INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN ECONOMICS & MANAGEMENT SCIENCES

OPEN ACCESS JOURNAL

Vol 11, Issue 2, (2022) E-ISSN: 2226-3624

Modelling and Forecasting the Volatility and Price of Malaysian Stock Market

Nurul 'Izzumi Nadhirah Ibrahim[1] and Nurul Nisa' Khairol Azmi[2]

[1]Faculty of Computer and Mathematical Sciences, Universiti Teknologi Mara, 40450, Shah Alam, Selangor, Malaysia, [2]Faculty of Computer and Mathematical Sciences, Universiti Teknologi Mara, Cawangan Negeri Sembilan, Seremban Campus,70300, Seremban, Malaysia

Email: [1]izzumi.nadhirah@gmail.com, [2]nurulnisa@uitm.edu.my

To Link this Article: http://dx.doi.org/10.6007/IJAREMS/v11-i2/12304 DOI:10.6007/IJAREMS/v11-i2/12304

Published Online: 15 May 2022

Abstract

Modelling and forecasting volatility of a financial time series has been a significant area of research in recent years, owing to the fact that volatility is regarded as an essential notion in many economic and financial applications. Because volatility is not directly observable, financial analysts are especially eager to obtain an accurate estimation of this conditional variance process. As a result, a number of models have been developed that are specifically suited to estimate the conditional volatility of financial instruments, with the most well-known and widely used model being the conditional heteroscedastic models. The objectives of this research are; to model and forecast the volatility and price of the Malaysian stock market; to assess the performance of competing models and to simulate the volatility and price of Malaysian stock market. This research estimates and examines the performance of ARIMA and GARCH family type models, standardized GARCH and GARCH-M for symmetric model and EGARCH and GJR-GARCH for asymmetric model using daily return price data. For GARCH models, two distributions were used which were the normal distribution and the t-distribution. The Malaysian stock market which is Kuala Lumpur Composite Index (KLCI) was studied using daily data over an 11-years period beginning from 1st January 2010 and ending on 31st December 2020. The results showed that the ARIMA method is not suitable in forecasting long term data since the ARCH effect is present. While, the performance of asymmetric GARCH models (GJR-GARCH), especially when the fat-tailed densities are taken into account in the conditional volatility, are better than symmetric GARCH. In addition, the student-t distribution performs better than the normal distribution. Moreover, it was found that the AR (1) GJR-GARCH model provides the best forecast for the Malaysian stock market, KLCI. Thus, it was concluded that the asymmetric AR (1) GJR-GARCH model coupled with the student-t distribution, performed well in modeling KLCI dataset. The forecasting process resulted in three different outcomes where the first one foresees a stationary trend of RM1,825.00 all over the years. While the second forecast indicates a fluctuation of around RM1,825.00 to RM1,625.00 and the last outcome had forecast the best result where the trend shows a positive upward trend of over RM1,850.00 all over the year of 2021.

Keywords: GARCH Models, Asymmetric, Stock Market Indices, Volatility Modelling, Heteroscedasticity

Introduction

Time series data are often contaminated with volatility clustering. Forecasting will be far off from what we predict if the data used contains high volatility. The research on volatility identification is essential since their existence can lead to measurement error of the parameter, or inaccuracy in forecasting. As volatility is often perceived as a measure of risk, one is of course interested in forecasting the volatility.

Asymmetric phenomena sometimes arise in data series which tend to behave differently when the economy moves into recession rather than coming out. Many data series have shown periods of stability, followed by periods of instability with high volatility. This implies that over a period of time, volatility is irregular. Since volatility may be used to calculate return and price, it is necessary to estimate volatility as accurately as possible. A time series model can be used to represent the volatility of asset returns.

The employed time series model however must comply with the property of heteroscedasticity. Heteroscedasticity explained the changes of volatility over the time horizon. The GARCH model proposed by Bollerslev (1986) is brought upon as it is effective with data containing high volatility. Many researchers had been widely using the GARCH model in order to estimate the volatility of time series data or stock market index.

Therefore, the objectives of this study are

1. To model the volatility of Kuala Lumpur Composite Index (KLCI) stock market by employing different univariate specifications of GARCH type models for daily observations on the index returns series of the market over the period of 1st January 2010 to 31st December 2020;
2. To measure the performance of competing models; and
3. To simulate the volatility and price of KLCI.

Literature Review

Forecasting via ARIMA Model

Deepika et al (2012) sought to examine the predicting of gold prices in the short-term using the ARIMA model and monthly gold prices. However, the forecast for the three-month period demonstrated a downward trend, and the Ljung Box Q test rejected the null hypothesis of residuals being pure white noise, suggesting that ARIMA was an inadequate fit since residuals were contaminated. They concluded that, with too many complex factors influencing gold prices and the dynamics of supply and demand for gold changing, using the ARIMA framework to forecast gold prices is inappropriate.

The research was supported by Guha and Bandyopadhyay (2016) who investigated the implementation of the ARIMA time-series model to predict the potential gold price based on past data from November 2003 to January 2014 to minimize the probability of purchasing of gold and thus, to provide the consumer with guidance on the purchase or selling of yellow metal. The analysis of performance of the gold prices had found that the best model in predicting the future values of gold was ARIMA (1, 1, 1). The authors had also concluded that the limitation of this approach was that it was only used for short runs to pinpoint minor variations in data. In the event of unexpected change in the data set (when the variation is large), a change in government policies, or economic instability (structural break), it becomes burdensome to capture the exact change, and thus this model becomes inefficient to forecast in those particular circumstances. Furthermore, forecasting with the ARIMA method, is based on the assumption of linear historic data. However, there is no solid proof that the price of gold is linear in nature.

Forecasting via GARCH Model

Kingsley (2019) in his study, investigated the volatility in equity prices of insurance stocks traded on the floor of the Nigerian Stock Exchange. Excluding weekends and public holidays, the time series data covered almost five years starting from 4th of March 2011 to 31st of December 2015 resulting in approximately 1106 observations. The study concluded that the best model in capturing the presence of volatility in the insurance stocks through the information criteria of Akaike, Bayesian, Shibata and Hanna Quinn were the GARCH (0, 3) which was the same as ARCH (3) and GARCH (1, 1).

