ENHANCED KNOWLEDGE IN SCIENCES AND TECHNOLOGY
e-ISSN: 2773-6385

EKST

Vol. 6 No. 1 (2026) 160-168
https://publisher.uthm.edu.my/periodicals/index.php/ekst

Prediction of Rice Production in Malaysia Using State Space Model

Hanani Farinah Honorius[1], Kek Sie Long[1*]

[1] *Department of Mathematics and Statistics, Faculty of Applied Sciences and Technology, UTHM Kampus Cawangan Pagoh, Hab Pendidikan Tinggi Pagoh, KM 1, Jalan Panchor, 84600 Pagoh, Muar, Johor, MALAYSIA.*

*Corresponding Author: slkek@uthm.edu.my
DOI: https://doi.org/10.30880/ekst.2026.06.01.020

Article Info

Received: 4 January 2026
Accepted: 21 January 2026
Available online: 9 July 2026

Keywords

State Space Model, Rice Production Prediction, Parameter Estimation, Mean Square Error (MSE), Time Series Analysis, Forecasting

Abstract

Rice is not only a staple food in Malaysia but also the primary agricultural product, influencing the food security, employment rate, and economic stability. Hence, an accurate prediction of rice production can help in farm planning and policy construction. However, rice production is constantly evolving due to various factors, including climate change, uncertain pricing, and limited technological advancements. This paper proposes a state-space model to predict rice production, capturing the underlying dynamics of rice production data. A historical rice production dataset from 1980 to 2019 was analysed, and descriptive statistics were calculated to understand the trend in rice production. The state space model is formulated after standardizing the data. The model parameters are estimated through an optimization process to reduce the prediction error. The accuracy of the proposed model is evaluated using a mean square error (MSE) metric. The results showed that the proposed model accurately fit the historical rice production data, with the smallest MSE of 1.96×10^{-6}, compared to the simple moving average and the exponential smoothing method. Therefore, the state space model demonstrates its effectiveness in predicting rice production, supporting data-driven decision-making in agricultural planning and policy formation.

1. Introduction

Rice or *Oryza sativa* is a staple food for most of the world's population, especially in Asia and Africa [1]. In Malaysia, rice has not only become the most important staple food, but it has also become the main product in the agricultural sector that influences Malaysia's food security, employment rate, and economic stability. Furthermore, rice also defines a key feature of Malaysian culture and identity [2], where Malaysian dishes such as *nasi lemak* and *nasi goreng* are well-known among tourists. In consideration of the increasing population each year and the unpredictability of the weather, a consistent and sufficient rice production is the utmost priority. Regarding planning and decision-making, accurate forecasting of rice production can significantly assist the government, policymakers, and agricultural industries in meeting the demand and supply of rice. Hence, many studies have used statistical modelling with time series analysis to predict rice production in Malaysia. However, the majority of the existing studies mainly focus on conventional time series approaches such as trend analysis, regression models or ARIMA that assume stable relationships that may not effectively capture dynamic changes in agricultural production data. This indicates a research gap in the application of a more flexible modelling framework that can capture dynamic changes in Malaysia's rice production.

A state space model is an impressive method for analysing and modelling time series data [3]. State space models comprise two primary components: the state equation and the observation equation. The state equation can capture the dynamics of the underlying process and denote the hidden states that influence the observed time series. In contrast, the observation equation defines the relationship between the hidden states and the actual observation. The benefits of using state space models are evident in their flexibility in handling complex time series patterns, their ability to estimate the unobserved components of unseen variables, and their assistance in forecasting and prediction [4]. Compared to conventional methods, state space models offer a noticeable advantage by explicitly modelling the unobserved factors. For example, the underlying trends and gradual structural changes, which are common in agricultural systems affected by climate variability, policy and technological advancement. This property makes the state space model outstandingly convenient for predicting in the agricultural scope, where production data may change due to seasonal or external influences.

