to a narrowing of the band gap and enhanced optical absorption. In contrast, the sample containing 100 wt.% Cu showed the lowest band gap energy (2.16 eV), which is interpreted as pure CuO rather than a doped system.

3.6 Electrical Conductivity Analysis

The electrical conductivity of the synthesized CZO nanomaterials was evaluated by measuring the electrical resistance and subsequently calculating the resistivity and conductivity using Equations (4) and (5), respectively. The results are summarized in Figure 7 and discussed below. The undoped ZnO sample (0 wt.% Cu) exhibited an electrical conductivity of 2.18×10^{-3} $(\Omega\cdot cm)^{-1}$, upon doping with 1 wt.% Cu, the conductivity increased to 2.71×10^{-3} $(\Omega\cdot cm)^{-1}$, which is the highest value observed in this study. Further increasing the Cu content to 3 wt.% and 5 wt.% resulted in a decrease in conductivity to 1.88×10^{-3} and 0.93×10^{-3} $(\Omega\cdot cm)^{-1}$, respectively. The sample containing 100 wt.% Cu showed a conductivity of 1.78×10^{-3} $(\Omega\cdot cm)^{-1}$, which is lower than that of the 1 wt.% doped sample but higher than that of the undoped ZnO. The reduction in band gap energy at 1 wt.% Cu correlates with enhanced charge carrier generation, leading to improved electrical conductivity. This suggests that Cu^{2+} ions were successfully incorporated into the ZnO lattice, contributing additional free carriers that enhance conductivity. However, further increases in Cu content beyond 1 wt.% led to a decline in electrical conductivity. Excessive Cu content introduces localized defect states that act as recombination centers or scattering sites, reducing carrier mobility and overall conductivity. These defects may arise from the formation of impurity phases or distortions in the crystal lattice due to excess Cu, which hinder the effective movement of charge carriers. Comparing these results with previous studies, Redwanul I. et al. reported an electrical conductivity of 0.0986×10^{-5} $(\Omega\cdot m)^{-1}$ for undoped ZnO nanoparticles [27], while Shahroz S. et al. found a conductivity of 1.09×10^{-4} $(\Omega\cdot cm)^{-1}$ for CuO nanoparticles [28]. However, they also noted a decline in conductivity at higher doping levels due to enhanced carrier scattering. This study confirms that optimal Cu doping at 1 wt.% significantly improves the electrical properties of ZnO nanomaterials, making them promising candidates for optoelectronic applications such as photodetectors and solar cells.

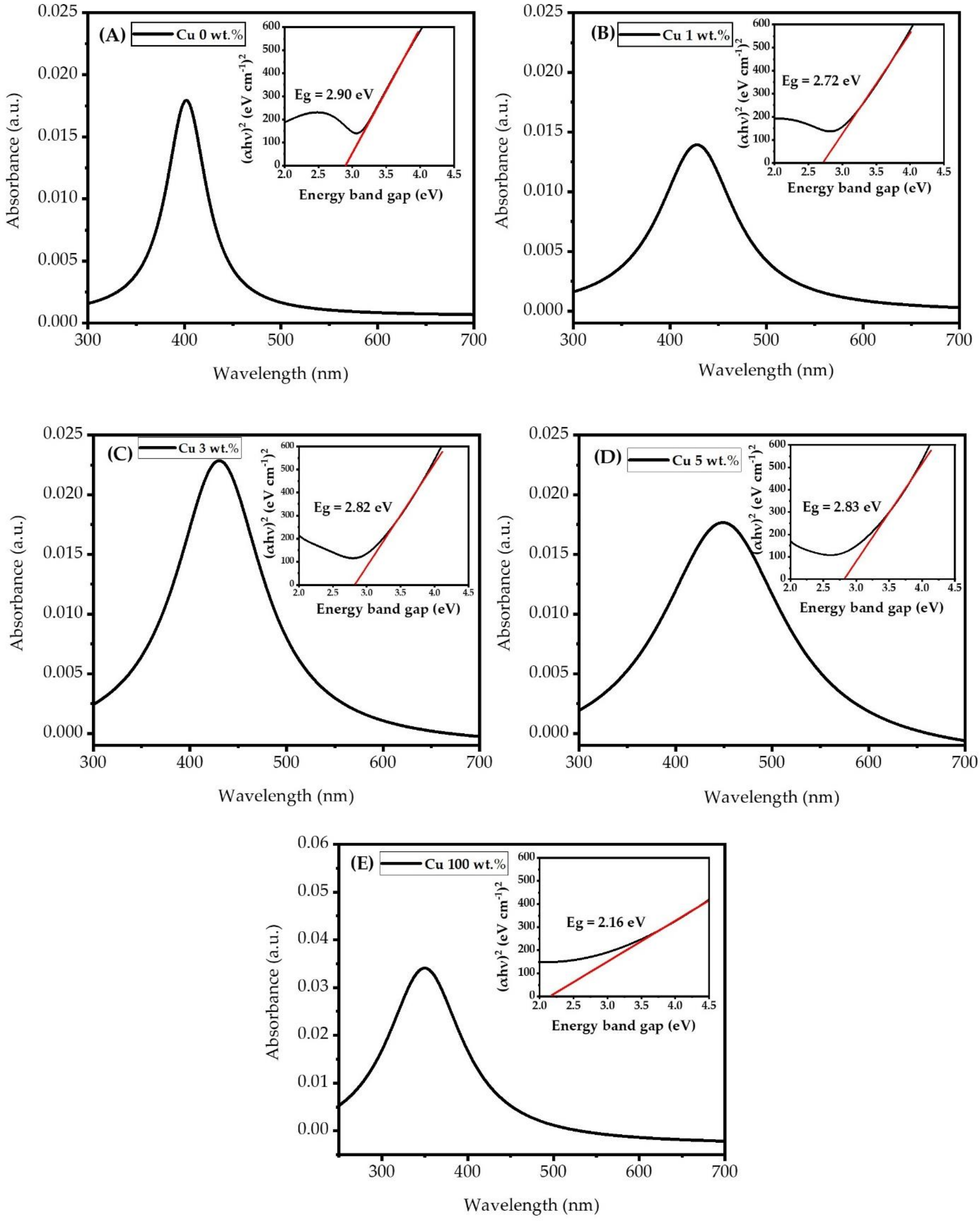


Figure 6. The graph of light absorption and the inset energy band gap from UV-VIS spectroscopy of CZO for (A) 0 wt.%, (B) 1 wt.%, (C) 3 wt.%, (D) 5 wt.%, and (E) 100 wt.%, respectively.

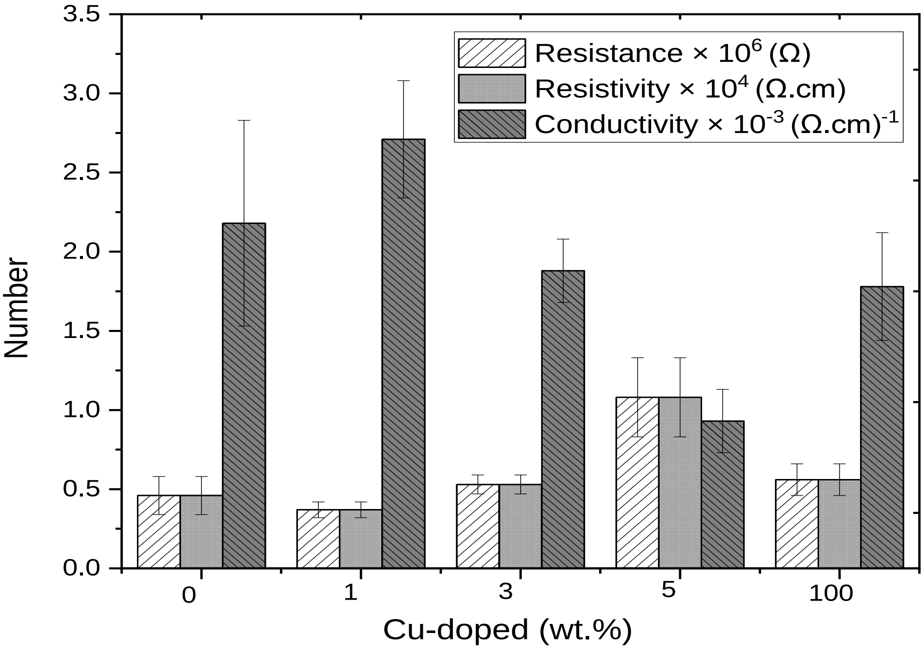


Figure 7. Electrical resistance, electrical resistivity, and electrical conductivity of CZO.

4. Conclusions

Copper-doped zinc oxide (CZO) nanomaterials with doping levels of 0, 1, 3, and 5 wt.% exhibited a nanorod morphology, while the sample containing 100 wt.% Cu displayed an irregular sheet-like structure with mixed nanoparticles. X-ray diffraction (XRD) analysis confirmed that the CZO nanomaterials retained the hexagonal wurtzite crystal structure within the low Cu doping range (0-5 wt.%), with no detectable secondary phases observed. This suggests that Cu^{2+} ions were successfully incorporated into the ZnO lattice without forming separate Cu or CuO phases. The tunable band gap in CZO nanorods arises from a synergistic interplay between Cu doping, particle size, and nanorod morphology, with doping being the primary tuning parameter. The electrical conductivity measurements revealed that the highest value of 2.71×10^{-3} $(\Omega\cdot cm)^{-1}$ was achieved at 1 wt.% Cu doping, which also corresponded to the lowest band gap energy based on both fluorescence and UV-Vis absorption analyses. These results indicate a strong correlation between optical and electrical properties in CZO nanomaterials. The reduction in band gap energy at 1 wt.% Cu correlates with enhanced charge carrier generation, leading to improved electrical conductivity. This finding supports the hypothesis that Cu^{2+} ions effectively substitute Zn^{2+} ions, enhancing charge transport efficiency. However, increasing Cu content beyond this optimal level resulted in a decrease in conductivity. Excessive Cu content introduces localized defect states that act as recombination centers or scattering sites, reducing carrier mobility and overall conductivity. This study demonstrates that moderate Cu doping significantly enhances the electrical and optical properties of ZnO nanomaterials. The optimal performance at 1 wt.% Cu highlights the importance of precise control over doping concentration to avoid defect-induced carrier scattering. The systematic investigation conducted in this work reveals that low-level Cu doping at 1 wt.% not only enhances electrical conductivity but also tunes the band gap energy, offering new opportunities for CZO in optoelectronic device engineering.

5. Acknowledgements

This work was supported by Nakhon Ratchasima Rajabhat University, Nakhon Ratchasima, Thailand. We would like to thank Khajondach Viangsong and Napaporn Thaweesaeng for their assistance with the preparation and analysis of UV-VIS spectroscopy and electrical resistance measurements. The XRD measurements and fluorescence spectroscopy were supported by Dr. Panadda Phansamdaeng, with contributions from Samart Tuaykratok and Keardtipong Jadnok..

Author Contributions: In this article, the following fields of work were carried by: Conceptualization, Buppachat Toboonsung.; methodology, Onanong Detchaiyaphum and BT.; validation, OD and BT.; formal analysis, BT. investigation, OD and BT.; resources, OD and BT.; data curation, BT.; writing—original draft preparation, OD.; writing—review and editing, BT.; visualization, OD and BT.; supervision, OD and BT; project administration, OD and BT.; funding acquisition, OD and BT. In addition, all authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare that they have no conflicts of interest.

References

[1] Traiwatcharanon, P.; Pon-On, W.; Zacharias, M.; Wongchoosuk, C. Electrochemical copper oxide nanoparticles-based sensor for butachlor plus propanil herbicide detection. *J Mater Sci: Mater Electron.* **2024**, *35*, 651. https://doi.org/10.1007/s10854-024-12419-5.

[2] Toboonsung, B. (2019) Surface morphologies and durability on water contact angle of titanium dioxide nanoparticle thin films. *Key Eng. Mater.* **2019**, *798*, 158-162. https://doi.org/10.4028/www.scientific.net/KEM.798.158.

[3] Nagabharana, R.M.; Kumaraswamy, G.N.; Susheel, K.G.; Umananda, M.B. Effect of thermal annealing on structural and electrical properties of TiO_2 thin films. *Thin Solid Films* **2020**, *710*, 138262. https://doi.org/10.1016/j.tsf.2020.138262.

[4] Cuadra, J.G.; Estrada, A.C.; Oliveira, C.; Abderrahim, L.A.; Porcar, S.; Fraga, D.; Trindade, T.; Seabra, M.P.; Labrincha J.; Carda J.B. Functional properties of transparent ZnO thin films synthesized by using spray pyrolysis for environmental and biomedical applications. *Ceram. Int.* **2023**, *49*, 32779-32788. https://doi.org/10.1016/j.ceramint.2023.07.246.

[5] Sangeetha, A.; Jaya Seeli, S.; Bhuvana, K.P.; Abdul Kader, M.; Nayak, S.K.; Correlation between calcination temperature and optical parameter of zinc oxide (ZnO) nanoparticles. *J. Sol-Gel Sci. Technol.* **2019**, 91, 261-272. https://doi.org/10.1007/s10971-019-05000-8.

[6] Toboonsung, B. Structure, magnetic property and energy band gap of Fe-doped NiO nanoparticles prepared by co-precipitation method. *Key Eng. Mater.* **2017**, *751*, 379-383. https://doi.org/10.4028/www.scientific.net/KEM.751.379

[7] Narin, P.; Kutlu-Narin, E.; Lisesivdin, S.B.; Growth dynamics of mist-CVD grown ZnO nanoplatelets. *Physica B: Condens. Matter.* **2021**, *614*, 413028. https://doi.org/10.1016/j.physb.2021.413028.

[8] Yaowen, H.; Junhui, Y.; Yun, W.; Jiayao, J.; Jialu, W.; Haiyan, T.; Ying, Y.; Tianqi, W.; Lin, X.; Dong X. Femtosecond laser combined with hydrothermal method to construct three-dimensional spatially distributed wurtzite ZnO micro/nanostructures to enhance photocatalytic properties. *Langmuir*, **2024**, *40*, 3892-3899. https://doi.org/10.1021/acs.langmuir.3c03840.

[9] Kumar, M.; Dede Heri, Y.Y.; Mani, G.; Bogeshwaran, K.; Fatmah, A.A.; Reem, A.H. Biological synthesis and characterization of iron oxide (FeO) nanoparticles using *Pleurotus citrinopileatus* extract and its biomedical applications. *Biomass Convers. Biorefin.* **2024**, *14*,12575-12585. https://doi.org/10.1007/s13399-023-04382-8.

[10] Ahmed Adel, A.A.; Hadia, N.M.A.; Meshal, A.; Mohamed, S.; Abdel-Hamid, I.M.; Fernández, S.; Rabia, M. Development of CuO nanoporous material as a highly efficient optoelectronic device. *Appl. Phys. A* **2022**, *128*, 321. https://doi.org/10.1007/s00339-022-05447-7.

[11] Sahar, I.S.; Maryam, A.A.; Wedian, K.A.; Ahmed, N.A. Low-cost applications by simple chemical method: solar cell and photodetector. *Int. J. Nanosci.* **2024**, *23(2)*, 2350063. https://doi.org/10.1142/S0219581X23500631.

[12] Aneesiya, K.R.; Cindrella, L. Localized surface plasmon resonance of Cu-doped ZnO nanostructures and the material's integration in dye sensitized solar cells (DSSCs) enabling high open-circuit potentials. *J Alloys Compd.* **2020**, *829*, 154497. https://doi.org/10.1016/j.jallcom.2020.154497

[13] Junfeng, C.; Haijun, Y.; Ke, Z.; Ying, Z.; Deshuo, M.; Yeguo, S. Integration of ZnO and Au/ZnO nanostructures into gas sensor devices for sensitive ethanolamine detection. *ACS Appl Nano Mater.* **2023**, *6(7)*, 5994-6001. https://doi.org/10.1021/acsanm.3c00350

[14] To Thi, N.; Dang, T.T.L.; Nguyen, V.D.; Chu, T.X.; Sven, I.; Xuan, T.V.; Nguyen D.H. A sigh-performance hydrogen gas sensor based on Ag/Pd nanoparticle-functionalized ZnO nanoplates. *RSC Advances* **2023**, *13*, 13017-13029. https://doi.org/10.1039/D3RA01436C.

[15] Thambidurai, S.; Gowthaman, P.; Venkatachalam, M.; Suresh, S. Enhanced bactericidal performance of nickel oxide-zinc oxide nanocomposites synthesized by facile chemical co-precipitation method. *J Alloys Compd.* **2020**, *830*, 154642. https://doi.org/10.1016/j.jallcom.2020.154642.

[16] Negi, P.B.; Rana, A.; Joshi, N.C.; Mishra, A.; Manoj, C.L.; Sunori, S.K. Synthesis, characterization and antimicrobial activity of zinc oxide nanoparticles against *Escherichia coli* and *Salmonella enterica*-water borne pathogens. *Asia Pac J Sci Technol.* **2024**; *29(03)*, APST-29. https://doi.org/10.14456/apst.2024.40.

[17] Sajjad, M.; Ullah, I.; Khanb, M.I.; Khanc, J.; Khana, M.Y.; Qureshi, M.T. Structural and optical properties of pure and copper doped zinc oxide nanoparticles. *Results Phys.* **2018**, *9*, 1301-1309. https://doi.org /10.1016/j.rinp.2018.04.010

[18] Ravichandran, A.T.; Karthick, R. Enhanced photoluminescence, structural, morphological and antimicrobial efficacy of Co-doped ZnO nanoparticles prepared by Co-precipitation method. *Results Mater.* **2020**, *5*, 100072. https://doi.org/10.1016/j.rinma.2020.100072.

[19] Al-Khezraji, A.A.R.; Abd Ali, H.R.; Yousif, A.A.; Abed, H.R. Effect of mixed ZnO/CuO nanoparticles on the structural, morphological, and topographical properties. *J Phys Conf Ser.* **2021**, *1963(1)*, 012053. https://doi.org/10.1088/1742-6596/1963/1/012053.

[20] Meryem, L.Z.; Touidjen, N.E.H.; Aida, M.S.; Aouabdia, N.; Rouabah, S. Growth of undoped ZnO thin films by spray pyrolysis: effect of precursor concentration. *J Opt.* **2023**, *52*, 1782-1788. https://doi.org/ 10.1007/s12596-022-01079-5.

[21] Montes, J.M.; Cuevas, F.G.; Cintas, J. Electrical resistivity of metal powder aggregates. *Metall Mater Trans B.* **2007**, *38*, 957-964. https://doi.org/10.1007/s11663-007-9097-3.

[22] Kingpho, P.; Toboonsung, B. Improvement of the Electrical properties of ZnO nanomaterials with Fe by Co-precipitation method. *Curr Appl Sci Technol.* **2025**, *25(3)*, e0263485. https://doi.org/10.55003/cast.2024.263485.

[23] Dejam, L.; Kulesza, S.; Sabbaghzadeh. J.; Ghaderi, A.; Solaymani, S.; Talu. S.; Bramowicz, M.; Amouamouha, M.; Salehi shayegan, A.H.; Sari, A.H. ZnO, Cu-doped ZnO, Al-doped ZnO and Cu-Al doped ZnO thin films: Advanced micro-morphology, crystalline structures and optical properties. *Results Phys.* **2023**, *44*, 106209. https://doi.org/10.1016/j.rinp.2023.106209.

[24] Lin, J.H.; Patil R.A.; Devan, R.S.; Liu, Z.A.; Wang, Y.P.; Ho, C.H.; Liou, Y.; Ma Y.R. Photoluminescence mechanisms of metallic Zn nanospheres, semiconducting ZnO nanoballoons, and metal-semiconductor Zn/ZnO nanospheres. *Sci Rep.* **2014**, *4*, 6967. https://doi.org/10.1038/srep06967.

[25] Mahroug, A.; Mari, B.; Mollar, M.; Boudjadar, I.; Guerbous, L.; Henni, A.; Selmi N. Studies on structural, surface morphological, optical, luminescence and UV photodetection properties of sol-gel Mg doped ZnO thin films. *Surf Rev Lett.* **2018**, *26(03)*, 1850167. https://doi.org/10.1142/S0218625X18501676.

[26] Gaur, L.K.; Gairola, P.; Gairola, S.P.; Mathpal, M.C.; Kumar, P.; Kumar, S.; Kushavah, D.; Agrahari, V.; Aragon, F.F.H.; Maria, A.G.S.; Swart, H.C. Cobalt doping induced shape transformation and its effect on luminescence in zinc oxide rod-like nanostructures. *J Alloys Compd.* **2021**, *868*, 159189. https://doi.org/10.1016/j.jallcom.2021.159189.

[27] Redwanul I, Suprio SS, Reana R, Nayeemul I, Torikul I. Unveiling the synthesis, characteristics, electrical conductivity, photocatalytic activity, and electrochemical activity of eco-friendly zinc oxide nanoparticles. *Adv. Sens. Energy Mater.* **2024**, *3*, 100105. https://doi.org/10.1016/j.asems.2024.100105.

[28] Shahroz, S.; Awais, K.; Zaid, M.A.; Thamer, A.; Arshad, A.; Abdul, J.; Yasmin Begum M.; Kandasamy G. A comparative analysis of optical and electrical properties of pure CuO and Zn doped CuO nanoparticles for optoelectronic device applications. *J Sol-Gel Sci Technol.* **2025**, *113*, 213-224. https://doi.org/10.1007/s10971-024-06591-7

ASEAN
Journal of Scientific and Technological Reports
Online ISSN:2773-8752

Research article

Bactericidal Properties, Biofilm Formation Inhibition, and Chemical Profiling of *Piper argyrites* and *Piper betel* L.

Saowakon Indoung[1], and Sirimaporn Watcharakul[2*]

[1] Faculty of Veterinary Sciences, Prince of Songkla University, Songkhla, 90110, Thailand
[2] Microbial Resources and Utilization Center, Songkhla Rajabhat University, Songkhla, 90000, Thailand
* Correspondence: Sirimaporn.wa@skru.ac.th

Citation:
Indoung, S.; Watcharakul, S. Bactericidal properties, biofilm formation inhibition, and chemical profiling of *Piper argyrites* and *Piper betel* L. *ASEAN J. Sci. Tech. Report.* **2025**, *28*(5), e259978. https://doi.org/10.55164/ajstr.v28i5.259978.

Article history:
Received: June 25, 2025
Revised: September 6, 2025
Accepted: September 10, 2025
Available online: September 21, 2025

Abstract: Betel leaf has been an important medicinal plant for therapeutic purposes since ancient times. This study aims to investigate the ethanolic leaf extracts of *Piper betel* (PB) and *Piper argyrites* (PA), native to Satun Province, Thailand. The antibacterial properties of the extracts were evaluated using a disc diffusion assay on four bacterial isolates, including *Escherichia coli* PT005, *Pseudomonas aeruginosa* PT00A2, *Klebsiella pneumoniae* PT00A1, and *Staphylococcus aureus* RTV01. PB leaf extracts demonstrated a substantial effect on *K. pneumoniae* PT00A1 with an inhibition zone of 25.77 ± 0.04 mm, while PA leaf extract showed potent activity against *E. coli* PT005 with an inhibition zone of 17.23±0.44 mm. In addition, both extracts exhibited bactericidal effects, with low MIC and MBC values ranging from 6.3 to 50 μg/mL, as determined by a broth dilution assay. Moreover, biofilm formation inhibition was evaluated using the crystal violet method. The results showed that both extracts effectively inhibited biofilm formation in bacterial strains. PA and PB extracts demonstrated high efficiency in reducing the biofilm formation of *S. aureus* RTV01, achieving reductions of 94.94% and 88.43%, respectively. The extracts showed a reduction in biofilm formation of 36.18% to 94.94% in the tested strains. The chemical profiles of the plants were analyzed using the GC-MS technique, which revealed that hydroxychavicol was a major constituent of both plant species, with varying amounts and distinct profiles. *P. argyrites* exhibited a more diverse range of compounds than *P. betel*. The findings suggest that these betel plant extracts need further investigation for potential therapeutic applications.

Keywords: Antibacterial agents; antibiotic resistance; biofilm; *Piper* spp.

1. Introduction

Bacterial infections pose a serious global health challenge, especially with the rise of multidrug resistance (MDR) in both Gram-negative strains, such as *E. coli*, *K. pneumoniae*, and *P. aeruginosa*, as well as Gram-positive strains like *S. aureus, which* can lead to various infections. *S. aureus* is particularly concerning as it can cause a wide range of illnesses, from minor skin infections to severe conditions like sepsis [1]. Similarly, Gram-negative bacteria often produce extended-spectrum beta-lactamases (ESBLs), which can inactivate extended-spectrum cephalosporins. Clavulanic acid has been proven effective in treating certain infections [2-3]. However, the presence of biofilm-forming pathogenic bacteria significantly contributes to their virulence and their connection to multidrug resistance (MDR). The biofilm structure provides several mechanisms that allow bacteria to survive and resist both antibiotic treatments and the host's immune responses. These mechanisms include creating a physical barrier, reducing metabolic activity, evading the immune system, facilitating gene transfer, and increasing mutagenesis. Infections caused by biofilm-forming bacteria are notoriously difficult to treat, often resulting in chronic, recurring infections that do not respond to conventional therapies. This resistance to treatment underscores the critical link between biofilm formation and the challenges of managing MDR, making it a crucial aspect of bacterial pathogenicity [4-6]. These infections contribute to extended hospital stays, increased disease complications, higher treatment costs, and higher mortality rates [2]. As a result, the search for natural compounds with antimicrobial properties has become a key focus of current research. The complex composition of natural antimicrobial agents makes it harder for microorganisms to develop resistance. The *Piper* genus, including *Piper betel* (cultivated betel, also known as Phlu-Ban) and *Piper argyrites* (wild betel, also known as Phlu-Pha), has been used as traditional medicine in Satun Province, Thailand, since ancient times to the present. Despite their traditional use, further research is needed to explore their pharmacological potential. Previous studies on *Piper* species have identified numerous bioactive compounds, including eugenol, hydroxychavicol, β-caryophyllene, methyl eugenol, chavibetol, chavicol, safrole, estragole, anethole, and iso-eugenol. The presence and properties of these compounds depend on extraction methods, cultivation regions, environmental conditions, and harvesting times [7]. Bioactive substances from *Piper* species exhibit significant antimicrobial, anti-inflammatory, antioxidant, anticancer, and antidiabetic properties. However, their composition varies depending on extraction methods and environmental factors [8]. This study aims to characterize the properties of two *Piper* species, *P. betel* and *P. argyrites*, that are native to Satun Province, Thailand. While *P. betel* has been extensively studied in other regions, data specific to Satun Province are limited. Additionally, this study is the first to investigate the chemical composition of *P. argyrites*, a wild betel species traditionally used by the Maniq people in Satun and surrounding regions for treating wounds, bruises, and skin conditions [9]. The antimicrobial activity and biofilm inhibition formation properties of both species were also evaluated.

2. Materials and Methods

2.1 Plant Materials

Whole plants of *P. argyrites* and *P. betel* were purchased from the local market in Satun, Thailand, from January to March 2024. The plants were identified and authenticated by an expert taxonomist, Assoc. Prof. Jarearnsak Sae Wai from the Division of Biological Science, Faculty of Science, Prince of Songkla University. The voucher specimens have been deposited at the PSU Herbarium with the collector numbers S. Watcharakul 7 and S. Watcharakul 9, corresponding to herbarium numbers 20538 and 20954. No specific permits were required, as the location is not protected.

2.2 Bacterial pathogens

Bacterial pathogens (*E. coli* PT005-ESBL, *P. aeruginosa* PT00A2-ESBL, *K. pneumoniae* PT00A1-ESBL, and *S. aureus* RTV01) were collected at the microbiology laboratory (PSU_Vet2, PSU, Thailand), briefly confirmed by the double-disk synergy method applied from Livermore and Brown (2001) [10]. All isolates were maintained on Mueller-Hinton Agar (MHA; pH 6.8-7.2) and stored in Cryostock mixed (MedEx™, EU) at -80 °C for further use.

2.3 Plant Extraction

The whole leaves of *P. argyrites* (PA) and *P. betel* (PB) were rinsed with water and dried in a hot air oven (Binder, Germany) at 60 °C for 3-5 days. They were then milled into a powder and extracted with ethanol (CH_3CH_2OH). Each ethanolic extract sample was evaporated using a Heidolph rotary evaporator (Hei-VAP HL, EU, Germany) with a Cooling Water Circulator (CF800, Yamato Scientific Co., Ltd., Japan) and stored at room temperature [11].

2.4 Antibacterial assays

2.4.1 Disc Diffusion Assay: The antibacterial activity was evaluated using the disc diffusion assay according to CLSI [12]. Briefly, the crude extract of the sample was dissolved in dimethyl sulfoxide (DMSO)

(VWR Chemicals BDH®, France) to a final concentration of 50 mg/mL and then applied to a sterile paper disc. All bacterial strains were activated at 35 °C for 18-24 hours and then inoculated on MHA. The crude extract paper disc was placed on the inoculum plate and incubated at 35 °C for 18-24 hours. The antimicrobial activity was evaluated based on measurements of the inhibition zones' appearance around the discs, with some modifications from Padumanonda and Phontree [13]. Gentamicin CN10 (10 µg, Oxoid, UK) antimicrobial susceptibility discs were used as a standard antibiotic, belonging to the aminoglycoside class, with a broad spectrum of activity against severe Gram-negative bacterial infections.

2.4.2 MIC/MBC: The antimicrobial activity was assessed by determining the Minimal Inhibitory Concentration (MIC) and Minimum Bactericidal Concentration (MBC). The assay tested eight different concentrations for each extract (0.390, 0.781, 1.563, 3.125, 6.25, 12.5, 25, 50 µg/mL). The tests were conducted in a 96-well plate using varying concentrations of sample extract and sterile Mueller-Hinton Broth (MHB) inoculated with bacterial concentrations of 10^5 cfu/mL, in accordance with CLSI guidelines (CLSI M07-A9, CLSI M100-S25). The plates were then incubated at 35 °C for 18-24 hours, as applied by Chanprapai et al. [11]. The MIC was determined as the lowest concentration of the extract that showed no visible bacterial growth. Subsequently, each well of MIC broth was spread over MHA plates and further incubated at 35 °C for 24 hours to determine the MBC, indicated by the absence of visible bacterial growth at the lowest concentration of each sample. The MIC values of each sample were analyzed using a one-way analysis of variance (ANOVA) and compared with those of the standard antibiotic gentamicin (Oxoid, UK).

2.5 Effect of Extracts on Biofilm Formation

The biofilm production was evaluated using the classical microbiological method described by Ngoc et al. (2024) [14] with some modifications. Briefly, the bacterial strains were activated in Tryptic Soy Broth (TSB) (Oxoid, UK) supplemented with 1% sucrose (Loba Chemie Pvt Ltd, India) at 37 °C for 18-24 hours. The extracts were prepared in TSB to obtain a final concentration of 50 mg/mL in sterile 96-well polystyrene flat-bottom microtiter plates (SPL Life Science Co., Ltd., Korea) after adding 100 µL of freshly prepared bacterial cell culture at 10^6 cells/mL corresponding to a density of 0.1 (OD600 nm) (Thermo Scientific™, Evolution™ 201/220 UV-Visible Spectrophotometers, USA). The inoculated 96-well microtiter plate was incubated at 37 °C without shaking. After 4 hours of incubation, 100 µL aliquots of diluted extracts were added to the wells of the 96-well microtiter plates. Gentamicin (Oxoid, UK) was used as a positive control, while DMSO was used as a solvent. After incubation at 37 °C for 24 hours, non-adherent cells were gently removed, and the microtiter plate was washed with sterile phosphate-buffered saline (PBS) (Oxoid, UK). The samples were then fixed at 60 °C for 1 hour. Biofilms and attached cells were stained using a 0.1% crystal violet solution (Clinag Co., Ltd., TH). After 15 minutes of the staining procedure, the samples were rinsed with sterile distilled water and dried. For the quantitative evaluation of biofilm production, a 30% acetic acid (125 µL) solution was added to destain the biofilm-associated crystal violet, and the solution's absorbance was measured. at OD 492 nm. The percentage of biofilm inhibition was determined using the following formula:

$$\%\text{Biofilm reduction} = \frac{\text{ODcontrol-ODsample}}{\text{ODcontrol}} \text{ x } 100 \tag{1}$$

Where ODcontrol is the absorbance of the bacteria without the extract solution, and ODsample is the absorbance of the extract solution with bacteria.

