MJFAS
Malaysian Journal of Fundamental and Applied Sciences

Chuan *et al.* | Malaysian Journal of Fundamental and Applied Sciences, Vol. 22 (2026) 955–973

RESEARCH ARTICLE

A Modified Single-Stacked Ensemble Regression Framework for Data-Driven Rice Yield Forecasting in Southeast Asia

Zun Liang Chuan[a,*], Tham Ren Sheng[b], Abraham Lim Bing Sern[a], David Lau King Luen[a], Tan Chek Cheng[a], Chong Yeh Sai[c]

[a]Centre for Mathematical Sciences, Universiti Malaysia Pahang Al-Sultan Abdullah, Lebuh Persiaran Tun Khalil Yaakob, 26300 Kuantan, Pahang, Malaysia; [b]Faculty of Industrial Management, Universiti Malaysia Pahang Al-Sultan Abdullah, Lebuh Persiaran Tun Khalil Yaakob, 26300 Kuantan, Pahang, Malaysia; [c]Ever AI Holdings Sdn Bhd, 12, Jalan Anggerik Aranda 31/170C, 40460 Kota Kemuning, Shah Alam, Selangor, Malaysia

Abstract Food insecurity in Southeast Asia has intensified in recent years, highlighting the urgent need for reliable and interpretable quantitative methods for agricultural yield forecasting under data-limited conditions. This study develops a modified single-stacked ensemble statistical framework that integrates Simple Linear Regression (SLR) and Multiple Linear Regression (MLR) to enhance predictive robustness and interpretability. The proposed framework is implemented within the Cross Industry Standard Process for Data Mining (CRISP-DM), ensuring systematic data preparation, algorithm construction, and validation across five Southeast Asian nations, utilizing 28 potential determinants. From a numerical statistics perspective, the ensemble configuration is designed to improve predictive stability and reduce variance while maintaining algorithm transparency under constrained sample conditions. Comparative evaluation against selected conventional Artificial Intelligence (AI) algorithms demonstrates that the proposed regression-based ensemble achieves competitive and, in several cases, superior forecasting performance in terms of accuracy, generalizability, and computational efficiency. The results further quantify the interconnected effects of geographical characteristics, climate variables, and macro-structural factors on rice yield dynamics. By providing a scalable and computationally efficient modeling approach, this framework offers a practical foundation for applied mathematical and statistical applications in data-driven agricultural decision-making. The findings also offer actionable insights to support evidence-based policymaking and related strategic investment, contributing indirectly to the advancement of regional food security and related Sustainable Development Goals (SDGs), including SDG1 (No Poverty), SDG2 (Zero Hunger), SDG3 (Good Health and Well-being), and SDG8 (Decent Work and Economic Growth).

Keywords: Rice yield prediction, Ensemble regression, Data-driven forecasting, Food security, Sustainable Development Goals.

***For correspondence:**
chuanzl@umpsa.edu.my

Received: 02 March 2026
Accepted: 28 Julai 2026

©Copyright **Chuan**. This article is distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use and redistribution provided that the original author and source are credited.

Introduction

In agricultural science and food security research, rice production refers to the total quantity of rice harvested over a substantial measurement period, such as monthly, quarterly, or annually. In contrast, rice yield refers to the amount of rice produced per unit area of land, typically measured in kilograms or tons per hectare, over a similar period. Rice yield is influenced by determinants such as cultivated land area, resource inputs, and management practices. Although these terms are frequently utilized

interchangeably in the literature, they represent distinct concepts. Understanding this distinction is essential for developing predictive algorithms that analyze key determinants affecting food security.

The need for reliable and interpretable predictive algorithms has become increasingly urgent as food insecurity rises. The Food and Agriculture Organization (FAO) of the United Nations has reported an increasing trend in the 3-year average prevalence of moderate or severe food insecurity (percent of the total population) from 2018 to 2023, particularly in lower-middle- and upper-middle-income Southeast Asian nations, including Cambodia (KHM), Laos (LAO), Myanmar (MMR), the Philippines (PHL), Vietnam (VNM), Malaysia (MYS), and Thailand (THA). Given that rice is a staple food in Asia, addressing these challenges requires practical and data-driven solutions capable of both accurate prediction and meaningful interpretation for informing policy decisions and guiding strategic investments to enhance regional food security.

In the literature, various predictive algorithms have been explored in South, East, and Southeast Asia, ranging from panel econometrics and statistical modeling to machine learning (ML), deep learning (DL), and hybrid ML-DL approaches. Many nations in these regions are among the world's top ten rice producers in metric tons in 2024 [1]. However, determinants of rice production vary both within and across nations due to differences in geographical characteristics, climate conditions, economic policies, and agricultural strategies. This variation has hindered the development of a universal predictive algorithm for rice production. Moreover, most existing studies primarily emphasize agronomic and atmospheric determinants, frequently neglecting other dimensions of food security that encompass all four dimensions of food security. Such a narrow focus limits broader applicability to sustainable agriculture, food security, and economic outcomes. In addition, each study presents specific limitations, further restricting generalizability. Addressing these limitations is critical for advancing more comprehensive and actionable predictive algorithms.

To address these challenges, this article introduces pioneering statistical predictive algorithms designed to monitor and forecast rice yield under limited data availability. The proposed approach offers significant advantages for the Southeast Asian region. Specifically, the algorithms integrate all four dimensions of food security, including availability, accessibility, stability, and utilization, and are developed utilizing the well-established Cross Industry Standard Process for Data Mining (CRISP–DM) framework. The predictive framework is evaluated across four major rice-producing Southeast Asian nations, with MYS included despite not being among the top producers, to contrast national rice yield determinants and demonstrate applicability across diverse agricultural and economic contexts. Notably, the proposed single-stacked ensemble approach is a data mining and soft computing method that leverages statistical learning principles to provide robust and interpretable predictions under limited data conditions. While this study focuses on Southeast Asia, the structured methodology provides a scalable foundation for broader global food security applications

By enhancing both predictive accuracy and interpretability, the findings offer actionable, data-driven insights that support evidence-based policy and investment decisions. These insights contribute to improved food security while aligning with key Sustainable Development Goals (SDGs), including SDG1 (No Poverty), SDG2 (Zero Hunger), SDG3 (Good Health and Well-being), and SDG8 (Decent Work and Economic Growth). The remainder of this article is organized as follows: The next section reviews the related literature and identifies the research gaps. The followed by a section describing the research methodology and theoretical framework. The subsequent section presents and discusses the empirical findings. The final section concludes the study and outlines directions for future research.

Related Works

Accurate prediction of rice yield plays a crucial role in ensuring food security and supporting evidence-based policy interventions in Southeast Asia. While rice production research has traditionally emphasized total harvested quantities, increasing attention has been given to modeling rice yield per unit area in order to better capture the effects of agronomic, climatic, and socioeconomic determinants. A wide range of predictive methodologies has been applied across South, East, and Southeast Asian nations, including panel econometrics, statistical modeling, ML, DL, and hybrid ML–DL approaches. However, many existing studies tend to focus primarily on agronomic and atmospheric determinants, with relatively limited integration of broader food security dimensions such as availability, accessibility, stability, and utilization. In addition, some studies implicitly assume that test set evaluation reflects future forecasting performance without generating actual yield projections, while others do not clearly demonstrate forecasting outcomes based on their best-performing predictive algorithms. These methodological limitations may reduce practical applicability for sustainable agriculture, economic

planning, and policy making. To address these gaps, this section reviews recent rice yield prediction research in the region, highlighting methodological approaches, key determinants, and areas for improvement that motivate the development of the more robust, interpretable, and data-driven framework proposed in this study. By integrating comprehensive food security determinants and emphasizing interpretability under limited data conditions, the proposed single-stacked ensemble algorithm aims to provide a scalable solution for rice yield forecasting across diverse Southeast Asian contexts.

In South Asia, Sathya & Gnanasekaran [2] proposed a hybrid Multiple Linear Regression–Long Short-Term Memory (MLR–LSTM) predictive algorithm to predict rice yield in Thanjavur district, India, utilizing crop, weather, and soil determinants. Their findings indicated that the MLR–LSTM outperforms several ML and DL predictive algorithms, including Random Forest (RF), Support Vector Machine (SVM), and LSTM. Nevertheless, certain methodological aspects could be further clarified. Feature selection was conducted utilizing a Python function without a detailed explanation of its statistical rationale, which may limit transparency. In addition, the study did not provide explicit forecasting results based on the proposed predictive algorithm. Furthermore, the hybrid structure may present interpretability challenges typically associated with black-box algorithms, potentially constraining its applicability in policy-oriented contexts.

Sahoo et al. [3] compared several AI-based predictive algorithms, including Cubist, Gradient Boosting Model (GBM), Multivariate Adaptive Regression Spline (MARS), RF, SVM, and Extreme Gradient Boosting (XGB), to predict rice yield in West Bengal, India. The study incorporated elevation, soil moisture, precipitation, temperature, soil temperature, and actual evapotranspiration as determinants, thereby capturing multiple dimensions relevant to food security. However, the reliance on spatially simulated climatic datasets (1990–2030) may influence the robustness and interpretation of results, particularly with regard to real-world variability.

Sharma et al. [4] utilized stepwise regression analysis, discriminant function analysis (DFA), and ordinal logistic regression (OLR) to predict rice yield in Karnal district, India, utilizing weather determinants. While the study explored alternative statistical approaches, DFA and OLR are primarily classification techniques, which may not be ideally suited for continuous outcome forecasting, such as rice yield. Transforming a time-dependent yield dataset into categorical outcomes may reduce its capacity to capture temporal structure. Additionally, distinctions between multivariable and multivariate analysis could have been clarified to avoid conceptual ambiguity. These issues may affect the interpretability and robustness of the findings.

