APPLICATIONS OF
MODELLING AND SIMULATION
http://arqiipubl.com/ams eISSN 2600-8084 VOL 9, 2025, 315-323

Deep Learning Python-Based Time-Series Model for Oil Palm Yield Prediction

Yuhao Ang[1], Helmi Zulhaidi Mohd Shafri[2*], Yang Ping Lee[3], Shahrul Azman Bakar[3], Haryati Abidin[3], Mohd Umar Ubaydah Mohd Junaidi[4], Hwee San Lim[5], Rosni Abdullah[6], Yusri Yusup[7], Syahidah Akmal Muhammad[7] and Teh Sin Yin[8]

[1]Faculty of Sustainable Agriculture, Universiti Malaysia Sabah Sandakan Campus, Locked Bag No. 3, 90509 Sandakan, Sabah, Malaysia
[2]Department of Civil Engineering and Geospatial Information Science Research Centre (GISRC), Faculty of Engineering, Universiti Putra Malaysia (UPM), 43400 Serdang, Selangor, Malaysia
[3]Geoinformatics Unit, FGV R&D Sdn Bhd, FGV Innovation Centre, PT35377, Lengkuk Teknologi, 71760 Bandar Enstek, Negeri Sembilan, Malaysia
[4]FGV Agri Services Sdn Bhd, Level 9, West, Wisma FGV, Jalan Raja Laut, 50350 Kuala Lumpur, Malaysia
[5]School of Physics, Universiti Sains Malaysia (USM), 11800 Gelugor, Penang, Malaysia
[6]School of Computer Sciences, Universiti Sains Malaysia (USM), 11800 Gelugor, Penang, Malaysia
[7]School of Industrial Technology, Universiti Sains Malaysia (USM), 11800 Gelugor, Penang, Malaysia
[8]School of Management, Universiti Sains Malaysia (USM), 11800 Gelugor, Penang, Malaysia

*Corresponding author: helmi@upm.edu.my

Submitted 19 April 2025; Revised 11 June 2025; Accepted 04 July 2025; Available online 22 July 2025.

Abstract: The application of time-series analysis in agricultural yield forecasting has gained much attention in recent years with the use of machine learning models. However, the utilization of Python-based time-series models for predicting oil palm yield remains limited. This study aims to explore the potential of two Python-based time-series models, Neural Prophet and Facebook Prophet, as alternatives to the conventional Autoregressive Integrated Moving Average (ARIMA) model for oil palm yield prediction. The study utilized historical yield data from two oil palm estates, Estate A and Estate B, covering a total of 100 data points from 2015 to 2022 on a monthly basis. The results demonstrate that the Neural Prophet and Prophet models outperformed the ARIMA model in terms of predictive accuracy. For Estate A, the Neural Prophet model achieved the highest accuracy, with a Mean Absolute Error (MAE) of 0.16, a Root Mean Square Error (RMSE) of 0.18, and a Mean Absolute Percentage Error (MAPE) of 0.14. Similarly, in Estate B, the Neural Prophet model obtained an MAE of 0.17, an RMSE of 0.21, and a MAPE of 0.10. The superior performance of the Neural Prophet model can be attributed to its ability to capture the complex patterns and nonlinear relationships inherent in the time-series data, owing to the adoption of deep learning principles. The fast implementation and robust forecasting capabilities of the Neural Prophet and Fb Prophet models make them viable alternatives to the conventional ARIMA model for oil palm yield prediction. The time-series predictive models developed in this study can assist plantation management in making informed decisions concerning yield forecasting, which is crucial for effective management of inputs, ultimately leading to cost optimization and enhanced sustainability in oil palm cultivation.

Keywords: ARIMA; Facebook prophet; Neural prophet; Oil palm; Time-series analysis; Yield prediction.

1. INTRODUCTION

Oil palm, *Elaeis guineensis*, is a valuable cash crop that is important for the production and demand of foods, supplying about 40% of all traded vegetable oil [1]. Generally, palm oil can be used to make cooking oil, which is a source of foods that are quite demanding for those people who live in low-income countries since the price for palm oil is relatively low compared to other vegetable oils [2]. Besides, oil palm cultivation is much more productive than other vegetable oil crops because one hectare of oil palm planting typically generates 3.3 tonnes per hectare, which is significantly four to eight times higher than other vegetable oil crops such as soybean, rapeseed and sunflower [3]. The prediction of oil palm yield plays an important role in optimising the production for plantation management. Accurate yield forecasting is crucial for sustaining oil palm production, particularly in the face of environmental factors like extreme weather events such as El Niño [4]. The shifting of climate conditions has exacerbated production and led to declines in oil palm yield [5]. Therefore, conducting rapid yield forecasting is essential for effectively monitoring the dynamic changes in oil palm yield.

In recent years, the application of time-series analysis in agricultural yield prediction has gained significant attention due to its ability to capture temporal dependencies and patterns inherent in the data [6], [7], [8]. Autoregressive Integrated Moving Average (ARIMA) is the common model used in agricultural applications, especially crop yield prediction [9], [10]. The recent development of time series forecasting for agricultural yield prediction mainly focuses on advanced machine learning and optimised hybrid time-series models. The XGBoost and ARIMA models were used to forecast annual rice production in Bangladesh from 1961 to 2020 [11]. The results showed that the XGBoost model outperformed the ARIMA model, with the Mean Absolute Percentage Error (MAPE) of the test set being 5.38% for XGBoost compared to 7.23% for ARIMA. [12] proposed a Genetic Algorithm (GA)-based ARIMA model to predict crop yield. Results show that the GA-based ARIMA model achieved superior performance compared to the Reinforced Random Forest Algorithm (RRFA) model, with a Mean Absolute Error (MAE) of 0.80%, Root Mean Square Error (RMSE) of 3.75%, and overall accuracy of 80%.

