International Journal of Academic and Applied Research (IJAAR)
ISSN: 2643-9603
Vol. 5 Issue 12, December - 2021, Pages:36-41

Forecasting The Jakarta Composite Index (JCI) Value During The COVID-19 Pandemic Using Comparison of The ARIMA Model and The Nonparametric Regression Model With Fourier Series Estimator

Christopher Andreas

Department of Mathematics, Faculty of Science and Technology, Universitas Airlangga, Surabaya, Indonesia
christopher.andreas-2018@fst.unair.ac.id

Abstract: *The conditions of the Covid-19 pandemic had a significant impact on the economic sector. In Indonesia, the stock market is under great pressure, so that the value of the Jakarta Composite Index (JCI) fluctuates greatly. Various policies have been carried out by the government to maintain the stability of the JCI value. Forecasting the value of JCI, especially during the Covid-19 pandemic, is very important to do in order to evaluate applicable policies. For this reason, forecasting is carried out based on the ARIMA approach and nonparametric regression with the Fourier series estimator. ARIMA modeling shows insignificant parameters. Meanwhile, forecasting with the Fourier series shows satisfactory results. Forecasting on out-sample data shows a good MAPE-out-sample value of 9.34%. Meanwhile, the MAPE-in-sample is 1.57%. Thus, it is predicted that the JCI value will still fluctuate, but not significantly so that the existing policies can be maintained.*

Keywords—ARIMA, Forecasting, Fourier Series, Jakarta Composite Index (JCI), Nonparametric Regression**.**

1. INTRODUCTION

Economic stability is one of the keys to creating sustainable economic growth. This is the focus of the study in the Sustainable Development Goals (SDGs). However, the Coronavirus Disease 2019 (Covid-19) pandemic has created uncertainty that has disrupted economic stability. Furthermore, the Covid-19 pandemic has also had a serious impact on the economic and financial sectors which was marked by an economic recession in several countries, including Indonesia [1].

In the financial sector, Indonesia's capital market conditions experienced a number of pressures during the Covid-19 pandemic. The movement of the Jakarta Composite Index (JCI) value during 2020 was very volatile and several times reached its lowest value in the last few years. In fact, the government through the Financial Services Authority has tried to maintain the stability of the value of JCI by issuing some policies. These policies include imposing a trading halt, changing the auto-rejection limit, and reducing the BI 7-Days Repo Rate [2-4]. A number of these policies were taken to reduce pressure on the Indonesia stock market.

Forecasting is an important element that is always considered in making decisions or determining a policy [5]. Forecasting the value of JCI during the Covid-19 pandemic is very important as a basis for evaluating Indonesian stock market policies. The existence of JCI value forecasting makes it easier for policymakers to maintain or adjust current policies so that the stability of the Indonesian stock market can be maintained. For this reason, this study was conducted to predict the value of JCI during the Covid-19 pandemic.

In this case, forecasting is done using a time series analysis approach through the Autoregressive Integrated Moving Average (ARIMA) model. This method is widely used in modeling and forecasting data in various fields, especially the economic and financial sectors such as gold price forecasting, oil price forecasting, and economic growth forecasting [6-8]. The ARIMA model can explain the relationship between the value of observations in a certain period which is influenced by the value of the observations of the previous period. This is following the characteristics of the JCI value which tends to be influenced by the JCI value in several previous periods. However, the ARIMA model has several assumptions that must be met [9-10].

In this study, forecasting is also done using a nonparametric regression approach with a Fourier series estimator. The nonparametric regression approach is more flexible because it is not based on several assumptions that must be met as in the ARIMA model. In addition, with the Fourier series estimator, data patterns that are diverse, fluctuating, or those that form seasonal patterns can be captured clearly [11].