Meanwhile, Amarasinghe et al. [5] compared MLR, SVM, k-Nearest Neighbor (kNN), and RF for predicting seasonal rice yield in Sri Lanka. The study found RF to outperform other algorithms, with weather determinants identified as key determinants. While the findings contribute to sustainability objectives aligned with the SDGs, several aspects could benefit from further refinement. For instance, the stochastic nature of RF and limited feature selection procedures may affect interpretability and forecasting stability. The utilization of both hold-out validation and five-fold cross-validation may also introduce redundancy if not clearly justified. Additionally, the selection of the best-performing algorithm could potentially be strengthened through multi-criteria decision-making (MCDM) approaches to address conflicting Goodness-of-Fit (GoF) measures.

In East Asia, Mo et al. [6] proposed an LSTM-based DL approach to predict rice yield in Guangxi Province, China, utilizing meteorological determinants. The study reported superior performance compared to support vector regression (SVR). However, the claim of outperforming conventional recurrent neural networks (RNNs) was not supported by detailed comparative results. The relatively small dataset and reliance solely on Mean Squared Error (MSE) as the evaluation metric may limit the comprehensiveness of performance assessment. Furthermore, test set performance was interpreted as indicative of forecasting capability without generating explicit future projections. As with many DL approaches, interpretability considerations remain relevant for practical policy applications.

On the other hand, research on rice yield prediction incorporating key determinants in lower-middle-income Southeast Asian nations remains limited. For instance, Pheakdey et al. [7] applied an Ordinary Least Squares-Multiple Linear Regression (OLS–MLR) predictive algorithm to predict seasonal rice yield in Kampot province, KHM, utilizing climate determinants. While the study provides useful insights, the analysis primarily relied on histogram-based distribution assessment without extensive residual diagnostics or formal OLS assumption testing. Addressing these aspects could strengthen confidence in coefficient interpretation and predictive reliability. Additionally, forecasting applications were not explicitly demonstrated.

In the PHL, David [8] compared MLR, RF, and Artificial Neural Network (ANN) predictive algorithms for rice yield prediction in the Pampanga River basin. RF achieved superior predictive accuracy without apparent overfitting. Nonetheless, limited feature selection across predictive algorithms may introduce redundancy and affect interpretability. In the ANN predictive algorithm, hyperparameter configuration, such as the number of neurons, hidden layers, and training epochs, could potentially influence predictive performance, warranting further optimization analysis.

Similarly, Baltazar [9] conducted rice yield forecasting for Central Luzon, PHL, utilizing MLR, RF, Gradient Boosting (GB), and ANN predictive algorithms. Although all predictive algorithms were reported to perform reasonably well, the GoF measures varied considerably. The reliance on a simulated climate dataset with merely 10 years of observed data (2009–2018) may limit the reliability of long-term projections, particularly for ANN predictive algorithms that typically benefit from larger datasets. Random splitting of time series data for cross-validation may also disrupt temporal structure, potentially affecting forecasting validity. In contrast, Nguyen et al. [10] employed a Feasible Generalized Least Squares (FGLS) panel predictive algorithm to analyze seasonal climate effects on rice yield in the Central Coast, VNM. While the study offers valuable insights into seasonal variability, it did not extend the analysis to forecasting applications. Moreover, focusing primarily on climate determinants may not fully capture broader food security considerations.

Research on modeling and forecasting rice yield in upper-middle-income Southeast Asian nations remains limited compared to studies on rice production, particularly regarding key determinants for food security. In IDN, Aziz et al. [11] proposed a Bayesian Deep Belief Network (BDBN) incorporating agronomic, climate, and topographic determinants. Although BDBN demonstrated strong predictive performance, feature selection was conducted subjectively utilizing a Pareto chart, which may affect reproducibility. As with other NN-based predictive algorithms, hyperparameter tuning remains critical to predictive algorithm stability and interpretability.

For MYS, Putri et al. [12] applied an MLR predictive algorithm for Sungai Besar, Selangor, utilizing chlorophyll and soil nutrient determinants. While aiming to enhance profitability and yield, reliance on filter-based feature selection and interaction-heavy predictive algorithm specification reduces interpretability despite statistical validity. Building on this, Tan et al. [13] developed a fixed-effect panel regression predictive algorithm for eight granary areas in Peninsular Malaysia, MYS. Although incorporating agronomic and climate determinants, certain methodological sequences, such as pre-screening and diagnostics, could be more clearly structured. Dependence on climate simulations and nearby station data may also introduce potential measurement bias.

More recently, Marong et al. [14] evaluated SVM, RF, and ANN predictive algorithms for rice yield prediction in MYS, utilizing environmental factors as determinants. While RF showed strong correlation-based performance, reliance primarily on correlation coefficients for predictive algorithm ranking may not fully capture predictive robustness across multiple GoF measures. Additionally, forecasting results based on the selected superior predictive algorithm were not extensively demonstrated.

In THA, Saithanu et al. [15] developed an MLR predictive algorithm to predict rice yield across 13 northern provinces, incorporating agronomic and climate determinants. However, the study primarily emphasized rice production rather than yield-specific dynamics. Additionally, the reliance on Root Mean Square Error (RMSE) as the sole evaluation metric may limit the comprehensiveness of performance assessment, potentially affecting the robustness of the predictive conclusions. The absence of explicit forecasting analysis may also reduce its practical relevance for sustainable agriculture planning, food security assessment, and economic decision-making.

Across South, East, and Southeast Asia, rice yield predictive modeling demonstrates methodological diversity and growing sophistication. However, many studies remain partially disconnected from broader food security dimensions and practical forecasting applications. The exclusion of determinants representing all four dimensions of food security may limit policy relevance. Additionally, interpretability and methodological transparency remain recurring considerations. These observations highlight the need for a predictive framework that integrates comprehensive food security determinants while maintaining methodological rigor and interpretability. The present study addresses these considerations by developing a robust and scalable predictive algorithm designed to enhance forecasting accuracy and support sustainable agricultural decision-making.

Theoretical Framework and Research Methodology

This study adopts a structured theoretical and methodological framework integrating the four dimensions of food security with a data-driven predictive algorithm approach. The theoretical foundation is grounded in the multidimensional nature of food security, including availability, accessibility, stability, and utilization, which collectively influence agricultural productivity and long-term sustainability. These dimensions provide the conceptual basis for determinant selection and algorithm construction.

To operationalize this framework, the study employs the CRISP-DM, a widely recognized methodology suitable for both academic research and applied data-driven problem-solving. Originally developed through industry collaboration, CRISP-DM was designed to address complex real-world challenges utilizing a systematic data mining approach. The framework consists of six iterative phases: business understanding, data understanding, data preparation, modeling, evaluation, and deployment. These phases ensure a structured yet flexible process that translates research objectives into actionable predictive solutions.

Two principal reasons justify the selection of CRISP-DM in this study. First, the framework has been successfully applied across diverse research domains, including agricultural economics and engineering [16, 17], computer science [18], educational economics [19, 20], energy economics [21, 22], healthcare [23], forensic science [24], and finance [25]. Second, compared to other established methodologies, such as Knowledge Discovery in Databases (KDD), and Sample, Explore, Modify, Model, and Assess (SEMMA), the CRISP-DM provides greater iterative flexibility, methodological transparency, and adaptability, making it appropriate for applied statistical and data mining research under data-limited conditions.

In this study, the CRISP-DM is implemented in a structured yet adaptive manner. The first three phases, including business understanding, data understanding, and data preparation, are presented and discussed in the first sub-section. These stages emphasize conceptual alignment with food security theory, data validation, and systematic pre-processing. The subsequent phases, including modeling, evaluation, and deployment, are detailed in the second sub-section, forming an iterative workflow that supports robust predictive performance and continuous refinement of feature selection and algorithm stability. The overall alignment between CRISP-DM phases and the proposed predictive framework is illustrated in Figure 1.

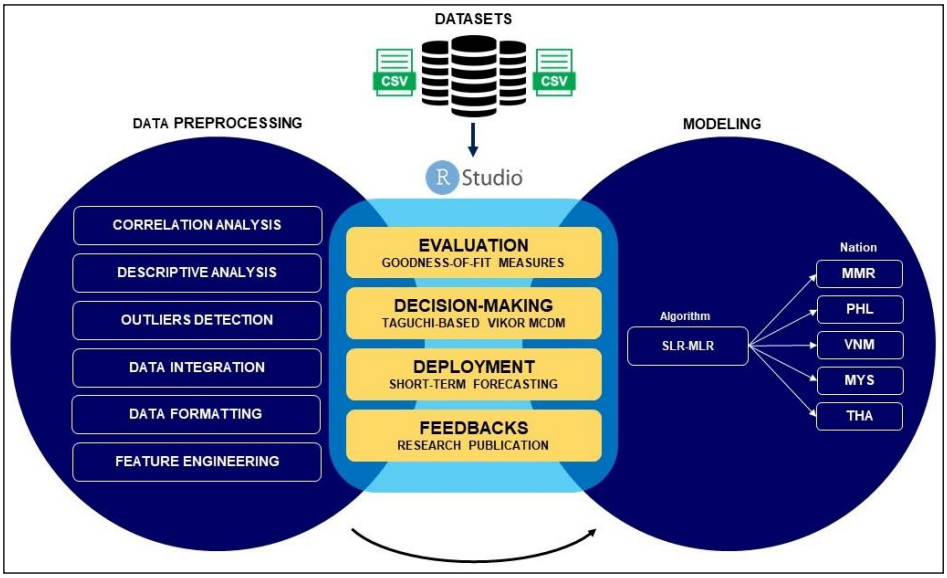


Figure 1. Schematic overview of the CRISP-DM framework in this study

Business understanding, data understanding, and data preparation

Business understanding constitutes the foundational phase of the CRISP-DM, during which the research objectives, scope, and practical constraints are clearly defined. In this study, the primary objective is to develop statistical predictive algorithms for monitoring and forecasting rice yield under data-limited conditions across five selected Southeast Asian nations, namely MMR, PHL, VNM, MYS, and THA. These nations were selected due to their strategic importance in regional rice production, with MYS included as a comparative reference to enhance cross-nation analytical robustness despite not being among the world's top rice producers. This phase also ensures that the theoretical grounding of food security is operationalized within the modeling framework. By explicitly incorporating availability, accessibility, stability, and utilization dimensions into determinant selection, this study establishes a structured linkage between agricultural productivity and broader socio-economic and environmental factors. This alignment ensures that predictive modeling remains conceptually consistent with food security theory rather than being purely data-driven.

