Time Series Analysis of Textile Product Testing Using ARIMA Model

M S Farabi[1,1], **I Surjandari**[1,2]

[1] Industrial Engineering Department, Faculty of Engineering, Universitas Indonesia, Kampus UI Depok, Indonesia
[1]moh.suhud@ui.ac.id, [2]isti@ie.ui.ac.id

Abstract. Product quality testing is essential to maintain consumer satisfaction in the textile company. Thus, the quality assurance industry should be aware of the shifting fashion trend. The application of forecasting has an impact on corporate planning, which can be used as a reference. In addition, good forecasting can make the cost of production efficient. The ARIMA model or Autoregressive Integrated Moving Average is a good time series data model to understand data for forecasting. ARIMA modelling aims to find an excellent statistical relationship between the predicted variables and the historical values of these variables so that this model can be used to make predictions. Forecasting accuracy is calculated using the value of Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE). The results show that the best model for forecasting the demand for textile product testing is ARIMA (2,1,0), with an RMSE of 195.88 and MAPE value of 7.01%.
Keywords: Forecasting, ARIMA, Time series, Textile.

1. Introduction

Product quality testing is essential in the industry to satisfy consumers. The importance of quality in the industry is increasing rapidly as competition rises [1]. In Indonesia, there are 323 textile companies registered at the ministry of industry [2]. Thus, Indonesia is one of the world's largest textile and apparel-producing and exporting countries [3]. Due to the high competition to meet the market for textile products worldwide, guaranteeing product quality must compete worldwide. The process of testing and certifying product quality is used to ensure the quality of the final product.

The demand for chemical testing of these textile products fluctuates and changes irregularly, merging into one combination. Changes influence these fluctuations in testing requests related to world fashion trends. The application of good forecasting is very impactful for corporate planning. In the company by which this study was conducted, the forecasting only focuses on the number of test request forms rather than based on the actual test data. It is because each test request form has different tests (depending on the type of textile being tested). However, the company where this research was conducted to optimize optimized the forecasting process in chemical testing. Inventors that occur cause delays in the testing process and increase the cost of procuring consumables. Therefore, the application of forecasting based on the actual number of chemical tests is expected to assist in planning consumables supplies. The aim is to avoid shortages or excess inventory so that inventory planning can be well prepared.

Based on the literature study, the fluctuating data forecasting model that can be used is ARIMA (Autoregressive Integrated Moving Average). The ARIMA model is a good time series data model to understand data for forecasting [4]. ARIMA modelling aims to find an excellent statistical relationship between the predicted variables and the historical values of these variables so that this model can be used to make predictions. However, the ARIMA approach only works well for short-term data forecasting, and its long-term forecasting accuracy is mediocre [5]. During the pandemic, some researchers used ARIMA to predict the COVID confirmation cases. Thus, it can help the government to prepare healthcare facilities and reduce the effect of the pandemic shortly [6][7]. Besides, ARIMA also has benefits in the economic area, such as predicting the market price [8] and predicting GDP growth [9][10][11][12]. Therefore, it helps the government to determine fiscal and monetary policies and plans future operating activities. Forecasting modelling also benefits private companies, especially in predicting sales and demand [13][14][15].

This study presents demand forecasting based on actual testing data with the ARIMA model. During the modelling process, the mean absolute percentage error (MAPE) and the root mean squared error (RMSE) are considered to select the optimum model.

2. Methods

2.1. Data source

In this study, the data is from a company engaged in analysis service, testing, and certification of textile products. The data used in the modelling are data demand for testing in chemical laboratories from January 2015 to October 2022. The data are divided into training and testing data. Data testing was used from January 2021 to October 2022, and data training from January 2015 to December 2020.

2.2. ARIMA models

ARIMA is often referred to as the Box-Jenkins time series method [16] because Box and Jenkins were introduced in the 1970s [17]. ARIMA model aims to find an excellent statistical relationship between the predicted variables and the historical values of these variables, so it can be used to make predictions [18]. This model can be explained by three-parameter arguments using the formula ARIMA (p, d, q), where q is the order of the moving average, p is the order of the autoregression, and d is the amount of the difference [19]. This model is classified into three groups, including autoregressive (AR) models, moving average (MA) models, and mixed models with the characteristics of the first two models (ARMA).