Meanwhile, Sharma, Aggarwal and Yadav (2020) had conducted a study on "Comparison of linear and non-linear GARCH models for forecasting volatility of select emerging countries". They had empirically investigated the volatility of financial markets of five major emerging countries (China, India, Indonesia, Brazil and Mexico) over a period of two decades from January 2000 to December 2019 using univariate volatility models namely GARCH (1,1), EGARCH (1,1) and TGARCH (1, 1). The results show that the GARCH (1,1) model outperforms non-linear GARCH models for forecasting volatility since the effect of leverage is negligible. China is thought to be the most volatile, followed by India, which is turbulent but positively biased, and Indonesia, which is the least volatile. The research can assist investors in improving the worldwide diversity of their portfolios and discovering the greatest hedging possibilities.

Forecasting the Kuala Lumpur Composite Index (KLCI)

Nor and Shamiri (2007) had modelled and forecasted the volatility of the Malaysian and the Singaporean stock indices using asymmetric GARCH models and the non-normal densities. In their paper, three GARCH (1, 1) models (GARCH, EGARCH and GJR-GARCH) were examined and estimated using daily price data. Using daily data of over a 14-years period, two Asian stock indices (KLCI and STI) were studied. Gaussian normal, Student-t and Generalized Error Distributions were applied on the competing models including GARCH, EGARCH and GJR-GARCH. The forecasting performances of the asymmetric GARCH models (GJR-GARCH and EGARCH) was estimated and shown to be better than the symmetric GARCH especially when fat-tailed densities were considered in the conditional volatility. Additionally, it was found that the best out-of-sample forecast for the Malaysian stock market was the AR (1)-GJR model. Meanwhile, better estimation for the Singaporean stock market was provided by the AR (1)-EGARCH.

Consequently, Shamiri and Isa (2009) conducted a study on modelling and forecasting volatility of the Malaysian stock markets. The performance of symmetric GARCH, asymmetric EGARCH and non-linear asymmetric NAGARCH models were compared with six error distributions which were normal, skew normal, student-t, skew student-t, generalized error distribution and normal inverse Gaussian. The distinction centered on two related aspects: the disparity between symmetrical and asymmetrical GARCH (GARCH versus EGARCH and NAGARCH) and the difference between normal tailed symmetrical, heavy tailed symmetrical distributions and both high-tailed and asymmetric distributions for forecasting KLCI stock-market index return volatility. As predicted, the KLCI leverage market shown by the EGARCH model was statistically significant with a negative sign suggesting that negative shocks mean a higher conditional variance over the next duration than positive shocks of the same sign, indicating that the presence of leverage effect was observed in the returns of the KLCI stock market index. However, the distinction between models for each density (normal versus non-normal) showed that, based on the various metrics used for the performance of the volatility prediction, the EGARCH model offered the best out-sample approximation for KLCI that was obviously superior to the symmetrical models. The results showed that non-normal distributions had better in-sample results than normal distributions. However, the out-of-samples findings showed less evidence of superior predictive capabilities. By looking at the overall results, it was arguable that an asymmetric model combined with student-t distribution would perform very well with the investigated dataset. The dynamics of the first and second moments of the KLCI model seem to be captured by the models. Thus, they concluded that rather than the choice of GARCH models, successful volatility model forecasts depended much more heavily on the choice of error distribution.

Methodology

Data Source

For this study, the time series forecasting methods are applied on the secondary data which is the Kuala Lumpur Composite Index (KLCI) data. The secondary data was obtained through the open-source Yahoo! Finance website. The daily data covers a period of 11 years which began on 1st January 2010 and ending on 31st December 2020. The total observations for the KLCI dataset are 2,695 representing the number of trading days. The data was converted into log return using the following formula:

$$R_t = (100) * (ln\, ln\, (P_t) - ln\, ln\, (P_{t-1}))$$

wherein is the natural logarithm operator; t is the time period in days (date for trading days of KLCI); R_t is the return for period t (log return for closing price of KLCI); P_t is the index closing price for period t (closing price of KLCI).

ARIMA Model

This study applied the Autoregressive Integrated Moving Average (ARIMA) model which was developed using Box-Jenkins's methodology. ARIMA model is a mathematical model that was designed by George Box and Gwilym Jenkins (1970) to forecast data from a specified time series using probabilistic approach. There are four basic models involved in the ARIMA modelling which are Autoregressive (AR), Moving Average (MA) model, Mixed Autoregressive Moving Average (ARMA) and Mixed Autoregressive Integrated Moving Average (ARIMA) model.

ARCH Model

Engle (1982) introduced a model time-varying conditional variance with autoregressive conditional heteroscedasticity (ARCH) model by using lagged disturbances. ARCH is a function of autoregression which assumes that the variance is not constant over time and also affected by past data. The idea behind this model is to see the relationship between the current and the previous random variable.

The ARCH model is built as: Let R_1, R_2, …, R_T be the sequence of random data, and be the set of random data up to time t, then ARCH model with degree q with respect to R_t is: $R_t | F_{t-1} \sim N(0,\sigma_t^2)$, where F_{t-1} is the information available at time t-1. Conditional variance of the residual ε_t which is σ_t^2, can be written as,

$$\sigma_t^2 = \omega + \alpha_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \cdots + \alpha_q \varepsilon_{t-q}^2 \qquad (1)$$

where ω and α are non-negative constant; the variance residual depends on the-q squares of residual, and is called ARCH. The ARCH model can be written as shown in Equation (2), (Brooks, 2014).

$$R_t = \omega + \sum_{i=1}^{p} \alpha_1 R_{t-i} + \varepsilon_t \tag{2}$$

Where $\varepsilon_t \sim N(0, \sigma^2)$; ω and α_1 are non-negative constant; ε_t denotes a discrete-time stochastic taking the form of $\varepsilon_t = z_t \sigma_t$ where $z_t \sim iid$ (0, 1), and σ_t is the conditional standard deviation of return at time t.

GARCH Model

The GARCH model is a generalized form of ARCH. This model is built to avoid the order of the ARCH model, which is too high. Bollerslev (1986) introduced the GARCH model which suggests that the time-varying volatility process is a function of both past disturbances and past volatility. The GARCH model not only observes the relationship among some residuals, but also depends on some past residuals. The GARCH model with degree p and q is defined as follows.