Similar to other empirical studies, this study is subject to data availability and resource constraints. These constraints influence dataset coverage, temporal consistency, and determinant selection. To mitigate these limitations, all datasets utilized in this study were obtained from three widely recognized open-access repositories: The Climate Change Knowledge Portal (CCKP) [26, 27], Our World in Data (OWID) [28, 29], and the World Bank [30, 31]. These repositories provide reliable and consistent long-term historical records that have been widely utilized in empirical studies. Although certain agronomic and soil-specific determinants could not be fully incorporated due to data availability constraints, the selected determinants provide a comprehensive representation of key dimensions influencing rice yield, including environmental, economic, human, and social factors. These determinant categories are also consistent with broader sustainable development considerations, including SDG1 (No Poverty), SDG2 (Zero Hunger), SDG3 (Good Health and Well-being), and SDG8 (Decent Work and Economic Growth). By bridging methodological and empirical gaps, this study contributes valuable insights for agriculture policy formulation, investment planning, and long-term food security strategies.

Following the business understanding phase, the data understanding phase focuses on systematic data collection, description, exploration, and quality assessment. The dataset comprises one endogenous variable, which is rice yield (tons per hectare), and 28 determinants representing the four food security dimensions and four SDG pillars, as summarized in Table 1. All datasets were acquired in Comma-Separated Values (CSV) format and selected based on domain knowledge and an extensive literature review to ensure theoretical and empirical relevance. To examine linear associations among continuous variables, the Pearson correlation coefficient was employed. Alternative non-parametric measures, such as Spearman's or Kendall's coefficients, were not adopted, as the primary objective was to evaluate linear associations consistent with the regression-based modeling framework applied in this study.

Table 1. Potential determinants aligned with the key pillars of the SDGs and food security dimensions

SDGs's Pillar	Variable (Measurement unit)	Notation	Food Security Dimensions
Environmental	Annual CO_2 emissions (per capita)	γ_1	Availability and Stability
	Average Annual Maximum Temperature (˚C)	γ_2	Availability and Stability
	Average Annual Minimum Temperature (˚C)	γ_3	Availability and Stability
	Average Annual Precipitation (millimeters)	γ_4	Availability and Stability
	Fertilizer utilizes: Nutrient Nitrogen (kilograms per hectare)	γ_5	Availability
	Fertilizer utilizes: Nutrient Phosphate (kilograms per hectare)	γ_6	Availability
	Fertilizer utilizes: Nutrient Potash (kilograms per hectare)	γ_7	Availability
	Land use (per capita)	γ_8	Availability
Economic	Annual Growth of GDP (per capita)	γ_9	Availability
	Area harvested (hectare per capita)	γ_{10}	Availability
	Consumer Price Index	γ_{11}	Accessibility
	Gross National Income Growth (Annual %)	γ_{12}	Accessibility
	Import of goods and services (constant 2015 US$)	γ_{13}	Accessibility

	Inflation in Consumer Price (annual %)	γ_{14}	Accessibility and Stability
	Trade (% of GDP)	γ_{15}	Availability
	World crude oil prices (US$)	γ_{16}	Accessibility and Stability
Human	Cereals allocated to human food (billion tonnes)	γ_{17}	Utilization
	Daily per capita calorie supply (kilocalories)	γ_{18}	Accessibility and Utilization
	Daily per capita fat supply (grams)	γ_{19}	Accessibility and Utilization
	Daily per capita protein supply: Animal Product (grams)	γ_{20}	Accessibility and Utilization
	Daily per capita protein supply: Vegetal Product (grams)	γ_{21}	Accessibility and Utilization
Social	Food supply (kilocalories per capita per day)	γ_{22}	Utilization
	Food supply (protein grams per capita per day)	γ_{23}	Utilization
	Food supply (fat grams per capita per day)	γ_{24}	Utilization
	Percentage of urban population (%)	γ_{25}	Stability
	Percentage of rural population (%)	γ_{26}	Utilization
	Percentage of urbanization (%)	γ_{27}	Stability
	Population	γ_{28}	Accessibility

Data quality assessment was conducted through frequency tabulation and outlier detection utilizing the 1.5 interquartile range (1.5IQR), and 3IQR rules, as defined in Equations (1) – (4). Observations beyond the lower inner fence (LIF) and upper inner fence (UIF) were classified as mild outliers, whereas values exceeding the lower outer fence (LOF) and upper outer fence (UOF) were identified as extreme outliers. Extreme outliers were treated utilizing a capping approach to preserve the structural integrity of the time series data. The dataset consists of annual observations covering the period from 1961 to 2021. Therefore, the temporal ordering of observations was preserved throughout the data preparation and modeling processes, and no random shuffling was performed during dataset construction. This procedure ensured that the original chronological characteristics of the dataset were retained prior to the subsequent modeling procedures. Determinants containing missing values were excluded to maintain dataset completeness and analytical reliability.

$$\text{LIF} = Q_1(\mathbf{\Phi}) - 1.5\text{IQR} \tag{1}$$

$$\text{UIF} = Q_3(\mathbf{\Phi}) + 1.5\text{IQR} \tag{2}$$

$$\text{LOF} = Q_1(\mathbf{\Phi}) - 3\text{IQR} \tag{3}$$

$$\text{UOF} = Q_3(\mathbf{\Phi}) + 3\text{IQR} \tag{4}$$

where IQR = $Q_3(\mathbf{\Phi})$ $Q_1(\mathbf{\Phi})$, and Q_1 and Q_3 denote the first and third quartiles of the random variable $\mathbf{\Phi}$. The 3IQR rule applied to LOF and UOF is conceptually aligned with the six-sigma methodology. Here, $\mathbf{\Phi}$ represents either the endogenous variable or γ_i, where i =1, 2, 3, ..., 28.

The data preparation phase involves data selection, cleaning, construction, integration, and formatting to ensure high-quality inputs for predictive modeling. Inadequate data preparation may adversely affect the generalizability, interpretability, and predictive stability. Data selection was conducted iteratively during the understanding phase due to incomplete or irrelevant variables. Notably, this study deliberately adopts a complete-case analysis approach, retaining merely fully observed time series data from 1961 to 2021. Although complete-case analysis may reduce sample size, it enhances statistical reliability and avoids potential bias introduced by imputation techniques under data-limited conditions.

In addition, determinant γ_{27} was constructed utilizing Equation (5), as it was not directly available from the original repositories. The processed datasets were subsequently integrated into a unified CSV file compiled from multiple sources. Exploratory Data Analysis (EDA) was conducted utilizing the first four L-moments (L-Means, L-Coefficient of variation (L-CV), L-Skewness, and L-Kurtosis), as defined in Equations (6) – (9). These measures are particularly suitable for macro-level time series data due to their

robustness to small samples and extreme values. Given the subjective interpretation of distributional shape based solely on L-skewness and L-kurtosis, the Shapiro-Wilk test was additionally employed to provide formal statistical evidence of normality.

$$\gamma_{27} = \frac{\text{TUP}}{\text{TUP} + \text{TRP}} \times 100\% \tag{5}$$

$$\text{L} - \text{Means} = \binom{n}{1}^{-1} \sum_{j=1}^{n} (\varphi_j)_{(j:n)} \tag{6}$$

$$\text{L} - \text{CV} = \frac{1}{2}\binom{n}{2}^{-1} \sum_{j=1}^{n} \left\{ \binom{j-1}{1} - \binom{n-j}{1} \right\} (\varphi_j)_{(j:n)} \tag{7}$$

$$\text{L} - \text{Skewness} = \frac{1}{3}\binom{n}{3}^{-1} \sum_{j=1}^{n} \left\{ \binom{j-1}{2} - 2\binom{j-1}{1}\binom{n-j}{1} + \binom{n-j}{2} \right\} (\varphi_j)_{(j:n)} \tag{8}$$

$$\text{L} - \text{Kurtosis} = \frac{1}{4}\binom{n}{4}^{-1} \sum_{j=1}^{n} \left\{ \binom{j-1}{3} - 3\binom{j-1}{2}\binom{n-j}{1} + 3\binom{j-1}{1}\binom{n-j}{2} - \binom{n-j}{3} \right\} (\varphi_j)_{(j:n)} \tag{9}$$

where TUP denotes the total urban population, and TRP denotes the total rural population.