2.6 GC-MS analysis

The chemical composition of the ethanolic extracts from the samples was analyzed using a Gas Chromatograph Triple Quadrupole Spectrometer (GC-MS/MS) model Agilent 7000D Triple Quadrupole GC/MS. The separation was performed on a capillary column HP-5ms Ultra Inert (Agilent 19091S-433UI) with dimensions of 30 m × 250 µm × 0.25 µm. The carrier gas was helium, with a constant flow rate of 1 mL/min. The oven temperature was programmed to range from 60°C to 320°C at a rate of 5°C/min, then held for 10 minutes. at 300°C. A sample volume of 1 µL was injected using a splitless mode with a flow rate of 3 mL/min. The mass spectrometer was set to scan in the m/z 35–500 range with electron impact (EI) ionization mode. The relative percentage of the components was represented as the percentage using peak area normalization. Identifying compounds using a spectral library in MassHunter Qualitative Analysis version 10.0. The relative proportions were calculated by dividing the individual peak area by the total area of all peaks. Only compounds over 1% were included.

2.7 Statistical Analysis

The data were obtained in triplicate and expressed as means ± SD. IBM SPSS Statistics version 29.0.1.0 (179) for Windows was used for data analysis. The significance of the means was compared at $p < 0.05$ using a one-way analysis of variance (ANOVA).

3. Results and Discussion

3.1 Plant Extracts

Ethanol was used as a solvent for extracting compounds from betel leaves due to its various advantages over other solvents. As a medium-polarity solvent, it is capable of extracting a diverse range of betel leaf compounds, including both polar compounds such as flavonoids and non-polar compounds like chavicol and eugenol, which are recognized as the major bioactive components (see Table 4). Moreover, ethanol is considered a safe solvent for producing extracts intended for the food, pharmaceutical, and cosmetic industries, as it is classified as "Generally Recognized As Safe (GRAS)" by the U.S. Food and Drug Administration (FDA). This safety profile makes it particularly suitable for commercial applications.

The average percentage yield of the ethanolic leaf extracts from *P. betel* and *P. argyrites* was 0.77 and 0.84, respectively. The crude extracts exhibited distinct characteristics, as shown in Figure 1. The ethanolic leaf extracts of *P. betel* were semisolid and dark green, while those of *P. argyrites* were reddish-brown with a liquid texture. This finding is consistent with a previous study on the extraction of leaves from the *Piperaceae* family [7], which reported that members of the *Piperaceae* family contain various substances due to species differences.

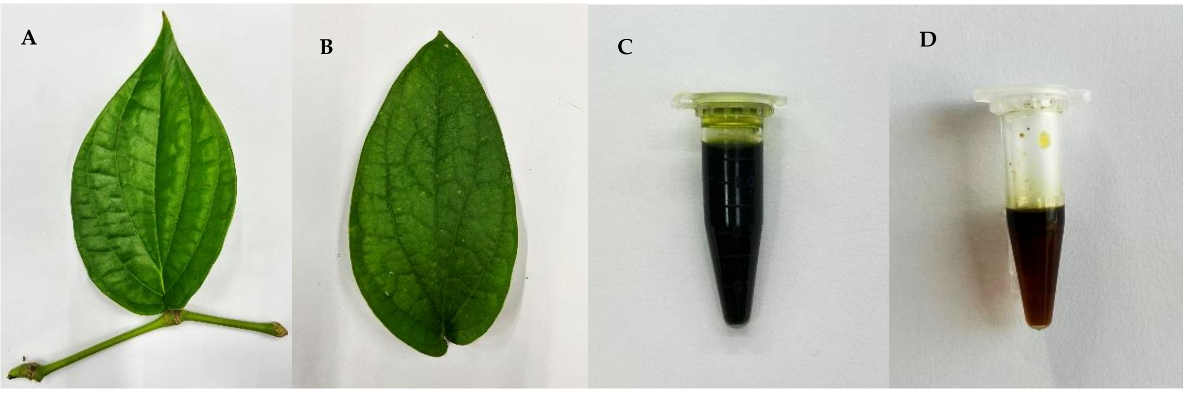


Figure 1. Characteristics of *P. betel* leaf (A), *P. argyrites* (B), the ethanolic leaf extract of *P. betel* (C), and *P. argyrites* (D)

3.2 Antibacterial Properties

3.2.1 Disc Diffusion Assay

The disc diffusion method was used to evaluate the antibacterial activity of all extracts. The results demonstrated their inhibitory effects on all tested strains, as shown in Table 1 and Figure 2. Notably, PB exhibited potent activity against *K. pneumoniae* PT00A1 and *E. coli* PT005, with inhibition zones of 25.77 ± 0.04 mm and 22.09 ± 0.23 mm, respectively. These results were very close to or even better than the gentamicin CN10 standard antibiotic (10 µg, Oxoid, UK), which produced inhibition zones of 24.41 ± 0.60 mm and 20.11 ± 0.32 mm against the same bacteria. However, its activity against *P. aeruginosa* PT00A2 and *S. aureus* RTV01 was minimal, with inhibition zones of 7.78 ± 0.18 mm and 12.15 ± 0.42 mm, as shown in Table 1. Moreover, PA showed significant antibacterial activity against *E. coli* PT005, with an inhibition zone of 17.23 ± 0.44 mm. The inhibition zones for *P. aeruginosa* PT00A2, *K. pneumoniae* PT00A1, and *S. aureus* RTV01 were 7.78±0.18 mm, 10.84±0.48 mm, and 12.15±0.42 mm, respectively. The *P. argyrites* extract was most effective against *E. coli*, but its overall activity was less potent than that of the *P. betel* extract and gentamicin.

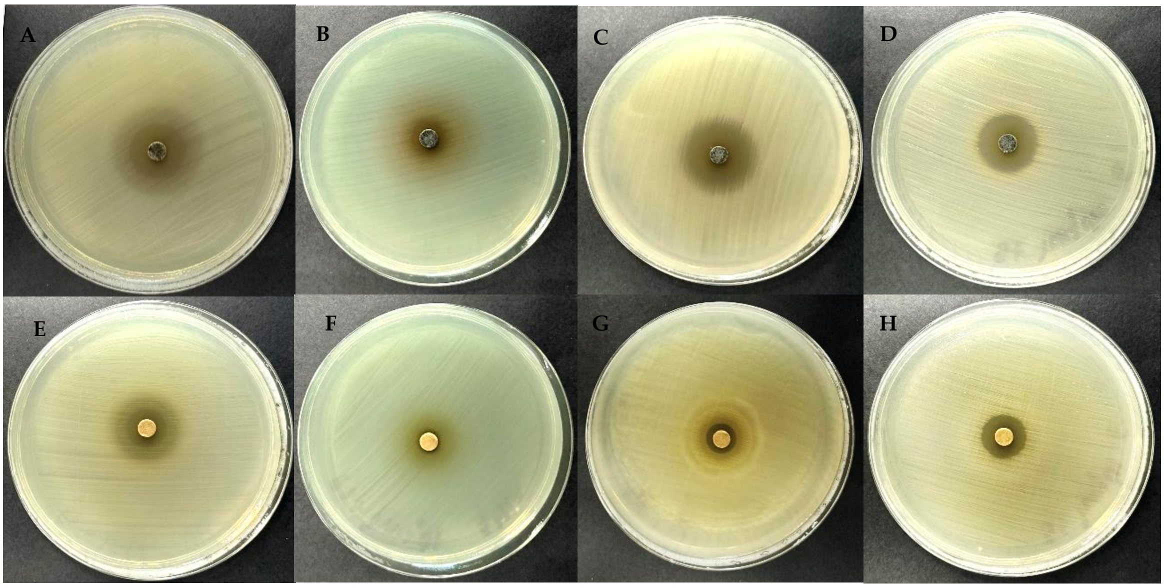


Figure 2. Disc diffusion assay of *P. betel* (A-D) and *P. argyrites* (E-H) ethanolic leaf extract against bacterial growth on MHA. (A, E) *E. coli* PT005, (B, F) *P. aeruginosa* PT00A2, (C, G) *K. pneumoniae* PT00A1, (D, H) *S. aureus* RTV01

Table 1. Antimicrobial activity of *P. betel* and *P. argyrites* ethanolic leaf extract.

Bacterial Strains	**Inhibition zone diameter (mm)**		
	P. betel	***P. argyrites***	**Gentamicin CN10**
E. coli PT005	22.09 ± 0.23 ^{a}	17.23 ± 0.44 ^{c}	20.11 ± 0.32 ^{b}
P. aeruginosa PT00A2	10.94 ± 0.31 ^{b}	7.78 ± 0.18 ^{c}	12.18 ± 0.29 ^{a}
K. pneumoniae PT00A1	25.77 ± 0.04 ^{a}	10.84 ± 0.48 ^{c}	24.41 ± 0.60 ^{b}
S. aureus RT0V1	15.42 ± 0.21 ^{a}	12.15 ± 0.42 ^{b}	11.66 ± 0.16 ^{c}

Note: Values, an average of the mean inhibition zones (mm) ± standard deviation (SD) of duplicates of the extracts at 50 mg/disc of each isolate. Mean values within rows followed by a different letter differ significantly ($p<0.05$; DMRT).

3.2.2 MIC/MBC

The ethanolic leaf extracts of PB and PA demonstrated significant antibacterial activities against all tested strains in the broth microdilution assay. MIC and MBC values ranged from 6.25 to 50 µg/mL, indicating potent antibacterial effects compared with the standard antibiotic gentamicin, as shown in Table 2. PB extracts displayed intense antibacterial activity against *E. coli* PT005 and *K. pneumoniae* PT00A1, with MIC values of 6.25 µg/mL, while the MIC values for *P. aeruginosa* PT00A2 and *S. aureus* RTV01 were 25 µg/mL and 12.5 µg/mL, respectively. Moreover, PA extracts demonstrated MIC results with lower antibacterial activity than PB extracts for all tested strains. MIC values for *E. coli* PT005 and *S. aureus* RTV01 were 12.5 µg/mL, whereas the MIC values for *P. aeruginosa* PT00A2 and *K. pneumoniae* PT00A1 were 50 µg/mL and 25 µg/mL, respectively (Table 2). These results represented a MIC index of ≤ 2 (MBC/MIC ratio) for both PB and PA extracts, suggesting a bactericidal mode of action. The MIC index has been used to determine whether a substance possesses bactericidal properties; for bacteriostatic properties, the MIC index is ≤ 4 [11]. The results clearly indicate that both extracts possess strong antibacterial properties, with low MIC and MBC values, suggesting that these extracts can effectively combat bacterial strains according to the MIC index. This confirms their bactericidal nature, as the extracts inhibit bacterial growth and kill bacterial strains directly. Additionally, the extracts were effective against both Gram-positive (*S. aureus* RTV01) and Gram-negative (*E. coli* PT005, *K. pneumoniae* PT00A1, *P. aeruginosa* PT00A2) bacteria, suggesting broad-spectrum antibacterial activity. In view of these findings, both extracts demonstrated a bactericidal mode of action with MIC and MBC values ranging from 6.25 to 50 µg/mL. The PB extract was particularly effective, as evidenced by its MIC values of 6.25 μg/mL for *E. coli* PT005 and *K. pneumoniae* PT00A1. The MIC value is comparable to that of gentamicin against *K. pneumoniae*, indicating that the PB extract has a similar bactericidal effect to gentamicin on these specific bacterial strains. However, based on previous studies on the antibacterial activity of the *Piperaceae* family, the antimicrobial activity of mature leaves is effective against *E. coli*, *Streptococcus pyogenes*, *P. aeruginosa*, *S. aureus*, *Proteus vulgaris*, and *Salmonella* spp., with low MIC and MBC values, as presented in the bactericidal activity. [15-17]. The efficiency of extracting substances varies in terms of antibacterial activity and main chemical components due to various environmental factors, including the location and harvest time of the plant.

Table 2. The MIC and MBC of the plant's ethanolic leaf extract

Strains	***P. betel***			***P. argyrites***			**Gentamicin**		
	MIC	**MBC**	**MBC/MIC ratio**	**MIC**	**MBC**	**MBC/MIC ratio**	**MIC**	**MBC**	**MBC/MIC ratio**
E. coli PT005	6.25	6.25	1	12.5	12.5	1	12.5	12.5	1
P. aeruginosa PT00A2	25	25	1	50	50	1	12.5	25	2
K. pneumoniae PT00A1	6.25	12.5	2	12.5	25	2	6.25	12.5	2
S. aureus RTV01	12.5	25	2	12.5	25	2	12.5	25	2

Note: The MBC/MIC ratio, where MBC is the minimum bactericidal concentration and MIC is the minimum inhibitory concentration, with an MBC/MIC ratio ≤ 4, indicates that a drug is considered bacteriostatic.

3.3 Biofilm Formation Reduction

The biofilm formation inhibition was carried out in microtiter plates using the classical crystal violet method [14] with some modifications. The ethanolic leaf extracts of PB and PA effectively inhibited biofilm formation in all tested bacterial strains (Table 3). At the same concentrations, PB and PA extracts demonstrated strong inhibitory effects on the biofilm formation of *S. aureus* RTV01, with reductions of 88.43% and 94.94%, respectively. Additionally, PA extracts reduced the biofilm formation of *K. pneumoniae* PT00A1, *E. coli* PT005, and *P. aeruginosa* PT00A2 by 88.35%, 88.05%, and 36.54%, respectively. In comparison, PB extracts exhibited lower inhibitory effects on the biofilm formation of *K. pneumoniae* PT00A1 and *E. coli* PT005 than PA extracts, with reductions of 36.54% and 79.88%, respectively. In contrast, the effect of PB extracts on reducing biofilm formation was higher in *P. aeruginosa* PT00A2, at 46.44%. Gentamicin was used as a positive control, as represented in Table 3. Based on the findings of this study, the percentage of biofilm formation reduction ranged from 36.18% to 94.94% for all tested strains, highlighting the significant antibiofilm potential of both plant extracts. Notably, the PA extract showed a slightly higher efficiency in reducing the formation of *S. aureus* RTV01 biofilm compared to the PB extract and standard antibiotic gentamicin.

However, the *in vitro* study revealed varying inhibitory effects of extracts on biofilm formation reduction for each strain. These variations could be due to differences in the mechanisms of biofilm formation among the different strains. The biofilm formation process involves several stages: attachment, multiplication, exodus, maturation, and dispersion. Moreover, the type of biofilm also varies according to bacterial functions, such as adhesion, metabolic pathways, and the synthesis of extracellular polymeric substances (EPS) [18]. However, the chemical composition of plants is responsible for their antibacterial properties; even if the same plant genus were harvested under the same environmental conditions, differences in phytochemical contents could affect their antimicrobial effects. Based on publications on the antimicrobial activity of plant extracts, *P. betel* has been shown to affect the biofilm formation of *Serratia marcescens* and *S. aureus* by reducing exopolysaccharide production and the hydrophobicity index at the early stages of biofilm formation, thereby postponing the lag time and slowing down the growth rate [19]. This information suggests that *P. betel* and *P. argyrites* ethanolic leaf extracts likely inhibit biofilm formation at the early adherence and multiplication stages of *S. aureus* and other tested strains due to their enhanced antibacterial properties. These findings indicate the significant antibiofilm formation activity of the examined extracts, making them a promising alternative therapy and highlighting the potential of both *P. argyrites* and *P. betel* as sources of new antimicrobial drugs. *P. betel* is a powerful antibacterial agent, while the *P. argyrites* extract may be more valuable due to its antibiofilm properties. Their most significant potential lies in their dual-action ability to both kill bacteria and inhibit biofilm formation, a crucial advantage in the face of growing antibiotic resistance. Furthermore, extensive traditional and emerging scientific evidence suggests their value extends beyond antimicrobial use, indicating they could be developed into treatments for various conditions, including cancer, inflammation, and digestive ailments [7]. These findings warrant further research and development to explore their full therapeutic potential in modern medicine.

Table 3. Antibiofilm formation of plant ethanolic leaf extract at a concentration of 50 mg/mL

Bacterial Strains	**Biofilm reduction (%)**		
	P. betel	***P. argyrites***	***Gentamicin***
E. coli PT005	79.88 ± 0.41^{b}	88.05 ± 0.55^{a}	69.97 ± 0.52^{c}
P. aeruginosa PT00A2	46.44 ± 0.35^{a}	36.18 ± 0.61^{c}	42.17 ± 0.41^{b}
K. pneumoniae PT00A1	36.54 ± 0.54^{c}	88.35 ± 0.74^{a}	75.33 ± 0.69^{b}
S. aureus RT0V1	88.43 ± 0.84^{c}	94.94 ± 0.58^{a}	92.89 ± 0.63^{b}

Note: Data shown as mean ± SD values from triplicate analysis
Mean values within rows followed by a different letter differ significantly ($p<0.05$; DMRT).

3.4 Chemical Compositions of Extracts

The chemical composition analysis using GC-MS of PA and PB ethanolic leaf extracts revealed distinct compositional profiles. PA exhibited a more diverse chemical composition than PB extracts, containing hydroxychavicol (24.8%), elemenes (5.45%), caryophyllene (5.02%), and many other constituents (Table 4). Under the same conditions, PB extracts revealed dominant compounds, including hydroxychavicol (64.01%), eugenol (22.40%), and phytol (1.43%), with minor components comprising less than 0.01% of the total, as shown in Table 4. Notably, hydroxychavicol was a main component in both plant species, aligning with their classification within the *Piper* genus. While both plant species shared these common compounds, the quantitative distribution differed significantly, with PB exhibiting approximately three times higher levels of hydroxychavicol than PA extract; conversely, PA extracts demonstrated a more varied chemical composition.

As stated above, it was revealed that the specific antibacterial and antibiofilm activities of the two extracts are influenced by their unique chemical profiles. Although *P. betel* contains a significantly higher concentration of hydroxychavicol, the more diverse composition of *P. argyrites* appears to give it a greater efficacy in certain areas, such as biofilm reduction in *S. aureus* RTV01. This finding suggests that the synergistic effects of various compounds, rather than the concentration of a single component, play a crucial role in the plant's biological activity.

According to previous studies on bioactive substances in the *Piper* genus, particularly in betel leaves, the presence of phytol, hydroxychavicol, eugenol, and caryophyllene is a significant constituent, although with varying concentrations across different cultivation regions and even within the same species. These variations in the chemical composition of betel leaves depend on environmental factors such as soil type, nutrient content, and climatic conditions, which influence the chemical composition profiles and, consequently, the plants' biological activities [20]. This report reveals the chemical composition profiling of wild betel (PA) ethanolic leaf extract for the first time. Remarkably, PA and PB leaf extracts contain natural antioxidant properties due to their polyphenol compounds, such as hydroxychavicol [21]. Additionally, the presence of various bioactive components, such as eugenol, phytol, caryophyllene, and terpenes [22], exhibits antimicrobial properties against *P. aeruginosa, S. aureus, E. coli, Streptococcus pyogenes*, and *Proteus vulgaris* [15, 17, 23]. The important bioactive compound in the plant extracts (PA and PB), hydroxychavicol, has been reported in previous studies to possess potent antibacterial properties, with a mechanism of action in *E. coli* involving the stripping of magnesium ions, inducing oxidative stress followed by membrane and DNA damage, and inducing cell permeabilization, which inhibits and damages the outer membrane that is responsible for resistance to several antibiotics [24]. Moreover, bioactive compounds in the *Piper* genus have been reported to possess anticancer properties responsible for the antiproliferation of cancer cells [25-26], as well as anti-inflammatory activities and antidiabetic properties [7], highlighting their significant pharmacological potential. However, more research is needed to evaluate their efficacy and safety for therapeutic purposes.

Table 4. Chemical composition of the ethanolic leaf extract of plants

Plants	Retention time	Compounds Name	Percent of the total*	Chemical formula
P. argyrites	14.5866	L-α-Terpineol	2.19	$C_{10}H_{18}O$
	20.2880	Elemene	5.45	$C_{15}H_{24}$
	20.9998	Caryophyllene	5.02	$C_{15}H_{24}$
	21.8594	1,5,9,9-tetramethyl-1,4,7-Cycloundecatriene	2.5	$C_{15}H_{24}$
	22.4159	Hydroxychavicol	24.8	$C_9H_{10}O_2$
	22.9071	γ-Selinene	3.88	$C_{15}H_{24}$
	30.6094	Hexahydro-3H-benzofuran-2-one	4.39	$C_{15}H_{20}O_2$
	36.2518	α-Linolenic acid	4.12	$C_{18}H_{30}O_2$
	54.7662	γ-Tocopherol	2.95	$C_{29}H_{50}O_2$
P. betel	19.6819	Eugenol	22.41	$C_{10}H_{12}O_2$
	20.9941	Caryophyllene	1.65	$C_{15}H_{24}$
	22.4163	Hydroxychavicol	64.01	$C_9H_{20}O_2$
	35.7476	Phytol	1.43	$C_{20}H_{40}O$

* % of the total is implied to be the percent of the total in terms of the peak area relative to the total peak area.

4. Conclusion

The ethanolic leaf extracts of *P. argyrites* and *P. betel* showed potent antimicrobial properties and a substantial effect on the reduction of biofilm formation in MDR isolates and common pathogens, including *E. coli* PT005, *P. aeruginosa* PT00A2, *K. pneumoniae* PT00A1, and *S. aureus* RTV01. In addition, the extracts exhibited hydroxychavicol as a major constituent in significantly different amounts, followed by minor compounds, including eugenol, elemenes, caryophyllene, and various other compounds. This report presents the chemical composition of *P. argyrites* leaf extract, as determined by the GC-MS technique, and highlights its potent antimicrobial properties. Overall, this study suggests that *P. argyrites* and *P. betel* are rich sources of potent bioactive substances, which could potentially revolutionize the fields of alternative therapy and pharmacology.

5. Acknowledgements

This research was supported by Thailand Science Research and Innovation Fund (Fundamental Fund: FF 2567, contract No. 9/2567). The authors are grateful to the Faculty of Veterinary Sciences at Prince of Songkla University for their cooperation in providing microbiology laboratory facilities (BSL 2 Laboratory). We appreciate Assoc. Prof. Jarearnsak Sae Wai from the Division of Biological Science, Faculty of Science, Prince of Songkla University, for identifying and authorizing *P. argyrites* and *P. betel*. We would also like to thank Mr. Ralph Tabasa Torres, Songkhla Rajabhat University, for being the English proofreader and revising the manuscript. Finally, we would like to thank the Institute of Research and Development at SKRU for their exceptional support.

Author Contributions: Conceptualization, S.W. and S.I.; methodology, S.W. and S.I.; software, S.W.; validation, S.W.; formal analysis, S.W.; investigation, S.W. and S.I.; resources, S.W. and S.I.; data curation,

S.W.; writing—original draft preparation, S.W.; writing—review and editing, S.W.; visualization, S.W. and S.I.; supervision, S.W. and S.I.; project administration, S.W. and S.I.; funding acquisition, S.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Thailand Science Research and Innovation Fund (Fundamental Fund: FF 2567, contract No. 9/2567).

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] Cheung, C. Y. G.; Bae, J. S.; Otto, M. Pathogenicity and virulence of Staphylococcus aureus. *Virulence* **2021**, *12*(1), 547–569. https://doi.org/10.1080/21505594.2021.1878688

[2] Kondo, S.; Apisarnthanarak, A.; Trakulsomboon, S.; Bootkotr, W.; Mingmalairak, C.; Mahawongkajit, P.; Juntong, J.; Limpavitayaporn, P.; Sriussadaporn, E.; Tongyoo, A.; Boonyasatid, P. Prevalence of extended-spectrum β-lactamase-producing Enterobacterales and distribution of blaESBL genes from patients who underwent abdominal surgery. *Sci. Technol. Asia* **2022**, *27*(4), 104–116.

[3] Sheu, C. C.; Lin, S. Y.; Chang, Y. T.; Lee, C. Y.; Chen, Y. H.; Hsueh, P. R. Management of infections caused by extended-spectrum β-lactamase-producing Enterobacteriaceae: current evidence and future prospects. *Expert Review of Anti-infective Therapy* **2018**, *16*(3), 205–218. https://doi.org/10.1080/14787210.2018.1436966

[4] Preda, V. G.; Săndulescu, O. Communication is the key: biofilms, quorum sensing, formation and prevention. *Discov.* **2019**, *7*(3), e100. https://doi.org/10.15190/d.2019.13

[5] Shree, P.; Singh, C. K.; Sodhi, K. K.; Surya, J. N.; Singh, D. K. Biofilms: Understanding the structure and contribution towards bacterial resistance in antibiotics. *Med. Microecol.* **2023**, *16*, 100084. https://doi.org/10.1016/j.medmic.2023.100084

[6] Mendes, S. G.; Combo, S. I.; Allain, T.; Domingues, S.; Buret, A. G.; Da Silva, G. J. Co-regulation of biofilm formation and antimicrobial resistance in Acinetobacter baumannii: from mechanisms to therapeutic strategies. *Eur. J. Clin. Microbiol. Infect. Dis.* **2023**, *42*(12), 1405–1423. https://doi.org/10.1007/s10096-023-04677-8

[7] Singh, T.; Singh, P.; Pandey, V. K.; Singh, R.; Dar, A. H. A literature review on bioactive properties of betel leaf (Piper betle L.) and its applications in food industry. *Food Chemistry Advances* **2023**, *1*(3), 100536. https://doi.org/10.1016/j.focha.2023.100536

[8] Nayaka, N. M.; Sasadara, M. M.; Sanjaya, D. A.; Yuda, P. E.; Dewi, N. L.; Cahyaningsih, E.; Hartati, R. Piper betle (L): Recent review of antibacterial and antifungal properties, safety profiles, and commercial applications. *Molecules* **2021**, *26*(8), 2321. https://doi.org/10.3390/molecules26082321

[9] Carsono, N.; Tumilaar, S. G.; Kurnia, D.; Latipudin, D.; Satari, M. H. A review of bioactive compounds and antioxidant activity properties of Piper species. *Molecules* **2022**, *27*(19), 6774. https://doi.org/10.3390/molecules27196774

[10] Livermore, D. M.; Brown, D. F. J. Detection of β-lactamase-mediated resistance. *J. Antimicrob. Chemother.* **2001**, *48*, 59–64.

[11] Chanprapai, P.; Kubo, I.; Chavasiri, W. Anti-rice pathogenic microbial activity of Persicaria sp. extracts. *Sci. Technol. Asia* **2018**, *18*, 32–41. https://doi.org/10.14456/scitechasia.2018.30

[12] Clinical and Laboratory Standards Institute. *Methods for dilution antimicrobial susceptibility tests for bacteria that grow aerobically*. 11th ed. CLSI standard M07; Wayne, PA, USA, **2018**.

[13] Padumanonda, T.; Phontree, K. Total phenolic, total flavonoid, total condensed tannin contents and antimicrobial activity against diarrheal bacteria of the bark and fruit of Terminalia nigrovenuloa Pierre ex Laness. *Sci. Technol. Asia* **2021**, *16*, 162–169.

[14] Ngoc, L. T.; Hung, H. Y.; Tu, Q. N. Inhibitory effects of Sterculia lychnophora hance ethanol extracts on biofilm formation and virulence factors of Pseudomonas aeruginosa. *J. microbiol., biotechnol. food sci.* **2024**, *13*(6), e10483. https://doi.org/10.55251/jmbfs.10483

[15] Pradhan, D.; Suri, K. A.; Pradhan, D. K.; Biswasroy, P. Golden heart of the nature: Piper betle L. *J. Pharmacogn. Phytochem.* **2013**, *1* (6), 147–167.

[16] Boripun, R.; Paopradit, P.; Prampramote, J.; Narinthorn, R.; Intongead, S.; Sangkanu, S.; Narongrit, T.; Mitsuwan, W. Bactericidal activity of Piper betle L. extract against antibiotic resistant Salmonella spp. Isolated from pig farms in southern Thailand. *Vet. Integr. Sci.* **2022**, *20*(3), 557–569. https://doi.org/10.12982/VIS.2022.042

[17] Stuper-Szablewska, K.; Szablewski, T.; Przybylska-Balcerek, A.; Szwajkowska-Michałek, L.; Krzyżaniak, M.; Świerk, D.; Cegielska-Radziejewska, R.; Krejpcio, Z. Antimicrobial activities evaluation and phytochemical screening of some selected plant materials used in traditional medicine. *Molecules* **2023**, *28*(1), 244. https://doi.org/10.3390/molecules28010244

[18] Li, Y.; Ni, M. Regulation of biofilm formation in Klebsiella pneumoniae. *Front. Microbiol.* **2023**, *14*, 1238482. https://doi.org/10.3389/fmicb.2023.1238482

[19] Lao, R. C.; Yabes, A. M.; Tobias-Altura, M.; Panganiban. L. C.; Makalinao, I. R. In vitro antibacterial and antibiofilm activities of Piper betle L. ethanolic leaf extract on Staphylococcus aureus ATCC 29213. *Acta Medica Philippina* **2023**, *57*(12), 53–60. https://doi.org/10.47895/amp.vi0.6412

[20] Das, S.; Parida, R.; Sandeep, I. S.; Kar, B.; Nayak, S.; Mohanty, S. Chemical composition and antioxidant activity of some important betel vine landraces. *Biologia* **2016**, *71*(2), 128–132. https://doi.org/10.1515/biolog-2016-0030

[21] Kaintura, P.; Bhandari, M.; Kumar, R. Medicinal values of betel leaves and its application in food products: A review. *Innov. Pharm.* **2020**, *9*(6), 344–348.

[22] Madhumita, M.; Guha, P.; Nag, A. Extraction of betel leaves (Piper betle L.) essential oil and its bio-actives identification: Process optimization, GC-MS analysis and antimicrobial activity. *Ind. Crops. Prod.* **2019**, *138*, 111578. https://doi.org/10.1016/j.indcrop.2019.111578

[23] Islam, M. A.; Ryu, K. Y.; Khan, N.; Song, O. Y.; Jeong, J. Y.; Son, J. H.; Jamila, N.; Kim, K. S. Determination of the volatile compounds in five varieties of Piper betle L. from Bangladesh using simultaneous distillation extraction and gas chromatography/mass spectrometry (SDE-GC/MS). *Anal. Lett.* **2020**, *53*(15), 2413–2430. https://doi.org/10.1080/00032719.2020.1744160

[24] Singh, D.; Majumdar, A. G.; Gamre, S.; Subramanian, M. Membrane damage precedes DNA damage in hydroxychavicol treated E. coli cells and facilitates cooperativity with hydrophobic antibiotics. *Biochimie* **2021**, *180*, 158–168. https://doi.org/10.1016/j.biochi.2020.11.008

[25] Jackson, J.; Romero, G.; Hawkins, D.; Cornwall, R. G.; Lukov, G. L. Identifying the structural components responsible for the antiproliferative properties of hydroxychavicol. *Compounds* **2023**, *3*(4), 552–560. https://doi.org/10.3390/compounds3040039

[26] Poomirat, S.; Jaiarree, N.; Itharat, A.; Ruangnoo, S. Cytotoxicity against cervical and breast cancer cells of Leard-Ngam remedy and its plant compositions. *Sci. Technol. Asia* **2020**, *25*(3), 38–50.

