ComTech: Computer, Mathematics and Engineering Applications, 14(1), June 2023, 55−67 P-ISSN: 2087-1244
DOI: 10.21512/comtech.v14i1.8462 E-ISSN: 2476-907X

The Implementation of Control Charts as a Verification Tool in a Time Series Model for COVID-19 Vaccine Participants in Pontianak

Nurfitri Imro'ah[1], Nur'ainul Miftahul Huda[2]*, and Abang Yogi Pratama[3]

[1]Statistics Department, Universitas Tanjungpura
Jln. Profesor Dokter H. Hadari Nawawi, Kalimantan Barat 78124, Indonesia
[2]Mathematics Department, Universitas Tanjungpura
Jln. Profesor Dokter H. Hadari Nawawi, Kalimantan Barat 78124, Indonesia
[3]Department of Communications and Informatics of Pontianak City
Jln. Rahadi Usman no. 3, Kalimantan Barat 78111, Indonesia
[1]nurfitriimroah@math.untan.ac.id; [2]nur'ainul@fmipa.untan.ac.id;
[3]abangyogipratama22@gmail.com

Received: 18th April 2022/ **Revised:** 13th October 2022/ **Accepted:** 14th October 2022

How to Cite: Imro'ah, N., Huda, N. M., & Pratama, A. Y. (2023). The Implementation of Control Charts as a Verification Tool in a Time Series Model for COVID-19 Vaccine Participants in Pontianak. *ComTech: Computer, Mathematics and Engineering Applications, 14*(1), 55−67. https://doi.org/10.21512/comtech.v14i1.8462

Abstract **-** Vaccines are the primary weapon used to stop the outbreak, especially amid the COVID-19 pandemic. Thus, supplying vaccines to control the COVID-19 pandemic is essential, especially in minimizing the incidence and achieving herd immunity to break the chain of COVID-19. West Kalimantan has taken firm anticipatory steps to prevent COVID-19 in the form of a vaccination program in Indonesia. The highest vaccination achievement occurs in Pontianak City, the province's capital. The research analyzed data on vaccine participants in Pontianak using time series analysis. In addition, the residuals from the time series model were used as observations in constructing the control chart. The research also analyzes the accuracy of the time series model using the Individual Moving Range (IMR) control chart. The results show that the ARIMA model (5,0,2) is the best because it fulfills the assumption of white noise. However, the ARIMA (5,0,2) model is inaccurate in making predictions because the residuals from the ARIMA (5,0,2) model are out of control (based on the IMR control chart). Hence, it is necessary to evaluate in determining the time series model. It can be analyzed using a control chart. Therefore, measuring the model's accuracy on the best model is essential in predicting several subsequent periods.

Keywords: control charts, verification tool, time series model, COVID-19, vaccine participants

I. INTRODUCTION

Since 2019, the world has been rocked by the COVID-19 pandemic, which started in Wuhan. There are now 133 million cases of COVID-19 infection worldwide and 1,5 million cases in Indonesia (World Health Organization (WHO), n.d.). The pandemic has resulted in a crisis that impacts all aspects of human life. Even though there are many policies for treating COVID-19 patients, the spike in positive cases and mortality is still happening. Therefore, the government is currently severely tackling the COVID-19 pandemic, starting by conducting an emergency Policy for the Enforcement of Community Activity Restrictions (Pemberlakuan Pembatasan Kegiatan Masyarakat - PPKM) on July 3, 2021, to making efforts to import the COVID-19 vaccine, which all Indonesian people use to pursue the herd immunity at 70%. However, the number of exposure cases continues to increase even though the policy for PPKM has been implemented. Then, 3M prevention efforts (wearing masks, maintaining distance, and washing hands) are considered insufficient to suppress the spread of this virus. Finally, the government set a target to carry out vaccinations for all Indonesian people, up to two million doses in one day. In addition, the government has promoted a vaccination program to restore conditions to pre-pandemic conditions (Kementerian Kesehatan Republik Indonesia, 2021).

*Corresponding Author

Since the first case in Indonesia was reported in early March 2020, the COVID-19 pandemic has still shown an increasing curve for both cases and deaths. COVID-19 is a disease caused by a virus. So, immunity is essential, considering that few antivirals are available. The interaction between the immune system of the virus determines whether the body will recover or get worse. Vaccines can provide a specific immune response to a particular type of disease. Vaccinations are classified as active immunity, usually lasting several years or throughout life. The COVID-19 vaccine is included in the inactivated vaccine type, or called inactivation. The COVID-19 vaccination is currently carried out too in Indonesia. Phase 1 vaccination is intended for human resources for health in all health facilities. Moreover, in the next stage, all Indonesian people also participate in this vaccination (Kementerian Kesehatan Republik Indonesia, 2021).

Vaccination aims to make a person's immune system recognize and quickly fight bacteria or viruses that cause infection. Like other vaccines, the COVID-19 vaccine can protect the body from diseases caused by COVID-19 by stimulating the body's specific immunity by administering the (Kementerian Kesehatan Republik Indonesia, 2021). Although it is not 100% able to protect a person from COVID-19 infection, this vaccine can reduce the possibility of severe symptoms and complications. In addition, the COVID-19 vaccination aims to encourage the formation of herd immunity. It is important because some people cannot be vaccinated for specific reasons. Hence, vaccines are the primary weapon used to stop an outbreak, especially in the COVID-19 pandemic.

All provinces in Indonesia, including West Kalimantan, have taken strict anticipatory steps to prevent COVID-19 in the form of a vaccination program. Vaccines are circulated periodically and according to the occupational risk or age who are easily exposed to the COVID-19 virus. On January 17, 2021, the Governor of West Kalimantan released the distribution of the Sinovac vaccine to three districts, namely Pontianak City, Kubu Raya Regency, and Mempawah. Each region got a different amount of vaccine. It distributed 10.400 doses for Pontianak City, 3.480 doses for Kubu Raya, and 2.000 doses for Mempawah. The first vaccination was carried out on Thursday, January 14, 2021. The recipients of this first batch of vaccines were health workers and the Indonesian National Police members. The achievement of the COVID-19 vaccination in West Kalimantan was only 42,23% until the end of November 2021. Nine districts were asked to be more aggressive because their vaccination achievement was 40%. The districts were Sambas, Sintang, Melawi, Mempawah, Ketapang, Kapuas Hulu, Kubu Raya, North Kayong, and Landak Regency. Meanwhile, the highest vaccination achievement was in Pontianak City, 66,80%, with a target of 473.070 people (Dinas Komunikasi dan Informatika Provinsi Kalimantan Barat, 2021).