Finally, feature selection was implemented utilizing a hybrid approach combining an automated wrapper-based stepwise procedure with Student's *t*-test for statistical filtering. The optimal determinant subset was first identified utilizing the Akaike Information Criterion (AIC), followed by the elimination of statistically insignificant determinants based on Student's *t*-test. Detailed discussions of this procedure are presented in the predictive modeling section, as it directly informs the proposed ensemble-based framework. Through systematic data preparation and rigorous statistical validation, this study establishes a reliable foundation for predictive modeling, supporting robust inference for agricultural policy development, investment planning, and long-term food security assessment.

The proposed framework differs from conventional autoregressive time-series algorithms, as it does not model rice yield utilizing lagged endogenous variables or autoregressive components. Instead, this study adopts a regression-based stacked forecasting framework, where determinant forecasting and multivariable prediction are integrated to estimate future rice yield. Therefore, stationarity and lag-order identification procedures typically associated with autoregressive algorithms, such as unit-root testing utilizing the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test, were not adopted as a prerequisite for algorithm development. Nevertheless, the independence assumption of the regression algorithms was assessed through residual diagnostics, including a run test to ensure that the fitted algorithms did not exhibit systematic dependency patterns. Detailed descriptions of the proposed predictive modeling framework are provided in the subsequent section.

Modeling, evaluation, and deployment

The high-quality dataset prepared during the data preparation phase is utilized to develop a modified single-stacked ensemble predictive algorithm integrating Simple Linear Regression (SLR) and Multiple Linear Regression (MLR), hereafter referred to as the SLR-MLR predictive algorithm. This procedure corresponds to the modeling phase of the CRISP-DM framework. Specifically, this study introduces a structured modification to the conventional stacked ensemble methodology, which typically consists of two hierarchical levels: Level 0 (L0) and Level 1 (L1). In conventional stacked ensemble architectures, multiple heterogeneous base learners are trained at L0, and their generated outputs are subsequently combined through a meta-learner at L1. Although such architectures can enhance predictive performance in complex and high-dimensional environments, their direct application to rice yield forecasting under limited-data conditions may introduce challenges related to model generalizability, interpretability, and transparency.

To address these challenges, this study proposes a simplified and domain-consistent single-stacked regression framework, in which predictive information is sequentially refined through two hierarchical levels. Unlike conventional stacking approaches that aggregate outputs from multiple heterogeneous algorithms, the proposed framework emphasizes determinant-specific forecasting at the initial stage and multivariable statistical integration at the subsequent stage. Therefore, the term "stacked" in this study refers to the hierarchical integration of predictive information across modeling levels rather than solely the aggregation of multiple heterogeneous learners, thereby maintaining interpretability and methodological transparency.

At L0, SLR predictive algorithms are utilized to generate initial forecasts for each selected determinant. These determinant-specific forecasts serve as preliminary predictive information for the subsequent modelling stage. Subsequently, at L1, MLR is employed as the meta-learner to integrate the projected determinants and simultaneously algorithm the combined effects of multiple determinants of rice yield prediction. Therefore, the potential limitation of utilizing individual determinant forecasting is addressed through the subsequent multivariable integration process. Furthermore, the MLR stage incorporates a hybrid wrapper-filter feature selection strategy to identify the optimal determinant subset and reduce redundant information before the final prediction. Mathematically, the MLR predictive algorithm for rice yield prediction is expressed in Equation (10).

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1\gamma_1 + \hat{\beta}_2\gamma_2 + \cdots + \hat{\beta}_p\gamma_p; \quad p \leq 28 \tag{10}$$

where $\hat{y}$ denotes the estimated rice yield; $\hat{\beta}_0,\ \hat{\beta}_1,\ \hat{\beta}_2, \ldots, \hat{\beta}_p$ represent estimated regression coefficients, $\gamma_1,\ \gamma_2, \ldots, \gamma_p$, denote candidate determinants; and $p \leq 28$ represents the maximum number of potential determinants considered. Additionally, feature selection initially employs an automated stepwise wrapper procedure that identifies the optimal subset of determinants by minimizing the AIC as defined in Equation (11).

$$\mathrm{AIC} = 2(p+1) - 2\log(\hat{L}) \tag{11}$$

where $\log(\hat{L})$ denotes the log-likelihood function associated with the fitted MLR predictive algorithm. Subsequently, to maintain parsimony and remove statistically insignificant determinants, a Student's *t*-test is applied to evaluate the statistical significance of each retained determinant. The corresponding test statistic is provided in Equation (12).

$$\text{Student's } t - \text{test} = \frac{\hat{\beta}_k}{\sqrt{\left(\frac{RSS}{n-(k+1)}\right)[(\gamma'\gamma)^{-1}]_{kk}}} \tag{12}$$

where *RSS* denotes the residual sum of squares of the fitted MLR algorithm, $[(\gamma'\gamma)^{-1}]_{kk}$ represents the diagonal element corresponding to the *k*-th retained determinant, *k* is the number of selected determinants in the algorithm, and $k \leq p$. Although stepwise selection combined with hypothesis testing provides computational efficiency and determinant reduction, the relatively small sample size ($n = 61$) increases the potential risk of overfitting, particularly in the presence of multiple candidate determinants. Given the limited temporal observations available for each nation, the study adopts internal validation, whereby the predictive algorithm is constructed and evaluated utilizing a similar dataset. This approach allows for assessment of explanatory strength and apparent predictive accuracy through GoF measures, including the Adjusted-R^2 (GoF_1).

To mitigate algorithm complexity under limited sample conditions, strict determinant parsimony is enforced through a hybrid wrapper-filter feature selection strategy consisting of AIC-based stepwise selection followed by statistical significance testing. The proposed framework does not directly incorporate all candidate determinants into the final prediction model; instead, merely the retained determinants are subsequently evaluated utilizing Variance Inflation Factor (VIF) diagnostics to identify potential multicollinearity and ensure coefficient stability prior to estimating the final SLR-MLR predictive algorithm. The emphasis on determinant parsimony, coefficient stability, statistical diagnostics, and hierarchical refinement at L1 further constrains model complexity, reduces the potential risk of over-parameterization, and maintains the interpretability of individual determinants.

Given the limited temporal observations available for each nation, this study adopts an internal validation strategy, whereby the predictive algorithm is developed and evaluated utilizing the available dataset. Although rolling-origin, walk-forward, and other out-of-sample validation strategies are widely adopted in large-scale time-series forecasting studies, their application under the present data-limited setting would substantially reduce the effective sample size available for model estimation at each validation iteration and consequently increase the variability of parameter estimates and performance evaluation. Therefore, internal validation was considered an appropriate evaluation strategy considering the objectives and data

constraints of this study. Furthermore, the forecasting capability of the proposed framework is assessed utilizing multiple predictive performance measures, including GoF_1, RMSE (GoF_2), Mean Absolute Error (GoF_3), and Mean Absolute Percentage Error (GoF_4), rather than relying solely on in-sample GoF evaluation.

It should be noted that the proposed framework follows a determinant-based multivariable forecasting paradigm rather than a conventional autoregressive time-series forecasting approach. Unlike Autoregressive Integrated Moving Average (ARIMA)-based algorithms that primarily exploit historical lag structures of the endogenous variable, the proposed framework estimates future rice yield through the integration of projected determinants and their multivariable associations with rice yield. Therefore, benchmark comparisons with purely autoregressive algorithms may not fully represent the methodological objective of this study. While internal validation does not constitute external generalization testing, it provides a statistically coherent evaluation framework under limited data conditions, while preserving the structural interpretability of the proposed regression-based ensemble framework.

At L0, a series of SLR predictive algorithms is constructed, with each algorithm corresponding to an individual statistically significant determinant retained from the feature selection procedure. These SLR predictive algorithms are not utilized as final rice yield forecasting algorithms; instead, they generate determinant-specific initial forecasts that provide predictive information for the subsequent modeling stage. This approach enables transparent extrapolation of individual determinants while preserving determinant parsimony under limited-data conditions.

Subsequently, at L1, the MLR predictive algorithm is developed to integrate the forecasted determinant values generated from L0 and estimate rice yield by simultaneously considering the combined effects of multiple determinants. Therefore, the limitation associated with modeling rice yield based solely on individual determinants is addressed through the multivariable integration process at L1. The hierarchical refinement between L0 and L1 establishes structural coherence between determinant forecasting and final rice yield prediction, aligning with the objectives of the CRISP-DM modeling phase.

The evaluation phase aims to assess the predictive performance and statistical validity of the proposed SLR-MLR predictive algorithm. Predictive performance is evaluated through internal validation framework utilizing four GoF measures, including GoF_1, GoF_2, GoF_3, and GoF_4. Lower values of GoF_2, GoF_3, and GoF_4 indicate superior predictive performance, whereas GoF_1 values closer to unity reflect stronger explanatory capability. The mathematical formulations of these GoF measures, which quantify the closeness between the predicted rice yield $(\hat{y}_j)$ and the observed rice yield (y_j), are presented in Equations (13) – (16).

$$\mathrm{GoF}_1 = 1 - \left(\frac{n-1}{n-(k+1)}\right)(1-R^2) \tag{13}$$

$$\mathrm{GoF}_2 = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(y_j - \hat{y}_j\right)^2} \tag{14}$$

$$\mathrm{GoF}_3 = \frac{1}{n}\sum_{j=1}^{n}\left|y_j - \hat{y}_j\right| \tag{15}$$

$$\mathrm{GoF}_4 = \frac{100}{n}\sum_{j=1}^{n}\frac{\left|y_j - \hat{y}_j\right|}{\left|y_j\right|} \tag{16}$$

where GoF_1 corresponds to the adjusted form of the coefficient of determination.