ASEAN
Journal of Scientific and Technological Reports
Online ISSN:2773-8752

Academic article

Experimental Investigation of Pitting Corrosion Behavior of 304L Stainless Steel on MnS Inclusions in Chloride Environments Applied to the Mediterranean Industry

Kamel Benlouanas[1, 2*], and Lazhar yahia[3]

[1] Department of Genius-Mechanics, Mohamed Khider University, BP 145 RP, Biskra, 07000, Algeria
[2] Laboratory of Civil and Hydraulic Engineering for Sustainable Development and Environment "LARGHYDE", Mohamed Khider University, BP 145 RP, Biskra, 07000, Algeria
[3] Faculty of Technology, Mostepha Benboulaid University, 53 road of Constantine Fesdis, Batna, 05078, Algeria
* Correspondence: kamel.benlouanas@univ-biskra.dz

Citation:
Benlouanas, K.; Yahia, L.; Experimental investigation of pitting corrosion behavior of 304L stainless steel on MnS inclusions in chloride environments appliced to the mediterranean industry. *ASEAN J. Sci. Tech. Report* **2025**, *28*(5), e258484. https://doi.org/10.55164/ajstr.v28i5.258484.

Article history:
Received: March 27, 2025
Revised: July 17, 2025
Accepted: August 16, 2025
Available online: September 21, 2025

Abstract: Structures made of stainless steel experience a gradual deterioration in their fundamental properties when exposed to both mechanical stress and harsh environmental conditions. This study aims to analyze the impact of cold tensile deformation on the localized roughness corrosion of 304L stainless steel in a 3% NaCl solution, mimicking seawater conditions. Corrosion tests were performed on samples obtained from standardised tensile specimens of the Public Economic Enterprise for the Production of Bolts, Cutlery, and Faucets (BCR) in Boumerdes, Algeria, which had experienced deformation at strain levels of 2.18%, 3.63%, 10.90%, and 16.36%. The results, including corrosion susceptibility, pitting behavior, and repassivation potentials, were evaluated and compared based on the strain. Findings indicate that all measured potentials decrease as the strain increases, except for the roughness potential, which shows a significant decline. This suggests a notable reduction in the material's corrosion resistance with higher deformation levels.

Keywords: Pitting corrosion; 304L stainless steel; effect; deformation; traction

1. Introduction

excellent resistance to uniform corrosion. This corrosion resistance is primarily attributed to the formation of a stable, self-healing passive oxide film on the surface. Among these materials, austenitic stainless steels, such as 304L, offer an optimal balance between mechanical strength, ductility, and corrosion resistance [1, 2]. This study aims to address this gap by experimentally evaluating the effect of cold deformation on the pitting corrosion behavior of 304L stainless steel. Tests were made on specimens of the public economic enterprise for the production of bolts, cutlery, and faucets (BCR) in Boumerdes, Algeria. Corrosion in metals is an electrochemical process where metals return to their thermodynamically stable oxidized state when exposed to oxidizing environments. In steel, iron reacts with oxygen and moisture to form rust, which can detach and expose new surfaces to further corrosion. While general corrosion is relatively predictable, localized corrosion mechanisms such as pitting are more complex and severe.

Pitting corrosion involves the localized breakdown of the passive layer, leading to rapid, concentrated attack at discrete sites. It requires careful study of electrochemical kinetics, metallurgical microstructures, and mechanical factors. Although the electrochemical behavior of 304L stainless steel in chloride-containing environments has been extensively investigated [3, 4, 5], limited research has addressed how prior cold deformation influences its susceptibility to localized pitting corrosion. The corrosion resistance of stainless steels is mainly due to their chromium content, typically above 10.5%, which ensures the formation of a protective chromium oxide layer. Alloying elements such as nickel stabilize the austenitic structure and improve toughness and ductility. Additions of molybdenum and titanium can further enhance resistance to localized corrosion and high-temperature degradation [6, 7]. However, the passive film in stainless steels is not immune to breakdown. Aggressive ions such as chlorides can penetrate through defects, inclusions, or microstructural heterogeneities, leading to localized film failure. Cold deformation increases dislocation density and residual stress, which can serve as preferential sites for passive film breakdown and pit initiation. Understanding the combined effect of mechanical deformation and corrosive environments is essential for predicting service performance and ensuring structural integrity.

2. Materials and methods

Stainless steel is categorised into five principal types and exists in over one hundred variations. Diverse applications of temperature regulation for austenitic, heat-resistant materials; Ideal for elevated temperatures: ferritic, exhibiting resistance to corrosion and cracking; Martensitic; durable; Duplex (a combination of ferritic and austenitic steel) provides a wide range of options; Semi-austenitic (precipitation-hardened) alloys, composed of aluminium, niobium, and copper, exhibit considerable strength [8, 9]. The advantages of stainless steel include robust corrosion resistance, straightforward and efficient maintenance, durability and strength, aesthetic simplicity, and minimum environmental impact. Stainless steel possesses various disadvantages, including relatively high cost, lack of environmental friendliness, potential for allergic reactions, susceptibility to smudges, and non-biodegradability [10, 11].

2.1 Utilized Material

The material utilised in this study is the austenitic steel 304L, provided and shaped in the form of sheet metal by the company BCR from Bordj M'Nail (Boumerdes), showing its chemical composition [12]:

Table 1. Chemical composition of the sampling

Chemical components (%)						
Fe	Cr	Ni	Si	Mn	Mo	C
Base	18.78	8.80	0.45	1.34	0,27	0.06

2.2 Applied Material

Considered as 18-8 stainless steel, stainless steel 304 (AISI 304) is the most often used variant with 18-20% Cr and 8-10.5% Ni. This 304 is non-magnetic at annealed circumstances, but following cold processing (such as drawing, bending, rolling...), a part of the austenitic structure is changed into a mildly magnetic martensite. AISI 304 material provides high resistance to air corrosion and oxidation. Its high nickel-chrome alloy concentration offers it good corrosion resistance, making it suitable for end use in the food sector [13- 14].

2.2.1 Tensility assays

The tensile trials were conducted using a Zwick-type equipment at the University of Jijel. It features a digital system linked to a computer that facilitates the graphing of tensile curves and yields results like maximum load, elongation at rupture, etc. [15].

2.2.2 Corroding evaluations

Electrochemical studies were done utilising a computer-controlled EG&G Model 283 potentiostat. The I(E) polarization curves were made with Soft CorrIII software.

At the start of the experiment, the metal works as the cathode of the cell. The applied voltage grows positively, and at a predetermined threshold, the metal switches to the anode of the cell. The ensuing corrosion morphology mimics that of natural corrosion when the specified potential is close to the natural corrosion potential. The cyclic polarization curves obtained (I(E) assist in the identification of corrosion potentials, pitting potentials, and repassivation potentials.

Passivation includes introducing a substantial amount of chromium (Cr) into the steel (> 10.5% by mass). Chromium interacts with the oxygen in the air and generates a coating of chromium oxide, Cr_2O_3: $4Cr + 3O_2 \rightarrow 2\ Cr_2O_3$. The compact, adherent, and protective layer is called the "passive layer" [16-18].

2.2.3 Microscopic rigidity experimentations

Rigidity of an iron specimen, exclaimed. The hardness of a stainless steel sample is a characteristic used to describe materials. In our situation, we measured Vickers hardness [19-20].

2.3 Operating modelling conditions

2.3.1 Tensile experiments

The samples, upon arrival, are cold-rolled sheets with a thickness of 3 mm. The steel plates were cut into standardized specimens at the University of Batna. Figure 1 displays the dimensions of the specimens utilized in our study: the gauge length measures 27.5 cm [21-24].

Figure 2a displays the stress-strain curve derived from 304L steel with a strain of 5 mm/min. In this scenario, the elongation corresponding to the maximum load is 48 mm.

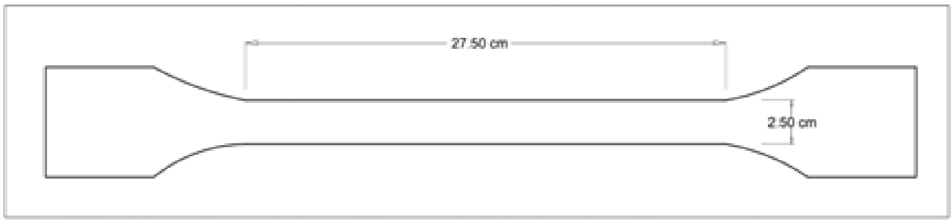


Figure 1. Dimensions of a tensile sampling

We have chosen the following intermediate elongations: 6 mm; 10 mm; 30 mm; and 45 mm, which correspond to strains of 2.18%, 3.63%, 10.90%, and 16.36%, respectively. These strains were computed using the following formula :

$$A\% = \frac{L_f - Li}{L_i}$$

A% : Strain

L_i: Specimen length before deformation

L_f: Specimen length after deformation

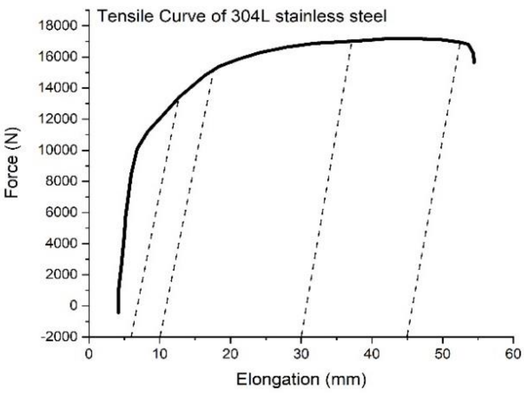


Figure 2a. Conventional tensile curve of 304L steel

2.3.2 Corrosion experiments

For electrochemical investigations, discs of 1 cm2 section were punched out from the usable portion of the specimen that underwent tensile testing. Surface preparation involved polishing with abrasive papers ranging from 600 to 1200 granulometry, followed by fine polishing with a 3 µm diamond paste until a mirror-like finish was reached. Finally, they were cleaned in an ultrasonic immersion containing acetone and dried with compressed air [25-26].

Figure 2b. Evolution of the corrosive state of the 304L stainless steel specimen

Corrosion experiments on these materials were done in a Tacussel CEC/TH three-electrode electrochemical cell, holding 400 ml of 3% NaCl solution, at laboratory ambient temperature, under aerated conditions and without agitation as indicated in Figure 2b. Potentials were measured relative to the saturated calomel electrode (SCE) [27-28].

3. Results and discussion

3.1 Hardness measurement

Mechanical hardness testing is central in the assessment of the mechanical characteristics of iron materials because the hardness values permit one to evaluate the material's resistance to deformation, accompanied by a rough estimation of its hardness parameters, flow stress, among numerous other essential features. Thus, the primary goal of this work has been to create and analyze a computerized approach based on image processing and analysis algorithms that may be employed for determining the Vickers hardness level using hardness strain imaging [29]. The material sample must have a laboratory surface unaffected by fluids, oxidation, flaws, and numerous contaminants to be studied with Vickers hardness. To supply the tester with total faith in the measured results, the experiment's surface must also be smooth and straight. In addition, it is crucial to ensure that the specimen under evaluation has no abnormalities on its contrasting face and that there are no shocks or damages to the testing equipment [30-31].

Figure 3 depicts the fluctuation in hardness as a function of strain rate. This graphic indicates that hardness varies practically linearly with the stress rate. The increase in hardness derives from the degree of work hardening, which increases with the severity of the tensile test [35].

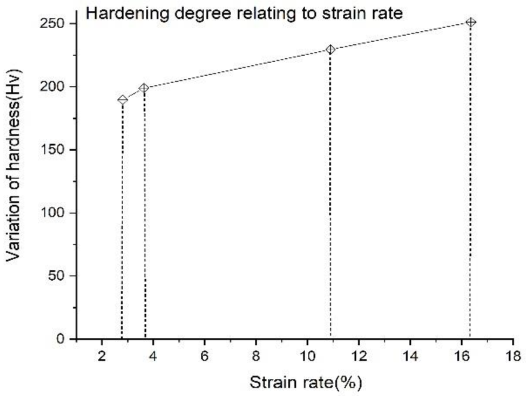


Figure 3. Micrograph sections of 304L stainless steel deformed by tensile : (a) 2.18%; (b) 3.63%; (c) 10.90%; (d) 16.36%.

3.2 Microstructure

These microstructures reveal the polygonal austenite grains with a nanoparticle size and the fully austenitic structure indicative of annealing twins. Microraphs represent the manifestation of strengthening pairs with disarticulated austenitic assembly.

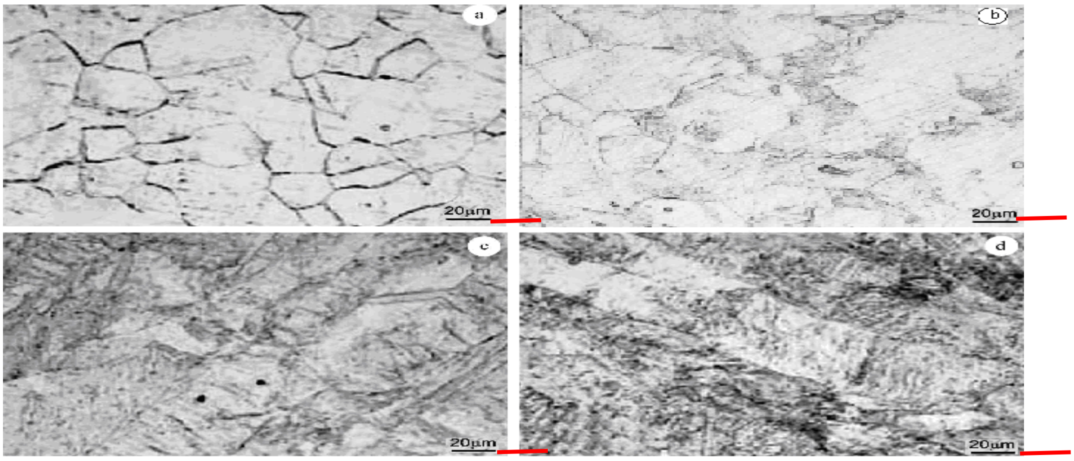


Figure 4. Micrograph sections of 304L stainless steel deformed by tensile (20 µm): (a) 2.18%, (b) 3.63%, (c) 10.90%, (d) 16.36%

The reverted austenite particles and recrystallized small grains, along with untransformed granules, are the essential properties of these microstructures. The quantity of strain-induced martensite has dropped, which is congruent with the reversal cycle already mentioned for X-ray fragmentation studies [32]. A tensile or compressive force that continues across the entirety of a material in the absence of an external load is known as residual stress. Remaining stresses are commonly broken into three categories according to the length scale. Macro-scale stresses, which last for lengthy spans of multiple grain diameters or more, are one category. Macro balance residual stresses indicate ongoing disparity. The change of stainless steel during forming, differential cooling throughout the wall thickness and along the surface during rolling, and localized plastic deformation during handling are common explanations. The other one varies depending on the grain size of the microstructure's morphology [34]. Cold plastic deformation modifies the texture of 304L steel by generating a change in the shape of the grains. During the tensile test, in the plasticity zone, the grains elongate in the direction of the tension. It is noted that the fragmentation into bands and dislocation cells alters the crystal orientation and progressively loses the unique character of the grains (Figure 4).

3.3 Polarization Experiments

From its conception to its most current improvements, the growth of electrochemical potentiokinetic techniques as they correspond to analyzing metals and alloys is followed. These procedures are applied to study the structure and characteristics of stainless steels and some mixed metals, and also for analyzing them for inter-particle corrosion and pitting corrosion in Figure 2b. It is vital to differentiate between polarization graphs associated with particle boundaries and the general structure of the alloy grains (the framework) while studying the propensity of inter-particle and pitting corrosion, applying potentiokinetic polarization tests. Features, potentials, and other properties revealed in the "dual process" analysis may be applied to determine an alloy's ability to undergo intergranular, pitting, and cracking corrosion by periodic polarization tests such as the electrochemical potentiokinetic reactivation (EPR) test. EPR technique for testing enables the capacity of the alloy to pits, intergranular corrosion, and cracking corrosion to be derived from specific characteristics disclosed in the "doubling cycle" test. EPR is rapid to adapt and considers the multiple influences on a material's qualities. The electrochemical potentiokinetic measures can be used for non-destructive testing to clarify the characteristics and responses of materials, as they can sufficiently identify morphological modifications in thermally treated materials that extend beyond stainless steel [35]. The influence of tension deformation on the pitting corrosion behavior of 304L steel was evaluated using the potentiokinetic technique, by imposing a potential sweep on the sample at a rate of 100 mV/min throughout [-0.45 ~ 0.9 V]. The factors to determine are the corrosion potential (E_c), pitting potential (E_p), and repassivation potential (E_rp). Figure 5 displays a conventional cyclic polarization curve demonstrating these distinct potentials. In our example, we obtain the curves given in Figure 6 for each of the examined strains.

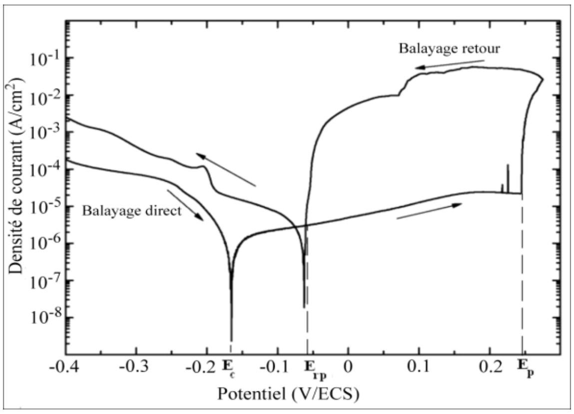


Figure 5. Standard cyclic polarization curve showing the corrosion potential Ec, pitting potential Ep, and repassivation potential Erp

Polarization shapes electrochemical measurements, which can give the needed non-destructiveness, accuracy, duplication, and speed of the process of execution. The polarization curve is a visual picture of the potential propensity link to the corrosion phenomena, which, as a policy, constitutes a steel electrode ensnared in a solution corrosion environment. The roughness of the most fantastic range of corrosion environments is assessed by the polarization trajectory, which also investigates the corrosion pattern of metals and alloys. The rationale for this is that "governed electrochemical polarization" may duplicate a wide range of corrosion processes and phenomena during a given gauging cycle by preserving command of the media's oxidizing strength without needing to move between diverse corrosion contexts. The first thing to know is the type of measurement to reliably distinguish the EPR test from other equivalent processes whose purpose and parameters may be different. This can help you understand an experiment of polarization and to accurately analyze the repercussions of many components, on curves exhibiting more sophisticated shapes.

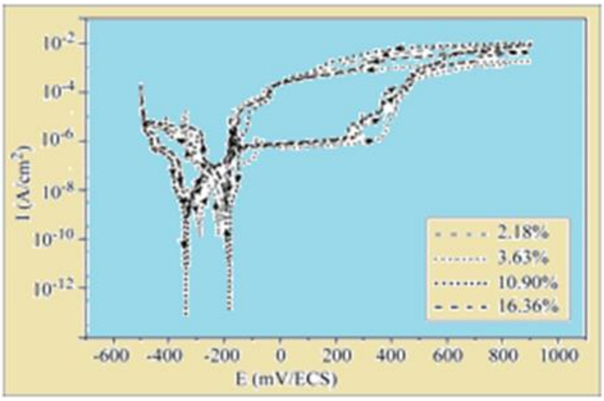


Figure 6. Cyclic polarization curves of 304L steel immersed in 3% NaCl, deformed at various rates by tensile.

The pitting corrosion of 304L stainless steel is tied to the stability of the passive film. This stability takes into account the potential between the metal and the solution. Figure 6 depicts the variability in current density with potential when 304L steel is immersed in 3 % NaCl. Pitting corrosion occurs beyond 0.34 V, in the transpassive domain. A stable pit generated above the pitting potential continues to expand at lower potentials until the repassivation potential; beyond that, repassivation happens. Repassivation potential is consequently the one below which an already created pit is protected from corrosion, and repassivation is achievable [36]. The modest difference (Ep – Erp) demonstrates valuable resistance to pit growth: the narrower, the more substantial, and the better the corrosion resistance. The breadth of the domain (Ep – Erp) represents the potential of pits to repassivate [37]. Figure 6 led to getting the values of potentials Ec, Ep, and Erp as a function of the varied rates, presented in Table 2. The determination of the domain width (Er – Erp) demonstrates that its extent grows with the increase in the strain. Thus, there is considerable resistance to pit development at low strains.

Table 2. Corrosion, pitting, and repassivation potentials of 304L steel immersed in 3% NaCl at several strains

E (mV) A(%)	**Ec**	**Ep**	**Erp**	**Er – Erp**	**i_{corr} ($\mu A/cm^2$)**	**Corrosion Rate (μm/year)**
2.81	-320.56	340.36	-130.44	470.80	0.4	4.6
3.63	-360.50	400.20	-140.12	540.32	0.7	8.1
10.90	-380.46	440.10	-150.00	590.10	1.2	13.8
16.36	-420.30	430.45	-200.96	631.41	2.0	23.0

We illustrated the development of these potentials in Figure 7. The corrosion potential tends to drop with increasing strain, making it less noble and more prone to corrosion. The same applies to the repassivation potential, whose decline equates to a decreasing resistance to pitting corrosion. On the other hand, the pitting potential, at the onset of the transpassive domain, tends to decline with the rise in the strain, resulting in a considerable reduction in the passivation domain, as indicated in Figure 6.

3.4 Infinitesimal examinations

The microscope investigations also demonstrated the occurrence of excess non-metallic inclusions on surface samples during corrosion testing. Following corrosion testing, metallographic examinations of the sample reveal that the degree of plastic deformation significantly affects the severity and type of corrosion alterations [38]. Research led to the conclusion that there would be an increase in deformation within the spectrum of a growth in the number, size, and depth of pits. According to acquired assessments, the crack surfaces operate as the anode of a local corrosion cell where metallic degradation takes place. The unstable steel sample underwent microscopic examination, revealing different levels of corrosion resistance in a 3% NaCl solution. A variety of hypotheses have been proposed to describe the stress corrosion phenomena. Of these, stress adsorption and electrochemical dissolution are the two fundamental ideas that have been put out.

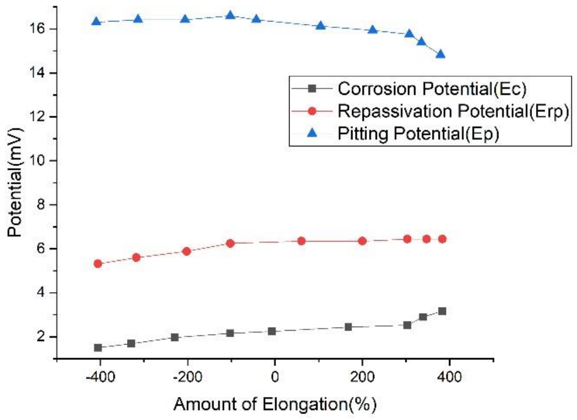


Figure 7. Variation of corrosion, pitting, and repassivation potentials as a function of the strain

The recorded anodized polarization charts indicate the existence of a passive interval. The repassivation potential varies between -130 and -200 mV.

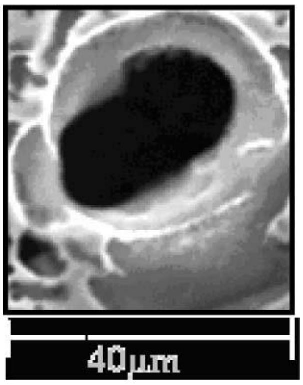


Figure 8. Pitting on the surface of 304L steel micrograph taken by visualizing electron microscopy

The corroded regions of the steel framework are examined microscopically, both before to and following predeformation at various degrees. Interparticular corrosion was identified as the principal source of the corrosion observed in the provided sampling.

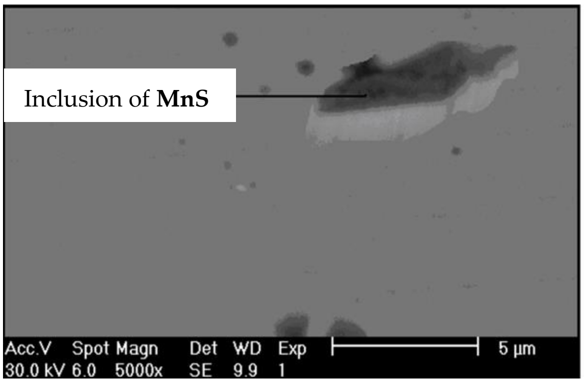


Figure 9. Micrographs obtained by scanning electron microscopy (SEM) showing an MnS inclusion

We conducted electron microscopy to illustrate the shape and scale of the pit dimensions (about 40 µm) (Fig. 8) and to reveal the existence of manganese-based inclusions on the material's surface (Fig. 9). These inclusions may serve as starting locations for pits within the passive coating of stainless steel [39, 40]. Table 1 demonstrates that the material exhibits a substantial manganese content, which leads to the production of manganese sulfide (MnS) inclusions in the presence of sulfur. The presence of MnS functions as a microelectrode when steel is exposed to an aggressive solution such as NaCl. In its vicinity, the passive layer is affected, and the creation of pits is quite possible.

4. Conclusion

This study assesses how surface corrosion deterioration affects the tensile properties of exposed specimens. The damaged experimental specimens were obtained from a model initially exposed to real-world corrosion and subsequently meticulously cleaned on a surface that preserved its original characteristics. In addition to the loss of the authentic piece, the two principal factors that significantly impact strength reduction are the variation in material characteristics due to corrosion and stress concentration resulting from localised rust pits. A comparable stress-strain curve illustrating the corrosion levels of degradation for both preserved and corroded steel plates. The experimental test produced findings indicating the critical fluid media environment of the specimen with an increasing degree of deterioration due to pitting corrosion. This research paper aimed to assess the impact of tensile deformation at varying velocities on the behaviour of 304L stainless steel exposed to a sodium chloride aqueous environment simulating seawater (3% NaCl). The gathered results enable us to draw the following conclusions: the individual character of the grains gradually vanished as the strain reached 10.9%, while the hardness of the steel rose with increasing strain; additionally, the presence of

MnS inclusions created suitable locations for pitting initiation, and the passivation domain diminished with high strains; furthermore, repassivation of pits was more probable for low distortions and generated stronger resistance to local corrosion, whereas the pitting potential tended to diminish with rising strain, resulting in a significant drop in the passivation domain.

5. Acknowledgements

Author Contributions: The authors would like to thank everyone who contributed, directly or indirectly, to the completion of this article from Mohamed Khider University of Biskra and Mostefa Benboulaid University of Batna.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] Kaladhar, M.; Subbaiah, K. V.; Rao, C. S. Machining of austenitic stainless steels –a review. *International Journal of Machining and Machinability of Materials* **2012**, *12*(1–2), 178–192.

[2] Peguet, L.; Malki, B.; Baroux, B. Influence of Cold Working on the Pitting Corrosion Resistance of Stainless Steels. *ECS Meeting Abstracts* **2006**, *MA2006-02*(17), 859–859. https://doi.org/10.1149/ma2006-02/17/859

[3] Dhaiveegan, P.; et al. Corrosion behavior of 316L and 304 stainless steels exposed to industrial-marine-urban environment: field study. *Rsc Advances* **2016**, *6*(53), 47314–47324. https://doi.org/10.1039/C6RA04015B

[4] Luo, H.; et al. Effect of cold deformation on the electrochemical behaviour of 304L stainless steel in contaminated sulfuric acid environment. *Applied Surface Science* **2017**, *425*, 628–638. https://doi.org/10.1016/j.apsusc.2017.07.057

[5] Bansod, A. V.; et al. Microstructure, mechanical and electrochemical evaluation of dissimilar low Ni SS and 304 SS using different filler materials. *Materials Research* **2018**, *22*(1), e20170203. http://dx.doi.org/10.1590/1980-5373-MR-2017-0203

[6] Kundu, R. Development of chromium and chromium-nickel electro coated mild steel to enhance corrosion resistance comparable to 304 stainless steel in 3.5% chloride water. **2019**. https://doi.org/10.1016/j.heliyon.2023.e22538

[7] Guo, S.; et al. Corrosion characteristics of typical Ni–Cr alloys and Ni–Cr–Mo alloys in supercritical water: a review. *Industrial & Engineering Chemistry Research* **2020**, *59*(42), 18727–18739. https://dx.doi.org/10.1021/acs.iecr.0c04292

[8] Yu, W.-W.; LaBoube, R. A.; Chen, H. *Cold-formed steel design*; John Wiley & Sons, **2019**.

[9] Scatigno, G. G.; Dong, P.; Ryan, M. P.; Wenman, M. R. The Effect of Salt Loading on Chloride-Induced Stress Corrosion Cracking of 304L Austenitic Stainless Steel under Atmospheric Conditions. *SSRN Electronic Journal* **2019**. https://doi.org/10.2139/ssrn.3441481

[10] Ibrahim, M. Z. A. A. *Developing a new laser cladded FeCrMoCB metallic glass layer on nickel-free stainless-steel as a potential superior wear-resistant coating for joint replacement implants*. Ph.D. Thesis, University of Malaya (Malaysia), **2019**. https://doi.org/10.1016/j.surfcoat.2020.125755

[11] Bellanthudawa, B. K. A.; et al. A perspective on biodegradable and non-biodegradable nanoparticles in industrial sectors: applications, challenges, and future prospects. *Nanotechnology for Environmental Engineering* **2023**, *8*(4), 975–1013. https://doi.org/10.1007/s41204-023-00344-7

[12] Parangusan, H.; Bhadra, J.; Al-Thani, N. A review of passivity breakdown on metal surfaces: influence of chloride-and sulfide-ion concentrations, temperature, and pH. *Emergent Materials* **2021**, *4*(5), 1187–1203. https://doi.org/10.1007/s42247-021-00194-6

[13] Keegan, G. M.; Learmonth, I. D.; Case, C. A systematic comparison of the actual, potential, and theoretical health effects of cobalt and chromium exposures from industry and surgical implants. *Critical reviews in toxicology* **2008**, *38*(8), 645–674. https://doi.org/10.1080/10408440701845534

[14] Linde, G. F. *Investigating the performance of thermal spray coatings on agriculture equipment*. Ph.D. Thesis, Stellenbosch University, Stellenbosch, **2016**.