In this study, a continuous state space model is applied to Malaysia's annual rice production data from 1980 to 2019, collected from the Department of Statistics Malaysia (DOSM) [11]. This study will rely on the historical rice production data to analyse the trend of rice production. It will not include external factors such as weather or policy indicators. The state space model is appropriate because it can identify latent trends and time varying behaviours that are embedded in the historical production data. The model consists of two equations: the state equation, which will be used to capture the underlying trend in rice production, and the output equation, which links this hidden trend to observed production values. Since rice production data often shows year-to-year changes due to its seasonal, environmental and structural factors, the state space model, which is well known for its flexibility to capture the time-varying behaviour [6], has made it an ideal method for producing reliable and interpretable forecasts.

2. Methodology

This section provides a detailed explanation of the methodology used to model and estimate Malaysia's rice production using a state space model. The study follows a framework that includes data collection, data preprocessing, descriptive statistical analysis, model formulation, parameter estimation and optimization, validation, and performance comparison. Firstly, historical rice production data were normalized to ensure the stability of the model, given the large data values [7]. Then, a continuous state space model was formulated to capture the underlying dynamics of rice production. Parameter estimation is performed through an optimization process to minimize prediction error. Lastly, the performance of the proposed state space model is evaluated using Mean Squared Error (MSE) and compared with Simple Moving Average (SMA) and Exponential Smoothing Method (ESM).

2.1 Data Collection

The data used in this study were annual rice production in Malaysia from 1980 to 2019. This dataset was obtained from the Department of Statistics Malaysia (DOSM). The dataset contains a total of 40 observations that were measured in metric tons.

2.2 Data Preprocessing

Before the modelling process, the data were normalized to ensure the model's stability [8]. Since the rice production values are large, the standardization was applied to the data using the formula shown in Equation (1),

$$y_s = \frac{y-\mu}{\sigma} \tag{1}$$

where:
y = original rice production data
μ = mean of the data
σ = standard deviation

2.3 Descriptive Statistical Analysis

Descriptive analysis was conducted to understand the characteristics and behaviour of Malaysia's rice production data. The descriptive measures that were considered in this study are mean, minimum value, maximum value and standard deviation of rice production. The mean value represents the average of rice production over the year. Minimum and maximum values indicate the range of the production levels. Then, the standard deviation reflects the variability of the data over time.

2.4 State Space Model Formulation

The state space model was selected due to its ability to capture the underlying behaviour of the data through its hidden state. In this study, a linear state space model was used to predict Malaysia's rice production over time.

The relationship between the state variable and the observed rice production is assumed to be linear. The model did not include the random noise in both the state and observation equations.

The proposed state-space model consists of two main components which are the state equation and observation equation. The formula of the state equation is shown in Equation (2) and the observation equation in Equation (3),

$$\dot{x} = Ax + b \tag{2}$$

$$y = Cx + d \tag{3}$$

where x is the state variable and y is the output variable, while A is the state transition matrix, b is the model constant vector, C is the output coefficient matrix and d is the output constant vector. Here, $\dot{x}$ is the state rate of change over time t. In this paper, $x(t)$ represents the state variable, while $\hat{x}(t)$ represents its estimated value. The observed rice production data is denoted by $y(t)$, and the model output is denoted by $\hat{y}(t)$.

2.5 Parameter Estimation and Optimization

Parameter estimation is an important step to ensure that the proposed state space model can accurately capture the behaviour of rice production data. In this study, parameters in the state space model are the state transition matrix A, the constant vector b, the output coefficient matrix C and the output constant vector d. The model's parameter was determined by considering a least squares optimization problem [5] as shown in Equation (4),

$$\mathit{minimize} \qquad J(x) = \frac{1}{2}\left(y - y_s\right)^T\left(y - y_s\right) \tag{4}$$

where J is the objective function and y_s is the historical data, whereas y is the model output measured from the state equation (2). The aim is to minimize the gap between the observed rice production data and the output generated by the model.