In addition to predictive performance assessment, diagnostic tests are conducted to verify the assumptions underlying OLS estimation in accordance with the Gauss-Markov theorem. These include the Shapiro-Wilk test for normality of residuals, the runs test for independence, the Breusch-Pagan test for homoscedasticity, and the Variance Inflation Factor (VIF) for multicollinearity detection. Although

minor violations of OLS assumptions may not substantially affect predictive performance, they may influence coefficient interpretability and statistical reliability. Emphasis is therefore placed on objective and reproducible hypothesis testing to strengthen methodological transparency.

To reconcile potential conflicts among multiple GoF measures, a modified Taguchi-based VIKOR multi-criteria decision-making (MCDM) framework is applied to identify the best-performing SLR-MLR predictive algorithm across Southeast Asian nations. All GoF results are normalized into a matrix, $\Omega = [\kappa_{uv}]_{U\times V}; u,(v) = 1,2,\ldots,U,(V)$, where U denotes the number of predictive algorithm alternatives, and V represents the number of evaluation criteria (GoF measures). Each element is transformed into a signal-to-noise ratio (SNR) utilizing Equation (17).

$$\mathrm{SNR}_{uv} = -10\log(\kappa_{uv}) \tag{17}$$

where SNR_{uv} represents the signal-to-noise ratio corresponding to the uth algorithm, and vth GoF measure. Subsequently, a modified VIKOR index, τ_u is computed utilizing Equation (18), incorporating a reverse-ranking procedure to correct inconsistencies observed in preliminary ranking results.

$$\tau_u = \alpha\left(\frac{C_u - \min\limits_u(C_u)}{\max\limits_u(C_u) - \min\limits_u(C_u)}\right) + (1-\alpha)\left(\frac{D_u - \min\limits_u(D_u)}{\max\limits_u(D_u) - \min\limits_u(D_u)}\right) \tag{18}$$

where $C_u = \sum_{v=1}^{V}\left(\frac{\hat{\sigma}_v^2\left(\pi_v^+ - \kappa_{uv}\right)}{\pi_v^+ - \pi_v^-}\right)$ and $D_u = \max\limits_v\left(\frac{\hat{\sigma}_v^2\left(\pi_v^+ - \kappa_{uv}\right)}{\pi_v^+ - \pi_v^-}\right)$ represents the utility and regret measures, respectively, which $\pi_v^+ = \max\limits_u\left(\kappa_{uv}\right)$, $\pi_v^- = \min\limits_u\left(\kappa_{uv}\right)$, $\sigma_v^2 = \dfrac{\frac{1}{U-1}\sum_{u=1}^{U}\left(\kappa_{uv} - \bar{\kappa}_{uv}\right)^2}{\sum_{v=1}^{V}\left(\frac{1}{U-1}\sum_{u=1}^{U}\left(\kappa_{uv} - \bar{\kappa}_{uv}\right)^2\right)}$, $\sum_{v=1}^{V}\sigma_v^2 = 1$, and $\bar{\kappa}_{uv} = \frac{1}{U}\sum_{u=1}^{U}\kappa_{uv}$. In the absence of prior preference information, the weighting parameter α is set to 0.5 to balance utility and regret measures. This approach is consistent with established multi-criteria decision-making literature [32, 33].

Based on the VIKOR rankings, the top-ranked and bottom-ranked SLR-MLR predictive algorithms for each Southeast Asian nation are selected and deployed to forecast rice yields over a five-year horizon. Direct benchmarking against previously proposed multivariable predictive algorithms is not conducted due to structural differences in modeling design and forecasting objectives. Many existing approaches emphasize short-term in-sample prediction or rely on complex ML architectures that prioritize predictive optimization over determinant interpretability and statistical transparency. In contrast, the present framework is specifically designed for data-limited conditions, emphasizing hierarchical determinant consistency, statistical robustness, and interpretable forecasting. Algorithm robustness is assessed via repeated internal validation procedures and multi-criteria evaluation rather than through direct comparison with structurally dissimilar algorithms.

Analysis Results and Discussion

All analyses presented in this section were conducted utilizing Microsoft Excel and R statistical software on a mid-range computing environment (Intel(R) Core (TM) i5-10210U CPU @ 1.60GHz, 4 Core(s), 8 Logical Processor(s)). This demonstrates that the proposed SLR-MLR predictive algorithm can be implemented without reliance on high-end computational infrastructure, which is frequently required by ANN-based predictive approaches. To ensure structural clarity, the analysis results are integrated with the discussion and organized according to the research methodology described in the preceding section. Within the data science framework, the first sub-section focuses on descriptive and partially inferential

analyses associated with data understanding and data preparation, while business understanding was addressed in the previous section. The second sub-section emphasizes inference analysis through algorithm development, evaluation, and deployment.

Business understanding, data understanding, and data preparation

This study employs time series datasets spanning from 1961 to 2021 with no missing observations. Consequently, no missing value treatment was required. All 28 potential determinants included in the analysis were selected based on complete data availability across the entire study period. Determinants with partial temporal coverage were excluded, despite the availability of imputation techniques, to avoid introducing additional uncertainty without a clearly justifiable imputation strategy. Although rice yield data for 2022 and 2023 are available, several potential determinants were merely observed up to 2021. To maintain consistency in the temporal scope of the analysis, rice yield data for 2022 to 2023 were utilized solely to validate the short-term forecasting results, as discussed in the subsequent section.

Table 2 presents descriptive and partially inferential analyses of rice yield for five Southeast Asian nations: MMR, PHL, VNM, MYS, and THA. Due to space constraints, detailed results for all 28 potential determinants are not fully reported. The results indicate that VNM recorded the highest average rice yield from 1961 to 2021, whereas THA exhibited the lowest. This distinction highlights the important conceptual difference between rice production and yield. While VNM also achieved the highest average rice production (24288232.55 tonnes), THA followed closely (22391553.42 tonnes), and MYS recorded the lowest (2005385.48 tonnes). Despite its relatively high production level, THA's rice yield remained the lowest among the five nations, illustrating the typical misconception of utilizing rice production and rice yield interchangeably.

Table 2. Descriptive and partially inferential analyses of rice yield in five selected Southeast Asian nations

Nation	L-Moments				Outlier(s)		Normality (p-value)
	L-Means	L-CV	L-Skewness	L-Kurtosis	Mild	Extreme	
MMR	2.9108	0.1679	-0.1018	-0.0390	No	No	No (0.0000)
PHL	2.7079	0.1998	-0.0377	-0.0140	No	No	No (0.0018)
VNM	3.5604	0.2340	0.1192	-0.0754	No	No	No (0.0000)
MYS	2.9775	0.1149	0.0777	0.0675	No	No	Yes (0.1229)
THA	2.3460	0.1233	0.1091	-0.0846	No	No	No (0.0000)

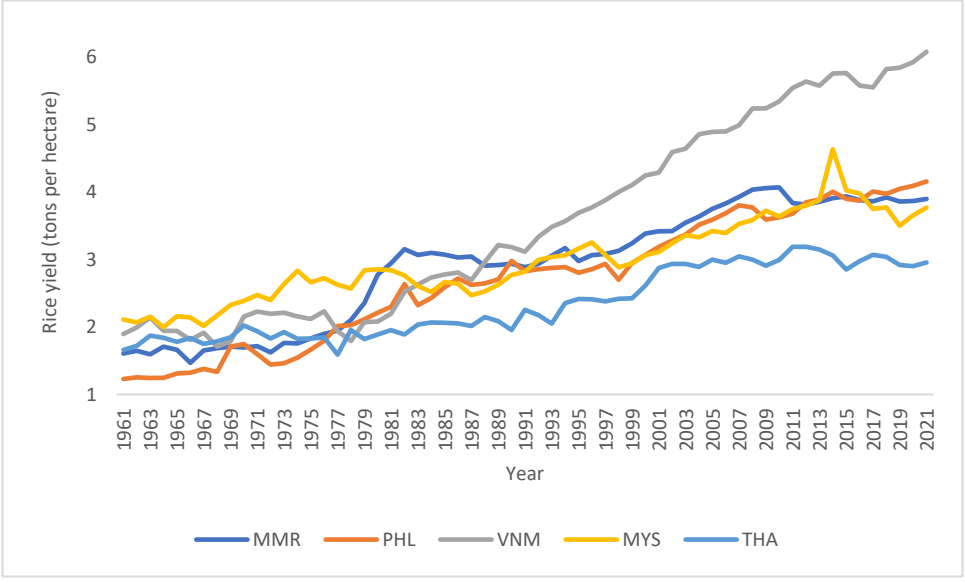


Figure 2. Rice yield trends for MMR, PHL, VNM, MYS, and THA over the period 1961–2021

Variability in rice yield was assessed utilizing the L-CV, where higher values indicate greater temporal fluctuations. As shown in Table 2, VNM exhibits the highest relative variation (23.40%), whereas MYS demonstrates the lowest (11.49%). These findings suggest that the higher variability in VNM may be influenced by uncontrolled or external factors, whereas MMR, PHL, MYS, and THA displayed moderate and comparatively stable yield patterns. This behaviour is further illustrated in Figure 2, where VNM's rice yield trajectory exhibits more pronounced fluctuations relative to the smoother trends observed in the other nations.