This study compares the accuracy of the model obtained from the ARIMA model and nonparametric regression model with a Fourier series estimator to obtain the most accurate model in predicting JCI values during the Covid-19 pandemic. Furthermore, the accuracy of the model will be reviewed from several indicators, namely Mean Square Error (MSE), Mean Absolute Deviation (MAD), coefficient of determination, and Mean Absolute Percentage Error (MAPE) [12]. This is a novelty in this research. The results of this study produce statistical models that are useful in predicting the value of JCI during the Covid-19 pandemic and the results of forecasting the value of JCI over a certain period. Thus, this research can be used as a reference in formulating appropriate policies to maintain the stability of the Indonesian capital market.

2. LITERATURE REVIEW

2.1 Jakarta Composite Index (JCI)

The Jakarta Composite Index (JCI) is one of the capital market indices used on the Indonesia Stock Exchange (IDX). The JCI is used as a reference for developments in the Indonesia capital market [13]. In addition, the JCI is used to measure the increase or decrease in stock prices. If a stock rises, it is positively correlated with an increase in the JCI. However, if a stock rises in price but the JCI goes down, then the stock has a negative correlation with the JCI.

2.2 Autoregressive Integrated Moving Average Model (ARIMA Model)

Basically, the ARIMA model is a model consisting of an autoregressive process, a moving average process, and a differencing process. This model is generally used in the modeling and forecasting of univariate time series data. In the ARIMA model, some assumptions must be met, such as stationary data, white noise residuals, normally distributed residuals, and no cases of heteroscedasticity. ARIMA model with order p in the autoregressive process, order q in the moving average process, and differencing as much as d is written as ARIMA (p,d,q). The complete ARIMA (p,d,q) model can be written as follows [9-10]:

$$\phi_p(B)(1-B)^d y_t = \Theta_q(B)\, e_t \qquad (1)$$

With,

$\phi_p(B)$: AR coefficient with order (p)
$\Theta_q(B)$: MA coefficient with order (q)
y_t : the actual data value on the t-th observation
e_t : residual

2.3 Nonparametric Regression

Nonparametric regression is used to detect the relationship between the response variable and the predictor variable when the nonparametric regression curve does not form a certain pattern. This is what distinguishes it from parametric regression, where parametric regression is used when the regression curve forms a certain pattern such as linear, quadratic, and so on. In general, the nonparametric regression model can be written as follows [11]:

$$y_i = m(x_i) + \varepsilon_i \qquad (2)$$

With,

$m(x_i)$: unknown regression function

In nonparametric regression, the regression residual is assumed to be random with mean 0 and variance σ^2. In addition, the assumptions underlying parametric regression can be ignored. In this case, the unknown regression function is estimated using a Fourier series estimator.

2.4 Fourier Series Estimator

Fourier series is one of the estimators that can be applied in nonparametric regression. This estimator is widely used when the data pattern is unknown or when seasonal data patterns are formed. The Fourier series estimator can be written as follows [14-15] :

$$m(x_i) = \beta_0 + \sum_{k=1}^{K}\left[a_k \cos\left(\frac{2\pi k(i-1)}{n}\right) + b_k \sin\left(\frac{2\pi k(i-1)}{n}\right)\right] \qquad (3)$$

With a_k and b_k as follows:

$$a_k = \frac{2}{n}\sum_{i=1}^{n} y_i \cos\left(\frac{2\pi k(i-1)}{n}\right) \qquad (4)$$

$$b_k = \frac{2}{n}\sum_{i=1}^{n} y_i \sin\left(\frac{2\pi k(i-1)}{n}\right) \qquad (5)$$

2.5 The Measure of The Goodness and Accuracy of The Model

A good model is a model that has high accuracy. In other words, forecasting results can represent actual data. To obtain a model with high accuracy, the error value should be kept as small as possible. The error value can be defined as follows:

$$e_i = y_i - \hat{y}_i \qquad (6)$$

With,
y_i : the actual data value on the i-th observation
$\hat{y}_i$: forecasting value of the i-th observation
e_i : the value of error on the i-th observation

