UNIVERSITI TEKNOLOGI MARA
Cawangan Kedah
Kampus Sungai Petani

2021 ICMS
INTERNATIONAL CONFERENCE ON COMPUTING, MATHEMATICS AND STATISTICS

Investment in Malaysia: Forecasting stock market using time series analysis

Nuzlinda Abdul Rahman, Chen Yi Kit, Kevin Pang, Fauhatuz Zahroh Shaik Abdullah and Nur Sofiah Izani

FAUHATUZ ZAHROH SHAIK ABDULLAH

Universiti Sains Malaysia

4 - 5 AUGUST 2021

CONTENTS

Introduction

- Stock market investment is an important and popular form of investment all across the world.

- 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.

Introduction

- For the main board of Bursa Malaysia, a commonly used index is the Kuala Lumpur Composite Index (KLCI).

- This index comprises of 30 largest companies on the Bursa's main board based on market capitalization.

- The daily KLCI index will rise or fall based on the weighted daily price performance of these 30 largest stocks.

Introduction

- Fortunes and lost have been made in stock markets all over the world.

- Therefore, 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.

Introduction

- 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 & Ayodeli et.al 2014).

Objectives

- To identify the most appropriate time series model for the KLCI
- To forecast the value of KLCI stock market for the coming three years
- To investigate whether investment in stock market is the best in terms of profit for the coming three years

Literature Review

AUTHOR, YEAR	EXPLAINATION
J. Zhang and S. Li (2016)	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. The result shows the Shanghai Composite Index have a small rise in the last two months of 2016.
A. Ayodele, A. Aderemi and A. Charles (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.

Literature Review

AUTHOR, YEAR	EXPLAINATION
Alkhazaleh and Hussein (2015)	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.

Methodology

- The model developed is known as ARMA model and defined as:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots \phi_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \theta_q \varepsilon_{t-q}$$

$$(1 + \phi_1 B + \phi_2 B^2 + \cdots \phi_p B^p) Y_t = \left(1 - \theta_1 B - \theta_2 B^2 - \theta_q B^q\right) \varepsilon_t$$

$$\phi_p(B) Y_t = \theta_q(B)\, \varepsilon_t$$

Methodology

- However, for the case of non-stationary ARMA model, Box and Jenkins proposed the Autoregressive Integrated Moving Average.

- The term (I) is integration referring as the differencing procedure with the notation *d* as the degree of differencing. The ARIMA model can be defined as:

$$\phi_p(B)(1-B)^d Y_t = \theta_q(B)\,\varepsilon_t$$

Methodology

- Four major steps involved in Box and Jenkins method:
  - -model identification
  - -parameter estimation
  - -model validation with diagnostic checking
  - -forecasting

Methodology

- Model identification involves:
  - -Logarithm transformation
  - -Autocorrelation Function (ACF)
  - -Partial Autocorrelation Function (PACF)
  - -Differencing
  - -Model selection based on the characteristic of the ACF and PACF.

Methodology

- Parameter estimation 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.