AR (p) model uses a linear combination between previous value variables ($y_{t-1}, y_{t-2}, \ldots, y_{t-p}$) and current residual ($\varepsilon_t$). Meanwhile, the MA (q) model use a linear combination of current and previous residual values ($\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots, \varepsilon_{t-2q}$). The AR(p) (1) and MA (q) (2) equation is written as:

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \cdots + \varphi_p y_{t-p} + \varepsilon_t \quad (1)$$

$$y_t = \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (2)$$

ϕ and θ interpreted as autoregressive and moving average Constanta. y_t represents the time value of response at time *t*, and ε_t represents white noise or random error at time *t*.

The ARMA (p, q) model will exist if the AR and MA models combine. The current time series response in ARMA (p, q) is linearly connected to its past values and the present and past residual series. The equation of ARMA can be presented as follows:

$$y_t = \alpha + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \cdots + \varphi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \quad (3)$$

The value of α is a constant, and the value of the previous random error was denounced as ε_{t-1}. Before continuing the process, data should be stationary in the mean, else the differentiation had to do. Thus, ARMA models become ARIMA (p, d, q), and the value of d is associated with a degree of differentiation [20]. The following four processes include model identification, parameter estimates, diagnostic verification, and building an ARIMA model fix time series (mean and variance stationary).

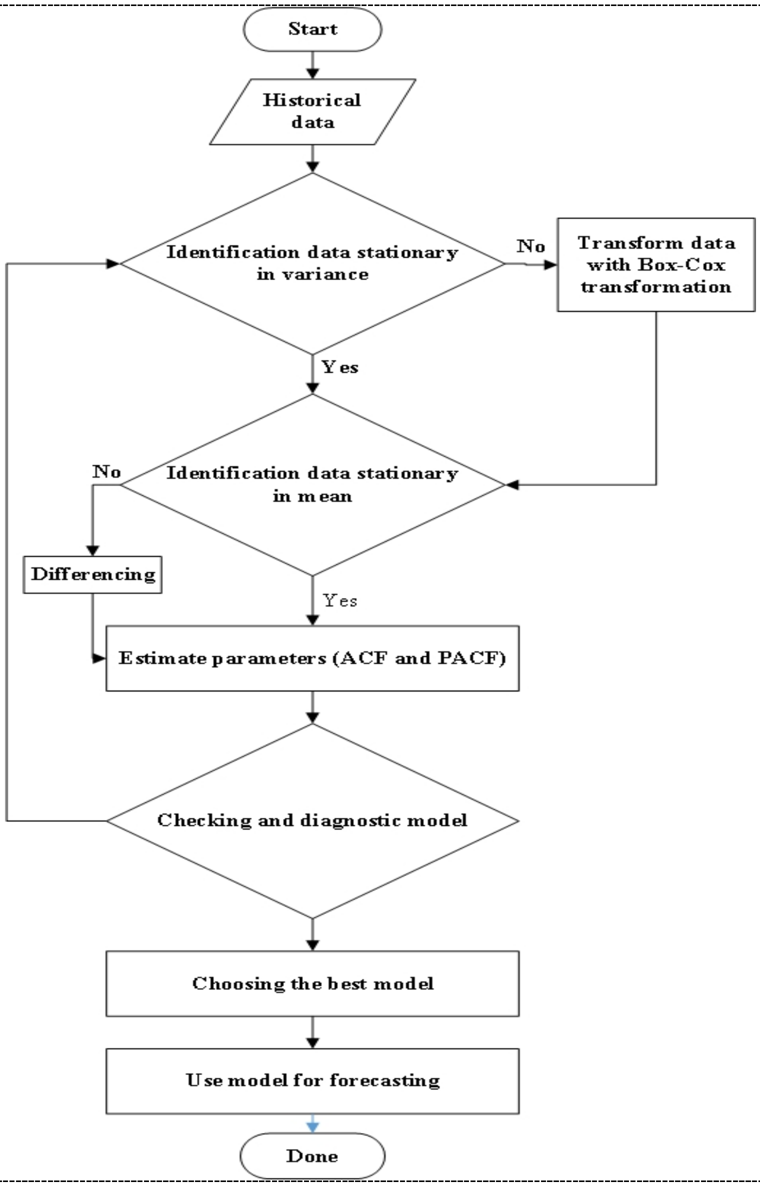

Figure 1. Flow Diagram of Research Methodology

2.3. Assessment and identification

Before estimating the model's parameter, ensure the variance and mean of time series data are stationed. It implies the variance and means data are constant over time. Static data in variance can be identified by Box-Cox plot, while stationary in the mean can be identified by using The Augmented Dickey-Fuller test (ADF). The transformation and differences process involved setting the variances and mean time series [21].