$R_t / F_{t-1} \sim N(0, \sigma_t^2)$

GARCH model allows the conditional variance based on the conditional variance of the previous lag. So, the equation of conditional variance becomes as presented by Equation (3),.

$$\sigma^2{}_t = \omega_0 + \sum_{i=1}^{q} \alpha_i \varepsilon^2{}_{t-i} + \sum_{j=1}^{p} \beta_j \sigma^2{}_{t-j} \tag{3}$$

Where ω_0 is the constant term; $\alpha_1, \alpha_2, \dots, \alpha_q$ represented the parameters of ARCH specifications; $\beta_1, \beta_2, \dots, \beta_p$ represented the parameters of GARCH specifications; p and q are the respective orders of ARCH and GARCH processes. The present values of the conditional variance are parameterized based on the q lag from the squares residual and the p lag of the conditional variance and is written as GARCH (p,q). So, the time-varying conditional variance of the GARCH model is heteroscedastic with both autoregression and MA (Wang, 2009). The simplest specification of this model is the GARCH (1, 1) model which can be written as shown in Equation (4), (Bollerslev, 1986).

$$\sigma_t^2 = \omega_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{4}$$

According to Bollerslev (1986), all the coefficients of the conditional variance specifications fulfilled the stationarity assumptions when 0 < β_1 < 1, 0 < α_1 < 1 and α_1 + β_1 < 1.

GARCH in Mean (GARCH-M) Model

Another well-known symmetric model is the GARCH in Mean (GARCH-M) model which was developed by Engle et. al (1987). In most of the financial markets, it is expected for risk to be compensated by a higher return and the return of a security may depend on its volatility. To model such a phenomenon, one might consider the GARCH-M model. This variant of the GARCH family allows the conditional mean of return series to depend on its conditional variance. A simple GARCH-M (1, 1) model can be defined by the two equations, the one for conditional mean is given by

$$r_t = \mu_t + \varepsilon_t \text{ where } \mu_t = \mu + \lambda\sigma_t^2 \quad (5)$$

The equation for conditional variance is the same as provided by the GARCH (p, q) model in Equation (3) and its specific case GARCH (1, 1) by Equation (4).

GJR-GARCH Model

This GJR-GARCH model was proposed by Glosten, Jagannathan and Runkle (1993). The generalized form was given by:

$$\sigma_t^2 = \omega_0 + \sum_{i=1}^{q} (\alpha_i {\varepsilon^2}_{t-i} + \gamma_i S_{t-i}^- {\varepsilon^2}_{t-i}) + \sum_{j=1}^{p} \beta_j {\sigma^2}_{t-j} \quad (6)$$

where α_0 and γ_i are non-negative constant term; S_t^- is a dummy variable. In this model, it is assumed that the impact of ${\varepsilon^2}_t$ on the conditional variance ${\sigma^2}_t$ is different when ε_t is positive or negative. The positivity of conditional variances is assured by $\omega > 0$, $\alpha \geq 0$, $\beta \geq 0$, and $\alpha + \gamma \geq 0$, the variances stationary is assured by $\alpha + \beta + 0.5\gamma < 1$, and S is an indicator function that expressed by

$$S_{t-i}^- = \{0 \; if \; R_{t-1} \geq 0, (good \; news) \; 1 \; if \; R_{t-1} < 0, (bad \; news)$$

Which can be interpreted as when $\gamma = 0$, the model reduces to the standard GARCH model which treats bad news ($R_{t-1} < 0$) and good news ($R_{t-1} \geq 0$) symmetrically: that is, bad news and good news have the same impact (αR_{t-1}^2) on the conditional variance σ_t^2. When $\gamma \neq 0$, the news impact is asymmetric: that is, bad news and good news have different impacts on the conditional variance. Bad news has an impact of $\alpha + \gamma$ on conditional variance, while good news has an impact of α on conditional variance. Hence, if $\gamma > 0$, bad news has a larger impact on conditional variance than good news.

It is for this reason that the dummy variable S_t^- takes the value '0' (respectively '1') when ε is positive (negative). It is worth noting that the TGARCH model of Zakoian (1994) is very similar to GJR. However, TGARCH models the conditional standard deviation instead of the conditional variance.

Exponential GARCH (EGARCH) Model

The first asymmetric GARCH model known as exponential GARCH model (EGARCH) was introduced by Nelson (1991). This model looks at the conditional variance and tries to accommodate for the asymmetric relation between stock returns and volatility changes. Nelson implements that by including an adjusting function g (z) in the conditional variance equation, it in turn becomes expressed by:

$$ln\, \sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i g(z_{t-i}) + \sum_{j=1}^{q} \beta_j ln(\sigma_{t-j}^2)\, t \qquad (7)$$

where, $z_t = \varepsilon_t / \sigma_t$ is the normalized residual series. The value of g (z_t) is a function of both the magnitude and sign of z_t and is expressed as:

$$g\,(z_t) = \underbrace{\theta_1 z_t}_{sign\,effect} + \underbrace{\theta_2[|z_t| - E|z_t|]}_{magnitude\,effect} \qquad (8)$$

Moreover, notice how $E|z_t|$ depends on the assumption made on the unconditional density. The EGARCH model differs from the standard GARCH model in two main aspects. First, it allows positive and negative shocks to have a different impact on volatility. Second, the EGARCH model allows large shocks to have a greater impact on volatility than the standard GARCH model.

Results

Descriptive Statistics of KLCI

Table 1 shows the descriptive statistics of KLCI series which consist of 2,695 samples of observation for all trading days. The data covers a period of 11 years, beginning on 1st January 2010 and ending on 31st December 2020. The mean return of 0% indicated that on the average, the return price is neither gaining profit or loss. However, it was accompanied by a volatility of 0.01%. This may be due to Malaysia's investors' normal approach which is to "wait-and-see" that is common for emerging markets and is consistent with the previous study (Ng, 2000; Mohd Nor and Shamiri, 2007).

