A Comparative Review on Various Method of Forecasting Crude Palm Oil Prices

Nur Fazliana Rahim*, Mahmod Othman and Rajalingam Sokkalingam

Fundamental and Applied Sciences Department, Universiti Teknologi PETRONAS,
32610 Seri Iskandar, Perak, Malaysia

*corresponding author’s e-mail address:nur_g03687@utp.edu.my

Abstract. Malaysia is very well-known as the world leader of palm oil in terms of production and export. As of today, palm oil has contributed the most in the Malaysian agricultural sector. This illustrates the importance of the palm oil industry in Malaysia. In situations of considerable uncertainty and therefore high risk, accurate and reliable price forecasting is necessary to facilitate decision making. In doing so, this study developed a CPO price forecasting method using Fuzzy Time Series with the proposed of Sliding Window Method. Besides, the concept of Fuzzy Rule Based Systems (FRBS) also was embedded in the application of Fuzzy Time Series. The aim of the proposed method is to enhance the effectiveness of time series forecasting and to provide higher predicting accuracy. The dataset of Crude Palm Oil (CPO) prices were taken from Malaysian Palm Oil Board (MPOB) and used to compare forecasting performance between several method such as Autoregressive Moving Average, ARIMA (Box-Jenkins model), Artificial Neural Network (ANN) and Hybrid ARIMA-ANN methods. The accuracy of all methods was compared to each other to determine the best forecasting method. In this study, the results showed that the forecast value of CPO price using all the method produces good and reliable CPO prices forecasting values. In spite of that, the forecast error of the proposed method and hybrid ARIMA-ANN method is lesser compared to other methods. This can be summarized that both these methods can reduce and perform better forecasting values, yet the hybrid ARIMA-ANN was known with its complexity of the method. Meanwhile, Fuzzy Time Series forecasting with the proposed method provides more proper and simpler method to forecast CPO price. Hence, the findings of this study could be used as an alternative method for CPO price forecasting to obtain a better forecast values.

1. Introduction

Palm oil has been introduced in Malaysia in 1911 and can be reflected a future economy of the world [1]. Palm oil is important because it is cheaper and efficient biofuel that mount the world population [2]. Its offer job opportunity to million people besides it improves Malaysia’s Gross Domestic Product (GDP). Not only that, oil palm has contributed to national income includes an investment of capital, expertise, foreign hands, and managing knowledge. However, the increasing competitive of oil palm industry was due to the growth productions of the oil palms from all over the world.

Crude Palm Oil Overview

Crude Palm Oil (CPO) is vitally important and well thought out as an encouragement for constructing the economy of evolving countries. It is realized that industry of oil palm can provide opportunity of
employment and growth earnings of poor population if it is systematically managed [3]. Besides, oil palms give high return, which motivated more Malaysian farmers to involve in the oil palm industry, rather than other industries. In addition, oil palm also produces better income compared to other fruit and vegetables. However, the researcher in [4] claims that the impact of oil palm to GDP is different from the researcher in [5], where its play as a key role in prompting infrastructural progress in producing area such as telephone service provider, clothes and education.

Due to instability of CPO prices, many researches have been done to forecast CPO price since it will give an impact to the organizational whose depend too much on this agricultural product. In [6], the researcher has summarized that Malaysia palm oil is subjected to significant price. However, the trend of CPO price is unclear and it is more unstable compared to other products. Therefore, the best forecasting method need to be finding in order to reduce the uncertainties and risks of palm oil trading world. The researcher in [7] mention that the increase in CPO price trends is important in the agriculture and industrial sector due to the higher cost of production.

Therefore, the instability of CPO prices is related to the significant price fluctuations. Due to this issue, there are risk and uncertainty that would be faced if too much depending on this agricultural production. Besides, many employments of workers to increase the oil palm production and high price of oil palm impacts more investment on this sector. The unstable of oil palm price becomes a threat for other developing countries that depend on the industry. The oil palm price does not only affect the oil palm producers, but it extends to many other industries.

Therefore, in order to overcome the risk, there is a need to find proper forecasting method to forecast CPO prices. Nevertheless, despite the numerous available models, a more precise and highly accurate forecasting model is still a major concern. Therefore, there is a need to adopt an approach in the quest to find a more accurate measurement that can facilitate the forecasting performance. It is evidenced that there has been some shortcoming in the common factor, thus, this study introduces a systematic method to determine the number of class intervals and provide a specific method to generate the rules in Fuzzy Time Series forecasting to minimize the prediction error. The results were compared with other forecasting methods according to the lowest errors produced from all the methods.