The number of people who vaccinate in a certain period can be considered time-series data. The Autoregressive Integrated Moving Average (ARIMA) method is one of the methods in time series analysis that can capture the necessary information regarding the number of people vaccinated during a specific period. ARIMA is a time series forecasting approach suited for predicting many variables fast, efficiently, inexpensively, and accurately. It only requires the variable data that will be forecasted (Wang et al., 2022). In addition to its more common name, the ARIMA method is sometimes referred to as the Box-Jenkins time series method. In 1970, George Box and Gwilym Jenkins significantly advanced this approach (Bagmar & Khudri, 2021). Time series data modeling can use three iterative stages of Box-Jenkins: model identification, parameter estimation, and diagnostic tests (Farimani, Parsafar, & Mohammadi, 2022). The best model in time series analysis is obtained at the diagnostic test stage, and the residual must be white noise. Measuring the model's accuracy on the best model is very important in predicting the following several periods (Alabdulrazzaq et al., 2021). Such a model must be measured for accuracy to be used for forecasting several periods in the future (Semenoglou, Spiliotis, Makridakis, & Assimakopoulos, 2021).

The measurement of the accuracy of the time series model can be seen from the residuals (Abanda, Mori, & Lozano, 2019). The time series model is accurate if the residuals are in a statistically controlled state (Herdiani, Fandrilla, & Sunusi, 2018). Furthermore, it can be seen on the control chart. Various researchers have contributed to the body of knowledge on control charts by utilizing time series data in their investigations. For example, it discusses using a double-moving average control chart for autocorrelated data (Arooj & Malik, 2022). It also has discussion of the performance evaluation of an Exponentially Weighted Moving Average (EWMA) chart for an exponentially dispersed process (Abbasi, Abid, Riaz, & Nazir, 2020). In addition, it discusses the opposition to integrated quality, maintenance, and production model based on the delayed monitoring under the ARMA control chart (Jafarian-Namin, Fallahnezhad, Tavakkoli-Moghaddam, Salmasnia, & Fatemi Ghomi, 2021).

The research analyzes the accuracy of the time series model obtained using a control chart. The case study uses data on vaccine participants in Pontianak City. The residual of the best model obtained is presented in the control chart (in this case, the Individual Moving Range (IMR) control chart). If the residual is in control, the time series model is accurate for another forecasting. On the other hand, the model is inaccurate if the residual is out of control. Even if the best available model has been obtained based on the criteria that have been established, the research results are expected to enhance the theories that have already been developed and increase knowledge about the usefulness of verifying the time series model. In addition to all of that, it is hoped that the research will provide a summary of the number of people who have

participated in the vaccination program in Pontianak City.

II. METHODS

The research uses information regarding the number of vaccination participants in Pontianak City between June 20, 2021, and November 19, 2021. The data collection takes place daily, and there are 153 total observations. The Department of Health in Pontianak City is the source of these statistics. Then, the data are analyzed using the ARIMA technique, consisting of three iterative Box-Jenkins stages. The model is verified with the help of an IMR control chart. It has been determined that the model with the fewest Akaike Information Criterion (AIC) and Root Mean Square Error (RMSE) values is the one that should be used.

The first step in the data analysis for the research is modeling the data on the number of people that have participated in the vaccination using ARIMA. This modeling step comprises stationarity testing, order identification using Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots, parameter estimation, and diagnostic verification. After choosing the best model, the model is subjected to verification procedures. The construction of an IMR control chart with residuals from the best ARIMA model is how verification is accomplished. If the plot demonstrates that things are in control, it can be concluded that the ARIMA model accurately predicts the following periods. Conversely, the model is inaccurate if the plot indicates an out-of-control condition. Figure 1 provides additional information regarding the steps involved in the research process.

If the sequence of random variable Z_t follows the ARIMA process, Z_t can be defined in Equation 1 (Benvenuto, Giovanetti, Vassallo, Angeletti, & Ciccozzi, 2020). It has $\phi_p(B) = \left(1 - \phi_1 B - \cdots - \phi_p B^p\right)$, $\theta_q(B) = \left(1 - \theta_1 B - \cdots - \theta_q B^q\right)$, and a_t is an error at time t. The model in Equation (1) is called the ARIMA model with the order (p, d, q), and p; d; q are the auto-regressive, differentiation, and moving average orders. The model is written as the ARIMA (p, d, q) model. If it is $p = 0$, the ARIMA (p, d, q) model is also called the IMA model with order (d, q). Likewise, if it is $q = 0$, the ARIMA (p, d, q) model is called an ARI model with order (p, d) (Katoch & Sidhu, 2021).