Minimizing this function will make the model output fit the historical data closely. To minimize the objective function, a gradient-based optimization method was employed. The gradient of the objective function with respect to the model parameters [5] was considered,

$$\frac{\partial J}{\partial A} = C^T\left(\hat{y} - y_s\right)\left(\hat{x}^T\right) \tag{5}$$

$$\frac{\partial J}{\partial b} = C^T\left(\hat{y} - y_s\right)\left(1^T\right) \tag{6}$$

$$\frac{\partial J}{\partial C} = \left(\hat{y} - y_s\right)\left(\hat{x}^T\right) \tag{7}$$

$$\frac{\partial J}{\partial d} = \left(\hat{y} - y_s\right)\left(1^T\right) \tag{8}$$

where $1 = (1,1, \dots ,1)^T$, $\hat{x}$ is the state estimate and $\hat{y}$ is the output estimate. By using these gradients, we update the model parameters [5] to be estimated as follows,

$$A^{i+1} = A^i - \alpha_1\left(\frac{\partial J}{\partial A}\right)^i \tag{9}$$

$$b^{i+1} = b^i - \alpha_2\left(\frac{\partial J}{\partial b}\right)^i \tag{10}$$

$$C^{i+1} = C^i - \alpha_3\left(\frac{\partial J}{\partial C}\right)^i \tag{11}$$

$$d^{i+1} = d^i - \alpha_4\left(\frac{\partial J}{\partial d}\right)^i \tag{12}$$

where α_1, α_2, α_3, and α_4 are learning rates, and i is the iteration number.

For the first iteration, the model parameters and the state variable will be given an initial value. Using these initial values, the state equation is evaluated to obtain the initial state estimate. Then, it will be used in the observation equation to obtain the model output. The difference between the observed rice production and the proposed model output is calculated and used to evaluate the objective function. Based on this, the gradients of the objective function are obtained and used to update the model parameters and the state variable. The process is repeated in the next iteration.

Then, the optimal model parameter can be obtained once the first-order necessary conditions [5] shown in Equation (13) are satisfied.

$$\frac{\partial J}{\partial A}=0, \frac{\partial J}{\partial b}=0, \frac{\partial J}{\partial C}=0, \frac{\partial J}{\partial d}=0 \tag{13}$$

In addition, we want to find an optimal state estimate, which updates the model solution to approach the historical data. The state estimate can be calculated from

$$\hat{x}^{i+1}=x^i-\alpha_5\left(\frac{\partial J}{\partial x}\right)^i \tag{14}$$

where α_5 is the learning rate and x is the solution to the state equation (2). The gradient with respect to the state variable is given by

$$\frac{\partial J}{\partial x}=C^T\left(\hat{y}-y_s\right) \tag{15}$$

and once the convergence is achieved, the optimal state estimates the results when the first-order necessary condition as shown in Equation (16) is presented.

$$\frac{\partial J}{\partial x}=0 \tag{16}$$

2.6 Prediction and Accuracy Evaluation

After parameter estimation, the prediction of the proposed state space model was implemented. By using the estimated parameters, the proposed state space model will generate an output of the predicted rice production. This process uses a modified state equation (17) and measured output (18) as follows,

$$\hat{x}_k=e^{A\Delta t}x_k-\left(I-e^{A\Delta t}\right)A^{-1}b \tag{17}$$

$$\hat{y}_k=C\hat{x}_k+d \tag{18}$$

where $\hat{x}_k$ and $\hat{y}_k$ are the state estimate and the output estimate at time step k, respectively.

To evaluate the performance of the proposed model, Mean Square Error (MSE) was used to measure the prediction accuracy. The formula of MSE is as shown in Equation (19),

$$MSE=\frac{1}{N}\sum_{k=1}^{N}\left(\hat{y}_t-y_t\right)^2 \tag{19}$$

where y_t is the actual number of rice production at year t, $\hat{y}_t$ is the predicted number of rice production in year t, and N is the actual rice production. This metric determines the model's accuracy in fitting the historical data, where a lower MSE indicates a better model accuracy [10].

2.7 Model Validation and Comparison

Model evaluation is conducted to assess the reliability of the proposed state space model when applied to unseen data [9]. The validity of the proposed state space model proves whether the proposed state space model can also be used in modelling and predicting, instead of just simply fitting the data. In this study, the MSE value after model estimation was compared with the MSE validation to see its capability.

