Cawangan Kedah
Kampus Sungai Petani

e-PROCEEDINGS

of The 5th International Conference on Computing, Mathematics and Statistics (iCMS2021)

4-5 August 2021

Driving Research Towards Excellence

e-Proceedings
of the 5th International Conference on Computing, Mathematics and Statistics (iCMS 2021)

Driving Research Towards Excellence

e-ISBN: 978-967-2948-12-4
DOI

Library of Congress Control Number:

Publication by
Department of Mathematical Sciences
Faculty of Computer & Mathematical Sciences
UiTM Kedah

TABLE OF CONTENT

PART 1: MATHEMATICS

PART 2: STATISTICS

PART 3: COMPUTER SCIENCE & INFORMATION TECHNOLOGY

PART 4: OTHERS

INVESTMENT IN MALAYSIA: FORECASTING STOCK MARKET USING TIME SERIES ANALYSIS

Nuzlinda Abdul Rahman[1], Chen Yi Kit[2], Kevin Pang[3], Fauhatuz Zahroh Shaik Abdullah[4] and Nur Sofiah Izani[5]

[1,2,3,4] School of Mathematical Sciences, Universiti Sains Malaysia, 11800 Minden, Pulau Pinang, Malaysia,
[5] Gelugor, 11700, Pulau Pinang, Malaysia
([4] fauhatuz@usm.my)

Abstract. Investment is the asset committed to generate additional income. Investing is often described as the process of laying out money for now in the expectation of gaining more money in the future. Every investment comes with a different magnitude of risk. This study aims to investigate the performance of the Kuala Lumpur Composite Index (KLCI) from 1997 to 2018, to forecast the value of KLCI stock market for the coming three years and to investigate whether investment in stock market is the best in terms of profit for the coming three years. The ARIMA model approach was applied to study the KLCI time series. Model building and diagnostic checking techniques were conducted to determine a statistically adequate model. Result indicate that the stock market is highly volatile. For an investor which is willing to take high risk, investment on KLCI stock market will be the suitable choice.

***Keywords*:** ARIMA

1. Introduction

Stock market investment is an important and popular form of investment all across the world. In general, anyone who wishes to invest in the stock market would need to trade stocks in markets often called stock exchanges. Companies which wish to list their shares in Malaysia can do so on the various board of the local stock exchange called Bursa Malaysia. Investors can then buy and sell the companies shares by trading on the Bursa. There are various indices to measure a stock markets performance. For example, the Dow Jones and S&P 500 in New York and the FTSE 100 in London.

Today, Bursa Malaysia has become one of the largest bourses in Asia and successfully marked down up to 1145 companies with a combination of about $235.28 billion in the market capitals in February 2014 (Arshad and Yahya, 2016). According to the official website of Bursa Malaysia, it is now one of the largest bourses in ASEAN and is home to more than 900 companies across 60 economic activities. Bursa Malaysia has used the FTSE Bursa Malaysia KLCI values as its main index, which consists of 30 largest companies listed on the main board of Bursa Malaysia. The daily KLCI index will rise or fall based on the weighted daily price performance of these 30 largest stocks. Thus, the KLCI can be regarded as a measure of performance of companies listed on the Bursa Malaysia.

The price performance over a period of time of companies constituting the KLCI clearly depends on a whole multitude of factors. It is well-known that these factors include general macroeconomic factors which are not directly related to a particular company. For example, financial crises such as the Asian financial crisis of 1997 and the US financial crises of 2008 had a serious negative impact on the KLCI (Goh and Lim, 2010). Fortunes and lost have been made in stock markets all over the world. It would be highly beneficial if one can forecast the performance of investments in companies listed on the stock market with a reasonable degree of accuracy.

There are several time series modeling techniques to obtain reliable forecasts of investments. An important forecasting technique introduced in 1970 called Box and Jenkins is the Autoregressive Integrated Moving Average (ARIMA) model. ARIMA model is good for short-term prediction (Li et.al, 2016 and Ayodeli et.al 2014).

The objectives of this study are to identify the most appropriate time series model for the KLCI, to forecast the value of KLCI stock market for the coming three years and to investigate whether investment in stock market is the best in terms of profit for the coming three years.

All the data set are obtained from the official website of Bank Negara Malaysia and Data Stream Database. The study period starts from January 1997 until December 2018 with a total duration of 22 years. Since the data is collected based on monthly basis, there are 264 observations available for analysis.

The time series analysis of Box and Jenkins model is used in this research to forecast whether the stock market gives a convincing and positive of return to the investors.

2. Literature

Li *et.al* (2016) conducted forecasting analysis of Shanghai stock index based on ARIMA model. The Shanghai Composite Index monthly closing price was collected from January 2005 until October 2016 and were used to build the ARIMA model. Forecasting for two months were carried out and compared with the actual value to investigate whether the ARIMA model fitted is adequate for the short-term Shanghai Stock Index prediction. Throughout the analysis, ARIMA (4,1,4) was found to be the most suitable model for the data. It was found that the ARIMA model has distinct advantages for forecasting Shanghai Stock Index. Since the study was aimed to forecast the closing price of the last two months in 2016, the ARIMA model is the ideal method because it is good for the short-term prediction. The result shows the Shanghai Composite Index have a small rise in the last two months of 2016.