Generally, this study reviewed the previous method that has been done by previous researcher for CPO price forecasting and compared with the proposed method. The results indicate that the proposed method produce more accurate forecasting value compare to other methods. The CPO price forecasting is vital as it can become a guideline to make better choice and improvement in oil palm industry. Furthermore, the forecasting in the CPO price offers a future planning advice for organizations.

2. Previous Forecasting Methods

The fluctuation of CPO price is very important to the industry, business, and government agencies. It triggers for the suitable method to be put in place to create a reliable and rigorous forecast for the CPO price. Various method have been introduced and the best method which can reduce error, and some predefined assumption were selected. The most widely method used to forecast the time series data is from the ARIMA family i.e. the Box-Jenkins time series model. This model is well-known to construct accurate forecast for small sample and could avoid the multivariate model problem. The ARIMA models are able to define the various descriptions of temporal rows and are flexible [8].

However, the Box-Jenkins model has the issue to determine the appropriate order in the model identification stage of ARIMA, for example the residual and parameters from the fitted model. The estimation of the autoregressive (AR) and moving average (MA) order will become difficult. The wrong model identification will give the incorrect model estimation stage and make the researcher to have to run the identification of the model once again. It relatively needs a large amount of data and wasting time when compared to traditional forecasting method [9]. This weakness increases the risk to opt the models with huge errors.

To overcome the limitation of ARIMA model, the researcher in [10] have proposed Artificial Neural Network (ANN) due to its ability in capturing the nonlinear estimation and give accurate for time series forecasting. However, ANN is found as not consistent because of the usage of nonlinear data, which is in contrast to the linear data usually found in statistical methods [11]. Moreover, ANN model has complex underlying time series characteristics. It needs large data samples because the model needs a big number of parameters. The problem to get the higher frequency of historical data series is about the same
encountered in ARIMA [10]. Besides, ANN can prevent capture in local minima and over fitting as mentioned in [12] and was supported by the researcher in [13] where the problem is prone to happen in neural network models. Compared with hybrid and fuzzy logic models, this becomes the major disadvantage of ANN model.

In forecasting, it was proven by the researcher in [14] that a hybrid model has more accuracy and is superior to the ANN and ARIMA models. Because of the impediment to the Box-Jenkins and ANN methods, a hybrid model was developed. It is a good strategy to solve these limitations due to its ability can simultaneously model either linear or nonlinear. Moreover, hybrid model produce better forecast compare to ARIMA and ANN methods [12]. Besides, the researcher in [15] proposed a hybrid ARIMA-ANN model for time series forecasting and the results gave a small percent error of forecasting. However, the hybrid model is difficult to be implemented because of the complexity of the model. Due to this issue, this research proposed fuzzy time series method, which designed an evolutionary algorithm, which provides proper method to determine the number of class intervals and specific method to generate rules in Fuzzy Time Series forecasting, indirectly able to minimize the prediction error.

3. Methodology

3.1. Fuzzy Time Series (FTS)

This study starts with data collection and pre-processing the historical monthly CPO price data. The past five-year data were taken from the Malaysian Palm Oil Board (MPOB), for the duration of January 2012 to December 2016. Fuzzy Time Series approach was implemented in this study in order to forecast CPO price. Figure 1 below depicted five steps of Fuzzy Time Series forecasting method. Sliding Window Method and Fuzzy Rule Based Systems was proposed to determine the number of class intervals and generate rules respectively. From the collected data, the universe, $U$ was defined. Then, the number of class interval obtained using the proposed Sliding Window Method was used to divide universe, $U$ into equal length of intervals and re-divide each intervals to acquire sub-intervals. Next, this study continue with identify rule of forecasting for each data using the proposed Fuzzy Rule Based Systems. The trend of forecasting can be defined through the classification of rule and the CPO price started to forecast.