Time-series Python libraries have demonstrated considerable efficacy in forecasting tasks and are extensively applied across various fields [13], [14], [15]. Using the time-series Python library for time series forecasting offers distinct advantages compared to traditional models such as ARIMA and machine learning-based approaches. This library provides a more user-friendly and streamlined implementation process, making it accessible to researchers and practitioners with varying levels of technical expertise. In addition, a time-series Python library creates a model that is able to handle the burden of manual feature engineering and enhance the accuracy of predictions across various domains. Their flexibility in capturing complex patterns, including nonlinear trends and sudden changes, surpasses the capabilities of traditional models. The comparison of Fb Prophet and ARIMA were made on the time-series forecasting of coronavirus disease [16]. The Fb Prophet model has higher accuracy in forecasting coronavirus disease for confirmed cases, with 91% precision. In comparison, the ARIMA model has the lowest accuracy, with 26% for the confirmed cases. The Fb Prophet model was applied to evaluate the predicted growth of wheat yield on historical data [17]. The result showed that the Fb Prophet model achieved the highest accuracy with a MAPE of 10.03% and an RMSE of 0.39 for summer wheat prediction when using yearly seasonality.

Recent studies focus on enhancing time-series Python-based model, particularly in deep learning, which has been proven to be effective in predicting the forecasting tasks in several fields [18] [19]. A study aimed at predicting ozone concentration was conducted using both the Prophet and Neural Prophet models [20]. The findings revealed that the Neural Prophet model significantly outperformed the traditional Prophet model, achieving a 70% improvement in accuracy. Specifically, the Neural Prophet model exhibited a Mean Absolute Error (MAE) of 1.97, a Mean Absolute Percentage Error (MAPE) of 6.41, and a Root Mean Square Error (RMSE) of 2.75. In contrast, the Prophet model recorded a MAE of 7.07, an MAPE of 21.85, and an RMSE of 7.38. Although few studies have adopted this approach, the studies on crop yield forecasting tasks are still limited by lack of experimentation and less adoption of advanced techniques. The study aims to address the gap in the application of advanced Python-based-time series models for predicting oil palm yields. We proposed neural prophet, a deep learning approach that enhances forecasting capabilities compared to traditional methods like ARIMA. The performance of this technique was subsequently evaluated and compared with ARIMA and Fb Prophet models. By optimizing a time-series predictive model specifically for oil palm yield forecasting, this study provides a robust framework that can assist plantation management in improving resource management and sustainability in oil palm cultivation.

2. MATERIALS AND METHODS

2.1 Study Area and Data

The study area (Figure 1) is located in two research plantations in Jerantut district, and it is within the state of Pahang, one of the states in Peninsular Malaysia. The historical yield data was recorded on a monthly basis and provided by the plantation management from 2015 to 2022. Yield data comprises two estates (Estates A and B), which have been used to develop time-series models for the latter prediction.

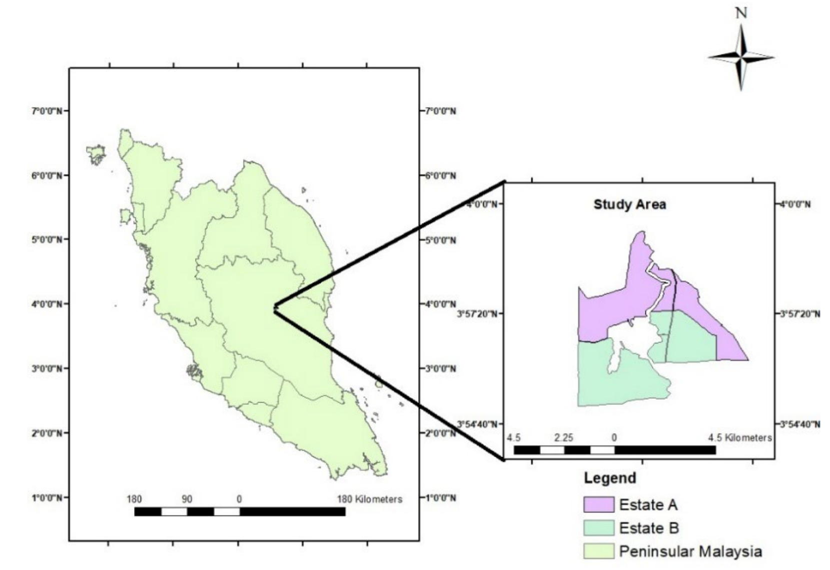


Figure 1. Study area.

2.2 Overall Methodology

The study begins with data input, where historical oil palm yield data from Estate A and B was collected. Next, in the pre-processing stage, the time series was plotted, and the autocorrelation function (ACF) and partial autocorrelation function (PACF) were analyzed. An Augmented Dickey-Fuller (ADF) test was then conducted to check for stationarity. In the processing stage, various models, including Neural Prophet, Facebook Prophet and ARIMA, were developed based on the pre-processed data. Finally, the models were evaluated for their performance and the prediction was made to forecast for the next 12-month period. The overall process (Figure 2) consists of data input, pre-processing, model development and evaluation.

2.3 Pre-processing

Checking the stationarity of the data is an important step when working with time series data before proceeding with any type of modelling. It is crucial that the data exhibits stationarity, which means that the mean, variance, and autocorrelation structure remain constant over time. Ensuring the absence of trends or seasonality in the data is necessary to avoid obtaining spurious results when forecasting time series. Plotting the time series can provide visual indications of patterns, trends and seasonality. An Augmented Dickey-Fuller test (ADF) was used to check for stationarity. The ADF test is a statistical test that examines the presence of a unit root in the data, which indicates non-stationarity [21]. It evaluates the null hypothesis that the data has a unit root against the alternative hypothesis of stationarity. By calculating the ADF test statistics and comparing it to critical values, we can determine the stationarity of the data. If the p-value associated with the test is below a specific significance level (e.g., 0.05), we reject the null hypothesis and conclude that the data is stationary. If the data is non-stationary, it can be transformed into a stationary form through differencing. Besides, Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots were generated and tested using monthly yield data in this analysis. The ACF plot measures the correlation between observations of a time series at different time lags. This can help identify the presence of any underlying patterns or dependencies within the data. The PACF plot shows the partial correlation between observations of a time series at different lags after removing the influences of the shorter lag periods. This helps to identify the direct relationship between observation and its lagged values.

