Oil Palm Industry Economic Journal
DOI: https://doi.org/10.21894/opiej.2025.01

MALAYSIAN PALM OIL PRICE PREDICTION USING ARIMA – ARCH MODEL

FARAH HAYATI MUSTAPA[1*]

ABSTRACT

Palm oil is one of Malaysia's primary industries. In 2022, Malaysia was responsible for 25.14% of global palm oil production, totalling 18.55 million tonnes, and 31.65% of worldwide palm oil exports, amounting to 15.71 million tonnes (MPOB, 2024; Trend Economy, 2024). This article presents a study for Malaysian palm oil price forecasting using the traditional time series model, with three objectives, to identify the behaviour of palm oil prices in Malaysia, to determine the best model for predicting palm oil prices using the ARIMA model, and to perform a short-term forecast of the future palm oil price in Malaysia. The analysis was conducted using monthly Malaysian palm oil prices for a period of 10 years, 2013–2023. EViews software was used to run the data analysis. It is found that ARIMA (1,1,4) – ARCH (1) model is the most parsimonious model for forecasting the crude palm oil prices in Malaysia. Overall, predicting the monthly price of Malaysian palm oil provides valuable insights that support decision-making across the entire palm oil value chain, from producers and traders to policymakers and investors.

Keywords: ARIMA-ARCH model, forecasting, Malaysian palm oil, price volatility.

Received: 9 August 2024**; Accepted:** 17 December 2024**; Published online:** 4 March 2025.

INTRODUCTION

Commodity prices, such as those for oil, metal, agricultural products and energy, are critical to the global economy, impacting inflation, trade balances and overall economic growth. Within this broader context, the Malaysian palm oil industry is a cornerstone of the agricultural commodities market, playing a pivotal role in global economies, food security and environmental sustainability. In 2022, Malaysia produced 25.14% of world's palm oil production (18.55 million tonnes). It accounted for 31.65% of the world's palm oil exports (15.71 million tonnes), with major buyers including India, China, the European Union, Türkiye and Kenya (MPOB, 2024; Trend Economy, 2024). The agriculture sector, particularly oil palm, provides 10% of total employment in Malaysia in 2023, engaging over 3 million individuals, including nearly 500,000 smallholders (MPOC, 2024). However, the industry's market dynamics are affected by global demand, geopolitical events and environmental regulations, making accurate forecast palm oil prices essential for stakeholders – producers, traders, policymakers and consumers – who rely on such insights for informed decision-making regarding production levels, inventory management and resource allocation. Price volatility introduces financial risks, underscoring the importance of reliable forecasts for risk management and investment assessments, while also enabling policymakers to understand market trends and design supportive interventions to stabilise the sector.

Forecasting models, particularly the Autoregressive Integrated Moving Average (ARIMA) model, is a widely adopted methodology in time series forecasting, particularly suited for capturing the complex dynamics of financial and economic data. By analysing historical price trends, ARIMA models can identify underlying patterns,

[1] Department of Computer and Mathematical Sciences, Universiti Teknologi MARA Cawangan Pulau Pinang, Permatang Pauh Campus, 13500 Pulau Pinang, Malaysia.

* *Corresponding author e-mail: farah.hayati@uitm.edu.my*

trends, and seasonal variations that affect Malaysian palm oil prices. This approach not only facilitates short-term predictions but also supports strategic planning and risk management in the palm oil industry.

The Malaysian palm oil industry has experienced significant price volatility in recent years, driven by diverse economic, environmental and geopolitical factors. Despite the extensive research into price forecasting techniques, there remains a need to develop and validate robust ARIMA models specifically tailored to predict Malaysian palm oil prices accurately. This study seeks to address this gap by investigating the effectiveness of ARIMA models in forecasting palm oil prices, identifying key time series components (such as trends, seasonality and cyclical patterns), and assessing forecasting performance over different time horizons. The objective is to develop a more robust and optimal ARIMA-ARCH model for Malaysian palm oil prices and conduct a short-term forecast to compare observed and fitted values.

Research on palm oil price forecasting has utilised a range of methodologies. In time series modelling, researchers typically use historical price data to identify trends, seasonality and cyclical patterns. A comparative study by Khin *et al.* (2013) found that the Multivariate ARMA model is the best predictive model than ARIMA and Vector Error Correction model (VECM) for forecasting the spot palm oil price. Ahmad *et al*. (2014) combined ARIMA with GARCH to model price volatility, finding a hybrid ARIMA-GARCH model most effective for the Malaysian CPO price from January 1999 to May 2014. Hamid and Shabri (2017) preferred the Autoregressive Distributed Lag (ARDL) model over ARIMA, while Khalid *et al*. (2018) identified ARIMA with exogenous variables (ARIMAX) as superior to ARDL and ARIMA. Zaidi *et al*. (2021) employed Support Vector Autoregression (SVAR) technique to assess the impact of substitute goods on palm oil prices.

