INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN ACCOUNTING, FINANCE AND MANAGEMENT SCIENCES

www.hrmars.com
ISSN: 2225-8329

Comparative Performance of Arima and Garch Models in Modelling and Forecasting Volatility of Kuala Lumpur Composite Index

Nasuhar Ab. Aziz, Siti Nurhafizah Mohd Shafie, Mohd Noor Azam Nafi

To Link this Article: http://dx.doi.org/10.6007/IJARAFMS/v13-i1/16213 DOI:10.6007/IJARAFMS /v13-i1/16213

Received: *13* December 2022*,* ***Revised:*** *11* January 2023*,* ***Accepted:*** *26* January 2023

Published Online: *20* February 2023

In-Text Citation: (Aziz et al., 2023)

To Cite this Article: Aziz, N. A., Shafie, S. N. M., & Nafi, M. N. A. (2023). Comparative Performance of Arima and Garch Models in Modelling and Forecasting Volatility of Kuala Lumpur Composite Index. *International Journal of Academic Research in Accounting Finance and Management Sciences*, *13*(1), 330–343.

Copyright: © 2023 The Author(s)

Published by Human Resource Management Academic Research Society (www.hrmars.com)

This article is published under the Creative Commons Attribution (CC BY 4.0) license. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this license may be seen at: http://creativecommons.org/licences/by/4.0/legalcode

Vol. 13, No. 1, 2023, Pg. 330 - 343

http://hrmars.com/index.php/pages/detail/IJARAFMS **JOURNAL HOMEPAGE**

Full Terms & Conditions of access and use can be found at

http://hrmars.com/index.php/pages/detail/publication-ethics

INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN ACCOUNTING, FINANCE AND MANAGEMENT SCIENCES

HRMARS MANAGEMENT ACADEMIC RESEARCH SOCIETY
EXPLORING INTELLECTUAL CAPITAL
www.hrmars.com
ISSN: 2225-8329

Comparative Performance of Arima and Garch Models in Modelling and Forecasting Volatility of Kuala Lumpur Composite Index

Nasuhar Ab. Aziz, Siti Nurhafizah Mohd Shafie, Mohd Noor Azam Nafi

Mathematical Sciences Studies, College of Computing, Informatics and Media, Universiti Teknologi MARA (UiTM) Kelantan Branch, Kota Bharu Campus, 15150 Kota Bharu, Kelantan, Malaysia

Corresponding Author's Email: nasuha978@uitm.edu.my

Abstract

Time series is a set of observations in sequence over time. Time series modelling is used to create an applicable model that defines the necessary arrangement of the series by study the previous information of a time series. The past information of a time series is used to generate forecast value for the series. It is well acknowledged that a time series are regularly affected with outliers. Outliers may impact the forecasting where the tendency in parameter estimates created by extreme observation will reduce its effectiveness because the optimum predictor for an Autoregressive Integrated Moving Average (ARIMA) model is determined by its parameters. Thus, the occurrence of extreme observations might have a huge effect on predictions value. Therefore, Generalized Autoregressive Conditional Heteroskedastic (GARCH) model has been used to compare the result obtain from ARIMA model. This study used ARIMA and GARCH to compare the best model for forecasting Kuala Lumpur Composite Index (KLCI) when the outlier exists. The best models of ARIMA and GARCH were evaluated using Mean Square Error (MSE), Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE). It can be concluded that GARCH model performed better compared to Box-Jenkins ARIMA in forecasting KLCI.

Keywords: ARIMA, Forecast, GARCH, Time Series, Outlier

Introduction

Time series is a set of observations in sequence over time. The sequence may be represented by the values of $y_1, y_2, ..., y_t$ where t refers to the period of time (Anderson, 1977). Time series modelling is used to create an applicable model that defines the necessary arrangement of the series by study the previous information of a time series. The past information of a time series is used to generate forecast value for the series (Raicharoen et al., 2003).

Box-Jenkins is one of the methods in time series that appropriate for analyzing the time series data that have long-series data (Box & Pierce, 1970). It is used to convert non-stationary series

