

Comparative Study on Forecasting Crude Palm Oil Price using Time Series Models

Azme Bin Khamis, Raed Hameed, Maria Elena Nor, Norziha Che Him,
Rohayu Mohd Salleh & Siti Noor Asyikin Mohd Razali

Department of Mathematics and Statistics,
Faculty of Applied Science and Technology,
Universiti Tun Hussein Onn Malaysia

DOI: 10.31364/SCIRJ/v6.i12.2018.P1218588

<http://dx.doi.org/10.31364/SCIRJ/v6.i12.2018.P1218588>

Abstract: Forecasting crude palm oil price is important, particularly when the investors encounter with the increasing risks and uncertainties in the future. Therefore, the aim of this study is to forecast the price of palm oil in Malaysia for the next years based on price for the period of 31 years. The objective of the research is to propose an appropriate model to forecast the CPO price. Thus, this study proposes three types of models, which are namely: Autoregressive Integrated Moving Average (ARIMA), Autoregressive Conditional Heteroskedasticity (ARCH) and Generalized Autoregressive Conditional Heteroskedasticity (GARCH). Akaike Information Criterion (AIC) and Hannan-Quinn Criterion (H-Q) statistic were used to obtain the best model. It was found that ARIMA (2, 1, 5) performed better compared to ARCH and GARCH models. It is concluded that ARIMA (2, 1, 5) can be used as an alternative model to forecast the CPO price.

Keywords: *crude palm oil price, ARIMA model, ARCH model, GARCH model*

I. INTRODUCTION

Malaysia is one of the largest producers and exporters of palm oil and its products. Currently, the industry is flourishing and at the same time providing employment to its people. Thus, the oil palm industry is a major contributor to Malaysian export revenues. Crude palm oil (CPO) is produced from oil palm which then further processed and fractionated to produce a wide range of food and non-food palm products.

Malaysia has played an important part in supporting consumption and remaining competitive in the world's oils and fats market. As, the main consumer and business market for palm oil is the food industry and for this, the major importers are India, China and the European Union. India is the largest and leading consumer of palm oil worldwide, importing nearly 7.8 million tons in 2014. The European Union (EU27) is the second biggest importer of palm oil and China become third largest importer of palm oil importing about 5.7 million tons in 2014 (USDA, 2014). The current production of the world palm oil suggests an increased by 32% to nearly 60 million tons by 2020. Thus, the modelling and forecasting of CPO prices is significant so as to hold valuable data relating to the future of CPO prices. Therefore, the forecasting of crude palm oil (CPO) price is imperative for the decision makers in order to make the right decisions and also to enhance the managerial decision-making process in several areas, such as industry, farming. Moreover, the high or low CPO price can cause economic catastrophe among dependent countries in all aspects of life.

Box-Jenkins approach was applied to forecast monthly crude palm oil price (Fatimah and Ghaffar, 1986). Khin et al; 2013), CPO price is volatile where the conditional variance of the price series changes between high and low values. The current study attempts to model CPO using the popular univariate model of Autoregressive Integrated Moving Average (ARIMA) and improve the forecasts the ARIMA model with Generalized Autoregressive Conditional Heteroskedasticity (GARCH) (Christ, *et. al.*, 2010).

The objective of this study is to propose forecasting model for palm oil price in Malaysia. Statistical software such as Minitab 16 and Gretl 10.2 was used to analyze the data. There are three models are considered in this study, which are ARIMA (autoregressive integrated moving average), ARCH (autoregressive conditional heteroskedasticity) and GARCH (Generalized Autoregressive Conditional Heteroskedasticity) models. Nanda (1998) and Khin (2011) studied on multiple forecasting models, included ARIMA, to predict the short-term future prices of palm oil in Malaysia where the short term ex ante and ex post forecasting was being generated. Nochai & Titida Nochai (2006), studied on ARIMA model for Forecasting oil palm price and the model accuracy was measured using minimum of mean absolute percentage error (MAPE). Mohd Arshad *et al.* (1986) studied on crude palm oil price forecasting using univariate Box-Jenkins model. Karia *et, al* (2012) conducted a study on forecast of crude palm oil price using ARIMA and Artificial

