



ANALYSIS OF THE BURSA MALAYSIA STOCK PRICES AND FORECASTING

¹Izzad Irfan Anis and ²Arifah Bahar

^{1,2}Department of Mathematical Sciences
Faculty of Science, Universiti Teknologi Malaysia,
81310 Johor Bahru, Johor, Malaysia.

e-mail: ¹izzad@graduate.utm.my, ²arifah@utm.my

Abstract This study aims to analyse the Bursa Malaysia KLCI stock market and generate a model to forecast future price of stocks. A dataset is collected from Yahoo! Finance from January 2019 to December 2020. The technical analysis done shows the performance of Bursa Malaysia KLCI drops in the beginning of 2020. An ARIMA model is then generated to analyse and forecast the future value. A structural break test is also run on the dataset to avoid any invalid conclusions and inaccurate forecasts. The analysis will then be done by using single exponential smoothing after the fourth structural break date and generate forecasts.

Keywords stock market forecasting; technical analysis; ARIMA; structural break; single exponential smoothing

1 Introduction

Malaysia stock market is a popular index that many people trade in around the world. As the case with all stock market investments investors will look to exploit and make a profit from buying low and selling high. To trade on Bursa Malaysia, investors must first open a Central Depository System (CDS) account with a stockbroking business or investment bank. A list of licenced stockbroking businesses may be found on the Bursa Malaysia website. The CDS is a financial instrument that is used to reflect security ownership and movement. It allows investors to buy and sell equities while also monitoring their performance. Purchased stocks will be credited to your CDS account, while sold securities will be subtracted. Aside from that, investors must open a trading account, which will be done simultaneously with the CDS account.

Bursa Malaysia operates on Monday to Friday, where the market opens at 9 a.m. and closes at 5 p.m.. Investing to stocks in Malaysia requires a minimum of 1 lot is, where 1 lot is equivalent to 100 shares. The 3 major markets for share trading in Bursa Malaysia are MAIN, ACE and LEAP where all markets are for different reasons. In some ways, they can be seen as 3 different levels of markets depending on the sizes and differences of the companies.

2 Overview of stock market analysis and forecasting

There are many statistical techniques and tools in predictive analytics to estimate and predict stock prices including linear regression, analysis of variation (ANOVA), and others. The focus of this article is on a real-world situation in the stock market. The stock market's highlight is the seasonal trend and flow. The ARIMA model and structural break analysis will be used to anticipate the results of a collection of data in this study. ARIMA, an algorithmic technique to convert the series, is better than predicting directly, according to the study, and it also produces more accurate findings [1]. Technical analysis is the evaluation of a company's underlying stock's price behaviour. It uses a variety of charts and statistical indicators to determine price, range, and trends by finding historically relevant price patterns and behaviours that may be used to anticipate the stock's probable direction. This technique solely considers the price of the stock, not the company's activities.

In this research, we will investigate the trend of the stock prices in Bursa Malaysia as well as forecasting the stock prices. By properly forecasting stock values, it will be able to make better-informed judgments about how much risk to take for how much possible profit. [2]. This is useful in predicting value of the stocks in the future for the knowledge of their investors. The information about the reliability of stock recommendations can help unsophisticated investors who are interested in absolute stock returns to improve their decision-making and useful for firms to optimize their investors' policy [3]. Forecasting stock prices can be done using various forecasting methods or tools. However, due to the absence of several important factors, some of the forecasting tools can trick user on visualizing the forecasted outcomes. In the area of stock prices forecasting, there are three categories of forecasting methods such as statistical methods, technical analysis and artificial analysis. By technical analysis, the presence of the human mind help investors to get in-depth and adequate information [4].