Econometric models incorporate economic variables such as crude oil prices, exchange rates, and global demand into the forecasting process. These models often use regression techniques to establish relationships between palm oil prices and relevant economic indicators (Ahmad *et al*., 2014). Vector Autoregression (VAR) models are also employed to capture the dynamic interactions among multiple time series variables affecting palm oil prices.

Machine learning techniques have recently gained popularity for modelling and forecasting palm oil prices (Kanchymalay *et al.*, 2017). Algorithms such as Support Vector Machines (SVM), Artificial Neural Networks (ANN) (Karia *et al.*, 2013), and Random Forests (Li *et al.*, 2016) were applied due to their ability to handle complex non-linear relationships and large datasets. These models often outperform traditional statistical approaches when sufficiently large data is available for training and validation. Rahim *et al*. (2018) found that the Fuzzy Time Series-based forecasting model was a simpler method to forecast CPO price than the ARIMA-ANN method. Moreover, Yee and Samsudin (2021) found that ANN outperformed ARIMA in forecasting palm oil price movement between January 2008 and December 2018.

In conclusion, modelling and forecasting Malaysian palm oil prices requires a multidisciplinary approach that integrates economic theory, statistical methods and advanced computational techniques. While machine learning algorithms is useful for modelling a complex, and non-linear patterns, the ARIMA model remains to be a competitive, robust, and practical tool, particularly in scenarios where the underlying data has a more linear or well-known patterns. To advance research in this field, it is essential to address data challenges and sustainability considerations.

MATERIAL AND METHOD

The Box-Jenkins model (1976) was named after the statisticians who developed systematic procedures for identifying and estimating ARIMA models. This model is favoured for its ability to capture complex patterns, trends, seasonality and random noises in time series data. It includes both the autoregressive (AR) component, which models the relationship between an observation and several of its lagged values (previous observation), and the moving average (MA) component, which represents the relationship between an observation and the residual errors from a moving average process applied to the lagged observations. The methodology encompasses statistical tests for residual diagnostic and short-term forecasting.

The stationarity of data series is a fundamental requirement for applying the ARMA (p, q) model. A stationary time series does not exhibit trends, seasonal effects, or any structure that evolves over time. However, most economic variables are non-stationary and have trends. If the variable is non-stationary, it needs to be differenced to remove the trend. ARIMA (p, d, q) model, in contrast to the ARMA (p, q) model, includes the 'I' for integration, where d is the number of differencing operations required to achieve stationarity. The equation of ARIMA (1,1,1) is given by (Lazim, 2005):

$$w_t = \mu + \phi_1 w_{t-1} + \theta_1 \epsilon_{t-1} + \epsilon_t \tag{1}$$

where, $w_t = y_t - y_{t-1}$ represents the first difference of the series.

Traditional econometrics models assume that residuals (errors) are homoscedastic, meaning they have constant variance over time. However, many economic time series face periods of volatility higher than usual, which violates the homoskedasticity assumption. This can result in biased parameter estimates and standard errors, impacting the model's reliability and reducing forecast accuracy. To address heteroscedasticity, models such as, Autoregressive Conditional Heteroscedasticity (ARCH) are used. The ARCH model, introduced by Engle (1982) in his article "Autoregressive Conditional Heteroscedasticity with estimates of the variance of United Kingdom Inflation" earned him the 2003 Nobel Memorial Prize in Economic Sciences. The model was later extended in 1986 by Bollerslev with the introduction of Generalised Autoregressive Conditional Heteroscedasticity (GARCH) models. Thus, ARCH/GARCH models capture conditional heteroscedasticity, allowing for the modelling of financial time series with time-varying volatility and volatility clustering. While GARCH is generally preferred for its flexibility, efficiency and forecasting accuracy, ARCH models can be suitable for simpler problems or datasets with limited volatility structure, where the added complexity of GARCH might not provide substantial benefits; therefore, for short time series or limited data, ARCH may still be a reasonable choice.

The error terms in an ARCH (1) model are normally distributed with a mean of 0 and the conditional variance of the error terms is represented by Equation (2):

$$Var(\epsilon_t) = \sigma_t^2 = \alpha_0 + \alpha_1\, \sigma_{t-1}^2 \qquad (2)$$

The above model is known as ARCH (1), since the conditional variance depends on only one lagged squared error.