Neural Network (ANN). Md Nor *et al.*, (2014) studied on forecasting of palm oil price in Malaysia used ARIMA models, neural networks and fuzzy logic systems. Ahmad *et al.*, (2014), studied on volatility modeling and forecasting of CPO prices. Philbertha & Achmad (2014), studied on volatility analysis international CPO used Arch, Arch-M and GARCH models. Zainudin (2013), studied GARCH model, meanwhile You & Yeap (2015) used Box Jenkins and regression model for forecasting purposes. It is really essential for investors and portfolio managers have a superior forecast of the volatility as this is a sound starting point for assessing investment. Figure 1 showed the fluctuation in Malaysian palm oil price.

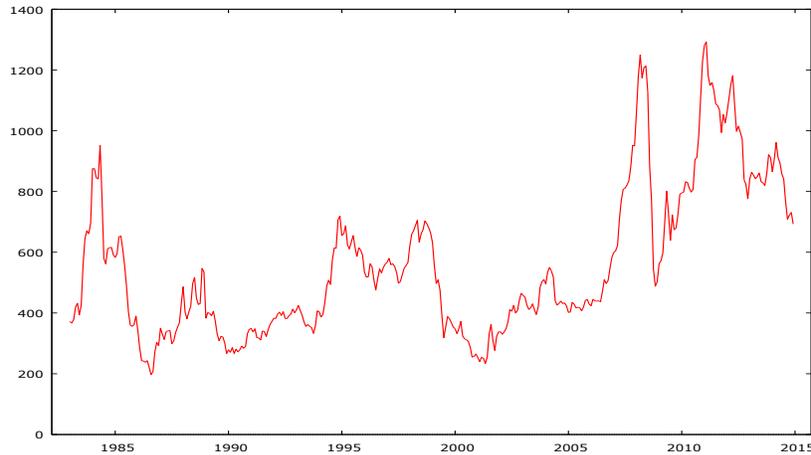


Figure 1: Malaysia CPO Price (Jan 1983-Dec 2015)

However, the instability of palm oil price can create significant risks to producers, suppliers, consumers, and other stakeholders. In risky conditions and price instability, an accuracy of palm oil price analysis is very important in supporting policymakers to make informed decisions and development of their own countries (Chuangchid, 2013).

II. RESEARCH METHODOLOGY

This study will consider three models to forecast the CPO prices, which is ARIMA (p, d, q), ARCH (p) and GARCH (p, q) models.

(i) ARIMA Model

ARIMA model is a popular time series modeling developed by Box and Jenkins and can be written as ARIMA (p, d, q) which allocates seasonality can be written as follows (Weber, 2007); (Harvey & Todd, 1983)

$$\phi_p(B)\phi_p(B^S)\nabla^d\nabla_S^p y_t = \theta_q(B)\theta_q(B^S)a_t \quad (1)$$

The autoregressive process of order p is denoted AR (p), and process is more generally defined by

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t \quad (2)$$

Where ϕ_1, \dots, ϕ_p are fixed constants and ε_t is a sequence of independent (or uncorrelated) random variable with mean 0 and variance σ^2 .