into stationary. Stationary is when the probability distribution is same for all initial values of t. Otherwise, it is not stationary when a series shows a simple trend because the values of the series depend on t. The time series can be approached by an Autoregressive Moving Average (ARMA) and Autoregressive Integrated Moving Average (ARIMA) model and has to fulfill assumption of the variance is consistent for all values of t. However, the time series data particularly in financial data shows that the variance of returns is not constant over time or the volatility is clustering (Chong et al., 1999).
Engle (1982) developed Autoregressive Conditional Heteroskedastic (ARCH) model to reflect the characteristics of volatility where this model considers the time-varying conditional variance of financial time-series using lagged disturbances. However, this model requires a large number of parameters to describe the conditional variance. Later, the work of Engle (1982) was extended by Bollerslev (1986) to reduce number of parameters in the model. Since then, most of the researchers have used these models to financial time series data (Omar & Halim, 2015; Husna et al., 2016; San et al., 2011; Hussin et al., 2021; Merabet et al., 2021).
It is well acknowledged that the time series are regularly affected with outliers. However, when the outliers exist in the time series data, the researchers always relied on assumptions of the residuals independently integrated and identically distributed (IID) to deal with outliers (Fox, 1972). Thus, the tendency in parameter estimates created by outliers will reduce its effectiveness and will have a huge effect on predictions value (Bianco et al., 2001). Later, to rectify this problem, Fox (1972) proposed two types of outliers in an autoregressive (AR) model for time series data (i.e., innovational outlier (IO) and additive outlier (AO)). Afterward, many studies on detecting and modeling outlier have been extended to the other classes of models including ARIMA, ARCH, GARCH and other models (Ljung, 1993; Chan & Cheung, 1994; Zainol, 2010; Gourieroux, 1997; Dijk et al., 1999; Shi & Chen, 2008).

Currently, GARCH time series techniques have been used to forecast Kuala Lumpur Composite Index (KLCI) data. However, to date, based on our knowledge, none of the research work have consider an outlier exist in the estimate the volatility KLCI stock market data. Therefore, this study aims to clarify the types of outlier presence in Kuala Lumpur Composite Index (KLCI) which are innovational outlier (IO) and additive outlier (AO) that proposed by previous researchers. This study uses daily volatility of stock prices from period 1 January 2011 to 31 December 2018. The study also intends to compare the best model between ARIMA and GARCH model to be incorporated into KLCI daily volatility returns to forecast a future value for KLCI 2019.

Data Sources and Method

The data of Kuala Lumpur Composite Index (KLCI) are used in this study. The data can be categorized as quantitative data and were obtained from Financial Times Stock Exchange (FTSE) Malaysia KLCI. This indicator is extracted from 100 companies that Bursa Malaysia has selected from a cross sector of the total companies in Malaysia. However, the process before it is selected to be one of the hundred is much more complicated and a company must meet many requirements. This index is taken as an indicator of stock market's performance and thus it provides with a standard that reproduces the improvement of Malaysia's economy. The data consisted of 1978 price index from January 2011 to December 2018 and were used in the process of identification, estimation and forecasting.

Stationary of the Series

The series should be check either it stationary and the seasonality is analysed before further study. Seasonality can be caused by different factors such as weather, holidays and consists of intermittent, repeated, frequent and predictable patterns in a time series. On a weekly, quarterly or monthly basis, seasonality can be repeated. Seasonal pattern can be identified by checking the data series time plot. Furthermore, stationary of the data is checked by plotting the raw data or using statistical test which is Augmented Dickey-Fuller (ADF) test. If the ADF statistic is more negative than the table value, reject the null hypothesis of a unit root. The more negative the DF test statistic, the stronger the evidence for rejecting the null hypothesis of a unit root. The non-significant results of the tests indicated that the series not stationary. The non-stationary series is differenced to make it stationary. The ADF test statistic as (1).

$$DF_{\tau} = \frac{\hat{\gamma}}{SE(\hat{\gamma})} \tag{1}$$

The unit root test is then carried out under the null $\gamma = 0$ against the alternative hypothesis of $\gamma < 0$.

Outliers

Identification of outliers plays an important role in statistical analysis. Most of the data collected for analysis and interpretation contains one or two observations which not identical to the rest of the data. Outliers could have arisen naturally or due to human error in data collection or theoretical error in model selection. These outliers must properly treat to avoid misleading conclusions. Hence, such outliers should be identified and treated properly to draw proper conclusions from the data (Deneshkumar & Kannan 2011). An outlier can be analysed when a data point is lying far away from the other point of the data (Aguinis et al. 2013). In the time series model, Fox (1972) first identifies outliers and classifies outliers in two categories which is Additive Outliers (AO) and Innovative Outliers (IO).

i. Additive Outier (AO)

Additive outlier impact the phase at a single point at which it takes place and tends to be a relatively large or small value for a single observation. The presence of outliers in these findings could have a significant impact on forecasts. There are two ways an additive outlier influences the forecasts, firstly through carry-over effect and by attempting to influence the forecasts, which means that incorrect parameter values are used in the forecast calculation. The error of estimation depends on the type of outliers, an AO will significantly increase the predictor error and AO's presence will seriously influence coefficient estimates and variance.

ii. Innovational Outier (IO)

Innovational outlier is the form of outliers that affects subsequent observations starting from their location such as arising from normal randomness. The presence of AO can severely influence the estimates of the ARMA coefficients and variance, while IO has a much smaller effect in general (Chang & Tiao, 1983). An IO affects only residual at the outlier date and affects the next residual, inflating two residuals in a row. This effect has several implications for any further residual analysis. According to some experts, if an AO type outlier occurs in an observation set, its effects should be excluded. However, if an IO type outlier exists, it should be recognized as a result of natural randomness and its effects should not be eliminated.