L-Skewness and L-Kurtosis are typically employed descriptive measures for assessing distributional shape. However, reliance solely on these indicators may lead to less robust conclusions. To address this limitation, the Shapiro–Wilk normality test was conducted to provide formal statistical evidence via *p*-values. According to Table 2, rice yield does not follow a normal distribution in MMR, PHL, VNM, and THA, whereas normality is not rejected for MYS. Similarly, non-normal distributions were observed for several of the 28 potential determinants across nations (MMR: γ_1, γ_5, γ_6, γ_7, γ_8, γ_9, γ_{10}, γ_{16}, γ_{17}, γ_{18}, γ_{19}, γ_{20}, γ_{21}, γ_{22}, γ_{23}, γ_{24}, γ_{28}; PHL: γ_5, γ_6, γ_7, γ_8, γ_9, γ_{10}, γ_{11}, γ_{13}, γ_{14}, γ_{16}, γ_{17}, γ_{19}, γ_{20}, γ_{21}, γ_{23}, γ_{24}, γ_{25}, γ_{26}, γ_{27}, γ_{28}; VNM: γ_1, γ_4, γ_5, γ_6, γ_8, γ_{10}, γ_{16}, γ_{17}, γ_{18}, γ_{19}, γ_{20}, γ_{21}, γ_{22}, γ_{23}, γ_{24}, γ_{25}, γ_{26} γ_{27}, γ_{28}; MYS: γ_1, γ_3, γ_5, γ_6, γ_7, γ_8, γ_9, γ_{10}, γ_{11}, γ_{12}, γ_{13}, γ_{14}, γ_{15}, γ_{16}, γ_{17}, γ_{18}, γ_{19}, γ_{20}, γ_{22}, γ_{23}, γ_{24}, γ_{25}, γ_{26}, γ_{27}, γ_{28}; THA: γ_1, γ_5, γ_6, γ_7, γ_8, γ_9, γ_{10}, γ_{11}, γ_{13}, γ_{14}, γ_{15}, γ_{16}, γ_{17}, γ_{18}, γ_{19}, γ_{20}, γ_{22}, γ_{23}, γ_{24}, γ_{25}, γ_{26}, γ_{27}, γ_{28}).

Importantly, the identification of non-normality in individual **Φ** at this stage does not violate the assumptions of the OLS estimation. A typical misconception in the literature is that each determinant must follow a normal distribution. In fact, the normality assumption in OLS pertains to the distribution of the residuals rather than individual **Φ**. Consequently, the normality assessment here serves the purpose of data understanding rather than acting as a strict diagnostic requirement for algorithm estimation.

Consistent with the research methodology, descriptive analysis was extended to detect mild and extreme outliers. No mild or extreme outliers were identified in rice yield across the five nations. In contrast, several mild and extreme outliers were observed among **Φ** for each nation. Specifically, the number of mild outliers detected in **Φ** for MMR, PHL, VNM, MYS, and THA were 8 (γ_1, γ_5, γ_6, γ_7, γ_9, γ_{16}, γ_{17}, γ_{20}), 10 (γ_1, γ_4, γ_7, γ_9, γ_{13}, γ_{14}, γ_{16}, γ_{17}, γ_{21}, γ_{23}), 5 (γ_1, γ_{16}, γ_{17}, γ_{21}, γ_{23}), 7 (γ_9, γ_{12}, γ_{14}, γ_{16}, γ_{18}, γ_{21}, γ_{22}), and 6 (γ_2, γ_4, γ_9, γ_{14}, γ_{16}, γ_{17}), respectively, while extreme outliers detected was 4 (γ_1, γ_5, γ_6, γ_7) for MMR, 2 (γ_9, γ_{14},) for PHL, none for VNM, 3 (γ_9, γ_{12}, γ_{14},) for MYS, and 1 (γ_{14}) for THA.

Despite the presence of the extreme outliers among **Φ**, no capping or transformation was applied. Preserving the temporal structure of the time series data was prioritized. Moreover, the presence of an extreme observation does not necessarily imply a high influence point (HIP) capable of substantially distorting parameter estimation. Accordingly, all observations were retained, with further diagnostics for HIPs deferred to the post-modeling evaluation stage.

Modeling, evaluation, and deployment

In accordance with the research methodology, Equations (19) – (23) present the proposed nation-specific rice yield predictive algorithms for MMR, PHL, VNM, MYS, and THA, respectively. These algorithms are derived from statistically significant determinants identified through the hybrid wrapper-filter feature selection procedure. The selected determinants differ across the five Southeast Asian nations, reflecting heterogeneity in geographical characteristics, climate conditions, economic structures, and agricultural policy frameworks. Such heterogeneity suggests that econometric approaches assuming structural homogeneity, most notably panel regression analysis, are not appropriate in this context.

$$\hat{y}_{\mathrm{MMR}} = 3.6610 + (7.123\mathrm{E}-1)\gamma_1 - (7.281\mathrm{E}-4)\gamma_8 + (9.156\mathrm{E}-4)\gamma_{16} - (1.289\mathrm{E}-7)\gamma_{17} - (7.222\mathrm{E}-2)\gamma_{20} + (7.634\mathrm{E}-2)\gamma_{23} - (2.499\mathrm{E}-1)\gamma_{25} + (1.049\mathrm{E}-7)\gamma_{28} \quad (19)$$

$$\hat{y}_{\mathrm{PHL}} = -0.7536 - (6.552\mathrm{E}-1)\gamma_1 - (1.985\mathrm{E}-2)\gamma_{11} + (3.242\mathrm{E}-4)\gamma_{16} + (7.290\mathrm{E}-8)\gamma_{28} \quad (20)$$

$$\hat{y}_{\mathrm{VNM}} = 1.3110 + (2.720\mathrm{E}-1)\gamma_1 + (5.306\mathrm{E}-3)\gamma_6 + (1.095\mathrm{E}-7)\gamma_{17} + (4.039\mathrm{E}-2)\gamma_{19} - (1.531\mathrm{E}-1)\gamma_{25} + (3.839\mathrm{E}-8)\gamma_{28} \quad (21)$$

$$\hat{y}_{\mathrm{MYS}} = -3.4930 + (2.508\mathrm{E}-3)\gamma_8 - (2.229\mathrm{E}-3)\gamma_{15} + (1.332\mathrm{E}-3)\gamma_{16} - (7.914\mathrm{E}-7)\gamma_{17} + (4.976\mathrm{E}-2)\gamma_{20} + (8.916\mathrm{E}-2)\gamma_{21} + (1.564\mathrm{E}-7)\gamma_{28} \quad (22)$$

$$\hat{y}_{\mathrm{THA}} = -0.2985 + (2.582\mathrm{E}-4)\gamma_4 - (1.050\mathrm{E}-2)\gamma_{14} + (8.818\mathrm{E}-3)\gamma_{15} + (3.197\mathrm{E}-2)\gamma_{21} + (8.322\mathrm{E}-9)\gamma_{28} \quad (23)$$

Panel regression techniques rely on the assumption that the association between the endogenous variable and explanatory determinants is structurally similar across cross-sectional units. This assumption is unlikely to hold in the present study, given the clear nation-specific determinants structures identified for rice yield. This reasoning is consistent with empirical findings reported by Chuan et al. [34], who applied a random-effect panel regression predictive algorithm to C3 crop production in Malaysia. Their results indicated a low coefficient of determination, merely one statistically significant determinant, and multiple violations of key algorithm assumptions. These outcomes suggest that panel regression algorithms may experience degraded explanatory and predictive performance when applied to highly heterogeneous systems. Consequently, panel regression analysis was not pursued further in this study.

The estimated regression coefficients in Equations (19) – (23) were subjected to diagnostic evaluation based on residual analysis to verify OLS assumptions. The Shapiro-Wilk test was utilized to assess residual normality (MMR: *p*-value=0.3584; PHL: *p*-value=0.1017; VNM: *p*-value=0.2475; MYS: *p*-value=0.0006; THA: *p*-value=0.0004). The runs test examined independence (MMR: *p*-value=0.0045; PHL: *p*-value=0.0020; VNM: *p*-value=0.1217; MYS: *p*-value=0.6069; THA: *p*-value=0.0709), and the Breusch-Pagan test assessed homoscedasticity (MMR: *p*-value=0.5361; PHL: *p*-value=0.2980; VNM: *p*-value=0.1811; MYS: *p*-value=0.3554; THA: *p*-value=0.8018). Multicollinearity was evaluated utilizing the VIF, as no formal hypothesis test exists for this diagnostic.

The diagnostic results indicate violations of OLS assumptions in four of the five predictive algorithms. Specifically, the independence assumption is rejected for MMR and PHL, while residual normality is rejected for MYS and THA. Additionally, VIF values exceeding conventional thresholds were observed for several determinants: five in MMR (γ_{17}, γ_{20}, γ_{23}, γ_{25}, γ_{28}), two in PHL (γ_{11}, γ_{28}), six in VNM (γ_1, γ_6, γ_{17}, γ_{19}, γ_{25}, γ_{28}), five in MYS (γ_8, γ_{17}, γ_{20}, γ_{21}, γ_{28}), and five in THA (γ_8, γ_{15}, γ_{17}, γ_{20}, γ_{28}), indicating the presence of multicollinearity.

Although transformation techniques such as the Box-Cox method mitigate violations related to non-normality or heteroscedasticity, such transformations alter the natural structure of the time series data. Given the study's objective of preserving temporal integrity and maintaining interpretability, no data transformation was applied. As a consequence, interpretation focuses primarily on the directional effects of statistically significant determinants rather than the precise magnitude of coefficient estimates. These determinants collectively represent the four SDG pillars and the four dimensions of food security, and their effects vary in both direction and magnitude across nations. This further reinforces the structural heterogeneity of the system and the inappropriateness of panel-based modeling strategies.