2.4 Model Development

2.4.1 Neural Prophet

Neural Prophet is an advanced time-series forecasting Python package that builds on top of PyTorch and acts as an extension for the Facebook Prophet package developed by Facebook [22]. It uses a deep learning principle that combines artificial neural networks with the autoregressive model. It offers additional features such as automated differencing, enhancing the capabilities of Facebook's Prophet. Furthermore, it also leverages external variables and their future values to enhance forecast accuracy. By incorporating additional components such as regressive, autoregressive and lag terms, Neural Prophet is able to effectively capture underlying patterns present in crop data, including those involving more complex relationships [23]. The working mechanism of the Neural Prophet is presented in Equation (1):

$$Y_t = T(t) + S(t) + E(t) + F(t) + A(t) + L(t) \tag{1}$$

where $T(t)$ is trend at time, $S(t)$ is seasonal effect at time, $E(t)$ is event and holiday effect at time, $F(t)$ is regression effect, $A(t)$ is autoregression effects and $L(t)$ is lag effect.

The model was trained for 250 epochs, and the learning rate was set to 0.01, which enhances the convergence and improves overall training performance. Various combinations of epochs and learning rates were evaluated through a systematic trial-and-error procedure to identify the values that produced optimal performance. Adam optimizer was used to optimize the performance of the Neural Prophet model during the training process, integrating weight decay regularization with the optimization process to prevent overfitting [24]. The formula for the Adam optimizer is written in Equation (2).

$$w_t = -\frac{n}{\sqrt{v_t - \varepsilon}}\left(\frac{m_t}{\sqrt{\hat{u}_t + \varepsilon}} + \lambda w_t\right) \tag{2}$$

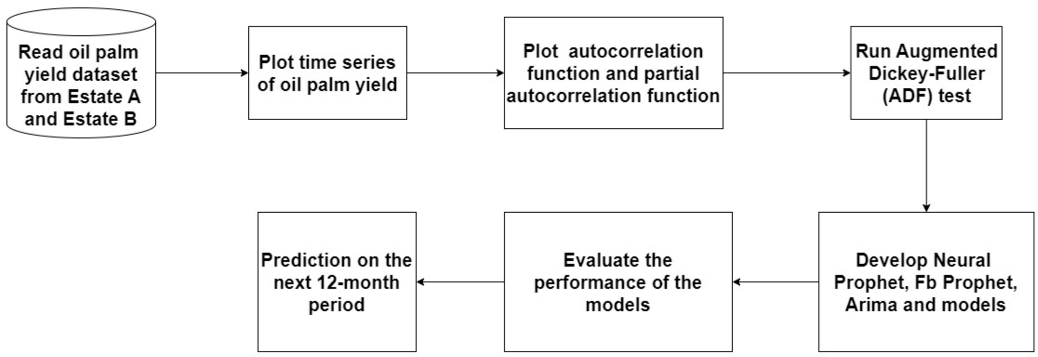


Figure 2. Methodology flowchart.

where w_t represents the update to the weights at time step t, n denotes the learning rate, m_t represents the first-order moment estimate of the gradients at time step t, $\hat{u}$ represents the exponentially decaying average of the squared gradient at time step t, v_t represents the exponentially decaying average of squared weights at time step t, λ represents the weight decay coefficient, and ϵ is a small value introduced for ensuring numerical stability.

2.4.2 Facebook Prophet (Fb Prophet)

Fb Prophet is an open-source library built by Facebook's core data science team in 2017 for time series forecasting [25]. The Fb Prophet approach is able to deal with missing data, outliers, and seasonality. Because of its simplicity and automatic process, the Fb Prophet approach recently gained popularity and has proven to develop a good time-series model, which has widely been applied in various fields [26], [27]. The Fb Prophet model includes several key components, such as trends, seasonality and holiday effects, that give it flexibility and predictive power. The Fb Prophet model can be represented mathematically in the Equation (3).

$$Y_t = g(t) + s(t) + h(t) + \varepsilon_t \tag{3}$$

where $g(t)$ describes a piecewise linear or logistic growth curve for modelling non-periodic changes in time series trend, $s(t)$ is periodic changes in weekly or yearly seasonality, $h(t)$ is the holiday effects as simple dummy variables, and ε_t is a white noise error term.

2.4.3 ARIMA

The ARIMA model is a widely used statistical technique for time series analysis and forecasting [28]. The ARIMA model combines three components: autoregression (AR), differencing (I), and moving average (MA). The AR component models the relationship between the current observation and past observations, capturing the autocorrelation structure of the time series. The differencing (I) component is used to remove trends or seasonality and achieve stationarity in the time series. The MA component accounts for the impact of past forecast errors on the current observation, capturing short-term fluctuations. The ARIMA model is denoted as ARIMA (*p, d, q*), where the parameters *p, d, and q* represent the orders of the AR, differencing, and MA components, respectively. These parameters are determined based on the characteristics of the time series, such as the autocorrelation structure and the presence of trends or seasonality. The equation of ARIMA is shown as follows:

$$Y_t = a + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \ldots + \beta_p Y_{t-p} \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_2 \varepsilon_{t-2} + \ldots + \phi_q \varepsilon_{t-q} \tag{4}$$

where Y_t is the current observation, Y_{t-1} are the lagged observations, are the past forecast errors where $j = 1, 2, \cdots, q$, while a, β and φ are the model parameters. The parameters p and q represent the orders of the AR and MA components, respectively, and were determined based on the characteristics of the time series.

The appropriate orders of the AR (p) and MA (q) terms were identified using the ACF and PACF plots, following the guidelines outlined by Box [29].