Autoregressive Integrated Moving Average (ARIMA)

Autoregressive Integrated Moving Average (ARIMA) or Box-Jenkins models are mixed of Autoregressive (AR) model and Moving Average (MA) model where the series deal with difference since the data are non-stationary. The general terms of ARIMA can be presented as ARIMA where indicates the AR order, is the MA order and d is the number of data that needs to be differenced to make the series stationary. The order of is the lagged of dependent or current value and it can be identified using partial autocorrelation (PACF). The number of spikes in PACF will determine the order of . While the order of is referring to the number of lagged time period in the model where it is determined by the number of spikes in autocorrelation (ACF). ARIMA is written as (2).

$$(1-\phi_1\beta-\phi_2\beta^2-\ldots-\phi_p\beta^p)(1-\beta)y_t=(1-\theta_1\beta-\theta_2\beta^2-\ldots-\theta_q\beta^q)e_t \quad (2)$$

Generalized Auto-Regressive Conditional Heteroscedasticity (GARCH)

Modeling and forecasting of asset return volatility is an important area of research in academics as well as in finance companies. This necessitates the need for selecting a suitable model from a class of GARCH models. GARCH model is well-known as a model of heteroscedasticity which is not constant in variance. This model has been used widely in financial and business areas, since the data of these areas tend to have variability or highly volatile throughout the time. GARCH model is written as GARCH (p,q) model where is the number of (MA) terms and is the number of (AR) terms. GARCH (p,q) model can be represented as (3).

$$\sigma_t^2=\alpha_0+\alpha_i\varepsilon_{t-i}^2+\beta_i\sigma_{t-i}^2 \quad (3)$$

Volatility of a data needs to be checked before using GARCH model. One of the methods is by computing histogram for a stationary series and checking the distribution of data. Kurtosis is the measure of peakness of the data distribution and skewness is the measure of symmetrical of the distribution about the mean. When the value of kurtosis is greater than three and it is skewed either to the left or right, then the series is volatile.

Model Evaluation

This study used the data of the daily KLCI dataset where it started from 1st January 2011 to 31st December 2018. It consists of 1978 number of observations. The data were divided into two parts which are 70% for estimation and 30% for evaluation (Cerqueira et al. 2019). In estimation part, this research included the observations from 3rd January 2011 to 26th July 2016 which contained 1385 number of observations. For evaluation part, this research included the observations from 27th July 2016 to 31st December 2018 which contained 593 number of observations. There are five typical measures used to evaluate the ARIMA and GARCH models which are the Akaike's Information Criteria (AIC), Bayesian Information Criterion (BIC), Mean Squared Error (MSE), Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE).

i. Akaike's Information Criteria

The common measure of the suitability for the model is Akaike's Information Criteria (AIC). These criteria deal with the penalty on the likelihood for each extra term is incorporated in the model. If the additional term does not develop the likelihood more than the penalty amount, therefore, it is not worth computing into the model. The AIC equation is given as (4)

$$AIC = e^{\frac{2k}{T}} \frac{\sum_{t=1}^{T} e_t^2}{T} \quad (4)$$

where $k = p + q + P + Q$ represents the number estimated parameters in the model, p and q are the usual respective terms of the AR and MA parts, and T are the seasonality parts of the ARIMA model and is the total number of observations in the data of time series. In fact, $e^{\frac{2k}{T}}$ constitutes a penalty function whose intention is to avoid model is over fitting. The purpose of the test is to decide on the values of p, q, P, Q such that the value of AIC is minimized. In other words, a model is considered as having a better fit among all other competing models if the value of its AIC is the smallest. However, the principal of parsimony is still being held when choosing the best model where the model has the least number of parameters.

ii. Bayesian Information Criterion

The Bayesian Information Criterion (BIC), also generally known as the Schwarz Criterion (SBC) was established by Schwarz in 1978. This criterion aims the method of selecting models that accomplish the most precise out-of-sample forecasts by balancing between the models' complexity and goodness-of-fit. On the other hand, to define the best model for SBC, it will be choosing the one that has the lowest SBC value. It needs to be restated that the BIC is usually used as a model selection criterion when no firm theoretical or empirical reasons are offered to decide one model over the other. The BIC equation is given as (5).

$$BIC = T^{\frac{k}{T}} \frac{\sum_{t=1}^{T} e_t^2}{T} \quad (5)$$

iii. Mean Squared Error (MSE)

Most specialists used this error measure or standard criterion for evaluating the model's fitness to a specific series of data that are available by most statistical software. This method is usually used for comparing the model's forecasting achievement. Furthermore, if this method is being used outside the sample, it usually matches the within sample criterion. The MSE is given as (6).

$$MSE = \frac{\sum_{t}^{n} e_t^2}{n} \quad (6)$$

for which $e_t = y_t - \hat{y}_t$, where y_t indicates actual observed value at time t and $\hat{y}_t$ is the value of fitted at time t.