There are many indicators that measure the accuracy of the model such as Mean Square Error, Mean Absolute Deviation (MAD), coefficient of determination, and Mean Absolute Percentage Error (MAPE). In full, the four indicators can be defined as follows [9-10,12] :

Mean Square Error (MSE)

$$MSE = \frac{1}{n}\sum_{i=1}^{n} e_i^2 \qquad (7)$$

Mean Absolute Deviation (MAD)

$$MAD = \sum_{i=1}^{n} \frac{|e_i|}{n} \qquad (8)$$

Coefficient of Determination

$$R^2 = 1 - \frac{\sum_{i=1}^{n} e_i}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \qquad (9)$$

Mean Absolute Percentage Error (MAPE)

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|e_i|}{y_i}\,100\% \qquad (10)$$

3. MATERIAL AND METHOD

3.1 DATA SOURCES AND RESEARCH VARIABLES

This study is based on a quantitative method that carries out an analysis of time series data. The data used is sourced from the Yahoo Finance, which is weekly data from January 2020 to November 2020 (46 weeks). The variable is the Jakarta Composite Index in rupiah currency.

In addition, the data used is divided into 2 parts, namely data from 1st week of January 2020 to 4th week of Oktober 2020 as the first section and data from 4th week of Oktober 2020 to 3rd week of November 2020 as the second section. The first section of the data is called in-sample data, and the second section is called out-sample data. In-sample data serves to determine model estimates. Meanwhile, testing the prediction accuracy of the formed model is done through out-sample data.

3.2 Steps of Analysis

In this study, data analysis was carried out using two approaches, namely ARIMA approach and nonparametric regression approach with the Fourier series estimator. In the ARIMA approach part, the analysis steps start from making sure the data is stationary, displaying the ACF and PACF curves, performing ARIMA modeling, testing the significance of parameters, and testing the prevailing assumptions. Whereas in the nonparametric regression approach with Fourier series estimator, the analysis step starts from determining the optimal lambda with the Generalized Cross Validation (GCV) method and estimating the model. The best model from the comparison of the two techniques above is used to predict the JCI value in the out-sample data. In this study, the error rate used was 5%.

4. RESULT AND DISCUSSION

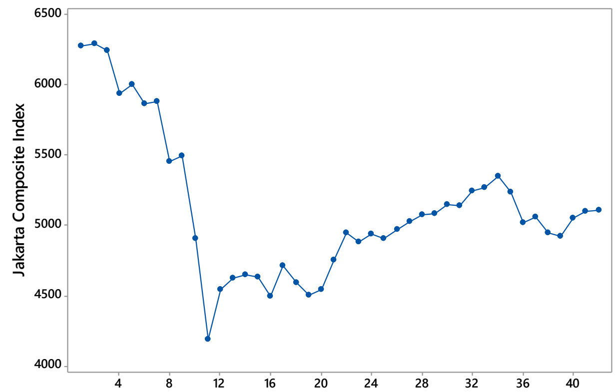
Descriptive statistics from the Jakarta Composite Index (JCI) data such as the total count, mean, median, StDev, and so on, will be shown in Table 1. Based on Table 1, the JCI data has an average of Rp. 5121.0. The standard deviation of JCI is 513.2. The standard deviation value is high shows the level of JCI fluctuations are quite high during the Covid-19 pandemic. A positive skewness value indicates the distribution of data sticks out to the right.

From Figure 1, it is seen that the JCI data fluctuates from time to time. Since Covid-19 pandemic, the JCI value decreased sharply and shows a fluctuating trend. This is illustrated by the very high ranges of maximum and minimum values. This shows that the pressure on the Indonesia capital market experienced during the Covid-19 pandemic is quite large.

Table 1: The descriptive statistics of JCI data.

Total Count	Mean	Median	StDev
42	5121.0	5042.5	513.2
Minumum	**Maximum**	**Skewness**	**Kurtosis**
4194.9	6291.7	0.84	0.23

Fig 1: JCI data plot.