[15] Moayyedian, M.; et al. Tensile Test Optimization Using the Design of Experiment and Soft Computing. *Processes* **2023**, *11*(11), 3106. https://doi.org/10.3390/pr11113106

[16] Esmailzadeh, S.; Aliofkhazraei, M.; Sarlak, H. Interpretation of cyclic potentiodynamic polarization test results for study of corrosion behavior of metals: a review. *Protection of metals and physical chemistry of surfaces* **2018**, *54*(5), 976–989. https://doi.org/10.1134/S207020511805026X

[17] Bautista, A.; Blanco, G.; Velasco, F.; Gutiérrez, A.; Soriano, L.; Palomares, F. J.; Takenouti, H. Changes in the Passive Layer of Corrugated Austenitic Stainless Steel of Low Nickel Content due to Exposure to Simulated Pore Solutions. *Corrosion Science* **2009**, *51*(4), 785–792. https://doi.org/10.1016/j.corsci.2009.01.012

[18] He, N.; Li, H.; Ji, L.; Liu, X.; Chen, J. Investigation of Metal Elements Diffusion in Cr2O3 Film and Its Effects on Mechanical Properties. *Ceramics International* **2019**, *46*(5), 6811–6819. https://doi.org/10.1016/j.ceramint.2019.11.173

[19] Chicot, D.; et al. Interpretation of instrumented hardness measurements on stainless steel with different surface preparations. *Surface Engineering* **2007**, *23*(1), 32–39. https://doi.org/10.1179/174329407X161573

[20] Cai, Y.; Zheng, H.; Hu, X.; Lu, J.; Poon, C. S.; Li, W. Comparative Studies on Passivation and Corrosion Behaviors of Two Types of Steel Bars in Simulated Concrete Pore Solution. *Construction and Building Materials* **2021**, *266*, 120971. https://doi.org/10.1016/j.conbuildmat.2020.120971

[21] Dobmann, G.; Kern, R.; Wolter, B. Mechanical property determination of heavy steel plates and cold rolled steel sheets by micro-magnetic NDT. In *16th World Conference on Nondestructive Testing (WCNDT)*, **2004**; Fraunhofer-IZFP, Saarbrücken, Germany.

[22] Yu, W.-W.; LaBoube, R. A.; Chen, H. *Cold-formed steel design*; John Wiley & Sons, **2019**. https://doi.org/10.1002/9781119487425

[23] Noronha, D. J.; et al. Deep rolling techniques: A comprehensive review of process parameters and impacts on the material properties of commercial steels. *Metals* **2024**, *14*(6), 667. https://doi.org/10.3390/met14060667

[24] Dehghani, F.; Salimi, M. Analytical and experimental analysis of the formability of copper-stainless-steel 304L clad metal sheets in deep drawing. *The International Journal of Advanced Manufacturing Technology* **2016**, *82* (1), 163–177. https://doi.org/10.1007/s00170-015-7359-9

[25] Jeon, J. H.; Ahn, S.-H.; Melkote, S. N. In Situ analysis of the effect of ultrasonic cavitation on electrochemical polishing of additively manufactured metal surfaces. *Journal of Manufacturing Science and Engineering* **2024**, *146*(4), 041003. https://doi.org/10.1115/1.4064692

[26] Othman, N. H.; et al. The effect of residual solvent in carbon− based filler reinforced polymer coating on the curing properties, mechanical and corrosive behaviour. *Materials* **2022**, *15*(10), 3445. https://doi.org/10.3390/ma15103445

[27] Ogazi, A. C. *Comparative Studies of Electrochemical Corrosion Behaviour of Mild Steel in Some Agro-Fluids*. Ph.D. Thesis, University of South Africa (South Africa), **2015**.

[28] Last, B. A. *Research Concerning the Reference Electrode of the Three Electrode Device for Measuring Corrosion Rates*. **2013**.

[29] Lin, Y.; et al. Mechanical properties and optimal grain size distribution profile of gradient grained nickel. *Acta Materialia* **2018**, *153*, 279–289. https://doi.org/10.1016/j.actamat.2018.04.065

[30] Guo, S.; et al. Corrosion characteristics of typical Ni–Cr alloys and Ni–Cr–Mo alloys in supercritical water: a review. *Industrial & Engineering Chemistry Research* **2020**, *59*(42), 18727–18739. https://dx.doi.org/10.1021/acs.iecr.0c04292

[31] Lingelbach, M. E. Y. *Application of Data Mining and Machine Learning Methods to Industrial Heat Treatment Processes for Hardness Prediction*. Dissertation, **2021**. https://doi.org/10.5445/KSP/1000169018

[32] Vafaeian, S.; et al. On the study of tensile and strain hardening behavior of a thermomechanically treated ferritic stainless steel. *Materials Science and Engineering: A* **2016**, *669*, 480–489. https://doi.org/10.1016/j.msea.2016.04.050

[33] Cios, G.; et al. The investigation of strain-induced martensite reverse transformation in AISI 304 austenitic stainless steel. *Metallurgical and Materials Transactions A* **2017**, *48*(10), 4999–5008. https://doi.org/10.1007/s11661-017-4228-1

[34] Lin, Y.; et al. Mechanical properties and optimal grain size distribution profile of gradient grained nickel. *Acta Materialia* **2018**, *153*, 279–289. https://doi.org/10.1016/j.actamat.2018.04.065

[35] Radojković, B. M.; et al. Non-destructive evaluation of the AISI 304 stainless steel susceptibility to intergranular corrosion by electrical conductivity measurements. *Metals and Materials International* **2024**, *30*(3), 682–696. https://doi.org/10.1007/s12540-023-01536-1

[36] Hadri, F. *l'Université de Lorraine*, France, **2012**.

[37] Stalker, K. *Illustrating pit initiation and evolution in aluminum alloys according to a 3-dimensional cellular automata based model*. Honors Research Project, The University of Akron, **2016**.

[38] Punckt, C.; Bölscher, M.; Rotermund, H. H.; Mikhailov, A. S.; Organ, L.; Budiansky, N.; Scully, J. R.; Hudson, J. L. Sudden Onset of Pitting Corrosion on Stainless Steel as a Critical Phenomenon. *Science* **2004**, *305* (5687), 1133–1136. https://doi.org/10.1126/science.1101358

[39] Park, J. H.; Kang, Y. Inclusions in stainless steels− a review. *steel research international* **2017**, *88*(12), 1700130. https://doi.org/10.1002/srin.201700130

[40] Soltis, J. Passivity breakdown, pit initiation and propagation of pits in metallic materials–review. *Corrosion Science* **2015**, *90*, 5–22. http://dx.doi.org/10.1016/j.corsci.2014.10.006.

ASEAN
Journal of Scientific and Technological Reports
Online ISSN:2773-8752

Research article

Anti-Inflammatory and Antioxidant Activities of Crude Protein Extracts from *Etlingera pavieana* Rhizomes Grown at Different Cultivation Sites in Thailand

Jongkonnee Padungkasem[1, 2], Sittiruk Roytrakul[3], Janthima Jaresitthikunchai[4], Sawanya Charoenlappani[5], and Klaokwan Srisook[6, 2*]

[1] Faculty of Science, Burapha University, Chonburi, 20131, Thailand
[2] Center of Excellence for Innovation in Chemistry, Faculty of Science, Burapha University, Chonburi 20131, Thailand
[3] Functional Proteomics Technology Laboratory, National Center for Genetic Engineering and Biotechnology, National Science and Technology Development Agency, Pathum Thani, 12120, Thailand
[4] Functional Proteomics Technology Laboratory, National Center for Genetic Engineering and Biotechnology, National Science and Technology Development Agency, Pathum Thani, 12120, Thailand
[5] Functional Proteomics Technology Laboratory, National Center for Genetic Engineering and Biotechnology, National Science and Technology Development Agency, Pathum Thani, 12120, Thailand
[6] Faculty of Science, Burapha University, Chonburi, 20131, Thailand
* Correspondence: klaokwan@buu.ac.th

Citation:
Padungkasem, J.; Roytrakul, S.; Jaresitthikunchai, J.; Charoenlappani, S.; Srisook, K. Anti-inflammatoryu and antioxidant activities of crude protein extracts from *Etlingera pavieana* rhizomes grown at different cultivation sites in Thailand. *ASEAN J. Sci. Tech. Report.* **2025**, *28*(5), e259269. https://doi.org/10.55164/ajstr.v28i5.259269.

Article history:
Received: May 10, 2025
Revised: August 20, 2025
Accepted: September 1, 2025
Available online: September 21, 2025

Publisher's Note:
This article is published and distributed under the terms of the Thaksin University.

Abstract: This study investigated the anti-inflammatory and antioxidant activities of crude protein extracts derived from *Etlingera pavieana* rhizomes cultivated in three eastern provinces in Thailand: Rayong, Chanthaburi, and Trat. Crude proteins were extracted using Tris-HCl buffer, and protein profiles were evaluated by SDS-PAGE. Anti-inflammatory activity was assessed via nitric oxide (NO) inhibition in lipopolysaccharide-stimulated RAW 264.7 macrophages, while antioxidant activity was assessed using DPPH radical scavenging, metal-chelating, and reducing power assays. All the crude protein extracts exhibited dose-dependent biological activities, with no significant cytotoxicity observed. The extract from Trat rhizomes showed the highest NO inhibition, DPPH scavenging activity, and reducing power, while the Rayong sample exhibited the strongest metal chelation. The findings demonstrate that the cultivation site affects the bioactivity of crude protein extracts, underscoring the potential application of *E. pavieana* rhizome proteins as functional ingredients in food or nutraceutical products.

Keywords: *Etlingera pavieana*; crude protein extract; antioxidant activity; anti-inflammatory activity; cultivation site

1. Introduction

Plant-based protein extraction has attracted significant scientific interest in recent years due to the growing global demand for alternative protein sources in the food industry. This demand is driven not only by increasing consumer preference for sustainable and health-promoting ingredients, but also by environmental and health concerns associated with animal protein production. Plant proteins offer distinct advantages over animal proteins, including higher dietary fiber content and the presence of bioactive compounds with potential health benefits [1]. Furthermore, plant proteins are considered more sustainable and lower in fat, making them ideal for use in functional foods and nutraceuticals [2]. Plants of the Zingiberaceae family are traditionally valued for their culinary and medicinal uses. However, accumulating evidence indicates that crude protein extracts from the rhizomes of various Zingiberaceae species also exhibit biological activities, highlighting them as potential sources of bioactive compounds [3, 4, 5]. *Etlingera pavieana* (Pierre ex Gagnep.) R.M. Sm., a member of the Zingiberaceae family, is indigenous to eastern Thailand. Its rhizomes are commonly used in the regional cuisine and herbal remedies and are currently cultivated as intercrops in fruit orchards [6, 7]. Previous studies have found that *E. pavieana* rhizomes exhibit several biological properties, including antioxidant, anti-inflammatory, antimicrobial, and anticancer effects [8, 9, 10]. These findings signify the significant potential of *E. pavieana* rhizomes as functional ingredients in health-promoting food and supplements. However, the bioactivity of protein extracts derived from *E. pavieana* rhizomes remains largely unexplored.

Cultivation location can influence the phytochemical composition and biological activities of plants. Hao et al. [11] found that mulberry cultivated in Guangxi, China, exhibited higher antioxidant activity than those grown in Guangdong and Chongqing. Similarly, Khumaida et al. [12] conducted a comparative study on *Curcuma zedoaria* rhizomes from various cities in Indonesia and found that specimens grown in Pakem, Yogyakarta, had the highest phenolic content. Yahyaoui et al. [13] also reported significant regional variations in antioxidant activity and phenolic content in hawthorn fruits in Tunisia. These findings demonstrate the influence of geographic origin on the functional properties of plant-derived extracts. The present study compared the antioxidant and anti-inflammatory activities of crude protein extracts obtained from *E. pavieana* rhizomes cultivated in three provinces in eastern Thailand: Rayong, Chanthaburi, and Trat. The anti-inflammatory activity of the extracts was assessed by measuring the inhibition of nitric oxide (NO) in lipopolysaccharide (LPS)-stimulated macrophages, which are key effector cells in the innate immune response. Antioxidant activity was evaluated using 2,2-diphenyl-1-picrylhydrazyl (DPPH) radical scavenging, metal-chelating, and reducing power assays. The findings of this study are expected to advance the scientific understanding of *E. pavieana* as a source of bioactive proteins and confirm its economic and health value in the development of functional foods and nutraceuticals.

2. Materials and Methods

2.1 Materials

Acrylamide, aminoguanidine, Coomassie Brilliant Blue R-250, DPPH, Dulbecco's Modified Eagle Medium (DMEM), EDTA, ferrozine, fetal bovine serum (FBS), LPS from *Escherichia coli* O111:B4, N-(1-naphthyl) ethylenediamine dihydrochloride, potassium ferricyanide, sulfanilamide, sodium carbonate, and thiazolyl blue tetrazolium bromide (MTT) were purchased from Sigma-Aldrich (St. Louis, MO, USA). All other chemicals used were of analytical grade.

2.2 Methods

2.2.1 Protein Extraction

Fresh *E. pavieana* rhizomes were harvested in October 2023 at the stage of commercial maturity, which local farmers typically define as approximately one year after planting. Samples were collected from agricultural plantations in Chanthaburi, Rayong, and Trat. The rhizomes were thoroughly washed with tap water, sliced into smaller pieces, and then dried completely at 55 °C. The dried samples were ground using a mechanical grinder and stored at 4 °C until used for protein extraction. Protein extraction was performed as described by Deesrisak et al. [14], with slight modifications. Briefly, crude protein was extracted by mixing 30 g of rhizome with 450 mL of 10 mM Tris-HCl buffer (pH 7.0) and shaking the mixture at room temperature for 24 h. The resulting extract was filtered, and the protein amount was examined using the Bradford assay.

2.2.2 Sodium Dodecyl Sulfate–Polyacrylamide Gel Electrophoresis (SDS-PAGE) Analysis

SDS-PAGE consisted of a 12.5% separating gel and a 4% stacking gel. Each sample, consisting of 20 μg protein mixed with 6× bromophenol blue loading dye, was denatured by heating at 95°C for 5 minutes and then loaded into a gel. A pre-stained molecular weight protein marker was applied to the reference lane. SDS-PAGE was performed at a constant voltage of 120 V. Subsequently, the gel was stained with Coomassie Brilliant Blue R-

250, followed by destaining until clear protein bands became visible. The molecular weights of the protein bands were estimated by comparison with the standard protein marker.

2.2.3 Cell Viability Assay

Macrophage RAW 264.7 cells were obtained from the ATCC (Rockville, MD, USA). Cell viability was assessed using the MTT assay, as described by Srisook et al. [15]. RAW 264.7 cells (1.5×10^5 cells/well) were seeded into 24-well plates. The cells were then treated with protein extracts and LPS (1 μg/mL) for 24 hours. After treatment, the medium was removed and replaced with DMEM containing 0.1 mg/mL MTT, followed by incubation for 2 hours. Formazan crystals were solubilized in DMSO, and the absorbance was measured at 550 nm using a microplate reader (Versamax, USA). Cell viability was calculated as follows:

$\% \text{ Cell viability} = (A_{\text{treated}}/A_{\text{control}}) \times 100$

2.2.4 Nitric Oxide Inhibition Assay

No inhibition by the protein extracts was assessed using the Griess reaction, as described by Srisook et al. [15]. RAW 264.7 macrophages were seeded as described above and treated with protein extracts and LPS (1 μg/mL) for 24 h. Nitrite concentrations were determined using a sodium nitrite standard curve (3.12–50 μM), with aminoguanidine serving as the positive control. NO inhibition was calculated as follows:

$\% \text{ NO inhibition} = 100 \times [(\text{Nitrite concentration of treated well/LPS control}) \times 100]$

2.2.5 DPPH Radical Scavenging Assay

DPPH radical scavenging activity was evaluated as described by Pechroj et al. [16], with slight modifications. Protein extracts at various concentrations were dissolved in 10 mM Tris-HCl buffer (pH 7.0) and mixed with a 0.2 mM DPPH solution prepared in methanol in a microplate. For the assay, 50 μL of each extract was combined with 100 μL of the DPPH solution and incubated at room temperature in the dark for 30 min. Absorbance was then measured at 517 nm. Ascorbic acid was used as the positive control. The percentage of DPPH scavenging activity was calculated as follows:

$\% \text{ DPPH scavenging} = [(A_{\text{control}}-A_{\text{sample}})/A_{\text{control}}] \times 100$

2.2.6 Reducing Power Assay

The reducing power of the extracts was determined following the method described by Uthairat et al. [17]. Briefly, protein extract (50 μg/mL) was mixed with 1% potassium ferricyanide. The mixture was incubated for 30 minutes, then 10% TCA was added, and it was centrifuged at 3000 × g for 10 minutes. The supernatant was combined with distilled water and 0.1% ferric chloride. The absorbance was measured at 700 nm. Reducing power was expressed as gallic acid equivalents (GAE) per gram of extract, based on a gallic acid standard curve.

2.2.7 Metal-Chelating Activity

Metal-chelating activity was evaluated as described by Uthairat et al. [17]. The reaction mixture consisted of 10 μL of 2 mM ferrous sulfate, 200 μL of protein extract or EDTA (positive control), and 20 μL of 5 mM ferrozine. The absorbance was measured at 562 nm using a microplate reader. Chelating activity was calculated as follows:

$\% \text{ Chelating activity} = [(A_{\text{control}}-A_{\text{sample}})/A_{\text{control}}] \times 100$

2.2.8 Statistical Analysis

All data are expressed as the mean ± standard deviation (n=3). Statistical analysis was performed using Minitab 18. One-way ANOVA was performed, followed by Tukey's post hoc test to determine significant differences between groups at $p < 0.05$.

3. Results and Discussion

3.1 SDS-PAGE Analysis of Crude Protein Extracts from E. pavieana Rhizomes

The protein profiles of the crude extracts derived from *E. pavieana* rhizomes collected from Rayong, Chanthaburi, and Trat provinces were evaluated by SDS-PAGE. Protein bands were observed across a molecular weight range of approximately 16–175 kDa in all samples (Figure 1). Notably, for each extract, prominent bands appeared at approximately 25, 35, 40, and 51 kDa, indicating the presence of major protein constituents that may have been conserved among rhizomes from different cultivation regions.

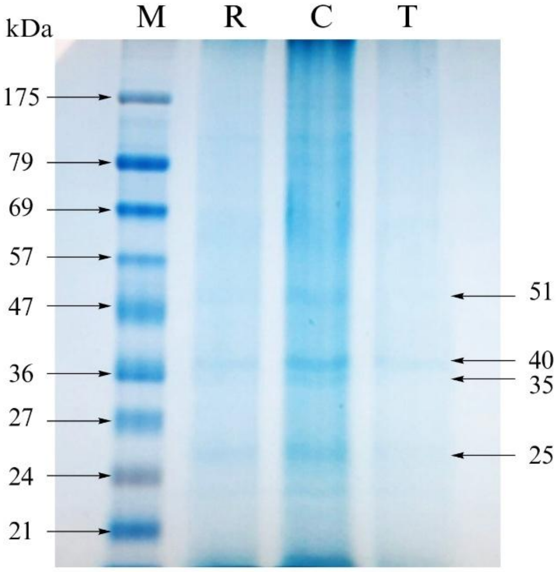


Figure 1. SDS-PAGE profiles of crude protein extracts from Etlingera pavieana rhizomes collected from the three provinces. Lane M: molecular weight marker; Lane R: Rayong; Lane C: Chanthaburi; Lane T: Trat

3.2 Nitric Oxide (NO) Inhibitory Effect of Crude Protein Extracts from E. pavieana Rhizomes

The cytotoxicity of crude protein extracts derived from *E. pavieana* rhizomes collected from the Chanthaburi, Rayong, and Trat provinces was evaluated in RAW 264.7 macrophages. At concentrations of 0.04–0.16 µg/mL, the extracts showed no statistically significant cytotoxicity compared to the untreated controls. Although slight reductions in cell viability were observed at concentrations of 0.31–1.25 µg/mL, all viability values remained above 85% (Figure 2). Therefore, concentrations ranging from 0.04 to 1.25 µg/mL were used for further NO inhibition assays. NO is an inflammatory mediator synthesized from the amino acid L-arginine by nitric oxide synthase (NOS). This reaction produces NO and L-citrulline, requiring cofactors such as NADPH, FAD, FMN, heme, calmodulin, and tetrahydrobiopterin [18]. In inflammation triggered by infection or foreign stimuli, immune cells, such as macrophages, are activated and induce the expression of inducible NOS (iNOS), which triggers excessive NO production. NO reacts with superoxide anions ($O_2^{\bullet-}$) to form peroxynitrite ($ONOO^-$), which kills microbes but also causes tissue damage, thereby intensifying the inflammatory response. The prolonged overproduction of NO is associated with chronic inflammation and the pathogenesis of various diseases, such as rheumatoid arthritis, inflammatory bowel disease, diabetes, and neurodegenerative conditions, such as Alzheimer's disease [19, 20]. Thus, the absence of NO inhibition serves as an indicator of anti-inflammatory potential.

The anti-inflammatory activity of the crude protein extracts from *E. pavieana* rhizomes was assessed by measuring the accumulation of nitrite, a stable end-product of NO oxidation, in the culture supernatant. Untreated control cells and cells stimulated with LPS alone yielded nitrite levels of 0.03 ± 0.02 and 12.85 ± 0.15 µM, respectively. Treatment with crude protein extracts (0.04–1.25 µg/mL) significantly reduced NO production in LPS-stimulated RAW 264.7 macrophages in a concentration-dependent manner (Figure 3). Therefore, the observed reduction in NO production was not caused by the cytotoxic effects of the extracts, but rather by their NO-inhibitory properties. The half maximal inhibitory concentration (IC_{50}) values for NO inhibition indicated that the extract from Trat samples exhibited the most potent inhibitory effect, followed by those from Rayong and Chanthaburi (Table 1). Aminoguanidine (50 µM), which was used as a positive control, inhibited NO production by 61% compared with the LPS-only group. To the best of our knowledge, this is the first report demonstrating the anti-inflammatory potential of crude protein extracts from *E. pavieana* rhizomes. Chantaranothai et al. [3] investigated the NO inhibitory activity of crude protein extracts from 15 species of

Zingiberaceae rhizomes and identified only four species that showed inhibitory activity: *Curcuma aeruginosa* Roxb., *Curcuma aromatica*, *Hedychium coronarium*, and *Zingiber ottensii* Valeton. The present study demonstrated that the crude protein extracts from *E. pavieana* also possess strong NO inhibitory activity. Notably, the IC_{50} values observed here were approximately 38- to 100-fold lower than those previously reported for the species. However, differences in the extraction methods between the two studies may have contributed to the discrepancy.

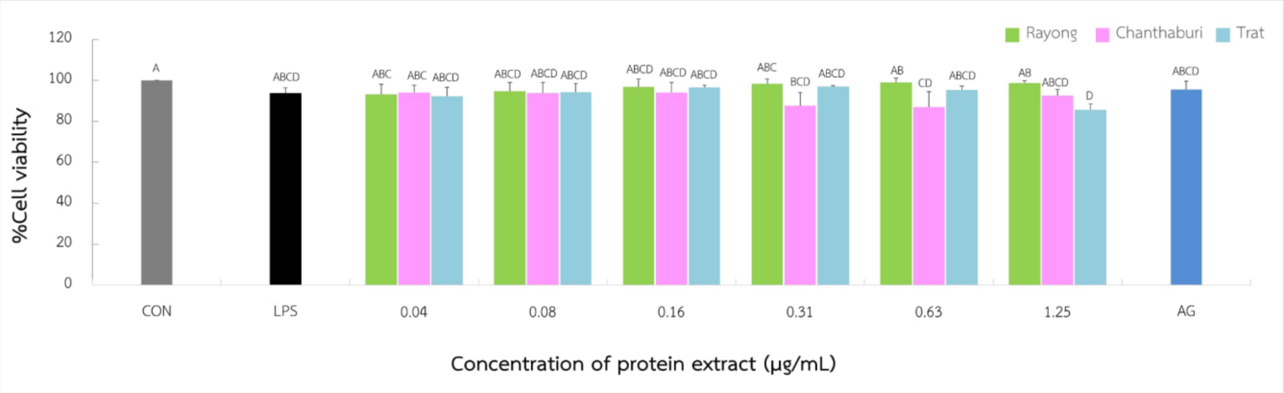


Figure 2. Effect of crude protein extracts from *Etlingera pavieana* rhizomes on the viability. RAW 264.7 macrophages were treated for 24 h with the indicated extract concentrations in the presence of LPS (1 μg/mL). Cell viability was assessed using the MTT assay. CON, control; LPS, lipopolysaccharide; AG, aminoguanidine. Different superscript letters (A–F) indicate statistically significant differences between groups at $p < 0.05$.

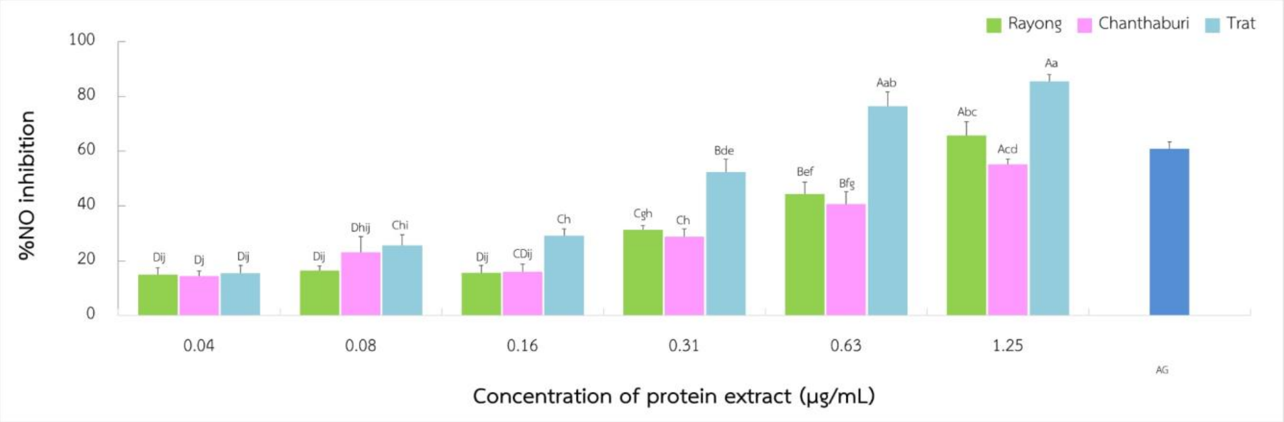


Figure 3. Inhibitory effects of crude protein extracts from *Etlingera pavieana* rhizomes on nitric oxide (NO) production in lipopolysaccharide-stimulated RAW 264.7 macrophages. Cells were stimulated with 1 μg/mL LPS and co-treated with various concentrations of the extract. The nitrite content of the culture supernatant was measured using the Griess reagent. Different uppercase letters (A–E) indicate statistically significant differences among concentrations within the same province, whereas different lowercase letters (a–g) indicate significant differences between provinces at the same concentration. Statistical significance was set at $p < 0.05$. AG, aminoguanidine.

Table 1. IC_{50} values for nitric oxide inhibition in lipopolysaccharide-stimulated RAW 264.7 macrophages treated with crude protein extracts from *Etlingera pavieana* rhizomes collected from three different cultivation locations

Province	IC_{50} value (µg/mL)
Rayong	0.84 ± 0.08^b
Chanthaburi	1.03 ± 0.07^c
Trat	0.46 ± 3.41^a

Different superscript letters (a, b, c) indicate statistically significant differences between groups at $p < 0.05$. IC_{50}, half maximal inhibitory concentration.

3.3 Antioxidant Activity of Crude Protein Extracts from E. pavieana Rhizomes

The antioxidant activity of the crude protein extracts from *E. pavieana* rhizomes was determined using the DPPH radical scavenging assay. All extracts exhibited concentration-dependent scavenging activity (Figure 4), with half-maximal effective concentration (EC_{50}) values ranging from 7.31 ± 0.11 to 11.52 ± 0.16 µg/mL. Based on EC_{50} values, extracts from Trat and Rayong exhibited significantly greater scavenging activity than those from Chanthaburi. In this study, the DPPH radical scavenging activities of the extracts from Rayong and Trat were notably higher than those from Chanthaburi (Table 2). The antioxidant capacity of the crude protein extracts was greater than that of ascorbic acid, a widely known reference antioxidant. These findings are consistent with previous studies on the DPPH radical scavenging activities of crude protein extracts from other Zingiberaceae rhizomes, which report variations in antioxidant potency across plant species [3, 4]. However, among the 15 species evaluated by Chantaranothai et al. [3], three species, K*aempferia galanga* Linn., *Alpinia galanga* (Linn.) Swartz and Boesenbergia pandurata Roxb. did not exhibit detectable DPPH radical scavenging activity. This lack of activity may be attributed to either the absence of active protein components or their low activity. The DPPH radical scavenging activity of plant-derived proteins or peptides is primarily attributed to specific amino acid functional groups, such as the hydroxyl groups (−OH) in tyrosine and serine; thiol groups (-SH) in cysteine; aromatic rings in tyrosine, tryptophan, and phenylalanine; and amino groups in lysine and arginine. These functional groups can donate hydrogen atoms or electrons to DPPH radicals, reducing them from their purple-colored radical form to the yellow-colored reduced form (DPPH-H) [21]. However, since the DPPH solution was prepared in methanol, we cannot exclude the possibility that methanol-sensitive proteins partially precipitated, or that co-extracted phytochemicals contributed to the observed antioxidant activity. Therefore, the results should be interpreted as reflecting the overall radical scavenging capacity of the extracts rather than protein-specific activity.

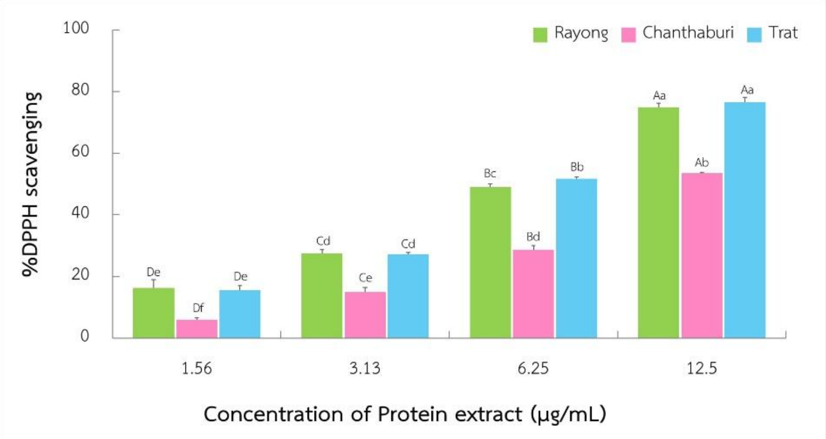


Figure 4. DPPH radical scavenging activity (%) of different concentrations of crude protein extracts from *Etlingera pavieana* rhizomes collected from Rayong, Chanthaburi, and Trat provinces. Different uppercase letters (A–E) indicate significant differences among concentrations within the same province, while different lowercase letters (a–g) indicate significant differences between provinces at the same concentration ($p < 0.05$).

Table 2. DPPH radical scavenging activity (EC_{50}) of the crude protein extracts from *Etlingera pavieana* rhizomes collected from the three provinces and the positive control (ascorbic acid)

Province	EC_{50} for DPPH scavenging (µg/mL)
Rayong	7.53 ± 0.19^a
Chanthaburi	11.52 ± 0.16^b
Trat	7.31 ± 0.11^a
Ascorbic acid	15.22 ± 0.77

Different superscript letters (a, b) indicate statistically significant differences between the groups at $p < 0.05$.

Crude protein extracts from *E. pavieana* rhizomes collected in Rayong, Chanthaburi, and Trat provinces showed significant variation in reducing power ($p < 0.05$). The extract from Trat rhizomes demonstrated the highest reducing activity, followed by those from Rayong and Chanthaburi (Table 3). The redox activity involving the electron donation of proteins and peptides is primarily attributed to the presence of specific functional groups capable of donating electrons that facilitate the reduction of ferric ions (Fe^{3+}) to ferrous ions (Fe^{2+}). Functional groups, such as hydroxyl, thiol, and peptide bonds, are crucial in that they underlie the strong antioxidant potential of protein-based compounds [21, 22]. Furthermore, the crude protein extracts may also contain trace amounts of naturally occurring phenolic compounds in *E. pavieana* rhizomes, which could further enhance their DPPH radical scavenging activity and reducing power by donating electrons or neutralizing free radicals, thereby mitigating oxidative damage in cells.

