

# Crude Palm Oil Price Forecasting in Malaysia: An Econometric Approach (Peramalan Harga Minyak Sawit Mentah di Malaysia: Satu Pendekatan Ekonometrik)

**Norlin Khalid**

Universiti Kebangsaan Malaysia

**Hakimah Nur Ahmad Hamidi**

Universiti Kebangsaan Malaysia

**Sharmila Thinagar**

Universiti Kebangsaan Malaysia

**Nur Fakhzan Marwan**

Universiti Teknologi MARA (UiTM) Kedah

## ABSTRACT

*This paper aims to forecast the performance of crude palm oil price (CPO) in Malaysia by comparing several econometric forecasting techniques, namely Autoregressive Distributed Lag (ARDL), Autoregressive Integrated Moving Average (ARIMA) and Autoregressive Integrated Moving Average with exogenous inputs (ARIMAX). Using monthly time series data spanning from 2008 to 2017, the main results revealed that ARIMAX model is the most accurate and the most efficient model as compared to ARDL and ARIMA in forecasting the crude palm oil price. The results also show that the spot price of palm oil is highly influenced by stock of palm oil, crude petroleum oil price and soybean oil price. The empirical findings provide some insights for decision making and policy implementations, including the formulation of strategies to help the industry in dealing with the price changes and thus enable the Malaysian palm oil industry to continue dominating the international market.*

*Keywords: ARDL; ARIMA; ARIMAX; forecasting; crude palm oil prices*

## ABSTRAK

*Kertas ini bertujuan untuk meramal prestasi harga minyak sawit mentah di Malaysia dengan membandingkan beberapa teknik peramalan ekonometrik seperti kaedah auto regresif lat tertabur (ARDL), auto regresif purata bergerak (ARIMA) dan auto regresif purata bergerak berintegrasi dengan input eksogenous (ARIMAX). Dengan menggunakan data siri masa bulanan yang meliputi tahun 2008 hingga 2017, hasil kajian mendapati bahawa model ARIMAX adalah model yang paling tepat dan paling berkesan dalam meramalkan prestasi harga minyak sawit mentah di Malaysia berbanding dengan model ARDL dan ARIMA. Hasil kajian juga menunjukkan harga minyak sawit mentah dipengaruhi oleh stok minyak sawit, harga minyak petroleum mentah dan harga minyak kacang soya. Dapatan kajian ini memberikan beberapa implikasi penting kepada penggubalan dasar baharu bagi membantu industri sawit Malaysia untuk berdepan dengan kesan daripada perubahan harga minyak sawit mentah supaya industri ini terus teguh menguasai pasaran antarabangsa.*

*Kata Kunci: ARDL; ARIMA; ARIMAX; peramalan; harga minyak sawit mentah*

## INTRODUCTION

The Malaysian palm oil industry plays an important role in the economic development as well as enhancing the country's socioeconomic level. This has been proven by the contributions of Malaysian palm oil industry where, for example in 2015, it is the fourth largest contributor to the national income, accounting for RM63 billion of Malaysia's Gross National Income (GNI). With the experience in the palm oil industry for over 100 years, Malaysia has a comparative advantage in the international market and thus becomes the market leader in terms of productivity and research and development (R&D).

The government has planned and carried out various initiatives and frameworks in order to increase the domination of the Malaysian palm oil industry in the international market and therefore indirectly increase the country's income. Besides being a vital income source, Malaysia Palm Oil Council (MPOC) also reported that the Malaysia palm oil industry provides major employment opportunity for 650,000 smallholders and labor, and a further 3 million people in Malaysia whose livelihood is dependent upon this industry.

In order to heighten the contribution of this industry, there are few fundamental factors that need to be managed. One of the factors is the price of crude palm





FIGURE 1. Trend of Malaysia CPO Price

oil. This is important because price plays a major role in export activities where the revenues of the export will contribute to the national economic growth. Figure 1 shows that the price of CPO fluctuates over period. For example, in 1983, the price marks at RM2000/tonne, then it falls down to RM500/tonne in 1986. This scenario repeatedly happens along the time period and the gap of the price falls is quite big. Starting in 2008, the price trend shows a high volatility in which indicates the sensitiveness of palm oil prices to market changes. This fluctuation trend also raises concern for those that dealing with risks and uncertainties in the oil palm business and may harm the income of smallholders and will further affect the country's revenue. Therefore, it is crucial to understand the crude palm oil price behavior and its affecting factors.

Therefore, understanding the theory of price is relevance in which it stated that price for any specific good or service is based on the relationship between supply and demand (Friedman 1962). The main goal of this theory is to achieve equilibrium in which the quantities of goods or services provided match the corresponding market desire and ability to acquire the good or service. Hence, the fluctuations trend of Malaysian crude palm oil prices are due to the supply and demand factors. This is in line with the findings by Abdullah and Wahid (2010) and Ab Rahman (2012) in which both studies found that supply and demand factors are crucial in understanding the fluctuations of crude palm oil price.

Past literature have identified two main factors namely demand and supply side that affects the movement of crude palm oil price. Specifically, factors from demand side are, first, price of other vegetable oils (Amiruddin et al. 2005; Kumar et al. 2014; Sehgal et al. 2013), second, the economic growth (Amiruddin et al. 2005; Rosa & Vasciaveo 2008; Nazlioglu & Soytaş 2012), third, the exchange rates (Bui & Pippenger 1990; John 2007; Rosa & Vasciaveo 2008; Headey & Fan

2008; Cespedes & Velasco 2012) and lastly, the price of crude petroleum oil (Abdullah & Wahid 2010; Nazlioglu & Soytaş 2011). Interestingly, exchange rate is proved to be an important factor affecting the crude palm oil price. It is said that different regime of exchange rates gives different impact on the commodity prices. For instance, under the flexible exchange rate regime, the nominal exchange rate is allowed to move freely in response to supply and demand conditions in the foreign exchange market thus the CPO also behaves accordingly.

Meanwhile, the factors from the supply side are, first, trade barriers (Hasan et al. 2001; Obado et. al. 2009; Kelly et al. 2014; Abdulla et al. 2014), second, weather or climate change (John 2007; Rosa & Vasciaveo 2008; Headey & Fan 2008; Ong 2017) and third, production of palm oil (Rahman 2012 & 2013). It is found that the trade barrier such as the implementation of export tax may have a negative effect to the price competitiveness in the global market. Therefore, it is crucial to handle the export tax efficiently in order to guarantee the stabilization of the price while ensuring the future profitability. In terms of weather and climate change, the production of commodities could decline due to the adverse weather such as drought or flood in major producing and exporting countries. In some tropical climate countries like Malaysia and Indonesia, El Nino and La Nina phenomenon have possibly crippled the production of palm oil hence the price will be eventually increased due to adverse supply shock.

Furthermore, price forecasting is also important in order to ensure a stable future price and to help making wise decisions, therefore it is imperative to examine the determinants of price change, thus the fluctuations in the price over time could be properly anticipated. However, looking at the forecasting techniques which have been used in the past studies, most researchers were still using the conventional forecasting techniques like exponential smoothing and Box-Jenkins method (Arshad

& Ghaffar 1986; Shamsudin & Arshad 1991; Ab Rahman 2012). Nevertheless, recent studies have used different forecasting techniques in which it yields more reliable findings. For example, Pektas and Cigizoglu (2013) and Catalano et al. (2016) has used multivariate ARIMA (ARIMAX) and Artificial Neural Networks (ANN) models in their forecasting studies.