Another study on stock price prediction using the ARIMA model was carried out by Ayodele *et.al* (2014). The researchers develop the ARIMA model based on the Nokia stock index from April 2005 to February 2011 with 3990 observations and Zenith Bank stock index from January 2006 to February 2011 with 1296 observations. The results revealed that the performance of the ARIMA model is quite good since the predicted values are fairly related to the actual value. They concluded that the ARIMA model is good for short-term prediction and could compete with any other forecasting method in predicting the short-term trend.

Alkhazaleh and Hussein (2015) have also conducted a study on forecasting insurance sector volatility on the Amman Stock Exchange using ARIMA model. The researchers wished to predict the volatility on the Amman Stock Exchange using Box-Jenkins model. Weekly data of Amman Stock Exchange were accumulated using historical indices from January 2005 to April, 2010. In this study, the ARIMA model has shown its advantages in forecasting the stock market data.

A similar study on forecasting banking volatility on the Amman Stock exchange in Jordan by using ARIMA model was conducted by Alkhazaleh (2018). The purpose of this study is to forecast the banking sector volatility on the Amman Stock Exchange which is an emerging market. It aimed to obtain an adequate ARIMA model to predict the volatility characteristic of the stock market. Data on a weekly basis from January, 2010 to April, 2015 was collected from the "Amman Stock Exchange" website. Unit root test was implemented for checking of stationarity of the data, then indicators such as minimum mean square error and t-statistics value were selected as the criteria for selecting the most adequate ARIMA model. Akaike's Information Criterion (AIC) was employed as the indicator to evaluate the ARIMA model chosen. Based on the results, ARIMA (2,0,2) was found to be the best model for this data. ARIMA model is a good statistical tool for forecasting oscillating variables.

Zhang and Li (2016) studied financial time series analysis model for stock index forecasting. This research implemented the wavelet analysis together with the ARMA model for time series modelling and forecasting. They aimed to improve the original mining technologies by building an adequate time series model for transactions in the stock market. Wavelet decomposition is used to identify and separate the hidden periods and nonstationary factors and its features are fully applied on Elman dynamic neural network and ARMA model. The ARMA model is adopted for scale transformation series whereas neural network is adopted for wavelet transformation series. The integration of predicting results on all the scale regions by wavelet reconstruction is used to acquire the final forecasting. 1426 observations from stock index transaction in China were collected from October 31, 2013, to January 31, 2014.

In another study, Chu et al. (2009) conducted the stock index forecasting using fuzzy dual-factor time series. They suggested a dual-factor modified fuzzy time series model, which account stock index and trading volume to forecast stock index. TAIEX (Taiwan stock exchange capitalization weighted stock index) and NASDAQ (National Association of Securities Dealers Automated Quotations) are chosen as the experimental data. Besides, they have chosen two multiple-factor models, Chen's (2000) and Huarng and Yu's (2005) as comparison models. The result of the study shows that the fitted model outperforms the listing models. The stock index, employed factors, and the volume technical indicator are the important criteria in stock index forecasting of this study.

Several other studies on ARIMA and forecasting are Gupta *et.al* (2021), Pillay (2020) and Sun (2020). Based on representative past studies, ARIMA model is selected as the forecasting method in this research as this model is better in predicting for short-term period.

Based on representative past studies, ARIMA model is selected as the forecasting method in this research as this model is better in predicting for short-term period

3. Methodology

The main purpose of time series modelling is to study the past observations of a time series so as to generate an appropriate model describing the inherent structure of the series. In this particular study, time series statistical technique is applied to obtain the most adequate model for forecasting purpose of KLCI. All the analysis of time series will be conducted with the aid of SPSS, EViews and Minitab software.

In 1970, Box and Jenkins introduced the ARIMA model. It is one of the statistical methods implanted for analyses and forecast the time series data. The ARIMA model is a composition of autoregressive model (AR), integration (I), and moving average model (MA). A combination of the autoregressive and moving average model together may produce a better stimulation on the series and produce a more accurate forecasting result. It considers the parameters of the both models. The model developed is known as ARMA model and defined as:

$$\begin{gathered} Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots \phi_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \theta_q \varepsilon_{t-q} \\ (1 + \phi_1 B + \phi_2 B^2 + \cdots \phi_p B^p) Y_t = \left(1 - \theta_1 B - \theta_2 B^2 - \theta_q B^q\right) \varepsilon_t \\ \phi_p(B) Y_t = \theta_q(B)\, \varepsilon_t \end{gathered} \tag{1}$$

However, for the case of non-stationary ARMA model, Box and Jenkins proposed the Autoregressive Integrated Moving Average, ARIMA (p, d, q) model to counter the time correlated modelling. The term (I) is integration referring as the differencing procedure with the notation *d* as the degree of differencing. It is introduced for a non-stationary ARMA model to solve the stationary problem of the series. The ARIMA model can be defined as:

$$\phi_p(B)(1-B)^d Y_t = \theta_q(B)\, \varepsilon_t \tag{2}$$

Basically, there are four major steps involved in Box and Jenkins method which are model identification, parameter estimation, model validation with diagnostic checking and lastly forecasting.
Model identification involves logarithm transformation, Autocorrelation Function (ACF), Partial Autocorrelation Function (PACF), differencing and model selection which is based on the characteristic of the ACF and PACF. Estimating the parameters for the Box-Jenkins models is quite complicated, therefore, high quality software program that fits Box-Jenkins models is employed for the parameter estimation. Main approaches to fitting Box-Jenkins models are non-linear least squares and maximum likelihood estimation. Model validation with diagnostic checking phase aimed to identify whether the estimated model is statistically adequate. The diagnostic checking is implemented based residuals, which are defined by

$$Residual = Observation - Fitted\ value \tag{3}$$

Diagnostic checking involves residual analysis, ACF and PACF of the residuals, Breusch-Godfrey Lagrange-Multiplier test, heteroskedasticity test, GARCH (Generalized ARCH) model, Exponential GARCH (EGARCH) model, GARCH-in-mean (GARCH-m) model, overfitting of the model and Bayesian Information Criterion (BIC).