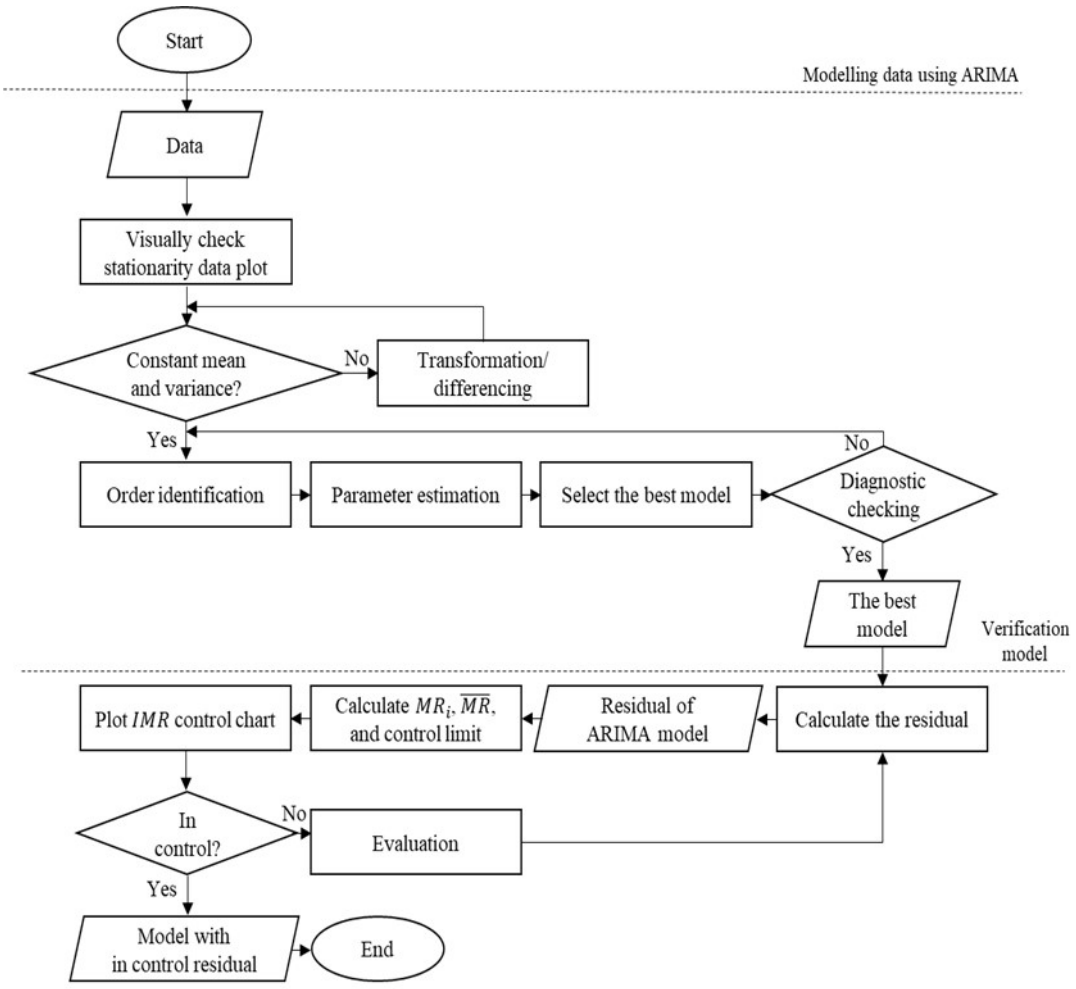

Figure 1 Flowchart of Implementation Control Chart in Time Series Model

$$\phi_p(B)(1-B)^d Z_t = \theta_0 + \theta_q(B)a_t \quad (1)$$

There are three main steps in time series modeling (Sahu, 2021). The first step is model identification. The selection of the order *p, d, q* in the ARIMA model is determined by observing ACF and PACF patterns. The second step is parameter estimation. Several often used parameter estimation methods are Maximum Likelihood Estimation (MLE) and Ordinary Least Square (OLS). The last step is the diagnostic test. Diagnostic examination of the model includes residual white noise and residual normality tests. After the three main steps are carried out, the best model is obtained and used as a reference for *h*-future predictions.

The control chart is a statistical quality control tool in the form of a graphical display of quality characteristics that have been measured or quantified in a sample according to time or observations (García, Peñabaena-Niebles, Jubiz-Diaz, & Perez-Tafur, 2022). The control chart was first introduced by DR. Walter Andrew Shewhart of Bell Telephone Laboratories, United States, in 1924 to eliminate abnormal variations by separating variations caused by common causes. All processes display variations, but the production process must be controlled by eliminating the particular causes of variations from the process so that, generally, causes only cause variations (Koutras & Triantafyllou, 2020). The control chart contains the Centre Line (CL), representing the average value of the quality characteristics corresponding to the state under control. The other two horizontal lines, namely at the top, are the Upper Control Limit (UCL), and the bottom is the Lower Control Limit (LCL) (Golilarz et al., 2019). An example of a control chart for some conditions can be seen in Figures 2−5 (Fan, Xue, Yi, & Xu, 2021). For the first condition, a process is said to be out of control if there is a point beyond the control limits. An example of this condition can be seen in Figure 2.

The second condition is said to be out of control if at least seven consecutive points are located above or below CL which is presented in Figure 3. This pattern suggests that something has taken place that has caused the process average to rise. It implies a unique variable at work here. Discovering specific

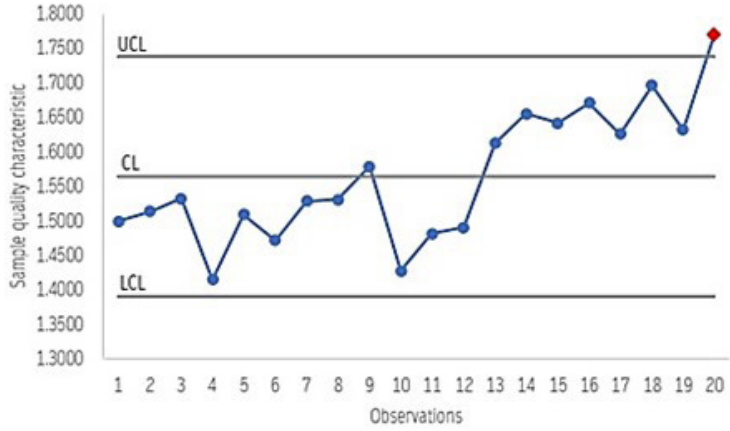

Figure 2 The Example of a Control Chart Showing an Out-of-Control Process Caused by a Point that is Beyond the Control Limits

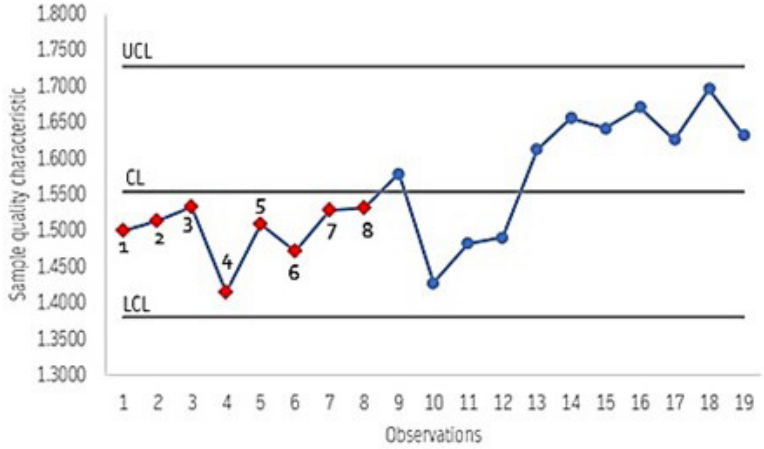

Figure 3 The Example of a Control Chart Showing an Out-of-Control Process Caused by At Least Seven Consecutive Points Located Above or Below CL

causal causes requires several different steps, one of which is to recognize patterns in processes. Each rule of a control chart is a pattern found on a control chart and serves to identify the existence of sources of variation. A control chart's zone determines the behavior of certain of these patterns.