Despite diagnostic violations limiting the strict inferential interpretation, they do not invalidate the predictive capability of the proposed algorithms. As reported in Table 3, the predictive performance of Equations (19) – (23) demonstrates strong algorithm adequacy. The adjusted coefficient of determination (GoF_1) approaches unity for all five predictive algorithms, indicating substantial explanatory strength. Meanwhile, GoF_2 and GoF_3 values remain close to zero, and GoF_4 values are consistently below the 10% threshold, collectively indicating satisfactory apparent predictive accuracy under internal validation.

Table 3. Predictive performance of rice yields predictive algorithms across five selected Southeast Asian nations

Nation	Number of Determinants	GoF				Rank
		GoF_1	GoF_2	GoF_3	GoF_4	
MMR	8	0.9765	0.1223	0.0981	3.6389	2
PHL	4	0.9875	0.1001	0.0756	3.3428	1

VNM	6	0.9919	0.1236	0.1010	3.5478	5
MYS	7	0.9490	0.1258	0.0924	3.1105	4
THA	5	0.9683	0.0867	0.0632	2.7805	3

To rank overall algorithm performance across nations, the modified Taguchi-based VIKOR MCDM framework defined in Equation (18) was employed. This approach is particularly appropriate due to the conflicting nature of the evaluation attributes, including the number of determinants, GoF_1, GoF_2, GoF_3, and GoF_4. As these criteria inherently involve trade-offs between parsimony and predictive accuracy, the ranking derived from Equation (18) provides a structured compromise solution.

The rankings presented in Table 3 are logically consistent and interpretable. For instance, the top-ranked algorithm for PHL employs the fewest determinants while achieving the third lowest average of GoF_2, GoF_3, and GoF_4. In contrast, the bottom-ranked algorithm for VNM involves a relatively larger determinant set and exhibits comparatively higher error-based measures. Although the ranking procedure involves equal weighting ($\alpha = 0.5$), its primary objective is normative optimization, but the identification of representative best- and worst-performing algorithms for illustrative short-term forecasting, as shown in Figures 3 – 4. This distinction clarifies a typical misconception in the literature between forecast evaluation and the forecasting task itself.

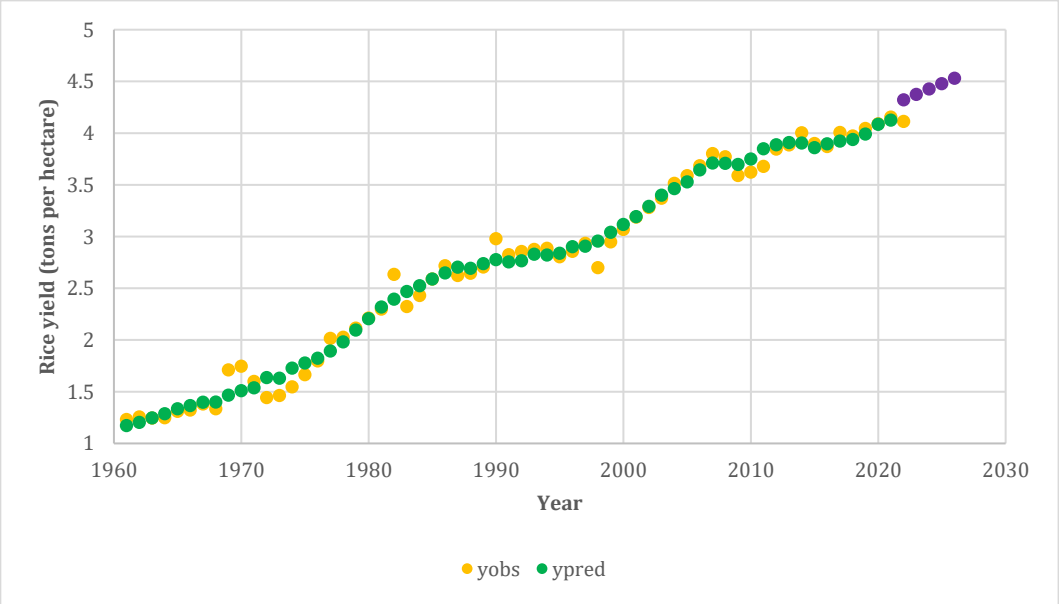


Figure 3. Short-term rice yield forecasting in PHL utilizing the top-ranked proposed predictive algorithm

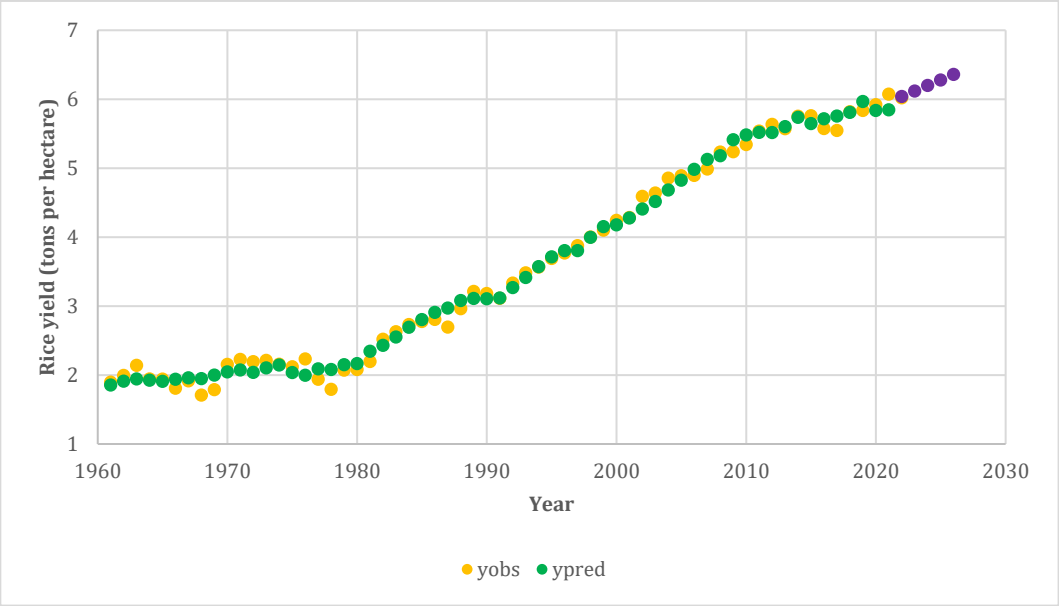


Figure 4. Short-term rice yield forecasting in VNM utilizing the bottom-ranked proposed predictive algorithm

To further assess forecasting reliability, observed rice yield data for 2022 to 2023 were utilized for validation in PHL and VNM, as 2024 data were unavailable at the time of analysis. For PHL, the observed rice yields in 2022 and 2023 were 4.1120 and 4.1654 tons per hectare, respectively, while the forecasted values were 4.3208 and 4.3728 tons per hectare, resulting in percentage errors of 5.0778% and 4.9791%, both within the 10% acceptability threshold. For VNM, the observed rice yields in 2022 and 2023 were 6.0180 and 6.1135 tons per hectare, respectively, while the forecasted values were 6.0373 and 6.1172 tons per hectare, corresponding to minimal percentage errors of 0.3207% and 0.0605%.

These findings indicate that even the bottom-ranked algorithm maintains satisfactory short-term forecasting performance. More importantly, the forecasted trajectories are not interpreted in isolation but examined in relation to ongoing structural and policy developments within the respective nations. In PHL, the forecasted moderate upward trend in rice yield is consistent with policy innovations aimed at enhancing productivity through input modernization and institutional support. The Rice Competitiveness Enhancement Fund [35], implemented under the Agri-Pinoy framework [36], emphasizes seed distribution, mechanization, credit assistance, and extension services. Several of these components are structurally related to determinants retained in the predictive algorithm, particularly those associated with agricultural inputs and productivity-enhancing mechanisms. The alignment between forecasted yield improvements and policy-driven productivity measures, therefore, reflects structural consistency rather than post-hoc narrative interpretation.

Similarly, in VNM, national initiatives, such as Climate-Smart Rice Cultivation [37], the 1 Million Hectares High-Quality Low-Carbon Project [38], and participation in the Sustainable Rice Platform [39] emphasize sustainable intensification, input efficiency, and climate adoption. Determinants retained in the VNM predictive structure include variables reflecting environmental conditions, production efficiency, and social dimensions of agricultural development, suggesting that the projected stability and incremental improvement in rice yield are compatible with ongoing structural transformation in the agricultural sector. It is important to clarify that the predictive algorithm does not establish causal attribution between specific policies and yield outcomes. Instead, the observed coherence between forecasted trajectories and documented policy directions supports the external plausibility and contextual validity of the algorithm outputs. This alignment enhances the practical relevance and credibility of the proposed predictive algorithm framework within real-world agricultural policy environments.

Conclusions and Future Recommendations

Recent statistics reported by the FAO indicate that the 3-year average prevalence of moderate or severe food insecurity has increased across both lower-middle- and upper-middle-income Southeast Asian nations between 2018 and 2023. This trend underscores the need for reliable and interpretable rice yield forecasting frameworks that support evidence-based agricultural planning under resource and data constraints. While numerous predictive algorithms have been proposed in the literature, several approaches either rely on data-intensive ML architectures with limited interpretability or encounter instability when applied to small sample environments. These limitations reduce their practical suitability for policy-oriented applications.