Residual analysis is the first step of diagnostic checking. Once the model is found and fitted to the time series, it is advisable to check the adequacy of the model by carry out residual analysis. The residuals of a good model should satisfy the white noise assumptions of the error term, such that the error term is a sequence of uncorrelated random variables and identically distributed with zero mean and a finite variance.

To fulfil the white noise assumption of the error term, the ACF and PACF of the residuals for the potential model needs to be 95% statistically insignificant, especially for the values at small lags. In other words, both correlogram of the residuals should have no pattern and falls inside the Bartlett confidence interval. The insignificance values of the residuals indicate that the potential model did successfully capture the autocorrelation among the observations. If there exist significant values at the low lags, the presence of autocorrelation among the observations is detected and thus the model needs to be improved.

Breusch-Godfrey Lagrange-Multiplier (LM) test is an alternative test for autocorrelation of the residuals for testing serial correlation. This test implemented with the null hypothesis of no autocorrelation exist in the residuals against the alternative hypothesis of autocorrelation present in the residuals by running auxiliary regression. The regression of the estimated residuals is carried out on the selected variables of the model with the residuals of lag up to order *p*. The LM test statistic is asymptotically distributed as $\chi^2(p)$ with *p* degree of freedom.

Heteroskedasticity occurs when the variance of the error terms differs across the observations. In the standard Box and Jenkins model, the error term, ε_t is assumed to be identically distributed. However, most of the real-time data shows the presence of heteroskedasticity. Hence, heteroskedasticity test should be conducted to test for the specifications of heteroskedasticity in the residuals of the model.

ARCH-LM test is one of the useful tests to detect the presence of autoregressive conditional heteroskedasticity (ARCH) effect in the residuals. Ignoring the ARCH effect may result in loss of efficiency. This test is computed from an auxiliary regression conducted on the square of residuals with null hypothesis of there is no ARCH effect up to order *k* in the residuals such that

$$\hat{\varepsilon}_t^2 = \alpha_0 + \alpha_1\hat{\varepsilon}_{t-1}^2 + \alpha_2\hat{\varepsilon}_{t-2}^2 + \cdots + \alpha_k\hat{\varepsilon}_{t-k}^2 + \upsilon_t \qquad (4)$$

where $\hat{\varepsilon}_t^2$ is the residuals. The LM test statistic is asymptotically distributed as $\chi^2(p)$ where p is the number of degrees of freedom. In addition, the presence of ARCH effect in the residuals could also be detected using the correlograms of the squared residuals. The autocorrelation of the squared residuals should be zero at all lags if no ARCH found in the residuals.

GARCH model was specially developed to model and forecast conditional variance. It is an improvement based on the ARCH model by including a smoothing-averaging term to generate a more parsimonious specification. The ARCH process could be generalized to a process which the variance depends on an infinite number of lags of $\hat{\varepsilon}_{t-j}^2$ where

$$\sigma_t^2 = \lambda_0 + \sum_{i=1}^{\infty}\lambda_i\varepsilon_{t-i}^2 = \lambda_0 + \lambda(B)\varepsilon_{t-i}^2 \qquad (5)$$

$\lambda(B)$ is chosen to be parameterized as the ratio of two finite-order polynomials:

$$\lambda(B) = \frac{\alpha_s(B)}{1 - \beta_r(B)} = \frac{\alpha_1 B + \alpha_2 B^2 + \cdots + \alpha_s B^s}{1 - \beta_1 B - \beta_2 B^2 - \cdots - \beta_r B^r} \qquad (6)$$

where the roots of $1 - \beta_r(B)$ = 0 is assumed to be lied outside a unit circle. The process is substituted and rearranging to form

$$\sigma_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \alpha_2\varepsilon_{t-2}^2 + \cdots + \alpha_s\varepsilon_{t-s}^2 + \beta_1\sigma_{t-1}^2 + \beta_2\sigma_{t-2}^2 + \cdots + \beta_r\sigma_{t-r}^2 \qquad (7)$$

This is the generalized autoregressive conditional heteroskedasticity process of order r and s, GARCH (r, s). The coefficient β_r represents the impact if previous conditional variance made on the current conditional variance. The sum of α_s and β_j measures the persistence of volatility. The higher the value of sum, the greater the level of volatility persistency.

EGACRH model is another form of GARCH model developed by Nelson in 1991.It could be defined as

$$log\sigma_t^2 = \omega + \beta \log(\sigma_{t-1}^2) + \alpha \left|\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right| + \tau \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \tag{8}$$

where $log\sigma_t^2$ is the log of the conditional variance. This suggested that the leverage effect is more on exponential compared to quadratic. This also implies that the conditional variance will always be remained as nonnegative. The presence of leverage effects can be tested by the hypothesis of $\tau < 0$. The impact is asymmetric if $\tau \neq 0$. The EGARCH model will gives different impact on volatility based on the type and size of certain news.