2.4. Model Parameter Estimation

After the data is stationary, the autocorrelation function (ACF) and partial autocorrelation function should be used for determining orders of autoregressive terms (AR) and moving mean (MA) then the ARIMA model will be formed.

2.5. Diagnostic checking

The Ljung-Box test (Q*) checked the prediction model's feasibility. This test scans that residual does not occur autocorrelation [22].

Shapiro-Wilk test (w) is beneficial for testing normal distribution in statistics. This test evaluates the null hypothesis that a sample was drawn from a population with a normal distribution [23].

The element of uncertainty (residual) always involves prediction models. However, it can be predicted to find the best models. This study used two types of performance criteria, namely Root Means Square Error (RMSE) and Mean Absolute Percentage Error (MAPE), to check the ARIMA model performance accuracy.

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(e_t)^2} \quad (5)$$

$$\text{MAPE} = \frac{1}{N}\sum_{t=1}^{N}\left|\frac{e_t}{y_t}\right| \quad (6)$$

Where e_t = the residual at time t, N = the number of observations, and y_t = the observed response at time t, lower RMSE and MAPE values indicate a more accurate prediction model.

3. Result and discussion

3.1. Exploration data

Data exploration is done by determining the data demand for testing textile products throughout a monthly period of 96 data. Figure 2 shows a time series plot for the requested textile product testing from January 2015 to December 2022. The data looks non-stationary in both variance and mean.

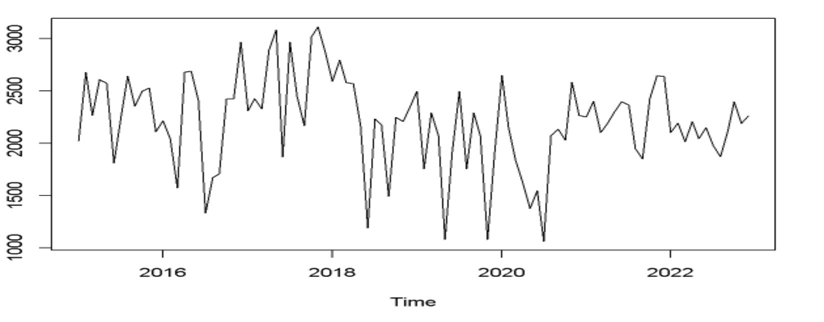

Figure 2. Data Actual Testing Textile Product at Laboratory

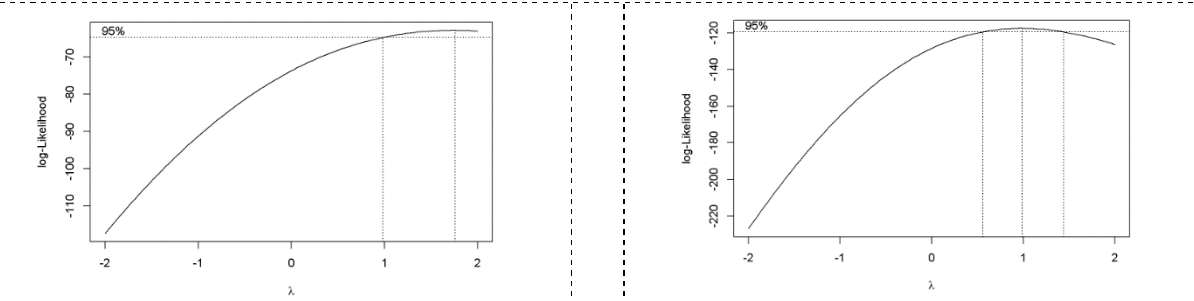

Figure 3. Box-Cox Plot Actual Testing at Laboratory Non-Stationary In Variance	**Figure 4.** Box-Cox Plot Transform The Data

Therefore, it is necessary to check the data using a Box-Cox plot to check stationary variance [24]. The ADF test checks stationary in the mean [25]. Figure 2 shows the result of checking stationarity in variance using a Box-Cox plot. According to Figure 3, it yields a lambda value of 1.767. This indicates that the data variance is not stationary since its value is too far from one. Therefore, a box-cox transformation must be performed. Figure 4 shows the result of approaching a lambda value close to one, which means that the variance of the transformed data is stationary.

39.

