Volume 5, Issue 11, pp. 12-15, 2021.
International Journal of Scientific Engineering and Science
ISSN (Online): 2456-7361

Comparison between Artificial Neural Network and ARIMA Model in Forecasting Palm Oil Price in Malaysia

Kee Wei Yee, Humaida Banu Samsudin
Actuarial Science Programme, Department of Mathematical Sciences, Faculty of Science & Technology, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia
Email address: humaida(at)ukm.edu.my

Abstract— *Malaysia is one of the largest producers of palm oil in the world. Malaysia exported palm oil to almost 160 countries and contributed about RM50 billion or 3.5 per cent to Gross Domestic Product (GDP) in 2019. The palm oil industry is the fourth largest contributor to Malaysia's export earnings. However, palm oil price keeps fluctuating over time. Therefore, accurate prediction of palm oil price is important as investors deal with risks and uncertainties in the future. This study forecasts palm oil price in Malaysia using artificial neural network (ANN) and autoregressive integrated moving average (ARIMA) model. Comparisons between the two models are made and the most accurate model is selected. Monthly palm oil price data from January 2008 to December 2018 are used to build the forecasting models. The models are used to forecast the price of palm oil in Malaysia for year 2019. The predicted values are used to compare with the actual values. The main result reveals that artificial neural network model is more accurate compared to ARIMA model in forecasting the palm oil price although both models give good forecasting performance.*

Keywords— *Prediction; forecast; palm oil price.*

I. INTRODUCTION

Malaysia is one of the largest producers of palm oil in the world. Malaysia has exported palm oil to almost 160 countries with China, India and European Union as its largest market. From year 2017 to year 2019, India is the country that imported the most palm oil from Malaysia, followed by China, European Union and Pakistan. Southeast Asian countries that are near Malaysia such as Philippines, Vietnam and Singapore also imports palm oil from Malaysia.

Director General of the Malaysian Palm Oil Board (MPOB), Dr Ahmad Parveez Ghulam Kadir stated that the demand for Malaysian palm oil is expected to continue increasing in line with the world population which will increase to nine billion people by 2043. The Malaysian palm oil industry has produced almost 20 million tonnes of crude palm oil a year with an export value reaching RM65 billion in 2019. It contributes about RM50 billion or 3.5 per cent to Gross Domestic Product (GDP). In Malaysia, the palm oil industry is the fourth largest contributor to the country's export earnings after industry of electrical and electronics, petroleum products and chemical products (Zuki 2020). Therefore, palm oil industry is important for Malaysia’s economy as it plays an important role in the country’s export activities. The export income will directly contribute to the country’s economic growth.

However, palm oil prices have fluctuated without showing any clear trends and cyclical patterns over the past three decades. Food commodity prices including palm oil prices change rapidly over time due to price volatility for food commodities (Arshad & Hameed 2013). Ali (2019) stated that the price trends of palm oil are influenced by demand and supply, and other technical factors. Factors that influence the palm oil prices should be identified to avoid losses. Nevertheless, there is a need to address the low income earned by the 650,000 smallholders whose livelihood depends solely on oil palm output. The household income of these smallholders shrank significantly when the price of palm oil fell.

Hence, the accurate forecasting for palm oil prices is significant so that all relevant parties are able to make right decisions to avoid losses. There are several studies have been conducted on palm oil price forecasting. Most of the forecasting techniques used are conventional forecasting techniques such as exponential smoothing and autoregressive integrated moving average (ARIMA) model (Abdul Hamid & Shabri 2017; Khalid et al. 2018). Besides that, there are also other forecasting techniques used such as artificial neural network (ANN) (Karia, Bujang & Ahmad 2013; Md Nor, Sarmidi & Hosseinidoust 2014; Rahim, Othman & Sokkalingam 2018).

This study has two main objectives, firstly, to develop palm oil price forecasting model using artificial neural network and ARIMA model. Secondly is to identify the best model in palm oil price forecasting by comparing the forecasting performance between two models.

II. MATERIALS AND METHODS

A. Data

The monthly palm oil price data for Malaysia from year 2008 to year 2019 from Malaysian Palm Oil Council (MPOC) website are used for this study. 132 palm oil prices in Malaysia are used as sample data in estimating model parameters as well as model selection. The remaining 12 palm oil prices in Malaysia are used as out sample data and will be used to evaluate the model’s forecasting performance.