In real world, the return of a security may depend on its volatility. Capital Asset Pricing Model (CAPM) stated that the risk-averse agents will require compensation for taking a risky asset. In other words, investors will get higher return for holding higher risks. GARCH-m model is generated to model such a phenomenon. A simple example of GARCH (1, 1)-m model could be defined as

$$\begin{gathered} Y_t = \phi_1 Y_{t-1} + \lambda\sigma_\varepsilon^2 + \varepsilon_t \\ \sigma_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \beta_1\sigma_{t-1}^2 \end{gathered} \tag{9}$$

The CAPM is exist if $\lambda > 0$. A positive value of λ indicates that the return is related to its volatility positively.

Overfitting is the process of adding parameters in the previous model to form a new model. Overfitting of the model is conducted for two purposes which are to generate a new model for replacement of the inadequate tentative model found and to identify the tentative model found is the most adequate model.

Based on this case study, overfitting of the model is implemented to ensure the appropriate model is the best model could be found. Most of the cases have more than one model which is eligible to be the fitted model. Overfitting of the model takes place by comparing the adequacy of the original model and the overfitted model. If the overfitted model found to be more adequate, overfitting will continue until no further potential model acquired. Notice that when overfitting, the orders of the AR and MA components must not be increased simultaneously. A simple example of ARMA (1,1) is used to illustrate the procedure of overfitting of the model. ARMA (1,1) will be overfitted and form ARMA (2,1) and ARMA (1,2) to compare the adequacy of the model.

Generally, BIC is also known as Schwarz Information Criterion (SIC). The BIC idea is motivated by placing a uniform prior distribution on the number of parameters, and the model with the least BIC value is identified as the best model. Although there exist other criteria in model selection such as Akaike Information Criterion (AIC), BIC is still selected as the indicator of this study as BIC is more superior in finding model for large samples. Once the most adequate model is identified, then only it can be used to generate accurate forecasting.

Forecasting is a statistical approach to predict the future value after the model specification has been done. Dynamic forecast has been selected over the static forecast as it can performs multiple step forecasting based on the previous data. The process flow chart of Box and Jenkins modelling procedure is shown in figure 1:

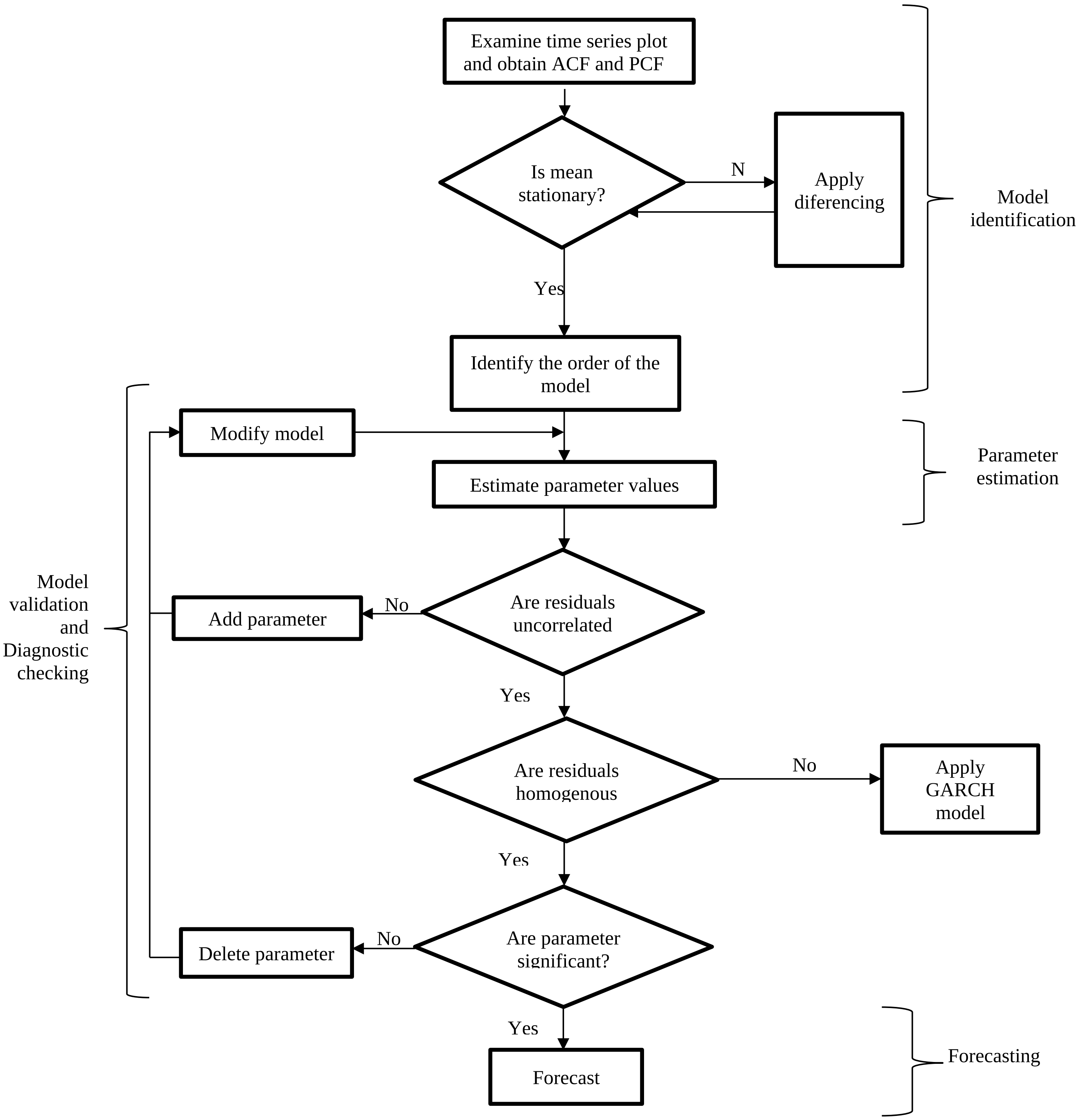

Figure 1: Flow chart of Box and Jenkins modelling procedure

4. Results and Analysis

In this study investment approach, stock market are selected to be investigated for their return of investment from 2019 to 2021. Monthly data selected ranges from January 1997 to December 2018. The data fitted to an appropriate model and then the future stock value is forecasted. For stock market indices, KLCI is selected as the indicator due to its high accuracy to represent the Malaysian stock market performance as it comprises of thirty largest companies from the main market in Bursa Malaysia.

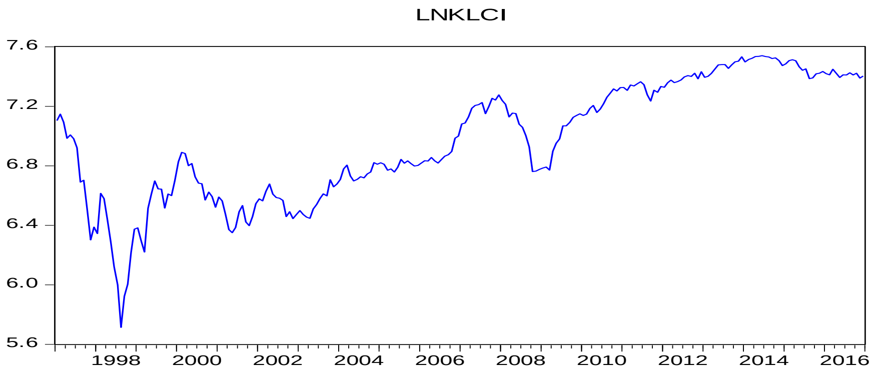

Figure 2: Time series plot of KLCI

Figure 2 shows the time series plot of monthly KLCI from January 1997 to December 2016. By inspection, there are decreasing trend starting from 1997, which is due to the Asian financial crisis. After that period, the series showed positive increment for approximately 18 years, suggesting that the series exhibit a long-term stochastic trend. Thus, the series is suspected to be non-stationary in mean. The Box and Jenkins model, is only applicable to a stationary series, therefore differencing method is employed to obtain a stationary series.

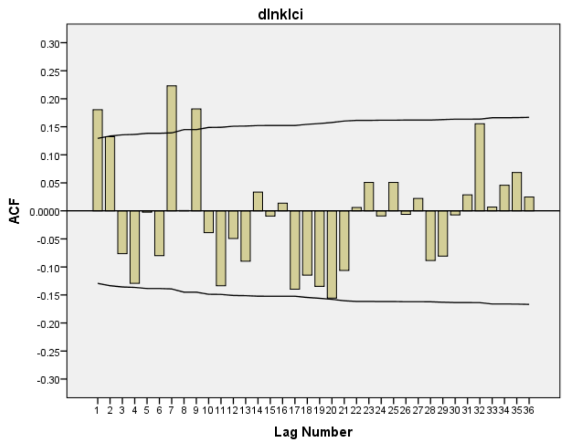

Figure 3: ACF of 1st differenced series for KLCI

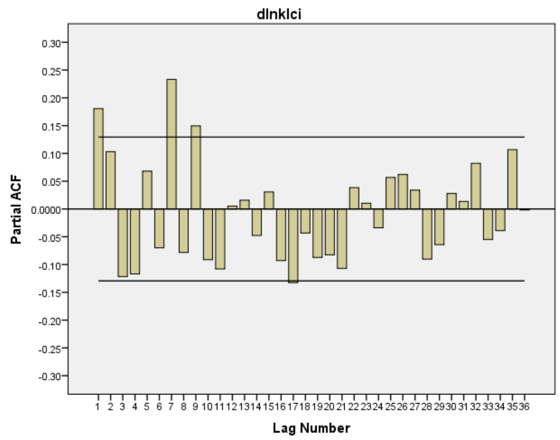

Figure 4: PACF of 1st differenced series for KLCI

Based on Figure 3 and 4, the model specification is determined based on the ACF and PACF of the differenced series. The first differencing seems has stabilized the mean and the characteristic of non-stationarity has been removed. The differenced series mean is now stationary. Although there exist a few vertical bars that are found to be significant at higher, the series could still be considered as stationary since the main purpose is to deal with the short-term forecasting. Three simplest time series models, which are AR (1), MA (1) and ARMA (1,1) are generated to determine the most appropriate model for the differenced series.