iv. Root Mean Squared Error (RMSE)

Root mean square error (RMSE) is the standard deviation of the prediction errors. Prediction errors are a calculation of how far from the regression line data points are. RMSE is a measure of how spread out these residuals is. In other words, it also determines that the concentrated data is around the line of best fit. The smaller the value of RMSE, the better is the model. The RMSE is given as (7).

$$RMSE = \sqrt{\frac{\left\{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2\right\}}{n}} \tag{7}$$

where $\sum$ is summation, $(y_t - \hat{y}_t)^2$ is differences squared and n is sample size.

v. Mean Absolute Percentage Error (MAPE)

The mean absolute percentage error (MAPE), also known as the mean absolute percentage deviation (MAPD), is a measure of prediction accuracy of a forecasting method in statistics. As an example in trend estimation, it is also used as a loss function for regression problems in machine learning. It usually expresses accuracy as a percentage and it is defined as (8).

$$MAPE = \left(\frac{\left[\Sigma_{t-1}^{n}\left|\frac{y_t - \hat{y}_t}{y_t}\right|\right]}{n}\right) x100\% \tag{8}$$

Where y_t is the actual value and $\hat{y}_t$ is the value of forecast. The absolute value in this calculation is summed for every forecasted point in time and divided by the number of fitted points n. Multiplying by 100% makes it a percentage error. The smaller the value of MAPE, the better is the model.

Results and Discussion

This research used the data of the daily Kuala Lumpur Composite Index (KLCI) dataset collected from database of Thomson Reuters Data stream. It consists of 1978 number of observations, where it started from January 2011 to December 2018. Figure 1 shows the plot of KLCI. However, the data of price shows that, it has droped sharply in 2011. Besides that, the figure shows that the volatility changes over time.

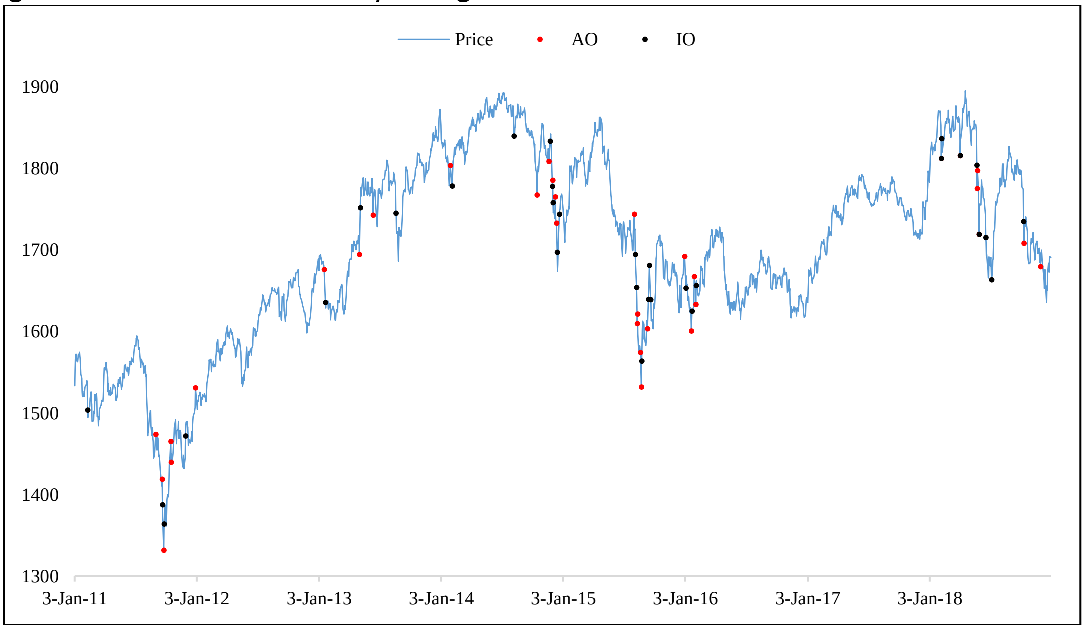

Figure 1. Kuala Lumpur Composite Index from January 2011 to December 2018 with outliers plot.