![Figure 1. Fuzzy Time Series forecasting](image-url)
3.1.1. Sliding Window Method (SWM). Based on Figure 1, the propose Sliding Window Method is embedded in Fuzzy Time Series forecasting. Figure 2 below showing how the sliding window was constructed in determining the intervals in step 2 and 3 as in Figure 1. The predicted intervals obtained from the Sliding Window Method were used to divide the universe, $U$ in Fuzzy Time Series forecasting of CPO prices for validation purposes.

![Figure 2. Sliding Window Algorithm](image)

3.1.2. Fuzzy Rule Based Systems (FRBS). FRBS is the production or expert system, which is a continuation of the classical rule-based system. FRBS contains the IF-THEN fuzzy rules,

$$\text{IF a set of conditions are satisfied}$$

$$\text{THEN a set of consequents can be inferred}$$

The knowledge base composed of two parts, which are a data base and a rule base. Fuzzification is a structure converting crisp data into fuzzy sets. With the knowledge base, the inference system uses both the data base and the rule base to create inference using a reasoning method. Decoding the fuzzy rule action by forming defuzzification will result in the real action. The most widely used the FRBS’s type is Mamdani FRBS because it can model the dynamic structure of systems where the strength of this method has high power of approximation of nonlinear function [16].

Fuzzy IF-THEN Rules. A fuzzy IF-THEN rule is distinguished from the linguistic variable by its element of rule. The fuzzy IF-THEN rule is written “IF $x$ is $A$, THEN $y$ is $B$”, where both $A$ and $B$ each is a fuzzy set. The types of linguistic fuzzy model are the Mamdani and Takagi-Sugeno. The most regularly in use is the Mamdani, for various types of applications [17] and it represents the structure where:

$$\text{IF $X_{1i}$ is } A_1 \text{ and } \ldots \text{ and } X_n \text{ is } A_n ,$$

$$\text{THEN } Y_1 \text{ is } B_1 \text{ and } \ldots \text{ and } Y_m \text{ is } B_m .$$

where; $X_{1i}$: Fuzzy input variables, $Y_j$: Fuzzy output linguistic variables, and $A_i$ and $B_j$: Linguistic fuzzy sets that characterise $X_i$ and $Y_j$ [18]. The Mamdani expansion is given as:
IF $X_{i1}$ is $\tilde{A}_1$, ..., and $X_{n}$ is $\tilde{A}_n$

THEN $Y$ is B where $\tilde{A}_1 = A_{i1}$ OR ... OR $A_{ik}$ and $\tilde{A}_n = A_{n1}$

The Takagi-Sugeno model has linguistic antecedents but the consequence takes the form of linguistic variables function. It represents the form where:

$$\text{IF } X_{i1} \text{ is } A_1 \text{ and } \ldots \text{ and } X_{n} \text{ is } A_n, \text{ THEN } Y_1 = p_{11}^1 X_1 + \ldots + p_{1n}^1 X_n + p_{1n}^0 \text{ and } \ldots \text{ and } Y_m = p_{m1}^1 X_1 + \ldots + p_{mn}^1 X_n + p_{mn}^0,$$

where $p$ is real parameter, $X_i$ is fuzzy input linguistic variables, and $Y_i$ is the fuzzy output value. The Fuzzy Complete Rules are defined as a set of fuzzy rules involving $p$ conditional variables $x_i \square X_i$, $i = 1, 2, \ldots, p$ and one conclusion $y \square Y$, where:

Rule 1: IF $x_1$ is $F_1^1$ and ... and $x_p$ is $F_p^1$, THEN $y$ is $G^1$, where $l = 1, 2, \ldots, n$, and $n$ is the total number of rules in the set. If only a subset of $p$ inputs is used to describe the rules, then these rules are the Fuzzy Incomplete Rules. The fuzzy rules, complete and incomplete types, both are the subclasses of the Mamdani. In the fuzzy rules antecedents’ interpretation, the connector ‘OR’ is used for disjunction, ‘AND’ for conjunction and ‘NOT’ for complement. The rules with quantifier are grouped as Quantifier Rules, for example, ‘some’ and ‘all’.

