Fractionally integrated ARMA for CPO prices prediction: case of potentially over differenced

Abdul Aziz Karia, Imbarine Bujang and **Ismail Ahmad**
Universiti Teknologi MARA (UiTM) Shah Alam

Key Words
CPO, ARFIMA, ARIMA, RMSE, MSE, MAD, MAPE, R^2 and SI.

Abstract
Dealing with stationarity remains an unsolved problem. Some of the time series data, especially CPO prices persistence towards nonstationary in the long run data. This dilemma forces the researchers to conduct first order difference. The basic idea is that to obtain the stationary data that is considered as a good strategy to overcome the nonstationary counterparts. An opportune remark as it is, this proxy may lead into overdifference. The power of frequency elements has not been attenuated but nearly annihilated. With regard to the above matter, this paper presents the usefulness of autoregressive fractionally integrated moving average (ARFIMA) model as the solution towards the nonstationary persistency of crude palm oil (CPO) prices in the long run data. In this study, we employed daily historical free-on-board CPO prices in Malaysia. A comparison was made between the ARFIMA over the existing autoregressive integrated moving average (ARIMA) model. Statistical evaluation criteria was used to assess the performance of employed models, such as root mean square error (RMSE), mean square error (MSE), mean absolute deviation (MAD), mean absolute percentage error (MAPE), coefficient of determination (R^2) and scatter index (SI). The general conclusion that can be derived from this paper is that the usefulness of ARFIMA model outperformed the existing ARIMA model.

Introduction
One of the most significant issues among researchers and econometrician is the evidence of persistence towards the nonstationary. Relying to the assumption of the Box and Jenkins methodology, the time series assumed to be stationary. With regard, most of the time series data is nonstationary and it is considered the norm. Where such assumption is not met, then the necessary procedures such as differencing ($\Delta y_t = y_t - y_{t-1}$) is performed in order to achieved the stationary time series data (Lazim, 2011). The effort of differencing seems to be a good solution toward the nonstationary counterparts. An opportune remark as it is, this proxy may lead to overdifference (Erfani & Samimi, 2009). The difference stationary is eliminating too far of the trend like characteristics and the value of the level difference indicates that there is no influence that is close to zero frequency. The power of frequency elements has not been attenuated but nearly annihilated.

Overdifference could also form unintentional issues of differencing which could significantly contribute a problem in the parameter estimation. Hurvich and Ray (1995) revealed that the time series which potentially consisted of overdifference would be bias in long memory time series analysis. As a result, the time series prediction could be insubstantial and lose its effectiveness on parameters estimation. Worst come to worst, the forecasting performance one-step-ahead degrades (Erfani & Samimi, 2009). Another important finding found by Xiu and Jin (2007) is that the problem of overdifference is also accountable for the tendency of lost of important information of the time series and this problem also affecting the model construction. With regard to this matter, the existing method of the ARMA models which is extremely important in many years has gradually given way to the model, in this manner the ARFIMA model is emphasised to deal with the time series which persistence towards the nonstationary.

It is important to give special attention toward the ARFIMA package introduced by Doornik and Ooms (2004) which has the capability to adopt the Maximum Likelihood Estimation (MLE) to the long memory time series data. Prior studies have noted the main weakness of adopting the MLE towards the ARFIMA estimation procedure and the problem has essentially been solved by Hosking Hosking (1981) and Sowell (1987). However, Ooms and Doornik (1999) list the reason why some problems remained unsolved. There will be problems in variance matrix into account which is totally inappropriate for extensions with regression parameters. This is the explanation why MLE estimation is difficult to be adopted in ARFIMA model.

The purpose of the current study was to compare the predictability power of each forecasting techniques for ARFIMA and existing ARIMA model to predict the past historical value of CPO prices in Malaysia in which consist the long memory. Additionally, this study set out to establish the use of fractionally integrated to be an alternative in treating the time series which is persistence towards nonstationary in long memory high frequency time series data.

The paper is organized as follows: the first section is the Introduction which summarizes issues for the CPO prices and purpose of this study conducted. The second sections deal with the literature review which highlights the general reference to previous scholarly activity such as reference to current state of knowledge, research gap and etc. Third sections deal with the descriptive summery of the used data. Next sections deal with methodology applied include the ARFIMA and ARIMA to predict the CPO prices, followed by the fifth sections which cover establishing the ARFIMA and ARIMA models including the results and discussions. Finally, the last section offers the concluding explanations for this study.

Literature Review

Most of the time series especially in the high frequency data exhibits the long memory and it is vital to focus and further the study on it. The previous study reported the main limitation of existing ARMA model is incapable to portray long run time series data precisely (Arino & Marmol, 2004; Erfani & Samimi, 2009; Karia & Bujang, 2011; Kisi et al., 2012; Mostafaei & Sakhabakhsh, 2011; Xiu & Jin, 2007). The possible explanation for this due to the ARMA model is good in capturing short run prediction, while ARIMA model can be better options in nonstationary series with small sample sizes (Arteche, 2012). Difficulties arise, how if the large sample sizes of series indicates the nonstationary? Is the ARIMA prediction still reliable? With regard to this matter, the study might have been far more convincing if the we adopted the ARFIMA model. The long memory models such as ARFIMA provide better predictions especially when dealing with long run time series data. The rudimentary knowledge of ARFIMA model originate by Granger and Joyeux (1980) and the extension and profound effect of this study done by Granger (1980). However, far too little attention has been paid to ARFIMA until Baillie (1996) conducts on the whole assessment on the long memory and ARFIMA model in process and this extension attracting substantial attention in econometrics time series studies.