However, studies that focus on the behavior of palm oil price is quite limited in the case of Malaysia. Existing studies only consider a few variables in their research without considering other important variables that may have a big impact on palm oil price. Therefore, this present study tries to fill the gap by including relevant determinants such as total demand of palm oil, stock of palm oil, exchange rate as well as other substitution's price oils such as price of soybean oil and crude petroleum oil price in order to understand the Malaysian crude palm oil price movement. Hence, in contrast to most of previous literatures that used univariate approach in forecasting palm oil price, this study uses multivariate approach in which several new explanatory variables are included in forecasting the movement of CPO. Furthermore, in order to forecast the crude palm oil price in the future, both conventional and new techniques are utilised. Results from each method are then compared to identify the best model in forecasting the crude palm oil price. This study is important since Malaysia is one of the major exporting country for palm oil, and therefore the result may provide some insight in ensuring the competitiveness of Malaysia's palm oil sector in the global market. Relevant policy recommendations are then proposed based on the results to enhance export performance of Malaysian palm oil.

## LITERATURE REVIEW

Time series forecasting is a major challenge in many real world applications especially in agricultural commodities. Numerous studies have been carried out in agricultural commodities based on forecasting such as on cocoa bean (Assis et al. 2010; Fatimah & Roslan 1986), soybean oil (Gallagher 1987; Johnson 2014; Kenyon et al. 1993; Kumar et al. 2017), corn (Darekar & Reddy 2017a; Johnson 2014; Kenyon et al. 1993), grains (Darekar & Reddy 2017b; Marchezan & Souza 2010) and cottons (Darekar & Reddy 2017c). According to Zabid and Abidin (2015), in the world's oil and fats market, palm oil is one of the important agricultural commodities where both Indonesia and Malaysia are the top producing countries. However, the price of palm oil keeps fluctuating all over the time. Therefore, the accuracy of forecasting CPO is an important insight for the investors and government agencies in dealing with the associated risks and uncertainties (Arasim & Karia 2015; Karia et al. 2013). Hence, an appropriate model should be implemented in order to generate a reliable forecast

for CPO which can be a direction for policy makers. Many aspects should be taken into consideration for selection of the methods which includes the availability of data, financial support for software, expertise, and accuracy of the model as well as the significance level (Dewi et al. 2011; Md Nor et al. 2014). There are various and different forecasting techniques that have been introduced to analyze the time series prediction. The forecasting approach can be divided into two categories which are statistical and artificial intelligence (AI) based techniques. Karia et al. (2013) use an Artificial Intelligence (AI) approach to forecast the CPO prices. This study also suggests that the utilization of the Artificial Neural Network (ANN) and Autoregressive Fractionally Integrated Moving Average (ARFIMA) models to predict the aforementioned parameter with increased accuracy. The study shows that the ANN gives out accurate result compared to ARFIMA's. However, according to Karia and Bujang (2011), both models possess weakness where they require larger amounts of data. This issue is supported by Rahim et al. (2018) who state that excessive problems occur more frequently in neural network models compare to other traditional statistical models.

Due to that, many researchers opt to employ the traditional statistical models to forecast the CPO price (Abdul Hamid & Shabri 2016; Ahmad et al. 2014; Dewi et al. 2011; Karia et al. 2013; Karia & Bujang 2011; Khin et al. 2013; Rahim et al. 2018; Zhang et al. 1999). Ahmad et al. (2014) use the traditional Box-Jenkins approach to forecast monthly CPO prices. Then, the Autoregressive Integrated Moving Average (ARIMA) is applied in finding the suitable time series to its past values but the results of the residuals are orthogonal and not normal. Also, it is found that there are serial correlations in the series and there are also problems in deciding the proper order of model identification stage of ARIMA, such as parameters and residuals from the fitted model. Thus, re-identification of a model becomes necessary because an incorrectly identified model gives rise to inaccurate estimations. Khin et al. (2013) attempt to explore which is the best comparative forecasting model between the VECM, MARMA and ARIMA based on the RMSE, MAE, RMPE and U-Theil criteria for the Malaysian monthly spot oil palm price. The result reveals that the MARMA is the most accurate and efficient model in forecasting the spot oil palm price. Besides that Abdul Hamid and Shabri (2016), use the ARDL model to compare it with ARIMA model. The data structure is modified to enable to compare the result from this experiment to an ARIMA model and found out that this model outperforms ARIMA model in term of forecasting accuracy.

Furthermore, most of the previous studies focus on the factors affecting the palm oil prices (Ab Rahman et al. 2012; Asari et al. 2011; Hassan & Balu 2016; Mohammadi et al. 2015;). According to Ab Rahman et al. (2012), the error-correction model result shows that the changes in the lagged crude palm oil future prices

do effectively influence the changes in the spot price. Whereas, Hassan and Balu (2016) revealed that an increase in palm oil price will produce negative responses in soybean oil price, total export and production in the short-term. This study also suggests that the information on the trends of total production and total export of palm should be included in determining the palm oil price. This is in line with the study done by Asari et al. (2011) where the result shows that the production of palm oil in Malaysia can influence its price level. Besides that, the simulation result by Mohammadi et al. (2015) indicates that the Malaysian CPO price is significantly affected by the local CPO production as well as the soybean oil prices. It is also proven that an increase in the soybean price will push the CPO price further as the demand for CPO increases. Therefore, this study suggests that the need for forecasting the CPO prices will significantly facilitate the efficient decisions by the investors and policy makers.

## RESEARCH METHODOLOGY

### DATA AND VARIABLES

In order to forecast the prices for the Malaysian palm oil industry, a time series data are being employed in this study. In line with past studies such as Amiruddin et al. (2005), Rosa and Vasciaveo (2008), Abdullah and Wahid (2010), Nazlioglu and Soytaş (2011) and Kumar et al. (2014), this present study include the following variables (see Table 1) that are expected to highly influenced the spot price of palm oil. The frequency of data set are monthly time series spanning from 2008 to 2017 in which are collected from the Bloomberg database. All the variables are then expressed in terms of natural logarithms.