2.5 Performance Evaluation and Prediction for the Next 12-Month Forecast

Several performance metrics were used to provide quantitative measures of the accuracy and reliability of time series models. The data from 2022 was used to evaluate the performance of each model and make forecasts for the next 12 months. Various evaluation metrics used in the evaluation for this study are MAE, RMSE, and MAPE. MAE measures the average absolute difference between the predicted values and the actual values in a time series. It provides a straightforward interpretation of the average magnitude of errors. The RMSE metric penalizes larger errors more heavily compared to smaller errors. RMSE metric increases its sensitivity to outliers because of the large errors caused by those outliers. The MAPE provides the average percentage of the errors for that model, indicating the relative accuracy or fitness of the result with lower error percentages. These measures can be written in the equations below:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|\hat{y}_i - y_i\right| \tag{5}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{y}_i - y_i\right)^2} \tag{6}$$

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\hat{y}_i - y_i}{y}\right| \times 100 \tag{7}$$

where n shows the number of samples, $\hat{y}_i$ denotes the predicted yield value, y_i indicates the actual yield value, while $\hat{y}_i$-y_i indicates the error value. Lower values of these values (MAE, RMSE, and MAPE) indicate better fitness or higher accuracy of the model's results. A paired t-test was conducted to assess the statistical significance of difference in model performance among ARIMA, Facebook Prophet and Neural Prophet.

3. RESULT AND DISCUSSION

3.1 Trend Analysis Using Time Series Plot

Figures 3 and 4 show the trend of the monthly yield over time from 2015 to 2022 for Estates A and B, respectively. The yield lies between 0.68 tonnes per hectare and 2.23 tonnes per hectare for Estate A, while Estate B has a yield between 0.67 tonnes per hectare and 3.33 tonnes per hectare. There is an overall downward trend in the yield from 2015 to 2022. For Estate A, the highest yield was seen in September 2020, which recorded 2.23 tonnes per hectare. The lowest yield occurred in December 2019, which was 0.68 tonnes per hectare. For Estate A, the yield pattern shows a cyclical pattern, with periods of rising and falling yields. These cycles appear to have a duration of approximately one to two years.

The overall trend in the data appears to be more volatile and exhibits significant fluctuation in Estate B. There are several sharp spikes and drops in the yield across the 2015-2022 period. The highest yield was depicted in June 2020, which was 3.33 tonnes per hectare. The lowest yield (0.67 tonnes per hectare) was presented in September 2021. The cyclical pattern observed earlier is still present, but the cycles are more irregular in duration, ranging from approximately six months to two years.

The ACF plot Figure 5 (a) shows the correlation coefficients on the y-axis and the lags on the x-axis. Each bar on the plot represents the correlation coefficient at a specific lag. The PACF plot (Figure 5(b)) shows the partial correlation coefficients on the y-axis and the lags on the x-axis. Each bar on the PACF plot represents the partial correlation coefficient at a specific lag. The ACF plot (Figure 5(a)) shows a significant spike at lag 1 followed by a quick drop-off, meaning it is mostly an MA ($q = 1$) process. The PACF plot shows a significant spike at lag 1, indicating an AR ($p = 1$) process. Estate B also exhibits a similar significant spike at lag 1 for ACF (Figure 6(a)) and PACF (Figure 6(b)) plots. Therefore, the values of p and q are 1 and the value of d is 0 for both estates A and B.

Table 1 shows the result of the ADF test for Estate A and Estate B. For both Estate A and Estate B, the p-values are less than the typical significance level of 0.05 (5%). This means we can reject the null hypothesis and conclude that both time series are stationary. The more negative the ADF statistics, the stronger the evidence against the null hypothesis of non-stationarity. Both the ADF statistics are sufficiently negative, further confirming the stationarity of these time series for Estate A and Estate B.

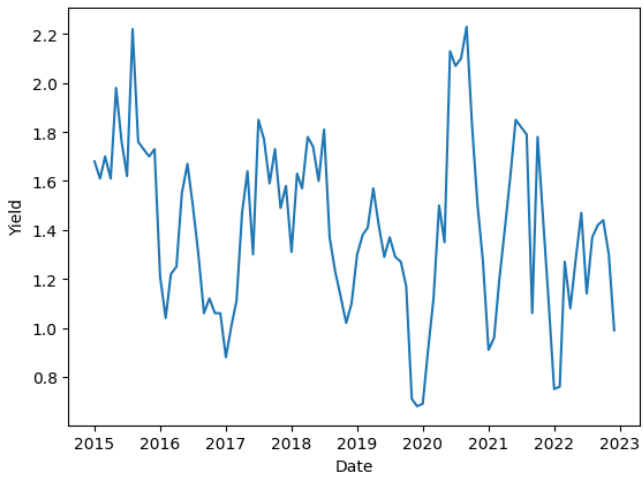


Figure 3. Time series plot for oil palm yield in Estate A (2015 - 2022).

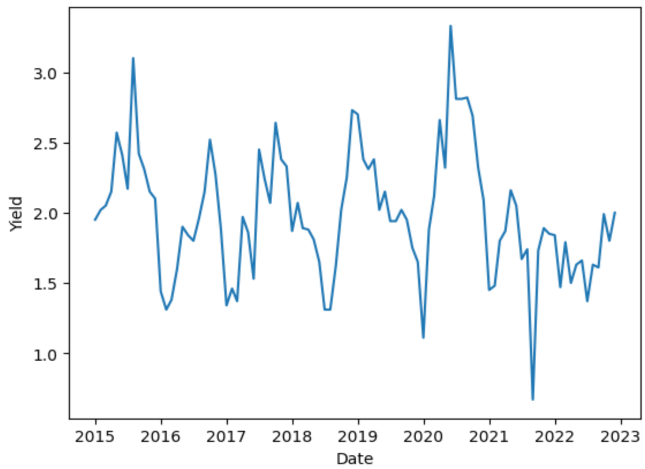


Figure 4. Time series plot for oil palm yield in Estate B (2015 - 2022).