There are large number of literatures which deliberates the contradictions between ARFIMA and existing ARMA model in time series forecasting. Baillie and Chung (2002) found the ARFIMA model is superior and remarkable successful in predicting time series data compared to the existing ARMA model. Consistent with the previous findings, Reisen and Lopes (1999) found the ARFIMA model is efficient compared to existing ARMA model in term of mean of forecasted errors and its lower up to five steps ahead and yet, the variance of the errors and MSE were statistically equal. There are similarities between the characteristics expressed by ARFIMA in the study of Erfani and Samimi (2009) and those described by Baillie and Chung (2002), and Reisen and Lopes (1999). The application of ARFIMA has been found to be successful compared to existing ARMA model. However, it is also depends with the memory parameters. The long memory models such as ARFIMA do not adequately clarify the systems that operate with short memory parameters.

Ellis and Wilson (2004) is probably the best known critic of the ARFIMA model which argues the applicability of the ARFIMA model is the main question. The ARFIMA model has been found to produce poor out-of-sample result since it fail to outperform forecast and perform barely well to forecast based on the last pragmatic value or generally known as random walk model. At the same time, the application of ARFIMA produces high prediction variance. Therefore, they confirm that the ARFIMA model turn out to be poor out-of-sample performance. This assertion of the finding is supported by Xiu and Jin (2007). The

ARFIMA model found to be poor and ineffective in predicting the Hang Sheng index. However, the matters of characteristics of nonlinear systems might put the time series analysis into ineffectiveness. There is the possibility due to the tiny distinction of initial conditions may produced to completely different outcomes. This turn the projected result might not reliable.

Ellis and Wilson analysis has been criticised by a number of writers. Wang and Wu (2012), for example, point out that the ARFIMA model which takes long memory into account can outperform in term of out-of-sample projection. This findings is in agreement with Bhardwaj and Swanson (2006) findings which showed that the ARFIMA model does not produce poor out-of-sample performance. Basically, considering the mean square error (MSE) Diebold and Mariano (1995) perspective, the ARFIMA model sometimes progress much better out-of-sample forecast compared to the alternative forecasting techniques. The ARFIMA model is not going to descend into the "empty box". The ARFIMA characteristics for the most part regularly outperform simpler linear process at longer prediction scope. Additionally, the matters of d totally useful when constructing the prediction models by fine data sets, such as estimators based on shrinking the loss in predictive error.

The evidence provide by Ellis and Wilson point has been devastatingly critiqued by Chortareas et al. (2011). It is found that high frequency out-of-sample data forecasting by ARFIMA model recover compared to other alternative forecasting techniques. The ARFIMA model is said to be more accurate as the time series is in minute basis. The present findings seem to be consistent with research done by Koopman et al. (2005), which applied the ARFIMA model in S&P100 stock index, the result of out-of-sample produce by ARFIMA indicates the most accurate forecasts compared to its rivals. These findings further support the idea of Chu (2008), the ARFIMA model perform well for the time series in which comprise the economic and political shocks yet the model found to be a victor amongst the rival models during the tranquil period. Hence, the ARFIMA is said to be better than the ARIMA due to its estimation procedure in which treating the d as a noninteger as an effort to impart the stationary time series data. In piece of information, the ARFIMA model progress the prediction accuracy by more than a few percentage points depending in the lead which rival models compared with.

Data

3.1 Used Data

In this paper, the methodology is applied on daily CPO prices (free-on-board Ringgit Malaysia U$/Tonne). Data sample consist 2087 observations (from 1 January 2004 to 31 December 2011) of daily CPO prices records. Table 1 shows the descriptive summery statistics of the CPO prices in Malaysia. The mean of CPO prices is about RM2284.55 per tonne. The Jarque-Bera test result confirms that null hypothesis which the distribution is normal is rejected. Meanwhile, Figure 1 shows the CPO prices movement at free-on-board Ringgit Malaysia U$/Tonne.

Table 1: Descriptive Summery Statistics of CPO Prices

Statistics	CPO
Mean	RM 2284.55
Median	RM 2220.96
Maximum	RM 4300.67
Minimum	RM 1272.50
Standard Deviation	RM 753.24
Skewness	0.3702
Kurtosis	1.9155

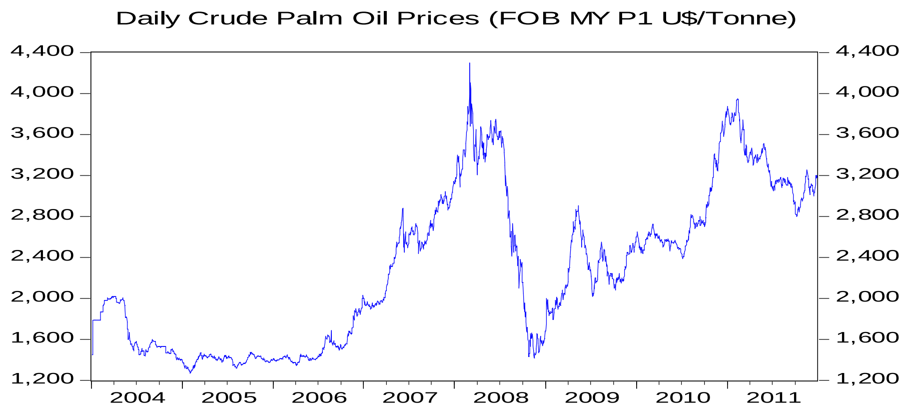


Figure 1: Observed CPO Prices Jan 1, 2004 - December 30, 2011

Methodology

In this section, we consider to apply previous empirical forecasting models such as ARFIMA and ARIMA models to forecast daily CPO prices record. Figure 2 illustrates the Box and Jenkins model-building strategy which was applied in this study.