TABLE 1. Description of data

Variables	Description
SPPO	Spot Price of Palm Oil (MYR/Ton)
TDPO	Total Demand of Palm Oil ('000 Ton)
STPO	Stock of Palm Oil ('000 Ton)
CPO	Crude Petroleum Oil Price (USD/Barrel)
REER	Real Effective Exchange Rate (RM/USD)
PSBO	Price of Soybean Oil (MYR/Ton)

### ECONOMETRIC SPECIFICATION

In terms of methodology, this study uses three different methods to analyze the determinants of palm oil spot price and to forecast the spot price in the future. The details of each method are discussed as follows;

### Autoregressive Distributed Lag (ARDL) Model

The ARDL model was developed by Pesaran and Shin (1995, 1999), Pesaran et al. (1996) and Pesaran (1997). Basically, an ARDL model is given as follows;

$$\begin{aligned} \Delta LSPPO_t = & \alpha_0 + \sum_{i=1}^p \phi_i \Delta LSPPO_{t-i} + \sum_{i=0}^q \theta_i \Delta LTDPO_{t-i} \\ & + \sum_{i=0}^r \lambda_i \Delta LSTPO_{t-i} + \sum_{i=0}^s \gamma_i \Delta LCPO_{t-i} \\ & + \sum_{i=0}^t \sigma_i \Delta LREER_{t-i} + \sum_{i=0}^u \beta_i \Delta LPSBO_{t-i} \\ & + \delta_1 LSPPO_{t-1} + \delta_2 LTDPO_{t-1} + \delta_3 LSTPO_{t-1} \\ & + \delta_4 LCPO_{t-1} + \delta_5 LREER_{t-1} + \delta_6 LPSBO_{t-1} \\ & + v_t \end{aligned} \tag{1}$$

where,  $\Delta$  is first-difference operator and  $p, q, r, s, t, u$  are the optimal lag length. Then, the first step is to estimate the existence of a long-run relationship (cointegration) among the variables. The null hypothesis for no cointegration among variables in the equation (1) is  $H_0: \delta_1 = \delta_2 = \delta_3 = \delta_4 = \delta_5 = \delta_6 = 0$ , against the alternative hypothesis  $H_1: \delta_1 \neq \delta_2 \neq \delta_3 \neq \delta_4 \neq \delta_5 \neq \delta_6 \neq 0$ . If the estimated F-statistic value is found to exceed the upper limit critical value, the null hypothesis is rejected, in which a long-term relationship (cointegration) among the time series variables is established. On the other hand, if F-statistics is less than the lower bound critical value, this indicates the null hypothesis is failed to be rejected. In addition, if the value of the estimated F-statistic is between the lower limit and the upper limit, then it cannot be identified (inconclusive) whether there is a cointegration or not because the integral degree of integration of the variable is unknown.

If there is an evidence of a long-run relationship (cointegration) of the variables, then the long-run model is estimated as follows;

$$\begin{aligned} LSPPO_t = & \alpha_1 + \sum_{i=1}^p \phi_i LSPPO_{t-i} + \sum_{i=0}^q \theta_i LTDPO_{t-i} \\ & + \sum_{i=0}^r \hat{\phi}_i LSTPO_{t-i} + \sum_{i=0}^s \hat{\gamma}_i LCPO_{t-i} \\ & + \sum_{i=0}^t \hat{\phi}_i LREER_{t-i} + \sum_{i=0}^u \omega_i LPSBO_{t-i} + \mu_t \end{aligned} \tag{2}$$

For the last step, an ARDL specification of the short-run dynamics will be derived by constructing an error correction model (ECM) of the following form:

$$\begin{aligned} \Delta LSPPO_t = & \alpha_2 + \sum_{i=1}^p \phi_{2i} \Delta LSPPO_{t-i} + \sum_{i=0}^q \theta_{2i} \Delta LTDPO_{t-i} \\ & + \sum_{i=0}^r \hat{\phi}_{2i} \Delta LSTPO_{t-i} + \sum_{i=0}^s \hat{\gamma}_{2i} \Delta LCPO_{t-i} \\ & + \sum_{i=0}^t \hat{\phi}_{2i} \Delta LREER_{t-i} + \sum_{i=0}^u \omega_{2i} \Delta LPSBO_{t-i} \\ & + \psi ECM_{t-1} + \vartheta_t \end{aligned} \tag{3}$$

where  $ECM_{t-1}$  is an error correction model that measures the speed of adjustment ( $\psi$ ) in the direction of long-term equilibrium, which is the time taken by the dependent variable to converge to long-run equilibrium. In addition,  $ECM_{t-1}$  can also explain the long run causality between all explanatory variables on spot price of palm oil.

#### *Autoregressive Integrated Moving Average (ARIMA) Model*

Autoregressive Integrated Moving Average (ARIMA) model has been used extensively in time series analysis. The model has also been proven to be the most prominent method in financial forecasting, where it has efficient capability to generate the short-term forecast (Meyler et al., 1998; Pai and Lin, 2005; Merh et al., 2010). It has constantly outperformed other complex structural models with regards to the short-term prediction. In ARIMA model, the future value of a variable is a linear combination of past values and past errors. The ARIMA ( $p, q$ ) model is expressed as follows;

$$Y_t = \delta + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots + \alpha_p Y_{t-p} + \varepsilon_t - \beta_2 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \dots + \beta_q \varepsilon_{t-q} \quad (4)$$

where,  $Y_t$  is the actual value,  $\varepsilon_t$  is the random error at time  $t$ ,  $\alpha_i$  and  $\beta_i$  are the coefficients,  $p$  and  $q$  are integers that are often referred to as autoregressive and moving average, respectively.

#### *Autoregressive Integrated Moving Average with exogenous input (ARIMAX) Model*

An ARIMAX model is generalized from univariate Autoregressive Integrated Moving Average (ARIMA) model. The notation is given as ARIMAX ( $p, q, b$ ), in which  $p$  are autoregressive terms,  $q$  are moving average terms and  $b$  are exogenous input terms. In the other words, an ARIMAX model is a linear regression model that includes exogenous or independent variable in the ARIMA type model for residuals. The model can be expressed as follows;

$$Y_t = \alpha_0 + \sum_{i=1}^p \delta_i Y_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i} + \sum_{i=1}^b \sigma_i X_{t-i} + \varepsilon_t \quad (5)$$

where,  $p$ ,  $q$  and  $b$  are integer or autoregressive, moving average and exogenous variable respectively. Four main statistics can be derived from equation (5) namely Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percent Error (MAPE) and Theil inequality coefficients. These statistics are used to test the accuracy of the forecasting models.

The best forecasting model are chosen based on the smallest values of RMSE, MAE, MAPE and Theil coefficient. All four statistics are calculated as follows;

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T (y_t - \hat{y}_t)^2} \quad (6)$$

$$MAE = \frac{1}{T} \sum_{t=1}^T |y_t - \hat{y}_t| \quad (7)$$

$$MAPE = \frac{1}{T} \sum_{t=1}^T \left| \frac{y_t - \hat{y}_t}{y_t} \right| \quad (8)$$

$$Theil = \frac{\left[ \sum_{t=1}^T (\hat{y}_t - y_t)^2 \right]^{1/2}}{\left[ \sum_{t=1}^T y_t^2 \right]^{1/2}} \quad (9)$$

where  $\hat{y}$  is the forecast value,  $y$  is the actual value in time  $t$ .

## EMPIRICAL RESULT

### DETERMINANTS OF PALM OIL PRICES

All variables are tested for their order of integration using Augmented Dickey-Fuller (ADF) and non-parametric Phillips-Perron (PP) test. The unit-root properties of the series are crucial for the cointegration and causality analysis. The results of unit-root tests are presented in Table 2. It is clear that the null hypothesis of a unit roots for LSPPO, LCPO, LREER and LPSBO at log-level based are failed to be rejected since the statistic values are all insignificant. However, at their first differences, the null hypothesis can be rejected as the statistic values are all significant at 1% significance level. This indicates that LSPPO, LCPO, LREER and LPSBO are stationary and integrated of order 1, or I(1). In the case of LTDPO, the null hypothesis of a unit roots at log-level base is rejected at the 1% significance level for both tests, in which indicates that the series of LTDPO is stationary at level or integrated of order 0, I(0). Therefore, it is found that all the variables tested is a mixture of I(0) and I(1) and consequently, the ARDL model can be used in estimating the baseline model.