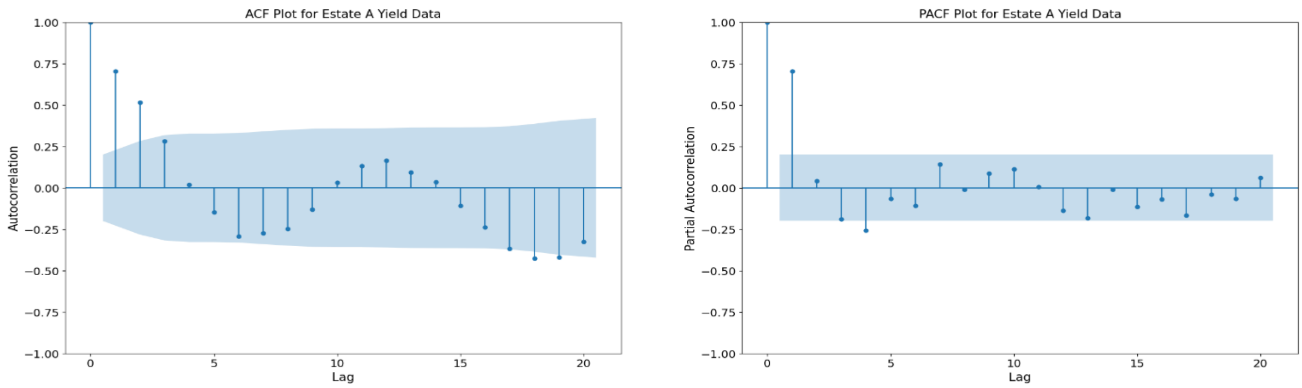


Figure 5. (a) Autocorrelation function plot (ACF plot) for Estate A; (b) Partial autocorrelation function plot (PACF plot) for Estate A.

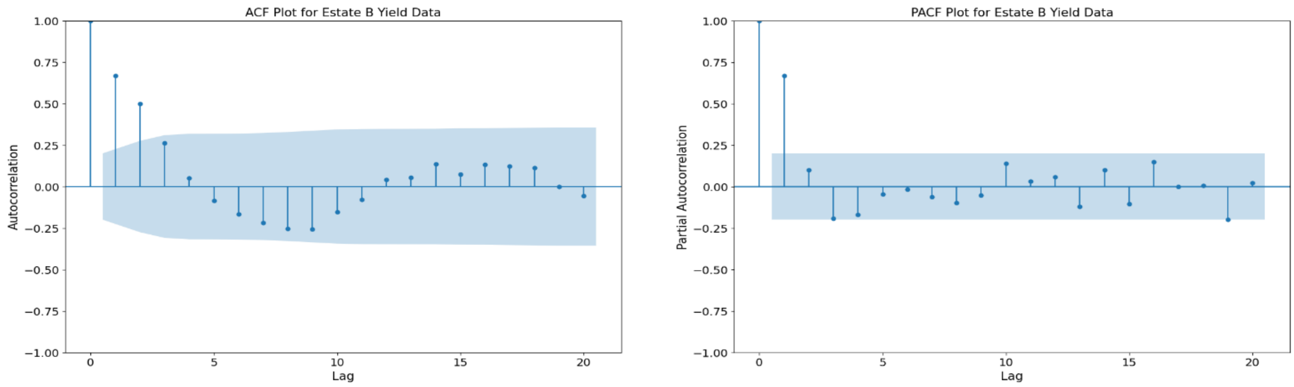


Figure 6. (a) Autocorrelation function plot (ACF plot) for Estate B; (b) Partial autocorrelation function plot (PACF plot) for Estate B.

Table 1. Augmented Dickey-Fuller (ADF) test.

Information	Estate A	Estate B
ADF Statistic	-4.722842	-3.970785
p-value	0.000076	0.001571

3.2 Evaluating Model Performance and Displaying the Next 12-month Forecast

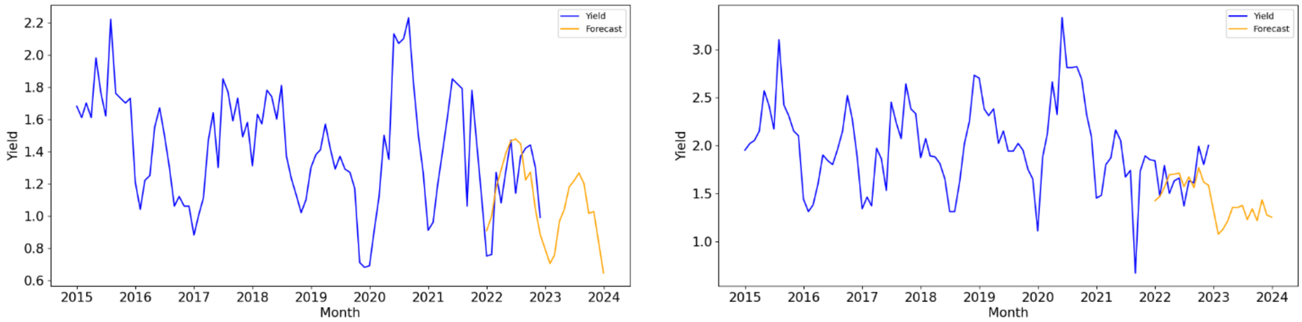
Table 2 presents the evaluation of time-series models for the oil palm yield prediction. Neural Prophet attained the best performance among the other models tested, both in Estate A and Estate B. In Estate A, Neural Prophet recorded a MAE of 0.16, a RMSE of 0.18, and a MAPE of 0.14. Meanwhile, in Estate B, Neural Prophet obtained a MAE of 0.17, a RMSE of 0.21, and a MAPE of 0.10. In comparison, the Fb Prophet model also demonstrated better performance than the ARIMA model. In Estate A, Fb Prophet had a MAE of 0.16, a RMSE of 0.20, and a MAPE of 0.15, which outperformed ARIMA with a MAE of 0.20, RMSE of 0.27, and MAPE of 0.21. A similar trend was observed in Estate B, where Fb Prophet recorded a MAE of 0.21, a RMSE of 0.24, and a MAPE of 0.13, compared to ARIMA with a MAE of 0.36, RMSE of 0.46, and MAPE of 0.19. Based on the analysis of statistical test, the neural prophet performed statistically significantly better than ARIMA with p-value less than 0.05. The result also shows there is no significant difference between Facebook Prophet and neural prophet since p-value larger than 0.05.