Table 1

Summary Statistics for Daily Returns 1 January 2010 – 31 December 2020

Statistics	Sample	Mean	Median	Standard Deviation	Skewness	Kurtosis	Q(10)	$Q^2(10)$	LM
Values	2,695	0	0	0.01	-0.2	9.38	23.64**	11.65**	14.93**

** Significant at 5% significance level

The statistical features such as the skewness, kurtosis and their tests are shown in the table. The Ljung-Box Q-statistics, Q(10) and $Q^2(10)$ are reported under the null hypothesis of non-serial correlation tests in residuals of return and squared residuals of return respectively. At 5% significance level, the null hypothesis of non-serial correlation is rejected respectively. This time series has the typical features of stock returns as fat-tail, spiked peak and persistence in variance.

With the evidence of ARCH effects (Humala and Rodriguez, 2010) which is indicated by the Lagrange Multiplier (LM) test, it is possible to proceed to the second step of the analysis focused on the GARCH modelling of market's volatility.

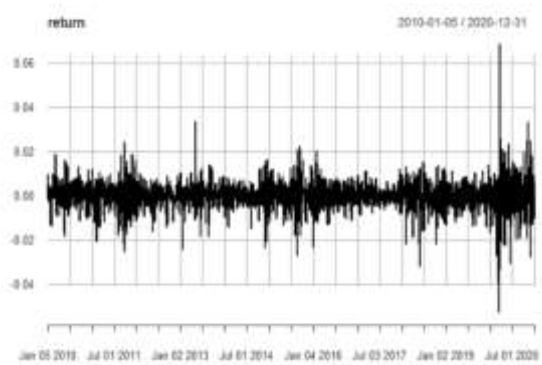


Figure 1: KLCI Log Returns

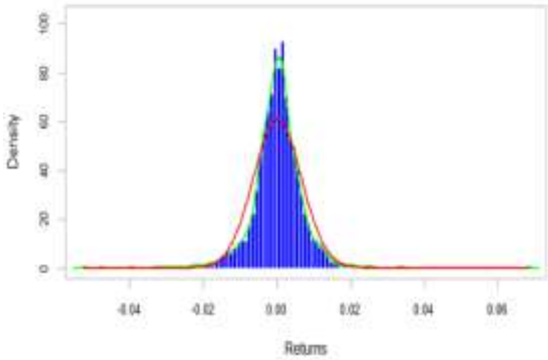


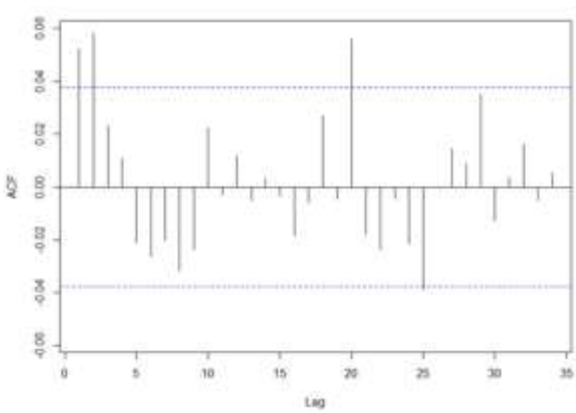
Figure 2: Histogram of Log Returns Distributions

Figure 1 shows the behavior of the KLCI log returns, over the sample period. There is evidence of volatility clustering and that large or small asset price changes tend to be followed by other large or small price changes of either sign (positive or negative). This implies that stock return volatility changes over time (Gallant et. al, 1991). In general, the series demonstrates the existence of the ARCH effect (also known as heteroscedasticity) ubiquitous in various financial time series data.

Figure 2 shows the distribution of KLCI log returns. The curve for returns is a bit taller and the green line which indicates the density has higher values implying that the tails are thicker. This means that there are days where higher or lower returns are obtained compared to the expectations of the normal distribution indicated by the red line.

ARIMA Model

The Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) were plotted to observe the stationarity of the time series model.



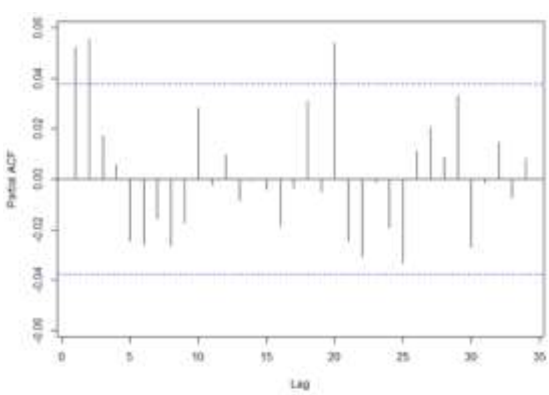


Figure 3: The ACF for Daily Return of KLCI   Figure 4: The PACF for Daily Return of KLCI

The ACF in Figure 3 shows two significant spikes at lag 1 and 2, then the rest taper off to zero as the lag increases. While the PACF in Figure 4 shows most of the spikes are not significant after the first and second lag. These behaviors indicate that the daily return of KLCI data meet stationary assumptions.

Table 2
The Augmented Dickey-Fuller (ADF) Test for Log Return Price of KLCI

Index	ADF Statistics	Critical Values (5% Significance Level)
KLCI	-14.068	0.01

From Table 2, the ADF test indicates that the log return is stationary. Hence, the null hypothesis of a unit root test at all conventional levels for the series can be rejected. Thus, it is concluded that the log return series is stationary over a period.
The ACF in Figure 3 shows significant spikes at lag 1 and 2 that suggest the order of MA is 2. On the other hand, two significant spikes were found at lag 1, 2 and 20 on the PACF plot as depicted in Figure 4. The significant spikes at lag 20 can be ignored since only the current event that is significantly affected are taken into consideration. Hence, two orders of AR might be suitable for the daily return of KLCI. Overall, the suggested ARIMA model to be fitted to daily return of KLCI are ARMA (1, 0), ARMA (1, 1), ARMA (1, 2), ARMA (2, 0), ARMA (2, 1) and ARMA (2, 2).

Table 3
Performance of ARMA Model on Daily Return of KLCI

ARMA Model	ARMA (1, 0)	ARMA (1, 1)	ARMA (1, 2)	ARMA (2, 0)	ARMA (2, 1)	ARMA (2, 2)
AIC	-19444.96	-19449.32	-19450.15	-19452.68	-19450.01	Non-invertible
BIC	-19427.26	-19425.72	-19420.65	-19434.98	-19420.51	Non-invertible

Table 3 shows the performance of competing models of ARMA based on AIC and BIC. ARMA (2, 0) is selected as the best ARMA model to forecast daily return of KLCI. For ARMA (2, 2), the parameter cannot be estimated since the matrix has become singular. Thus, the ARMA (2, 0) was chosen to be the best fit.