Data Description

The historical monthly price data of Malaysian palm oil was retrieved from MPOB (2024), for a period of 10 years, ranging from February 2013 to February 2023. There is a total of 121 observations. All the data is used for model fitting. To forecast Malaysian palm oil prices using an ARIMA model, there are four major stages, namely model identification, parameter estimation, diagnostic checking as well as forecasting. EViews software, 10th edition was used to run data analysis.

Initial Data Investigation and Model Identification

In the very first step, the data will be examined for the presence of any trend, seasonality or non-stationarity. This is done by plotting the raw data using a simple line chart, analysing the patterns of Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots and conducting a unit root test. If the data is stationary at the level, an ARMA model will be estimated. If the data is non-stationary, differencing will be applied to remove the trend. Usually, a first difference (d =1) is sufficient, but if the data remains non-stationary, a second differencing may be required. Once the data becomes stationary, an ARIMA (p, d, q) model will be used. The next step involves examining the correlogram to identify the appropriate values for p and q in the ARIMA model. The number of significant spikes of the ACF and PACF plots will guide the selection of potential models. The PACF will be used to determine the AR component while the ACF will assist in identifying the MA component.

Model Estimation

In this step, the parameters of the ARIMA model will be estimated based on the identified characteristics of the data by using the least squares estimator. Then the significance of the coefficients will be evaluated and the Akaike-Information-Criterion (AIC) and Bayesian or Schwartz-Criterion (BIC) scores obtained will be compared. The optimal model will be chosen based on the significance of coefficients, lowest volatility, lowest AIC and BIC values, and highest adjusted R-squared. The model with the least volatility will increase the power of prediction.

Model Diagnostic

Once a potential model is identified, diagnostic checks will be performed to ensure that it meets the requirement for a stable univariate process. It is important to verify that the model's residuals are white noise, meaning they have independent observations with constant mean and variance and exhibit no autocorrelation between consecutive residuals in the data series. This is assessed using the Ljung-Box Q-statistic test for autocorrelation in ARIMA models. A test for the ARCH effect will also be conducted. If the ARCH effect is detected, ARCH/GARCH models should be used to model the conditional variance explicitly. If all conditions are met, we can proceed with forecasting using the estimated model. Otherwise, we need to re-estimate and explore another potential model.

Forecasting

Once the most parsimonious model is identified, it will be used for generating future forecasts. In this context, the dynamic forecast option in EViews will be employed. Dynamic forecasting assumes that the forecasts will be made without using any data from beyond the sample period (out-sample data) even if relevant data is available.

Evaluation metrics such as Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE) are commonly used to assess the accuracy of ARIMA forecasts (Kanchymalay *et al.*, 2017; Khalid *et al.*, 2018; Yee & Samsudin, 2021; Zhou *et al.*, 2018). RMSE measures the average magnitude of forecast errors, with larger errors receiving more weight due to the squaring of differences. A lower RMSE indicates better predictive accuracy. RMSE for one-step-ahead forecast is written as:

$$\text{RMSE} = \sqrt{\frac{\sum_{j=1}^{n} (y_j - \hat{y}_j)^2}{n}} \tag{3}$$

Unlike RMSE, MAPE focuses on the relative size of errors as a percentage of the actual values, with lower percentages indicating closer predictions to actual values. A MAPE value below 10% is considered as good. MAPE is written as:

$$\text{MAPE} = \frac{1}{n} \sum_{j=1}^{n} \left| \frac{y_j - \hat{y}_j}{y_j} \times 100 \right| \tag{4}$$

Cross-validation techniques may be employed to validate the robustness of the ARIMA models and identify potential overfitting issues.

RESULTS AND DISCUSSION

At the initial stage, a preliminary data analysis was performed to identify the basic pattern of the series and to detect any unusual observations or characteristics. *Figure 1* shows Malaysian palm oil prices fluctuations over the 10 years. A cursory observation of *Figure 1* illustrates that the series is not stationary. Although the series does not show any seasonal effects, there are notably large values in mid-2022, which could be interpreted as irregularities. The palm oil price surge in mid-2022 was caused by a combination of supply disruptions caused by the Russia–Ukraine war which started in February 2022, Indonesia's temporary export ban from April to May 2022, labour shortages and production constraints in Malaysia and strong global demand for palm oil as a substitute for other edible oils (Voora *et al.*, 2023).

Stationarity Test

The ACF and PACF were plotted to determine the stationarity of the series. Based on *Figure 2*, the ACF plot tailing off extremely slowly in a linear fashion. The autocorrelations remain high over many lags and decline at a steady rate, suggesting that the series exhibits a trend and is non-stationary.

The Augmented Dickey-Fuller (ADF) test is also employed to assess whether the series is non-stationary. According to the results presented in *Table 1*, the *p*-value = 0.262 exceeds the 5% significance level, leading to a failure to reject the null hypothesis. This indicates that the series unsurprisingly possesses a unit root, meaning the data is non-stationary at level.