Figure 7 shows the Neural Prophet model that generates predicted yield for twelve months in Estate A and Estate B. The blue line represents the actual yield, while the orange line represents the predicted yield. The Neural Prophet model appears to perform better in forecasting the oil palm yield for Estate A, where it generally captures the overall trends and fluctuations. However, some discrepancies still exist in predicting the magnitude of the yield changes. For Estate A, the predicted yield in 2022 generally follows the trend of the actual yield. Still, it tends to underestimate the magnitude of the fluctuations, particularly towards the end of the year. In contrast, for Estate B, the model struggles significantly, consistently underestimating the extreme spikes and drops in the actual yield and failing to anticipate the sharp swings in the later years. The model struggled much more to forecast the yields for Estate B compared to Estate A, demonstrating the challenge of accurately predicting oil palm production, especially for estates with high variability.

The Fb Prophet model's (Figure 8) performance in forecasting the oil palm yields for Estate A and Estate B presents notable differences. For Estate A, the model appears to do reasonably well in capturing the overall trends and fluctuations, with the forecasted yield aligning relatively well with the actual yield, particularly in the short term. For 2022, the Fb Prophet model's forecasted yield aligns quite well with the actual observed yield, capturing the overall trend and direction of the yield changes throughout the year. However, the model still exhibits some limitations in accurately predicting the magnitude of the yield changes, struggling to fully anticipate the range of the fluctuations. In contrast, the Fb Prophet model fails to forecast the oil palm yield for Estate B, which exhibits sharp spikes and drops in the actual yield profile. The model struggles to fully capture the magnitude of these significant changes, often deviating quite substantially from the actual yield patterns. This suggests the Fb Prophet model has limitations in handling the complexity and unpredictability of the yield dynamics for Estate B when compared to the relatively more stable Estate A.

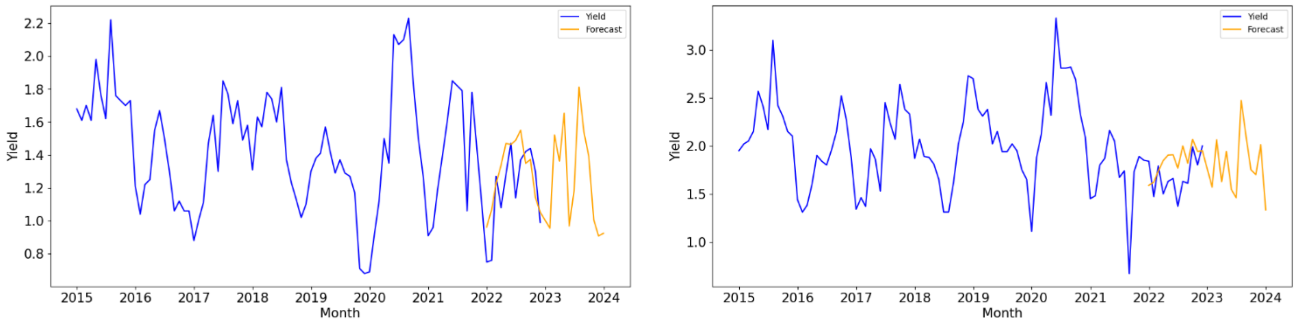
Table 2. Time-series models on the oil palm yield prediction.

Estates	A			B		
Models	MAE	RMSE	MAPE	MAE	RMSE	MAPE
ARIMA	0.20	0.27	0.21	0.36	0.46	0.19
Fb Prophet	0.16	0.20	0.15	0.21	0.24	0.13
Neural Prophet	0.16	0.18	0.14	0.17	0.21	0.10



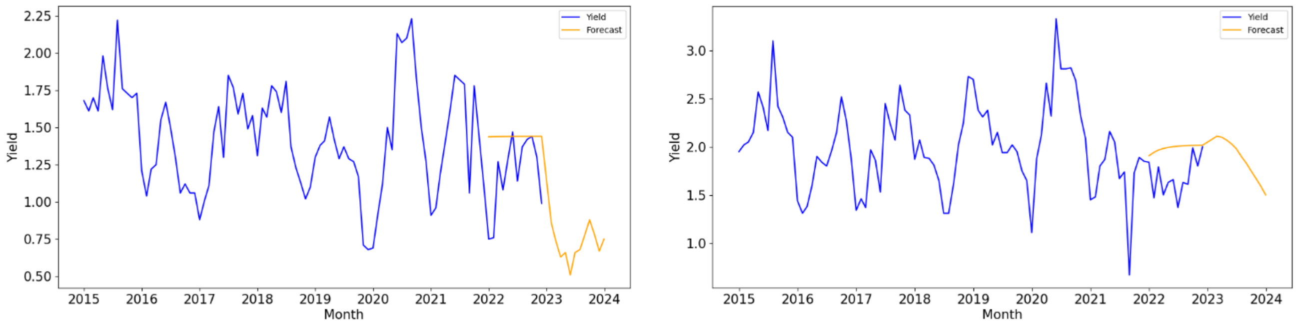
(a) (b)

Figure 7. Neural Prophet for: (a) Estate A; (b) Estate B.



(a) (b)

Figure 8. Fb Prophet model for: (a) Estate A; (b) Estate B.



(a) (b)

Figure 9. ARIMA model for: (a) Estate A; (b) Estate B.

Figures 9(a) and 9(b) show the actual and prediction trends generated by the ARIMA model. The ARIMA model fails to properly capture the oil palm yield patterns in both Estate A and Estate B. The model struggles to track the high volatility, and rapid fluctuations present in the actual yield data, frequently failing to predict the sharp spikes and dips. The model's forecasts also deviate significantly from the historical trends, projecting a much smoother pattern. This mismatch between the model's predictions and the real data indicates that the ARIMA approach is not flexible enough to adequately model the complex dynamics driving the oil palm yields over time.