4.1 ARIMA Modeling

Before doing ARIMA modeling, the data must be stationary. Based on Figure 2, it can be seen that the Autocorrelation Function (ACF) value of the data drops slowly indicating that the data is not stationary.

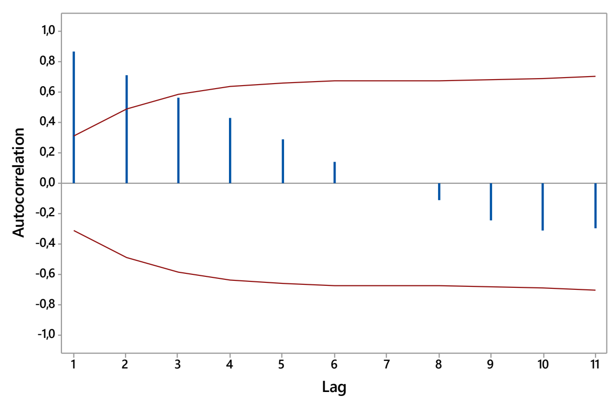


Fig 2: The ACF plot.

Visually, it has been obtained that the data is not stationary through the ACF plot. However, more formal testing is needed to identify the stationarity of the data. In this case, the data stationarity test was carried out through the Box-Cox transformation and the Augmented Dickey-Fuller (ADF) test. The Box-Cox transform is useful for knowing whether the

International Journal of Academic and Applied Research (IJAAR)
ISSN: 2643-9603
Vol. 5 Issue 12, December - 2021, Pages:36-41

data is stationary in the variance. While the ADF test is useful to find out whether the data is stationary in the mean.

Based on Figure 3, the rounded value of Box-Cox transformation is very close to 3 indicated that the data is not stationary in the variance. Furthermore, it was concluded that the data were not stationary in the mean based on the result of ADF test. Thus, in this case, the transformation and differencing process of the data is needed.

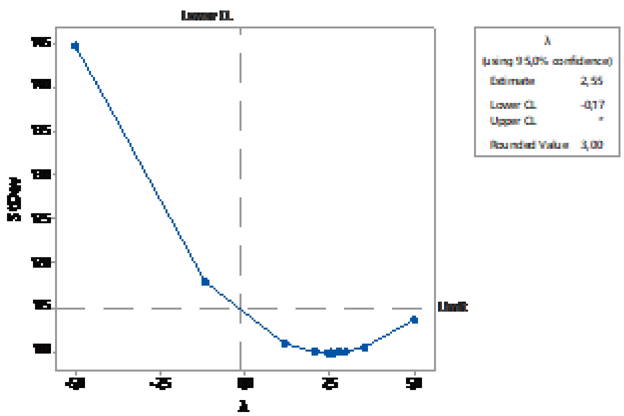


Fig 3: Result of Box-Cox Transformation.

Table 2: Result of ADF Test.

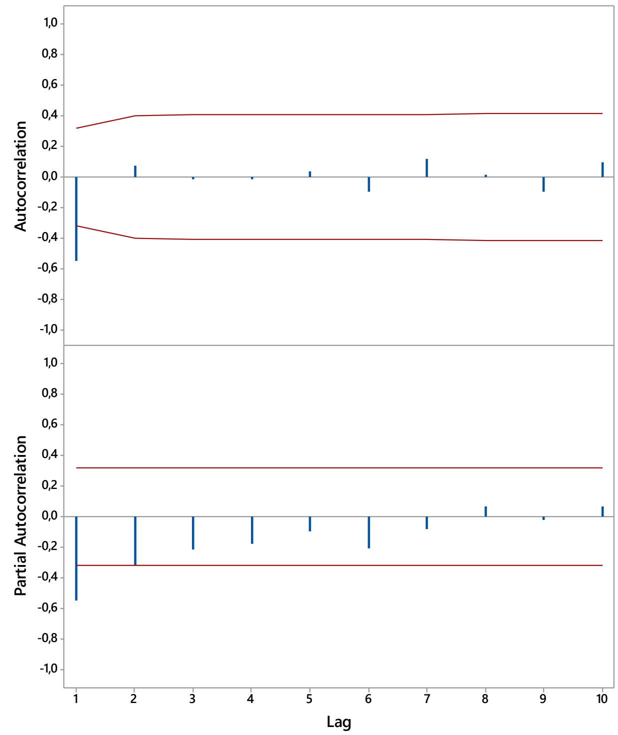
	Dickey-Fuller	Lag Order	P-value
Before Differencing	-2.1917	3	0.4981
After Differencing (1 Times)	-2.6989	3	0.2981
After Differencing (2 Times)	-4.6212	3	0.01