Figure 4 presents the third condition of the out-of-control process. It has six or seven points in a row which keeps going up or down. This condition represents a process that is trending in one direction. A basic condition of thumb is when a run chart exhibits seven or eight points successively up or down, a trend is clearly present in the data and needs process improvement. This condition does not care whether the consecutive points are above, below, or crossing the median.

If the process is out of control, it is necessary to investigate and analyze the cause (Alevizakos, Chatterjee, & Koukouvinos, 2021). Meanwhile, the fourth condition is that the process is in control if the sample points are within the control limits. An example of this fourth condition is presented in Figure 5.

Let w be a statistical sample that measures some quality characteristic, and u_w be the mean of w and the standard deviation of w denoted σ_w. Then, the control limit (UCL, CL, and LCL) are expressed In Equations (2)−(4) (Anwar, Aslam, Zaman, & Riaz, 2021). It shows k as the distance of the control limit from CL.

$$UCL = \mu_w + k\sigma_w \tag{2}$$

$$CL = \mu_w \tag{3}$$

$$LCL = \mu_w - k\sigma_w \tag{4}$$

Control charts are developed according to these principles called the Shewhart control chart (Nguyen, Nguyen, Tran, & Ho, 2019). The Shewhart control chart can be divided into variable and attribute control charts based on quality characteristics. The attribute control chart is used if the quality characteristics of the sample are stated in the appropriate form or not according to specifications (Zhou, Cheng, & Zheng, 2019). Besides that, the attribute control chart is used

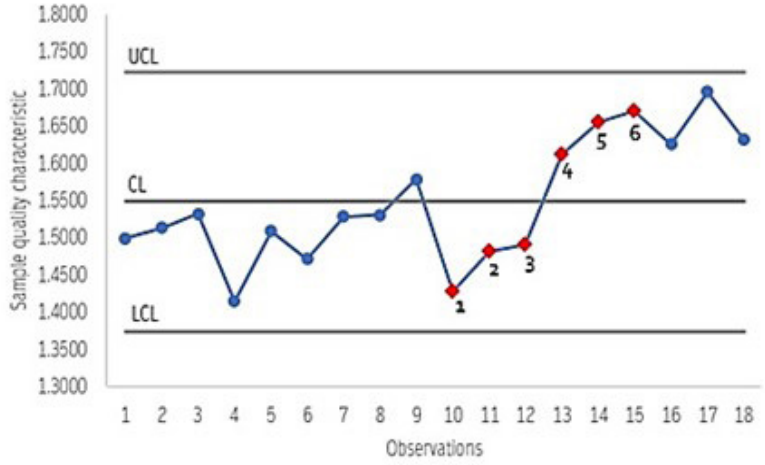

Figure 4 The Example of a Control Chart Showing an Out-of-Control Process Caused by Six or Seven Points in a Row that Keep Going Up or Down

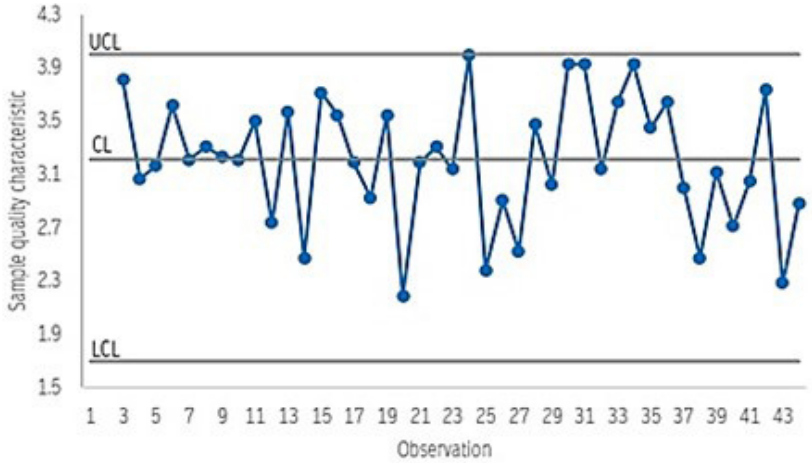

Figure 5 The Example of a Control Chart Showing the Process in Control

to control the process using attribute data, such as the number of units that fail to produce (reject), the number of absent employees, and the number of defective components. Attribute data only have two values or choices: yes or no, present or absent, and suitable or defective components. Attribute data are obtained from non-conformance units with the specified attribute specifications. Attributes in quality control indicate quality characteristics by specifications. Attributes are used when there are measurements that are not possible to do, such as scratches, errors, colors, or missing parts. In addition, attributes are used when measurements can be made but are not made for reasons of time, cost, or need. Statistical process quality control for attribute data is used instead of statistical process quality control for variable data. The types of attribute control charts include *np*-Chart, *p*-Chart, *c*-Chart, and *u*-Chart (Aslam, 2019).

A variable control chart is used if the characteristics of sample quality can be measured and expressed in numbers, such as measurements of temperature, weight, and volume (Mandal, Roychowdhury, & Bhattacharya, 2021). There are three types of variable control charts: $\bar{X} - R$ (for small sample size or ≤ 10), $\bar{X} - s$ (for large sample size or > 10), and IMR for single sample size (Fan et al., 2021). According to Ahsan, Mashuri, Kuswanto, Prastyo, and Khusna (2018), $\bar{X} - R$ is a variable control chart where data are collected in each observation in the form of a subgroup of 2–10. This control chart is used to know the stability of a process if the data are variable. Each datum is collected in subgroups of 2–10, such as tablet hardness, active ingredient content, and dissolution rate.

Meanwhile, $\bar{X} - s$ is a variable control chart where data are collected for each observation in subgroups of 10 or more. This control chart is used to know the stability of a process if the data are variable, and each data collected is in the form of a subgroup whose size is more than 10. For example, it can be the diameter of the ampoule cut and tablet weight (Srinivasa Rao, Raza, Aslam, AL-Marshadi, & Jun, 2019).

IMR is a variable control chart where the data collected in each observation are one/single. It is called moving range because the range is obtained from moving data from one test to the next (Leonov, Shkaruba, Vergazova, Golinitskiy, & Antonova, 2020). This control chart is used in conditions to know the stability of a process if the data are variable, and each data collected is individual data. Generally, it is used in industries running 24 hours, such as cement and chemical fertilizers. The results are relatively homogeneous in each data collection, so it is enough to take one sample. It can also be used when only a small amount of testing can be performed due to cost factors or available minimal production, such as disintegration time, pH, viscosity, and density (Arshad, Azam, Aslam, & Jun, 2022).