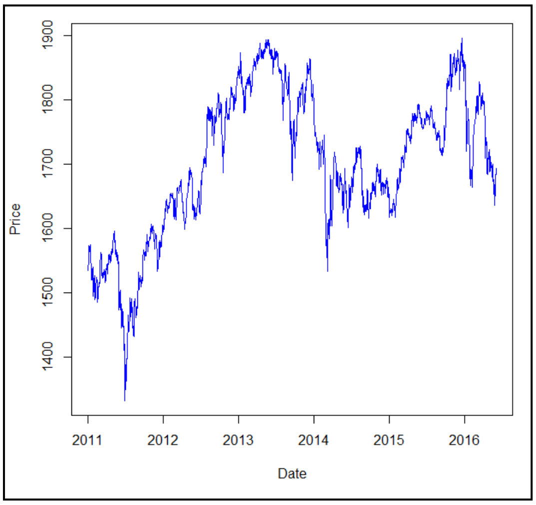

Figure 2. Kuala Lumpur Composite Index

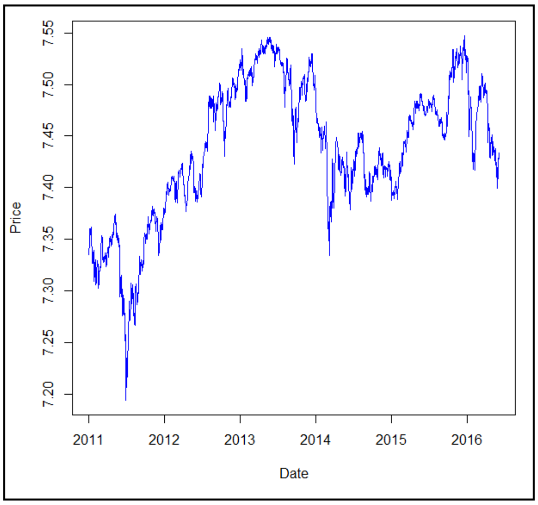

Figure 3. Log Kuala Lumpur Composite Index.

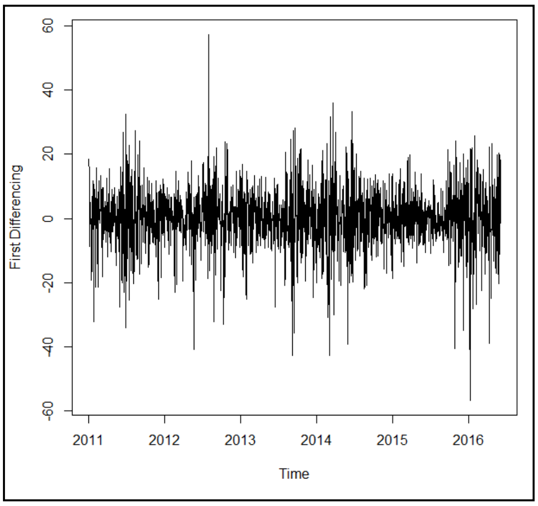

Figure 4. First differencing Kuala Lumpur Composite Index.

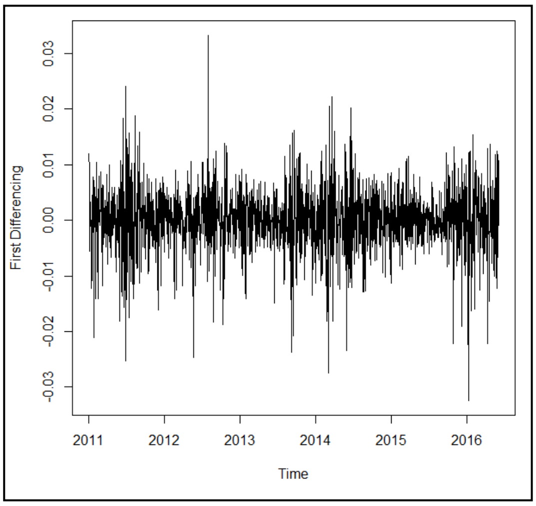

Figure 5. First differencing Log Kuala Lumpur Composite Index.

Figure 1 shows the plot of the data KLCI and it seems to have an outlier in the series. Therefore, the outliers have been identified and found that there are 32 of AO and 31 of IO. It showed that the data were non-stationary since there was an increasing trend, cyclical component and had irregularities which indicated that the data were not yet stationary. Figure 2 shows the KLCI plot while Figure 3 shows the log of KLCI and both are non-stationary. Therefore, first differencing is used to make the series stationary as shows in Figure 4 and Figure 5. Another way to check the stationary is by using ADF test. The result of ADF test is shown in Table 1.

Table 1
Test statistics for ADF test.

	Original Series	**First Difference Series**
Statistics	-2.0918	-12.046
p-value	0.5395	<0.01

The result for the ADF test in Table 1 showed that the value of original series for statistics was -2.0918 and the p-value was 0.5395 which were highly not significant. Because the p-value was higher than the significance level 0.05, the null hypothesis was accepted which concluded that the series were non-stationary. Therefore, the first differencing of the series was conducted and the stationary was tested again. It was found that the series became stationary where the ADF test showed the p-value was less than 0.01.

As for GARCH models, the histogram at first difference level need to plot to check either it can be used or not. GARCH model only can be used when the data is volatile. Figure 6 shows the histogram and Table 2 shows the descriptive statistics at first differencing.