Other than comparing the MSE value of validation with estimation, MSE validation was also compared with other time-series statistical methods which are the Simple Moving Average (SMA) and the Exponential Smoothing Method (ESM). This became a benchmark to figure out if the proposed model can perform as well as other methods.

3. Result and Discussion

This section discusses the results obtained from this study. The descriptive analysis results were the first to be discussed to explain an impression of rice production in Malaysia. Then, the representation of the estimation and prediction results obtained from the proposed model. After that, the result of proposed model was also discussed. Lastly, the finding of validation of the proposed model through comparison.

3.1 Descriptive Analysis of Paddy Production Data

Table 1 shows that rice production in Malaysia has a moderate variability where the value of the production has an average of 1.420 million tons. This reflects that the trend of rice production is smooth and has a gradual change rather than sudden fluctuations throughout 1980 to 2019. Since the rice production data is large, normalization was applied to ensure numerical stability and consistency of the model.

Table 1 Descriptive statistics of Malaysia's annual rice production

Statistic	Value (Tons)
Mean	1,416,594.83
Minimum	1,010,279
Maximum	1,876,922
Standard Deviation	225,029,61

Fig. 1 illustrates the normalized annual rice production throughout the year. The overall trend shows that there is a gradual increase in rice production. However, there are still some fluctuations that occur in specific years. This trend confirms that rice production has generally improved each year.

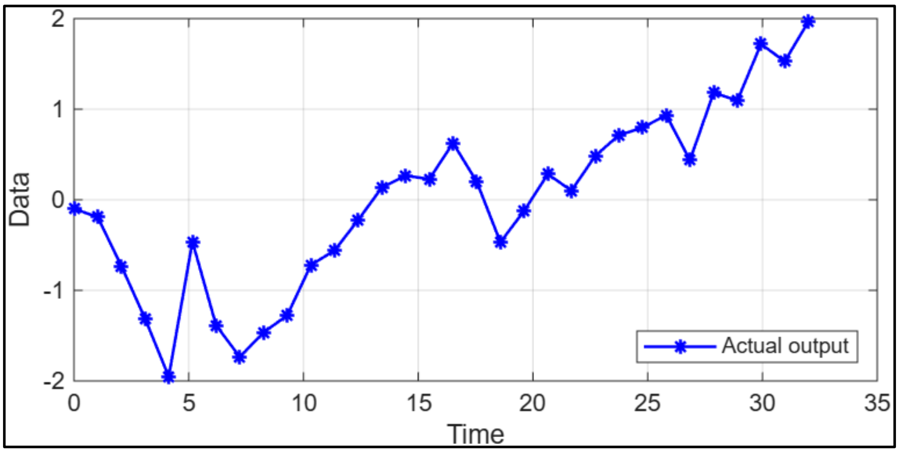

Fig. 1 Normalized rice production over time

3.2 State Space Modul Output

Fig. 2 shows the initial output of the state space model after parameter estimation. This model converges smoothly over time, which indicates that the model is stable. Therefore, Fig. 2 mainly shows the stability of the state space model, while the accuracy of the model is evaluated using Mean Square Error (MSE) and by comparing the estimated model output and the actual rice production data output, as shown in Fig. 3.