International Journal of Scientific Engineering and Science
Volume 5, Issue 11, pp. 12-15, 2021. ISSN (Online): 2456-7361

B. ARIMA model

ARIMA model is a widely used statistical method for time series forecasting. In ARIMA (*p*,*d*,*q*) model, parameter *p* represents autoregressive (AR), parameter *d* represents the number of differencing and parameter *q* represents moving average (MA). Based on Khalid et al. (2018), model ARIMA (*p*,*q*) can be written as:

$$Y_t = \delta + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \cdots + \alpha_p Y_{t-p} + \varepsilon_t - \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \cdots + \beta_q \varepsilon_{t-q}$$

where
Y_t = dependent variable at time t
δ = constant
$\alpha_1, \ldots, \alpha_p$ = order of autoregressive (AR)
$\beta_1, \ldots, \beta_q$ = order of moving average (MA)
ε_t = error term at time t

C. Artificial neural network

Artificial neural networks are forecasting methods that are based on simple mathematical models of the brain. Artificial neural networks consist of interconnected neurons. Each neuron or nodes are interconnected independently. Based on Karia and Bujang (2011), it can be shown by equation:

$$Y = f[\sum(x_1 w_1 + x_2 w_2 + \cdots + x_t w_t) + \beta]$$

where
Y = output from the neuron
x_t = input values
w_t = connection weights
β = bias value (threshold)
f = transfer function, typically known as sigmoidal function, $f(x) = \frac{1}{1+e^{-(x)}}$

There are many applications for time series prediction in neural network. In this study, the type of artificial neural network used is nonlinear autoregressive (NAR) model. This form of prediction is only involved one series. Therefore, the future values are predicted from the past value of its own. Based on Ruiz et al. (2016), the equation for NAR model for time series forecasting can be written as:

$$y(t) = h[y(t-1), \ldots, y(t-p)] + \epsilon_t$$

where
$y(t)$ = prediction value at time t
p = past values of data series
$y(t-1), \ldots, y(t-p)$ = feedback delays
ϵ_t = error of the approximation at time t

D. Accuracy Test

Model accuracy test include mean root mean square error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE).

III. RESULTS AND DISCUSSION

A. Descriptive Analysis

Figure 1 shows the graph of monthly palm oil price in Malaysia from year 2008 to year 2018. Based on Figure 1, the graph shows that the data is not stationary. In 2008, the price dropped significantly due to global financial crisis. The highest price was recorded in year 2011 as it was supported by positive sentiments related to world supply tightness of vegetable oils and low domestic palm oil stocks level. Palm oil price increased in year 2010, 2015 and 2016 affected by the El Nino and La Nina climate pattern. The production of palm oil declined due to the extreme weather. This reduced the supply of palm oil to the market which causes the price to increase.

During the first and second quarter of year 2012, price rose due to world supply tightness of vegetable oils, especially soyabean oil. Price dropped during third quarter of year 2012 resulting from unresolved Euro-zone financial crisis that leads to poor demand for oils and fats. In year 2013, the price trend showed an improved upward price performance as a result of the Government's measure to restructure the export duty. Price was traded higher during the first half of year 2014 due to tight domestic palm oil supply. During second half of year 2014, the lower prices was in line with the weaker soyabean oil prices. In 2017, price increased due to the increase in production of palm oil. Price continued to drop for year 2018 because of lower palm oil export demand coupled with weaker prices of other vegetable oil in the world market.

Palm oil price also shows seasonal pattern because the production of palm oil has a seasonal pattern every year. Production of palm oil will start to decline in November until February every year due to monsoon season. Production of palm oil will affect the supply of palm oil to the market and then affect the price.

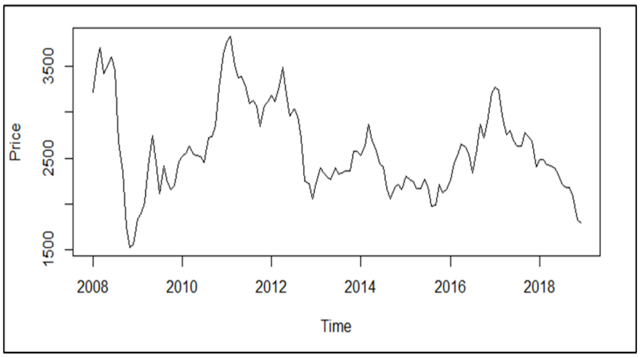

Fig. 1. Monthly palm oil price in Malaysia from year 2008 to year 2018.

B. Palm Oil Price Forecasting With ARIMA model

Augmented Dickey Fuller (ADF) test is conducted to test for stationary data. Table I shows the result for ADF test. The original data is non-stationary. Therefore, first differencing is carried out to obtain stationary data.