Table 3 reports the results of cointegration test (long-run relationship) using the bound test. The choice of optimum lagged that used in the estimation model is determined by minimizing the Akaike Information Criterion (AIC). As can be seen in in Table 3, the calculated F-statistics is 3.92 which is higher than the upper bound critical value 3.79 at 5% significance level. This implies that the null hypothesis of no cointegration among the variables is rejected and thus, the existence of a long run relationship among the variables are confirmed. In the other words, these variables are moving together and would not move too far from each other in the long run. Moreover, it also shows that each independent variable plays a significant role in influencing the movement of palm oil spot price.

Table 4 shows the empirical result of the long run model in which obtained by normalizing on the spot price

TABLE 2. Unit root test

Variables	ADF		PP		Test Equation
	I(0)	I(1)	I(0)	I(1)	
LSPPO	-2.3527	-7.9074***	-2.8926**	-7.9378***	Constant
	-2.2912		-2.8459		Constant with Trend
LTDPO	-5.9509***	-11.8119***	-4.4163***	-16.1745***	Constant
	-5.9807***		-4.3755***		Constant with Trend
LSTPO	-3.4798**	-8.0012***	-2.4727	-7.7969***	Constant
	-3.4773**		-2.4309		Constant with Trend
LCPO	-1.7969	-6.9362***	-1.5577	-6.8967***	Constant
	-1.3739		-1.8943		Constant with Trend
LREER	-1.1368	-8.7282***	-0.9179	-8.5752***	Constant
	-1.9547		-1.7621		Constant with Trend
LPSBO	-1.9889	-8.2103***	-2.3867	-8.2103***	Constant
	-1.9243		-2.3884		Constant with Trend

Note: \*, \*\* and \*\*\* indicates a significance at 10%, 5% and 1% respectively.

TABLE 3. Bound cointegration test

F-statistics	Lag	Significance Level	Bound Critical values	
			I(0)	I(1)
3.92	(1,0,2,0,1,2)	1%	3.41	4.68
		5%	2.62	3.79
		10%	2.26	3.35

Note: Bound critical values are based on Pesaran et al. (2001)

of palm oil. The most significant variables that affect the price of palm oil (SPPO) are stock of palm oil (STPO) and crude petroleum oil (CPO). These are the two common factors that affecting the price of palm oil (Abdullah & Wahid 2010). The negative relationship between stock of palm oil and its price indicates that when the stock of palm oil is increasing, it will push down the prices of the palm oil. Moreover, the development in the petroleum sector and its impact on the environment has caused many countries to consider using alternative renewable energy, for example vegetable oils. Consequently, this created additional demand for vegetable oils, including palm oil. Therefore, if the price of petroleum decrease, it may discourage the government from implementing

biodiesel using palm oil, in which will affect the demand and the prices in the market.

The result of the error correction model for price of palm oil is presented in Table 5. The error correction term ( $ECT_{t-1}$ ) for SPPO is found to be negative and statistically significant at 1% significance level, in which indicates the existence of the stable long-run equilibrium. The value of  $ECT_{t-1}$  (-0.192554) indicates that a deviation from the equilibrium level of palm oil price during the current period will be corrected by 19% in the next period and it takes about five years and two months to stabilize and to achieve the equilibrium when there is a shock. Diagnostic tests are also carried out and there are no evidence of serial correlation and heteroscedasticity effect in the disturbance terms. The model also passes the Jarque-Bera normality test in which suggests that the errors are identically and normally distributed (iid). In terms of stability, CUSUM and CUSUM square tests have been performed and the graphs plotted as shown in Figure 2 and Figure 3, in which indicate that the proposed model using ARDL(1,0,2,0,1,2) is stable.

FORECASTING THE FUTURE PRICES OF PALM OIL

The model specification was determined by looking at the autocorrelation function (ACF) and partial autocorrelation function (PACF) behavior as shown in

TABLE 4. Long-run model

Dependent variables	Independent variables				
	LTDPO	LSTPO	LCPO	LREER	LPSBO
LSPPO	-0.1738	-0.2938*	0.2609**	-0.9024	0.2402

Note: \*, \*\* and \*\*\* indicates a significance at 10%, 5% and 1% respectively.

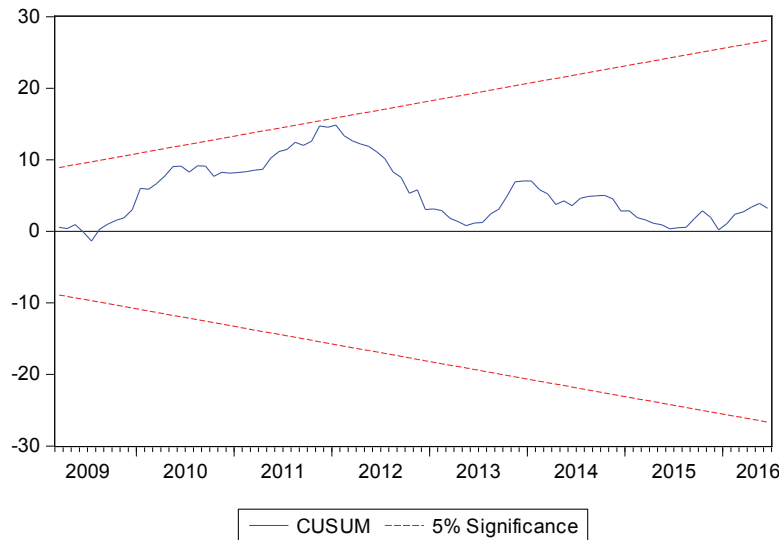


FIGURE 2. CUSUM

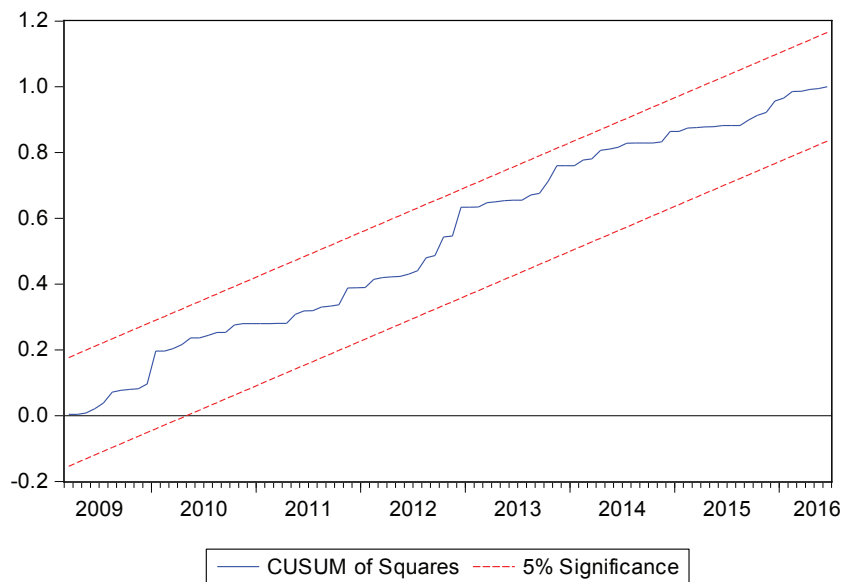


FIGURE 3. CUSUM of Square

TABLE 5. ECM-ARDL cointegration

	$ECT_{(t-1)}$	LM	ARCH	Jarque-Bera	CUSUM
$\Delta LSPPO$	-0.1926 (0.0001)	0.4326 (0.7847)	0.7212 (0.5796)	0.0012 (0.9993)	Stable

Note: Value in the parenthesis indicates the p-values.  $ECT_{(t-1)}$  is an error-correction term which measure the long-run relationship. LM is Lagrange multiplier for serial correlation test, ARCH is for heteroscedasticity test and JB is Jarque-Bera normality test. Null hypothesis for LM, ARCH and JB are the residuals are not serially correlated, residuals are homocedastic and residuals are normally distributed respectively.