Some of the factors that influence the fluctuation movement of stock price of a company are the incoming news articles and updates of the respective company. This news ranges from the company's most recent earnings, management announcements, dividends announcements, editorial news and expert trader analysis, and local and international news on the company's present state [5]. Besides that, on a worldwide scale, given that oil and gas (O&G) and plantation are the two largest industries on the local bourse, price fluctuations of two types of oil, crude oil and crude palm oil, have a considerable influence on the local market. When both sectors are down, Bursa Malaysia will be less enticing to investors, particularly foreign investors [6]. In light of recent events, the COVID-19 pandemic has also affected the global financial market, which includes Malaysia. In addition, Ozili [7] The COVID-19 pandemic has had two effects on the stock market, according to the report. The pandemic first caused the closure of businesses and corporate operations, then had an impact on the financial market. Second, investors' investment decisions were impacted by the COVID-19 instances, resulting in increased stock market volatility [7]. Overall, each trader should consider these aspects before making a stock purchase choice. Technical analysis is the science of recording, usually in graphic form, the actual history of trading such as the price changes and volume of transactions in a certain stock and then deducing from that history to the probable future trend [8]. It uses tools like statistical analysis to assist traders and investors bridge the gap between intrinsic values, which is a measurement of an asset's worth based on a financial model, and market pricing. The top-down method and the bottom-up approach are two alternative approaches to approach technical analysis. The top-down method is a macroeconomic study that considers the entire economy before focusing on specific assets. Prior to investing in stocks, this strategy focuses first on economies, then on business sectors, and last on firms. The bottom-up method focuses on individual stocks, and it entails examining a fundamentally intriguing stock for prospective entry and exit opportunities. This strategy appeals to investors who are looking for a good deal and want to keep the stock for a long period [9].

3 Data and method of analysis for Bursa Malaysia KLCI

3.1 Data collection

In this study, the dataset for the KLCI was collected from the Yahoo! Finance website. The selected dataset of the daily close price of gold in the financial market from 1st January 2019 to 30th December 2020. This study will gather the data of 30 stocks of the highest market capitalization of companies publicly listed in Bursa Malaysia from the past 2 years from the site.

3.2 Autoregressive Moving Average analysis

ARIMA model is classified as ARIMA (p,d,q) where p indicates autoregressive of the data, d refers to the integer of the data and q as the moving average of the data set [10].

3.2.1 Equation of ARIMA Model

$$Y_t = C + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \dots + \theta_q \epsilon_{t-q}$$

Where ϵ_t is an uncorrelated innovation process where mean is zero,

C – Constant term

AR – Non-seasonal AR coefficients ($\phi_1 \dots \phi_p$)

MA – Non-seasonal MA coefficients ($\theta_1 \dots \theta_q$)

AR Lags – Lags corresponding to nonzero, non-seasonal AR coefficients

MA Lags – Lags corresponding to nonzero, non-seasonal MA coefficients

D - Degree of non-seasonal differencing, D [if D has value 0 meaning no non-seasonal integration]

Variance – Scalar variance of the innovation process (σ_ϵ^2)

Distribution - Distribution of the innovation process

3.2.2 Autocorrealation Function (ACF)

Using ACF gives big advantage of measuring the amount of linear dependence between observations in a time series that are separated by a lag, k [11].

$$r_k = \frac{\sum_{t=1}^{n-k} (y_t - \bar{y})(y_{t+k} - \bar{y})}{\sum_{t=1}^n (y_t - \bar{y})^2} \tag{3.1}$$

$$\text{The } t_{r_k} \text{ statistics is } t_{r_k} = \frac{r_k}{s_{r_k}} \text{ where } s_{r_k} = \sqrt{\frac{1 + 2 \sum_{j=1}^{k-1} r_j^2}{n}} \tag{3.2}$$

3.2.3 Partial Autocorrealation Function (PACF)

PACF is used to measure the degree of association between Y_t and Y_{t-k} , when the effects of other time lags ($1, 2, 3, \dots, k-1$) are removed. The sample PACF is given by [11];

$$r_{kk} = \frac{r_k - \sum_{j=1}^{k-1}(r_{k-1,j})(r_{k-j})}{1 - \sum_{j=1}^{k-1}(r_{k-1,j})(r_j)} \tag{3.3}$$

The $t_{r_{kk}}$ statistics is $t_{r_{kk}} = \frac{r_{kk}}{s_{r_{kk}}}$ where $s_{r_{kk}} = \sqrt{\frac{1}{n}}$ (3.4)

3.3 Structural break test

A structural break in econometrics and statistics refers to an unexpected change in the parameters of regression models over time, which can result in large forecasting errors and model unreliability in general [12]. When there is a rapid shift in the relationship being studied, a structural break can develop in time series data or cross sectional data. Non-stationary data is defined as having a unit root or a structural break, before and after which the data exhibits distinct patterns [13].