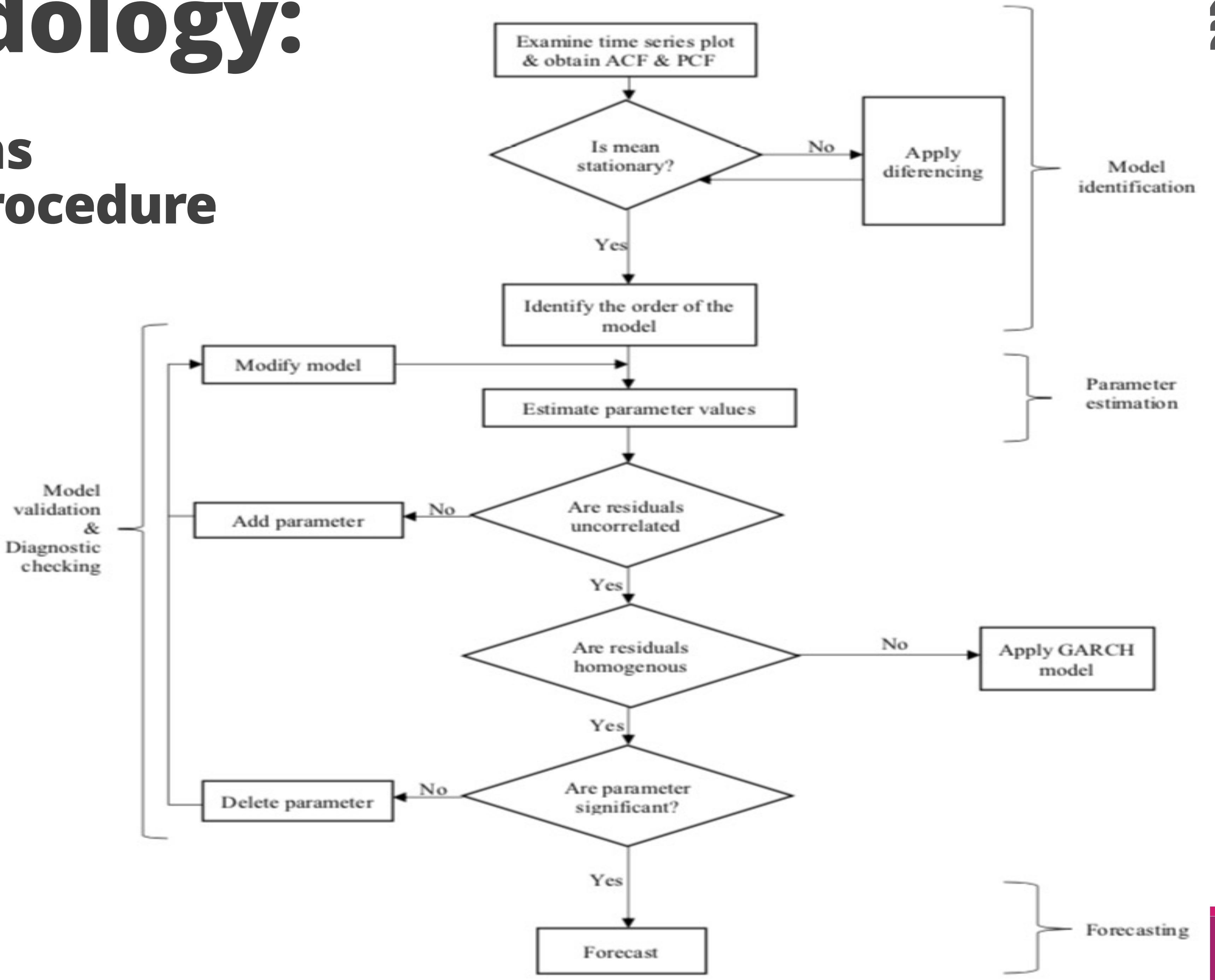
Methodology

- 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
  - -Bayesian Information Criterion (BIC)

- Once the most adequate model is identified, forecasting can be generated

Methodology:

Flowchart
Box & Jenkins
modelling procedure



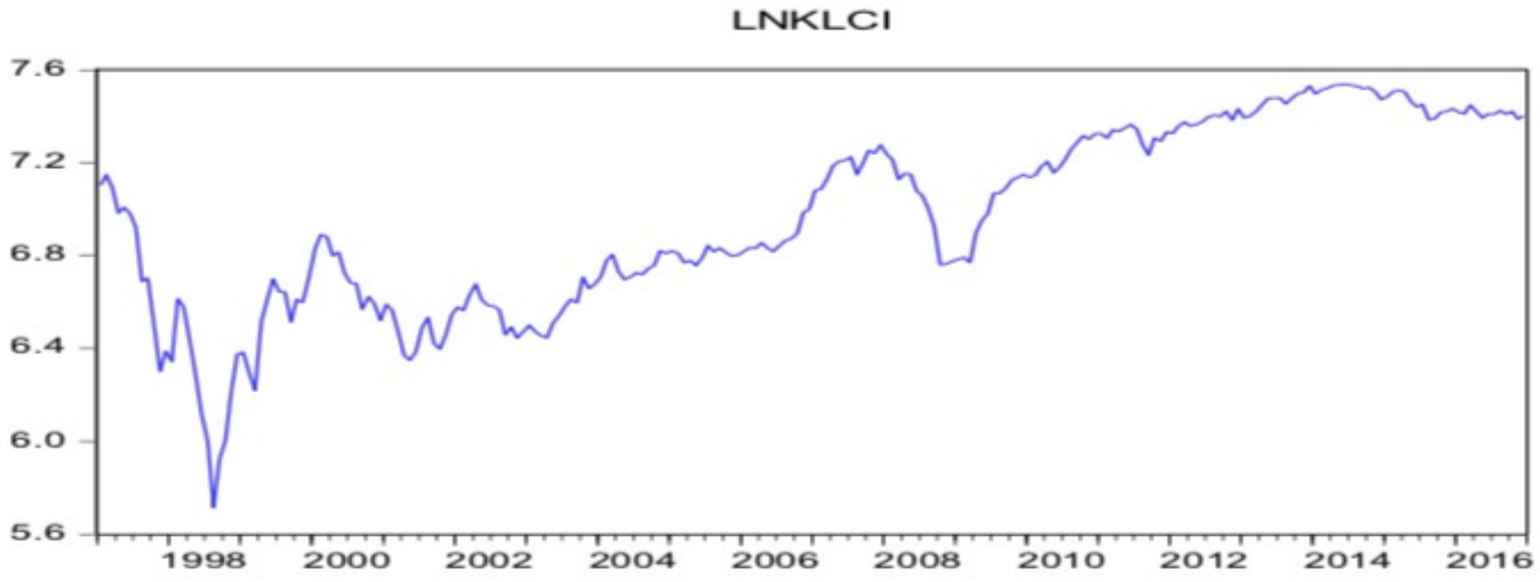
Results and Discussions

- Monthly data Kuala Lumpur Composite Index (KLCI) selected ranges from January 1997 to December 2018 .

- Data fitted to an appropriate model and then the future KLCI return of investment from 2019 to 2021 is forecasted.

- 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.