TABLE I. Result for ADF test.

Data	Statistic-*t*	Critical Value (1%)	Conclusion
Original	-3.2906	-3.46	Do not reject hypothesis null. Data is not stationary.
First Differencing	-7.4198	-3.46	Reject hypothesis null. Data is stationary.

There are three parameters required by an ARIMA (*p*,*d*,*q*) model. Parameter *d* represents the number of differencing, thus *d*=1. To determine parameters *p* and *q*, ACF and PACF plots are conducted. ACF plot gives estimation for parameter *q* which is the moving average (MA), while PACF plot gives estimation for parameter *p* which is the autoregressive (AR). The corresponding values for parameter *q* are 0,1,5 and 6 while the corresponding values for parameter *p* is 0,1 and 5.

International Journal of Scientific Engineering and Science
Volume 5, Issue 11, pp. 12-15, 2021. ISSN (Online): 2456-7361

The model with the lowest AIC value is the best model as shown in Table II.

TABLE II. AIC values.

ARIMA model	AIC value
(0,1,0)	1744.853
(0,1,1)	1734.027
(0,1,5)	1736.046
(0,1,6)	1731.645
(1,1,0)	1735.316
(1,1,1)	1736.119
(1,1,5)	1728.119
(1,1,6)	1731.645
(5,1,0)	1730.174
(5,1,1)	1728.910
(5,1,5)	1731.645
(5,1,6)	1731.645

Based on Table II, ARIMA (1,1,5) has the lowest AIC value. Residual graph, residual histogram, ACF and PACF plots for residual and Ljung-Box test are carried out to test the adequacy of the model. Based on the tests, ARIMA (1,1,5) has no autocorrelation. Thus, ARIMA (1,1,5) is selected to be the model for forecasting.

C. Palm Oil Price Forecasting With Artificial Neural Network

Data is divided into 3 parts: 70% for training, 15% for validation and 15% for testing. The trial and error method is implemented to achieve the optimal network structure. Figure 2 shows the architecture of NAR neural network. The best fits of the number of hidden layer size and the feedback delays are 10 and 1:2 respectively.

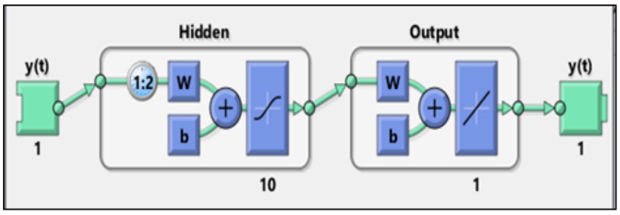

Fig. 2. Architecture of NAR neural network.

Levenberg-Marquardt backpropagation (LMBP) algorithm is used as the training function for ANN network training. The best validation performance for this model is obtained at the epoch point of 4 iterations as shown in Figure 3.

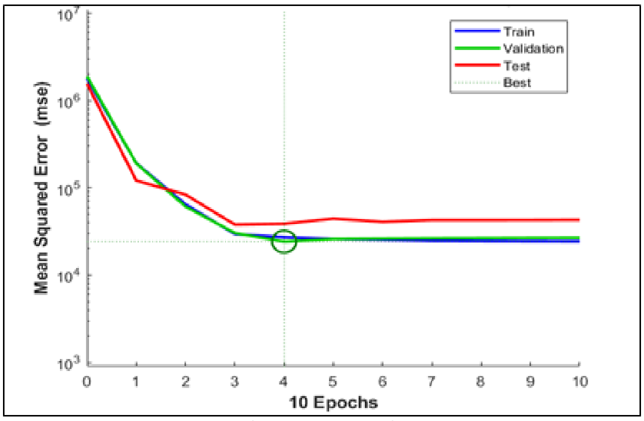

Fig. 3. Performance graph for NAR model

ACF plot and regression analysis for NAR model are carried out to test the adequacy of the model. Based on the validation tests, the model is adequate and has no autocorrelation. Hence, the model will be used for forecasting purpose.

D. Forecasting Performance Between ARIMA model and Artificial Neural Network

Table III shows the summary for forecasting performances of ARIMA (1,1,5) and NAR model. Based on the RMSE, MAE and MAPE values, model NAR is more accurate than ARIMA (1,1,5) as the values of RMSE, MAE and MAPE is smaller. However, based on Lewis (1982), both models show acceptable and highly accurate forecasting performance because the MAPE values for both models are less than 10%.

TABLE III. Summary for forecasting performance.