Table 6. Both ACF and PACF did not show any clear pattern, whether it is tails off or cuts off, so several simplest model using ARIMA(1,1,0) and ARIMA(0,1,1) are tested. Although both models are significant, the residuals are not white noise. The Q-statistics test for serial correlation and Q-squared statistic test for the

presence of heteroscedasticity are both significant, in which indicates that the residual of the estimated model suffers from autocorrelation and heteroscedasticity problem. This result suggests that an ARIMA(1,1,0) and ARIMA(0,1,1) are inappropriate model to forecast the palm oil price. Then, various possible models are

estimated by increasing the lags and using an ARMA model. However, the AR and MA parameters are not significant and the residuals are also found do not follow white noise behavior.

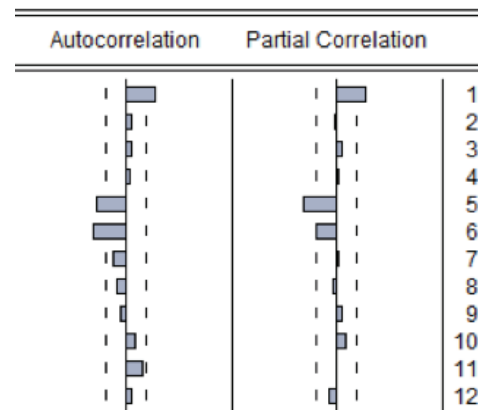
Then, an ARIMA (|1,5,6|,1,0) model as ACF is estimated and PACF is significant at lag 1, 5 and 6. The coefficients of ARIMA (|1,5,6|, 1,0) model are all significant and the value of log-likelihood is also higher than AR(1) model and MA(1) model, in which indicates that ARIMA (|1,5,6|,1,0) model may be an adequate model. To make sure that the model fits the data well, its adequacy is checked using the the residual diagnostic tests. Both Q-statistics and Q-squared statistics are insignificant, in which imply that the residuals of ARIMA (|1,5,6|,1,0) model are free from autocorrelation and heteroscedasticity problems. In other words, the residuals of ARIMA (|1,5,6|,1,0) model is white noise. Therefore, it is confirmed that ARIMA (|1,5,6|,1,0) model is an appropriate model to forecast the spot price of palm oil. The ARIMA (|1,5,6|,1,0) model can be mathematically written as follows:

$$\begin{aligned}
 D(LSPPO)_t = & -0.003170 - 0.244461D(LSPPO)_{t-1} \\
 & (0.6266) \quad (0.0043) \\
 & - 0.240360D(LSPPO)_{t-5} \\
 & (0.0119) \\
 & - 0.189751D(LSPPO)_{t-6} \quad (10) \\
 & (0.0477)
 \end{aligned}$$

Note: value in parenthesis ( ) indicates a p-value.

For ARIMAX model, ARIMAX (|2|,1,|2|) is found to be the best model as the AR and MA parameters are statistically significant and the residuals follow a white noise. Equation (11) represents ARMAX (|2|,1,|2|) model mathematically. Based on the equation (11), the price of soybean oil (LPSBO) is the most important variables in forecasting the spot price of palm oil as it is statistically significant at 1% significance level. According to Prasetyo et al. (2017), the soybean oil price reveals a strong influence towards crude palm oil exports as it holds the position as substitute of crude palm oil. This finding is in line with Prasetyo et al. (2017), where crude palm oil price is found to be strongly influenced by soybean oil price as both commodities are substitutes for each other. The high demand of soybean oil will be reduced if its price increases, leading to consumption switching to palm oil. Hence the import demand of palm oil will rise, causing the exporting countries such as Indonesia and Malaysia to increase their palm oil production. This shows that soybean oil is a long term competitor of palm oil. There are a few number of studies in which have highlights the issue of CPO price and soybean price. For example, study by Chuangchid et al. (2012), Kochaphum et al. (2013), and Talib & Darawi (2002) have found that there is a positive relationship between palm oil and soybean oil price.

TABLE 6. ACF & PACF



$$\begin{aligned}
 D(LSPPO)_t = & -0.161259 - 0.022607LTDPO_t \\
 & (0.6110) \quad (0.6125) \\
 & - 0.117215LSTPO_t + 0.028528LCPO_t \\
 & (0.2016) \quad (0.5795) \\
 & + 0.335418LREER_t + 1.086882LPSBO_t \\
 & (0.3793) \quad (0.0000) \\
 & - 0.894850D(LSPPO)_{t-2} \\
 & (0.0015) \\
 & + 0.845420\varepsilon_{t-2} \quad (11) \\
 & (0.0142)
 \end{aligned}$$

Note: value in parenthesis ( ) indicates a p-value.

To double check the compatibility and stability of the models chosen, an AR/MA roots graph have been plotted. The graphs are shown in Figure 4 and Figure 5 . It could be concluded that the ARIMA (|1,5,6|,1,0) and ARIMAX (|2|,1,|2|) models are stable since the roots lies within a unit circle.

In order to identify the best forecasting model among these three models, the forecasting accuracy needs to be performed. This study utilizes the value of Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percent Error (MAPE) and Theil inequality coefficient to test the accuracy of the forecasting model. The results are shown in Table 7. The model that has the lowest error value is said to be the best model for forecasting. Thus, based on the result in Table 7, ARIMAX model shows a smallest error value as compared to ARIMA and ARDL models. Finally, in order to confirm our results, a graph using the three forecasting models is plotted against the actual values. Figure 6 shows that ARIMAX forecast value falls closer to the actual values as compared to ARIMA and ARDL models. Therefore, it is confirmed that ARIMAX (|2|,1,|2|) is the most accurate and efficient to forecast the spot price of palm oil.