Table 4
ARIMA (2, 0, 0) with Zero Mean Model

	ϕ_1	ϕ_2
Coefficient	0.0496	0.0559
Standard Error	0.0192	0.0192
AIC=-19452.68 BIC=-19434.98		

Table 4 shows the result obtained for the best fit ARIMA model. The best model is ARIMA (2, 0, 0) with no-zero mean AIC value of -19452.68. The mean equation for the following ARIMA (2, 0, 0) model is as follow (SE in parenthesis):

$$\hat{R}_t = 0 + \underbrace{0.0496}_{(0.0192)} R_{t-1} + \underbrace{0.0559}_{(0.0192)} R_{t-1}$$

The ARIMA (2, 0, 0) was chosen to be the best fit by the auto arima function as it has the lowest value of AIC and BIC.

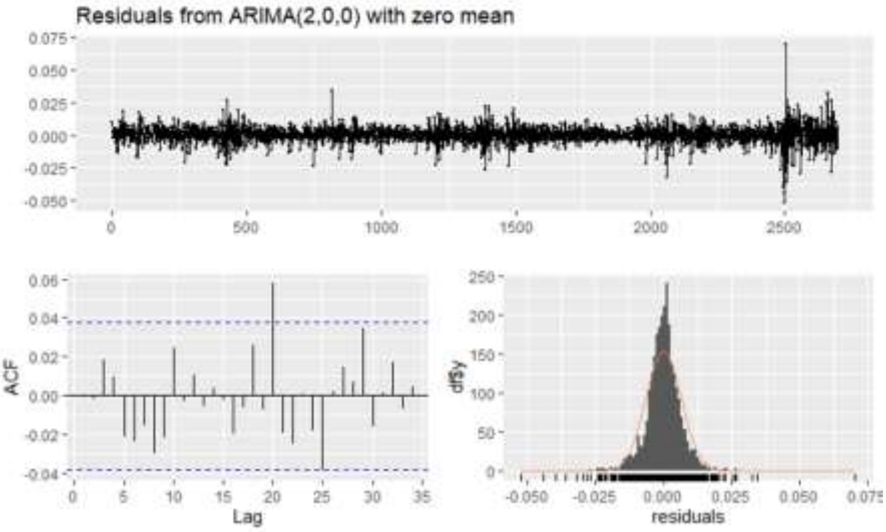


Figure 5: Residuals from ARIMA (2, 0, 0)

Figure 5 shows that the residuals from ARIMA (2, 0, 0) are well-behaved. The residual plot shows that the residuals are fluctuating around zero with high volatility. The ACF plots of residual show two significant spikes at lag 20 and lag 25. The histogram also seemed to be somewhat normally distributed. Next, the test for the ARCH effect was performed by applying the Ljung-Box test on the squared residuals of the ARIMA (2, 0, 0) model.

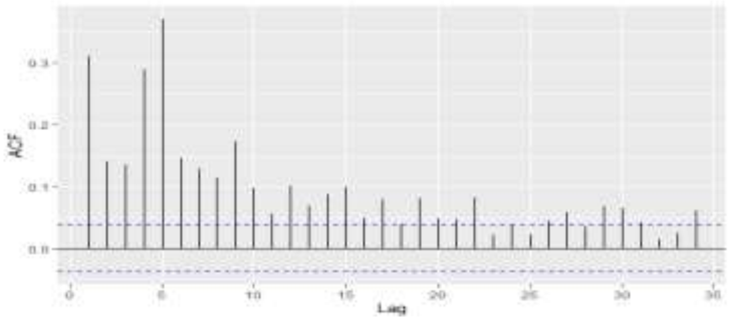


Figure 6: The ACF for the Squared Residuals of ARIMA (2, 0, 0) Model

From Figure 6, it can be observed that the ACF of squared residuals shows many significant lags. Hence, it can be concluded that there are indeed ARCH effects and thus the modelling of the volatility should be performed. In addition, the existence of spikes which were represented by the significant lags further proved that using the ARIMA model is not an appropriate and suitable approach.

GARCH Model

Table 5 shows the estimation results of the GARCH models. The application of asymmetric GARCH models appears to be appropriate. At standard levels, all asymmetric coefficients are significant. Furthermore, the Akaike Information Criteria (AIC) and Bayesian Information

Criterion (BIC) values emphasize the fact that the GJR-GARCH and EGARCH models outperform the standard and traditional GARCH.

Table 5
GARCH Models Estimation

	Normal GARCH (1, 1)	Student-t GARCH (1, 1)	GARCH-M (1, 1)	GJR-GARCH	AR (1) GJR-GARCH	EGARCH
μ	0.000155** (0.0001)	0.000084** (0.000097)	-0.000178** (0.000119)	0.000026** (0.000098)	0.000015** (0.000103)	0.000037** (0.000092)
ω	0.000001** (0.000001)	0.000001** (0.000001)	0.000001** (0.000000)	0.000001** (0.000000)	0.000001** (0.000000)	-0.163954** (0.028964)
α_1	0.100062** (0.016318)	0.096497** (0.015497)	0.049793** (0.006397)	0.047929** (0.006279)	0.046326** (0.006551)	-0.059914** (0.010962)
β_1	0.878525** (0.016483)	0.887765** (0.014925)	0.903469** (0.006973)	0.905707** (0.010596)	0.902941** (0.011451)	0.984061** (0.002916)
γ_1			0.058826** (0.016713)	0.064614** (0.016802)	0.072040** (0.017717)	-0.160916** (0.054748)
λ			7.963907** (3.050976)			
φ_1					0.053803** (0.019919)	
Skew		0.901984	0.903267	0.899720	0.900596	0.901582
Shape		6.227997	6.284410	6.359192	6.532823	6.326640
AIC	-7.4926	-7.5575	-7.5624	-7.5617	-7.5637	-7.5610
BIC	-7.4839	-7.5443	-7.5448	-7.5464	-7.5641	-7.5457
Q^2 (10)	8.475	8.934	8.934	9.681	8.136	8.136
LM (5)		2.7325	1.6097	1.3073	1.1757	1.46045

Standard errors are given in parentheses. **: Significant at 5% respectively

As is typical of GARCH model estimates for financial asset returns data, the sum of the coefficients on the lagged squared error (α_1) and the lagged conditional variance (β_1) is close to unity 1.00 and 0.99 with the normal and student-t error term respectively, this implies that shocks to the conditional variance will be highly persistent indicating that large changes and small changes tend to be followed by small changes, this mean volatility clustering is observed in KLCI financial returns series.