Time series control charts can be used for time series models because the essential step in constructing this control chart is taking the residuals from the model. One of the basic assumptions of using a control chart is that the data must be independent (Costa & Fichera, 2021). There is no autocorrelation, so the control chart in the time series analysis is constructed on the residual model. The control chart that fulfills these conditions is the IMR control chart, in which the calculations can be seen in Table 1 (Hieu, Chou, Fang, & Hoang, 2018).

In Table 1, $\overline{MR}$ is the average of MR_i with $MR_i = |x_i - x_{i-1}|$, x_i as i-th observation, and d_2 and D_4 are constant. The following steps construct a time series control chart (Nguyen et al., 2021). The first step calculates the moving range and average moving range. The second step computes UCL, CL, and LCL for individual and moving range plots (based on Table 1). The third step constructs an IMR control chart consisting of individual and moving range plots. If the residuals are out of control, the model obtained has poor accuracy. Meanwhile, if the residuals are in control, the time series model obtained can be used for prediction, or it can be said that the model has good accuracy (Oprime, Lizarelli, Pimenta, & Achcar, 2019).

Table 1 The Formula of IMR Control Chart

Individual plot	Moving Range plot
$UCL = 3 \times \frac{\overline{MR}}{d_2}$	$UCL = D_4 \times \overline{MR}$
$CL = 0$	$CL = \overline{MR}$
$LCL = -3 \times \frac{\overline{MR}}{d_2}$	$LCL = 0$

III. RESULTS AND DISCUSSIONS

The research uses data on the number of participants who have been vaccinated in Pontianak City from June 20, 2021, to November 19, 2021. The data collection period is daily, with 153 observations. The data are obtained from the Department of Health in Pontianak City. The average number of people in Pontianak City vaccinated during that period is 3259,11 (see Figure 6). It makes Pontianak City the city with the highest vaccination achievement rate in West Kalimantan. With this achievement, Pontianak City also focuses on accelerating the achievement of the second dose of vaccination. The government carries out a vaccination program to realize communal immunity against COVID-19 and control the transmission of the disease caused by the SARS-CoV-2 type coronavirus. The COVID-19 vaccination is available at the vaccination service center. Vaccination services are also provided in public places to make it easier for residents to access services. The distribution of vaccines is carried out by the government continuously. It was recorded that in September 2021, the government distributed 56,1 million doses of COVID-19 vaccines. This number was more than in August 2021, with only 42,8 million doses. Additionally, the highest number of vaccine participants also occurred in September 202, with 8.180. Figure 6 presents descriptive statistics from the data used, including time series plots which also display the maximum, minimum, and average values.

The histogram of the total vaccine participants is presented in Figure 7. Based on the histogram presented in Figure 7, it is possible to deduce that the data distribution is not symmetrical. It indicates that the mean and median values are not the same thing at all. In addition, looking at the histogram, it can be observed that the group receiving the most vaccinations consists of forty individuals.

Based on Figure 8, the shape of the data distribution is skewed to the right so that the right side contains more data than the left. It means that data are accumulated more in smaller values. In addition, no outliers are identified in the data. It can be seen in the Box Plot in Figure 8. Hence, the analysis can be continued using time series analysis.

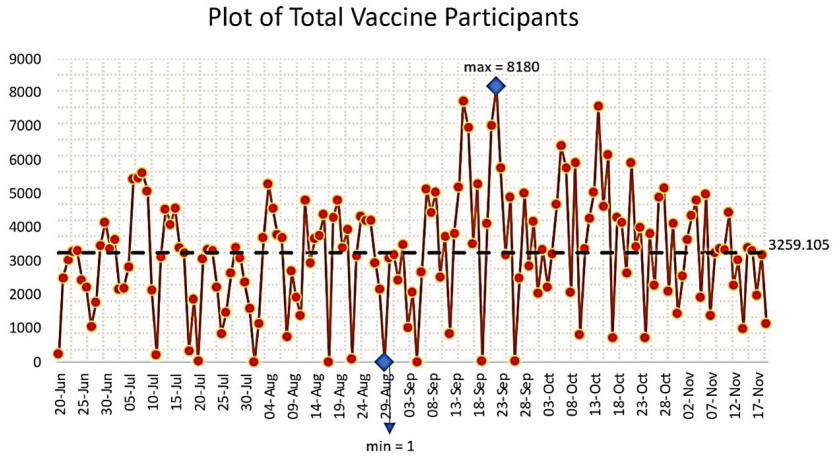

Figure 6 Time Series Plot of Total Vaccine Participants

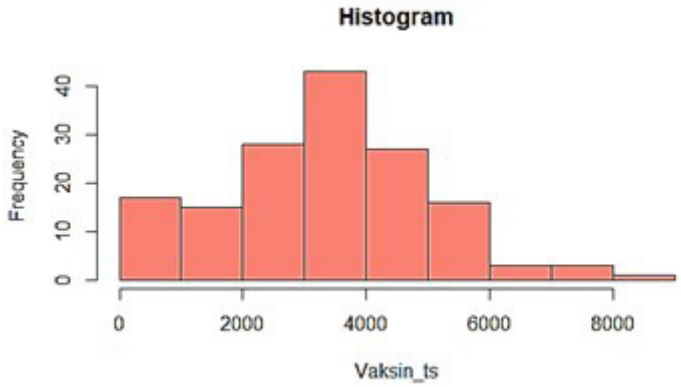

Figure 7 Histogram of Total Vaccine Participants

The time series plot of the data can be seen in Figure 6. Based on the time series plot, the data are stationary in the mean and variance because the values are constant or independent of time. After the data are stationary, it can proceed to the three main steps in time series modeling. First, it identifies the model using ACF and PACF plots, as presented in Figures 9 and 10. The ACF value at the cut-off after the second lag, along with the PACF plot, may be determined to generate a damped sine wave by referring to the ACF plot and the PACF plot, respectively. These plots are based on the ACF plot. In addition, it is possible to deduce that the PACF ceases to exist after the fifth lag, and the ACF instead takes the shape of a dampened sine wave. The results of this interpretation enable the development of other possible models.