### SUMMARY AND CONCLUSIONS

This study has utilized three different methods in investigating the determinants of crude palm oil price



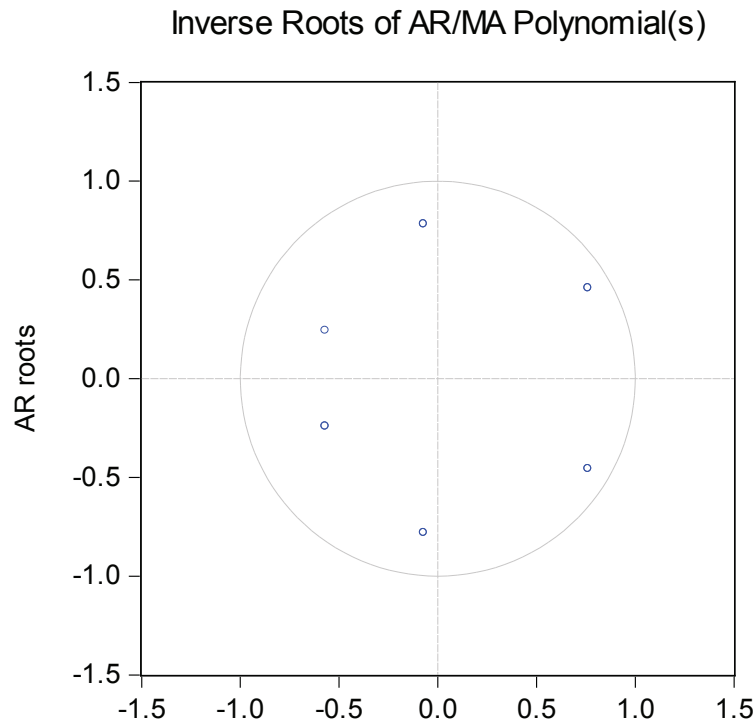


FIGURE 4. AR/MA roots graph for ARIMA (|1,5,6|,1,0)

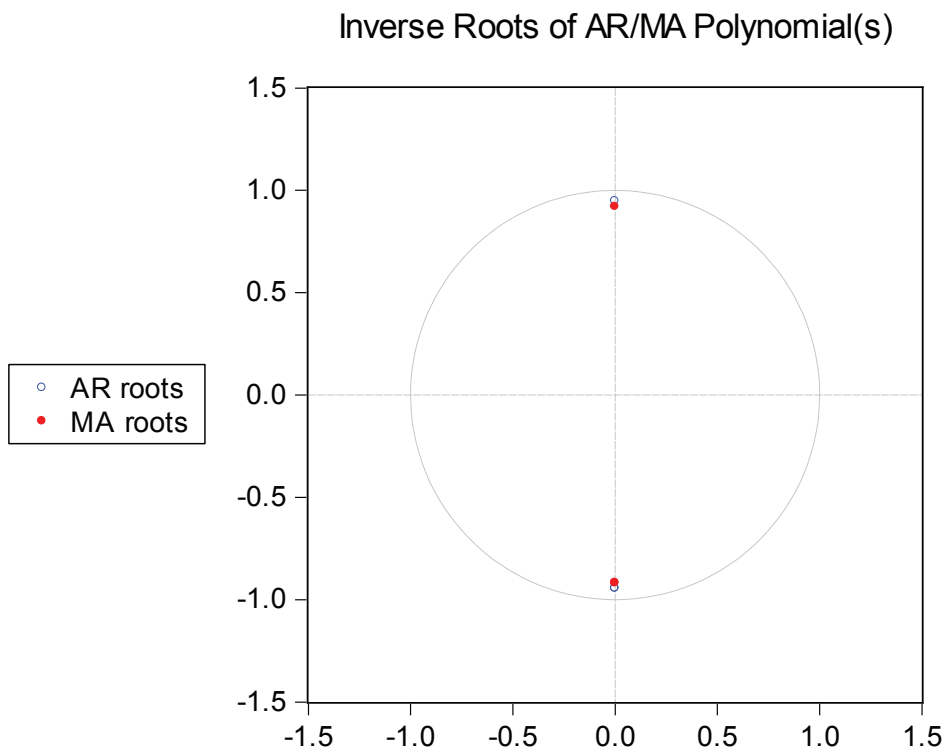


FIGURE 5. AR/MA roots graph for ARIMAX (|2|,1,|2|)

and to forecast the crude palm oil price in future. Based on the result using ARDL model, the long run relationship among the variables is confirmed, in which revealed that all independent variables play a significant role in

determining the movement or behavior of crude palm oil price. Interestingly, the result also shows that the most significant variables that affect the crude palm oil price are the stock of palm oil and the crude petroleum oil

TABLE 7. Model specification

Model	AIC	SIC	Log Likelihood	Q-Stats	Q <sup>2</sup> -Stat
ARIMA(1,1,0)	-2.3001	-2.2225	119.1574	15.989 (0.007)	18.043 (0.006)
ARIMA(0,1,1)	-2.2969	-2.2193	118.9978	15.643 (0.008)	17.026 (0.009)
ARIMA (  1,5,6  ,1,0)	-2.4182	-2.3146	126.1189	2.0712 (0.558)	10.578 (0.102)
ARIMAX (  2  ,1,  2  )	-3.1546	-2.9216	168.3096	3.7772 (0.437)	8.5198 (0.202)

Note: value in parenthesis ( ) indicates a p-value. Symbol || || means the model are test on that specific lag. AIC denotes the Akaike's Information Criterion, SBC denotes Schwartz Bayesian Criterion, Q-statistics is the modified portmanteau test with null hypothesis are residuals exhibit no serial correlation and Q<sup>2</sup>-statistics is a test for heteroscedasticity with null hypothesis are residuals are homocedastic.

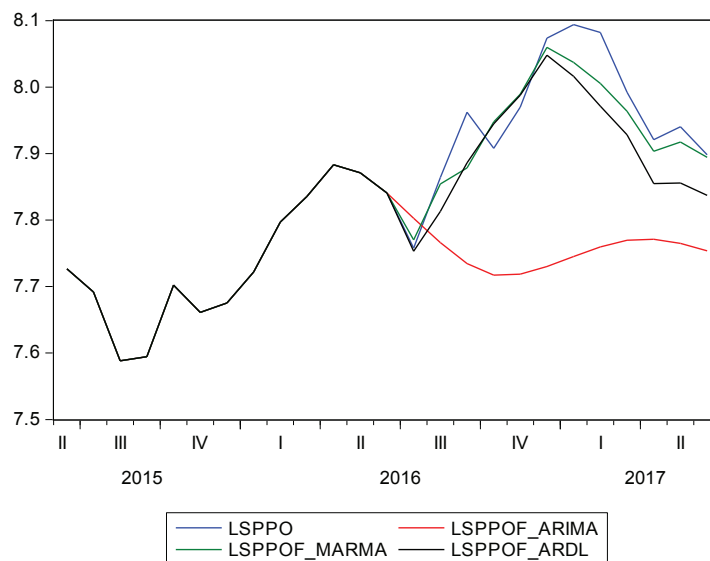


FIGURE 6. Model evaluation of ex-ante forecast of SPPO from July 2016 to June 2017

price. This finding shows that palm oil has an advantage to be used in biodiesel production and has a potential to attract more consumers in the future. Therefore, the government and policy makers may use this information in order to enhance the growth of palm oil industry in the future.

For forecasting, ARIMAX model seems to be the best model to forecast the Malaysian palm oil price. The fluctuations of crude oil price may lead policy makers to alter their budgetary plans for more investment grants and incentives in the palm oil market. The ARIMAX results also indicate that the spot price of palm oil (SPPO) is highly influenced by the soybean oil price (PSBO). Thus, it is important to consider the relationship between palm oil and soybean oil market in order to design a prudent strategy to stabilize the palm oil price due to movement of soybean price.