Autoregressive Fractionally Integrated Moving Average (ARFIMA)

This sections presents the skeleton of the autoregressive fractionally integrated moving average model, ARFIMA (p,d,q). It is also known as fractionally differenced ARMA model which the time series model that generalized the integrated value of ARIMA model by permitting non integer values of integrated. It is very useful for time series analysis which consist the long memory. Special attention is given towards a package introduced by Doornik and Ooms (2004), the ARFIMA model that possible for MLE for long memory time series data. The ARFIMA model which consist the elements of d for the ranging between ($0.0 \leq d \leq 0.5$) is good to capture the time series data that persist towards the nonstationary and has been considered by numbers of literature in many field of time series study. Additionally, the climax of ARFIMA model building is fractional integrated. However, in empirical studies, there will be difficulties in constructing the fractional differencing.

There is an issue in treating nonstationary time series data into stationary, in this case, becoming the foremost delinquent in the area of time series study. This issue worsens when the difference stationary of series become overdifference. This type of time series data is persistence towards the nonstationary data. Therefore, the proposed model of fractionally integrated ARMA model has the capability to deal with time series that shows the strong persistence level (Mostafaei & Sakhabakhsh, 2011). Please refer Doornik and Ooms (2004) for complete explanation on MLE.

$$\Phi(L)(1-L)^d(y_t-u_t)=\Theta(L)\varepsilon_t, t=1,..,T$$

(1) Where the autoregressive parts determine by the equation as follows

$$\Phi(L)=(1-\emptyset_1 L-..-\emptyset_p L^P)$$

(2) In one hand, the moving average part displayed as follows

$$\Theta(\mathrm{L})=(1+w_1 L+\cdots w_1 L^q)$$

(3) The p and q are integers while the d is real. The major player in this model is $(1-L)^d$ which is the frictional difference operator and define as the binomial equation as follows.

$$(1-L)^d=\sum_{j=0}^{\infty}\delta_j L^j=\sum_{j=0}^{\infty}\binom{d}{j}(-L)^j$$

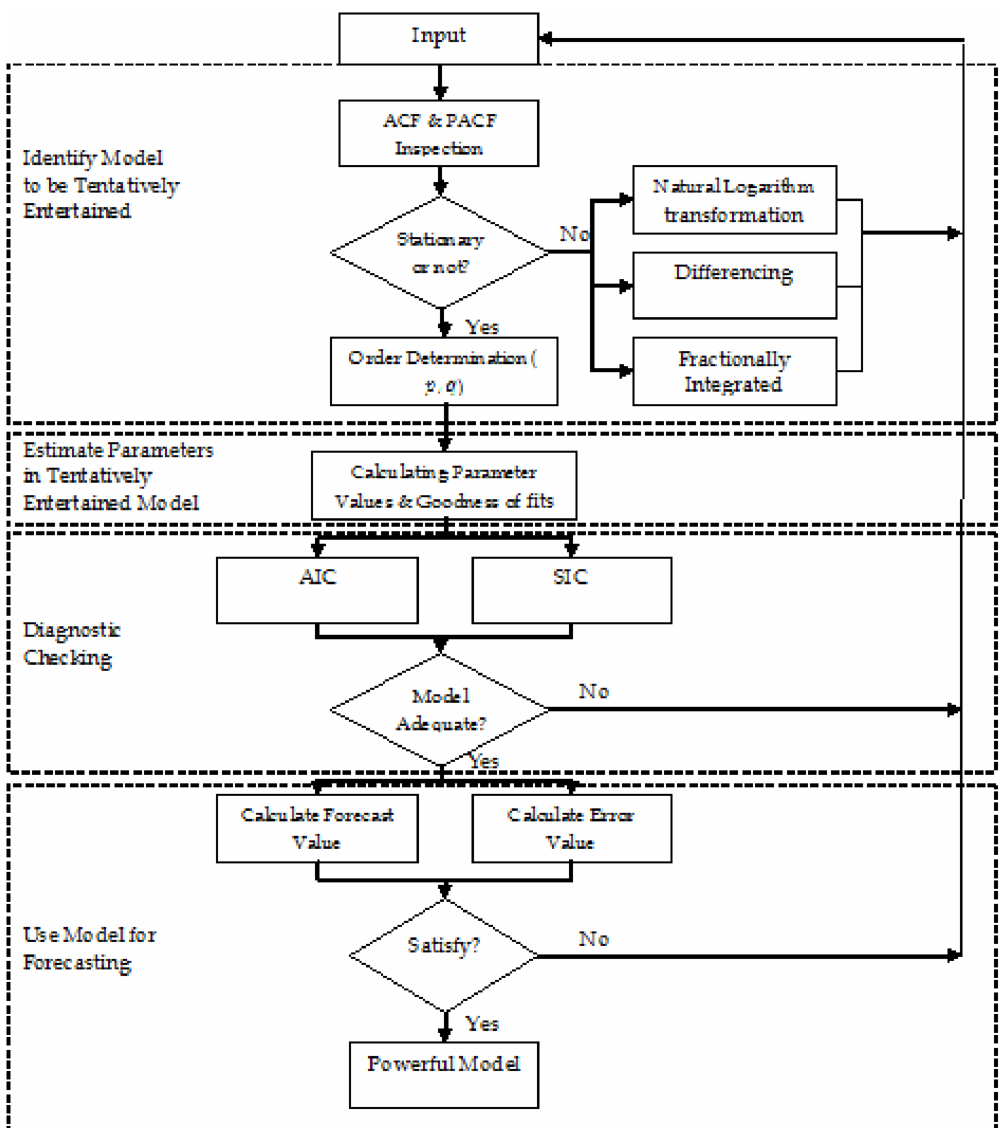
Autoregressive Integrated Moving Average (ARIMA)