Table 1: Output of AR (1), MA (1) and ARMA (1,1)

Model	Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR (1)	C	0.001049	0.005192	0.202095	0.8400
	AR(1)	0.180471	0.063975	2.820977	0.0052
MA(1)	C	0.001288	0.004864	0.264852	0.7914
	MA(1)	0.143708	0.064293	2.235198	0.0263
ARMA(1,1)	C	0.001110	0.005495	0.202096	0.8400
	AR(1)	0.377022	0.309850	1.216791	0.2249
	MA(1)	-0.196744	0.327946	-0.599928	0.5491

Based on Table 1, AR (1) and MA (1) model indicates the stationary and invertibility condition of the models are fulfilled since both the AR and MA components of the model has p-value greater than 0.05. Hence, only AR (1) and MA (1) model are considered as the candidate model.

For the next step, diagnostic checking is implemented to validate the adequacy of the models. To verify the independence of the error term, serial correlation LM test is conducted to check for the presence of autocorrelation and the results showed that serial correlation still exist in the series. Hence, the models failed to achieve the independence of the white noise assumption.

Therefore, overfitting is implemented to produce a better model based on AR (1), MA (1), and ARMA (1,1) model. Overfitting will produce four new models which are AR (2), MA (2), ARMA (1,2), and ARMA (2,1) model as shown in Table 2. Based on Table 2 only MA (2) model has significant value which means further examination and diagnostic checking is again performed to MA (2) to verify this tentative model is statistically adequate for the data.

Table 2: Output of AR (2), MA (2), ARMA (1,2) and ARMA (2,1)

Model	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	0.164600	0.064772	2.541217	0.0117
AR(2)	0.102963	0.064747	1.590244	0.1131
MA(1)	0.180956	0.063558	2.847083	0.0048
MA(2)	0.197484	0.063683	3.101032	0.0022
AR(1)	-0.108832	0.289735	-0.375627	0.7075
MA(1)	0.285057	0.282829	1.007878	0.3145
MA(2)	0.226511	0.072226	3.136157	0.0019
AR(1)	-0.113776	0.488838	-0.232748	0.8162
AR(2)	0.166011	0.096472	1.720809	0.0866
MA(1)	0.278675	0.494120	0.563983	0.5733

Then, ARCH-LM test was done whereby implied that the heteroskedasticity effect exists within the residuals. Thus, GARCH (1,1) is proposed to model the volatility of the data. Table III shows the GARCH (1,1) model indicates the appropriateness of fitting the model.

Table 3: Output of MA (2)-GARCH (1,1)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
MA(1)	0.073722	0.071277	1.034295	0.3010
MA(2)	0.109619	0.065877	1.663991	0.0961
C	3.68E-05	2.19E-05	1.680978	0.0928
RESID(-1)^2	0.115841	0.035656	3.248895	0.0012
GARCH(-1)	0.861465	0.036752	23.43993	0.0000

To ensure that MA (2)-GARCH (1,1) is the most adequate model that could be found, overfitting was done again and another four new models are produced, which are MA (3)-GARCH (1,1), ARMA (1,2)-GARCH (1,1), MA (2)-GARCH (2,1), and MA (2)-GARCH (1,2).

Table 4: Summary table of overfitted model for MA (2)-GARCH (1,1)

Model	*p*th coefficient	*q*th coefficient	ARCH	GARCH
MA (3)-GARCH (1,1)	-	NS	**	**
ARMA (1,2)-GARCH (1,1)	NS	*	**	**
MA (2)-GARCH (2,1)	-	*	*	*
MA (2)-GARCH (1,2)	-	NS	*	**

*NS-Not significant, *- significant at 10% level, **significant at 5% level*

Table 5: Summary table of BIC for MA (2)-GARCH (1,1) and MA (2)-GARCH (2,1)

Model	BIC
MA (2)-GARCH (1,1)	-3.1625
MA (2)-GARCH (2,1)	-3.1462

Referring to the Table 4, MA (2)-GARCH (2,1) may also be a tentative model for the data since all the coefficients are significance at 10% significance level. In such case, model with least BIC in Table 5 will be chosen as best adequate model acquired for the data which are MA (2)-GARCH (1,1).

The characteristic of the stock index has made it become hard to predict, which leads to higher risks with greater uncertainties. In such circumstances, the leverage effect on the series, which means the negative correlation of the asset's volatility on the asset's returns has become an interesting topic to be investigated. Thus, MA (2) model is fitted with the EGARCH (1,1) model and the results are shown Table 6.

Based on the result the coefficient of MA (2) component has become significant suggesting that MA (2) model is compatible with both GARCH and EGARCH model. On the other hand, EGARCH model is found to be more vigorous for description of the data since its BIC value is much smaller compared to the GARCH model. This may be explained by the difference of violation of assumption for these two models. In real world, the volatility of the stock index is expected to have asymmetric effect on both good and bad news. GARCH model assumes the impact of unexpected increase and decrease in stock index have symmetric effect whereas EGARCH model captures the phenomena by assuming the impacts on the volatility asymmetrically. The coefficient of (C5), is significant at 5% significance level, indicating the presence of the asymmetric effect. Besides, the negative sign of the coefficient represents the existence of leverage effect on the series. This shows that KLCI will tends to have a bigger impact on the volatility in same magnitude.