Overall, Neural Prophet and Prophet models are able to forecast the oil palm yield for the next 12 months. However, the ARIMA model appears to have struggled in accurately forecasting the future trend of the oil palm yield compared to other models like Prophet and Neural Prophet, as evident from the significant divergence between the forecasted and actual yield values, particularly in the latter part of the years. Fb Prophet is proven effective in capturing seasonal factors, which account for various influences on crop yield [18]. Verification will be done with plantation management in the future to identify which factors influence the oil palm yield. The finding was in agreement with the work of [30], who proved that the Fb Prophet model is fast to implement and achieved superior performance over the ARIMA model for the time-series forecasting of daily new cases in Indonesia. The Fb Prophet model is able to discover the implicit seasonal effect that occurred previously.

The results of this study indicate that the Neural Prophet model achieved a MAE of 0.16, RMSE of 0.18 and MAPE of 0.14 for Estate A, outperforming both the Fb Prophet (MAE of 0.16; RMSE of 0.20 and MAPE of 0.15) and ARIMA (MAE of 0.20; RMSE of 0.27 and MAPE of 0.21). The findings corroborate with the study of [31], who proved the neural prophet model surpassed the traditional model in predicting yield of sugar beet with the R^2 of 0.95. Our study has successfully proven that the Neural Prophet achieved the best performance compared to the ARIMA and Fb Prophet model because of its ability to handle the complexity and unpredictability of yield dynamics. The Neural Prophet model that works under a neural network was able to better capture the complex patterns and the underlying dynamics of the time series data, especially in scenarios where traditional models struggled [32]. Even though the yield can be affected by different aspects, this study provides a quick prediction of the future projection of the yield in the near term. The study is limited by a lack of temporal data and insufficient generalization. Future studies will focus on adding more temporal data and applying the model to other estates, which can enhance the robustness and accuracy of the prediction model. Future studies will also incorporate other important factors, such as weather or management variables, and compare different advanced algorithms, including deep learning to assess model performance.

4. CONCLUSION

Overall, this study provides a useful framework for understanding the historical trends and anticipating the future performance of the oil palm yield, which can support decision-making and optimize operational strategies for plantations. Both Neural Prophet and Fb Prophet models are considered fast implementation and convenient, as they do not require any transformation prior to transformation to stationary time series before modelling and are able to handle complexities in the prediction of oil palm yield in terms of efficiency. The findings suggest that the Neural Prophet model is the most effective time-series approach for predicting oil palm yields among the models evaluated in Estate A and Estate B. In Estate A, the Neural Prophet model achieved the highest accuracy, with a MAE of 0.16, a RMSE of 0.18, and a MAPE of 0.14. Similarly, in Estate B, the Neural Prophet model demonstrated a MAE of 0.17, a RMSE of 0.21, and a MAPE of 0.10. Future studies will focus on the comparison and additional substitution of advanced algorithms to improve the overall model performance. Other factors will also be considered in future studies to increase the reliability of the prediction through verification with the plantation management, which could represent the current situation in the plantation. The ability to forecast the oil palm yield for the upcoming 12-month period can be valuable for plantation and inventory management and planning, contributing to effective decision-making for the oil palm industry.

ACKNOWLEDGEMENT AND FUNDING

This study is funded by the Malaysian Research University Network's (MRUN) Long Term Research Grant Scheme (LRGS), under the Ministry of Higher Education (MOHE) in Malaysia. The research program “A Big Data Analytics Platform for Optimizing Oil Palm Yield Via Breeding by Design” provided support for this project with grant number 203.PKOMP.6770007. Funding was given for this study under the project “Geoinformatics Data for Palm Oil Yield Prediction Using Machine Learning” (Vote No: 6300268-10801). The authors expressed gratitude to the team members from FGV R&D Sdn. Bhd. for their valuable expertise.

DECLARATION OF CONFLICTING INTERESTS

The authors declare no potential conflicts of interest with respect to the research and publication of this article.

REFERENCES

[1] D. J. Murphy, K. Goggin and R. R. M. Paterson, Oil palm in the 2020s and beyond: Challenges and solutions, *CABIAgriculture and Bioscience*, 2(1), 2021, 39.
[2] W. P. Aung, E. Bjertness, A. S. Htet, H. Stigum, V. Chongsuvivatwong, P. P. Soe, and M. K. R. Kjøllesdal, Fatty acid profiles of various vegetable oils and the association between the use of palm oil vs. peanut oil and risk factors for non-communicable diseases in Yangon Region, Myanmar, *Nutrients*, 10(9), 2018, 1193.
[3] E. Meijaard, T. M. Brooks, K. M. Carlson, E. M. Slade, J. Garcia-Ulloa, D. L. A. Gaveau, J. S. H. Lee, T. Santika, D. Juffe-Bignoli, M. J. Struebig, S. A. Wich, M. Ancrenaz, L. P. Koh, N. Zamira, J. F. Abrams, H. H. T. Prins, C. N. Sendashonga, D. Murdiyarso, P. R. Furumo, N. Macfarlane, R. Hoffmann, M. Persio, A. Descals, Z. Szantoi and D. Sheil, The environmental impacts of palm oil in context, *Nature Plants*, 6(12), 2020, 1418-1426.
[4] J. F. Khor, L. Ling, Z. Yusop, R. J. Chin, S. H. Lai, B. H. Kwan and D. W. K. Ng, Impact comparison of El Niño and ageing crops on Malaysian oil palm yield, *Plants*, 12(3), 2023, 424.
[5] A. Abubakar, M. Y. Ishak and A. A. Makmom, Nexus between climate change and oil palm production in Malaysia: a review, *Environmental Monitoring and Assessment*, 194(4), 2022, 262.
[6] Y. Wang, Q. Zhang, F. Yu, N. Zhang, X. Zhang, Y. Li, M. Wang and J. Zhang, Progress in research on deep learning-based crop yield prediction, *Agronomy*, 14(10), 2024, 2264.
[7] C. Chengzhi, W. Sha, D. Shengnan and C. Wenfang, Potential yield of potato under global warming based on an ARIMA-TR Model, *Potato Research*, 2024, 1-17.
[8] B. Pandey, A. Shukla and A. Khamparia, An optimized hybrid ARIMA-LSTM model for time series forecasting of agricultural production in India, In *Microbial Data Intelligence and Computational Techniques for Sustainable Computing*, A. Khamparia, B. Pandey, D. K. Pandey, and D. Gupta, Eds., Singapore: Springer, 47, 2024, 107-119.
[9] Pushpa, Chetna, Aditi and Urmil Verma, ARIMA and ARIMAX models for sugarcane yield forecasting in Northern Agro-climatic zone of Haryana, *Journal of Agrometeorology*, 24(2), 2022, 200-202.
[10] F. Shahrin, L. Zahin, R. Rahman, A. J. Hossain, A. H. Kaf and A. K. Abdul Malek Azad, Agricultural analysis and crop yield prediction of Habiganj using multispectral bands of satellite imagery with machine learning, *2020 11th International Conference on Electrical and Computer Engineering (ICECE)*, 2020, 21-24.