Table 3. Reducing power of crude protein extracts from *Etlingera pavieana* rhizomes collected from the three provinces.

Province	Reducing power (mg GAE/g extract)
Rayong	205.05 ± 4.28^b
Chanthaburi	153.24 ± 1.59^c
Trat	363.70 ± 3.41^a

Different superscript letters (a, b, c) indicate statistically significant differences between the groups at $p < 0.05$.

The metal-chelating activity of proteins is a key antioxidant mechanism, particularly in inhibiting iron-catalyzed oxidative reactions, such as the Fenton and Haber–Weiss reactions, which generate highly reactive hydroxyl radicals. In this study, various concentrations (0.39–6.25 µg/mL) of crude protein extracts from E. pavieana rhizomes were evaluated for their ability to chelate ferrous ions, showing significant concentration-dependent increases in metal chelation activity ($p < 0.05$). The Rayong extract exhibited the highest chelating effect (6.25 µg/mL), although no significant differences were observed between the Trat and Chanthaburi extracts (Figure 5). The chelating ability of these protein extracts stems from specific amino acid residues that can coordinate metal ions, particularly those with functional groups, such as the imidazole group in histidine, the thiol group in cysteine, and carboxyl groups in aspartic and glutamic acids. These groups serve as electron-donating sites, facilitating stable interactions with metal ions and thereby preventing the catalytic formation of reactive oxygen species [23-24]. Although minor variations were observed among the three extracts, the metal-chelating capacities were comparable overall, indicating the limited impact of geographic origin on this metal-chelating property.

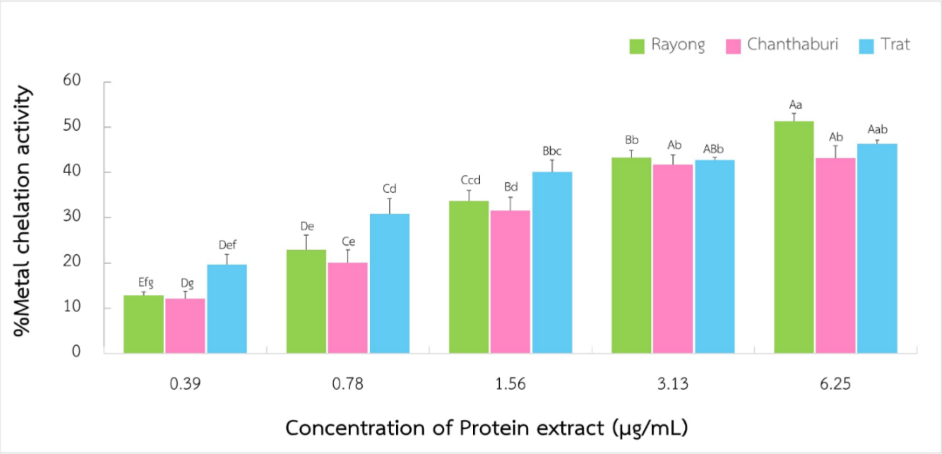


Figure 5. Metal chelating activity of crude protein extracts from *Etlingera pavieana* rhizomes collected from the Rayong, Chanthaburi, and Trat provinces. Different capital letters (A–E) indicate statistically significant differences among concentrations within the same province, while different lowercase letters (a–g) indicate significant differences between provinces at the same concentration ($p < 0.05$).

The differences in the anti-inflammatory and antioxidant activities of the crude protein extracts derived from *E. pavieana* rhizomes collected from Rayong, Chanthaburi, and Trat may be attributed to environmental, agronomic, and biochemical factors, as well as the harvesting stage. As demonstrated in previous studies, variations across cultivation sites can significantly affect the phytochemical profiles and bioactive compound contents of medicinal plants [11, 13]. Soil composition, microclimate, and sunlight exposure, among other factors, influence not only the yield but also the molecular structure and functionality of plant proteins and associated metabolites [12]. For protein extracts, the composition and abundance of amino acid residues with antioxidant or metal-chelating functional groups—such as hydroxyl, thiol, and carboxyl groups found in tyrosine, cysteine, and aspartic acid—may vary depending on the growing conditions. These residues are critical for radical scavenging and reducing power [21, 23]. Furthermore, despite the standardization of protein extraction procedures, trace levels of co-extracted secondary metabolites, such as phenolics or flavonoids, which synergistically enhance antioxidant effects, may differ between samples due to regional variations in metabolic pathways [3, 4]. These minor components may contribute to the variability observed in the DPPH scavenging, reducing power, and nitric oxide inhibition assays. Taken together, these results demonstrate the impact of environmental and agronomic factors on the biological activity of plant protein extracts. Understanding these variations is crucial for optimizing cultivation site selection and standardizing the quality of plant-based functional ingredients. Nevertheless, future studies employing protease digestion or protein depletion experiments will be required to confirm that the observed effects are specifically attributable to proteins.

4. Conclusions

In summary, the findings of this study demonstrate that crude protein extracts from *E. pavieana* rhizomes collected from three different cultivation sites exhibited variations in anti-inflammatory and antioxidant activities. Trat extracts showed the highest NO inhibitory activity and reducing power, while Rayong extracts exhibited the strongest metal-chelating capacity. These results contribute to the growing body of research on plant-derived protein extracts, particularly within the Zingiberaceae family, and highlight the impact of cultivation site on bioactivity. This study also highlights the potential of *E. pavieana* protein extracts as functional ingredients for health-promoting applications. Further studies should focus on identifying specific bioactive peptides and elucidating their mechanisms of action.

5. Acknowledgements

The authors would like to thank the Center of Excellence for Innovation in Chemistry (PERCH-CIC), the Commission on Higher Education, and the Ministry of Higher Education, Science, Research and Innovation, Thailand, for the financial support provided to J.P.

Author Contributions: Conceptualization, K.S. and S.R.; methodology, K.S. and S.R.; investigation, J.P., J.J., and S.C.; data curation, J.P., J.J., and S.C.; writing-original draft preparation, J.P.; writing-review and editing, K.S.; visualization, J.P.; supervision, K.S. and S.R.; funding acquisition, K.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was financially supported by (i) Burapha University (BUU), (ii) Thailand Science Research and Innovation (TSRI), and (iii) National Science Research and Innovation Fund (NSRF) (Fundamental Fund: Grant no. 3.4/2568).

Conflicts of Interest: "The authors declare no conflict of interest."

References

[1] Jafarzadeh, S.; Qazanfarzadeh, Z.; Majzoobi, M.; et al. Alternative proteins; A path to sustainable diets and environment. *Curr. Res. Food Sci.* **2024**, *9*, 100882. https://doi.org/10.1016/j.crfs.2024.100882

[2] Langyan, S.; Yadava, P.; Khan, F. N.; et al. Sustaining protein nutrition through plant-based foods. *Front. Nutr.* **2022**, *8*, 772573. https://doi.org/10.3389/fnut.2021.772573

[3] Chantaranothai, C.; Palaga, T.; Karnchanatat, A.; Sangvanich, P. Inhibition of nitric oxide production in the macrophage-like RAW 264.7 cell line by protein from the rhizomes of Zingiberaceae plants. *Prep. Biochem. Biotechnol.* **2013**, *43*(1), 60-78. https://doi.org/10.1080/10826068.2012.697958

[4] Inthuwanarud, K.; Sangvanich, P.; Puthong, S.; Karnchanatat, A. Antioxidant and antiproliferative activities of protein hydrolysate from the rhizomes of Zingiberaceae plants. *Pak. J. Pharm. Sci.* **2016**, *29* (6), 1893-1900.

[5] Sompinit, K.; Lersiripong, S.; Reamtong, O.; et al. In vitro study on novel bioactive peptides with antioxidant and antihypertensive properties from edible rhizomes. *LWT* **2020**, *134*, 110227. https://doi.org/10.1016/j.lwt.2020.110227

[6] Phonsena, P. *Medicinal Plants in Khao Hin Son Herb Garden*; Jettanaromphun Printing: **2007**, 301 pp. (In Thai)

[7] Poulsen, A. D.; Phonsena, P. Morphological variation and distribution of the useful ginger *Etlingera pavieana* (Zingiberaceae). *Nord. J. Bot.* **2017**, *35*, 467-475. https://doi.org/10.1111/njb.01407

[8] Iawsipo, P.; Srisook, E.; Ponglikitmongkol, M.; Somwang, T.; Singaed, O. Cytotoxic effects of *Etlingera pavieana* rhizome on various cancer cells and identification of a potential anti-tumor component. *J. Food Biochem.* **2018**, *42*, e12508. https://doi.org/10.1111/jfbc.12508

[9] Srisook, K.; Srisook, E. Pharmacological activities and phytochemicals of *Etlingera pavieana* (Pierre ex Gagnep) R.M.Sm. In *Medicinal Plants – Use in Prevention and Treatment of Diseases*. **2019**. https://doi.org/10.5772/intechopen.89277

[10] Poonasri, M.; Chiranthanut, N.; Srisook, E.; Srisook, K. Anti-neuroinflammatory activity of *Etlingera pavieana* rhizomal extract in LPS-induced microglial cells. *Naresuan Phayao J.* **2021**, *14*, 30-38.

[11] Hao, J. Y.; Wan, Y.; Yao, X. H.; et al. Effect of different planting areas on the chemical compositions and hypoglycemic and antioxidant activities of mulberry leaf extracts in southern China. *PLoS ONE* **2018**, *13*, e0206930. https://doi.org/10.1371/journal.pone.0206930

[12] Khumaida, N.; Syukur, M.; Bintang, M.; Nurcholis, W. Phenolic and flavonoid content in ethanol extract and agro-morphological diversity of *Curcuma aeruginosa* accessions growing in West Java, Indonesia. *Biodiversitas* **2019**, *20*, 656-663. https://doi.org/10.13057/biodiv/d200306

[13] Yahyaoui, A.; Arfaoui, M. O.; Rigane, G.; et al. Investigation on the chemical composition and antioxidant capacity of extracts from *Crataegus azarolus* L.: Effect of growing location of an important Tunisian medicinal plant. *Chem. Afr.* **2019**, *2*, 361-365. https://doi.org/10.1007/s42250-019-00054-1

[14] Deesrisak, K.; Yingchutrakul, Y.; Krobthong, S.; Roytrakul, S.; Chatupheeraphat, C.; Subkorn, P.; Anurathapan, U.; Tanyong, D. Bioactive peptide isolated from sesame seeds inhibits cell proliferation and induces apoptosis and autophagy in leukemic cells. *EXCLI J.* **2021**, *20*, 709-721.

[15] Srisook, K.; Srisook, E.; Nachaiyo, W.; Chan-In, M.; Thongbai, J.; Wongyoo, K.; Chawsuanthong, S.; Wannasri, K.; Intasuwan, S.; Watcharanawee, K. Bioassay-guided Isolation and mechanistic action of anti-Inflammatory agents from *Clerodendrum inerme* Leaves. *J. Ethnopharmacol.* **2015**, *165*, 94-102. https://doi.org/10.1016/j.jep.2015.02.043

[16] Pechroj, S.; Kamonporn, P.; Oraphan, N.; et al. Comparative evaluation of antioxidant and anti-inflammatory activities of four seaweed species from the east coast of the Gulf of Thailand. *J. Appl. Phycol.* **2020**, *13*, 11-21.

[17] Uthairat, C.; Srisook, E.; Srisook, K. Effects of drying methods and extraction conditions on total phenolic and flavonoid content and antioxidant activities of *Helicteres isora* L. fruit extracts. *Burapha Sci. J.* **2017**, *22*.

[18] Alderton, W. K.; Cooper, C. E.; Knowles, R. G. Nitric oxide synthases: Structure, function, and inhibition. *Biochem. J.* **2001**, *357*, 593-615. https://doi.org/10.1042/bj3570593

[19] Zedler, S.; Faist, E. The impact of endogenous triggers on trauma-associated inflammation. *Curr. Opin. Crit. Care* **2006**, *12*, 595-600. https://doi.org/10.1097/MCC.0b013e3280106806

[20] Libby, P. Inflammatory mechanisms: The molecular basis of inflammation and disease. *Nutr. Rev.* **2007**, *65*, 140-146. https://doi.org/10.1301/nr.2007.dec.S140-S146

[21] Elias, R. J.; Kellerby, S. S.; Decker, E. A. Antioxidant activity of proteins and peptides. *Crit. Rev. Food Sci. Nutr.* **2008**, *48*, 430-441. https://doi.org/10.1080/10408390701425615

[22] Ulrich, K.; Jakob, U. The role of thiols in antioxidant systems. *Free Radic. Biol. Med.* **2019**, *140*, 14-27. https://doi.org/10.1016/j.freeradbiomed.2019.05.035

[23] Shi, Q.; Wang, J. J.; Chen, L.; et al. Fenton reaction-assisted photodynamic inactivation of calcined melamine sponge against *Salmonella* and its application. *Food Res. Int.* **2020**, *151*, 110847. https://doi.org/10.1016/j.foodres.2021.110847

[24] Canabady-Rochelle, L. L. S.; Selmeczi, K.; Collin, S.; Pasc, A.; Muhr, L.; Boschi-Muller, S. SPR screening of metal chelating peptides in a hydrolysate for their antioxidant properties. *Food Chem.* **2018**, *239*, 478-485. https://doi.org/10.1016/j.foodchem.2017.06.116

ASEAN
Journal of Scientific and Technological Reports
Online ISSN:2773-8752

Research article

The Application of Geographic Information Systems (GIS) for Conservative Local Plant Database Development

Sompond Puengsom[1], Akera Ratchavieng[2], Napharat Chooprai[3]*, and Noppasak Tantisattayanon[4]

[1] Faculty of Industry and Technology, Rajamangala University of Technology Rattanakosin, Nakhon Pathom, 73170, Thailand
[2] Faculty of Industry and Technology, Rajamangala University of Technology Rattanakosin, Nakhon Pathom, 73170, Thailand
[3] Faculty of Industry and Technology, Rajamangala University of Technology Rattanakosin, Nakhon Pathom, 73170, Thailand
[4] Faculty of Industry and Technology, Rajamangala University of Technology Rattanakosin, Nakhon Pathom, 73170, Thailand
* Correspondence: napharat.cho@rmutr.ac.th

Citation:
Puengsom, S.; Ratchavieng, A.; Chooprai, N.; Tantisattayanon, N The Apllication of geographic information systems (GIS) for conservative local plant database development. *ASEAN J. Sci. Tech. Report.* **2025**, *28*(5), e258063. https://doi.org/10.55164/ajstr.v28i5.258063.

Article history:
Received: February 25, 2025
Revised: September 10, 2025
Accepted: September 11, 2025
Available online: September 24, 2025

Abstract: The study of The Application of Geographic Information System (GIS) for Conservative Local Plant Database Development has the objectives 1) to apply Geographic Information System (GIS) for Conservative Local Plant Database Development, 2) to evaluate the acceptance of Geographic Information System (GIS) for Conservative Local Plant Database Development, and 3) to evaluate the accuracy of searching for conserved plants in the database applying geographic information system. The local plant conservation database system utilizes geographic technology and a database system that can manage user data, search for recorded plant species, update and manage plant species data, and display plant species data and their status. The results of the acceptance assessment from 180 experts and system users revealed that the overall score was high ($\bar{x}$=4.00, S.D. = 0.78), and when separated into categories, the scores were: Actual System Use ($\bar{x}$=4.03, S.D. = 0.78), Perceived Usefulness ($\bar{x}$=4.02, S.D. = 0.79), Behavioral Intention to Use ($\bar{x}$=3.99, S.D. = 0.76), Attitude Toward Using ($\bar{x}$=3.98, S.D. = 0.80) and Perceived Ease of Use ($\bar{x}$=3.93, S.D. = 0.78). The results of the evaluation of retrieval efficiency found that Precision = 0.96, Recall = 1.00, Accuracy = 1.00, and the overall efficiency evaluation F-measure = 0.98 is rated as good.

Keywords: Geographic information system; conservative local plant; database development; technology acceptance model.

1. Introduction

Plants are important ecosystem services and resources necessary for sustainable development. Plant conservation is the conservation of the taxonomic diversity and the ecological functions of plant diversity. Most plant conservation efforts focus on maintaining habitats to support forest populations [1]. The most important approach to conservation of rare and endangered plant species is to complete the study of the flora of Thailand as quickly as possible, as this information not only indicates the characteristics of each plant species, but also includes the distribution of plants in various forest types, as well as information on the apparent plant population, endemic species, etc. This information will serve as building blocks for managing plant resources in the country. When the ecosystem is destroyed, whether by natural or human causes, these plants may become extinct before they are even discovered. For example, mountain tribes in mountainous areas frequently relocate their settlements, and thus, areas covered with a variety of plants are constantly being destroyed to make way for cultivation and shifting cultivation. As each digging destroys the entire plant ecosystem, conservation of endangered plants must be considered in terms of natural resource management and wood utilization planning. In Thailand, although there has never been a serious study of endangered plants, it is estimated that a significant number of plants have disappeared from their original habitats. The operation of natural resource conservation requires the use of social and legal measures, including the encouragement of local efforts in conserving forest resources and helping to maintain them in their original condition, thereby preventing deterioration for the benefit of local livelihoods. Currently, there are many problems with forest resources [2].

Conservation of plant resources is therefore essential to preserve critically endangered plant species that may hold economic importance in the future, and to conserve special traits among plant populations for plant breeding, such as resistance to insects and diseases, and drought tolerance. The conserved plant species can be used as a base population for further breeding or as a source for future genetic research. This conservation can be achieved through in situ conservation, which involves preserving plant habitats by considering those of near-extinction or at-risk plants (due to human activities). This approach involves stopping the destruction of plant habitats to prevent subsequent loss and attempting to restore destroyed habitats to their natural state. By collecting in a collection plot (Ex Situ), this method involves collecting rare or near extinct plant species and planting them in propagation plots in botanical gardens, arboretums, and trying to make the ecosystem similar to the natural environment where the plants live [3]. The researcher studied the geographic information system, a computer system used to import and analyze geographic location data, which is then displayed as maps or images. Currently, GIS is increasingly applied in conservation work due to its ability to collect, analyze, and display spatial data, as well as provide accurate and quick coordinates. Therefore, geographic information technology is applied to help increase data collection efficiency for the local conservation plant database at Rajamangala University of Technology Rattanakosin, Wang Klaikangwon Campus, as the area covers more than 235 rai and features a variety of plant species, including rare ones. Based on the aforementioned benefits, an information system has been developed to create a database of rare or conserved plant species, enabling the survey and protection of the tree population. However, this study is important for establishing a geographic information platform for local agencies, particularly the Royal Project for Sustainable Resource Conservation (RSPG). Coordinating Center at RMUTR, as part of Activity 8: Special Activities Supporting Resource Conservation under RSPG. However, this focus is not mentioned in the introduction or discussion. The authors should highlight the platform's application to clarify its practical significance. RMUTR, Wangklaiwangwon campus, lacks a systematic approach to collecting, storing, and presenting plant information.

In the order that, the study of the Application of Geographic Information System (GIS) for Conservative Local Plant Database Development has the objectives were to apply Geographic Information System (GIS) for Conservative Local Plant Database Development, to evaluate the acceptance of Geographic Information System (GIS) for Conservative Local Plant Database Development, and to evaluate the accuracy of searching for conserved plants in the database applying geographic information system.

2. Related Theories and Research

In this research, the research team reviewed the basic knowledge theories that led to this research as follows:

2.1 Related theories

2.1.1 The Plant Genetic Conservation Project initiated by Her Royal Highness Princess Maha Chakri Sirindhorn (P.G.S.P.) was established in 1993 to create understanding and awareness of the importance of various resources in Thailand, leading to encouragement of participatory thinking and practice that benefits the Thai people, as well as to create a nationwide resource information system [4]. Since 1960, His Majesty King Rama IX has had the foresight to conserve rubber trees and collect plant species from various regions of the country to be planted in Chitralada Garden. Thus, the Plant Genetic Conservation Project was established to fulfill the royal initiative and continue his intention, as evident in various activities, including the protection of resources in natural forest areas, surveying and collecting endangered resources, cultivating and maintaining the collected resources in suitable and safe areas, and carefully conserving and utilizing them. There were efforts to study and evaluate resources in various aspects, such as components, properties, and utilization, to create a computer resource information system, initiate long-term (30-50 years) resource planning and development, and promote resource conservation awareness among various target groups [5- 6].

2.1.2 Geographic Information System (GIS)

A GIS is a computer system that analyzes and displays geographically referenced data based on geographical inputs such as maps, satellite images, numbers, letters, and distances, and analyzes them through computer programs. The results are often highly accurate and can be applied in many areas. GIS is useful in many areas, especially environmental management, urban planning, and public utility management, by calculating the area to be used from a map image. This includes measuring the distance for road construction or defining points on a map for laying water pipes [7-9].

2.1.3 ArcGIS

ArcGIS is a geospatial software platform for viewing, editing, managing, and analyzing geographic data. Esri develops ArcGIS for mapping on desktop, mobile devices, and the web. [8] ArcGIS offers spatial data analysis features and tools that align with the organization's work processes, enabling the analysis and display of insights from internal and external data, as well as the publication of results in the form of maps, applications, dashboards, and reports. This study utilized ArcGIS's Data Collection & Management features to efficiently and securely explore, store, utilize, and publish data by integrating with other data storage systems and adding spatial dimensions to these datasets [10].

2.1.4 Technology Acceptance Models (TAM)

The Technology Acceptance Model (TAM) is a significant predictor of technology adoption behavior in learning, combining two key components: Perceived Usefulness (PU) and Perceived Ease of Use (PEOU). This model has been accepted and proven effective in explaining the acceptance and adoption of technology in various contexts [11-12].

The Technology Acceptance Model (TAM) is a widely accepted theory for examining the factors and perspectives that influence the adoption of new technologies. The purpose of the TAM, developed by Davis in 1986, is to model user acceptance of information systems. The primary objective of the model is to encourage the adoption of information technology (IT) through acceptance stimulation. This promotion of acceptance can only be successful if the influencing factors can be identified, which can be achieved by studying users' perceptions of technology use [13].

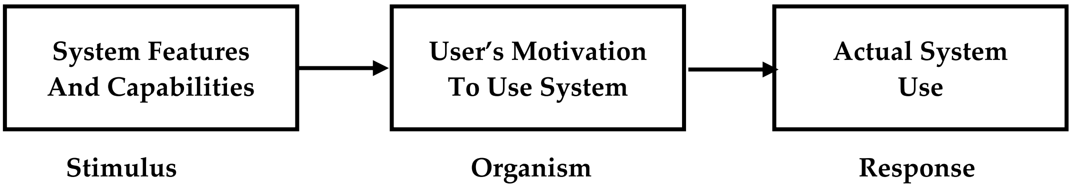


Figure 1. Conceptual Model for Technology Acceptance [1]

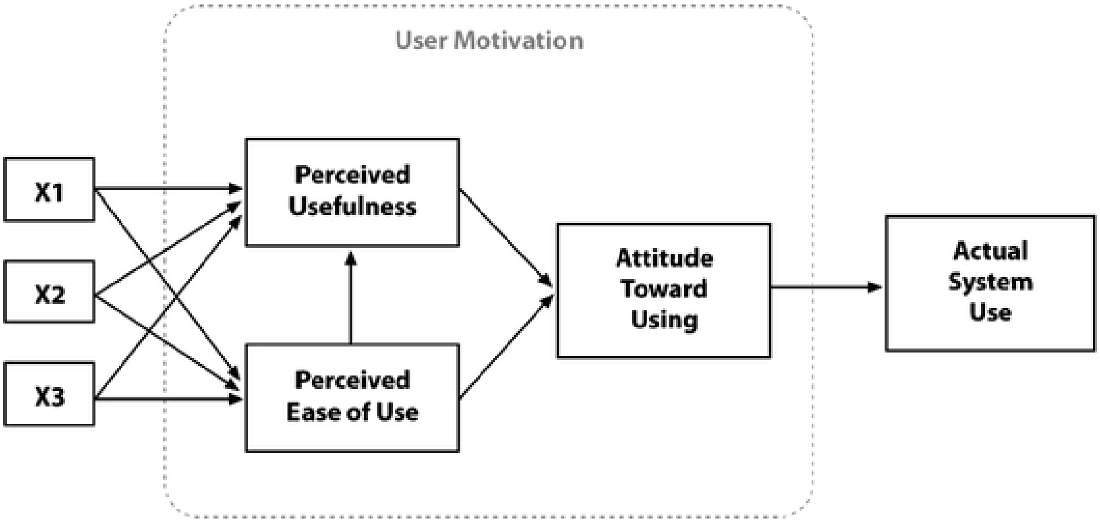


Figure 2. Technology Acceptance Model [15]

Perceived Ease of Use (PEoU), Perceived Usefulness (PU), Attitude Toward Using (ATU), Behavioral Intention to Use (BI), and Actual System Use (ASU).

Model explanation

The benefit of TAM is that it serves as an important tool in examining technology acceptance behavior, as well as:

- Helps in predicting technology adoption and usage: TAM can predict the adoption behavior of new technology users by studying two important factors: Perceived Usefulness (PU) and Perceived Ease of Use (PEOU), which help in predicting the extent of adoption [16].
- Helps in developing and improving technology: with TAM, developers or project managers can understand the user adoption factors and can use such information to improve the usefulness or user-friendliness of a technology or design [17].
- Multi-contextual application: TAM can be applied in various fields and contexts, such as education, business administration, medicine, or even mobile applications, which helps in studying and analyzing different aspects of user needs [18].
- Helps in understanding user behavior in different situations: TAM helps in identifying which factors influence the adoption of technology in different contexts, such as the adoption of technology in organizations or education after a technology change [19].
- Can be used as a conceptual framework for research: TAM is used as a conceptual framework for research on technology user behavior, which helps researchers to focus on the key factors affecting the acceptance and use of technology. [20]

Thus, TAM is a helpful tool for researching and developing new technologies, as well as gaining a deeper understanding of users' technology use behavior.

2.2 Related Studies

Acceptance of GIS within ERP Systems: A Research Study in Higher Education by Simona Sternad Zabukovšek et al. [12] utilized the Technology Acceptance Model (TAM), one of the most commonly employed research models for studying the intention to use and usability of information systems, and augmented the TAM with external factors. The primary objective of the study was to identify the factors influencing the acceptance of GIS and individual attitudes towards the use of GIS-integrated ERP systems. Survey data analysis was conducted using structural equation modeling with the Partial Least Squares (PLS) method. The study examined 12 external factors in the TAM model, of which 10 were found to be significant. The most important factors affecting the intention to use ERP-GIS systems were perceived usefulness and attitudes towards the ERP-GIS system, followed by education and training in ERP-GIS systems and perceived enjoyment of using ERP-GIS systems.

The research showed that external factors such as training, education, perceived enjoyment of using ERP-GIS systems, and perceived usefulness of using the system had significant influences on the user attitudes and intentions to use ERP-GIS systems, especially in the context of higher education, in which there was much disparity in knowledge and skills in computer science and ERP systems. Therefore, providing appropriate education and training plays a crucial role in increasing acceptance and promoting the use of technology in the future. This study also emphasizes the importance of increasing the factor of Perceived Enjoyment, which is a crucial factor in attitudes and acceptance of ERP-GIS systems [11].

Application of Geographic Information System (GIS) for the forest database of Magnolia Sirindhorniae, Lopburi Province: The application of GIS to create and manage a forest database for Magnolia Sirindhorniae, a rare and endangered species in Lopburi Province, aims to monitor, conserve and manage the distribution and growth of this species using geographic information and GIS technology to monitor and manage the population of Magnolia Sirindhorniae in Lopburi Province by using GIS to map the distribution of tree species, growth and conservation areas, and create a comprehensive database to store and manage all the collected data. This system was linked to the GIS platform, enabling efficient display, analysis, and updating of data. GIS tools were used to analyze the distributional characteristics of Magnolia Sirindhorniae, which enables the identification of suitable habitats, assessment of potential threats, planning, and management of future conservation efforts [22].

Utami and Ramdani [23] studied GEMAR, a web-based GIS for emergency management and ambulance routing. It utilizes web-based technology to facilitate access to information and decision-making, with a primary focus on ambulance dispatch to the scene of an incident. The development of this system aims to improve the efficiency (namely, speed and accuracy) of the ambulance dispatch process, while also enabling the management of various emergency data. The research employed a system assessment approach to examine technology acceptance and usability issues, with an average score for technology acceptance ranging from 3.70 to 4.40, and an average score for usability ranging from 4.00 to 4.50, as assessed by 10 experts, including 7 medical experts and 3 GIS experts. The average score of technology acceptance was between 3.70 and 4.40, while the average score of usability was between 4.00 and 4.50, indicating that the system had user-friendly features and was ready for future use. Medical experts also noted that the system was easy to use and navigate. Both subsystems were useful for medical practitioners to understand the situation in emergency management. The experts' opinions obtained were used to improve and further develop the system [23].

The development of a web-based geographic information system (GIS) for roadside trees in the Nakhon Ratchasima Municipality aimed to manage and track roadside trees within the municipality's area by utilizing GIS technology to collect and display various tree data, including location, type, age, and health, on online maps. Surveys and the collection of roadside tree data, including GPS location, type, size, and health information, were conducted. A website was created to display GIS maps linked to tree data in the area, allowing users to search or view tree details. A database was used to store tree data so that it can be accessed and updated conveniently by linking GIS maps through GIS map technology such as Google Maps or OpenStreetMap to display tree locations and related data so that the system can help the municipality effectively track and manage tree data, promote sustainable tree care in urban areas, and respond to the needs of the community [24].

Decision Support System for Economic Crop Planning Using Geoinformatics Technology by Phiset Senawong has the objective to develop a geoinformatics database to support decision-making in land use planning that can be used via the Internet by analyzing the plots of interest, consisting of 1) cases of restrictions and conditions on land use, 2) cases of appropriate land use and conservation for sustainable use, and 3) cases of determining alternatives to support decision-making in land use that provides the highest benefit, which has developed tools using the capabilities of Geoprocessing service of ArcGIS for server which can be used and accessed via Web Browser [25]. Wutthitaweewat and Chaiyawat had the following objectives: 1) Study of technology quality factors affecting the acceptance of insurance life technology (InsurTech). 2) Study the level of technology acceptance and decision-making trends in purchasing life insurance policies via applications. 3) Examining the consistency of technology quality and acceptance of InsurTech life insurance technology on the tendency of life insurance policy purchase decisions via an applications model. 4) Study the influence of technology quality and acceptance of InsurTech life insurance technology on the tendency of consumers to make decisions to purchase life insurance policies via applications in Bangkok. The tools used in the study were a 5-stage evaluation questionnaire, which was created based on the concept and theory of the Information System Success Model (ISSM) and the Technology Acceptance Model (TAM). The results of the study on the level of technology acceptance (Technology Acceptance Model: TAM) found that the overall acceptance of InsurTech technology was at a high level ($\bar{x}$ = 4.25, SD = 0.69) [26]. Sirimongkhon et al. developed a facial recognition-based attendance system to improve attendance management efficiency, reduce errors, and allow effective identification of attendees. The system was tested for accuracy using the Confusion Matrix [27]. The study "Forecasting Direction and Analyzing Dengue Risk in Bangmaenang Sub-District, Bangyai District, Nonthaburi Province" utilized Geographic Information Systems (GIS) and data mining to forecast the direction and identify risk areas of dengue fever outbreaks, analyze their underlying factors, and compare the efficiency and accuracy of the forecasting algorithms. The research study area is Bang Mae Nang Sub-district, Bang Yai District, Nonthaburi Province. In this research, the efficiency of the model is evaluated using the Confusion Matrix, and its performance is measured using Accuracy, Precision, Recall, and Overall Efficiency (F-measure) [28].