If the stationary assumption of the time series data is not met, we adopt the ARIMA model. Therefore, it is necessary to conduct first order difference. The general term of the model as ARIMA(p, d, q), where p donates as the number of autoregressive term. Meanwhile q represent the number of moving average term. The main player in the ARIMA model is the value of d which represent the first order difference for the nonstationary solutions. Therefore, the general formulation of the ARIMA model can be express as equation (5) as follows.

$$w_t = \mu + \emptyset_1 w_{t-1} + \emptyset_2 w_{t-2} + \dots \emptyset_p w_{t-p} - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} \dots \theta_q \varepsilon_{t-q} + \varepsilon_t$$

(5) Where w_t represent the first order difference $(y_t - y_{t-1})$ of the time series data and assumed to be stationary. The constant terms donated as μ. Here the $\emptyset_p$ and w_{t-p} are the coefficient of autoregressive to be estimated and the response variable at time lags respectively. In addition, the terms donated as θ_q and ε_{t-q}, represent the coefficient of moving average to be estimated, and the error terms which is assumed to be independently distributed over time.

Figure 2: The Box and Jenkins Model-Building Strategy



Statistical Evaluation Criteria

In this study, we utilized six statistical evaluation criteria in order to measure the performance of the applied models. These criteria are the mean square error (MSE), mean absolute deviation (MAD), mean absolute percentage error (MAPE), root mean square error (RMSE), coefficient of determination(R^2), and scatter index (SI). The formulation of the statistical evaluation criteria express respectively as follows.

$$MSE = \frac{\sum_t^n e_t^2}{n}$$

$$MAD = \frac{\sum_t^n |e_t|}{n}$$

$$MAPE = \sum_t^n \frac{|(e_t/y_t)*100|}{n}$$

$$RMSE = \sqrt{MSE} = \sqrt{\frac{\sum_t^n e_t^2}{n}}$$

$$R^2 = \frac{[\sum_i (y_i-\bar{y})(\hat{y}_i-\bar{\hat{y}})]^2}{[\sum_i (y_i-\bar{y})^2][\sum_i (\hat{y}_i-\bar{\hat{y}})^2]}$$

$$SI = \frac{RMSE}{Mean\ of\ Observational\ Values}$$

Results and Discussion

The present study aims to compare the predictability power for the ARFIMA and ARIMA models to forecast daily historical CPO prices in Malaysia. Additionally, this study also set to establish the use of fractionally integrated to be an alternative in treating the time series which persistence towards the nonstationary. Before we proceed to forecast daily CPO prices, it is vital to inspect the ACF and PACF.

The ACF and PACF Inspections

Figure 3 presents the results obtained from the autocorrelation function (ACF) and partial autocorrelation function (PACF) from the preliminary analysis of the CPO prices. As can be seen from the Figure 3(a), we found that a covariance stationary of the CPO prices exhibits a statistically significant dependence between the observations. The illustration from the ACF inspection indicates that it decays hyperbolic rate or sluggish than the short memory time series data. In this case, the time series present the evidence of long memory (Arouri et al., 2012; Diebold & Inoue, 2001; Kwan et al., 2012; Perron & Qu, 2007; Tan et al., 2012; Xiu & Jin, 2007).

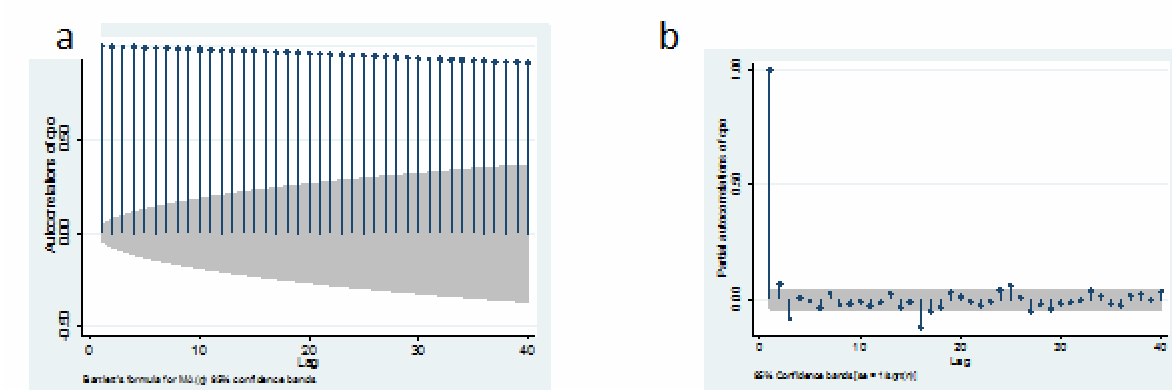


Figure 3: (a) Autocorrelation Function (ACF); (b) Partial Autocorrelation Function (PACF)