40. After the data had been transformed, it was checked for stationarity in the mean used ADF test. ADF test resulted in a p-value of 0.2908, which means the data is not stationary in the mean. Therefore, it is necessary to do differencing to transform data stationary into mean. The ADF test was performed to ensure the data was stationary in the mean after differencing at the first level. ADF test for the first level resulted in a p-value of 0.01, which means a p-value below a significant level of 5%. Figure 5 shows the data has been stationary in variance and mean, so the data can be identified and estimated the parameters using the ARIMA model.

41.

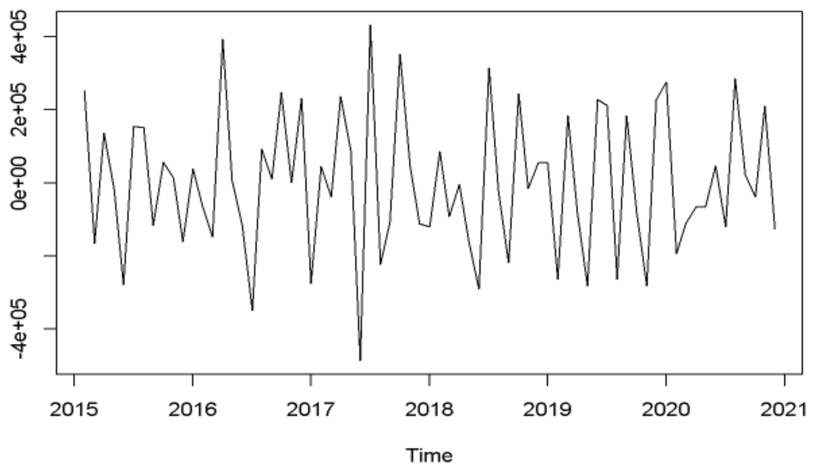

Figure 5. Data Actual Testing Textile Product at Laboratory

3.2. Identification and estimation ARIMA model

Identification and estimation parameters of the ARIMA model can be done by analyzing ACF and PACF plots. Figure 6 shows the ACF plot; it is known that lag 1 and 3 are out of significant level. Then, the MA estimator is 1 and 3. Figure 7 shows the PACF plot, which shows that lag 1 is out of significant level, so the AR estimator is 1. The estimating parameters are obtained based on ACF and PACF plots, as shown in Table 1.

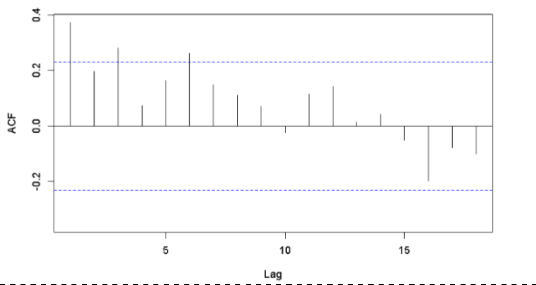

Figure 6. ACF Plot

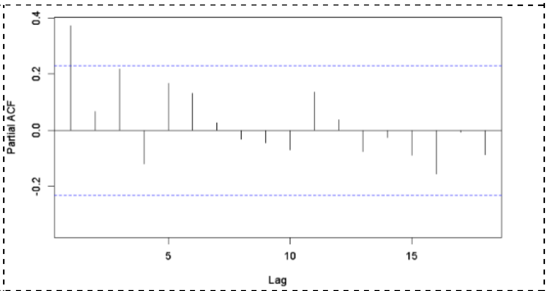

Figure 7. PACF plot

Table 1 shows three ARIMA models in which all the estimating parameters are significant at the 5% level. The significant models are ARIMA (1,1,0), ARIMA (0,1,1), and ARIMA (2,1,0).

Table 7 Estimated ARIMA Model

Model	Parameter	P-value	Description
ARIMA (1,1,1)	MA (1)*	<2e-16	Significant
	AR (1)	0.1473	Not significant
ARIMA (1,1,3)	MA (1)	0.8656	Not significant
	MA (2)*	0.0007	Significant
	MA (3)	0.9945	Not significant
	AR (1)	0.1359	Not significant
ARIMA (1,1,0)	AR(1)*	0.0013	Significant
ARIMA (0,1,1)	MA (1)*	5.27E-15	Significant
ARIMA (0,1,3)	MA (1)*	1.38E-06	Significant
	MA (2)*	0.0246	Significant
	MA (3)	0.2733	Not significant
ARIMA (2,1,0)	AR (1)*	7.62E-06	Significant
	AR (2)*	0.0004	Significant

*Parameters that significant at the 5% level

3.3. Model diagnostic

A significant model based on Table 1 will be diagnosed by Ljung-Box and Shapiro-Wilk tests. It is important to check model feasibility. Table 2 presents the outcomes of the Ljung-Box and Shapiro-Wilk tests.