According to the Box-Pierce statistics for the squared standardized residuals with lag 10, all of the models appear to do a reasonable job of capturing the dynamic of the first two moments of the series, which are all non-significant at the 5% level. The conditional heteroscedastic that occurred when the test was done on the pure return series (Table 1) was removed when the LM test for the presence of ARCH effects was performed at lag 5.

In EGARCH, since the value of α_1 (-0.059914) < 0, the leverage effect is significant, implying that the volatility reacts more heavily to negative shocks. Also, the indicator for asymmetric volatility, estimates shows that the coefficient for asymmetric volatility, γ_1, is negative. This indicates that negative shocks imply a higher next period conditional variance than positive shocks of the same sign.

However, the goodness-of-fit test revealed that at 5% significance level, the p-value of GARCH (1, 1) using normal distribution is significant thus, rejecting the hypothesis that a normal distribution is appropriate. This result is congruent with the research by Shamiri and Isa (2009).

From the AIC and BIC, it can be concluded that the best model is AR (1) GJR-GARCH with t-distribution as the model has the lowest AIC and BIC value of -7.5637 and -7.5641. The statistical result implied that the parameter of estimations for AR (1) GJR-GARCH model, the ARCH coefficients (α1) and the GARCH coefficients (β1) in the conditional variance equation of the AR (1) GJR are highly significant with a p-value equal to 0.0000 for both parameters. The term for γ1 is also significant which indicates that negative shocks imply a higher next period conditional variance than the positive shocks. This suggests the existence of leverage effect in returns of the KLCI stock market index.

In fact, the asymmetric model of AR (1) GJR-GARCH paired with student-t distribution will perform very well with the data series that is being investigated. Thus, AR (1) GJR-GARCH model was chosen as the best GARCH model to be used to simulate the stock prices of KLCI. This result supports the previous study by Nor and Shamiri (2007) which also concludes the AR (1) GJR-GARCH model as the best out-of-sample forecast for the Malaysian stock market.

Table 6
Simulation for GARCH Model

Parameter	True Estimation Value	Simulation		
		N = 100	N = 1000	N = 10000
GARCH (1, 1)				
μ	0.000084	0.000234	0.00017941	0.00006735
ω	0.000001	0.000000522	0.000000022	0.00000001981
α	0.096497	0.060067	0.0024789	0.0040096
β	0.887765	0.9260562	0.9964374	0.9948905
GARCH-M(1,1)				

μ	-0.000178	0.000712	0.000036	-0.000147
λ	7.963907	-5.430123	-4.208556	5.948482
ω	0.000001	0.000003	0.000001	0.000001
α	0.049793	0.116405	0.026738	0.038880
β	0.903469	0.771762	0.876713	0.907505
γ	0.058826	-0.011623	0.127619	0.069123
EGARCH(1,1)				
μ	0.00001145	-0.00222238	0.00034293	-0.00001236
ω	-0.2365437	-0.2001033	-0.2844630	-0.16406185
α	-0.0681282	-0.2414796	-0.0515247	-0.0694388
β	0.9772659	0.97691880	0.97274822	0.9838404
γ	0.2035275	-0.2278832	0.13874991	0.1749741
GJR-GARCH(1,1)				
μ	0.000026	0.0001084	0.0000954	0.00005382
ω	0.000001	0.000001133	0.000000802	0.000000687
α	0.047929	0.062697	0.052671	0.051782
β	0.905707	0.754522	0.876096	0.893016
γ	0.064614	0.356331	0.131057	0.078657
AR (1) GJR-GARCH(1,1)				
μ	0.00001463	0.0001688	0.0001226	0.00006402
φ	0.05380269	-0.2276255	0.0125386	0.03578909
ω	0.000000653	0.000001085	0.000000857	0.0000007132
α	0.04632612	0.02927954	0.0528735	0.05108231
β	0.9029412	0.7642686	0.8717048	0.8899644
γ	0.07203987	0.4047523	0.1369588	0.08520867

The simulations are carried out by generating 100, 1,000 and 10,000 returns data respectively from each model using the true parameter values presented in the tables. The initial values for each parameter of GARCH (1, 1), GARCH-M, EGARCH, GJR-GARCH and AR (1) GJR-GARCH model were set according to the respective values obtained from the model estimation of each GARCH model.

As observed from Table 6, the larger the sample, the closer the estimation approximates to the true parameters. The μ parameter denotes the average mean return. As a result, its estimate should converge to zero as $n \rightarrow \infty$. Furthermore, the results also indicated that all extended models fit significantly better than the standard model. This result shows that all extended GARCH models have the potential to outperform the standard GARCH (1, 1) model.

Forecasting Daily Return of KLCI Price

The best model, AR (1) GJR-GARCH was used to forecast the daily return of KLCI. A 10-days-ahead forecast was performed on the daily return of KLCI data. The date for every closing price is 4th January for every stock market since the initial 3 days were excluded as public holidays (Samuelsson, 2021).

Table 7
Forecast Value and Sigma Value of Daily Return of KLCI Price

	Date	Return	Sigma
T + 1	4/1/2021	-0.0005489	0.009033

T + 2	5/1/2021	-0.00001569	0.008996
T + 3	6/1/2021	0.00001300	0.008959
T + 4	7/1/2021	0.0001454	0.008923
T + 5	8/1/2021	0.0001462	0.008888
T + 6	11/1/2021	0.0001463	0.008852
T + 7	12/1/2021	0.0001463	0.008818
T + 8	13/1/2021	0.0001463	0.008783
T + 9	14/1/2021	0.0001463	0.008750
T + 10	15/1/2021	0.0001463	0.008716

A negative return was forecasted for the initial two days. Then, the return starting shows an increasing trend towards positive returns. The expected returns for the next 8-days then were expected to be positive. Thus, the amount of loss was also expected to decrease as the forecast value begins to increase positively.