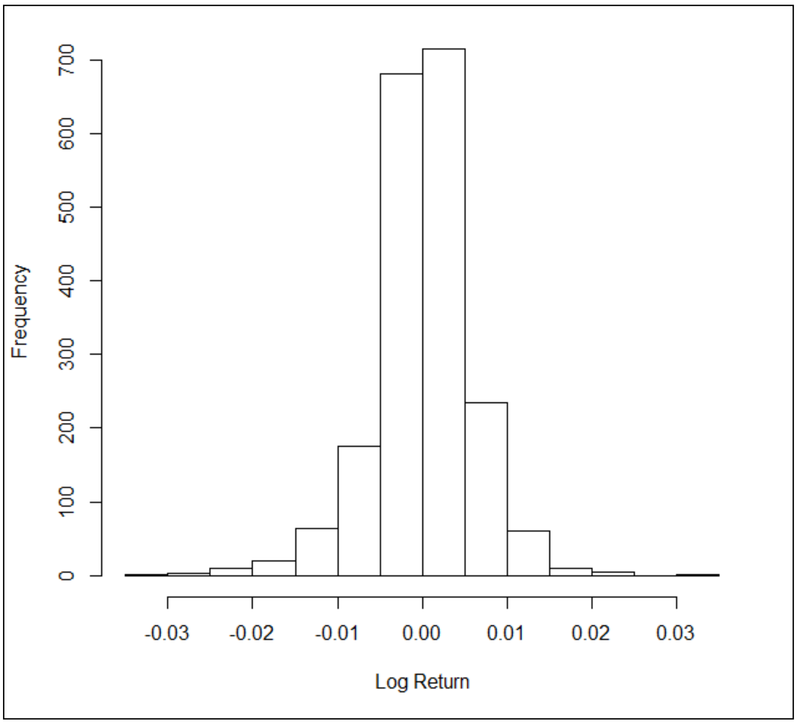

Figure 1. Histogram for Kuala Lumpur Composite Index at first difference level.

Table 1
Descriptive statistics for KLCI.

Skewness	Kurtosis
-0.395275	5.96736

The histogram in Figure 6 shows skew to the left and the values of kurtosis is more than three which indicates the GARCH model can be used. The data used for GARCH model is transformed using log transformation and taken from first differencing of the series.

Model Identification and Diagnostic Checking

To identify the ARIMA model, autocorrelation (ACF) and partial correlation (PACF) graph of the series has been used. Figure 7 and Figure 8 show the plot of ACF and PACF after first differencing.

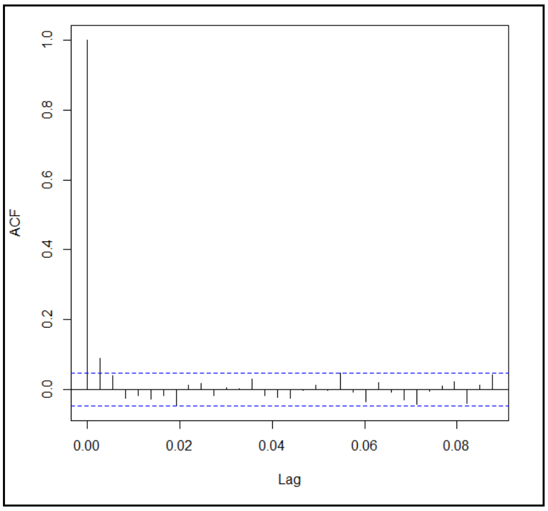

Figure 7. First differencing of ACF.

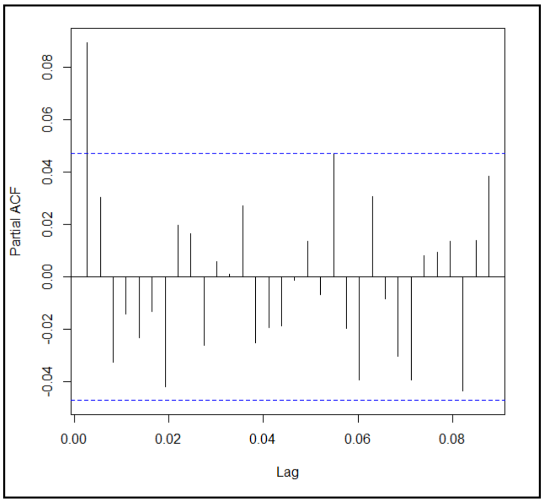

Figure 8. First differencing of PACF.

The ACF plot will determine the order for MA(q) while PACF plot will determine the order of AR(p). As shows in Figure 7 and Figure 8, there are two spikes for both ACF and PACF plot after differencing. Therefore, the model for ARIMA model is denoted by ARIMA (p,d,q) and the model are ARIMA(1,1,0), ARIMA(1,1,1), ARIMA(2,1,0) and ARIMA(2,1,1). While for the GARCH model are GARCH(1,1), GARCH(1,2) and GARCH(1,3).

ARIMA and GARCH performance

Table 3 shows the result of estimation and evaluation for ARIMA model while Table 4 shows for the GARCH model.