Results and Discussions

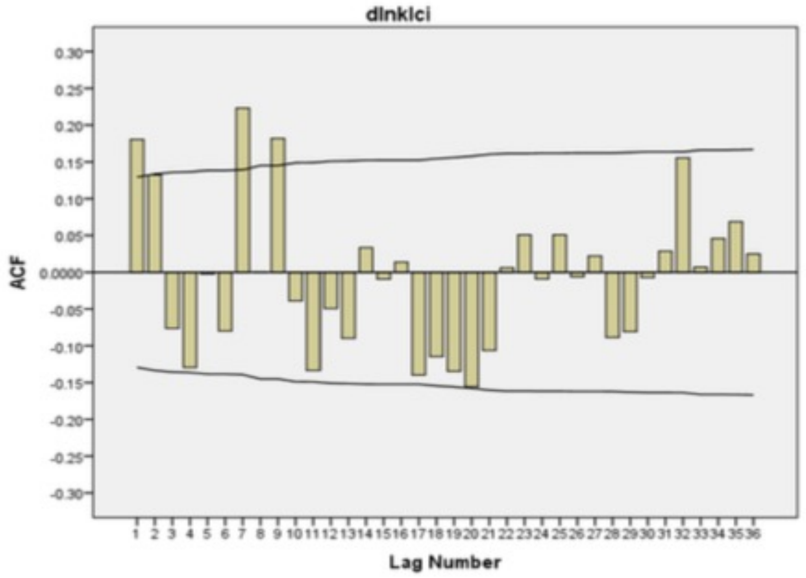


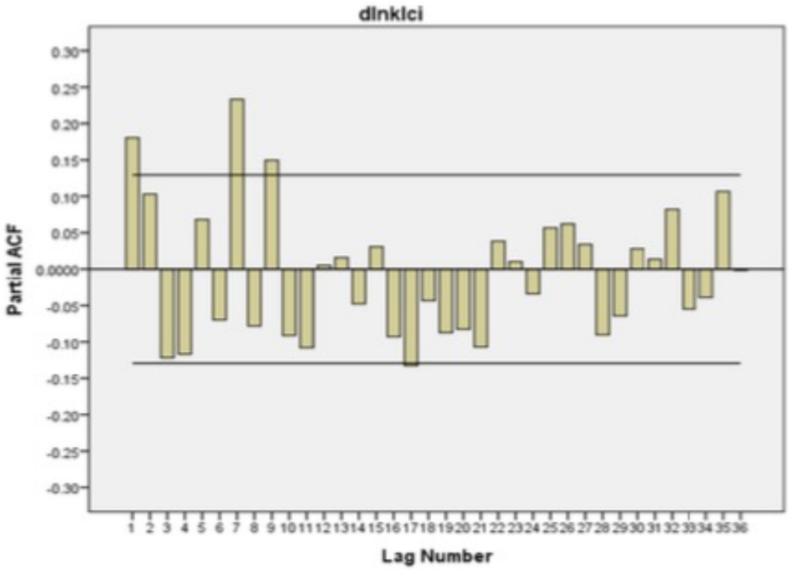
- Figure above shows time series plot of monthly KLCI from January 1997 to December 2016.
- Decreasing trend from 1997, which is due to the Asian financial crisis.
- Then, positive increment for approximately 18 years.

Results and Discussions

- The series is suspected to be non-stationary in mean.

- Therefore differencing method is employed to obtain a stationary series.





Results and Discussions

- The model specification is determined based on the ACF and PACF of the differenced series.

- The differenced series mean is now stationary.

- The application of GARCH model is done due to the time varying variance and heteroskedasticity was suspected to exist in the differenced.

- Three simplest time series models, AR (1), MA (1) and ARMA (1,1) are generated to determine the most appropriate model for the differenced series.

Results and Discussions

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

- The coefficients of AR (1) and MA (1) are significant at 5% significance level whereas the coefficients of ARMA (1,1) are not significant.

- Only AR (1) and MA (1) model are considered as the candidate model.

Results and Discussions

- Next, 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.

- Based on the results generated, the white noise assumptions are not fulfilled depicting the serial correlation still exist in the series.

- Overfitting is implemented to produce a better model which are AR (2), MA (2), ARMA (1,2), and ARMA (2,1) model.

Results and Discussions

- Based on the overfitted models, only MA (2) has significant value of coefficients.
- Therefore, all models except MA (2) are discharged from further examination and diagnostic checking is performed once again.
- 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.
- Then 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).

Results and Discussions

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	*	**

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

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

- Model with least BIC will be chosen as best adequate model acquired for the data which are MA (2)-GARCH (1,1).

Results and Discussions

- Characteristic of the stock index has made it become hard to predict whereby the higher the risks the greater the uncertainties of leverage effect on the series.

- Thus, MA (2) model is fitted with the EGARCH (1,1) model.

- EGARCH model captures the phenomena by assuming the impacts on the volatility asymmetrically.

Results and Discussions

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
Variance Equation				
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
Adjusted R-squared	0.035216	S.D. dependent var	0.066510
S.E. of regression	0.065328	Akaike info criterion	-3.257190
Sum squared resid	1.011460	Schwarz criterion	-3.169915

- MA (2) component has become significant at 5% significance level, suggesting that MA (2) model is compatible with both GARCH and EGARCH model.
- 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.

Results and Discussions

➢ 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.

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
Variance Equation				
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
Adjusted R-squared	0.022020	S.D. dependent var	0.066510
S.E. of regression	0.065773	Akaike info criterion	-3.238461
Sum squared resid	1.020968	Schwarz criterion	-3.151186

Results and Discussions

- The positive sign of the risk -return parameter, @SQRT(GARCH) indicates there is a positive relationship between the risk and return.

- This 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.

Results and Discussions

- Since 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.
- 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.
- 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.

Results and Discussions

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 beside shows comparison of actual value and predicted value of monthly KLCI from January 2017 to December 2018 with ARIMA (0,1,2)- GARCH (1,1).

Results and Discussions

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

- Based on the table above, 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, close to no return for investment in stock market from 2019 to 2021, which is not realistic as stock market is known to be highly volatile and would not fix at one value.

Results and Discussions

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

- Thus an alternative method for forecasting of KLCI is proposed in Table above to make a better predicted value of KLCI.
- The table 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.

Conclusions

- 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.
- 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.

Conclusions

- 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.
- 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.

Conclusions

- 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.

Conclusions

- 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, for investigating the properties of the series and provide useful informations for investors to make prediction. It can also be used as a reference for investors for their investment plan.

THANK YOU