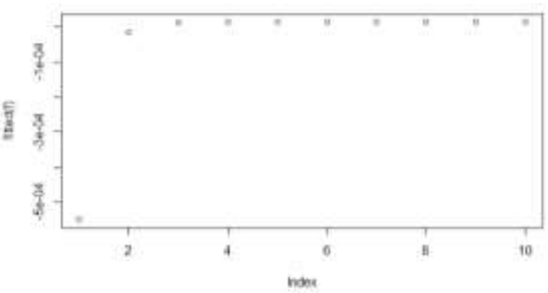


Figure 7: Forecast for 10-days-ahead of KLCI

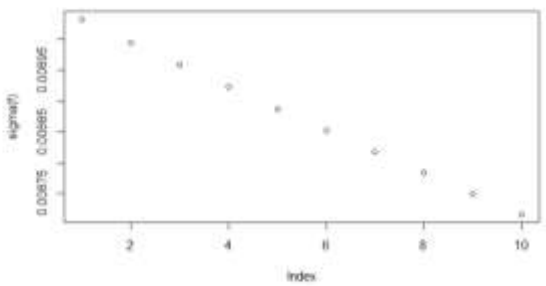


Figure 8: Variance of Forecasted KLCI

Figure 7 shows the forecasting period of 10-days-ahead for KLCI data. It was observed that the KLCI stock price will continue to increase in the future. Figure 8 depicted the variance of the KLCI data that was forecasted for a period of 10-days-ahead. It can be concluded that as the time increases, the risk for the investors to invest in the stock market will decrease in the future. Therefore, the investors can increase the weightage to the risky assets so that they can gain more profit

Table 8
Forecast Value and Sigma Value of Daily Return of KLCI Price

Date	Forecast Value	Actual Value	Forecast Error
4/1/2021	-0.0005489	-0.0005489	0.00%
5/1/2021	-0.00001569	0.003606707	0.01%
6/1/2021	0.000013	-0.01018435	0.01%
7/1/2021	0.0001454	0.006897115	0.01%
8/1/2021	0.0001462	0.018865217	0.01%
11/1/2021	0.0001463	-0.00976004	0.01%
12/1/2021	0.0001463	-0.003221518	0.01%

13/1/2021	0.0001463	0.015291184	0.01%
14/1/2021	0.0001463	-0.00059877	0.008%
15/1/2021	0.0001463	-0.005318791	0.01%
Average Forecast Error			0.009%

From Table 8, the forecasting error testing conducted on predicted data using AR (1) GJR-GARCH model found that the average percentage of forecasting error is 0.009%, this result is rational because the value did not exceed 5% margin of error (Khair, Fahmi, Al Hakim and Rahim, 2017). Thus, this implies that the AR (1) GJR-GARCH model is a reliable model in forecasting the KLCI.

Figure 9: Forecast Value vs Actual Value

Figure 9 shows the comparison between forecast value against the actual value of KLCI. Although the actual value of KLCI for the first 10 days were quite different from the forecast value, the forecast error is not large which was below the 5% margin of error. Thus, the value was deemed to be reliable for forecasting.

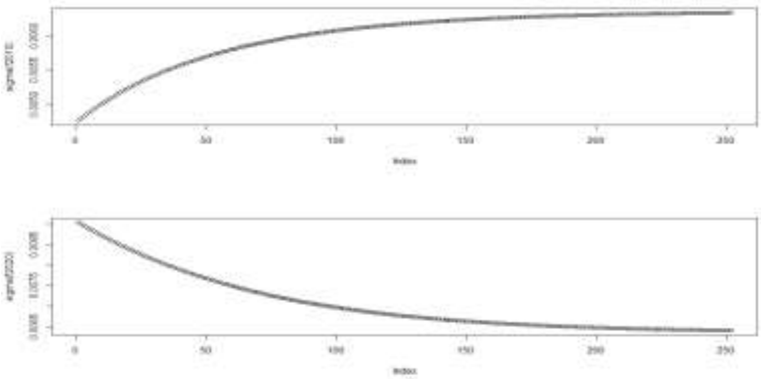


Figure 10: Forecasting Simulation for 2010 and 2020

The volatility was simulated for the period of 2010 and 2020. From Figure 10, low volatility was observed during 2010 where it was forecasted that for one year after 2010, the volatility was likely to rise. The result is consistent with the KLCI data obtained from Yahoo! Finance which was used in this study. At the end of 2020, the volatility was high. So, based on the forecast, it was expected that the volatility was going to decrease in 2021. Basically, it was expected that the volatility would fall below average.

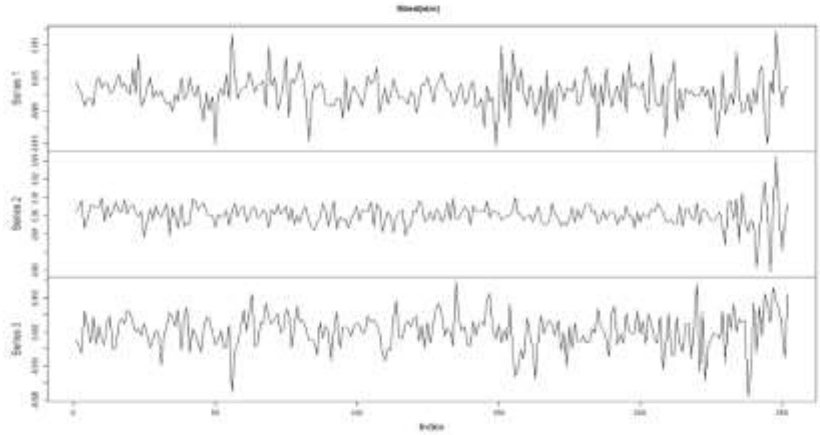


Figure 9: Forecasted KLCI for one Year

Figure 11 depicted three different time series forecasted for the simulated returns. This simulation was forecast for one year. For series 1, it can be observed that the time series shows a fluctuating trend with signs of volatility throughout the series. While for series 2, the plot suggested a more stationary and no trends or seasonal components before beginning to depict a variability towards the end of the year. For series 3, there is evidence of volatility clustering where large asset prices changes tend to be followed by other large price changes and vice versa.