Three possible models fit the ACF and PACF plots: ARIMA (5, 0, 2), ARIMA (5, 0, 0), and ARIMA (0, 0, 2). The second stage is parameter estimation. Table 2 presents the parameter estimation results for each model using MLE. Based on Table 2, the ARIMA (5, 0, 2) model has the smallest AIC and RMSE values. Hence, ARIMA (5, 0, 2) is the best model. Following the ARIMA (5, 0, 2) model, it has an equation as follows. It has Y_t as the total vaccine participant at time t and e_t as the noise process at time t.

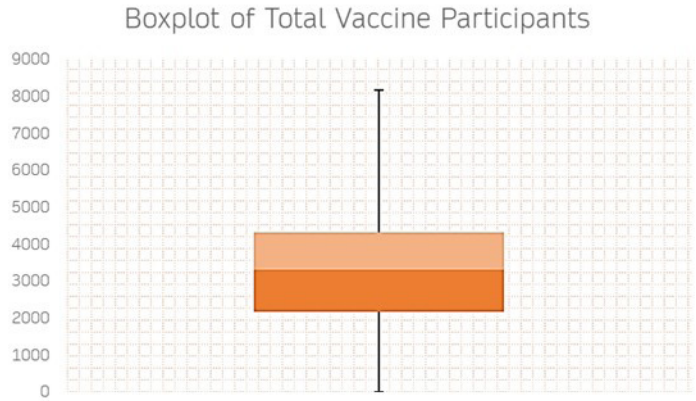

Figure 8 Box Plot of Total Vaccine Participants

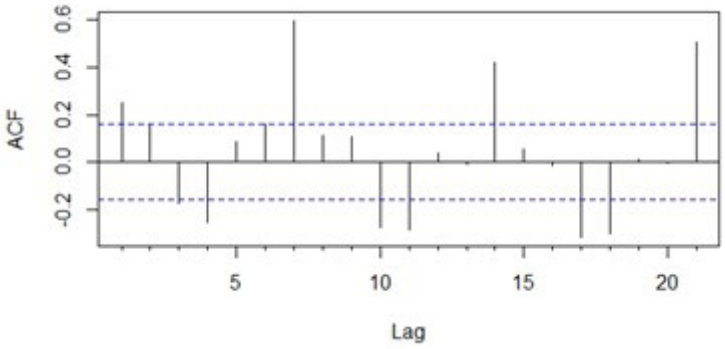

Figure 9 ACF plot of total vaccine participants

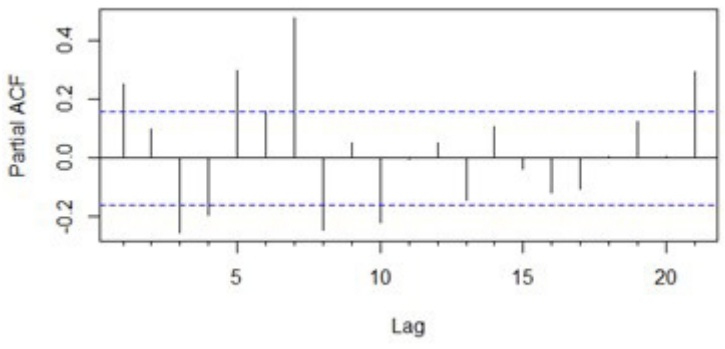

Figure 10 PACF Plot of total Vaccine Participants

$$Y_t = 3194{,}6247 + 0{,}2824Y_{t-1} + 0{,}7161Y_{t-2} - 0{,}4040Y_{t-3}$$
$$-0{,}2968Y_{t-4} + 0{,}5033Y_{t-5} - 0{,}0721e_{t-1} + 0{,}5900e_{t-2} + e_t \quad (5)$$

The final step in time series modeling is diagnostic checking. The correlation can be ignored if the residual autocorrelation is within the 5% significance limit. There is no correlation between the time lags. Figures 11−13 show the visualization of the residuals. Based on Figure 11, the ARIMA (5, 0, 2) model has an uncorrelated residual between time lags. Meanwhile, the normality of residuals can be seen in the histogram and Q-Q plot, presented in Figures 12 and 13, respectively. Figures 12 and 13 show that the residuals are normally distributed. So, it can be concluded that the ARIMA (5, 0, 2) model fulfills the white noise assumption.

Table 2 Parameter Estimation and Error Measures for ARIMA (5, 0, 2), ARIMA (5, 0, 0), and ARIMA (0, 0, 2) Models

Model	Parameter Estimation		AIC	RMSE
ARIMA (5, 0, 2)	ϕ_1	0,2824	2666,98	1382,672
	ϕ_2	0,7161		
	ϕ_3	-0,4040		
	ϕ_4	-0,2968		
	ϕ_5	0,5033		
	θ_1	0,0721		
	θ_2	-0,5900		
	μ	3194,6247		
ARIMA (5, 0, 0)	ϕ_1	0,2651	2683,15	1481,404
	ϕ_2	0,2530		
	ϕ_3	-0,2622		
	ϕ_4	-0,2570		
	ϕ_5	0,3027		
	μ	3247.8982		
ARIMA (0, 0, 2)	θ_1	0,2467	2698,14	1589,792
	θ_2	0,3582		
	μ	3240,8257		

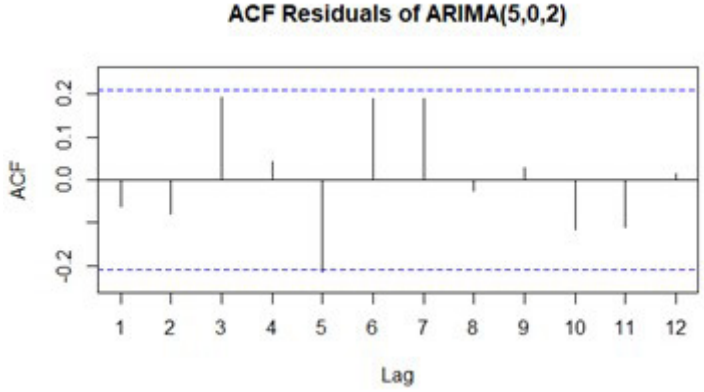

Figure 11 ACF Plot of Residuals in ARIMA (5, 0, 2) Model

Based on the steps of time series modeling, it is found that the ARIMA (5, 0, 2) is the best model. However, one more step is needed to verify the best model. This verification stage is carried out to measure the model's accuracy on the best model that has been obtained. Measuring the model's accuracy on the best model is very important in predicting the following several periods. The accuracy of a model is seen from the residual value. The smaller the residual is, the more accurate the model is. Alternatively, the residual is in a statistically controlled state. Such a situation can be seen in the control chart. The ARIMA model residuals are used to construct the IMR control chart, presented in Figures 14 and 15.