Model	RMSE	MAE	MAPE
ARIMA (1,1,5)	238.432	206.012	9.739
NAR	221.507	140.905	6.043

Figure 4 shows the comparison graph between actual palm oil price and forecasting price of ARIMA model and NAR model. The result shows that NAR model is more accurate in forecasting. Based on the graph, forecasting price using NAR model is more approached to the actual price. Therefore, in this study, model NAR has better forecasting performance.

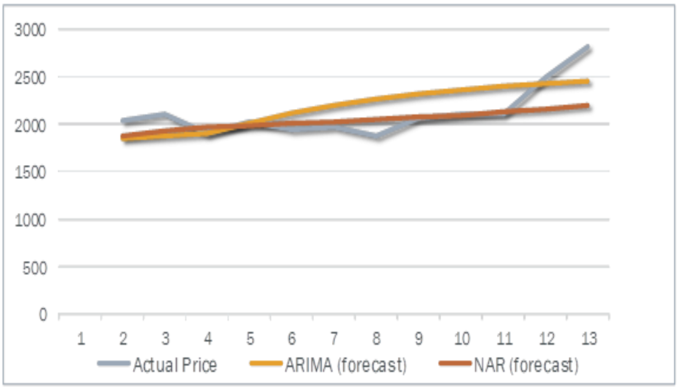

Fig. 4. Comparison graph between actual palm oil price and forecasting price of ARIMA model and NAR model.

IV. CONCLUSION

Palm oil prices fluctuate over time. Factors that influence the palm oil prices are demand and supply factor, extreme weather, price of vegetable oils especially soy bean oil, financial crisis and introduction of new policy. For forecasting, NAR model is more accurate than ARIMA model because it has lower RMSE, MAE and MAPE values. However, forecasting performance for both models are very accurate based on Lewis (1982), as MAPE values for both models are lower than 10%. Based on the graph comparison between actual palm oil price and forecasting price from ARIMA model and NAR model, NAR model shows better forecasting performance as it is more approached to the actual price. In this study, NAR model is the best model in forecasting palm oil price.

REFERENCES

[1] A. K. Ab Rahman. “Impact of Palm Oil Supply and Demand on Crude Palm Oil Price Behavior,” *Malaysia Palm Oil Board (MPOB)*, 2012.

[2] M. F. Abdul Hamid and A. Shabri. “Palm Oil Price Forecasting Model: An Autoregressive Distributed Lag (ARDL) Approach,” *American Institute of Physics*, 2017.

[3] F. Ali. “Reasons for price fluctuations,” *New Straits Times*, pp. 17, 2019.

[4] F. M. Arshad and A. A. A. Hameed. “Crude Oil, Palm Oil Stock and Prices: How They Link,” *Review of Economics & Finance*, vol. 3, pp. 48–57, 2013.

[5] M. A. M. Isa, A. T. Baharim, S. Mohamed, M. K. A. Noh, F. Nasrul, W. M. F. W. Ibrahim and S. S. Hassan. “Crude Palm Oil Price Fluctuation in Malaysia,” *International Journal of Academic Research in Business and Social Sciences*, vol. 10, issue 5, pp. 879–892, 2020.

[6] A. A. Karia and I. Bujang. “Progress Accuracy of CPO Price Prediction: Evidence from ARMA Family and Artificial Neural Network Approach,” *International Journal of Finance and Econometrics*, issue 64, pp. 66–79, 2011.

[7] A. A. Karia, I. Bujang and I. Ahmad. “Forecasting on Crude Palm Oil Prices Using Artificial Intelligence Approaches,” *American Journal of Operations Research*, vol. 3, issue 2, pp. 259–267, 2013.

[8] N. Khalid, H. N. A. Hamidi, S. Thinagar and N. F. Marwan. “Crude Palm Oil Price Forecasting in Malaysia: An Econometric Approach,” *Jurnal Ekonomi Malaysia*, vol. 52, issue 3, pp. 247–259, 2018.

[9] A. H. S. Md Nor, T. Sarmidi and E. Hosseinidoust. “Forecasting of Palm Oil Price in Malaysia using Linear and Nonlinear Methods,” *Statistics and Operational Research International Conference (SORIC 2013)*, pp. 138–152, 2014.

[10] N.F. Rahim, M. Othman and R. Sokkalingam. “A Comparative Review on Various Method of Forecasting Crude Palm Oil Prices,” *Journal of Physics: Conference Series*, 2018.

[11] N. M. Zuki. “Sumbangan Besar Minyak Sawit Malaysia,” *Sinar Harian*, 2020.