4. Results and Discussion

This section discusses the analysis results. The predicted interval obtained from the proposed Sliding Window Method is equal to six. This method is repeated to acquire the sub-interval. A comparison was made among the forecast values and other forecasting methods’ values. Table 1 illustrated the comparison between Fuzzy Time Series forecasting of CPO prices for the year 2016 that has been determined by using the proposed method and the other forecasting method by [1], [8], [14].

| Actual Price (RM) | Forecast Value (RM) |
|------------------|---------------------|
|                  | Proposed method     | ARIMA (BJ) | ANN | Hybrid ARIMA-ANN |
| 421.40           | 477.25              | 497.25     | 470.75 | 478.52 |
| 475.80           | 508.42              | 521.38     | 513.75 | 511.38 |
| 506.72           | 561.1               | 581.1      | 567.13 | 560.4 |
| 560.89           | 545.5               | 555.5      | 535.75 | 544.21 |
| 542.42           | 521.13              | 531.13     | 505.25 | 522.47 |
| 518.39           | 511.38              | 521.38     | 514.75 | 508.52 |
| 509.16           | 545.5               | 525.5      | 555.25 | 548.72 |
| 548.57           | 571.5               | 541.5      | 562.5  | 571.22 |
| 571.88           | 529.9               | 519.9      | 535.75 | 530.67 |
| 530.11           | 529.9               | 509.9      | 540.63 | 532.89 |
| 530.96           | 571.5               | 591.5      | 580.5  | 572.49 |
| 565.50           | 567.54              | 543.78     | 550.45 | 567.48 |

Referring to Table 1, no forecast value for the first month because the previous data were not being used. The forecast value of Fuzzy Time Series with the proposed method looks like similar to hybrid ARIMA-ANN method compared to another two methods. Forecast value of CPO price for both methods
seen to be much closer to actual price of CPO. These mean that these two methods produced more accurate forecast values. As mention earlier, many researchers agreed that the hybrid model can perform better and give superior results, but it involves a complicated model. While the proposed method shows the results produced is more accurate and this method provides simpler algorithm that have been specifically chosen which best suit Fuzzy Time Series forecasting. The forecasting performance showed using the proposed method gives excellent results for time series forecasting. Next, Table 2 summarizes the Mean Square Error and Root Mean Square Error (MSE and RMSE) results by using different methods.

| Forecasting Method       | Proposed method | ARIMA | ANN | Hybrid ARIMA-ANN |
|--------------------------|-----------------|-------|-----|------------------|
| MSE                      | 0.0009          | 0.06  | 0.015 | 0.0013           |
| RMSE                     | 0.55            | 5.52  | 2.6  | 0.61             |

Table 2 shows that the results from MSE and RMSE are smaller compared to other method which are of the value 0.0009 and 0.55 respectively. Followed by Hybrid ARIMA-ANN, the results of MSE and RMSE are 0.0013 and 0.61 respectively. Meanwhile, the results of MSE and RMSE for ARIMA method are 0.06 and 5.52 respectively and ANN method is 0.015 and 2.6 respectively. As seen in Table 2, the proposed method and hybrid ARIMA-ANN methods can reduce forecast error. Therefore, this method provides an alternative for CPO price forecasting in producing good time series forecasting. The graph in Figure 3 shows the proof.

![Figure 3](image)

**Figure 3.** Comparison of the CPO actual price and the forecast price.

The actual price is plotted using the red line, while the blue line plots the Artificial Neural Network method, and the green line refers to the Autoregressive Moving Average (ARIMA) method. While the purple line refer to the proposed method and the dotted line refer to hybrid ARIMA-ANN method. Figure 3 above depicted that all the forecasting method provides significant CPO price forecast value. However, the CPO prices forecasting values of the proposed method and hybrid ARIMA-ANN method is closer to the actual price values compared to another two methods. This shows that the proposed method and hybrid ARIMA-ANN methods are able to give better prediction in performing the forecast of CPO prices accurately.
5. Conclusion
The CPO price forecasting uses the Fuzzy Time Series and Sliding Window Method and Fuzzy Rule Based Systems. The Sliding Window Method in time series CPO price forecasting was used to decide the number of the universe (U) class interval. While the proposed of Fuzzy Rules Based Systems was implemented to generate rules of forecasting. This study uses data starting from year 2012 until 2016 which was taken from MPOB. The results of the proposed method were comparing with other forecasting methods that have been done before. As discussed in the analysis, the forecast value for all forecasting method can be considered as good. However, the results show the value of MSE and RMSE for the proposed method and hybrid ARIMA-ANN method is less compare to other method. The results proved that the proposed method and hybrid ARIMA-ANN method give more effective and accurate forecasting value. Nevertheless, the complexity of hybrid ARIMA-ANN method need large amount of time. While the proposed method gives proper and systematic algorithm in the application of Fuzzy Time Series for CPO price forecasting. Therefore, this method should be able to reduce error and provide more accuracy in forecasting CPO price in time series forecasting.