After transforming the data and the differencing process twice, the data obtained are stationary in terms of mean and variance. In other words, the assumption of stationarity in the ARIMA model has been met. To form the ARIMA model, we need order values for the autoregressive process and the moving average process. The order can be obtained by looking at the ACF and PACF curves from stationary data.

Based on Figure 4 and Figure 5, the possible orders for ARIMA are ARIMA (1,2,1), ARIMA (0,2,1), and ARIMA (1,2,0). After being analyzed, none of the ARIMA models meet the parameter significance test. This shows that the ARIMA model is not suitable for use. However, forecasting from the ARIMA model can still be done. Referring to the research results of Kostenko, et. al. which shows that modeling with insignificant parameters can still be used for forecasting purposes as long as the model has a good level of accuracy [16]. Even so, the model formed which is ARIMA (0,2,1) has a fairly good measure of the goodness and accuracy of the model. The estimation results of ARIMA (0,2,1) are presented in Table 3.

Fig 4: ACF of Data that has been stationary.
Fig 5: PACF of Data that has been stationary.

Table 3: Estimation Result of ARIMA(0,2,1).



Variable	MA(1)
Coefficient	0.1000
Std. Error	0.1679
t-Statistic	0.6000
P-value	0.5550

The modeling results can be written as follows.

$$\hat{y}_t = 2\,\hat{y}_{t-1} - \hat{y}_{t-2} + e_t - 0.1000\,e_{t-1} \qquad (11)$$

Based on Figure 6, it is clear that the ARIMA model is able to predict the value of in-sample data well. This is in line with the goodness and accuracy of the model as shown in Table 4.

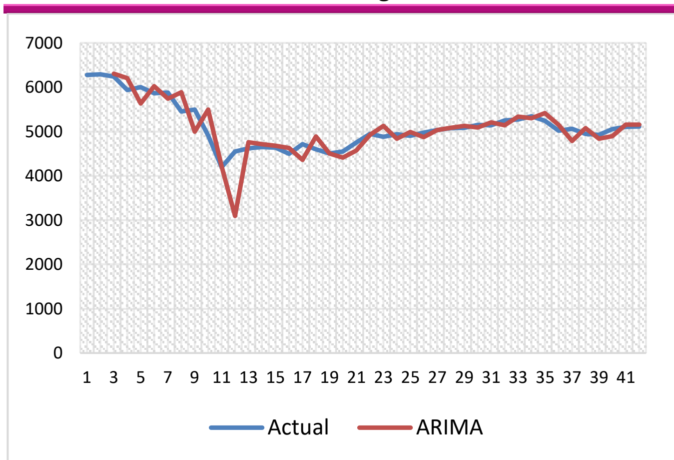


Fig 6: Comparison of forecasting results from the ARIMA model with actual data.

Table 4: Measures of Goodness and Accuracy of ARIMA Model.