To address these concerns, this study developed a structured statistical predictive framework based on a modified single-stacked ensemble SLR–MLR predictive algorithm within the CRISP-DM process model. The framework emphasizes determinant consistency, hierarchical refinement, and statistical transparency rather than purely predictive optimization. Utilizing 28 candidate determinants, statistically significant determinants were identified through a hybrid wrapper-filter feature strategy, leading to the construction of nation-specific predictive algorithms for five Southeast Asian nations. The empirical results reveal substantial cross-national heterogeneity in determinant structures, reflecting differences in climatic exposure, socioeconomic conditions, and agricultural policy environments. This structural heterogeneity suggests that modeling strategies assuming homogeneous slope structures across cross-sectional units may not be appropriate in this context.

Diagnostic evaluation indicated partial violations of classical OLS assumptions in several of the estimated algorithms. Consistent with econometric theory, such violations primarily affect the efficiency and inferential reliability of coefficient estimates rather than the mechanical computation of fitted values. Accordingly, this study distinguishes clearly between inferential interpretation and predictive application. Coefficient interpretation is restricted to directional insights, while predictive adequacy, evaluated via internal validation and multi-criterial performance assessment, remains the principal objective. Under this framework, all proposed algorithms demonstrate satisfactory apparent forecasting performance within the sample and acceptable short-term validation against recent observations.

From a practical standpoint, the findings provide structured and interpretable insights for agricultural planning in data-limited conditions. The results highlight the combined influence of environmental, production-efficiency, and socioeconomic determinants in shaping rice yield dynamics, while emphasizing that their relative importance differs across national contexts. Rather than offering universal policy prescriptions, the framework supports context-specific evaluation and transparent determinant prioritization. In this regard, the study contributes to ongoing discussions surrounding agricultural development and food security in Southeast Asia. These insights may inform evidence-based agricultural planning and are consistent with broader SDGs, particularly SDG1 (No Poverty), SDG2 (Zero Hunger), SDG3 (Good Health and Well-being), and SDG8 (Decent Work and Economic Growth).

The proposed framework is scalable in structure and adaptable to other regions characterized by limited and heterogeneous datasets. Nevertheless, several directions for future research remain. Alternative estimation strategies, including Bayesian regression and regularized multivariate approaches, may improve parameter stability under assumption violations. Additionally, future studies may examine hybrid statistical ML configurations that enhance forecasting robustness while maintaining interpretability. External validation utilizing independent datasets or rolling-origin evaluation procedures would further strengthen generalization assessment and should be considered in subsequent research.

Conflicts of Interest

The authors declare that there are no conflicts of interest associated with the publication of the article. Additionally, a portion of the analysis presented herein was previously included in an extended abstract published in the International Journal of Innovation, Volume 2, Issue 1, pages 179–183. However, the present article contains substantial new content, with at least 30% differing from the previously published

extended abstract.

Acknowledgment

The authors sincerely thank the Food and Agriculture Organization (FAO) of the United Nations, the World Bank Group (Climate Change Knowledge Portal), and Global Change Data Lab for granting access to the open-access datasets employed in this research. Appreciation also extended to Universiti Malaysia Pahang Al-Sultan Abdullah for supporting this work through project funding (Ref: RDU220393, 2022-2024). The authors are especially grateful to the anonymous reviewers for their insightful comments and constructive recommendations, which have notably enhanced the quality of this article.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this manuscript, the authors utilized ChatGPT to enhance the clarity, readability, and language of the text. The authors carefully reviewed and edited all AI-assisted content to ensure accuracy and appropriateness. The authors take full responsibility for the final content presented in this publication.

References

[1] U.S. Department of Agriculture, Production-rice, https://www.fas.usda.gov/data/production/commodity/0422110
[2] P. Sathya, P. Gnanasekaran, Appl. Artif. Intell. 37(1) (2023) 2175113.
[3] S. Sahoo, C. Singha, A. Govind, J. Agric. Food Res. 18 (2024) 101424.
[4] A. Sharma, J. Kumar, M. Redhu, P. Kumar, M. Godara, P. Ghiyal, P. Fu, M. Rahimi, Sci. Rep. 14 (2024) 12626.
[5] A. Amarasinghee, I. Sangarasekara, N. De Silva, M. Ariyaratne, R. Amarasinghe, J. Bogahawatte, J. Alawatugoda, D. Herath, Discov. Appl. Sci. 6 (2024) 603.
[6] H. Mo, Y. Zhang, F. Liu, Y. Zheng, in: Proc. Journal of Physics: Conference Series, 1952 (2021) 042033.
[7] D. V. Pheakdey, T. D. Xuan, T. D. Khanh, AIMS Geosci. 3(4) (2017) 561-575.
[8] M. C. V. David, Univ. J. Agric. Res. 11(5) (2023) 836-848.
[9] R. G. Baltazar, Int. J. Agric. Nat. Resour. 51(1) (2024) 10-26.
[10] T. Nguyen, P. T. Ho, H. X. Pham, Cogent Econ. Finance 12(1) (2024) 2421894.
[11] Aziz, A. Aziz, Komariah, D. P. Ariyanto, Sumani, Sci. Horiz. 26(7) (2023) 149-159.
[12] R. E. Putri, A. Yahya, N. M. Adam, S. Abd Aziz, Food Res. 3(2) (2019) 171-176.
[13] B. T. Tan, P. S. Fam, R. B. R. Firdaus, M. L. Tan, M. S. Gunaratne, Agriculture 11 (2021) 569.
[14] M. Marong, N. A. Husin, M. Zolkepli, L. S. Affendey, Int. J. Adv. Comput. Sci. Appl. 15(3) (2024) 362-370.
[15] K. Saithanu, J. Mekparyup, P. Phodjanawichaikul, Glob. J. Pure Appl. Math. 11(4) (2015) 2181-2186.
[16] Z. L. Chuan, D. C. T. Wei, A. S. B. A. Aminuddin, S. -F. Fam, T. L. Ken, in: AIP Conf. Proc. 3150 (2024) 050008.
[17] Z. L. Chuan, T. R. Sheng, T. C. Cheng, A. L. B. Sern, D. L. K. Luen, C. Y. Sai, Multidiscipl. Appl. Res. Innov. 6(2) (2025) 1-17.
[18] J. Garcia-Arteaga, J. Zambrano-Zambrano, J. Parraga-Alava, J. Rodas-Silva, Comput. Speech Lang. 84 (2024) 101579.
[19] Z. L. Chuan, N. Japashov, S. K. Yuan, T. W. Qing, N. Ismail, Akademika 94(2) (2024) 232-251.
[20] Z. L. Chuan, D. C. T. Wei, A. Akhmedov, L. T. Man, T. E. Hiae, A. A. Hamizula, J. Kejuruteraan 37(2) (2025) 967-976.
[21] Z. L. Chuan, O. S. Jie, T. Y. Hin, S. N. S. B. M. Zain, Y. A. B. A. Rashid, A. N. B. Kamarudin, Int. J. Built Environ. Sustain. 12(1) (2025) 9-21.
[22] Z. L. Chuan, L. K. Tan, A. W. C. Chyin, Y. H. Tham, S. J. Ong, J. Y. Low, C. Y. Sai, MATEMATIKA 40(3) (2025) 143-167.
[23] O. Lohaj, J. Paralič, P. Bednár, Z. Paraličová, M. Huba, Mach. Learn. Knowl. Extr. 5(4) (2023) 1266-1281.
[24] Z. L. Chuan, D. C. T. Wei, C. L. W. Yan, M. F. A. Nasser, N. A. M. Ghani, A. A. Jemain, C. -Y. Liong, Malays. J. Fundam. Appl. Sci. 19(4) (2023) 525-538.
[25] A. Cheng, Finan. Res. Lett. 55 (2023) 103966.
[26] A. Alodah, Sustainability 15(20) (2023) 14674.
[27] C. L. S. Cedric, W. Y. H. Adoni, R. Aworka, J. T. Zoueu, F. K. Mutombo, M. Krichen, C. L. M. Kimpolo, Smart Agric. Technol. 2 (2022) 100049.
[28] R. C. A. Achterbergh, I. McGovern, M. Haag, Vaccines 12(2) (2024) 216.
[29] T. Alamo, D. G. Reina, M. Mammarella, A. Abella, Electronics 9(5) (2020) 827.
[30] A. Costantiello, A. Leogrande, Discov. Glob. Soc. 2 (2024) 1.
[31] A. Johri, M. Asif, P. Tarkar, W. Khan, Rahisha, M. Wasiq, Humanit. Soc. Sci. Commun. 11 (2024) 361.
[32] B. Ceballos, D. A. Pelta, M. T. Lamata, Int. J. Inf. Technol. Decis. Mak. 17(2) (2018) 513-525.
[33] Z. L. Chuan, A. L. B. Sern, R. S. Tham, D. L. K. Luen, C. C. Tan, J. Qual. Meas. Anal. 21(4) (2025) 55-76.
[34] Z. L. Chuan, S.-F. Fam, Q. H. Lee, J. S. Kok, M. N. B. M. Azam, Data Anal. Appl. Math. 3(2) (2022) 1-12.

[35] Agricultural Training Institute, Rice Competitiveness Enhancement Fund, https://www.da.gov.ph/rice-competitiveness-enhancement-fund/
[36] Ministry of Agriculture, Fisheries and Agrarian Reform, Rice Program, https://mafar.bangsamoro.gov.ph/rice-program/
[37] Consultative Group on International Agricultural Research, Low-emission Technologies Transform Vietnam's Rice Sector, Annual Performance Report 2020.
[38] O. Nguyen, Vietnam Investment Review (2023, Nov. 30).
[39] V. T. Nhung, Rikolto (2025, June 17).