The Unit Root Tests

The ACF and PACF inspections on the CPO prices present a strong evidence of long memory. For that reason it is vital to conduct the unit root tests in this study. The empirical study in time series data assumes to be stationary. Therefore, time series data is believed to be nonstationary if containing the time varying mean, or time varying variance or both of the conditions. Likewise, in prediction commitment, the nonstationary time series data might be a slight practical worth. Worse comes into worse when nonstationary data risking the study into a spurious regression. Omitting the stationarity of the time series data absolutely misguided

.

Table 2: Result of Unit Root Tests

Test	Value of Statistic	1% Critical Value	5% Critical Value	10% Critical Value
ADF	-1.882636	-3.962451	-3.411965	-3.127886
PP	-1.965649	-3.962447	-3.411963	-3.127885
KPSS	0.236970	0.216000	0.146000	0.119000
DF GLS	-1.859572	-3.480000	-2.890000	-2.570000

There are numerous approaches to indicate the unit root for the time series data. Thus, there are no predetermined rules as to which approach to be adopted in a particular condition. Therefore, we test the unit root of the CPO prices based on Augmented Dickey and Fuller (1981) [ADF] test, Phillips and Perron (1988) [PP] test, Kwiatkowski et al. (1992) [KPSS] test, and Elliott et al. (1996) [DF GLS] test and show the result in Table 2. The ADF, PP, and DF GLS tests display the rejection of null hypothesis at 1%, 5%, and 10% critical values indicates that the time series is stationary. In one hand, for the case of KPSS test, the rejection of null hypothesis of stationary, suggesting diametrically opposed result from ADF, PP, and DF GLS tests. The results, as shown in Table 2, we fail to reject the null hypothesis of the ADF, PP, and DF GLS tests. Turning now to the KPSS test, the result reveals that we reject the null hypothesis of stationary time series data. With this, all unit root tests unambiguously indicate that there is a strong evidence of unit root in CPO prices. Since the stationary assumption is not met, then we consider the necessary procedures such as differencing and fractionally integrated to be conducted in this study.

Establishing the ARFIMA and ARIMA Models

Before applying the ARFIMA and ARIMA models, the series need to be stationary. Relying from the fact that, both of the ACF and PACF inspections and the unit root tests indicate that there is a strong evidence of nonstationary. Therefore, for the sake of comparison, the necessary procedures that we has to consider here is fractionally integrated and first order differencing. First, for the fractionally integrated parts, the parameter d must be determined. As mention previously, the parameter of d for the ranging between ($0.0 \leq d \leq 0.5$) is good to capture the series that persistence towards the nonstationary. Thus, we estimate the value of parameters of d through the package which is introduced by Doornik and Ooms (2004), we have d = 0.0016. To be specific, the models expose as $\Phi(L)(1-L)^{0.0016}(y_t - u_t) = \Theta(L)\varepsilon_t$ and depicted at Figure 4(a). From what we can observe from the depicted graph, there is not much loss of information as to conduct the fractionally integrated. Secondly, for the first order difference (d = 1), has been shown in Figure 4(b), from what we can observe, the first order difference is eliminating too far of the trend like characteristic. The power of frequency is not been attenuated but has nearly been annihilated. The difference stationary indicates that the power of frequency has no influence and fluctuates around to zero value. Besides, the depicted graphs (b) indicate that there is an evidence of overdifference. In this case, the difference stationary should be responsible for a loss of information in daily CPO prices record.

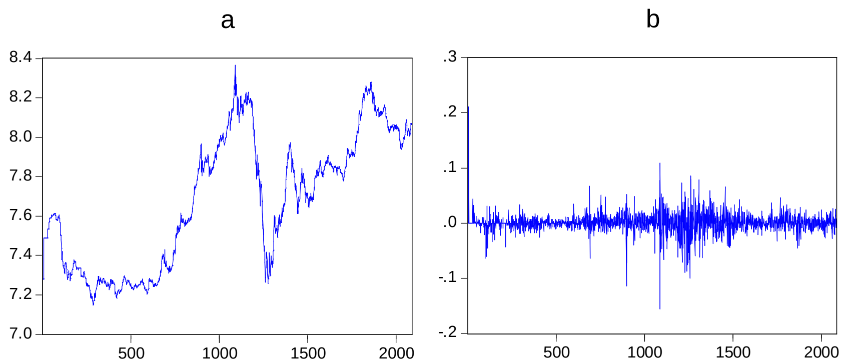


Figure 4: (a) Fractionally Integrated; (b) First order differenced

As we have achieve the stationary of the series as considering fractionally integrated and first order difference, we proceed with determining the order of p and q. Prior studies have noted that it is not easy to identify precisely an appropriate order of AR and MA based from the ACF and PACF spikes (Lazim, 2011). Therefore, we employed the "trial-and-error" method as one of the effort to reduce the risks of wrong model identification (Ahmad & Latif, 2011; Karia & Bujang, 2011).