The moving average process of order q is denoted MA (q) are determined by

$$X_t = \sum_{s=1}^q \theta_s \varepsilon_{t-s} \quad (3)$$

where $\theta_1, \dots, \theta_q$ are fixed constants, $\theta_0 = 1$, and ε_t is a sequence of independent (or uncorrelated) random variable with mean 0 and variance σ^2 . We remark that two moving average processes can have the same autocorrelation function. For instance, $X_t = \varepsilon_t + \theta\varepsilon_{t-1}$ and $X_t = \varepsilon_t + \frac{1}{\theta}\varepsilon_{t-1}$. Both have, $\rho_k = 0, |k| > 1$. However, the first gives

$$\varepsilon_t = X_t - \theta\varepsilon_{t-1} = X_t - \theta(X_{t-1} - \theta\varepsilon_{t-1}) = X_t - \theta X_{t-1} + \theta^2 X_{t-2} - \dots$$

This only valid for $|\theta| < 1$, a so-called invertible process. There are no two invertible processes have the same autocorrelation function.

The ARIMA process is the process $\{Y_t\}$ is said to be an autoregressive integrated moving average process, ARIMA (p, d, q). If the original process is not stationary, we can expect at the first order difference process

$$X_t = \nabla Y_t = Y_t - Y_{t-1} \quad (4)$$

or the second order differences

$$X_t = \nabla^2 Y_t = \nabla(\nabla Y_t) = Y_t - 2Y_{t-1} + Y_{t-2} \quad (5)$$

If $d = 0$, that means the data already stationary then ARMA (p, q) will be imposed.

(ii) *ARCH Model*

The ARCH process allowed the conditional variance to change over time as a function of past errors leaving the unconditional variance constant (Eagle, 1982; Brooks, 2014). Empirical evidence revealed that the ARCH model required a relatively long lag in the conditional variance equation and to ward off the problems with negative variance parameters, a fixed lag structure was imposed. ARCH model can be defined class ($P \geq I$) the following formula;

$$\begin{aligned} Y_t &= \mu + X_t \\ X_t &= \sigma_t \varepsilon_t \\ \sigma_t^2 &= \alpha_0 + \alpha_1 X_{t-1}^2 \dots \alpha_p X_{t-p}^2 \end{aligned} \tag{6}$$

Where Y_t represents return series, which is linked to a series of interrelated. μ represents the average (return series), either ε_t is independent of the series and identical distribution and follow the standard normal distribution with an average 0 and variance 1 and called on the Equation (6) mean equation .

Equation (7) represent a volatility equation which can be rewritten as

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2 \tag{7}$$

where as $\alpha_0 > 0$ and $\alpha_i \geq 0$ for $i > 0$ represent a parameter mode when $p = 1$ we have a model ARCH (1) of the first rank as:

$$\begin{aligned} Y_t &= \mu + X_t \varepsilon_t \\ X_t &= \sigma_t \varepsilon_t \\ \sigma_t^2 &= \alpha_0 + \alpha_1 X_{t-1}^2 \end{aligned} \tag{8}$$

Equation (8) is a special case when $p=1$ (Pereira, 2004).

(iii) *GARCH Model*

Bollerslev (1986) proposed a generalized ARCH (GARCH) to surmount the restrictions of the traditional ARCH model of Engle (1982). The GARCH model allowed for both a longer memory and a more flexible lag structure. In the ARCH process, the conditional variance is defined as a linear mapping of past sample variance only, whereas the GARCH process allows lagged conditional variances to enter in the model as well. Both the ARCH and GARCH models of Engle (1982) and Bollerslev (1986) could not differentiate how the variance of the return was influenced differently by positive and negative values.

General equation for GARCH models is

$$\begin{aligned} Y_t &= \mu + X_t \varepsilon_t \\ X_t &= \sigma_t \varepsilon_t \\ \sigma_t^2 &= \alpha_0 + \alpha_1 X_{t-1}^2 \dots \alpha_p X_{t-p}^2 + \beta_1 X_{t-p}^2 + \dots + \beta_q X_{t-q}^2 \end{aligned} \tag{9}$$

where Y_t represents the series return and μ average series, ε_t they are independent and identical distribution chain and keep track of the standard normal distribution with an average 0 and variance 1 and can rewrite the as equation (9) The following formula (Pereira, 2004) :

$$\sigma_t^2 = \alpha_0 \sum_{i=1}^p \alpha_i X_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2 \tag{10}$$

where $\alpha_0 \geq 0$, $\alpha_i \geq 0$, $\beta_1 \geq 0$ for $i > 0, j > 0$ represents the model parameters. When $p = 1$ and $q = 1$, and GARCH model can be written as

$$\begin{aligned} Y_t &= \mu + X_t \\ X_t &= \sigma_t \varepsilon_t \\ \sigma_t^2 &= \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \end{aligned} \tag{11}$$

The model was compared by using AIC and H-Q statistically to identify if the appropriate to be used in forecasting purposes.