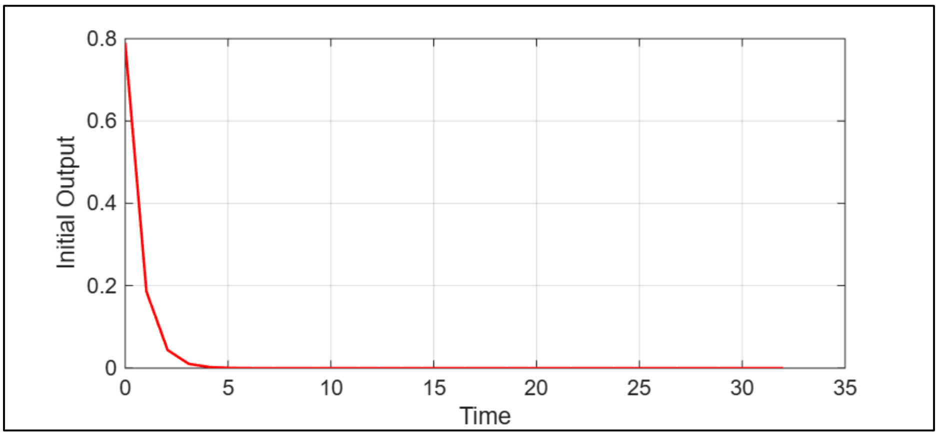

Fig. 2 State space model output

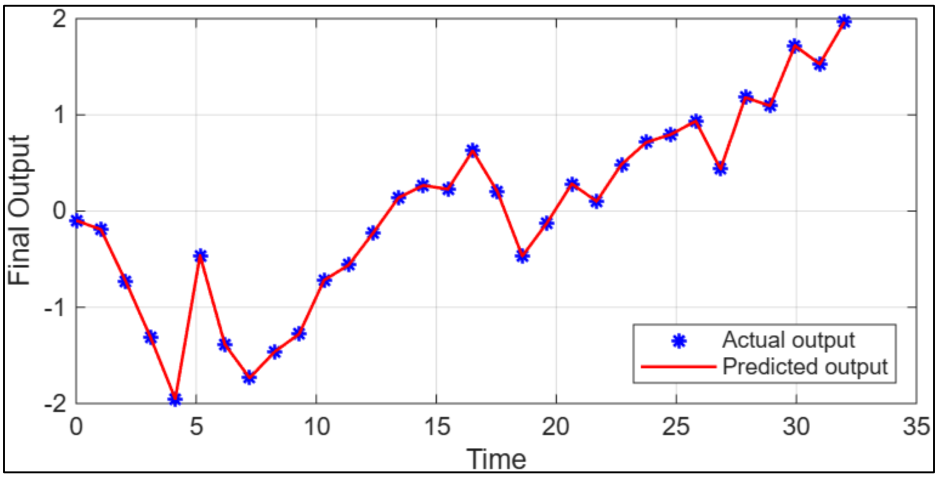

Fig. 3 Comparison of the state space model output and the actual data

Fig. 3 shows a comparison between the actual rice production data and the estimated output from the proposed state space model. The blue markers represent the actual rice production data, while the red line represents the estimated output from the proposed state space model. Fig. 3 illustrates that the estimated output closely follows the actual data throughout the year. This indicates that the proposed state space model had successfully captured the underlying trend and dynamic behaviour of rice production in Malaysia. However, there are still minor gaps between the estimated and actual output when it is looked at closely. These differences are relatively small and this is expected due to the presence of variability in real-world agricultural data. Generally, the close alignment between both curves demonstrates the effectiveness of the proposed state space model and is supported by a very small value of MSE that is shown in Table 2 and Fig. 4.

Table 2

Table 2 Descriptive statistics of Malaysia's annual rice production

Data	Number of Iteration	Elapsed Time (s)	Initial MSE	Final MSE
Rice Production (Normalized)	85	0.008852	0.9997	2.36×10^{-9}

Table 2 summarizes the parameter estimation performance of the proposed state space model. The result shows that the curve converged after 85 iterations within 0.008852 seconds. Therefore, after estimation, the MSE value reduces significantly from 0.9997 to 2.3640×10^{-9}. This value clearly displays that the parameter estimation was essential in helping to minimize the proposed state space model error.

Fig. 4 illustrates the prediction of error between actual rice production data and the estimated output from state space model. The figure shows that the curve is very close to zero, and the error value remains small

Enhanced Knowledge in Sciences and Technology Vol. 6 No. 1 (2026) p. 160-168

throughout the year. This indicates that the difference between the two curves is minimal. A small value of the prediction error indicates that the estimated output from the state space model follows the actual trend in rice production data. Hence, the result of MSE shown in Fig. 4 validates that the proposed state space model is capable of representing the historical rice production data.