To address the high autocorrelation and achieve stationary, we applied the first difference to the data. Upon reapplying the ADF test to the differenced series, we obtained a *p*-value of 0.000, which is below the 5% significance level. Therefore, we conclude that the data now is stationary in the mean, allowing us to proceed with estimating the model.

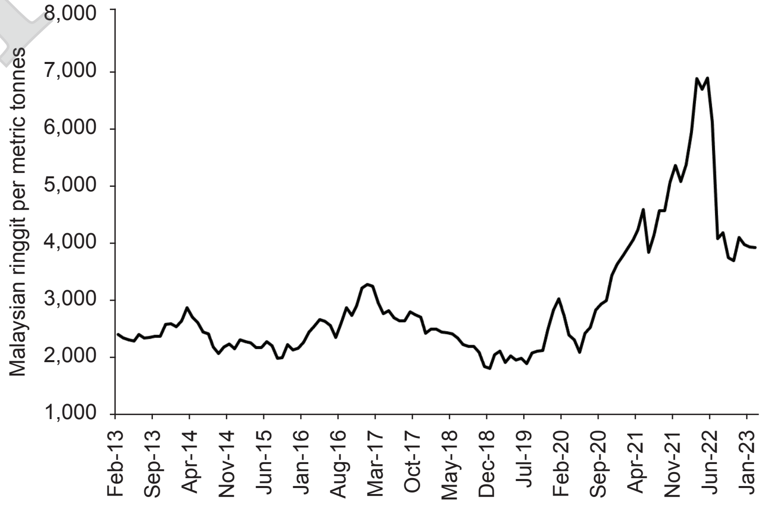
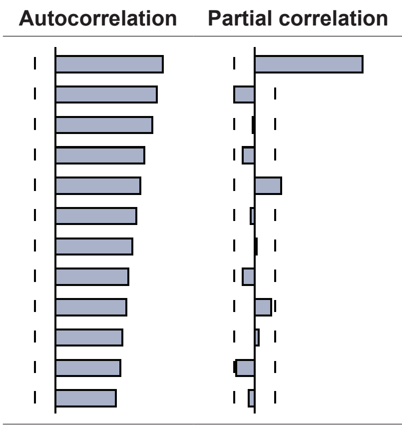


Figure 1. Malaysian palm oil price.



	AC	PAC	Q-stat	Prob
1	0.960	0.960	125.42	0.000
2	0.909	-0.173	238.55	0.000
3	0.856	-0.010	339.85	0.000
4	0.799	-0.102	428.69	0.000
5	0.759	0.223	509.47	0.000
6	0.724	-0.040	583.50	0.000
7	0.690	0.009	651.35	0.000
8	0.654	-0.108	712.80	0.000
9	0.624	0.138	769.17	0.000
10	0.601	0.040	821.90	0.000
11	0.568	-0.159	869.37	0.000
12	0.533	-0.052	911.52	0.000

Figure 2. The correlogram of Malaysian palm oil price at level.

TABLE 1. THE RESULTS OF AUGMENTED DICKEY-FULLER TEST

Model	Test for unit root at	Test statistic	*p*-value
Trend and intercept	Level	-2.644	0.262
	First difference	-8.543	0.000

Figure 3 graphically shows the fluctuations of the series around the zero line, indicating that stationarity in mean has been achieved after the first differencing. However, it is evident that there remains non-stationarity in variance of the series.

Model Identification and Estimation

The process of identifying the potential models for the data series involved analysing the ACF and PACF of the stationary series. As shown in *Figure 4*, both the PACF and ACF plots cut off at lag 1 and 4 respectively. This suggests that potential models to test could include ARIMA (1,1,1), ARIMA (1,1,4), ARIMA (4,1,1) and ARIMA (4,1,4).

As shown in *Table 2*, the ARIMA (1,1,4) was chosen as the preferred model because it fulfilled all the selection criteria.

Model Validation and Diagnostic Checking

In a well-fitted model, the residuals should ideally exhibit white noise characteristics. To verify this, diagnostic checks were performed by plotting the correlogram of the residuals. *Figure 5* demonstrates that all values in the ACF and PACF plots lie within the 95% confidence interval, indicating that the absence of significant autocorrelation in the residuals of the selected models. The null hypothesis that the residuals are white noise could not be rejected, suggesting that the ARIMA (1,1,4) model is adequate.