III. RESULTS AND DISCUSSION

The sample data consisted of 384 observations of the monthly price of palm oil in Malaysia. It covered a period of thirty one (31) years spanning from January, 1983 to December, 2014. The data were obtained from the Malaysia Palm Oil Broad (MPOB). The descriptive statistics of monthly CPO price. The time series data were measured in (USD) per tonne. The mean of the monthly price of palm oil was 556.19 with a standard deviation of 244.40. The maximum price was 1292 whilst the minimum price was 197.

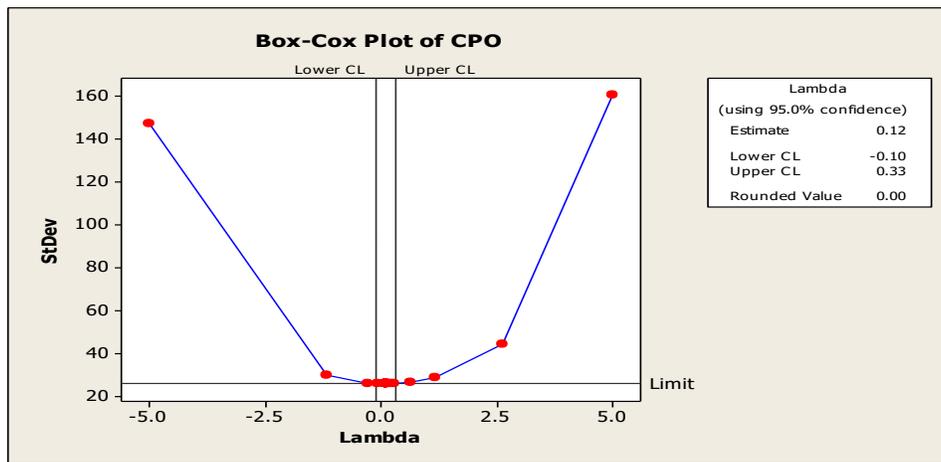


Figure 2 Box-Cox Plot of CPO Price

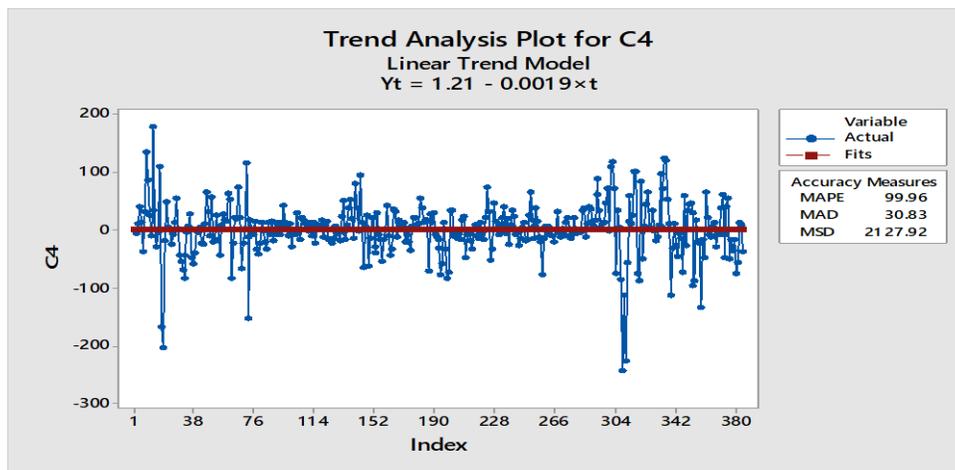


Figure 3: CPO time series plot when $d = 1$

In Figure 2 shows that the value of lambda (λ) value contains 0.12 in the 95% confidence interval, which is between -0.10 to 0.33. Figure 3 shows CPO plot after differencing and it is clear observation that the monthly CPO was stationary.