Based on Figure 14, all points are within the control limits. However, some points are outside the control limits (see Figure 15). So, it can be concluded that the process is said to be out of control. As a result, the ARIMA (5, 0, 2) model is not accurate for making predictions.

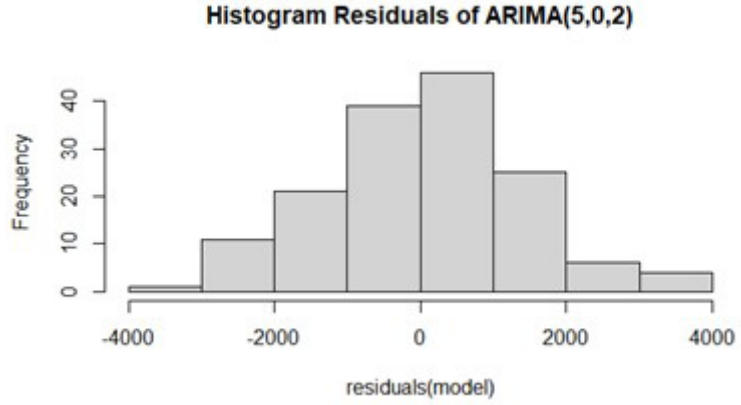

Figure 12 Histogram of Residuals in ARIMA (5, 0, 2) Model

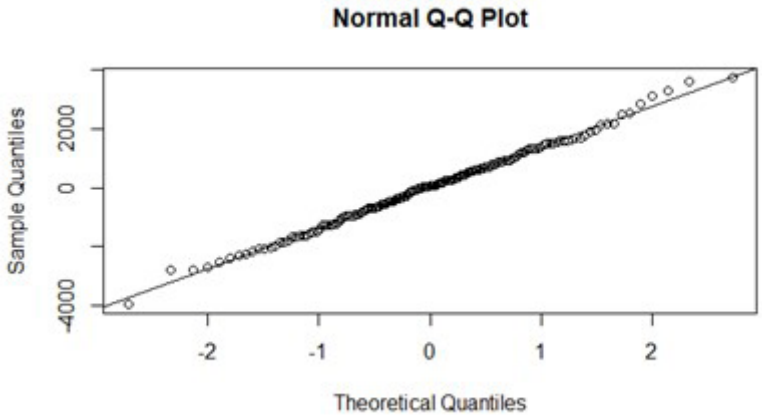

Figure 13 Normal Q-Q Plot of Residuals in ARIMA (5, 0, 2) Model

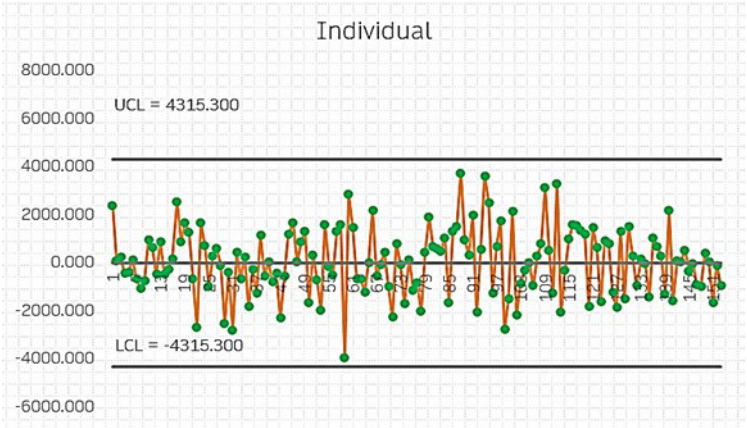

Figure 14 Individual Plot for Residuals in ARIMA (5, 0, 2) Model

To be more convincing, Figure 16 presents a plot of the original data versus the fitted value of the ARIMA model (5, 0, 2). It can be seen that the model obtained cannot capture some of the peaks. However, it can capture the essence of the data well. The plot of the estimated data does not follow the original data. Instead, it results in a sizeable residual value, which is the difference between the actual and estimated results.

IV. CONCLUSIONS

Based on the steps of time series data modeling on vaccine participant data in Pontianak City, it is found that the ARIMA model (5, 0, 2) is the best because it fulfills the white noise assumption. However, when verified using the IMR control chart, it can be concluded that the ARIMA model (5, 0, 2) is not accurate for predicting the following several periods. It is because the residuals in the model are out of control. Hence, the best model obtained at the time series modeling step is not necessarily an accurate model for predicting some time in the future. It is necessary to carry out a verification stage on the best model, one of which is by constructing a control chart from the best residual model.

Based on the research, it is also found that determining the best model in time series analysis is not sufficient only to reach the diagnostic test stage (for example, the residuals fulfill the white noise assumption). It is necessary to verify by constructing a control chart on the best model to analyze whether the residuals from the model are statistically controlled (in control) or not (out of control). Controlled residual or not is used to see the accuracy of the time series model in predictions. If the residual is in control, the time series model obtained is accurate and can be used for prediction. Nevertheless, if the residual is out of control, the time series model obtained is inaccurate and needs to be evaluated.

The research results can be considered or used as a reference point when deciding which time series model is preferred. In addition, the results can also contribute to parties related to COVID-19 vaccination in Pontianak City. However, the research only analyzes the accuracy of the time series model using a control chart. It can be generalized in other models, such as spatial and space-time analyses. Therefore, it is expected that the research will not stop here.