Finally, the need for increasing research and development (R&D) and commercialization prospects of the palm oil industry in Malaysia is crucial due to the immense biotechnological advancement in the competing

crops such as soybeans, and the subsidies provided by some of soybeans producing country have deteriorated the palm oil competitiveness in Malaysia. The problems that arise within the industry need to be resolved promptly so that the industry will continue to develop orderly. These actions are much needed by the oil palm industry to boost their productivity and efficiency as well as competitiveness.

#### ACKNOWLEDGMENT

We would like to thank Malaysia Palm Oil Board (MPOB) for funding this research via MPOB-UKM Endowment Chair Research Grant (EP-2015-068).

#### REFERENCES

Ab Rahman, N. M. N., Nawi, A. S. & Nik Muhd Naziman, Y. H. 2012. The Price Discovery of the Malaysian Crude

- Palm Oil Futures Markets. *Journal of Applied Finance and Banking* 2(4): 25-47.
- Ab Rahman, A. K. 2012. Impact of Palm Oil Supply and Demand on Crude Palm Oil Price Behavior. *Malaysia Palm Oil Board (MPOB)*.
- Abdul Hamid, M. F. & Shabri, A. 2016. Palm Oil Price Forecasting Model: An Autoregressive Distributed Lag (ARDL) Approach. *The 3rd ISM International Statistical Conference 2016 (ISM-III)*, doi:10.1063/1.4982864
- Abdulla, I., Arshad, F.M., Bala, B.K., Noh, K.M., & Tasrif, M. 2014. Impact of CPO export duties on Malaysian palm oil industry. *American Journal of Applied Sciences* 11 (8): 1301-1309.
- Abdullah, R., & Wahid, M. B. 2010. World Palm Oil Supply, Demand, Price and Prospects: Focus On Malaysian and Indonesian Palm Oil Industry. *Malaysia Palm Oil Board (MPOB)*.
- 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.
- Amiruddin, M. N., Rahman, A. K. A., & Shariff, F. 2005. Market Potential and Challenges for the Malaysian Palm Oil Industry in Facing Competition from Other Vegetable Oils. *Oil Palm Industry Economic Journal* 5(1): 17-27.
- Arasim, D. H. & Karia, A. A. 2015. Identifying and Forecasting the Factors that Derive CPO Prices in Malaysia Using NARX Model. *Case Studies Journal* 4(12): 4-14.
- Arshad, F.M., & Ghaffar, R.A. 1986. Crude palm Oil Price Forecasting: Box Jenkins Approach. *PERTANIKA* 9(3): 359 - 367.
- Asari, F.F.A.H., Abd Rahman, N.H., Razak, E.A.A., Ahmad, B.A.S., Harun, N.F.A & Jusoff, K. 2011. A Time Series Analysis of the Relationship between Total Area Planted, Palm Oil Price and Production of Malaysian Palm Oil. *World Applied Sciences Journal*: 34-40.
- Assis, K., Amran, A., Remali, Y. & Affendy, H. 2010. A Comparison of Univariate Time Series Methods for Forecasting Cocoa Bean Prices. *Trends in Agricultural Economics* 3(4): 207-215.
- Bui, N., & Pippenger, J. 1990. Commodity Prices, Exchange Rates and Their Relative Volatility. *Journal of International Money and Finance* 9: 3-20.
- Catalano, M., Galatioti, F., Bell, M., Namdeo, A., & Bergantino, A. S. 2016. Improving the Prediction of Air Pollution Peak Episodes Generated by Urban Transport Networks. *Environmental Science and Policy* 60: 69-83.
- Cespedes, L. F., & Velasco, A. 2012. Macroeconomic Performance During Commodity Price Booms and Busts. *IMF Economic Review* 6(4): 570-599.
- Chuangchid, K., Wiboonpongse, A., Sriboonchitta, S. & Chaiboonsri, C. 2012. Factors Affecting Palm Oil Price Based on Extremes Value Approach. *International Journal of Marketing Studies* 4(6): 54-65.
- Darekar, A. & Reddy, A. A. 2017a. Price Forecasting of Maize in Major States. *Maize Journal*, 6.
- Darekar, A. & Reddy, A. A. 2017b Forecasting of Common Paddy Prices in India. *Journal of Rice Research* 10(1): 71-75.
- Darekar, A. & Reddy, A. A. 2017c. Cotton Price Forecasting in Major Producing States. *Economic Affairs* 62(3): 373-378.
- Dewi, S., Arshad, F. M., Shamsudin, M. N. & Hameed, A. A. 2011. An Econometric Analysis of the Link between Biodiesel Demand and Malaysian Palm Oil Market. *International Journal of Business and Management* 6(2): 35-45.
- Friedman, M. 1962. *Price Theory*. Transaction Publisher, New Brunswick, New Jersey.
- Gallagher, P. 1987. U.S Soybean Yields: Estimation and Forecasting with Nonsymmetric Disturbances. *American Journal of Agriculture Economics*: 796-803.
- Hassan, A. & Balu, N. 2016. Examining the Long-term Relationships between the Prices of Palm Oil and Soyabean Oil , Palm Oil Production and Export: Cointegration and Causality. *Oil Palm Industry Economic Journal* 16(1): 31-37.
- Hasan, M. F., Reed, M. R. & Marchant, M. A. 2001. Effects of an Export Tax on Competitiveness: The Case of the Indonesian Palm Oil Industry. *Journal of Economic Development* 26(2): 77-90.
- Headey, D., & Fan, S. 2008. Anatomy of a Crisis: The Causes and Consequences of Surging Food Prices. *Agricultural Economics* 39: 375-391.
- Johnson, D. M. 2014. An Assessment of Pre- and Within-Season Remotely Sensed Variables For Forecasting Corn and Soybean Yields in the United States. *Remote Sensing of Environment* 141: 116-128.
- Karia, A. A. & Bujang, I. 2011. Progress Accuracy of CPO Price Prediction: Evidence From ARMA Family and Artificial Neural Network Approach. *International Research Journal of Finance and Economics* (64): 66-79.
- Karia, A. A., Bujang, I. & Ahmad, I. 2013. Forecasting on Crude Palm Oil Prices Using Artificial Intelligence Approaches. *American Journal of Operations Research* 3: 259-267.
- Kelly, K. S. W., Shamsudin, M. N., Mohamed, Z. & Sharifuddin, J. 2014. Effects of Export Duty Structure on the Performance of the Malaysian Palm Oil Industry. *Journal of Food Products Marketing* 20(1): 193-221.
- Kenyon, D., Jones, E. & McGuirk, A. 1993. Forecasting Performance of Corn and Soybean Harvest Futures Contracts. *American Journal of Agriculture Economics* 75: 399-407.
- Khin, A. A., Mohammed, 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.
- Kochaphum, C., Gheewala, S. H. & Vinitnantharat, S. 2013. Does Biodiesel Demand Affect Palm Oil Prices in Thailand? *Energy for Sustainable Development* 17(6): 658-670.
- Kumar, M. M., Acharya, D. & Suresh, B.M. 2014. Price Discovery and Volatility Spillovers in Futures and Spot Commodity Markets: Some Indian Evidence *Journal of Advances in Management Research* 11(2): 211-226.
- Kumar, M., Raman, R. K. & Kumar, S. 2017. Forecasting of Soybean Yield in India Through ARIMA Model. *International Journal of Pure & Applied Bioscience* 5(5): 1538-1546.
- Marchezan, A. & Souza, A. M. 2010. Forecasting the Price of Major Grains Produced in Rio Grande. *Ciencia Rural* 40(11): 2368-2374.
- Md Nor, A. H. S., Sarmidi, T. & Hosseinidoust, E. 2014. Forecasting of Palm Oil Price in Malaysia Using Linear and Nonlinear Methods. *Statistics and Operational*