Table 8 Diagnostic Parameters Test Results

Model	Ljung-Box test (p-value)	Shapiro-Wilk test (p-value)
ARIMA (1,1,0)	0.2469	0.7981
ARIMA (0,1,1)	0.2822	0.7974
ARIMA (2,1,0)	0.9748	0.6776

Table 2 shows the results of the Ljung-Box and Shapiro-Wilk tests. The Ljung-Box test was carried out to determine the presence of residual autocorrelation in the model formed [26]. From Table 2, it was found that the three models were free from autocorrelation because the p-value is greater than the significant level. Shapiro-Wilk is a test to find that the residual data is normally distributed [27]. The

Shapiro-Wilk test on the three ARIMA models has a normal distribution with a p-value more remarkable than the significant level.

Table 9 Forecasting Accuracy

Model	**Accuracy**	
	RMSE	**MAPE**
ARIMA (1,1,0)	215.24	8.04
ARIMA (0,1,1)	236.09	8.20
ARIMA (2,1,0)	195.88	7.01

Model accuracy is done by comparing the RMSE and MAPE values of the three ARIMA models. The model with the smallest RMSE and MAPE values is used for forecasting modelling. The forecasted findings are compared to 24 actual data from January 2021 to December 2022 to determine the accuracy of the prediction. Based on Table 3, ARIMA (2,1,0) has the smallest RMSE and MAPE. Then, the company can use ARIMA (2,1,0) as a model for forecasting the number of tests as a reference for planning consumable materials.

4. Conclusion

Indonesia has a high potential to meet the demand for textile products in the global market. This potential must be supported by the quality of textile products so that they can compete in the worldwide market. Product testing is required to guarantee the quality of textile products before being distributed to the worldwide market. This study focuses on forecasting the demand for textile product testing in a chemical laboratory in an analysis and certification service company. Based on the results obtained, for the number of requests for testing in chemical laboratories with the ARIMA model, the optimum model is ARIMA (2,1,0) with an RMSE is 195.88 and MAPE is 7.01%.

This forecasting using the ARIMA model is based on historical data on the number of request testing and is suitable for various data patterns. However, the accuracy could be better for long-term forecasting because it will produce a consistent result. In addition, consumer loyalty and export-import regulations can affect the number of test requests. Due to changes in test requests, the supply chain becomes unstable, and it is difficult to predict changes in demand. However, in general, these factors only sometimes occur. Therefore, future research is expected to include other elements and use other forecasting models to obtain a more appropriate model for this problem.

5. References

[1] Dhillon, B. S. (2004). *Reliability, quality, and safety for engineers*. *In Reliability, Quality, and Safety for Engineers*. CRC Press LLC.

[2] Prambadi, G. A. (2021, August 22), Industri tekstil Indonesia terus didorong ke arah lebih baik. *Republika.co.id.* Available: https://www.republika.co.id/berita/qy7oyl456/industri-tekstil-indonesia-terus-didorong-ke-arah-lebih-baik.

[3] Maskur, F. (2022, September 16). Negara produsen tekstil dunia: China terbesar, Indonesia keenam. *Bisnis Indonesia*. Available: https://bisnisindonesia.id/article/negara-produsen-tekstil-dunia-china-terbesar-indonesia-keenam.

[4] Rodrigues, P. C., Awe, O. O., Pimentel, J. S., & Mahmoudvand, R. (2020). Modelling the behaviour of currency exchange rates with singular spectrum analysis and artificial neural networks. *Stats*, *3*(2), 137-157.

[5] Alabdulrazzaq, H., Alenezi, M. N., Rawajfih, Y., Alghannam, B. A., Al-Hassan, A. A., & Al-Anzi, F. S. (2021). On the accuracy of ARIMA based prediction of COVID-19 spread. *Results in Physics*, *27*, 104509.

[6] Dehesh, T., Mardani-Fard, H. A., & Dehesh, P. (2020). Forecasting of covid-19 confirmed cases in different countries with arima models. *MedRxiv*.