Comparisons between the two models which ARFIMA and ARIMA were made using the forecasting errors measures such as the MSE, MAD, MAPE, and RMSE and the results illustrates in Table 3. In particular, the forecasting are then used to select the best ARFIMA model and the best ARIMA model. As we can observe, for ARFIMA models relatively, it is clearly comprehend the best fits models is the ARFIMA(1,0.0016,0) with the RMSE value of 0.017034. We found that the value of RMSE will increase as the order of AR and MA element increases. Looking into the ARIMA model, it is clearly stated that the ARIMA(1,1,0) with the value of RMSE equal to 0.025257 is slightly better than another ARIMA model. Overall, the Table 3 is quite revealing that the ARFIMA model is superior compared to the ARIMA model.

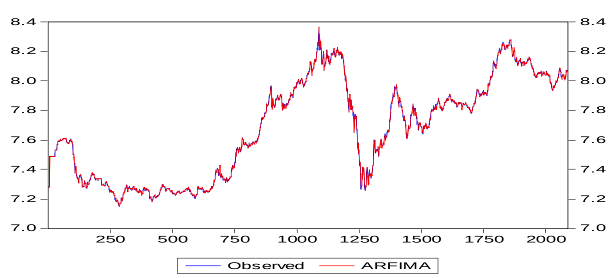


Figure 5: Observed and ARFIMA Prediction

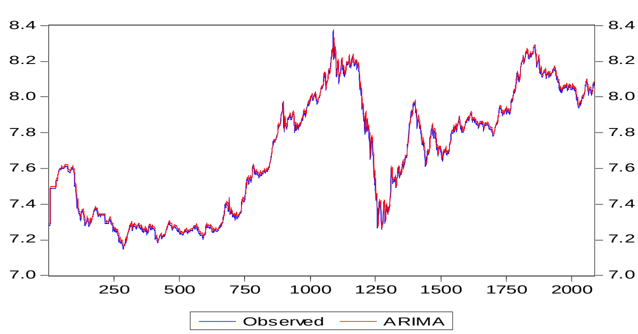


Figure 6: Observed and ARIMA Prediction

Interestingly, the results from the Table 3 presents evidence that the first order difference should be responsible for the loss of important information since its eliminating too far of the trend like characteristics which leads to the increase number of errors in CPO prices forecasting. Turn now to the depicted graph between the observed and forecasted value by ARFIMA(1,0.0016,0) and ARIMA(1,1,0) models respectively showed in Figure 5 and Figure 6. Comparing the two depicted graph, it is obviously can be seen that the most fits predictions is ARFIMA model. The forecasted values from the ARFIMA model are closed to the observed CPO prices than those of ARIMA model. The depicted graph is consistent to the results reported in Table 3.

Table 3: Forecasting Errors Performance

Model	MSE	MAD	MAPE	RMSE
ARFIMA(1,0.0016,0)	0.000290	0.010328	0.001337	0.017034
ARFIMA(1,0.0016,1)	0.000290	0.010329	0.001338	0.017036
ARFIMA(2,0.0016,0)	0.000290	0.010334	0.001338	0.017040
ARFIMA(2,0.0016,1)	0.000290	0.010335	0.001338	0.017042
ARFIMA(2,0.0016,2)	0.000295	0.010487	0.001358	0.017182
ARIMA(1,1,0)	0.000638	0.017653	0.002296	0.025257
ARIMA(1,1,1)	0.000729	0.018592	0.002417	0.027003
ARIMA(2,1,0)	1.132136	1.062787	0.138383	1.064019
ARIMA(2,1,1)	0.001789	0.028277	0.003670	0.042292
ARIMA(2,1,2)	0.001793	0.028590	0.003707	0.042349

Note: Using the Diebold and Mariano (1995) prospective stat ranging (-1.2 to +1.0) shows that all of the reported of RMSE are statistically significant apart from for ARIMA(2,1,0).

Further analysis on the prediction errors $(e = Y_t - f)$ generated by both of the models displayed in the Figure 7. It apparent from this Figure 7(b) exposed that the ARIMA model show an increasing level of noise compared to its rival models. Consistence with the previous statistical analysis, the prediction errors prove that ARFIMA is the best fits model to forecast the daily CPO prices.