Table 2
ARIMA Model

	ARIMA (1,1,0)	ARIMA (1,1,1)	ARIMA (2,1,0)	ARIMA (2,1,1)
(i) Estimation				
AIC	7.392646	7.393525	**7.391230**	7.392459
BIC	**7.396427**	7.401087	7.398796	7.403808
(ii) Evaluation				
MSE	**85.26118**	85.44725	85.72053	85.44825
RMSE	**9.233698**	9.243768	9.258538	9.243822
MAPE	**117.0019**	117.7493	118.8455	117.4827

Table 3
GARCH model.

	GARCH (1,1)	GARCH (1,2)	GARCH (1,3)
(i) Estimation			
AIC	-7.563834	**-7.563702**	-7.573070
BIC	-7.548711	**-7.544798**	-7.550384
(ii) Evaluation			
MSE	0.00002795	0.00002796	**0.00027888**
RMSE	0.005287	0.005288	**0.005281**
MAPE	118.3697	118.1990	**118.1037**

Based on the result given in Table 3 and Table 4, the best model for modelling is considered by the lowest AIC. Therefore, ARIMA(2,1,0) and GARCH(1,2) are the best model for modelling. Even though the result or ARIMA(1,1,0) has the lowest BIC, the values of AIC is more preferable to use. To find the best model for forecasting, the evaluation of the sample has been used. It can be conclude that the best model are ARIMA(1,1,0) and GARCH(1,3) with the lowest MSE, RMSE and MAPE.

Forecasting using best Performance

The best model between ARIMA and GARCH will be choose in order to forecast KLCI in year 2019. Table 5 shows the result of best model for both ARIMA and GARCH.

Table 4
Best model between ARIMA and GARCH.

	ARIMA (1,1,0)	GARCH (1,3)
MSE	0.000027899	**0.000027888**
RMSE	0.005282	**0.005281**
MAPE	**116.5162**	118.1037

For the value of error measures of ARIMA(1,1,0) in Table 5, it was different compared to the value in Table 3. It is because the observations were transformed using log transformation when compared with GARCH. The values for MSE, RMSE and MAPE in ARIMA(1,1,0) were 0.000027899, 0.005282 and 116.5162 respectively. As for GARCH(1,3), the values for MSE, RMSE and MAPE are 0.00002788, 0.005281 and 118.1037. It can be concluded that ARIMA(1,1,0) has the lowest MAPE while GARCH(1,3) has the lowest MSE and RMSE. However, most preferable in comparing using error measure is by MSE and RMSE. Therefore, GARCH(1,3) is been chosen as the best model for forecasting because the values of MSE and RMSE are lowest. Figure 9 shows the forecasting time series plot using GARCH(1,3).

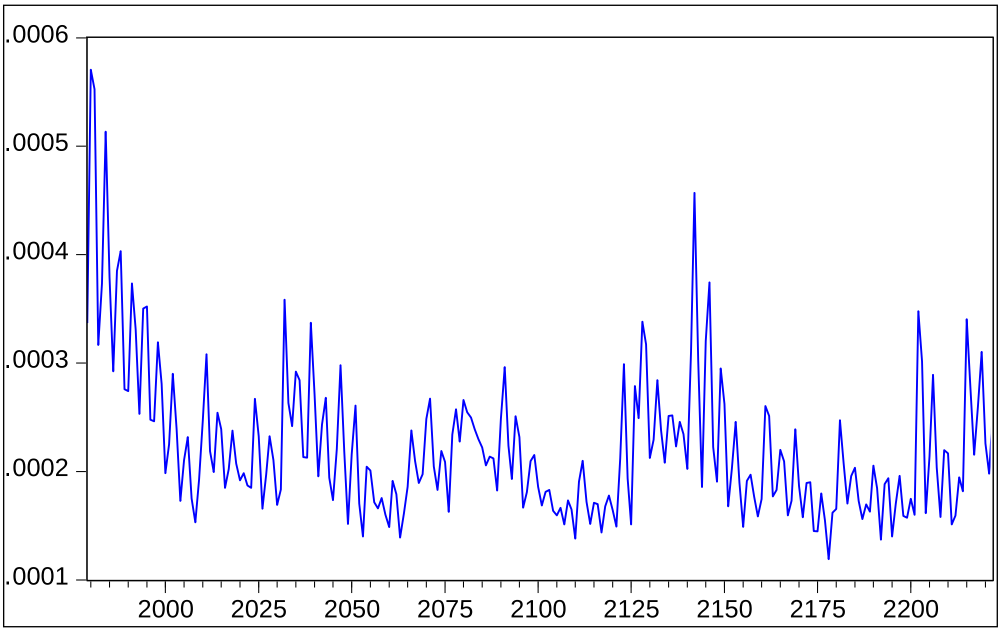

Figure 9. Forecasting time series plot of KLCI for year 2019.

Conclusions

Time series is a set of numbers that measure the status of some activities over time. It is well acknowledged that a series of data are regularly affected with outliers. The outliers may affect the results in forecasting since the tendency of parameter estimates created by outlier will reduce its effectiveness.