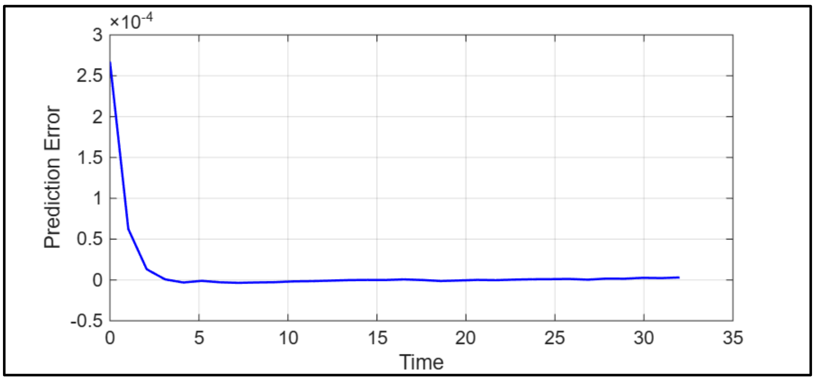

Fig. 4 Prediction error of state space model after parameter estimation

The estimated output of rice production that was obtained from the proposed state space model shows an increasing trend over the years. In agricultural planning perspective, this information can assist them in estimating future rice supply and preparing appropriate production strategies. While in policy decision making, this forecast can support any decision that relate with import planning, rice subsidy and food security management.

3.3 State Space Model Validation

Validation Model performance evaluated by validation proved that the model does not simply fit the historical data during the estimation process but is also reliable for modelling and predicting rice production data. A small value of validation MSE means that the model can maintain its predictive performance and does not suffer from severe overfitting.

Fig. 5 shows that the value of MSE validation was very small and remained below 0.001 throughout the year. This indicates that the difference between the actual rice data and the predicted output from the proposed model was little.

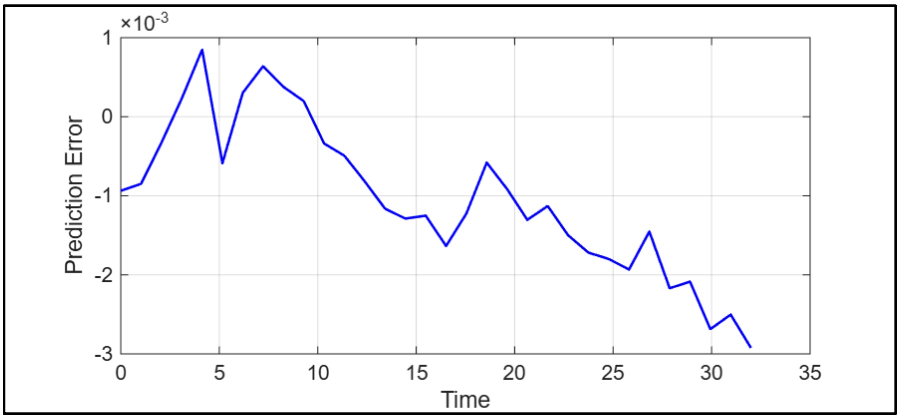

Fig. 5 MSE Validation of state space model

Table 3 presents the value of the validation SME for the proposed state space model together with the Simple Moving Average (SMA) and the Exponential Smoothing Method (ESM). It vividly shows that the proposed state space model achieves the lowest MSE value among the other compared methods. This demonstrates that the proposed state space model's ability to capture the dynamic behaviour of rice production is higher. Therefore, the validation results confirm that the proposed state space model is reliable and suitable for modelling and predicting rice production data. This enhance the trustworthiness of the proposed model as a tool that can be used to support agricultural planners and policy maker in evaluating future rice production situations.