[11] M. Noorunnahar, A. H. Chowdhury and F. A. Mila, A tree based eXtreme Gradient Boosting (XGBoost) machine learning model to forecast the annual rice production in Bangladesh, *PLOS ONE*, 18(3), 2023, e0283452.
[12] Y. Guo, Integrating genetic algorithm with ARIMA and reinforced random forest models to improve agriculture economy and yield forecasting, *Soft Computing*, 28(2), 2024, 1685-1706.
[13] A. Hasnain, Y. Sheng, M. Z. Hashmi, U. A. Bhatti, A. Hussain, M. Hameed, S. Marjan, S. U. Bazai, M. A. Hossain, M. Sahabuddin, R. A. Wagan and Y. Zha, Time series analysis and forecasting of air pollutants based on prophet forecasting model in Jiangsu Province, China, *Frontiers in Environmental Science*, 10, 2022, 945628.
[14] Z. Z. Oo and S. Phyu, Time series prediction based on Facebook Prophet: A case study, temperature forecasting in Myintkyina, *International Journal of Applied Mathematics Electronics and Computers*, 8(4), 2020, 263-267.
[15] W. K. H. Wan Khairul Amir, A. B. Md Soom, A. Mat Jasin, J. Ismail, Aszila Asmat and R. Abdul Rahman, Sales forecasting using convolution neural network, *Journal of Advanced Research in Applied Sciences and Engineering Technology*, 30(3), 2023, 290-301.
[16] C. B. Aditya Satrio, W. Darmawan, B. U. Nadia and N. Hanafiah, Time series analysis and forecasting of coronavirus disease in Indonesia using ARIMA model and PROPHET, *Procedia Computer Science*, 179, 2021, 524-532.
[17] M. Desai and A. Shingala, Time series prediction of wheat crop based on FB Prophet forecast framework, *ITM Web of Conferences*, 53, 2023, 02014.
[18] N. K. ChikkaKrishna, P. Rachakonda, and T. Tallam, Short—term traffic prediction using Fb-PROPHET and neural-PROPHET, *2022 IEEE Delhi Section Conference (DELCON)*, 2022, 1-4.
[19] S. Ghimire, R. C. Deo, S. Ali Pourmousavi, D. Casillas-Pérez and S. Salcedo-Sanz, Point-based and probabilistic electricity demand prediction with a Neural Facebook Prophet and Kernel Density Estimation model, *Engineering Applications of Artificial Intelligence*, 135, 2024, 108702.
[20] H. Chérif, H. Snoun, G. Bellakhal and H. Kanfoudi, Forecasting of ozone concentrations using the Neural Prophet model: Application to the Tunisian case, *Euro-Mediterranean Journal for Environmental Integration*, 8(4), 2023, 987-998.
[21] R. Mushtaq, Augmented Dickey fuller test, *SSRN Electronic Journal*, 2011.
[22] O. Triebe, H. Hewamalage, P. Pilyugina, N. Laptev, C. Bergmeir and R. Rajagopal, Neuralprophet: Explainable forecasting at scale, *arXiv preprint*, arXiv, 2111.15397, 2021.
[23] W. Kim and B. M. Soon, Advancing agricultural predictions: A deep learning approach to estimating bulb weight using neural Prophet model, *Agronomy*, 13(5), 2023, 1362.
[24] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, *arXiv preprint*, arXiv, 1412.6980, 2014,
[25] S. J. Taylor and B. Letham, Forecasting at scale, *The American Statistician*, 72(1), 2018, 37-45.
[26] K. K. R. Samal, K. S. Babu, S. K. Das and A. Acharaya, Time series based air pollution forecasting using SARIMA and Prophet Model, *Proceedings of the 2019 International Conference on Information Technology and Computer Communications*, New York, USA, 2019, 80-85.
[27] Y. Ning, H. Kazemi and P. Tahmasebi, A comparative machine learning study for time series oil production forecasting: ARIMA, LSTM, and Prophet, *Computers & Geosciences*, 164, 2022, 105126.
[28] P. Newbold, ARIMA model building and the time series analysis approach to forecasting, *Journal of Forecasting*, 2(1), 1983, 23-35.
[29] G. Box, Box and Jenkins: Time series analysis, forecasting and control, In *A Very British Affair: Six Britons and the Development of Time Series Analysis during the 20th Century*, T. C. Mills, Ed., London: Palgrave Macmillan UK, 2013, 161-215.
[30] Y. Wang, Z. Yan, D. Wang, M. Yang, Z. Li, X. Gong, D. Wu, L. Zhai, W. Zhang and Y. Wang, Prediction and analysis of COVID-19 daily new cases and cumulative cases: Times series forecasting and machine learning models, *BMC Infectious Diseases*, 22(1), 2022, 495.
[31] I. Mirpulatov, M. Gasanov and S. Matveev, Soil dynamics and crop yield modeling using the Monica crop simulation model and time series forecasting methods, *Agronomy*, 13(8), 2023, 2185.
[32] T. R. Noviandy, A. Maulana, G. M. Idroes, R. Suhendra, M. Adam, A. Rusyana and H. Sofyan, Deep learning-based bitcoin price forecasting using neural Prophet, *Ekonomikalia Journal of Economics*, 1(1), 2023, 19-25.