The ARCH-LM test is performed to determine the presence of heteroscedasticity in variance of the residuals. The null hypothesis states that there are no ARCH effects up to the specified lag order. If *p*-values are less than the significance level,

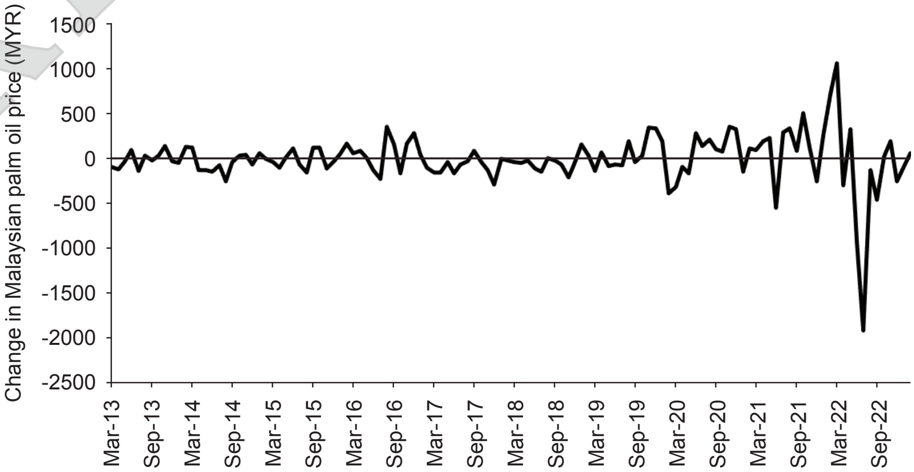
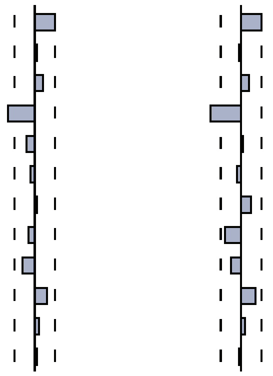


Figure 3. Malaysian palm oil price after first differencing.

the null hypothesis will be rejected. As shown in *Table 3*, before fitting ARCH (1) model into the ARIMA (1,1,4) model, the p-values are 0.0142, which is less than alpha = 0.05. Therefore, the null hypothesis is rejected, indicating the presence of heteroscedasticity in the residuals. Although the inclusion of additional lags was considered, the ARCH (1) model was found to adequately address the heteroscedasticity in the model. The ARCH test was re-conducted to verify the absence of ARCH effects after fitting the ARCH (1) model. As shown in *Table 3*, the *p*-values are greater than 0.05, confirming that heteroscedasticity is no longer present in the model.

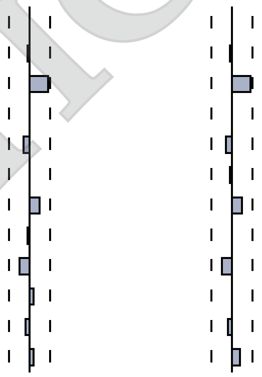


Autocorrelation	Partial correlation		AC	PAC	Q-stat	Prob
		1	0.177	0.177	4.2293	0.040
		2	0.013	-0.019	4.2524	0.119
		3	0.070	0.073	4.9186	0.178
		4	-0.236	-0.271	12.600	0.013
		5	-0.080	0.019	13.482	0.019
		6	-0.031	-0.038	13.620	0.034
		7	0.025	0.093	13.706	0.057
		8	-0.051	-0.149	14.080	0.080
		9	-0.113	-0.091	15.914	0.069
		10	0.103	0.128	17.454	0.065
		11	0.027	0.031	17.565	0.092
		12	0.012	-0.014	17.586	0.129

Figure 4. The collelogram of Malaysia palm oil price at first difference.

TABLE 2. TEST RESULTS OF ARIMA (*p*,1,*q*)

Model	Coefficient	Sigma2 (volatility)	Adj. R^2	AIC	BIC
ARIMA (1,1,1)	Significant	77617.37	0.0720	14.17352	14.26644
ARIMA (1,1,4)	Significant	74554.68	0.1086	14.12804	14.22096
ARIMA (4,1,1)	Significant	75503.02	0.0973	14.13988	14.23279
ARIMA (4,1,4)	Not significant	78300.60	0.0638	14.17625	14.26917



Autocorrelation	Partial correlation		AC	PAC	Q-stat	Prob
		1	-0.002	-0.002	0.0004	
		2	-0.019	-0.019	0.0465	
		3	0.161	0.161	3.2828	0.070
		4	-0.006	-0.006	3.2870	0.193
		5	-0.057	-0.053	3.7072	0.295
		6	0.006	-0.020	3.7118	0.446
		7	0.084	0.087	4.6204	0.464
		8	-0.024	-0.008	4.6987	0.583
		9	-0.093	-0.093	5.8328	0.559
		10	0.029	-0.001	5.9435	0.654
		11	-0.045	-0.041	6.2208	0.718
		12	0.033	0.075	6.3650	0.784

Figure 5. The Ljung-Box Test for ARIMA (1, 1, 4).