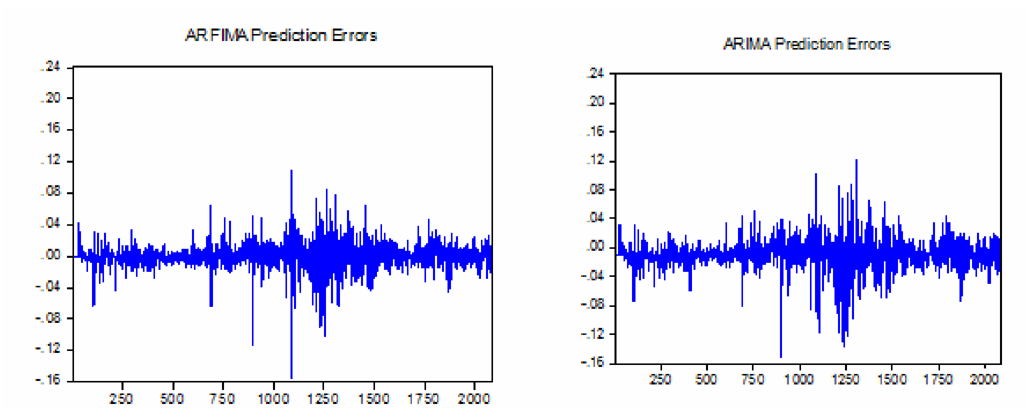


Figure 7: (a) ARFIMA Prediction Errors; (b) ARIMA Prediction Errors

In Figure 8, a comparison of scatter plots was used for predicted and the observed CPO prices values. Both of the scatter plots display the positive relationship between the forecasted and observed CPO prices. Looking at the both scatter plots, it is clearly appear that both of the models present close fits towards the regression line. In a particular, the ARFIMA model presents the closer fits towards the regression line than those of ARIMA model. A part of it, the ARIMA model present more scattered compared to its rival model. The coefficient of determination represent proof that the ARFIMA is fits than the ARIMA model, $R^2 = 0.997$ and $R^2 = 0.995$ respectively. With this, considering all of the statistical analysis and depicted graph, we can conclude that the ARFIMA model is highly better than the ARIMA model for the case of CPO prices.

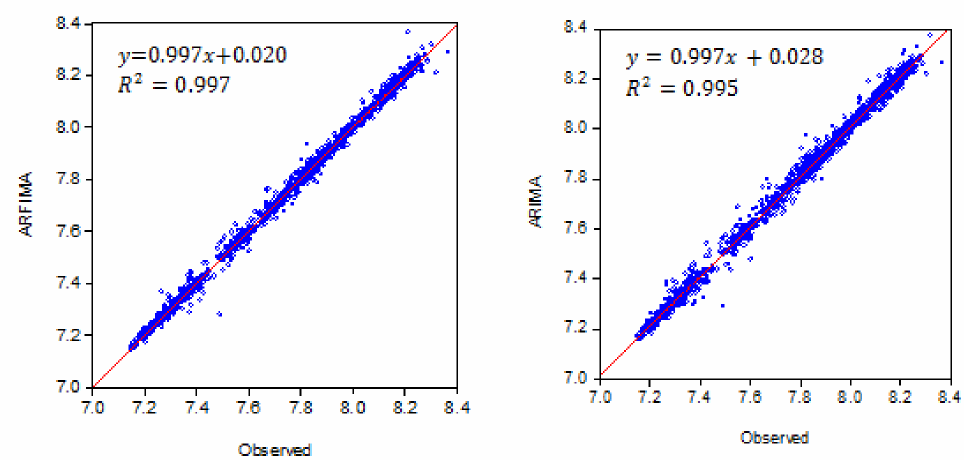


Figure 8: (a) Scatter Plots for Observed and; ARFIMA model (b) Scatter Plots for Observed and ARIMA model

Conclusion

The foregoing sections deal with the time series data towards the persistence of nonstationary. In this case, the daily CPO prices in Malaysia are consisting of strong long memory. Therefore, the necessary procedures proposed such as fractionally integrated and the first order difference toward achieving stationary of the series. We found that the forecasting performance of ARFIMA model produced high quality prediction and outperformed than that of ARIMA model by considering the statistical evaluation criteria, depicted graphs, and Diebold and Mariano (1995) prospective. With this, the ARFIMA model that possible for MLE introduced by Doornik and Ooms (2004) package is superior to forecast the CPO prices in Malaysia that consist the long memory. Additionally, we found that the ARIMA performance degrades as the time series is large sample sizes. The possible explanation for this is due to the necessary procedure to obtain stationary such as first order difference that usually indicates the overdifference for the time series. The first order difference should be responsible for the loss of important information since it's eliminating too far the trend like characteristics which leads to increase the level of noise. However, this study only represents proof for the in-sample forecasting analysis. Therefore, it is far more convincing with the extension towards out-of-sample forecasting analysis.

References

Ahmad, S., & Latif, H. A. (2011). Forecasting on the Crude Palm Oil and Kernel Palm Production: Seasonal ARIMA Approach. *IEEE Colloquium on Humanities, Science and Engineering Research*, 939-944.

Arino, M. A., & Marmol, F. (2004). A Permanent-Transitory Decomposition for ARFIMA Processes. *Journal of Statistical Planning and Inference 124*, 87-97.

Arouri, M. E. H., Hammoudeh, S., Lahiani, A., & Nguyen, D. K. (2012). Long Memory and Structural Breaks in Modeling the Return and Volatility Dynamics of Precious Metals. *The Quarterly Review of Economics and Finance 52*, 207-218.

Arteche, J. (2012). Parametric Vs. Semiparametric Long Memory: Comments on "Prediction from ARFIMA Models: Comparison between MLE and Semiparametric Estimation". *International Journal of Forecasting, 28*, 54-56.

Baillie, R. T. (1996). Long Memory Processes and Fractional Integration in Econometrics. *Journal of Econometrics, 73*, 5-59.

Baillie, R. T., & Chung, S. K. (2002). Modeling and Forecasting from Trend-Stationary Long Memory Models with Applications to Climatology. *International Journal of Forecasting 18*, 215-226.