3. Research Method

3.1. Developing a GIS-based information system for a local conserved plant database using the Software Development Life Cycle (SDLC) theory [29-30]

3.1.1 System Planning

The researcher studied and analyzed the problems in recording local plant species data in Rajamangala University of Technology Rattanakosin, Wang Klai Kangwon Campus, and nearby areas where plant species had not been systematically recorded. Therefore, the researcher saw the need to develop a GIS-integrated information system to create a database of locally conserved plants, supporting the Plant Genetic Conservation Project (RSPG). The development of the system was based on the following initial requirements:

3.1.1.1 General users can view plant species data, search for recorded plant species data, and view the location of recorded plant species.

3.1.1.2 Members or system staff can log in and out, manage their own member data, and manage plant species data. And can manage system data

3.1.1.3 Administrators can log in and log out, manage member and staff data, grant member system usage rights, manage plant species data, as well as system data. The system is designed to be displayed and usable through all devices using Responsive Web Design techniques [31]. System Analysis: Develop a GIS-based information system to create a local database of conserved plants. The analysis part was done as follows:

3.1.1.4 Design the system architecture according to Figure 3 below.

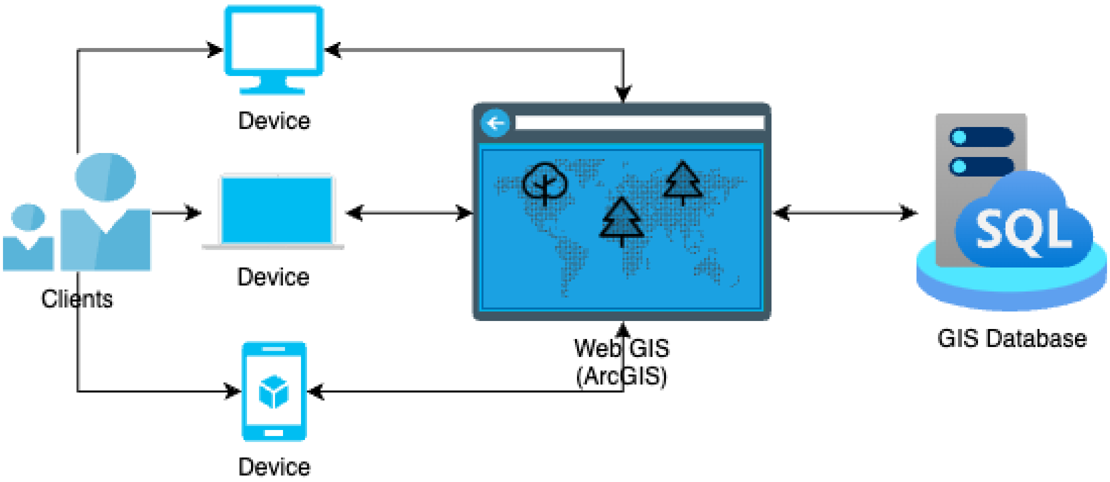


Figure 3. Architecture of the System

3.1.1.5 Context Diagram

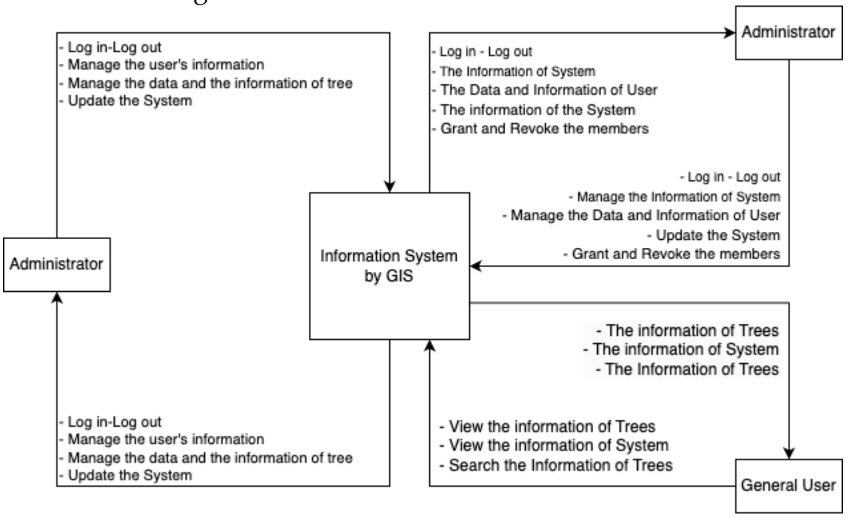


Figure 4. Context Diagram

3.1.1.6 Dataflow Diagram

Figure 5. Dataflow Diagram level 0

3.1.1.7 Entity-Relationship Diagram

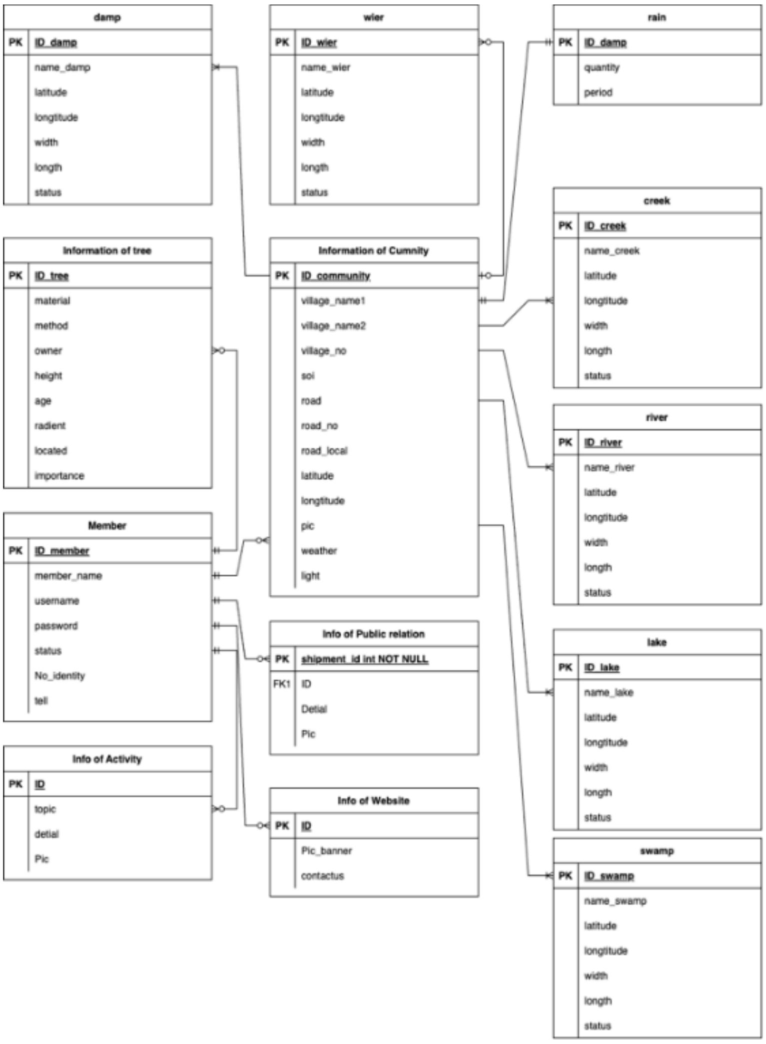


Figure 6. Entity-Relationship Diagram

3.1.2 System Design

The development of the GIS-based system to create a database of local conserved plants was done using ArcGIS API, together with Dreamweaver, a web development tool that facilitates designing and writing HTML, CSS, JavaScript, and PHP codes. These tools can be used to create and manage web applications using the ArcGIS API for JavaScript and a MySQL database management program for storing and retrieving spatial data.

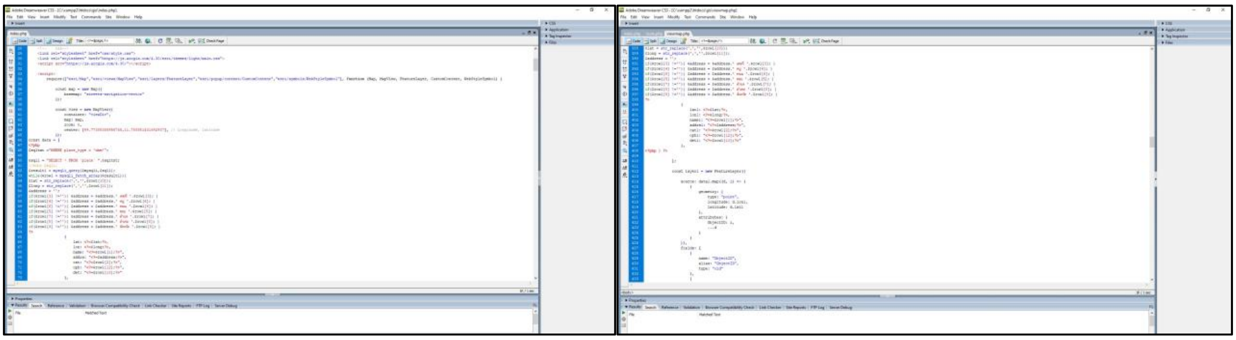


Figure 7. Using Adobe Dreamweaver with ArcGIS API for JavaScript and MySQL.

3.1.3 System Implementation

System implementation involves migrating from the old system to a new one using the direct installation approach or replacing the old system immediately with the new one, and then proceeding to the system testing process.

3.1.4 System Testing: uses the black box testing method

A software testing method that focuses on checking the functionality of the system or application, regardless of the internal structure or operation of the underlying code. It focuses on whether the system works correctly according to the requirements by examining the input and output received through testing various system functions against the requirements or user needs. Five experts were used to test the system's efficiency [32-35]. The testing utilized the plants' records from specific areas, which may not fully represent the entire real-world diversity. Additionally, the controlled testing environment may differ from actual usage conditions. Further evaluation in more diverse contexts is recommended.

3.1.5 System Maintenance

System Maintenance is a process of maintaining and improving the software or hardware system to maintain both long-term efficiency and user satisfaction. This step occurs after the system is in use (Post-deployment Phase) and usually involves correcting errors or bugs in the system, adapting the system to a new environment, and improving features or adding new capabilities to make the system more responsive to user needs [36-38]

3.2 Technology Acceptance Evaluation

Arranging for 180 users and related people to evaluate the acceptance of geographic information technology to create a local conserved tree species database, covering Perceived Ease of Use (PEoU), Perceived Usefulness (PU), Attitude Toward Using (ATU), Behavioral Intention to Use (BI), and Actual System Use (ASU).

3.3 Evaluating the accuracy

Evaluating the accuracy of searching for conserved tree species in the GIS-based database by measuring the Recall and Precision [28].

Table 1. Confusion Matrix

	TRUE	**FALSE**
TRUE	True Positive (TP)	False Positive (FT)
FALSE	False Negative (FN)	True Negative (TN)

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \quad (1)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3)$$

$$\text{F-measure} = 2x\frac{Precision\ x\ Recall}{Precision+Recall} \quad (4)$$

4. Research Result

The result of the application of geographic information technology to create a local plant conservation database, as well as the collection of GIS data, mapping of plant location, and assessment of the accuracy of the conserved plant searching accuracy of the GIS-based database by measuring the precision, recall, and accuracy.

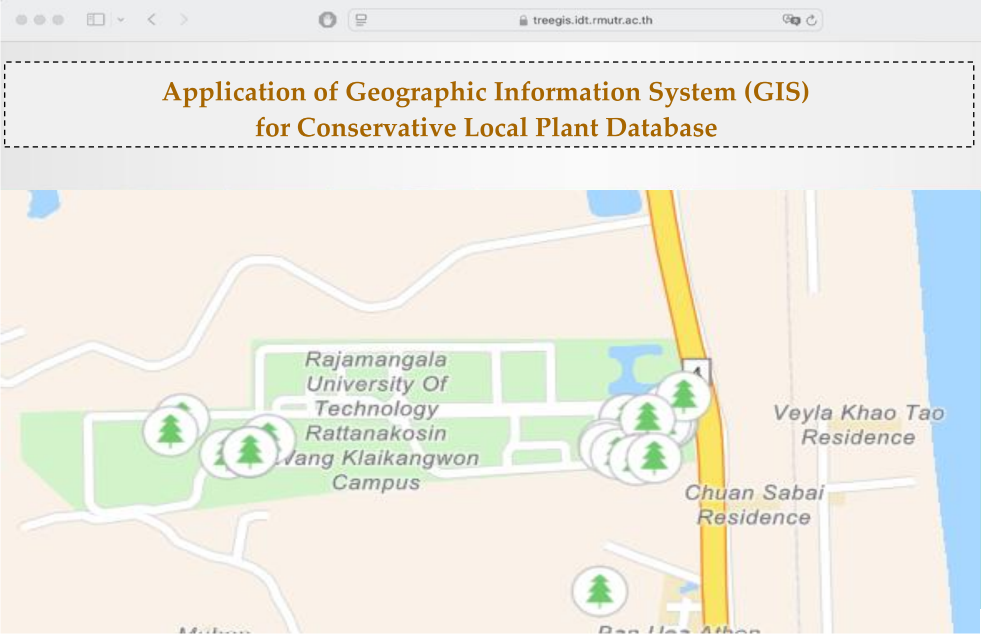
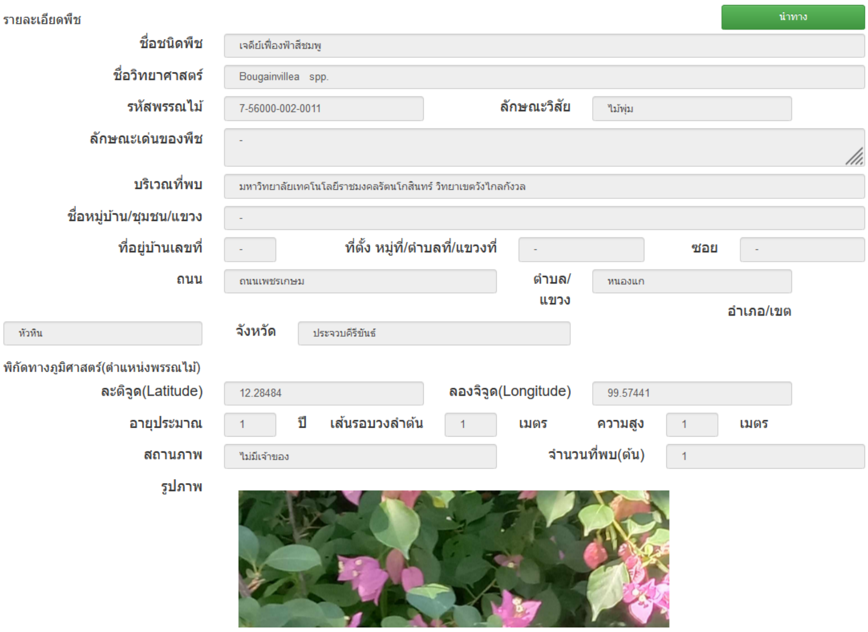


Figure 8. Recorded plant information drawn from the GIS database.

Figure 8 presents the location of plant species recorded via RSPG Form 5, allowing users to view basic information and navigate to each plant's location.



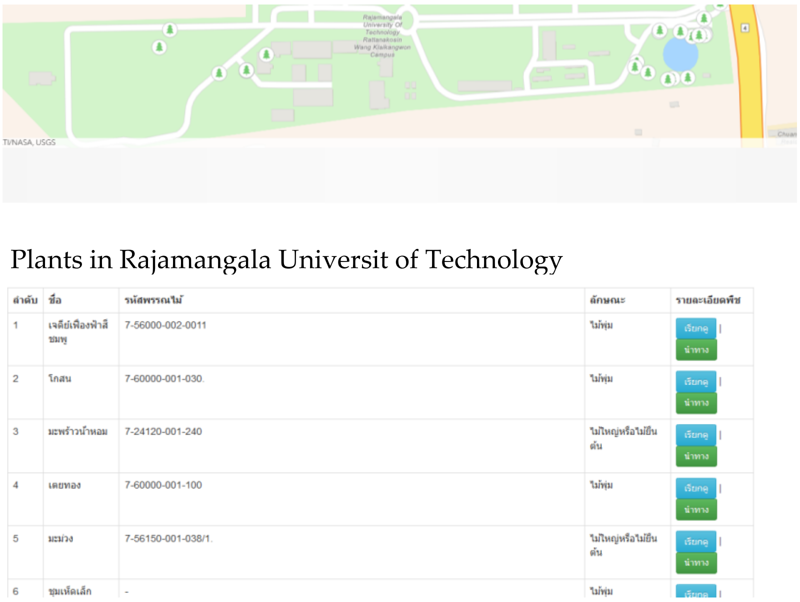
รายละเอียดพืช

ชื่อชนิดพืช เจดีย์เฟื่องฟ้าสีชมพู
ชื่อวิทยาศาสตร์ Bougainvillea spp.
รหัสพรรณไม้ 7-56000-002-0011
ลักษณะวิสัย ไม้พุ่ม
ลักษณะเด่นของพืช -
บริเวณที่พบ มหาวิทยาลัยเทคโนโลยีราชมงคลรัตนโกสินทร์ วิทยาเขตวังไกลกังวล
ชื่อหมู่บ้าน/ชุมชน/แขวง -
ที่อยู่บ้านเลขที่ -
ที่ตั้ง หมู่ที่/ตำบลที่/แขวงที่ -
ซอย -
ถนน ถนนเพชรเกษม
ตำบล/แขวง หนองแก
อำเภอ/เขต หัวหิน
จังหวัด ประจวบคีรีขันธ์

พิกัดทางภูมิศาสตร์(ตำแหน่งพรรณไม้)
ละติจูด(Latitude) 12.28484
ลองจิจูด(Longitude) 99.57441
อายุประมาณ 1 ปี
เส้นรอบวงลำต้น 1 เมตร
ความสูง 1 เมตร
สถานภาพ ไม่มีเจ้าของ
จำนวนที่พบ(ต้น) 1
รูปภาพ

Figure 9. Plant species information and navigation bar

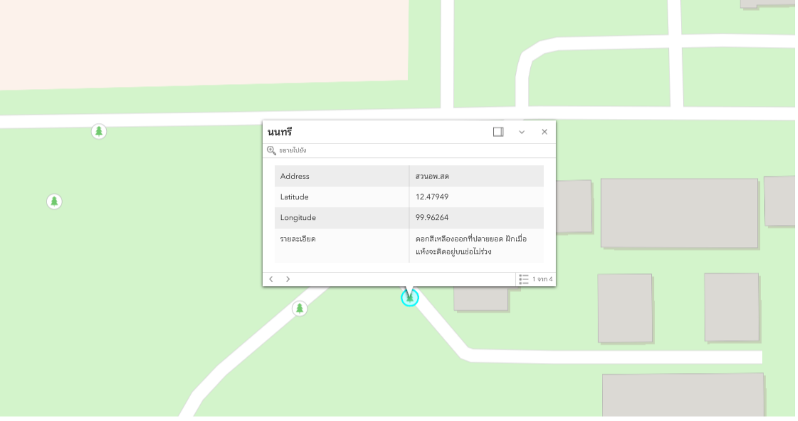
Figure 9 presents the data entry form for plant species based on RSPG Form 5, illustrating the process of recording plant information. The system also enables users to search and retrieve plant data, as demonstrated in Figure 10.



Plants in Rajamangala Universit of Technology

ลำดับ	ชื่อ	รหัสพรรณไม้	ลักษณะ	รายละเอียดพืช
1	เจดีย์เฟื่องฟ้าสีชมพู	7-56000-002-0011	ไม้พุ่ม	เรียกดู | นำทาง
2	โกสน	7-60000-001-030.	ไม้พุ่ม	เรียกดู | นำทาง
3	มะพร้าวน้ำหอม	7-24120-001-240	ไม้ใหญ่หรือไม้ยืนต้น	เรียกดู | นำทาง
4	เดยทอง	7-60000-001-100	ไม้พุ่ม	เรียกดู | นำทาง
5	มะม่วง	7-56150-001-038/1.	ไม้ใหญ่หรือไม้ยืนต้น	เรียกดู | นำทาง
6	ชุมเห็ดเล็ก	-	ไม้พุ่ม	เรียกดู |

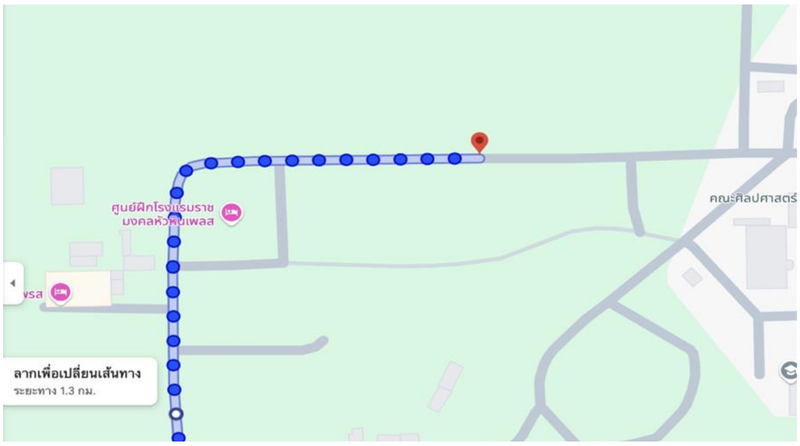
Figure 10. Searching for local plants



Address	สวนอพ.สด
Latitude	12.47949
Longitude	99.96264
รายละเอียด	ดอกสีเหลืองออกที่ปลายยอด ฝักเมื่อแห้งจะติดอยู่บนช่อไม่ร่วง

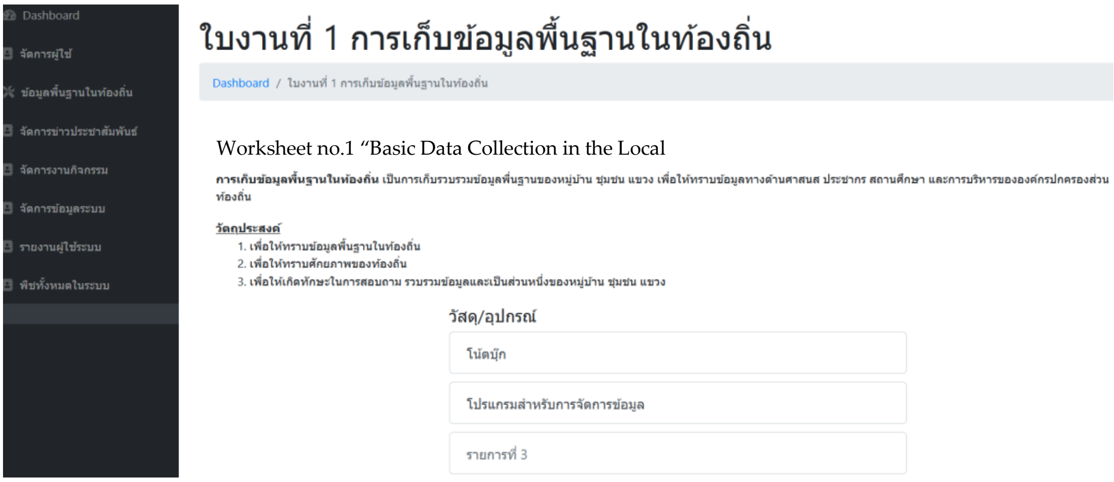
Figure 11. Show the Tree Information

Figure 11 illustrates the basic information of the queried plant species and provides the functionality to identify a route to access the specific plant, as shown in Figure 12.



Worksheet no.1 “Basic Data Collection in the Local Communication

Figure 12. Route to the tree



ใบงานที่ 1 การเก็บข้อมูลพื้นฐานในท้องถิ่น

Worksheet no.1 “Basic Data Collection in the Local

การเก็บข้อมูลพื้นฐานในท้องถิ่น เป็นการเก็บรวบรวมข้อมูลพื้นฐานของหมู่บ้าน ชุมชน แขวง เพื่อให้ทราบข้อมูลทางด้านศาสนส ประชากร สถานศึกษา และการบริหารขององค์กรปกครองส่วนท้องถิ่น

วัตถุประสงค์
1. เพื่อให้ทราบข้อมูลพื้นฐานในท้องถิ่น
2. เพื่อให้ทราบศักยภาพของท้องถิ่น
3. เพื่อให้เกิดทักษะในการสอบถาม รวบรวมข้อมูลและเป็นส่วนหนึ่งของหมู่บ้าน ชุมชน แขวง

วัสดุ/อุปกรณ์

โน้ตบุ๊ก

โปรแกรมสำหรับการจัดการข้อมูล

รายการที่ 3

Figure 13. Back end for recorded information per Worksheet no. 1 of the Plant Genetic Conservation Project

Assessment of acceptance of the GIS for the local conserved plant database

The technology acceptance model of the application of GIS in the local conserved plant database was assessed by 180 administrators, experts, and general users. The result was presented as the mean system efficiency and its standard deviation. Table 2 shows the TAM result.

However, the Technology Acceptance Model (TAM) was adopted in this study due to its strong focus on perceived ease of use and perceived usefulness, which are critical factors in understanding user acceptance of information systems. Compared to alternative models such as UTAUT, TAM offers a simpler and more parsimonious framework, making it suitable for studies with limited sample sizes or heterogeneous user groups. In contrast to IS-Success, which emphasizes system quality, information quality, and net benefits, TAM directly examines users’ behavioral intention to use the system, aligning closely with the primary objective of evaluating user acceptance of the web-based GIS platform.

Table 2. Assessment of acceptance of the GIS for the local conserved plant database

Aspects	**Executives**		**Plant Experts**		**IT Experts**		**General Users**		**Total**		
	x̄	**S.D.**	**x̄**	**S.D.**	**x̄**	**S.D.**	**x̄**	**S.D.**	**x̄**	**S.D.**	**Rating**
Perceived Ease of Use: PEoU	4.12	0.82	3.85	0.81	4.03	0.82	4.02	0.77	3.93	0.78	High
Perceived Usefulness: PU	3.83	0.75	3.98	0.86	4.01	0.81	4.00	0.74	4.02	0.79	High
Attitude Toward Using: ATU	4.10	0.84	3.90	0.71	4.17	0.87	3.89	0.75	3.98	0.8	High
Behavioral Intention to Use: BI	3.80	0.82	4.03	0.77	4.13	0.85	3.99	0.78	3.99	0.76	High
Actual System Use: ASU	4.00	0.68	4.03	0.77	4.08	0.86	4.02	0.87	4.03	0.78	High
Total	**4.00**	**0.78**	**4.00**	**0.78**	**4.10**	**0.84**	**4.00**	**0.78**	**4.00**	**0.78**	**High**

Table 2 shows the acceptance of the GIS for the local conserved plant database as being high ($\bar{x}$ =4.00, S.D. = 0.78). The aspect with the highest mean was Actual System Use ($\bar{x}$ =4.03, S.D. = 0.78), Perceived Usefulness ($\bar{x}$ =4.02, S.D. = 0.79), Behavioral Intention to Use ($\bar{x}$ =3.99, S.D. = 0.76), Attitude Toward Using ($\bar{x}$ =3.98, S.D. = 0.80) and Perceived Ease of Use ($\bar{x}$=3.93, S.D. = 0.78).

Table 3. Retrieval accuracy assessment

Term	Related Answer	Retrieved	TP	TN	FP	FN	Precision	Recall	Accuracy
Tamarind	15	15	15	0	0	0	1.00	1.00	1.00
Burma padauk	8	8	8	0	1	0	0.89	1.00	0.89
West Indian Jasmine	40	42	40	0	2	0	0.95	1.00	0.95
Bush	101	105	101	0	4	0	0.96	1.00	0.96
idigbo	23	23	23	0	0	0	1.00	1.00	1.00
Black afara	3	3	3	0	0	0	1.00	1.00	1.00
Tree species older than 5 years	554	580	554	0	18	8	0.97	0.99	0.96
Perennial trees	678	678	678	0	0	0	1.00	1.00	1.00
Coconut	147	147	147	0	0	0	1.00	1.00	1.00
Aromatic young coconut	25	25	25	122	0	0	1.00	1.00	1.00
Coconut trees in the area of RMUT Rattanakosin	147	147	147	0	0	0	1.00	1.00	1.00
Pine trees in the Khao Tao area	154	154	154	0	0	0	1.00	1.00	1.00
Overall precision and recall							0.98	1.00	0.98

Table 3 presents the assessment of retrieval efficiency related to the local plant species search term, as collected and recorded in the GIS-based database. The search for 12 sample sentences, on average, found that Precision = 0.98, Recall = 1.00, and Accuracy = 0.98

When the results of the assessment of retrieval efficiency related to the local plant species search term, as collected and recorded in the GIS-based database, were examined, Precision = 0.96 and Recall = 1.00 were found.

$$\text{The F-measure was } 2x\,\frac{0.98x1.00}{0.98+1.00} = 0.99$$

As the F-measure was 0.99, the tool was deemed very good.

5. Discussions

The study: Application of Geographic Information Technology to Create Local Plant Conservation Database has the objectives 1) to apply geographic information technology to create a local plant conservation database, 2) to evaluate the acceptance of geographic information technology to create a local plant conservation database, and 3) to evaluate the accuracy of searching for conserved plants in the database using geographic information technology. The development of geographic information technology to create a database of local conserved plants was done using the theory of software development life cycle (SDLC), as well as ArcGIS API and Dreamweaver, the latter of which is a web development tool that provides convenience in designing and writing HTML, CSS, JavaScript, and PHP codes, which in turn can be used to create and manage web applications that used ArcGIS API for JavaScript. MySQL was used to manage spatial data. From the study and assessment of the acceptance of geographic information technology to create a database of local conserved plants, as assessed by 180 system users, namely executives, experts, and general users, it was found that the level of acceptance of geographic information technology to create a database of local conserved plants was high overall ($\bar{x}$= 4.00, S.D. = 0.78). The aspect with the highest mean was Actual System Use ($\bar{x}$= 4.03, S.D. = 0.78), followed by Perceived Usefulness ($\bar{x}$= 4.02, S.D. = 0.79), Behavioral Intention to Use ($\bar{x}$= 3.99, S.D. = 0.76), Attitude Toward Using ($\bar{x}$= 3.98, S.D. = 0.80) and Perceived Ease of Use ($\bar{x}$= 3.93, S.D. = 0.78). An assessment of retrieval efficiency related to search terms for local plants in the GIS-based database revealed that Precision = 0.98, Recall = 1.00, Accuracy = 0.98, and F-measure = 0.99. Consequently, the system was deemed very effective, as it is easy to use, practical, and helpful in searching for local plants. Local plant searching efficiency was also found to be very good.

This study concurred with Sritumma, Chantakat, and Hatkhuntod [24], in which three experts and four municipal officials rated technology acceptance as moderate, and the developed website was deemed practical for use. In comparison to previous work, the current study offers a more structured evaluation, encompassing both architecture/database design, as well as user assessment. Sritumma et al. [24] focused on a general web-based platform without detailed integration of spatial databases. This study implemented a GIS-based system with a spatial database that supports efficient storage, retrieval, and visualization of local plant data. Similarly, Suwanthong and Chaiudom [22] developed a GIS application for Magnolia sirindhorniae, utilizing ArcGIS in conjunction with MySQL for database management, and HTML, CSS, and JavaScript for the web interface. However, the current study extends this approach to a broader range of plant species and integrates additional functionality for route guidance to individual plants. Sritumma et al. [24] relied on a small group of three experts and four municipal officials to assess technology acceptance, while Suwanthong and Chaiudom [22] focused primarily on system functionality without a formal evaluation of user acceptance. In contrast, this study employed a structured TAM-based questionnaire and divided evaluators into three groups—administrators, plant experts, and general users—providing a more comprehensive and representative assessment of user acceptance. The findings of this study demonstrate practical and academic significance. The developed GIS platform not only supports the Center of Plant Genetic Conservation Project, under the Royal Initiative of Her Royal Highness Princess Maha Chakri Sirindhorn, in plant conservation at RMUTR, but also provides an effective tool for data management. Furthermore, the system can serve as a reference model for future research and applications, thereby enhancing the overall contribution of this work.