MSE	Coefficient of Determination	MAPE	MAD
93558.63	99.98%	3.65%	182.8523

4.2 Nonparametric Regression Modeling

In nonparametric regression, the Fourier series estimator is used. By using the Generalized Cross Validation (GCV) method, it was found that the optimal lambda value was 8. Therefore, this modeling was carried out using a lambda value of 8. Thus, the following model was obtained. The model obtained based on nonparametric regression with a Fourier series estimator can be written as follows:

$$
\begin{aligned}
m(t_r) = 5121.03 &+ 468.30\, cos2\pi t_r \\
&- 65.01\, sin2\pi t_r \\
&+ 166.20\, cos4\pi t_r \\
&+ 355.62\, sin4\pi t_r \\
&+ 19.40\, cos6\pi t_r \\
&+ 255.42\, sin6\pi t_r \\
&+ 7.82\, cos8\pi t_r \\
&+ 119.10\, sin8\pi t_r \\
&+ 40.32\, cos10\pi t_r \\
&- 29.58\, sin10\pi t_r \\
&+ 107.19\, cos12\pi t_r \\
&+ 31.08\, sin12\pi t_r \\
&+ 23.95\, cos14\pi t_r \\
&+ 83.31\, sin14\pi t_r \\
&+ 13.28\, cos16\pi t_r \\
&+ 84.75\, sin16\pi t_r
\end{aligned} \tag{11}
$$

It can be seen that the model is able to detect actual data very well. The values for the goodness and accuracy of the model are presented in Table 5.

Table 5: Measures of Goodness and Accuracy of Nonparametric Regression Model.

MSE	Coefficient of Determination	MAPE	MAD
13459.6	100%	1.57%	80.5721

The modeling is illustrated along with the actual data as in Figure 7.

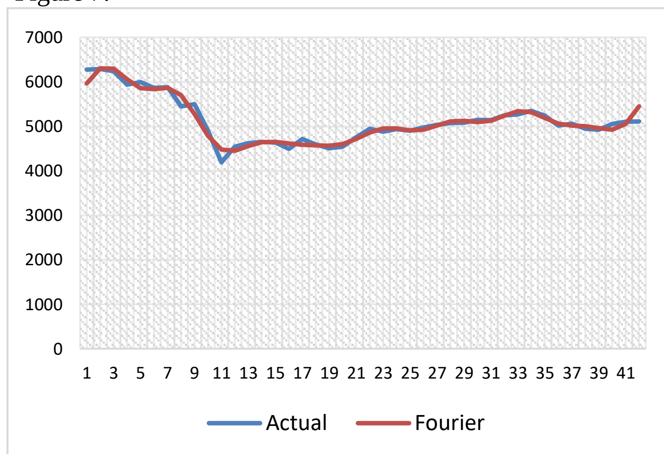
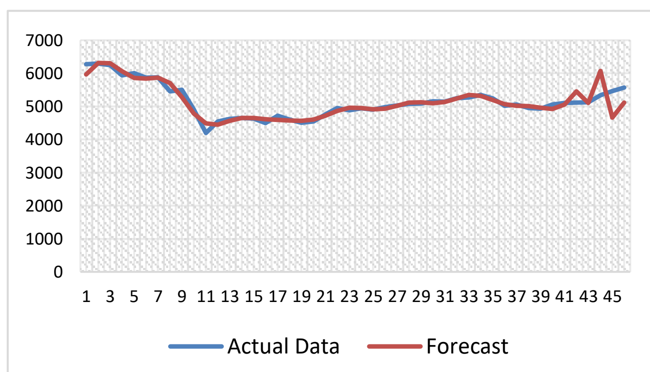


Fig 7: Comparison of forecasting results from the Nonparametric Regression model with actual data.

4.3 Forecasting

In this section, forecasting will be carried out using a nonparametric regression approach with the Fourier series estimator using previously obtained models. The plot between



the actual data and the forecasting results in the out-sample data is presented in Figure 8. It can be seen that the forecast results show fluctuations, but not too big. The level of accuracy of the model for the out-sample data is presented in Table 6. A MAPE value of 9.34% indicates that the model has a good level of accuracy in predicting data.

Fig 8: Forecasting Results Using Nonparametric Regression model with the Fourier Series Estimator.

Table 6. Forecasting Results and MAPE Value.