- Research International Conference (SORIC 2013)*: 138-152.
- Mohammadi, S., Arshad, F. M., Bala, B. K. & Ibragimov, A. 2015. System Dynamics Analysis of the Determinants of the Malaysian Palm Oil Price. *American Journal of Applied Sciences* 12(5): 355-362.
- N. Merh, V.P. Saxena, and K.R. Pardasani. 2010. A Comparison Between Hybrid Approaches of ANN and ARIMA For Indian Stock Trend Forecasting. *Journal of Business Intelligence* 3(2): 23-43.
- Nazlioglu, S., & Soytas, U. 2012. Oil Price, Agricultural Commodity Prices and The Dollar: A Panel Cointegration and Causality Analysis. *Energy Economics* 34: 1098-1104.
- Obado, J., Syaukat, Y. & Siregar, H. 2009. The Impacts of Export Tax Policy on the Indonesian. *Journal ISSASS* 15(2): 107-119.
- Obado, J., Syaukat, Y. & Siregar, H. 2009. The Impacts of Export Tax Policy on the Indonesian Crude Palm Oil Industry. *Journal International Society for Southeast Asian Agricultural Sciences*, 15 (2), 107-119.
- Ong, Z. 2017. Palm Oil Producers Report Better Numbers. *The Edge Malaysia*, 22 May.
- Othman, N., Yusop, Z. & Ismail, M. M. 2018. Identifying Export Competitiveness of Malaysia's Processed Palm Oil: Using Shift-Share Analysis. *International Journal of Accounting, Finance and Business* 3(12): 1-11.
- Pektas, A. O., & Cigizoglu, H. K. 2013. ANN Hybrid Model Versus ARIMA and ARIMAX Models of Runoff Coefficient. *Journal of Hydrology* 500: 21-36.
- Pesaran, H. M. 1997. The Role of Economic Theory in Modelling The Long-Run. *Economic Journal* 107: 178-191.
- Pesaran, H. M. & Shin, Y. 1999. Autoregressive Distributed Lag Modelling Approach to Cointegration Analysis. In: S. Storm (Ed.) *Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial Symposium*, chapter 11 (Cambridge: Cambridge University Press).
- Pesaran, M. H. & Shin, Y. & Smith, R. J. 1996. Testing for the Existence of a Long-run Relationship, *Cambridge Working Papers in Economics* 9622, Faculty of Economics, University of Cambridge.
- Pesaran, M.H. & Shin, Y. 1995. An Autoregressive Distributed Lag Modelling Approach to Cointegration Analysis. *Cambridge Working Papers in Economics* 9514, Faculty of Economics, University of Cambridge.
- Prasetyo, A., Marwanti, S. & . D. 2017. The Influence of Exchange Rate on CPO Exports of Indonesia. *Jurnal Ekonomi Pembangunan: Kajian Masalah Ekonomi dan Pembangunan* 18(2): 159
- Rahim, N. F., Othman, M. & Sokkalingam, R. 2018. A Comparative Review on Various Method of Forecasting Crude Palm Oil Prices. *Journal of Physics: Conference Series* doi:10.1088/1742-6596/1123/1/012043.
- Rahim, N. F., Othman, M., Sokkalingam, R. & Abdul Kadir, E. 2018. Forecasting Crude Palm Oil Prices Using Fuzzy Rule-Based Time Series Method. *IEEE Access*, 6, 32216-32224.
- Rosa, F. & Vasciaveo, M. 2008. Agri-Commodity Price Dynamics: The Relationship Between Oil and Agricultural Market. *International Association of Agricultural Economists (IAAE) Triennial Conference*, Brazil.
- Sazmand-asfaranjan, Y. & Ziaei-moayyed, A. 2012. Assessing Export Opportunity of Malaysia's Palm Oil Products : Using Shift-Share Technique. *European Journal of Scientific Research* 70(2): 323-333.
- Sehgal, S., Rajput, N. and Deisting, F. 2013. Price discovery and volatility spillover: Evidence From Indian Commodity Markets, *The International Journal of Business and Finance Research* 7(3): 57-75.
- Shamsudin, M. N., & Arshad, F.M. 1991. Short-term Forecasting of Crude Palm Oil Prices. *Paper Presented at the PORIM International Palm Oil Conference*, 9 September, Kuala Lumpur.
- Suherman, T. 2016. The Import Demand Analysis of Indonesian and Malaysian Palm Oil in Italy. *International Journal of Science and Research*, 5(10), 1326-1329.
- Talib, B. A. & Darawi, Z. 2002. An Economic Analysis of the Malaysian Palm Oil Market. *Oil Palm Industry Economic Journal* 2(1): 19-27.
- Zabid, F. & Abidin, N. Z. 2015. Palm oil industry : A Review of the Literature on the Modelling Approaches and Potential Solution. *Innovation and Analytics Conference and Exhibition (IACE 2015)*, doi:10.1063/1.4937023.
- Zhang, G., Hu, M. Y., Patuwo, B. E. & Indro, D. C. 1999. Artificial Neural Networks in Bankruptcy Prediction: General Framework And Cross-Validation Analysis. *European Journal of Operational Research* 116: 16-32.

Norlin Khalid\*

Fakulti Ekonomi dan Pengurusan  
Universiti Kebangsaan Malaysia  
43600 UKM Bangi Selangor  
MALAYSIA  
E-mail : nrlin@ukm.edu.my

Hakimah Nur Ahmad Hamidi  
Fakulti Ekonomi dan Pengurusan  
Universiti Kebangsaan Malaysia  
43600 UKM Bangi Selangor  
MALAYSIA  
E-mail : hakimahnur@hotmail.com

Sharmila Thinagar  
Fakulti Ekonomi dan Pengurusan  
Universiti Kebangsaan Malaysia  
43600 UKM Bangi Selangor  
MALAYSIA  
E-mail : vanessa91.sham@gmail.com

Nur Fakhzan Marwan  
Fakulti Perniagaan dan Pengurusan  
Universiti Teknologi MARA (UiTM) Kedah  
08400 Merbok, Kedah.  
MALAYSIA  
E-mail : nurfakhzan@kedah.uitm.edu.my

\*Corresponding author

## APPENDIX

TABLE 8. Comparison of the performance of forecasting techniques

Value Error Test	ARDL	ARIMA	ARMAX
RMSE	0.0635	0.2292	0.0411
MAE	0.0564	0.2099	0.0321
MAPE	0.7069	2.6268	0.4021
Theil's	0.0040	0.0146	0.0026