Table 3 Comparison of MSE values

Method	MSE Value
State Space Model (SSM)	1.96×10^{-6}
Simple Moving Average (SMA)	0.0603
Exponential Smoothing (ESM)	0.0383

4. Conclusion

In conclusion, this study proposed a state space model to predict rice production in Malaysia. This study proved that the proposed model is suitable for modelling and predicting rice production trends. During the parameter estimation process, MSE value was found to be decreased to the smallest value possible and the model curve converged. This indicates that the estimated parameters were correct and can be used for the model. Hence, the stability of the proposed model was also confirmed when the outcome of the curve was smooth and converged. During model validation process, the result of the MSE validation attained from the proposed state space model was compared to the MSE value of the Simple Moving Average (SMA) and the Exponential Smoothing Method (ESM). MSE value of the proposed state space model was the smallest among the three methods. This means that the proposed state space model performs better than the SMA and the ESM in this study. Even though the MSE validation value is higher than the value of MSE at the final estimation, the validation MSE still remained small and stayed within the acceptable range. This matter is expected since we know that the MSE validation was calculated by using the unseen data while the final MSE during the estimation process was calculated using the data used in the optimization process. Thus, the objectives of this study were successfully achieved and proved that the state space method was reliable for agricultural production forecasting. Nevertheless, this study is limited to univariate modelling that only use historical rice production data and did not include climate, economy or policy factors. Future study may consider to extend the proposed model by adding variables to improve the model forecasting performance and can provide a better insight for agricultural planning and policy making.

Acknowledgement

The author would like to express sincere gratitude to Universiti Tun Hussein Onn Malaysia for the support and facilities provided throughout the completion of this study.

Conflict of Interest

The authors declare that there is no conflict of interest regarding the publication of the paper.

Author Contribution

The authors confirm contribution to the paper as follows: ***study conception and design:*** *Hanani Farinah Honorius, Kek Sie Long;* ***data collection:*** *Hanani Farinah Honorius;* ***analysis and interpretation of results:*** *Hanani Farinah Honorius, Kek Sie Long;* ***draft manuscript preparation:*** *Hanani Farinah Honorius, Kek Sie Long. All authors reviewed the results and approved the final version of the manuscript.*

References

[1] World Population Review. (2025). *Rice consumption by country*. World Population Review. Retrieved December 21, 2025, from https://worldpopulationreview.com/country-rankings/rice-consumption-by-country

[2] Che Omar, S., Ahmad Shaharudin, A. A., & Tumin, S. A. (2019). *The status of the paddy and rice industry in Malaysia*. Khazanah Research Institute. https://www.krinstitute.org/publications/the-status-of-the-paddy-and-rice-industry-in-malaysia

[3] Durbin, J., & Koopman, S. J. (2012). *Time series analysis by state space methods* (2nd ed.). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199641178.001.0001

[4] Fiveable. (2024, July 23). 12.1 Introduction to state-space models – Intro to Time Series. https://fiveable.me/intro-time-series/unit-12/introduction-state-space-models/study-guide/Z7tvPs7qu6wkSoHP

[5] Ma, X., & Ding, F. (2015). Gradient-based parameter identification algorithms for observer canonical state space systems using state estimates. *Circuits, Systems, and Signal Processing*, 34(5), 1697–1709. https://doi.org/10.1007/s00034-014-9911-5

[6] IBM. (2025). *What are state space models?* IBM — Think. Retrieved December 21, 2025, from https://www.ibm.com/think/topics/state-space-model

[7] Hicks, S.C., Irizarry, R.A. *quantro*: a data-driven approach to guide the choice of an appropriate normalization method. *Genome Biol* 16, 117 (2015). https://doi.org/10.1186/s13059-015-0679-0
[8] Barabanov, N. E., Ortega, R., & Astolfi, A. (2005). *Is normalization necessary for stable model reference adaptive control?* IEEE Transactions on Automatic Control, 50(9), 1384–1390. https://doi.org/10.1109/TAC.2005.854625
[9] Cerqueira, V., Torgo, L., & Mozetic, I. (2019). *Evaluating time series forecasting models: An empirical study on performance estimation methods* (arXiv:1905.11744) [Preprint]. arXiv. https://doi.org/10.48550/arXiv.1905.11744
[10] Encord. (2025). *Mean Square Error (MSE).* Retrieved December 22, 2025, from https://encord.com/glossary/mean-square-error-mse/
[11] Department of Statistics Malaysia (2019). *Principal Statistics of Paddy, Malaysia (1980–2019)* [Dataset]. Malaysian Open Data Portal. https://archive.data.gov.my/data/dataset/principal-statistics-of-paddy-malaysia