3.3.1 Bai & Perron Test

Bai and Perron test considers estimating multiple structural changes in a linear model. The results are obtained under a general framework of partial structural changes which allows a subset of the parameters not to change. Bai-Perron multiple structural breakpoint test is a method based on the internal determination of structural breaks in series rather than unit root test. The methodology considers the following multiple structural break model with m breaks (m+1 regimes).

$$y_t = x'_t \beta + z'_t \delta_1 + u_t, t = 1, \dots, T_1 \tag{3.5}$$

$$y_t = x'_t \beta + z'_t \delta_2 + u_t, t = T_1 + 1, \dots, T_2 \tag{3.6}$$

$$y_t = x'_t \beta + z'_t \delta_{m+1} + u_t, t = T_m + 1, \dots, T \tag{3.7}$$

Where y_t is the observed dependent variable, both x_t and z_t are vectors of covariates with dimensions $p \times 1$ and $q \times 1$ respectively. β and δ_j are the corresponding vectors of coefficients and u_t is the disturbance term at time t while the break points (T_1, \dots, T_m) are treated as unknown, and are estimated together with the unknown coefficients when T observations are available.

The objective is to estimate the unknown regression coefficients and break dates ($\beta, \delta_j, \dots, \delta_{m+1}, T_1, \dots, T_m$) when T observations on (y_t, x_t, z_t) are available.

3.4 Single exponential smoothing

One of the easiest ways to forecast a time series is to use a single exponential smoothing. The main premise of this model is that the future will be roughly similar to the recent past. As a final estimate of the level, the exponential smoothing model will anticipate future demand. An exponential smoothing model's fundamental principle is that the model will learn from the most recent observation and record the last forecast it made at each period. To represent this,

$$f_t = \alpha \delta_{t-1} + (1 - \alpha) f_{t-1} \tag{3.8}$$

$$0 < \alpha < 1$$

Where,

- α represents the ratio of the importance of the model will allocate to most recent observation.
- $\alpha\delta_{t-1}$ represents the previous demand observation multiply with the learning rate.
- $(1 - \alpha)f_{t-1}$ represents how much the model remembers from its previous forecast value.

To forecast future values, the last forecast is simply extrapolated into the future. If we define f_{t^*} as the last forecast that we could make based on demand history, we simply have

$$f_{t>t^*} = f_{t^*} \tag{3.9}$$

4 Analysis on time series and structural break test

This section presents the experimental results of Bursa Malaysia KLCI analysis, autoregressive integrated moving average (ARIMA) forecasting analysis, structural break test and simple exponential smoothing.

4.1 Bursa Malaysia KLCI analysis

This study considers the data set of daily closing price of KLCI from 1st January 2019 to 31st December 2020. These time series are chosen to analyse the performance of KLCI due to the uncertainty of the economy before and after the COVID-19 outbreak.

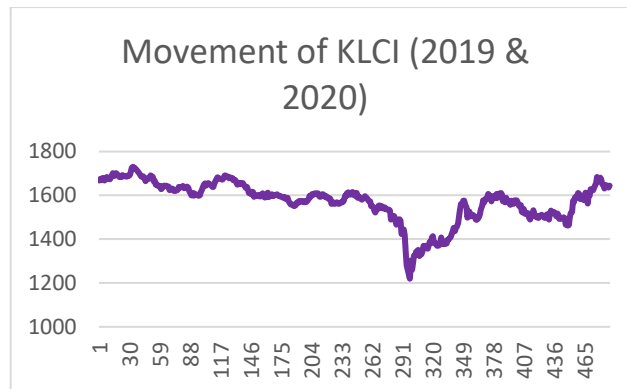


Figure 1: Movement of KLCI in 2019 and 2020

Figure above shows the movement of KLCI for the year 2019 and 2020 based on their close data. As we can see from the chart above, there exist an outlier data at certain points. To create a boundary outside of Q1 and Q3, we may utilise the Interquartile Range (IQR) method for detecting outliers. Outliers are any values that fall outside of this boundary. To construct this fence, multiply 1.5 by the IQR, then remove this value from Q1 and add it to Q3. This provides the lowest and maximum boundaries against which each observation is measured. Outliers are defined as observations that are more than (1.5 multiple IQR) below Q1 or more than (1.5 multiply IQR) above Q3. These numbers will be used to compute the limits for spotting outliers. The outliers does not represent the target of this study, hence we remove it. The computed values are shown in the output below.