TABLE 3. HETEROSCEDASTICITY TEST

Heteroscedasticity test: ARCH	Before fitting ARCH (1)		After fitting ARCH (1)	
ARIMA (1,1,4)	Prob. F (1,117)	0.0142	Prob. F(1,116)	0.9487
	Prob. Chi-Square(1)	0.0144	Prob. Chi-Square(1)	0.9482

In general, the equation for ARIMA (1,1,4) model can be written as:

$$w_t = \mu + \phi_1 w_{t-1} + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \theta_3 \epsilon_{t-3} + \theta_4 \epsilon_{t-4} + \epsilon_t \quad (5)$$

where, $w_t = y_t - y_{t-1}$ is the first difference of the series, μ is the constant term, ϕ_1 is the coefficient for the AR (1) term, ϵ_t is the error term at time t, and $\theta_1, \theta_2, \theta_3, \theta_4$ are the coefficients for the MA (4) term,

However, in this study, for the MA component, only one composite moving average term, θ is estimated, and this coefficient will affect all the lags of the error term simultaneously, not with separate coefficients for each lag.

$$w_t = \mu + \phi_1 w_{t-1} + \theta (\epsilon_{t-4} + \epsilon_{t-3} + \epsilon_{t-2} + \epsilon_{t-1} + \epsilon_t) \quad (6)$$

where, θ is a single parameter that reflects the cumulative effect of all four lags of the error term. This simplified model may be more stable and parsimonious, leading to more robust parameter estimates and significant coefficients.

The following is the estimation results of ARIMA (1, 1, 4) – ARCH (1) model.

This estimation result corresponds to the following specification,

$$w_t = 0.0553 + 0.2682 w_{t-1} - 0.1971 (\epsilon_{t-4} + \epsilon_{t-3} + \epsilon_{t-2} + \epsilon_{t-1} + \epsilon_t) \quad (7)$$

where the variance equation is,

$$\sigma_t^2 = 21242.68 + 0.7313 \sigma_{t-1}^2 \quad (8)$$

As we can see the α_1 value = 0.7313 which is close to 1 indicates that the persistence of the volatility is quite high. *Figure 8* demonstrates how the variance has increased significantly in the period of the first quarter of 2022.

Forecasting

The ARIMA (1, 1, 4) – ARCH (1) model for Malaysia palm oil price was estimated using the whole 121 observations (or monthly observations from February 2013 until February 2023). Since the model fulfilled all the conditions, the next step is to generate forecast and assess its accuracy. A forecast for the palm oil price will be made for the next 12 observations, extending through 2024. This process begins by adding 12 out-of-sample observations to the dataset.

The 95% confidence interval for the mean is a range of values that we are 95% confident contains the population mean. It is typically calculated as the sample mean plus or minus the margin of error. In this case, the forecasted values are computed by adding and subtracting two times the standard error from the sample mean. This accounts for approximately 95% of the area under the normal distribution curve. *Table 4* presents the forecast evaluations for the ARIMA (1,1,4) and ARIMA (1,1,4)–ARCH (1) models. The model with the lower RMSE and MAPE are generally preferred, as it indicates better predictive performance on the test data.

For the hybrid ARIMA (1,1,4) – ARCH (1) model, the RMSE of 237.64 is lower than that of the ARIMA (1,1,4) model, indicating superior forecast accuracy. The MAPE of 5.53% means the average difference between the forecasted and actual values is 5.53%, which is also an improvement over the ARIMA (1,1,4) model of 10.55%.

Figure 9 visually displays the actual and forecasted prices, along with the two confidence bands.

Variable	Coefficient	Std. error	z-statistic	Prob.
C	0.055319	19.52330	0.002834	0.9977
AR(1)	0.268136	0.088874	3.017016	0.0026
MA(4)	-0.197132	0.061340	-3.213772	0.0013
		Variance equation		
C	21242.68	4472.070	4.750078	0.0000
RESID(-1)^2	0.731298	0.173463	4.215873	0.0000
R-squared	0.108504	Mean dependent var		13.24370
Adjusted R-squared	0.093134	S.D. dependent var		295.31340
S.E. of regression	281.2256	Akaike info criterion		13.59191
Sum squared resid	9174189.	Schwarz criterion		13.70868
Log likelihood	-803.7189	Hannan-Quinn criter.		13.63933
Durbin-Watson stat	2.123342			
Inverted AR Roots	.27			
Inverted MA Roots	.67	-.00+.67i	-.00-.67i	-.67

Figure 7. The ARIMA (1, 1, 4) – ARCH (1) Model for Malaysian palm oil price.