The model was selected based on AIC and H-Q statistics. Table 1 shows the values of AIC and H-Q for each model. ARIMA (2, 1, 5) recorded the AIC and H-Q statistics as 3942.684 and 3956.789 respectively. While ARCH (1) model was recorded AIC value at 4843.502 and H-Q was 4848.789 and for GARCH (3, 1) was recorded AIC and H-Q values as 4839.736 and 4849.138 respectively.

Table 1. The values of AIC and H-Q for respective model

Model	AIC	H-Q
ARIMA (2,1,5)	3942.694	3956.789
ARCH (1)	4843.502	4848.203
GARCH (3,1)	4839.736	4849.138

Based on AIC and H-Q values, we can conclude that the ARIMA (2, 1, 5) model is the most appropriate representation of the data and therefore it would be used to forecast the future values of the monthly price palm oil. The equation of ARIMA (2, 1, 5) can be written as;

$$Y_t = 0.793 + 1.389Y_{t-1} - 0.771Y_{t-2} + \epsilon_t + 1.056\epsilon_{t-1} - 0.276\epsilon_{t-2} - 0.332\epsilon_{t-3} - 0.194\epsilon_{t-4} + 0.222\epsilon_{t-5}$$

The Ljung-Box statistics test is 6.0838 which is statistically significant and supports that the model is appropriate and can be used to forecast future CPO price. Figure 4 shows the CPO forecast for year 2015-2016. It can be seen that the forecast values are close to the actual values.

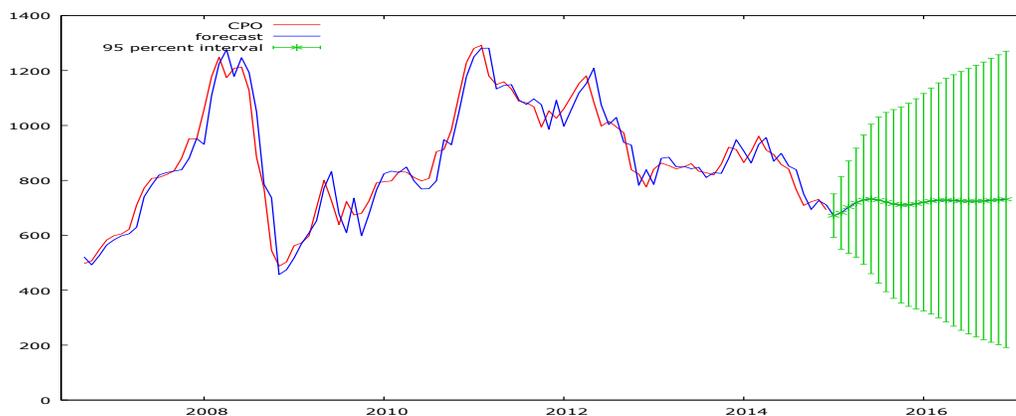


Figure 4. Forecast and actual values of CPO for year 2015-2016

IV. CONCLUSION

Even though heteroscedasticity problem exist in the CPO prices but ARCH and GARCH models failed to performance compared to ARIMA model. AIC and H-Q statistics confirmed that ARIMA model is outperformed and can be used as alternative model to forecast future CPO price.