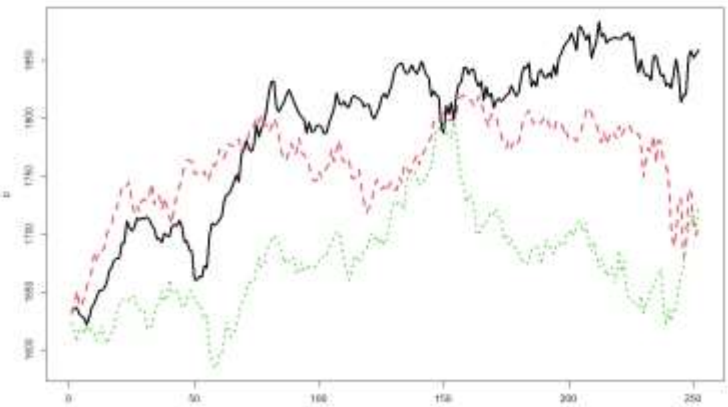


Figure 12: Forecast Value of KLCI Stock Market Prices in 2021

Figure 12 depicted the value for KLCI stock market prices forecasted for year 2021. The final closing value for 31st December 2020 was RM1,627.21 and then used as the starting value to forecast the prices for year 2021 as shown by the three forecasted series in the plot. The stock price for KLCI forecasted for the red series which touches around RM1,825.00 and then more or less hovers around the number before falling at the end of the year. As per the green line forecast, in fact it continued to rise over RM1,800.00 and then fell back towards RM1,625.00 before going higher towards the end of the year. Lastly, the black simulated line in fact provides the best forecast results out of the three. As per the black forecast line, KLCI may see a stock price of over RM1,850.00 before the end of 2021.

Conclusion

In conclusion, the study is carried out to model and forecast the price and volatility of the Malaysian stock market namely the Kuala Lumpur Composite Index (KLCI). For the first objective; to model the volatility of KLCI using ARIMA model and GARCH processes model. The findings conclude that the ARIMA model is not suitable in modelling the volatility of KLCI since it cannot cater to the large variations that exist in highly volatile data and the long-term changes. Thus, the GARCH processes model was used in modelling the volatility where it can be inferred that volatility clustering in KLCI is quite persistant. From the GARCH family type model parameter estimation, the KLCI data was concluded to be highly volatile as there exists a highly volatile clustering.

For the second objective, the GARCH processes model performance was assessed. The results concluded that the asymmetric model (AR (1) GJR-GARCH) had by far outperformed the symmetric model. The asymmetric AR (1) GJR-GARCH model coupled with the student-t distribution had performed very well with the KLCI dataset. The simulation process had further approved the findings by depicting that as the sample got larger during simulation, the closer the estimation approximates to the true parameters. The estimates of the AR (1) GJR-GARCH model show the nearest and most accurate simulation results.

Lastly, AR (1) GJR-GARCH, which is the best model, has been used to forecast the daily return of KLCI price. The forecasting process resulted in three different outcomes where the first one foresees a stationary trend of RM1,825.00 all over the years. While the second forecast indicates a fluctuation of around RM1,825.00 to RM1,625.00 and the last outcome had forecast the best result where the trend shows a positive upward trend of over RM1,850.00 all over the year of 2021.

As this research only considered data with high volatility, it is recommended for future studies to use data with high volatility together with the existence of outliers to tackle and solve more problems. Besides, the future researcher can also consider using other GARCH family types such as the threshold GARCH (TGARCH) and the power GARCH (PGARCH). Last but not least, it is recommended for the future researcher to employ different time series data other than the stock market data.

References

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

Deepika, M. G., Gautam, N., & Rajkumar, M. (2012). Forecasting price and analysing factors influencing the price of gold using ARIMA model and multiple regression analysis.

Engle, R. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica, 50*, 987 - 1007.

Gallant, A., Rossi, P., & Tauchen, G. (1992). Stock Prices and Volume. *Review of Financial Studies, 5*, 199 - 242. doi:https://doi.org/10.1093/rfs/5.2.199

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *The Journal of Finance, 48*(5), 1779-1801. doi:10.1111/j.1540-6261.1993.tb05128.x

Guha, B., & Bandyopadhyay, G. (2016). Gold Price Forecasting using ARIMA Model. *Journal of Advanced Management Science, 4*(2), 117 - 121.

Humala, A., & Rodriguez, G. (2010). Foreign exchange intervention and exchange rate volatility in Peru. *Applied Economics Letters, 17*(15), 1485-1491. doi:10.1080/13504850903049643

Khair, U., Fahmi, H., Al Hakim, S., & Rahim, R. (2017). Forecasting Error Calculation with Mean Absolute Deviation and Mean Absolute Percentage Error. *International Conference on Information and Communication Technology (IconICT), 930*. doi:10.1088/1742-6596/930/1/012002

Kingsley, A. (2019). Volatility Modelling using Arch and Garch Models (A Case Study of the Nigerian Stock Exchange). *International Journal of Mathematics Trends and Technology (IJMTT), 65*(4), 58 - 63. doi:10.14445/22315373/IJMTT-V65I4P511

Nor, M. A., & Shamiri, A. (2007). Modeling and Forecasting Volatility of the Malaysian and the Singaporean Stock Indices using Asymmetric GARCH Models and Non-normal Densities. *Malaysian Journal of Mathematical Sciences, 1*(1), 83 - 102.

Ng, A. (2000). Volatility spillover effects from Japan and the US to the Pacific-Basin. *Journal of International Money and Finance, 19*(2), 207 - 233.

Shamiri, A., & Isa, Z. (2009). Modeling and Forecasting Volatility of the Malaysian Stock Markets. *Journal of Mathematics and Statistics, 5*(3), 234 - 240.

Sharma, S., Agarwal, V., & Yadav, M. (2020). Comparison of linear and non-linear GARCH models for forecasting volatility of select emerging countries. *Journal of Advances in Management Research*. doi:10.1108/JAMR-07-2020-0152

Wang, P. (2009). *Financial Econometrics* (2nd ed.). New York: Routledge, Taylor and Francis Group.

Zakoian, J. M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and Control, 18*(5), 931-955. doi:10.1016/0165-1889(94)90039-6