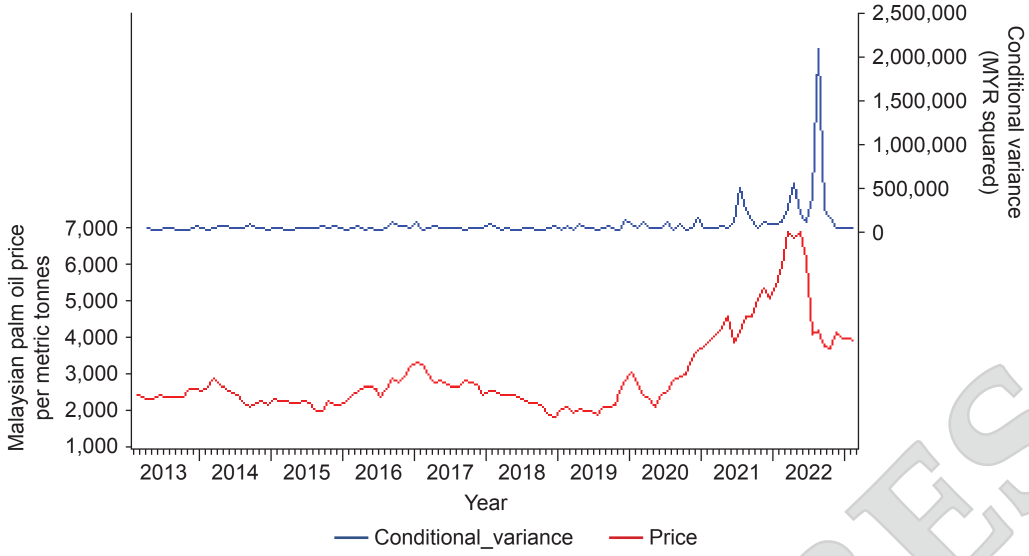


Figure 8. Conditional variance and Malaysian palm oil price.

TABLE 4. FORECAST EVALUATION FOR ARIMA (1,1,4) AND ARIMA (1,1,4)-ARCH (1) MODEL

Forecast: PRICEF ARIMA(1,1,4)		Forecast: PRICEF ARIMA(1,1,4)-ARCH(1)	
Actual: PRICE		Actual: PRICE	
Forecast sample: 2023M03 2024M02		Forecast sample: 2023M03 2024M02	
Included observations: 12		Included observations: 12	
Root mean squared error	419.3994	Root mean squared error	237.6398
Mean absolute error	395.7992	Mean absolute error	209.1486
Mean abs. percent error	10.55121	Mean abs. percent error	5.530055
Theil inequality coefficient	0.052587	Theil inequality coefficient	0.030631
Bias proportion	0.606486	Bias proportion	0.215054
Variance proportion	0.067194	Variance proportion	0.577258
Covariance proportion	0.326320	Covariance proportion	0.207688
Theil U2 coefficient	2.159479	Theil U2 coefficient	1.147275
Symmetric MAPE	9.991144	Symmetric MAPE	5.419211

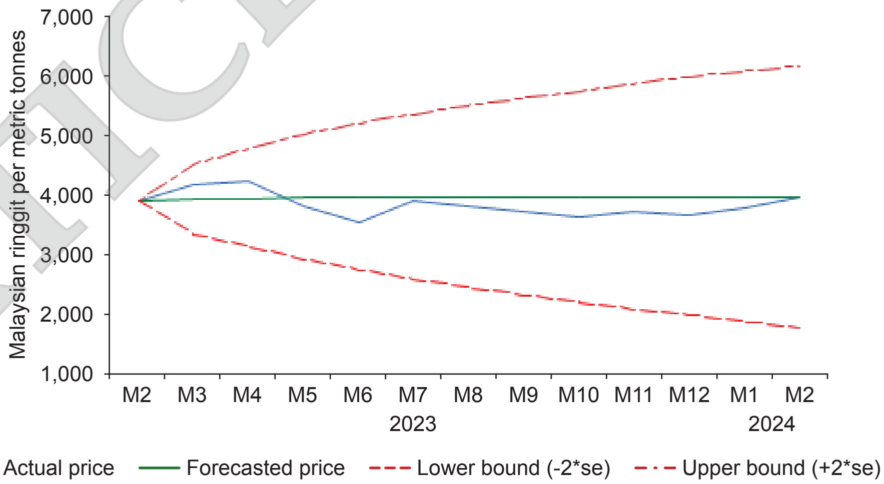


Figure 9. Malaysian palm oil price (Actual - Forecast).