Table 1: Values for interquartile range of the data

Variable	Value
1st Quartile	1523.785
3rd Quartile	1638.195
Inner Quartile Range (IQR)	114.41
Lower Bound	1352.17
Upper Bound	1809.81

From the values obtained in Table 1, we are able to determine the outliers from the data. The outliers are from 13th March 2020 to 6th April 2020 which dates back to the MCO period in Malaysia. The outliers does not represent the target of this study, hence we can remove the data points. Table 2 below shows the data points that we consider as an outlier.

4.2 Structural break test

In this part of the study, we will run a structural break test on the data from January 2019 to December 2020. Figure 4.11 below shows the result in which the dates obtain from the test are where the structural break occurs.

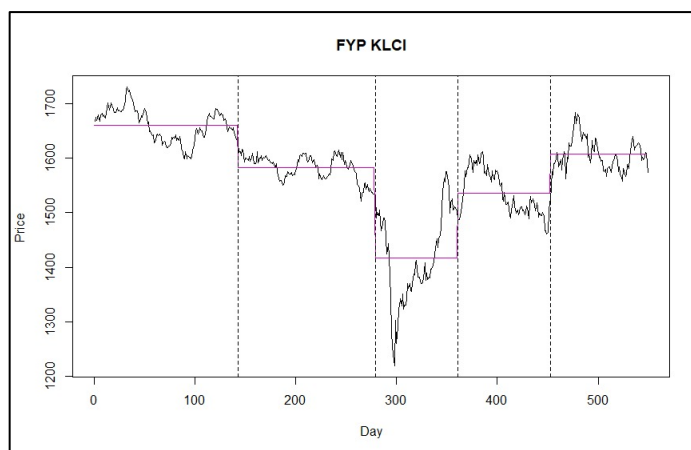


Figure 2: Structural break of the dataset

Table 3: Structural break dates of the dataset

Date	Open	High	Low	Close	Adj Close	Volume	Break
2/8/2019	1631.39	1631.42	1623.66	1626.76	1626.76	1.16E+08	n=143
21/2/2020	1529.01	1533.91	1527.57	1531.2	1531.2	1.07E+08	n=279
25/6/2020	1494.44	1498.61	1484.4	1489.2	1489.2	93460100	n=361
9/11/2020	1521.63	1533.05	1513.21	1524.32	1524.32	1.09E+08	n=453

Based on the Table 4.7, we are able to detect where does the structural break occurs in this time series based on the Bai & Perron test. Hence, since there exist structural breaks in our dataset, it indicates that dataset is non-stationary.

5 Forecasting using ARIMA and Single Exponential Smoothing

This section presents the result of forecasting by autoregressive integrated moving average (ARIMA) analysis and simple exponential smoothing.

5.1 Autoregressive integrated moving average (ARIMA) forecasting

In this part of the study, we will consider the data from 2019 and 2020, excluding the outliers to forecast the movement of KLCI, which amount to 472 observations. Firstly, we will check the data whether it is stationary or non-stationary.

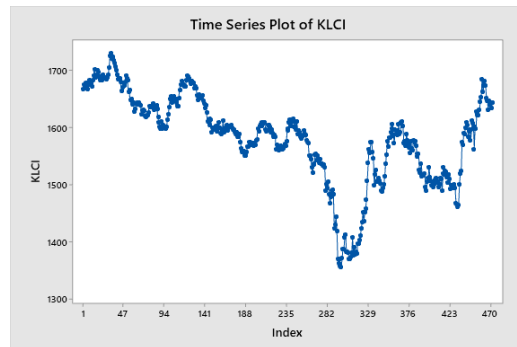
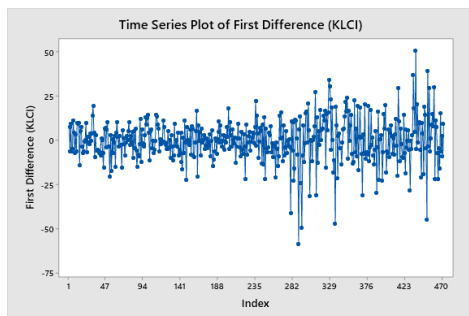
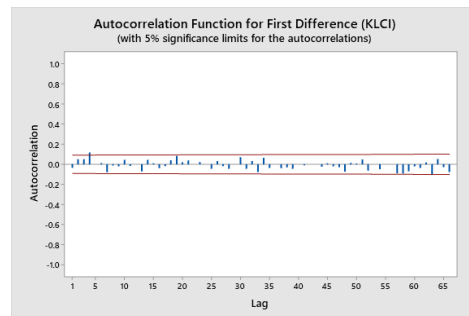