V. REFERENCE

- 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.
- Bollerslev, T. (1986). Generalized autoregressive condition heteroskedasticity. *Journal of Econometrics*, 31 (3), 307-327.
- Box, G. E., Jenkins, G. M., & Reinsel, G. C. (2011). *Time series analysis: forecasting and control* Vol. 734. John Wiley & Sons.
- Brooks, Chris (2014). *Introductory Econometrics for Finance* (3rd ed.) Cambridge: Cambridge University Press.
- Chuangchid, K., Sriboonchitta, S., Rahman, S., & Wiboonpongse, A. (2013). Predicting Malaysian palm oil price using extreme value theory. *International Journal of Agricultural Management*, 2(2), 91-99.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987-1007.
- Fatimah Mohd Arshad and Roslan A. Ghaffar, Crude Palm Oil Price Forecasting: Box-Jenkins Approach, *Pertanika*, 9 (3), (1986), 359 – 367.
- Go, Y. H., & Lau, W. Y. (2015). Evaluating the hedging effectiveness in crude palm oil futures market during financial crises. *Journal of Asset Management*, 16(1), 52-69.
- Hanssens, D. M., Parsons, L. J., & Schultz, R. L. (2003). *Market response models: Econometric and time series analysis* Vol. 12. 2nd Edition. Springer Science & Business Media. Kluwer Academic Publishers.
- Harvey, A. C., & Todd, P. H. J. (1983). Forecasting economic time series with structural and Box-Jenkins models: A case study. *Journal of Business & Economic Statistics*, 1(4), 299-307.
- Hyndman, R. J., & Athanasopoulos, G (2014) *Forecasting: principles and practice*. OTexts.org/fpp.
- Karia, A. A., Bujang, I., & Ahmad, I. (2012). Forecasting on crude palm oil prices using artificial intelligence approaches. *The Business & Management Review*, 3(1), 355.
- Khin A. A. (2011). *Econometrics forecasting models for short term natural rubber prices*. Saarbrcken, Germany: Lambert Academic Publishing.
- Khin Aye Aye, Mohamed Zainalabidin, Nambhi Malarvizhi, (2013), Price Forecasting Methodology of the Malaysian Palm Oil Market, *International Journal of Applied Economics & Finance*, 7 (1).
- Mohd Arshad, F., & A Ghaffar, R. (1986). Crude palm oil price forecasting: Box-Jenkins approach. *Pertanika*, 9(3), 359-367.
- Nanda, S. (1988). Forecasting: Does the Box-Jenkins method work better than regression? *Vikalpa*, 1, 53-62.
- Nochai, R., & Nochai, T. (2006). ARIMA model for forecasting oil palm price. In *Proc. 2nd IMT-GT Regional Conf. on Mathematics, Statistics and Applications, University Sains Malaysia, Penang*.
- Nor, A. H. S. M., Sarmidi, T., & Hosseinidoust, E. (2014). Forecasting of palm oil price in Malaysia using linear and nonlinear methods. In *Statistic and Operational Research International Conference* Vol.1613, pp. 138-152.
- Pereira, P. L. V. (2004). *How persistent is volatility? An answer with stochastic volatility models with Markov regime switching state equations*. Finance Lab, Insper Instituto de Ensino e Pesquisa.
- Philbertha Gentha Pramana & Achmad Herlanto Anggono (2014). Volatility analysis international crude palm oil price 2004 – 2014. *Proceedings of 7th Asia-Pacific Business Research Conference 25 - 26 August 2014, Singapore*
- United States Department of Agriculture (USDA), (2014) *Foreign Agricultural Service, World Markets and Trade*.
- Weber, R. (2007). *University Of Cambridge*. Retrieved from Statistical Laboratory:TimeSeries: [www.statslab. cam.ac.uk/~rrw1/timeseries/t.pdf](http://www.statslab.cam.ac.uk/~rrw1/timeseries/t.pdf)

Zainudin, R. (2013). The effect of regime shift in minimum variance hedging ratio: the evidence of the crude palm oil market. *Investment Management and Financial Innovations*, Vol. 10 (4).