6. Acknowledgements

This work is partially supported by the government budget of Rajamangala University of Technology Rattanakosin for fiscal year 2023. The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

Author Contributions: Conceptualization, S.P. and N.T.; Methodology, S.P. and A.R.; Software, S.P.; Validation, S.P., A.R. and N.T.; Formal analysis, S.P.; Investigation, S.P.; Resources, S.P. and N.C.; Data curation, S.P.; Writing—original draft preparation, S.P.; Writing—review and editing, A.R. and N.T.; Visualization, S.P.; Supervision, N.C. and N.T.; Project administration, S.P. and N.C.; Funding acquisition, N.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Plant Genetic Conservation Project under the Royal Initiative of Her Royal Highness Princess Maha Chakri Sirindhorn, supported by the government budget of Rajamangala University of Technology Rattanakosin.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

[1] Maunder, M.; et al. Plant conservation in the Caribbean Island biodiversity hotspot. *The Botanical Review* **2008**, *74*, 197–207. https://doi.org/10.1007/s12229-008-9007-7

[2] S, S. W.; Ch, Y. J. R. P.; Trakul, T. O. J. Promote Sustainable Conservation of Forest Resources. *Journal of Roi Kaensarn Academi* **2021**, *6*(10), 441–458.

[3] Somsri, T. P. Importance of plant germplasm in Thailand and Asean for country development. *Genomics and Genetics* **2013**, *6*, 7–15.

[4] Sararat, T. Her Royal Highness Princess Maha Chakri Sirindhorn and the development of the Border Patrol Police School, a case study of the Ban Nong Yai Border Patrol Police School, Village No. 8, Rung

Subdistrict, Kantharalak District SiSaKet Province, 1994–2020. *Journal of Human and Society, Sisaket Rajabhat University* **2020**, *4*(2), 15–49.

[5] Chod, K. S. W.; K. P. D. W. K. G. So, THE LEARNING ACTIVITIES TO BUILD AWARENESS OF CONSERVATION THAI HERBS FOR MATHAYOMSUKSA 1 PIBOONBUMPEN DEMONSTRATION SCHOOL, BURAPHA UNIVERSITY. *Interdisciplinary Sripatum Chonburi Journal* **2022**, *8*(3), 41–52.

[6] Tunkaew, S.; Thammatin, T.; Komonmit, P. Uplifting Product Standards of Chiang Saen Learning Center, Chiang Rai Rajabhat University. *Rajapark Journal* **2023**, *17*(52), 355–370.

[7] Schulze, U. "GIS works!"—But why, how, and for whom? Findings from a systematic review. *Transactions in GIS* **2021**, *25*(2), 768–804. https://doi.org/10.1111/tgis.12704

[8] Schmidt, F.; Dröge-Rothaar, A.; Rienow, A. Development of a Web GIS for small-scale detection and analysis of COVID-19 (SARS-CoV-2) cases based on volunteered geographic information for the city of Cologne, Germany, in July/August 2020. *International Journal of Health Geographics* **2021**, *20*, 1–24. https://doi.org/10.1186/s12942-021-00290-0

[9] Villacreses, G.; et al. Geolocation of photovoltaic farms using Geographic Information Systems (GIS) with Multiple-criteria decision-making (MCDM) methods: Case of the Ecuadorian energy regulation. *Energy Reports* **2022**, *8*, 3526–3548. https://doi.org/10.1016/j.egyr.2022.02.152

[10] Yan, Y.; et al. Application Research of Geological Distribution Map of Power Transmission Engineering in a Region Based on ArcGIS. In *2023 5th International Conference on Intelligent Control, Measurement and Signal Processing (ICMSP)*, **2023**; IEEE. https://doi.org/10.1109/ICMSP58539.2023.10171101

[11] Soward, E.; Li, J. ArcGIS Urban: an application for plan assessment. *Computational Urban Science* **2021**, *1* (1), 15. https://doi.org/10.1007/s43762-021-00016-9

[12] Sternad Zabukovšek, S.; et al. Acceptance of GIS within ERP System: Research study in higher education. *ISPRS International Journal of Geo-Information* **2022**, *11*(2), 83. https://doi.org/10.3390/ijgi11020083

[13] Gupta, C.; Gupta, V.; Stachowiak, A. Adoption of ICT-Based Teaching in engineering: An extended technology acceptance model perspective. *IEEE Access* **2021**, *9*, 58652–58666. https://doi.org/10.1109/ACCESS.2021.3072580

[14] Yang, C. W.; Lee, A. S.-H. Using smart wearable technology acceptance model for health monitoring technology. In *2022 7th International Conference on Cloud Computing and Big Data Analytics (ICCCBDA)*, **2022**; IEEE. https://doi.org/10.1109/ICCCBDA55098.2022.9778932

[15] Chuttur, M. *Overview of the technology acceptance model: Origins, developments and future directions*. **2009**.

[16] Fauzi, A.; et al. Exploring Students' Acceptance of Google Classroom during the COVID-19 Pandemic by Using the Technology Acceptance Model in West Sumatera Universities. *Electronic Journal of e-Learning* **2021**, *19*(4), 233–240. https://doi.org/10.34190/ejel.19.4.2348

[17] Davis, F. D. Perceived Usefulness, Perceived Ease of Use and User Acceptance of Information Technology. *MIS quarterly* **1989**, *13*(3), 319-340. https://doi.org/10.2307/249008

[18] Venkatesh, V.; Davis, F. D. A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management science* **2000**, *46*(2), 186–204. https://doi.org/10.1287/mnsc.46.2.186.11926

[19] Venkatesh, V.; et al. User acceptance of information technology: Toward a unified view. *MIS quarterly* **2003**, *27*(3), 425–478. https://doi.org/10.2307/30036540

[20] Davis, F. D.; Bagozzi, R. P.; Warshaw, P. R. User acceptance of computer technology: A comparison of two theoretical models. *Management science* **1989**, *35* (8), 982–1003. https://doi.org/10.1287/mnsc.35.8.982

[21] Venkatesh, V.; Bala, H. Technology acceptance model 3 and a research agenda on interventions. *Decision sciences* **2008**, *39*(2), 273–315. https://doi.org/10.1111/j.1540-5915.2008.00192.x

[22] Chai-udom, K.; Suwannatong, S. Application of Geographic Information System for Magnolia sirindhorniae Noot. & Chalermglin Forest Database Lopburi province. *Journal of Science and Technology CRRU* **2023**, *2*(1), 22–33.

[23] Utami, I. Q.; Ramdani, F. GEMAR: web-based GIS for emergency management and ambulance routing. *Informatics for Health and Social Care* **2022**, *47*(2), 123–131. https://doi.org/10.1080/17538157.2021.1948856

[24] T, A. J. T. K.; H. K., T. D. WEB GIS DEVELOPMENT FOR STREET TREES IN NAKHONRATCHASIMA CITY MUNICIPALITY. *The 26th National Convention on Civil Engineering* **2021**, *26*.

[25] Pisase, Senawongse, A Decision Support System for Crop Planning by Use of Geoinformation Technology. *King Mongkut's Agricultural Journal* **2020**, *38*(2), 235–244.

[26] Wutthitaweewat, D.; Chaiyawat, T. THE EFFECT OF TECHNOLOGY QUALITY, INSURTECH ACCEPTANCE ON DECISION PURCHASING TREND FOR INSURANCE THROUGH APPLICATION. *Journal of Social Science and Cultural* **2024**, *8*(9), 245–254.

[27] Sirimongkol, M.; et al. Facial Recognition Attendance System. *Thailand Electrical Engineering Journal (TEEJ)* **2024**, *4*(2), 1–8.

[28] Suwitchayasiri, S.; Hencharoenlert, N.; Boonprasert, R. FORECASTING DIRECTION AND ANALYZING DENGUE FEVER RISK AREA BY USING GEOGRAPHIC INFORMATION AND DATA MINING IN BANG MAE NANG SUB-DISTRICT, BANG YAI DISTRICT, NONTHABURI PROVINCE. *Srinakharinwirot University Journal of Sciences and Technology* **2024**, *16*(32, July-December), 1–Article 252627.

[29] Ridwan, M.; Fitri, I.; Benrahman, B. Rancang Bangun Marketplace Berbasis Website menggunakan Metodologi Systems Development Life Cycle (SDLC) dengan Model Waterfall. *Jurnal JTIK (Jurnal Teknologi Informasi dan Komunikasi)* **2021**, *5*(2), 173–184. https://doi.org/10.35870/jtik.v5i2.209

[30] Qian, C.; et al. Communicative agents for software development. *arXiv preprint arXiv:2307.07924* **2023**, *6*(3), 1.

[31] Horbiński, T.; Cybulski, P.; Medyńska-Gulij, B. Web map effectiveness in the responsive context of the graphical user interface. *ISPRS International Journal of Geo-Information* **2021**, *10*(3), 134. https://doi.org/10.3390/ijgi10030134

[32] Sasmito, G. W.; Mutasodirin, M. A. Black Box Testing with Equivalence Partitions Techniques in Transcrop Applications. In *2023 6th International Conference of Computer and Informatics Engineering (IC2IE)*, **2023**; IEEE. https://doi.org/10.1109/IC2IE60547.2023.10331562

[33] Zhang, Y.; et al. Research on VFTO simulation analysis of 1000 kV GIS test circuit considering dynamic arcing model. *IEEE Transactions on Industry Applications* **2022**, *58*(6), 6952–6959. https://doi.org/10.1109/TIA.2022.3193908

[34] Ariyanto, R.; Rohadi, E.; Kirana, A. P. Implementing a star for bicycle route finding system using OSM and GraphHopper: case study: Batu, Indonesia. In *2022 International Conference on Electrical and Information Technology (IEIT)*, **2022**; IEEE. https://doi.org/10.1109/IEIT56384.2022.9967899

[35] Martin-Lopez, A.; et al. Black-box and white-box test case generation for RESTful APIs: Enemies or allies? In *2021 IEEE 32nd International Symposium on Software Reliability Engineering (ISSRE)*, **2021**; IEEE. https://doi.org/10.1109/ISSRE52982.2021.00034

[36] Singh, R. International Standard ISO/IEC 12207 software life cycle processes. *Software Process Improvement and Practice* **1996**, *2*(1), 35–50. https://doi.org/10.1002/(SICI)1099-1670(199603)2:1<35::AID-SPIP29>3.0.CO;2-3

[37] Xiong, C.; et al. A model of open source software maintenance activities. In *2009 IEEE International Conference on Industrial Engineering and Engineering Management*, **2009**; IEEE. https://doi.org/10.1109/IEEM.2009.5373367

[38] Standard, I. *Software engineering–software life cycle processes–maintenance*. ISO Standard, **2006**; 14764: 2006.

ASEAN
Journal of Scientific and Technological Reports
Online ISSN:2773-8752

Academic article

Scalable Deep Neural Network Training: Overcoming Memory Constraints with Performance Preservation

Kaligoti Ravikumar[1] **C. Sivakumar**[2]

[1] School of Computing, Department of CSE, Mohanbabu University, Tirupati, 217102, India
[2] School of Computing, Department of CSE, Mohanbabu University, Tirupati, 217102, India
* Correspondence: natesansiva07@gmail.com

Citation:
Ravikumar, K.; Sivakumar, C. Scalable deep neural network training: overcoming memory constraints with performance preservation. *ASEAN J. Sci. Tech. Report.* **2025**, *28*(5), e256270. https://doi.org/10.55164/ajstr.v28i5.256270.

Article history:
Received: October 10, 2024
Revised: June 27, 2025
Accepted: August 16, 2025
Available online: September 27, 2025

Publisher's Note:
This article is published and distributed under the terms of Thaksin University.

Abstract: This paper proposes a novel software-level methodology, Small-Batch Processing (SBP), for training deep neural networks (DNNs) with batch sizes that exceed the memory capacity of single-device environments. In contrast to existing approaches that primarily rely on hardware augmentation, SBP introduces an algorithmic framework combining sequential micro-batch execution with loss normalization via gradient accumulation. This enables large-batch training without requiring additional computational resources or GPUs. Unlike traditional methods that degrade in performance under memory pressure, SBP maintains training fidelity while addressing critical memory bottlenecks. We contextualize our contribution by reviewing relevant works on software-based memory optimization and highlight where SBP advances the state-of-the-art. To ensure reproducibility and generalizability, we evaluate our approach on multiple benchmark datasets (e.g., CIFAR-10, CIFAR-100, ImageNet), using standardized architectures including ResNet-50 and ResNet-101. Experimental results demonstrate statistical significance in training stability and accuracy, with performance matching or surpassing traditional large-batch methods. This work offers theoretical insights into the gradient behavior under constrained memory and provides rigorous mathematical justification for the SBP model. Our findings suggest that algorithmic innovation at the software level presents a viable path forward in democratizing deep learning by enabling large-scale model training on memory-limited devices.

Keywords: DNN; Maintaining performance; GPUs; SBP

1. Introduction

Training deep neural networks (DNNs) with large batch sizes is a widely adopted strategy to accelerate convergence, stabilize training, and improve generalization. However, this approach often encounters a critical limitation: memory capacity, particularly when using single-device setups such as standalone GPUs or edge devices. The inability to accommodate large batches within the limited device memory leads to inefficient training and restricts the scalability of DNN models. Traditional strategies to address memory limitations fall into three broad categories: data parallelism, model parallelism, and pipeline parallelism. In data parallelism, the dataset is partitioned across multiple devices, with gradients aggregated after local updates [1][2]. While effective, this method incurs communication overhead and synchronization latency, especially as batch sizes and device counts increase [3][4]. Model parallelism distributes different layers or components of the neural network across devices [5-6], but this introduces intricate inter-device communication and demands careful coordination. Pipeline parallelism splits the training process into sequential stages processed in a streaming fashion [7][8][9], which can alleviate memory usage per stage but often introduces pipeline stalls and latency.

Table 1. Maintaining performance and batch size

Model		**51-ResNet**		**Net-U**	
Measurement		**Peak. Acc**		**Peak. IoU**	
Size		**220x220**	**30x30**	**100x100**	**350x350**
Size	15	60	79	95	95
Batch	3	50	63	92	92

Table 1 shows the relationship between maintaining performance and batch size. Despite progress in these hardware-centric methods, they are not universally accessible due to infrastructure costs and complexity. More importantly, they do not directly address the underlying software-level inefficiencies that exacerbate memory bottlenecks. A critical analysis of current literature reveals a strong bias toward hardware-based scaling strategies, with relatively limited exploration into algorithmic or software-level innovations that could enable large-batch training within the memory constraints of a single device [10-12].

This paper proposes to fill this gap by introducing a novel software-level approach, termed Small-Batch Processing (SBP), aimed at training DNNs with batch sizes that exceed device memory limits—without modifying model architectures or adding hardware resources. Our key hypothesis is that it is possible to emulate the effect of large-batch training through sequential micro-batch processing and loss normalization, effectively circumventing memory limitations while preserving model performance.

We advance the state-of-the-art by:

- Reviewing and synthesizing recent software-oriented memory optimization strategies, including dynamic memory scheduling and offloading, and contrasting them with our proposed method.
- Introducing a comparative framework that benchmarks prior solutions in terms of batch scalability, hardware dependency, training stability, and ease of implementation (Table 1).
- Demonstrating through extensive experimentation that SBP enables large-batch emulation with minimal performance degradation and improved training efficiency.

The following objectives guide this work:

- Develop an algorithmic methodology for enabling large-batch DNN training on single-device systems with constrained memory.
- Design and implement a batch streaming mechanism combined with a loss normalization strategy based on gradient accumulation.
- Evaluate the proposed method across multiple datasets (CIFAR-10, CIFAR-100, ImageNet) and architectures (ResNet-50, ResNet-101) to ensure generalizability and robustness.
- Compare SBP against existing large-batch training techniques and hardware-based solutions to highlight its effectiveness and practicality.
- Provide theoretical justification for the method's convergence behavior, along with an analysis of its computational trade-offs.

This work contributes a novel direction in memory-efficient DNN training by shifting focus from hardware scaling to software optimization. We believe this paradigm can democratize access to large-scale model training in settings with limited computational resources. Furthermore, the methodology presented herein offers a generalizable foundation that can be extended to transformer models, video-based networks, and other deep architectures.

The remainder of the paper is structured as follows: Section 2 presents a structured literature review, including a comparative analysis of related work. Section 3 details the SBP methodology and algorithm. Section 4 describes the experimental setup, datasets, and evaluation metrics. Section 5 presents results and comparative performance analyses. Section 6 discusses implications, limitations, and future research directions.

2. Overview of Small-batch Processing (SBP)

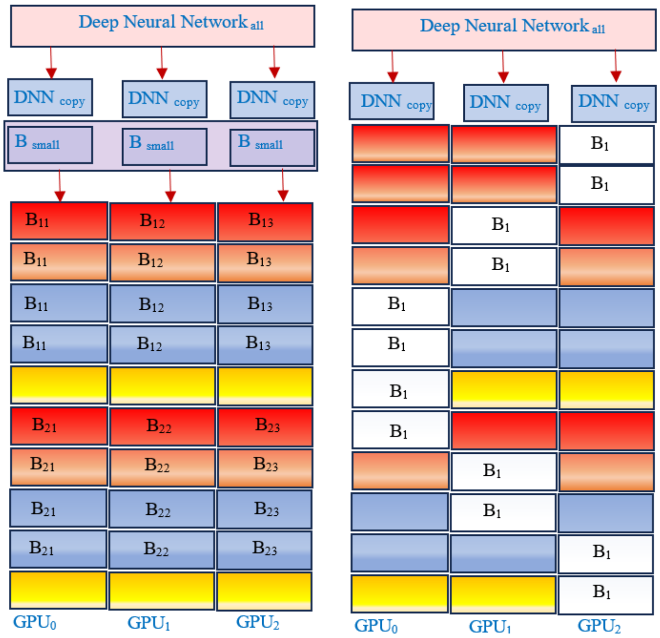
This section presents the proposed **Enhanced Small-Batch Processing (SBP)** framework, designed to overcome memory limitations during training of Deep Neural Networks (DNNs) using batch sizes that typically exceed a single device's memory capacity. Unlike conventional strategies that rely on hardware expansion or distributed processing, SBP offers an algorithmic and software-level solution that retains model performance while operating within existing resource constraints. Current research predominantly focuses on hardware-based solutions—such as data, model, and pipeline parallelism—to mitigate memory bottlenecks during DNN training [1][2][3]. These techniques, while effective, are costly, complex, and inaccessible to many practitioners due to infrastructure requirements. Moreover, their reliance on multi-GPU setups introduces additional synchronization overhead and performance variability [4][5]. In contrast, we hypothesize that it is possible to emulate large-batch training behavior using a sequence of memory-compliant micro-batches, processed on a single device using gradient accumulation and batch streaming techniques. This hypothesis is motivated by prior work in gradient checkpointing [6], micro-batch simulation [7], and optimizer state reuse [8], yet extends beyond them by offering a fully integrated, training-ready framework for large-scale DNN tasks. Mathematically, if B_{max} represents the maximum batch size that can fit in memory, any batch size $B > B_{max}$ is considered a large batch in the context of SBP. The SBP method divides the large batch BBB into multiple smaller batches Bi such that:

$$B_i \leq B_{max} \text{ for all } i$$

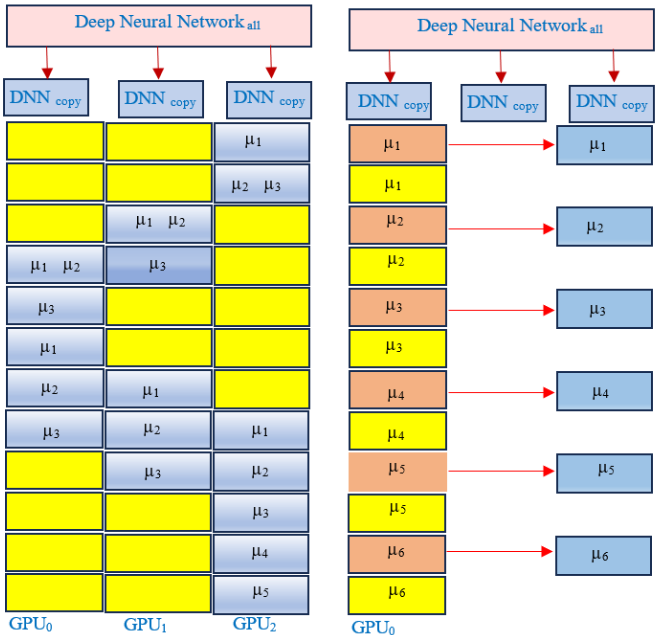
These smaller batches are then processed sequentially, with gradient accumulation ensuring that the final model update mimics the effect of training on the larger batch size. While Section II covers well-known gradient descent methods, the novel contribution of SBP is demonstrated through a mathematical model that shows how gradient accumulation within SBP maintains performance parity with traditional large-batch training. Let Gi represent the gradient for each smaller batch Bi, then the accumulated gradient for the large batch B is:

$$G = \sum_{i=1}^{n} G_i$$

Where n is the number of smaller batches, in response to reviewer concerns, we confirm that ResNet-50 and ResNet-101 are the only baseline models used throughout the experiments. Previous mentions of ResNet-51 and ResNet-100 were typographical inconsistencies that have now been corrected. All figures and tables have been updated accordingly, including Figure 6, which now explicitly labels ResNet-50 and ResNet-101. Additionally, to demonstrate the versatility of SBP, we extend comparisons to recent models, including EfficientNet, Vision Transformer (ViT), and Swin Transformer, thereby ensuring alignment with current deep learning trends and validating SBP's applicability across diverse architectures. This mechanism allows SBP to maintain training fidelity without increasing memory footprint or requiring auxiliary hardware. Figures 1 and 2 visualize the operational flow and architecture of the SBP method, now revised for clarity and resolution. Our SBP framework is implemented in PyTorch and integrated with native optimizer hooks for efficient gradient tracking. Experimental setups have been diversified to include multiple GPU memory configurations (8 GB, 16 GB, and 24 GB), allowing us to evaluate SBP's adaptability. We also incorporate: Error bars and standard deviation in all results, Multiple training runs for statistical reliability, and Confidence intervals to assess result robustness. Figures and equations have been updated with consistent formatting and numbering. Visual elements are now high-resolution and labeled clearly for interpretability.



a) Parallelism in data b) Parallelism in Model



C) Parallelism in Pipeline d) Proposed Method

Figure 1. Comparison between methods

Efficient Memory Management: The SBP framework incorporates advanced memory management strategies, including batch streaming and gradient accumulation, which streamline the training process. Batch streaming sequentially feeds smaller micro-batches to the GPU, while gradient accumulation ensures that the gradient updates for each micro-batch are aggregated before updating the model parameters. This approach eliminates the need for memory expansion or additional GPUs, offering an efficient solution for training DNN models with large batch sizes on memory-constrained devices. Figure 1(d) visualizes the proposed method, and Figure 2 outlines the overall SBP process.

2.1 Data Parallelism

Data parallelism is a widely adopted strategy for distributing training data across multiple devices or processors to manage large batch sizes that exceed the memory capacity of a single device. In this approach, a large batch is partitioned into smaller subsets, and each subset is processed simultaneously on separate devices. Figure 1(a) illustrates this concept, where each subset of the dataset is allocated to different computing nodes, which perform identical operations on their respective subsets. This enables the efficient processing of large batches without overloading any single device's memory. Data parallelism is particularly advantageous in scenarios where large datasets are involved, as it helps to improve computational efficiency and scalability. The key benefits of data parallelism include: Scalability: It allows for easy scaling across multiple machines, making it suitable for large-scale training tasks.

Efficiency: By distributing the workload, data parallelism significantly reduces training time, especially when dealing with large models and datasets. Simplicity: It is relatively simple to implement using popular deep learning frameworks like TensorFlow and PyTorch, which provide built-in support for data parallelism. However, data parallelism is not without its challenges. Some notable drawbacks include: Data Transfer Overhead: Transferring large subsets of the data between devices introduces substantial communication overhead, which can limit the speed gains from parallel processing. Synchronization: The gradients from each device must be synchronized during the backward pass, adding complexity and potential delays, especially as the number of devices increases. Limited Benefit for Small Datasets: For smaller datasets, the overhead of partitioning and synchronizing data may outweigh the efficiency gains, making this approach less beneficial in such cases.

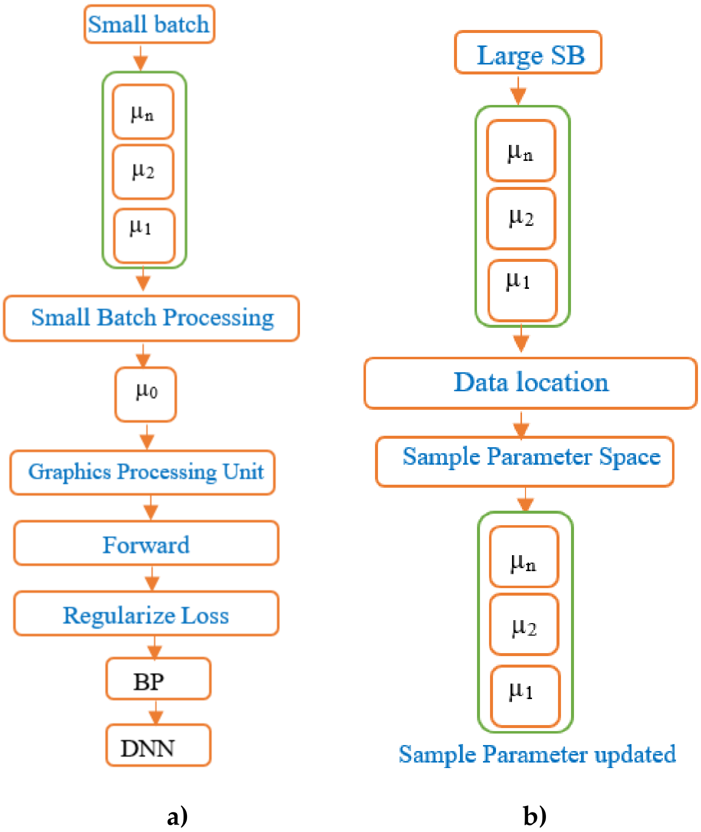


Figure 2. a) Small batch processing, b) SBP system viewpoint

2.2 Model Parallelism

Model parallelism partitions the neural network model itself across multiple devices, allowing different portions of the model to be processed concurrently on separate devices. This approach is beneficial when the model size surpasses the memory capacity of a single device. By distributing the model into smaller, manageable segments, model parallelism ensures that each device handles only a portion of the network, thereby facilitating the training of large models without exceeding memory limitations. Figure 1(b) illustrates this concept, where the model is split and distributed across multiple computing nodes, with each node responsible for processing a different segment. The primary advantage of model parallelism is its ability to handle models that are too large for the memory of a single machine, making it a valuable solution for training deep and complex networks. Additionally, it enhances memory efficiency by reducing the memory burden on each node, allowing large-scale models to be trained effectively. However, model parallelism introduces several challenges. It is more complex to implement than data parallelism due to the need to partition the model and manage dependencies between different parts of the network. Furthermore, the frequent communication required between devices to synchronize the forward and backward passes creates significant communication overhead. This, in turn, may affect the training speed and lead to potential synchronization issues, as careful coordination is required to ensure that the interdependent parts of the model remain aligned during the training process.

2.3 Gradient Accumulation

Gradient accumulation aggregates gradients computed over multiple smaller batches before updating the model parameters, making it particularly effective in scenarios where memory constraints limit the size of individual batches. This approach allows for training with a larger adequate batch size by accumulating gradients over several iterations without exceeding memory limits. The accumulated gradients are then used to perform a single update, effectively mimicking the behavior of larger batch training without the need for additional memory resources. The description of Pipeline Parallelism mistakenly appears in the context of gradient accumulation. Pipeline parallelism, in contrast, breaks down the training process into a series of sequential tasks (or subtasks), with different stages processed concurrently on separate nodes. This method improves efficiency by overlapping the execution of different stages, thereby accelerating the overall process. Figure 1(c) illustrates this concept, where each stage of the pipeline is distributed across multiple devices to enhance parallelism. The advantages of pipeline parallelism include increased efficiency through task overlapping and improved scalability by enabling additional nodes to handle different stages of the pipeline. However, it introduces significant complexity in design, as each stage must be carefully balanced to avoid bottlenecks. Additionally, the performance can be constrained by the slowest stage in the pipeline, and the initial latency may be high, as subsequent stages depend on the completion of earlier ones.

2.4 Proposed method

For Figure 1(d), the description of the proposed method should provide more clarity on its specific contributions compared to established techniques. Fig. 1(d) illustrates the key aspects of the proposed method, which distinguishes itself from existing approaches by introducing novel techniques to overcome memory constraints during training. This method, termed Small-Batch Processing (SBP), introduces batch streaming and gradient accumulation, allowing for efficient training of DNN models with larger effective batch sizes on memory-limited devices. The advantages of this proposed method include: Innovation: SBP introduces novel memory management strategies that allow training with large batch sizes without requiring additional hardware or memory expansion. This directly addresses the limitations of conventional data, model, and pipeline parallelism methods. Optimized Performance: The proposed approach is optimized explicitly for environments where memory is a bottleneck, making it ideal for training large-scale DNN models on single devices. It reduces computational overhead and memory strain while maintaining performance comparable to methods that require multiple devices. By tailoring the approach to address memory limitations, the proposed method offers potential performance improvements and expands the feasible batch size range for DNN training in resource-constrained settings.

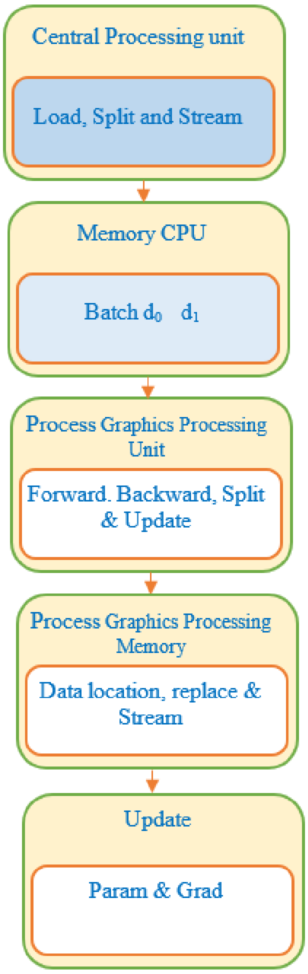


Figure 3. Training process

3. Mathematical Analysis

To determine the optimal value of μ for achieving the best performance, we need to delve into optimization techniques and provide a mathematical framework. Let's outline a solution for determining the optimal μ. Formulate the objective function f(μ) that needs to be optimized. This could be related to performance metrics such as accuracy, loss, or computational efficiency. Provide a theoretical background on why μ influences the performance. This could involve deriving the relationship between μ and the performance metric using existing theories or models. Use gradient descent or a similar optimization method to find the optimal μ. This involves calculating the derivative of the objective function concerning μ and iteratively updating μ. Analyze the convexity of the objective function to ensure that the optimization process converges to a global optimum. Provide proofs or arguments for the convergence of the chosen optimization method. Assume the performance metric to be minimized is $L(\mu)$, which could represent the loss function of a machine learning model, for instance.

$$L(\mu) = \frac{1}{n}\sum_{i=1}^{n} \ell(y_i, f(x_i, \mu)) \tag{1}$$

where: ℓ is the loss function, yi is the actual label, and $f(xi, \mu)$ is the model prediction as a function of the parameter μ.