Figure 3 : Time series plot of KLCI

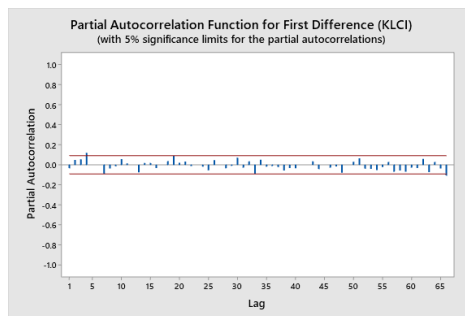
From Figure 2, we can see that the time series plot presented that it is non-stationary. Hence, we have to difference the data.



(i)



(ii)



(iii)

Figure (i) shows that the series is stationary after the first differencing which indicates that d is equal to 1. Next, we will find the PACF and ACF. Based on (iii), the lags for PACF is at 4

which indicates that $p=4$. For (ii), we can see that the lags are also at 4 which implies that $q=4$. Hence, we will have an ARIMA(4,1,4) model.

Table 2: Ljung Box Chi Square statistic

Lag	12	24	36	48
Chi-Square	1.63	10.91	21.62	28.24
DF	3	15	27	39
P-Value	0.654	0.759	0.757	0.899

To test the adequacy of ARIMA (4,1,4) model, we will use the Ljung-Box test. If the p -value for the lags (Table 3) are more than the critical value which is 0.05, then we say that this model is adequate. Based on Table 3, we obtained that the p-value for all of the lags are more than the critical value, hence this model is adequate and we can proceed to forecast.

APE	0.020994
MAPE	2.099355

The MAPE value obtained is 2.1% which implies that the percentage error of our model is very low

4.4 Single exponential smoothing

For this part of the study, we will forecast the movement of KLCI after the structural break by using simple exponential smoothing. Figure below shows the time series plot of the data after the fourth structural break which is on 9th November 2020.

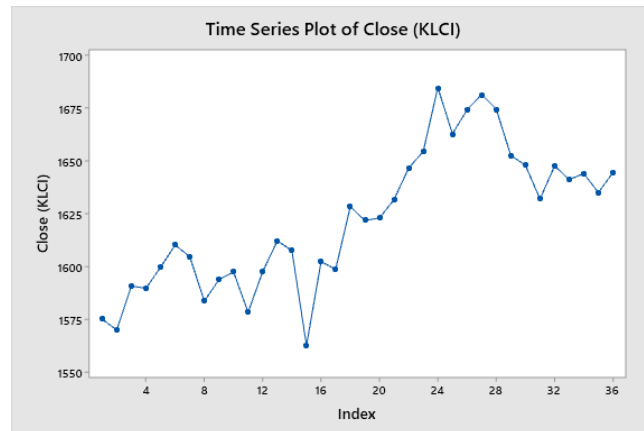


Figure 4: Time series plot of KLCI after fourth structural break

Based on Figure 5 below, we obtain the smoothing plot for single exponential method for KLCI index after the fourth structural break. Next, we will forecast the value of the KLCI index for the next 6 days. The forecasts from single exponential smoothing are very conservative because they are based solely on the latest value occurs, and no estimate of trend.

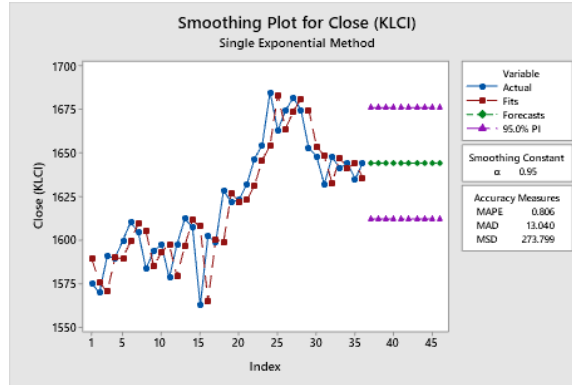


Figure 5: Exponential smoothing plot after fourth structural break

Table 4 shows the values of accuracy measures. Due to the forecasts from single exponential smoothing being constant, it is important that there is no trend in the data before the forecasts.

Table 4: Values of accuracy measures for the model

MAPE	0.806
MAD	13.040
MSE	273.799

5 Conclusion

The research on this paper includes the analysis of Bursa Malaysia index from 2019 and 2020 as well as technical analysis of the forecasting for Bursa Malaysia KLCI. This study has successfully met its objectives.