Acknowledgement
My appreciation goes to Universiti Teknologi PETRONAS for providing financial support under Grant STIRF entitled “Forecasting The Price Of Crude Oil Based on Fuzzy Time Series Model With Combination Of Sliding Window Method (Cost Centre: 0153AA-F33).

References
[1] Aziz A K and Bujang I 2011 Progress accuracy of CPO price prediction: Evidence from ARMA family and Artificial Neural Network approach Int. J. Financ. Econ. no 64, pp 66–79
[2] Hansen S B, Padfield R, Syayuki K, Evers S, Zakariah Z and Mastura S 2015 Trends in global palm oil sustainability research J. Clean. Prod, vol 100, pp 140–149
[3] Abdullah R 2011 World palm oil supply, demand, price and prospects: Focus on Malaysian and Indonesian palm oil industries J. Indus. Oil Palm Ind. Econ. vol 11(2) pp 13–25
[4] Noormahayu M N, Khalid A R and Elsadig M A 2009 Financial assessment of oil palm cultivation on peatland in Selangor, Malaysia Mires Peat vol. 5(2), pp 1–18
[5] Rahman N M N A 2012 The cointegration analysis on the spot prices of the Malaysian Crude Palm Oil futures market Int. J. Econ. Financ. vol 4(7)
[6] Muhammad N, Rahman N A, Nawi A S, Hayati Y and Naziman N M 2012 The price discovery of the Malaysian Crude Palm Oil futures markets J. Appl. Financ. vol 2(4) pp 25–47
[7] Vijaya S, Ma A N, Choo Y M and Meriam N I K 2008 Life cycle inventory of the production of Crude Palm Oil J. Oil Palm Res. vol 20(6), pp 484–494
[8] Ong C S, Huang J J and Tzeng G H 2005 Model identification of ARIMA family using genetic algorithms Appl. Math. Comput. vol 164(3), pp 885–912
[9] Gilliland M 2017 Changing the paradigm for business forecasting Foresight Int. J. Appl. Forecast. no 44 pp 29–35
[10] Khashei M, Bijari M and Ardali G A R 2009 Improvement of Auto-Regressive Integrated Moving Average models using Fuzzy logic and Artificial Neural Networks (ANNs) Neurocomputing J. vol 72 pp 956–967
[11] Khandelwal I, Adhikari R and Verma G 2015 Time series forecasting using hybrid ARIMA and ANN models based on DWT Decomposition Proc. Comp. Science vol 48(6) pp 173–179
[12] Maia A L S, De Carvalho F D A T and Ludermir T B 2008 Forecasting models for interval-valued time series Neurocomputing J. vol 71(16–18) pp 3344–3352
[13] Maia A L S and De Carvalho F D A T 2011 Holt’s exponential smoothing and neural network models for forecasting interval-valued time series Int. J. Forecast. vol 27(3)
[14] Khashei M and Bijari M 2011 A novel hybridization of artificial neural networks and ARIMA models for time series forecasting Applied Soft Computing Journal vol 11(2) pp 2664–2675
[15] Valenzuela O, Rojas I, Rojas F, Pomares H, Herrera L J, Guillen A, Marquez L and Pasadas M 2008 Hybridization of intelligent techniques and ARIMA models for time series prediction Fuzzy Sets Syst. vol 159(7) pp 821–845
[16] Rikhtegar N, Mansouri N, Oroumieh A A, Yazdani-Chamzini A, Zavadskas E K and Kildiene S
2014 Environmental impact assessment based on group decision-making methods in mining projects Econ. Res. Istraz. vol 27(1) pp 378–392
[17] Othman M, Saian R, Nazri M N, Safiah N, Kamal A and Liyana N 2013 Fuzzy forecasting for water level of flood warning system in Perlis Int. Symp. Math. Sci. Comput. Res. pp 6–7
[18] Mamdani E H 1977 Application of Fuzzy Logic to approximate reasoning using linguistic synthesis IEEE Trans. Comput. Vol 26(12), pp 1182–1191