Daily data of Kuala Lumpur Composite Index (KLCI) for the period ranging from January 2011 to December 2018 are used in this study. As an overview, the result of this study shows there are 32 observations that are additive outlier (AO) and 31 observations are innovational outliers (IO) in the data. Then, the stationary of the KLCI data set is examined and it shows that the series are non-stationary. Hence, differencing is performed to make the data series stationary. Box-Jenkins method is used to find the most optimal lags for AR and MA. The result shows that ARIMA(1,1,0) and GARCH(1,3) are chosen as the best models since the values of error measures are the lowest among others. For each of the model evaluated, the values of MSE, RMSE and MAPE are calculated. Furthermore, the values of MSE, RMSE and MAPE in evaluation are used to compare the best model between ARIMA and GARCH model. Most of the researcher also used the MSE, RMSE and MAPE to find the best model for comparison between ARIMA and GARCH where the result shows GARCH model is preferable when to forecast volatility.

References

Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, *16*(2), 270-301.

Anderson, O. D. (1977). The Box-Jenkins approach to time series analysis. *RAIRO-Operations Research*, *11*(1), 3-29.

Bianco, A. M., Ben, G. M., Martinez, E. J., & Yohai, V. J. (2001). Outlier detection in regression models with ARIMA errors using robust estimates. *Journal of Forecasting*, *20*(8), 565-579.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*(3), 307-327.

Box, G. E., & Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American statistical Association*, *65*(332), 1509-1526.

Cerqueira, V., Torgo, L., & Mozetic, I. (2019). Evaluating time series forecasting models: An empirical study on performance estimation methods.

Chan, W. S., & Cheung, S. H. (1994). On robust estimation of threshold autoregressions. *Journal of Forecasting*, *13*(1), 37-49.

Chang, I., & Tiao, G. C. (1983). Effect of exogenous interventions on the estimation of time series parameters. In *Proceedings of the American Statistical Association, Business and Economics Statistics Section* (pp. 532-537).

Chong, C. W., Ahmad, M. I., & Abdullah, M. Y. (1999). Performance of GARCH models in forecasting stock market volatility. *Journal of forecasting*, 18(5), 333-343.

Deneshkumar, V., & Kannan, K. S. (2011). Outliers in time series data. *Int. J. Agricult. Stat. Sci*, *7*(2), 685-691.

Dijk, D. V., Franses, P. H., & Lucas, A. (1999). Testing for smooth transition nonlinearity in the presence of outliers. *Journal of Business & Economic Statistics*, *17*(2), 217-235.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987-1007.

Fox, A. J. (1972). Outliers in time series. *Journal of the Royal Statistical Society: Series B (Methodological)*, *34*(3), 350-363.

Gourieroux, C. (1997). Univariate ARCH Models. In *ARCH Models and Financial Applications* (pp. 29-41). Springer, New York, NY.

Hasan, H., Chong, S. H., Ooi, W. T., & Ismail, M. T. (2016). Time Series Modelling on KLCI Returns in Malaysia.

Hussin, N. H., Yusof, F., & Norrulashikin, S. M. (2021). Forecasting Wind Speed in Peninsular Malaysia: An Application of ARIMA and ARIMA-GARCH Models. *Pertanika Journal of Science & Technology*, *29*(1).

Ljung, G. M. (1993). On outlier detection in time series. *Journal of the Royal Statistical Society: Series B (Methodological), 55*(2), 559-567.

Merabet, F., Zeghdoudi, H., Yahia, R. A., & Saba, I. (2021). Modelling of Oil Price Volatility Using ARIMA-GARCH Models.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, 347-370.

Omar, N. A. B., & Halim, F. A. (2015), May. Modelling volatility of Malaysian stock market using garch models. In 2015 *International Symposium on Mathematical Sciences and Computing Research* (iSMSC) (pp. 447-452). IEEE.

Raicharoen, T., Lursinsap, C., & Sanguanbhokai, P. (2003), May. Application of critical support vector machine to time series prediction. In *Proceedings of the 2003 International Symposium on Circuits and Systems, 2003. ISCAS'03.* (Vol. 5, pp. V-V). IEEE.

San, O. T., Chen, L. H., & Heng, T. B. (2011). An Evaluation Forecasting Techniques in Kuala Lumpur Stock Exchange (KLSE) Finance. *International Journal of Business Management & Economic Research*, *2*(6).

Shi, L., & Chen, G. (2008). Detection of outliers in multilevel models. *Journal of statistical planning and inference*, *138*(10), 3189-3199.

Zainol, M. S., Zahari, S. M., Ibrahim, K., Zaharim, A., & Sopian, K. (2010), January. Additive outliers (AO) and innovative outliers (IO) in GARCH (1, 1) processes. In *Proceedings of the 2010 American conference on Applied mathematics January* (pp. 471-479).