Bhardwaj, G., & Swanson, N. R. (2006). An Empirical Investigation of the Usefulness of ARFIMA Models for Predicting Macroeconomic and Financial Time Series. *Journal of Econometrics, 131*, 539-578.

Chortareas, G., Jiang, Y., & Nankervis, J. C. (2011). Forecasting Exchange Rate Volatility Using High-Frequency Data: Is the Euro Different? *International Journal of Forecasting 27*, 1089-1107.

Chu, F. L. (2008). A Fractionally Integrated Autoregressive Moving Average Approach to Forecasting Tourism Demand. *Tourism Management 29*, 79-88.

Dickey, D. A., & Fuller, W. A. (1981). Likelihood Ratio Statistics for Autoregressive Time Series with A Unit Root. *Econometrica 49*(4), 1057-1072.

Diebold, F. X., & Inoue, A. (2001). Long Memory and Regime Switching. *Journal of Econometrics 105*, 131-159.

Diebold, F. X., & Mariano, R. S. (1995). Comparing Predictive Accuracy. *Journal of Business & Economiics Statistics 13 No 3*, 3-25.

Doornik, J. A., & Ooms, M. (2004). Inference and Forecasting for ARFIMA Models with An Application to US and UK Inflation. *Studies in Nonlinear Dynamics & Econometrics 8*(2), 1-23.

Elliott, G., Rothenberg, T. J., & Stock, J. H. (1996). Efficient Test for An Autoregressive Unit Root. *Econometrica, 64*(4), 813-836.

Ellis, C., & Wilson, P. (2004). Another Look at the Forecast Performance of ARFIMA Models. *International Review of Financial Analysis 13*, 63-81.

Erfani, A., & Samimi, A. J. (2009). Long Memory Forecasting of Stock Price Index Using a Fractionally Differenced ARMA Model. *Journal of Applied Sciences Research, 5*(10), 1721-1731.

Granger, C. W. J. (1980). Long Memory Relationships and the Aggregation of Dynamic Models. *Journal of Econometrics 14*, 227-238.

Granger, C. W. J., & Joyeux, R. (1980). An Introduction to Long-Memory Time Series Models and Fractional Differencing. *Journal of Time Series Analysis 1*(1), 15-29.

Hosking, J. R. M. (1981). Fractional Differencing *Biometrika, 68*, 165-176.

Hurvich, C. M., & Ray, B. K. (1995). Estimation of the Memory Parameter for Nonstationary or Noninvertible Fractionally Integrated Processes. *Journal of Time Series Analysis, 16*, 17-41.

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.

Kisi, O., Shiri, J., & Nikoofar, B. (2012). Forecasting Daily Lake Levels Using Artificial Intelligence Approaches. *Computers & Geosciences, 41*, 169-180.

Koopman, S. J., Jungbacker, B., & Hol, E. (2005). Forecasting Daily Variability of the S&P 100 Stock Index Using Historical, Realised and Implied Volatility Measurements. *Journal of Empirical Finance, 12*(3), 445-475.

Kwan, W., Li, W. K., & Li, G. (2012). On the Estimation and Diagnostic Checking of the ARFIMA-HYGARCH Model. *Computational Statistics and Data Analysis 56*, 3632-3644.

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., & Shin, Y. (1992). Testing the Null Hypothesis of Stationarity Against the Alternative of A Unit Root. *Journal of Econometrics, 54*, 159-178.

Lazim, M. A. (2011). *Introductionary Business Forecasting* (Third ed.). Shah Alam, Selangor Darul Ehsan, Malaysia: UiTM Press.

Mostafaei, H., & Sakhabakhsh, L. (2011). Modelling and Forecastng of OPEC Oil Prices with ARFIMA Model. *International Journal of Academic Research 3*(1), 817-822.

Ooms, M., & Doornik, J. (1999). Inference and Forecasting for Fractional Autoregressive Integrated Moving Average Models, with An Application to US and UK Inflation. *Econometric Institute Report 9947/A, Erasmus University Rotterdam*.

Perron, P., & Qu, Z. (2007). An Analytical Evaluation of the Log-Periodogram Estimate in the Presence of Level Shifts. *Boston University*.

Phillips, P. C. B., & Perron, P. (1988). Testing for A Unit Root in Time Series Regression. *Biometrika 75*(2), 335-346.

Reisen, V. A., & Lopes, S. (1999). Simulations and Applications of Forecasting Long-Memory Time Series Models. *Journal of Statistical Planning and Inference 80*, 269-287.

Sowell, F. (1987). Maximum Likelihood Estimation of Fractionally Integrated Time Sereis Models. *Discussion Paper 87-07, Department of Economics, Duke University*, 1-31.

Tan, P. P., Galagedera, D. U. A., & Maharaj, E. A. (2012). A Wavelet Based Investigation of Long Memory in Stock Returns. *Physica A 391*, 2330-2341.

Wang, Y., & Wu, C. (2012). What Can We Learn from the History of Gasoline Crack Spreads?: Long Memory, Structural Breaks and Modeling Implications. *Economic Modelling, 29*, 349-360.

Xiu, J., & Jin, Y. (2007). Empirical Study of ARFIMA Model Based on Fractional Differencing. *Physica A, 377*, 138-154