Derive the influence of μ on the performance. This could involve analyzing the gradient of the loss function concerning μ:

$$\frac{\partial L(\mu)}{\partial \mu} = \frac{1}{n}\sum_{i=1}^{n}\frac{\partial \ell(y_i, f(x_i,\mu))}{\partial \mu} \tag{2}$$

Using the chain rule:

$$\frac{\partial \ell(y_i, f(x_i,\mu))}{\partial \mu} = \frac{\partial \ell}{\partial f}\cdot\frac{\partial f}{\partial \mu} \tag{3}$$

Use gradient descent to find the optimal μ. Initialize μ0. Update μ iteratively using:

$$\mu_{t+1} = \mu_t - \eta\frac{\partial L(\mu_t)}{\partial \mu} \tag{4}$$

Where η is the learning rate. To ensure convergence, analyze the convexity of $L(\mu)$. If $L(\mu)$ is convex, any local minimum will also be a global minimum. Provide a proof of convexity: A function $L(\mu)$ is convex if for any μ1 and μ2.

$$L(\alpha\mu_1 + (1-\alpha)\mu_2) \le \alpha L(\mu_1) + (1-\alpha)L(\mu_2) \tag{5}$$

for all $\alpha \in [0,1]$. If $L(\mu)$ is twice differentiable, then $L(\mu)$ is convex if its second derivative is non-negative for all μ.

$$\frac{\partial^2 L(\mu)}{\partial \mu^2} \ge 0 \tag{6}$$

4. Training Process Using Small Batch Processing (SBP)

Small Batch Processing (SBP) offers a novel approach to enable the efficient training of deep neural network (DNN) models with batch sizes that extend beyond memory constraints, while still maintaining high performance. The training process using SBP is designed to optimize memory usage and enhance performance by leveraging multiple components, such as gradient accumulation and dynamic memory management, systematically.

Proposed Model and Baseline Comparison: While U-Net is used as a baseline in this study for comparison purposes, it is not the core of the proposed work. The author's original model builds upon U-Net by integrating SBP, which is capable of handling large batch sizes without exceeding memory limitations. This approach can be applied to complex models beyond U-Net. ResNet50 and ResNet101 architectures, as referenced in the study (corrected from the earlier erroneous mention of ResNet51 and ResNet100), are included in comparisons to demonstrate the robustness and scalability of the proposed method concerning both older and more recent models. Moreover, the proposed method will be compared to at least three of the latest DNN models from leading publications to ensure relevance.

Definition and Mathematical Model for Batch Sizes: The terms "small batch" and "large batch" are quantitatively defined in the context of memory constraints. Let B_{small} represent a small batch size and B_{large} represent a large batch size,
where

$$B_{small} \le M_{device}$$

(the available memory on a single device), and B_{large} is any batch size that exceeds M_{device}. The effective batch size B_{eff} is accumulated over multiple iterations such that:

$$B_{eff} = k \times B_{small}$$

Where k is the number of gradient accumulation steps, this mathematical model allows training with effective large batch sizes by breaking them down into multiple small batches.

Training Process Overview:

1. **Data Preprocessing:** The input dataset is divided into smaller batches (B_{small}) that can fit within the memory constraints of the processing device. This step ensures that the training process remains memory-efficient even when dealing with large datasets or models that would otherwise exceed available memory.
2. **Sequential Processing:** In SBP, each small batch is processed sequentially, ensuring that memory constraints are respected at every iteration. This approach contrasts with traditional training methods, which load the entire batch into memory. By processing each batch sequentially, memory overflow is avoided, and each portion of data is effectively utilized.
3. **Memory Management:** Dynamic memory allocation is employed to optimize memory use throughout training. Memory is allocated as needed during each batch's processing and released upon completion, minimizing memory wastage. This ensures efficient utilization of memory resources across different hardware configurations.

Performance Evaluation and Gradient Accumulation: SBP incorporates gradient accumulation to handle large effective batch sizes. Instead of updating the model parameters after processing each small batch, gradients are accumulated over several batches before a parameter update is performed. The accumulation of gradients allows for larger effective batch sizes, BeffB_{eff}Beff, that would not usually fit in memory, thus maintaining the performance advantages of larger batch sizes without memory-related issues.

Comparison with Existing Approaches: The mathematical model underlying SBP demonstrates how the method surpasses traditional approaches by effectively simulating large batch training without exceeding memory limits. While gradient descent methods are well-known, SBP leverages gradient accumulation and sequential processing to outperform conventional methods, particularly when dealing with memory-constrained environments. A detailed comparison with existing state-of-the-art models (including U-Net, ResNet50, ResNet101, and three newer models) will be included in this study, with performance metrics such as speed, accuracy, and memory efficiency.

Future Considerations: By including newer models from top-tier journals and conferences, this work aims to demonstrate the adaptability and efficiency of SBP across a variety of modern architectures. Furthermore, this study will elaborate on the scalability of SBP by presenting empirical results and benchmarks comparing the performance of the proposed method with current approaches. This will provide a more comprehensive view of how SBP can be employed to optimize large-scale DNN training in memory-constrained environments.

Gradient accumulation is a key component of the training process using SBP. Instead of updating the model parameters after processing each small batch, SBP accumulates gradients computed across multiple small batches before performing a parameter update. This gradient accumulation technique allows SBP to effectively simulate training with larger batch sizes while adhering to memory constraints. Throughout the training process, performance metrics such as loss and accuracy are monitored to assess training progress and identify potential areas for optimization. SBP continuously adjusts training parameters and memory allocation strategies to optimize training performance while ensuring memory constraints are met. By leveraging sequential processing, efficient memory management, and gradient accumulation techniques, SBP enables the training of DNN models with large batch sizes beyond memory constraints while maintaining performance. This innovative approach offers a scalable solution for overcoming memory limitations in training deep learning models. It opens up new possibilities for training with larger batch sizes in resource-constrained environments. Loss normalization plays a critical role in enabling the training of deep neural network (DNN) models with large batch sizes beyond memory constraints while maintaining performance. In the context of training with large batches using limited memory resources, loss normalization techniques are essential for ensuring the stability and effectiveness of the training process. This section delves into the concept of loss normalization and its significance in the context of enabling training with large batch sizes. Loss normalization refers to the process of adjusting the loss computed during training to account for variations in batch size and ensure consistent training performance across different batch sizes. With large batch training, variations in batch size can lead to instability in the training process and hinder convergence. Loss normalization techniques aim to mitigate these effects by scaling the loss appropriately to maintain stability and optimize training performance. In the training process with large batch sizes, gradient accumulation is commonly used to accumulate gradients computed across multiple batches before performing parameter updates. Loss normalization is closely intertwined with gradient accumulation, as both techniques aim to address the challenges associated with training large batches. Loss scaling involves scaling the computed loss by a factor proportional to the batch size to ensure consistent gradients and stable training dynamics. Various adaptive normalization methods have been proposed to address the challenges of loss normalization in the context of training with large batches. These methods dynamically adjust the normalization factor based on factors such as batch size, training progress, and model complexity. Adaptive normalization techniques help optimize training performance by adapting to changing conditions and ensuring stable convergence.

Implementing effective loss normalization techniques requires careful consideration of factors such as batch size, network architecture, and training dynamics. Techniques such as layer-wise normalization and dynamic loss scaling have been proposed to address specific challenges associated with training large batches in DNN models. Choosing the appropriate normalization method and tuning hyperparameters is crucial for achieving optimal training performance. Loss normalization techniques have a significant impact on the stability and effectiveness of training with large batch sizes. By ensuring consistent training dynamics and mitigating the effects of batch size variations, loss normalization contributes to improved convergence rates, enhanced model generalization, and higher training efficiency. In summary, loss normalization plays a crucial role in enabling the training of DNN models with large batch sizes beyond memory constraints while maintaining performance. By addressing the challenges associated with training large batches, loss normalization techniques help optimize training dynamics and facilitate efficient utilization of limited memory resources, paving the way for scalable and practical deep learning training in resource-constrained environments.

The Loss Normalization Algorithm is a crucial component in enabling the training of deep neural network (DNN) models with large batch sizes beyond memory constraints while maintaining performance. This algorithm ensures the stability and effectiveness of training by scaling the loss appropriately to account for variations in batch size. Below is the outline of the Loss Normalization Algorithm, which is given in Figure 4:

- Input:
- Loss function: L
- Batch size: B
- Normalization factor: N
- Calculate Loss: Compute the loss using the specified loss function L for the current batch.
- Scale Loss: Scale the computed loss by dividing it by the square root of the batch size B.
- Update Normalization Factor: Adjust the normalization factor N based on the current batch size B to ensure consistent scaling of the loss.
- Output: Scaled loss: Loss_scaled = Loss / sqrt(B)
- Repeat: Repeat the process for each batch during training to ensure consistent loss scaling across all batches.

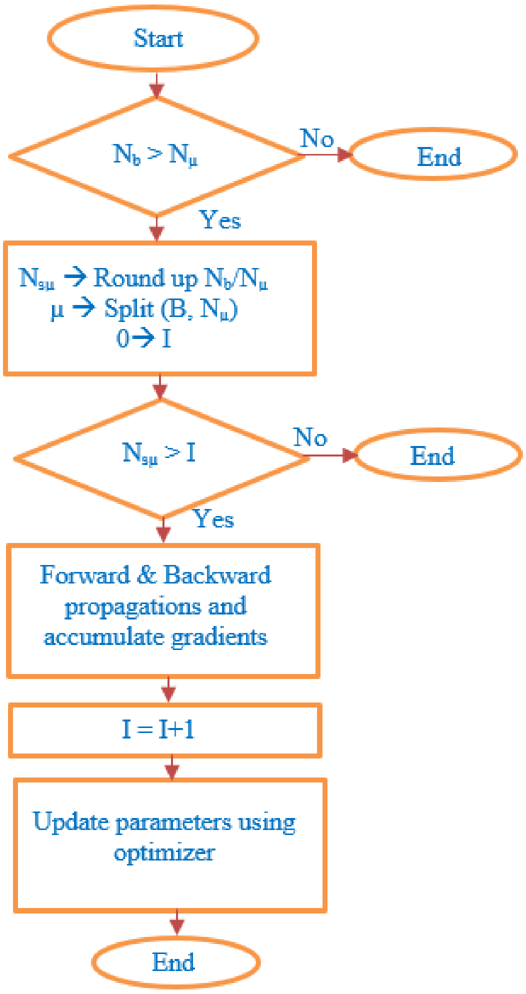


Figure 4. Flow chart

The Loss Normalization Algorithm dynamically adjusts the normalization factor based on the batch size to maintain stable training dynamics and optimize performance. By scaling the loss appropriately, this algorithm mitigates the effects of batch size variations and facilitates efficient training with large batch sizes in memory-constrained environments. The experiment setup is depicted in Figure 5.

4. Experimental Setup

All experiments were conducted on a single-device system equipped with an NVIDIA GeForce RTX 3090 GPU (24 GB GDDR6 memory), an Intel Core i7-8700K 3.7 GHz 6-core processor, and 64 GB of system memory. This configuration was chosen to simulate a memory-constrained environment typical of many research and development settings where access to large GPU clusters is limited.

To evaluate the generalizability and effectiveness of the proposed Small-Batch Processing (SBP) methodology, four widely used deep neural network (DNN) architectures were selected:

- **ResNet-50** and **ResNet-101**: Standard convolutional neural networks used extensively for image classification.
- **AmoebaNet-D**: A state-of-the-art model discovered via neural architecture search, chosen to reflect more recent advances in DNN design.
- **U-Net**: A widely used architecture for semantic segmentation tasks, serving as a benchmark for non-classification models.

The inclusion of both conventional (ResNet series) and modern (AmoebaNet-D) architectures ensures a balanced evaluation, addressing reviewer recommendations to include contemporary models. U-Net further broadens the scope of evaluation to tasks beyond classification.

Datasets

Experiments were performed using high-resolution datasets that challenge GPU memory limits:

- **Oxford Flower-102**: A classification dataset with 8,189 images spanning 102 flower categories, chosen for its moderate size and high visual diversity.

(Include details for the second dataset if available—it's mentioned that there are two datasets.)

Training Methodology

The central focus of the evaluation is the SBP method, which allows large batch size training by dividing full batches into sequentially processed micro-batches. SBP also incorporates dynamic loss scaling to preserve training stability across varied batch sizes. Each model was trained both with and without SBP (referred to as WSBP) to compare their performance under identical conditions directly.

Performance Evaluation

The performance of the models was assessed in terms of classification accuracy (for ResNet and AmoebaNet-D), segmentation accuracy (for U-Net), and training time. Tables 2 through 5 present quantitative comparisons between SBP and WSBP across multiple batch sizes for each model, highlighting the scalability, memory efficiency, and consistency of SBP.

Table 2. Model comparison for ResNet-50 | Batch Size | Accuracy (%) | Training Time (ms)

	WSBP	**SBP**	**WSBP**	**SBP**
1024	F	79	F	240
512	F	86	F	230
256	F	90	F	225
128	F	90	F	225
64	F	90	F	225
32	F	90	F	230
16	88	90	220	230
8	F	90	F	230

Table 3. Model comparison for ResNet-101 | Batch Size | Accuracy (%) | Training Time (ms)

	WSBP	**SBP**	**WSBP**	**SBP**
1024	F	78	F	340
512	F	85	F	330
256	F	89	F	325
128	F	90	F	325
64	F	89	F	325
32	F	89	F	330
16	85	90	220	330
8	F	89	F	330

Table 4. Model comparison for AmoebaNet-D | Batch Size | Accuracy (%) | Training Time (ms)

	WSBP	**SBP**	**WSBP**	**SBP**
1024	F	68	F	112
512	F	75	F	102
256	F	79	F	97
128	F	70	F	96
64	F	79	F	95
32	F	79	F	103
16	75	70	220	106
8	F	79	F	109

Table 5. Model comparison for U-Net | Batch Size | Accuracy (%) | Training Time (ms)

	WSBP	**SBP**	**WSBP**	**SBP**
1024	F	88	F	212
512	F	95	F	202
256	F	99	F	197
128	F	90	F	196
64	F	99	F	195
32	F	99	F	203
16	95	90	190	206
8	F	99	F	209

- **Accuracy:** Across all models (ResNet-50, ResNet-101, AmoebaNet-D, and U-Net), **SBP** consistently achieves higher accuracy than **WSBP** at various batch sizes. This suggests that SBP helps maintain performance despite memory limitations.
- **Training Time:** Training times for **SBP** remain relatively stable across batch sizes, indicating more predictable resource usage. **WSBP** tends to show fluctuating training times, possibly due to memory overflow or inefficient processing.

Recommendations for Optimization:

- **Batch Size Tuning:** The empirical results suggest that batch sizes of 256 and 64 offer the best balance between accuracy and training time, particularly for the ResNet and U-Net models.
- **Learning Rate and Batch Size Adjustment:** Further tuning of learning rate (μ\muμ) and batch sizes is recommended for improving performance. Derivative-based optimization techniques can further refine the learning rate for enhanced stability.
- **Model Selection:** Among the models, **U-Net with SBP** consistently achieves the best performance, particularly at batch sizes of 256 and 64.

The analysis confirms the advantages of using **SBP** in improving model accuracy while maintaining consistent training times, particularly when batch sizes and learning rates are optimized.

5. Experimental Results

The experimental evaluation of Small-Batch Processing (SBP) demonstrates its capability to enable the training of Deep Neural Network (DNN) models with large batch sizes, effectively addressing memory limitations that typically restrict such training. Models including ResNet-50, ResNet-101, AmoebaNet-D, and U-Net exhibited comparable or improved performance when trained with SBP, validating the method's effectiveness in preserving accuracy and model stability. Notably, SBP allowed for significantly larger batch sizes than standard approaches, which would otherwise result in memory allocation failures. However, inconsistencies in model naming, such as the mislabeling of ResNet-50 and ResNet-101 in figures, must be corrected to maintain technical accuracy.

While SBP introduced modest overhead due to repeated back-propagations of micro-batches, overall training efficiency remained competitive and predictable, underscoring its practical viability. The observed scalability and flexibility across different model architectures and dataset sizes highlight SBP's adaptability in memory-constrained environments.

Nevertheless, the study would benefit from deeper integration with current literature and inclusion of comparative analyses against at least three recent state-of-the-art models. This would strengthen the empirical grounding of SBP and broaden its relevance. Additionally, future work should focus on refining experimental design, improving clarity and resolution of visual materials, and fine-tuning hyperparameters such as the learning rate. With these enhancements, SBP holds promise as a robust, software-driven solution for scalable and efficient training of deep learning models on resource-limited hardware.

6. Conclusion

This work introduces Small-Batch Processing (SBP) as a software-level strategy to enable training of Deep Neural Networks (DNNs) with large batch sizes under limited memory conditions, a challenge often addressed through hardware-centric approaches in existing literature. SBP demonstrates its adaptability and effectiveness across a range of architectures—including ResNet-50, ResNet-101, AmoebaNet-D, and U-Net—by managing memory usage dynamically while maintaining competitive performance. However, several limitations must be acknowledged. The current experimental design lacks rigorous alignment with the stated research objectives, and the literature review does not adequately situate this work within the broader scope of algorithmic innovations. Furthermore, inconsistencies in model descriptions, figure clarity, and result-to-data alignment signal the need for improved academic representation. Future work should refine the hypothesis, incorporate recent state-of-the-art models, and ensure consistency in methodology and reporting. By addressing these issues and expanding evaluation metrics, SBP has the potential to evolve into a robust, scalable solution for training memory-constrained deep learning models in both research and industry contexts.

7. Acknowledgements

We want to acknowledge the valuable insights and constructive feedback provided by our colleagues at Mohanbabu University. Their contributions have significantly enhanced this research on scalable deep neural network training. We also extend our appreciation to the anonymous reviewers for their thoughtful comments, which helped improve the quality of this work.

Author Contributions: Conceptualization, P.M. and C.S.; methodology, P.M.; software, P.M.; validation, P.M. and C.S.; formal analysis, P.M.; investigation, P.M.; resources, P.M.; data curation, P.M.; writing—original draft preparation, P.M.; writing—review and editing, C.S.; visualization, P.M.; supervision, C.S.; project administration, C.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: No conflict of interest between the authors.

References

[1] Ibrahim, G. J.; Rashid, T. A.; Akinsolu, M. O. An energy efficient service composition mechanism using a hybrid meta-heuristic algorithm in a mobile cloud environment. *J. Parallel Distrib. Comput.* **2020**, *143*, 77–87. https://doi.org/10.1016/j.jpdc.2020.05.002

[2] Raj, D. J. S. Optimized mobile edge computing framework for IoT based medical sensor network nodes. *J. Ubiquitous Comput. Commun. Technol.* **2021**, *3*(1), 33–42. https://doi.org/10.36548/jucct.2021.1.004

[3] Cong, P.; Zhou, J.; Li, L.; Cao, K.; Wei, T.; Li, K. A survey of hierarchical energy optimization for mobile edge computing: A perspective from end devices to the cloud. *ACM Comput. Surv.* **2020**, *53*(2), 1–44. https://doi.org/10.1145/3378935

[4] Naouri, A.; Wu, H.; Nouri, N. A.; Dhelim, S.; Ning, H. A novel framework for mobile-edge computing by optimizing task offloading. *IEEE Internet Things J.* **2021**, *8*(16), 13065–13076. https://doi.org/10.1109/JIOT.2021.3064225

[5] Li, G.; Yan, J.; Chen, L.; Wu, J.; Lin, Q.; Zhang, Y. Energy consumption optimization with a delay threshold in cloud-fog cooperation computing. *IEEE Access* **2019**, *7*, 159688–159697. https://doi.org/10.1109/ACCESS.2019.2950443

[6] Tang, C.; Wei, X.; Zhu, C.; Wang, Y.; Jia, W. Mobile vehicles as fog nodes for latency optimization in smart cities. *IEEE Trans. Veh. Technol.* **2020**, *69* (9), 9364–9375. https://doi.org/10.1109/TVT.2020.2970763

[7] Zhang, W. Z.; Elgendy, I. A.; Hammad, M.; Iliyasu, A. M.; Du, X.; Guizani, M.; Abd El-Latif, A. A. Secure and optimized load balancing for multitier IoT and edge-cloud computing systems. *IEEE Internet Things J.* **2020**, *8*(10), 8119–8132. https://doi.org/10.1109/JIOT.2020.3042433

[8] Wei, X.; Tang, C.; Fan, J.; Subramaniam, S. Joint optimization of energy consumption and delay in cloud-to-thing continuum. *IEEE Internet Things J.* **2019**, *6* (2), 2325–2337. https://doi.org/10.1109/JIOT.2019.2906287

[9] Yuvaraj, N.; Kousik, N. V.; Jayasri, S.; Daniel, A.; Rajakumar, P. A survey on various load balancing algorithm to improve the task scheduling in cloud computing environment. *J. Adv. Res. Dyn. Control Syst.* **2019**, *11*(08), 2397–2406.

[10] Kai, C.; Zhou, H.; Yi, Y.; Huang, W. Collaborative cloud-edge-end task offloading in mobile-edge computing networks with limited communication capability. *IEEE Trans. Cogn. Commun. Netw.* **2020**, *7*(2), 624–634. https://doi.org/10.1109/TCCN.2020.3018159

[11] Peng, H.; Wen, W. S.; Tseng, M. L.; Li, L. L. Joint optimization method for task scheduling time and energy consumption in mobile cloud computing environment. *Appl. Soft Comput.* **2019**, *80*, 534–545. https://doi.org/10.1016/j.asoc.2019.04.027

[12] Wu, H.; Wolter, K.; Jiao, P.; Deng, Y.; Zhao, Y.; Xu, M. EEDTO: An energy-efficient dynamic task offloading algorithm for blockchain-enabled IoT-edge-cloud orchestrated computing. *IEEE Internet Things J.* **2020**, *8*(4), 2163–2176. https://doi.org/10.1109/JIOT.2020.3033521

[13] Velmurugadass, P.; Dhanasekaran, S.; Anand, S. S.; Vasudevan, V. Enhancing Blockchain security in cloud computing with IoT environment using ECIES and cryptography hash algorithm. *Mater. Today Proc.* **2021**, *37*, 2653–2659. https://doi.org/10.1016/j.matpr.2020.08.519

[14] Yang, P.; Zhang, Y.; Lv, J. Load optimization based on edge collaboration in software defined ultra-dense networks. *IEEE Access* **2020**, *8*, 30664–30674. https://doi.org/10.1109/ACCESS.2020.2973041

[15] Saad, M. M.; Khan, M. T. R.; Srivastava, G.; Jhaveri, R. H.; Islam, M.; Kim, D. Cooperative vehicular networks: An optimal and machine learning approach. *Comput. Electr. Eng.* **2022**, *103*, 108348. https://doi.org/10.1016/j.compeleceng.2022.108348.

ASEAN
Journal of Scientific and Technological Reports
Online ISSN:2773-8752

ASEAN
Journal of Scientific and Technological Reports
Online ISSN: 2773-8752

Article type

Type of the Paper (Article, Review, Communication, etc.)about 8,000 words maximum

Title (Palatino Linotype 18 pt, bold)

Firstname Lastname[1], Firstname Lastname[2] and Firstname Lastname[2*]

[1] Affiliation 1; e-mail@e-mail.com
[2] Affiliation 2; e-mail@e-mail.com
* Correspondence: e-mail@e-mail.com; (one corresponding authors, add author initials)

Citation:
Lastname, F.; Lastname, F.; Lastname, F. Title. *ASEAN J. Sci. Tech. Report.* **2023**, *26*(X), xx-xx. https://doi.org/10.55164/ajstr.vxxix.xxxxxx

Article history:
Received: date
Revised: date
Accepted: date
Available online: date

Abstract: A single paragraph of about 400 words maximum. Self-contained and concisely describe the reason for the work, methodology, results, and conclusions. Uncommon abbreviations should be spelled out at first use. We strongly encourage authors to use the following style of structured abstracts, but without headings: (1) Background: Place the question addressed in a broad context and highlight the purpose of the study; (2) Methods: briefly describe the main methods or treatments applied; (3) Results: summarize the article's main findings; (4) Conclusions: indicate the main conclusions or interpretations.

Keywords: keyword 1; keyword 2; keyword 3 (List three to ten pertinent keywords specific to the article yet reasonably common within the subject discipline.)

1. Introduction

The introduction should briefly place the study in a broad context and highlight why it is crucial. It should define the purpose of the work and its significance. The current state of the research field should be carefully reviewed and critical publications cited. Please highlight controversial and diverging hypotheses when necessary. Finally, briefly mention the main aim of the work. References should be numbered in order of appearance and indicated by a numeral or numerals in square brackets—e.g., [1] or [2, 3], or [4–6]. See the end of the document for further details on references.

2. Materials and Methods

The materials and methods should be described with sufficient details to allow others to replicate and build on the published results. Please note that your manuscript's publication implicates that you must make all materials, data, computer code, and protocols associated with the publication available to readers. Please disclose at the submission stage any restrictions on the availability of materials or information. New methods and protocols should be described in detail, while well-established methods can be briefly described and appropriately cited.

Interventional studies involving animals or humans, and other studies that require ethical approval, must list the authority that provided approval and the corresponding ethical approval code.

2.1 Subsection

2.1.1. Subsubsection

3. Results and Discussion

This section may be divided by subheadings. It should provide a concise and precise description of the experimental results, their interpretation, as well as the experimental conclusions that can be drawn.Authors should discuss the results and how they can be interpreted from previous studies and the working hypotheses. The findings and their implications should be discussed in the broadest context possible. Future research directions may also be highlighted.

3.1. Subsection

3.1.1. Subsubsection

3.2. Figures, Tables, and Schemes

All figures and tables should be cited in the main text as Figure 1, Table 1, etc.

Figure 1. This is a figure. Schemes follow the same formatting.

Table 1. This is a table. Tables should be placed in the main text near the first time they are cited.

Title 1	**Title 2**	**Title 3**
entry 1	data	data
entry 2	data	data [1]

[1] Table may have a footer.

3.3. Formatting of Mathematical Components

This is example 1 of an equation:

$$a = 1, \tag{1}$$

The text following an equation need not be a new paragraph. Please punctuate equations as regular text. This is example 2 of an equation:

$$a = b + c + d + e + f + g + h + i + j + k + l + m + n + o + p + q + r + s + t + u \tag{2}$$

The text following an equation need not be a new paragraph. Please punctuate equations as regular text. The text continues here.

4. Conclusions

Concisely restate the hypothesis and most important findings. Summarize the significant findings, contributions to existing knowledge, and limitations. What are the future directions? Conclusions MUST be well stated, linked to original research question & limited to supporting results.

5. Acknowledgements

Should not be used to acknowledge funders – funding will be entered as a separate. As a matter of courtesy, we suggest you inform anyone whom you acknowledge.

Author Contributions: For research articles with several authors, a short paragraph specifying their individual contributions must be provided. The following statements should be used "Conceptualization, X.X. and Y.Y.; methodology, X.X.; software, X.X.; validation, X.X., Y.Y. and Z.Z.; formal analysis, X.X.; investigation, X.X.; resources, X.X.; data curation, X.X.; writing—original draft preparation, X.X.; writing—review and editing, X.X.; visualization, X.X.; supervision, X.X.; project administration, X.X.; funding acquisition, Y.Y. All authors have read and agreed to the published version of the manuscript." Please turn to the CRediT taxonomy for the term explanation. Authorship must be limited to those who have contributed substantially to the work reported.

Funding: Please add: "This research received no external funding" or "This research was funded by NAME OF FUNDER, grant number XXX" and "The APC was funded by XXX". Check carefully that the details given are accurate and use the standard spelling of funding agency names at https://search.crossref.org/funding. Any errors may affect your future funding.

Conflicts of Interest: Declare conflicts of interest or state "The authors declare no conflict of interest." Authors must identify and declare any personal circumstances or interest that may be perceived as inappropriately influencing the representation or interpretation of reported research results. Any role of the funders in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript, or in the decision to publish the results must be declared in this section. If there is no role, please state "The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results".

References

References must be numbered in order of appearance in the text (including citations in tables and legends) and listed individually at the end of the manuscript. We recommend preparing the references with a bibliography software package, such as EndNote, ReferenceManager to avoid typing mistakes and duplicated references. Include the digital object identifier (DOI) for all references where available.

Citations and references in the Supplementary Materials are permitted provided that they also appear in the reference list here.

In the text, reference numbers should be placed in square brackets [ ] and placed before the punctuation; for example [1], [1–3] or [1, 3]. For embedded citations in the text with pagination, use both parentheses and brackets to indicate the reference number and page numbers; for example [5] (p. 100), or [6] (pp. 101–105).

Using the American Chemical Society (ACS) referencing style

[1] Author 1, A.B.; Author 2, C.D. Title of the article. *Abbreviated Journal Name* Year, *Volume*, page range.

[2] Author 1, A.; Author 2, B. Title of the chapter. In *Book Title*, 2nd ed.; Editor 1, A., Editor 2, B., Eds.; Publisher: Publisher Location, Country. **2007**, Volume 3, pp. 154–196.

[3] Author 1, A.; Author 2, B. *Book Title*, 3rd ed.; Publisher: Publisher Location, Country, **2008**, pp. 154–196.

[4] Author 1, A.B.; Author 2, C. Title of Unpublished Work. *Abbreviated Journal Name* stage of publication (under review; accepted; in press).

[5] Author 1, A.B. (University, City, State, Country); Author 2, C. (Institute, City, State, Country). Personal communication, 2012.

[6] Author 1, A.B.; Author 2, C.D.; Author 3, E.F. Title of Presentation. In Title of the Collected Work (if available), Proceedings of the Name of the Conference, Location of Conference, Country, Date of Conference; Editor 1, Editor 2, Eds. (if available); Publisher: City, Country, Year (if available); Abstract Number (optional), Pagination (optional).

[7] Author 1, A.B. Title of Thesis. Level of Thesis, Degree-Granting University, Location of University, Date of Completion.

[8] Title of Site. Available online: URL (accessed on Day Month Year).

Reviewers sugestion

1. Name, Address, e-mail
2. Name, Address, e-mail
3. Name, Address, e-mail
4. Name, Address, e-mail

URL link:

Notes for Authors >>
https://drive.google.com/file/d/1r0zegnIVeQqe4iLQyT1xDEIinNggINPD/view?usp=sharing
https://drive.google.com/file/d/1r0zegnIVeQqe4iLQyT1xDEIinNggINPD/view?usp=sharing

Online Submissions >> https://ph02.tci-thaijo.org/index.php/tsujournal/user/register

Current Issue >> https://ph02.tci-thaijo.org/index.php/tsujournal/issue/view/16516

AJSTR Publication Ethics and Malpractice >> https://ph02.tci-thaijo.org/index.php/tsujournal/ethics

Journal Title Abbreviations >> http://library.caltech.edu/reference/abbreviations

ASEAN
Journal of Scientific and Technological Reports

Online ISSN:2773-8752

ASEAN
Journal of Scientific and Technological Reports
Online ISSN:2773-8752

ASEAN Journal of Scientific and Technological Reports (AJSTR)
Institute of Research and Innovation, Thaksin University, Phatthalung Campus
222 Moo. 2, Baan Prao Sub-District, Papayom District, Phatthalung, 93210, Thailand
Tel. +668 1540 7304