It was found that the movement of KLCI in 2020 was at a significantly lower rate than in 2019. Thus, implying that the outbreak of COVID-19 had truly impacted Malaysia’s economy sector where the index points never reached the heights of what was achieved in 2019. Besides that, the changes in components of the KLCI also indicates the emergence of healthcare and medical supplies companies in 2020 where the demand is higher during this pandemic, hence led to an increase in their company’s stock prices.

The experimental results obtained with best ARIMA model, ARIMA(4,1,4) have been shown to be capable of accurately predicting stock values on a short-term basis. This might assist stock market participants in making lucrative investment selections. Next, detecting structural break prevents from the misleading of forecasting model. The structural break that occurs in the stock market can be supported by the economic events and political issues that has been happening in Malaysia as well as across the globe. The analysis done on the KLCI can be supported by the short period forecasting using simple exponential smoothing which provide us with better accuracy measures.

Lastly, the structural break analysis is more suitable in forecasting financial time series due to the volatility occurring in the market. Identifying structural breaks in models can lead to a better understanding of the true mechanisms that drives changes to the data.

References

- [1] S. A. Wadi, M. T. Ismail and S. A. A. Karim, "Selecting Wavelet Transforms Model in Forecasting Financial Time Series Data Based on ARIMA Model," *Applied Mathematical Sciences*, Vol. 5, 2011.
- [2] R. D. Edwards and J. Magee, *Technical Analysis Of Stock Trends*, California: Snowball Publishing, 2011.
- [3] P. Santosuosso, "Can We Trust Financial Analysts? Reliability of Stock Recommendations and Firm-Specific Characteristics," *International Journal of Economics and Finance*, Vol. 7, 2015.
- [4] C. Cheng, C. Su, T. Chen and H. Chiang, *Forecasting Stock Market Based on Price Trend and Variation Pattern*, Berlin: Springer-Verlag Berlin Heidelberg, 2010.
- [5] N. Godbole, M. Srinivasaiah and S. Skiena, "Large-Scale Sentiment Analysis for News and Blogs," *ICWSM 2007 - International Conference on Weblogs and Social Media*, 2007.
- [6] K. Fong and J. T. Liew, "6 factors that moved the markets," 8 January 2020. [Online]. Available: <https://www.theedgemarkets.com/article/6-factors-moved-markets>.
- [7] P. K. Ozili, "COVID-19 Pandemic and Economic Crisis: The Nigerian Experience and Structural Causes," 2020.
- [8] R. D. Edwards, J. Magee and W. Bassetti, *Technical Analysis of Stock Trends*, New York: Taylor & Francis Group, 2019.
- [9] S. Seth, "Technical Analysis Strategies for Beginners," 25 September 2020. [Online]. Available: <https://www.investopedia.com/articles/active-trading/102914/technical-analysis-strategies-beginners.asp>.
- [10] B. U. Devi, D. Sundar and D. P. Alli, "An Effective Time Series Analysis for Stock Trend Prediction Using ARIMA Model for Nifty Midcap-50," *International Journal of Data Mining & Knowledge Management Process (IJDKP)* Vol.3, No.1, 2013.
- [11] S. Zakaria, N. Al-Ansari, S. Knutsson and T. Al-Badrany, "ARIMA Models for weekly rainfall in the semi-arid Sinjar District at Iraq," *Journal of Earth Sciences and Geotechnical Engineering*, vol. 2, no. 3, 2012.
- [12] J. Antoch, J. Hanousek, L. Horváth, M. Hušková and S. Wang, "Structural breaks in panel data: Large number of panels and short length time series," *Econometric Reviews*, vol. 38, no. 7, pp. 828-855, 2018.
- [13] H. B. Allaro, B. H. Kotu and K. Belay, "A time series analysis of structural break time in the macroeconomic variables," *African Journal of Agricultural Research*, 2011.
- [14] "World Gold Council," [Online]. Available: <https://www.gold.org/goldhub/data/gold-prices>. [Accessed 16 January 2021].

- [15] A. A. Adebisi and A. O. Adewumi, "Stock Price Prediction Using the ARIMA Model," *2014 UKSim-AMSS 16th International Conference on Computer Modelling and Simulation*, 2016.
- [16] R. J. Hyndman and G. Athanasopoulos, *Forecasting : Principles And Practise*, OTexts, 2018.