[7] Sahai, A. K., Rath, N. S., & Singh, M. P. (2020). ARIMA modelling & forecasting of COVID-19 in top five affected countries. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews, 14*(5), 1419-1427.
[8] Meher, B. K., Hawaldar, I. T., Spulbar, C. M., & Birau, F. R. (2021). Forecasting stock market prices using mixed ARIMA model: A case study of Indian pharmaceutical companies. *Investment Management and Financial Innovations, 18*(1), 42-54.
[9] Abonazel, M. R., & Abd-Elftah, A. I. (2019). Forecasting Egyptian GDP using ARIMA models. *Reports on Economics and Finance, 5*(1), 35-47.
[10] Atanu, E. Y., Ette, H. E., Nwuju, K., & Nwaoha, W. C. (2020). ARIMA Model for gross domestic product (GDP): evidence from Nigeria. *Archives of Current Research International, 20*(7), 49-61.
[11] Eissa, N. (2020). Forecasting the GDP per Capita for Egypt and Saudi Arabia Using ARIMA Models. *Research in World Economy, 11*(1), 247-258.
[12] Muma, B., & Karoki, A. (2022). Modelling GDP Using Autoregressive Integrated Moving Average (ARIMA) Model: A Systematic Review. *Open Access Library Journal, 9*(4), 1-8.
[13] Permatasari, C. I., Sutopo, W., & Hisjam, M. (2018). Sales forecasting newspaper with ARIMA: A case study. *AIP Conference Proceedings. 1931*, p. 030017. AIP Publishing LLC.
[14] Salah, O. M., Mahdi, G. J., & Al-Latif, I. A. (2021). A modified ARIMA model for forecasting chemical sales in the USA. *Journal of Physics: Conference Series. 1879*, p. 032008. IOP Publishing.
[15] Udom, P., & Phumchusri, N. (2014). A comparison study between time series model and ARIMA model for sales forecasting of distributor in plastic industry. *IOSR Journal of Engineering, 4*(2), 32-38.
[16] Seneviratna, D. M. K. N., & Shuhua, M. (2013). Forecasting the Twelve Month Treasury Bill Rates in Sri Lanka: Box Jenkins Approach. *IOSR Journal of Economics and Finance (IOSR-JEF)*, *1*(1).
[17] Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: forecasting and control.* John Wiley & Sons.
[18] Fattah, J., Ezzine, L., Aman, Z., El Moussami, H., & Lachhab, A. (2018). Forecasting of demand using ARIMA model. *International Journal of Engineering Business Management*, *10*, 1847979018808673.
[19] Li, X., Zhang, C., Zhang, B., & Liu, K. (2019). A comparative time series analysis and modeling of aerosols in the contiguous United States and China. *Science of the Total Environment, 690*, 799-811.
[20] Zhang, H., Zhang, S., Wang, P., Qin, Y., & Wang, H. (2017). Forecasting of particulate matter time series using wavelet analysis and wavelet-ARMA/ARIMA model in Taiyuan, China. *Journal of the Air & Waste Management Association*, *67*(7), 776-788.
[21] Damesa, T. M., Möhring, J., Forkman, J., & Piepho, H. P. (2018). Modeling spatially correlated and heteroscedastic errors in Ethiopian maize trials. *Crop Science*, *58*(4), 1575-1586.
[22] Hussin, N. H., Yusof, F., & Norrulashikin, S. M. (2021). Forecasting Wind Speed in Peninsular Malaysia: An Application of ARIMA and ARIMA-GARCH Models. *Pertanika Journal of Science & Technology*, *29*(1).
[23] Moftakhar, L., Mozhgan, S. E. I. F., & Safe, M. S. (2020). Exponentially increasing trend of infected patients with COVID-19 in Iran: a comparison of neural network and ARIMA forecasting models. *Iranian Journal of Public Health*, *49*(Suppl 1), 92.
[24] Anggraeni, W., Andri, K. B., & Mahananto, F. (2017). The performance of ARIMAX model and Vector Autoregressive (VAR) model in forecasting strategic commodity price in Indonesia. *Procedia Computer Science*, *124*, 189-196.
[25] Nashwan, M. S., Ismail, T. A. R. M. I. Z. I., & Ahmed, K. A. M. A. L. (2019). Non-stationary analysis of extreme rainfall in peninsular Malaysia. *Journal of Sustainability Science and Management*, *14*(3), 17-34.

[26] Serra, R., & Rodríguez, A. C. (2012, September). The Ljung-Box test as a performance indicator for VIRCs. In *International Symposium on Electromagnetic Compatibility-EMC EUROPE* (pp. 1-6). IEEE.
[27] Unhapipat, S. (2018, June). ARIMA model to forecast international tourist visit in Bumthang, Bhutan. In *Journal of Physics: Conference Series* (Vol. 1039, No. 1, p. 012023). IOP Publishing.