Table 6: Results of MA (2) model with EGACRH (1,1)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
MA(1)	0.059264	0.072443	0.818073	0.4133
MA(2)	0.130362	0.064164	2.031686	0.0422
C(3)	-0.329664	0.106780	-3.087310	0.0020
C(4)	0.213895	0.064187	3.332377	0.0009
C(5)	-0.091701	0.031354	-2.924720	0.0034
C(6)	0.974383	0.013023	74.82141	0.0000
R-squared	0.039269	Mean dependent var	0.001254	R-squared
Adjusted R-squared	0.035216	S.D. dependent var	0.066510	Adjusted R-squared
S.E. of regression	0.065328	Akaike info criterion	-3.257190	S.E. of regression
Sum squared resid	1.011460	Schwarz criterion	-3.169915	Sum squared resid

The analysis is proceeded by investigating if the Capital Assets Pricing Model (CAPM) holds in the differenced series using the GARCH-m (1,1) model. From Table 7, the positive sign of the risk - return parameter, @SQRT(GARCH) indicates there is a positive relationship between the risk and return, which means higher risks will be compensated with higher return. However, since the coefficient of the parameter is not statistically significant at even 10% significance level, there is insufficient evidence to conclude there is a significant impact of volatility on the return.

Table 7: Results of MA (2) with GARCH-m

Variable	Coefficient	Std. Error	z-Statistic	Prob.
@SQRT(GARCH)	0.129456	0.083478	1.550788	0.1210
MA(1)	0.063849	0.074816	0.853423	0.3934
MA(2)	0.098443	0.068242	1.442565	0.1491
C	3.96E-05	2.24E-05	1.771641	0.0765
RESID(-1)^2	0.118975	0.035675	3.334930	0.0009
GARCH(-1)	0.856560	0.036957	23.17733	0.0000
R-squared	0.030238	Mean dependent var	0.001254	R-squared
Adjusted R-squared	0.022020	S.D. dependent var	0.066510	Adjusted R-squared
S.E. of regression	0.065773	Akaike info criterion	-3.238461	S.E. of regression
Sum squared resid	1.020968	Schwarz criterion	-3.151186	Sum squared resid

Hence ARIMA (0,1,2)-GARCH (1,1) fulfilled the white noise assumption and selected as the most adequate model for the series, it is applied for the forecasting purpose. For in-sample forecast, the predicted value of monthly KLCI from January 2017 to December 2018 is calculated and then are compared with the actual value to determine the accuracy of forecasted model.

Table 8 shows the predicted value increase slightly in the first two months and remain constant until December 2018. This situation are illogical since the stock market is always changing rapidly. However, the actual value still falls inside the predicted interval, indicating that this model still considered as reliable for forecasting.

Table 8: Comparison of actual value and predicted value of monthly KLCI from January 2017 to December 2018 with ARIMA (0,1,2)- GARCH (1,1)

Date	Actual value	Predicted value	Lower prediction interval	Upper prediction interval
Jan 2017	1671.54	1637.870	1552.828	1722.912
Feb 2017	1693.77	1640.546	1514.992	1766.099
Mar 2017	1740.09	1640.546	1478.250	1802.841
Apr 2017	1768.06	1640.546	1447.591	1833.500
May 2017	1765.87	1640.546	1420.474	1860.617
Jun 2017	1763.67	1640.546	1395.735	1885.357
Jul 2017	1760.03	1640.546	1372.724	1908.366
Aug 2017	1773.16	1640.546	1351.044	1930.047
Sep 2017	1755.58	1640.546	1330.424	1950.667
Oct 2017	1747.92	1640.546	1310.675	1970.416
Nov 2017	1717.86	1640.546	1291.658	1989.433
Dec 2017	1796.81	1640.546	1273.269	2007.823
Jan 2018	1868.58	1640.546	1255.423	2025.668
Feb 2018	1856.20	1640.546	1238.058	2043.034
Mar 2018	1863.46	1640.546	1221.119	2059.972
Apr 2018	1870.37	1640.546	1204.563	2076.528
May 2018	1740.62	1640.546	1188.355	2092.736
Jun 2018	1691.50	1640.546	1172.463	2108.628
Jul 2018	1784.25	1640.546	1156.863	2124.228
Aug 2018	1819.66	1640.546	1141.531	2139.560
Sep 2018	1793.15	1640.546	1126.449	2154.642
Oct 2018	1709.27	1640.546	1111.560	2169.492
Nov 2018	1679.86	1640.546	1096.968	2184.123
Dec 2018	1690.58	1640.546	1082.542	2198.549

Table 9: Values of investments in KLCI stock market

KLCI Stock Market	
Start: Jan 2019	End: Dec 2021
1690.609	1692.046

Table 10: Return of Investment per Segment

Segment of Period	Rate of Return
1997-2000	0.559
2001-2003	1.091
2004-2006	1.339
2007-2009	1.070
2010-2012	1.341
2013-2015	1.040
2016-2018	1.014

Table 9 showed the starting and end values of investments in KLCI stock market in the interval of three years are shown. The forecasted KLCI values for the coming three years only increase for one months, and subsequently remain constant value at index value 1692.046. The difference in value of only $1692.046 - 1690.609 = 1.437$ indicates there is little increase in stock market index for the coming three years, which is not realistic as stock market is known to be highly volatile and would not fix at one value.

Since the predicted value of KLCI showed increasing for one months and remain the same until December 2021 which is quite uncommon since the stock market value is frequently fluctuating, thus an alternative method for forecasting of KLCI is proposed in Table VII to make a better predicted value of KLCI.