Actual Data	Forecasting	Error	APE
5128.23	5100.638661	27.5913394	0.005380285
5335.53	6074.50754	-738.97754	0.138501244

5461.06	4658.285275	802.774725	0.146999799
5571.66	5109.351571	462.308429	0.082974989
	MAPE		0.093464079

5. CONCLUSION

The ARIMA approach is not suitable for use in JCI data. This is indicated by the insignificance of the parameters in the model. Meanwhile, the nonparametric regression model shows good results in modeling. This is shown by the level of accuracy and goodness of the model which is very good. The analysis shows that the JCI value will still fluctuate although not significantly. Thus, the existing policies can be maintained for the near term.

6. REFERENCES

[1] BBC. (2020). Indonesia in Recession For First Time in 22 Years. Retrieved 21 September 2021, from https://www.bbc.com/news/business-54819898

[2] Indonesia Stock Exchange. (2020). IDX Implements Provisions Regarding Trading Halt for Trading in the Exchange. Press Release No 022/BEI/SPR/03-2020

[3] Indonesia Stock Exchange. (2020). IDX Policy Explanation Regarding Changes to the Auto Rejection Limit on Stock Trading on the Exchange. Press Release No. 026/BEI/SPR/03-2020

[4] Andreas, C., Harianto, F. Y., Safitri, E. J., and Chamidah, N. (2021). Analyzing The Effect of BI 7-Days Repo Rate on The Jakarta Composite Index Using Nonparametric Regression Approaches Based on Least Square Spline Estimator. Jurnal Matematika, Statistika & Komputasi, 17(3): 447-461.

[5] Ulyah, S. M., Andreas, C., and Rahmayanti, I. A. (2021). Forecasting Gold and Oil Prices Considering US-China Trade War Using Vector Autoregressive with Exogenous Input. AIP Conference Proceedings 2329: 060020.

[6] Andreas, C., Rahmayanti, I. A., and Ulyah, S. M. (2021). The Impact of US-China Trade War in Forecasting The Gold Price Using ARIMAX Model. AIP Conference Proceedings 2329: 060011.

[7] Rahmayanti, I. A., Andreas, C., and Ulyah, S. M. (2021). Does US-China Trade War Affect The Brent Crude Oil Price? An ARIMAX Forecasting Approach. AIP Conference Proceedings 2329: 060010.

[8] Abonazel, M. and Abd-Elftah, A. (2019). Forecasting Egyptian GDP Using ARIMA Models. Reports on Economics and Finance Vol. 5 (1) : 35-47.

[9] Wei, W.W.S. (2006). Time Series Analysis: Univariate and Multivariate Methods. Addison-Wesley, Inc.

[10] Cryer, J. D., and Chan, K. S. (2008). Time Series Analysis with Application in R, 2nd Edition. New York : Springer.

[11] Takezawa, K. (2006). Introduction to Nonparametric Regression. New Jearsey : John Willey & Sons, Inc.

[12] Montgomery, D. C., Jennings, C. L., and Kulahci, M. (2008). Introduction to Time Series Analysis and Forecasting. John Wiley & Sons, Inc.

[13] Indonesia Stock Exchange. (2018). Index. Retrieved 21 September 2021, from https://www.idx.co.id/en-us/products/index/

[14] Mardianto, M., Kartiko, S., Utami, H. (2017). Forecasting Trend-Seasonal Data using Nonparametric Regression with Kernel and Fourier Series Approach. Proceedings of The Third International Conference on Computing, Mathematics and Statistics pp 343-349.

[15] Mardianto, m. F. F., Kartiko, s. H., and Utami, H. (2020) The performance of nonparametric regression for trend and seasonal pattern in longitudinal data. ARPN Journal of Engineering and Applied Sciences, 15(9), pp.1111 – 1115.

[16] Kostenko, A. V., and Hyndman, R. J. (2008). Forecasting without significance tests. Manuscript, Monash University, Australia.