CONCLUSION

This study elaborates the development of a statistical model for forecasting Malaysian palm oil prices using the Box-Jenkins methodology, specifically the ARIMA model. The best-fitting model identified is ARIMA (1,1,4)– ARCH (1). Utilising this model, forecasts for palm oil price over the next 12 months were generated. Given the volatility in global commodity markets and the recent fluctuations in Malaysian palm oil prices due to factors such as changing demand, geopolitical tensions and environmental policies, the ARIMA models have demonstrated their effectiveness in predicting Malaysian palm oil prices. These forecasts provide valuable insights that can help industry stakeholders, such as producers, exporters and policymakers, navigate the uncertainty in pricing trends and better align their strategies with market conditions. As the palm oil market faces continued challenges - such as price instability, fluctuating global demand and regulatory changes - further research and innovation in time series forecasting methodologies will be critical. Advances in these methods will enhance the accuracy and reliability of predictions, thereby supporting informed decision-making, effective risk management and the development of sustainable strategies for the Malaysian palm oil industry.

ACKNOWLEDGEMENT

The author wishes to express a special thanks to MPOB for providing data and statistics for this study. Also, an appreciation goes to Universiti Teknologi MARA Cawangan Pulau Pinang for providing the financial support necessary to conduct this study. The suggestions of the reviewers for improvement to the earlier version of this article are gratefully acknowledged.

REFERENCES

Ahmad, M. H., Ping, P. Y., & Mahamed, N. (2014). Volatility modelling and forecasting of Malaysian crude palm oil prices. *Applied Mathematical Sciences, 8*(124), 6159–6169.

Box, G., & Jenkins, G. (1976). *Time series analysis: Forecasting and control* (2nd ed.). Holden-Day.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica, 50*(4), 987–1007.

Hamid, M. F. A., & Shabri, A. (2017, May). Palm oil price forecasting model: An autoregressive distributed lag (ARDL) approach. In *AIP Conference Proceedings* (Vol. 1842, No. 1). AIP Publishing.

Kanchymalay, K., Salim, N., Sukprasert, A., Krishnan, R., & Hashim, U. R. A. (2017, August). Multivariate time series forecasting of crude palm oil price using machine learning techniques. *IOP Conference Series: Materials Science and Engineering, 226*(1), 012117. IOP Publishing.

Karia, A. A., Bujang, I., & Ahmad, I. (2013). Forecasting on crude palm oil prices using artificial intelligence approaches. *American Journal of Operations Research, 3*(2), 259–267.

Khalid, N., Hamidi, H. N. A., Thinagar, S., & Marwan, N. F. (2018). Crude palm oil price forecasting in Malaysia: An econometric approach. *Jurnal Ekonomi Malaysia, 52*(3), 263–278.

Khin, A. A., Mohamed, Z., Malarvizhi, C., & Thambiah, S. (2013). Price forecasting methodology of the Malaysian palm oil market. *The International Journal of Applied Economics and Finance, 7*(1), 23–36.

Lazim, M. A. (2005). *Introductory business forecasting – A practical approach*. Pusat Penerbitan Universiti (UPENA).

Li, B., Yang, G., Wan, R., Dai, X., & Zhang, Y. (2016). Comparison of random forests and other statistical methods for the prediction of lake water level: A case study of the Poyang Lake in China. *Hydrology Research, 47*(S1), 69–83. https://doi.org/10.2166/nh.2016.264

Malaysian Palm Oil Council. (2024). *Malaysian palm oil industry*. http://archive.mpoc.org.my/malaysian-palm-oil-industry

MPOB. (2024). *Malaysia prices of crude palm oil*. https://bepi.mpob.gov.my/admin2/price_local_daily_view_cpo_msia.php?more=Y&jenis=3M

Rahim, N. F., Othman, M., & Sokkalingam, R. (2018). A comparative review on various methods of forecasting crude palm oil prices. *Journal of Physics: Conference Series, 1123*(1), 012043. IOP Publishing.

Trend Economy. (2024). *Palm oil import & export 2022*. https://trendeconomy.com/data/commodity_h2/1511

Voora, V., Bermúdez, J., Farrell, J., Larrea, C., & Luna, E. (2023). *Palm oil prices and sustainability*. Global Market Report, 2023-06.

Yee, K. W., & Samsudin, H. B. (2021). Comparison between artificial neural network and ARIMA model in forecasting palm oil price in Malaysia. *International Journal of Scientific Engineering and Science, 5*(11), 12–15.

Zaidi, M. A. S., Abdul Karim, Z., & Zaidon, N. A. (2021). External and internal shocks and the movement of palm oil price: SVAR evidence from Malaysia. *Economies, 10*(1), 7.

Zhou, L., Zhao, P., Wu, D., Cheng, C., & Huang, H. (2018). Time series model for forecasting the number of new admission inpatients. *BMC Medical Informatics and Decision Making, 18*, 39.