Table 10 shows the investment on KLCI where investors had the highest loss from the period of 1997 to 2000 which is 44.1%. The highest rate of return per segment could goes up to 34.1 % whereas the lowest positive rate of return is 1.4 %. These phenomena indicating the stock market is highly volatile as the difference of profit is significantly higher or lower and furthermore investors might suffer huge loss when financial crisis occur.

5. Discussion and Conclusion

Forecasting the value of KLCI stock market and investigating whether stock market is the best in terms of profit for three years to come are the main objectives of this study.

The data is from January 1997 to December 2018. Note that the occurrence of financial crises is within the time period of study. The behaviour of the stock market in this time period with certain financial and economic indicators were identified.

Based on the findings, an appropriate time series model is determined for the KLCI series. The series is not stationary as it contains stochastic trend, but first differencing is sufficient to transform the series into a stationary series. In the end of the model selection, ARIMA (0,1,2) – GARCH (1,1) model was the most adequate model obtained for the stock index. To validate the appropriateness of this model, in-sample forecasting was carried out by comparing the predicted value and actual value throughout the period of the beginning of 2017 to December 2018. However, due to forecast properties of ARIMA (0,1,2)-GARCH (1,1), where the MA component is significant, the forecasted values converge to a constant value. The results shown looks illogical but since the actual value falls inside the predicted interval, this model is still reliable in its forecast ability.

The return rate of stock market is simply the ratio of unit price during selling over the unit price during buying. However, the forecast values of KLCI series remain at a constant value suggest little to no increment for the future three years, also indicates there will be no profit in investing in the stock market. This is not realistic as stock market contains high volatility and the values should be fluctuating over time. Therefore, an alternative way is proposed such that the historical values of KLCI is segmented into several 3-years period and the return of investment is calculated. This aim is to provide a general picture to the investors about the return for investing in stock market for the next three years. The alternative method shows findings that the return rate of stock market has a highest loss of 44.1 % and a highest gain of 34.1%, which clearly express the risk of investing in stock market. Therefore, for an investor willing to take high risk, investment on KLCI stock market will be the suitable choice as its return might goes several times higher as well as lower.

As an alternative for the risk averse, fixed deposit is an option of investment. A Fixed Deposit (FD) is a special type of bank savings account where a higher rate of interest is earned provided the deposit, a fixed amount, is not withdrawn over a fixed period. Typical periods are one month, three months, six months and a year. The interest is paid by the bank at the end of the stipulated period. Fixed deposit are popular in Malaysia because it is very safe and can earn better returns than an ordinary savings account. Fixed deposit rates are usually reference to a certain rates determined by Bank Negara Malaysia.

In conclusion, time series analysis can be used although they have their limitations, in investigating the properties of the series and provide useful informations for investors to make prediction. It can also be a reference for investors for their investment plan.

Acknowledgment

The research was financially supported by the Fundamental Research Grant Scheme (FRGS) offered by Ministry of Education Malaysia (203.PMATHS.6711666) and the fourth author gratefully acknowledges support from USM research university grant (1001.PMATHS.8011101).

References

Alkhazaleh, M. and Hussein, A. A. (2015). Forecasting Insurance Sector Volatility in Amman Stock Exchange Using ARIMA Model. *Arab Journal of Administration,* 35(1).

Alkhazaleh, M. (2018). Forecasting Banking Volatility in Amman Stock Exchange by Using ARIMA Model. *British Journal of Management,* 29(3).

Arshad, M. N. and Yahya, M. H. (2016). Relationship between stock market returns and exchange rates in emerging stock markets. *Journal of Islamic Economics and Business,* 1(2):131-143.

Ayodele, A., Aderemi, A., and Charles, A. (2014). Stock price prediction using the ARIMA model. *UKSim-AMSS 16th International Conference on Computer Modelling and Simulation*, Cambridge, United Kingdom.

Box, G. and Jenkins, G. (1970). Time Series Analysis: Forecasting and Control. *Holden-Day*, San Francisco.

Chu, H. H., Chen, T. L., Cheng, C. H. and Huang, C. C. (2009). Fuzzy dual-factor time-series for stock index forecasting. *Expert Systems with Applications,* 36(1):165-171.

Goh, S. K. and Lim, M. H. (2010). The impact of the global financial crisis: the case of Malaysia. *Third World Network*, 1-35. Penang.

Gupta, B. K., Mallick, M. K. and Hota, S. (2021). Survey on stock price forecasting using regression analysis. *Proceedings of Intelligent and Cloud Computing*, p.153. Conference Paper.

Pillay, S. (2020). Determining the Optimal Arima Model for Forecasting the Share Price Index of the Johannesburg Stock Exchange. *Journal of Management Information and Decision Sciences*, 23(5).

Sun, Z (2020). Comparison of Trend Forecast Using ARIMA and ETS Models for S&P500 Close Price. *4th International Conference on E-Business and Internet, ICEBI, Singapore*. Conference Paper.

Zhang, J. and Li, S. (2016). Financial Time Series Analysis Model for Stock Index Forecasting. *International Journal of Simulation Systems, Science & Technology*, 17(16).

2021 ICMS

INTERNATIONAL CONFERENCE ON COMPUTING, MATHEMATICS AND STATISTICS

e ISBN 978-967-2948-12-4

9 7 8 9 6 7 2 9 4 8 1 2 4