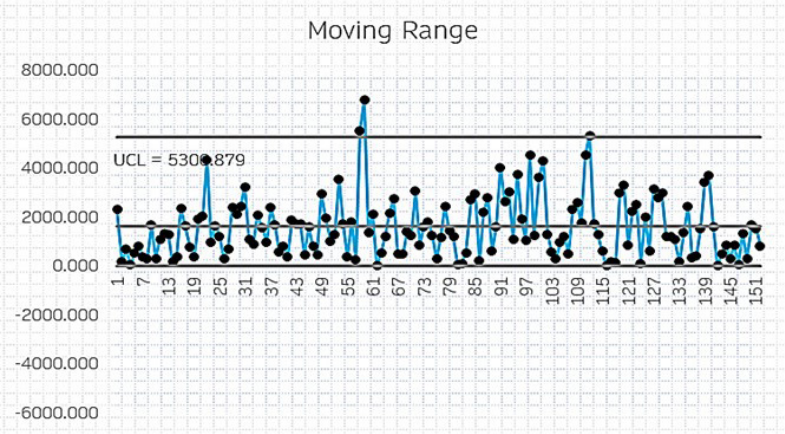

Figure 15 Moving Range Plot for Residuals in ARIMA (5, 0, 2) Model

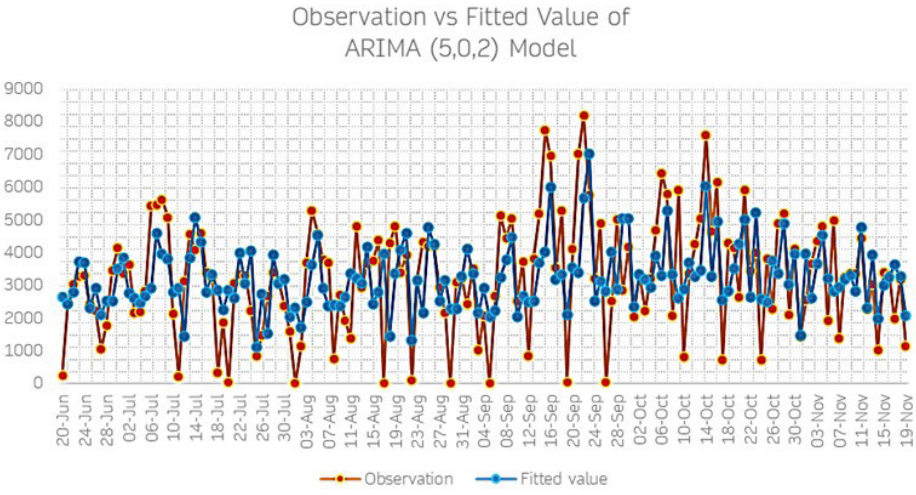

Figure 16 Plot of Observation Versus Fitted Values Based on ARIMA (5, 0, 2) Model

REFERENCES

Abanda, A., Mori, U., & Lozano, J. A. (2019). A review on distance based time series classification. *Data Mining and Knowledge Discovery*, *33*, 378–412. https://doi.org/10.1007/s10618-018-0596-4

Ahsan, M., Mashuri, M., Kuswanto, H., Prastyo, D. D., & Khusna, H. (2018). Multivariate control chart based on PCA mix for variable and attribute quality characteristics. *Production & Manufacturing Research*, *6*(1), 364–384. https://doi.org/10.1080/21693277.2018.1517055

Abbasi, S. A., Abid, M., Riaz, M., & Nazir, H. Z. (2020). Performance evaluation of moving average-based EWMA chart for exponentially distributed process. *Journal of the Chinese Institute of Engineers*, *43*(4), 365–372. https://doi.org/10.1080/02533839.2020.1719893

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*(August), 1–17. https://doi.org/10.1016/j.rinp.2021.104509

Alevizakos, V., Chatterjee, K., & Koukouvinos, C. (2021). The triple exponentially weighted moving average control chart. *Quality Technology & Quantitative Management*, *18*(3), 326–354. https://doi.org/10.1080/16843703.2020.1809063

Anwar, S. M., Aslam, M., Zaman, B., & Riaz, M. (2021). Mixed memory control chart based on auxiliary information for simultaneously monitoring of process parameters: An application in glass field. *Computers & Industrial Engineering*, *156*(June), 1–18. https://doi.org/10.1016/j.cie.2021.107284

Arshad, A., Azam, M., Aslam, M., & Jun, C. H. (2022). A resubmission-based variable control chart. *Communications in Statistics-Theory and Methods*, 1–13. https://doi.org/10.1080/03610926.2022

Arooj, H., & Malik, K. I. (2022). Double moving average control chart for autocorrelated data. *Scientific Inquiry and Review*, *6*(2), 1–20. https://doi.org/10.32350/sir.62.01

Aslam, M. (2019). Attribute control chart using the repetitive sampling under neutrosophic system. *IEEE Access*, *7*, 15367–15374. https://doi.org/10.1109/ACCESS.2019.2895162

Bagmar, M. S. H., & Khudri, M. M. (2021). Application of box-Jenkins models for forecasting drought in north-western part of Bangladesh. *Environmental Engineering Research, 26*(3), 1–7. https://doi.org/10.4491/eer.2020.294

Benvenuto, D., Giovanetti, M., Vassallo, L., Angeletti, S., & Ciccozzi, M. (2020). Application of the ARIMA model on the COVID-2019 epidemic dataset. *Data in Brief*, *29*(April), 1–4. https://doi.org/10.1016/j.dib.2020.105340

Costa, A., & Fichera, S. (2021). Economic-statistical design of adaptive ARMA control chart for autocorrelated data. *Journal of Statistical Computation and Simulation*, *91*(3), 623–647. https://doi.org/10.1080/00949655.2020.1825716

Dinas Komunikasi dan Informatika Provinsi Kalimantan Barat. (2021). *Distribusi vaksin COVID-19 di Kalimantan Barat*. Retrieved from https://diskominfo.kalbarprov.go.id/17/01/2021/video-distribusi-vaksin-covid-19-di-kalimantan-barat/

Fan, W., Xue, H., Yi, C., & Xu, Z. (2021). TQWT-assisted statistical process control method for condition monitoring and fault diagnosis of bearings in high-speed rail. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, *235*(2), 230–240. https://doi.org/10.1177/1748006X20958321

Farimani, N. M., Parsafar, P., & Mohammadi, S. (2022). Evaluation performance of time series methods in demand forecasting: Box-Jenkins vs artificial neural network (Case study: Automotive parts industry). *Journal of Statistical Computation and Simulation, 92*, 3639–3658. https://doi.org/10.1080/00949655.2022.2077945

García, E., Peñabaena-Niebles, R., Jubiz-Diaz, M., & Perez-Tafur, A. (2022). Concurrent control chart pattern recognition: A systematic review. *Mathematics*, *10*(6), 1–31. https://doi.org/10.3390/math10060934

Golilarz, N. A., Addeh, A., Gao, H., Ali, L., Roshandeh, A. M., Munir, H. M., & Khan, R. U. (2019). A new automatic method for control chart patterns recognition based on ConvNet and Harris Hawks meta heuristic optimization algorithm. *IEEE Access*, *7*, 149398–149405. https://doi.org/10.1109/ACCESS.2019.2945596

Herdiani, E. T., Fandrilla, G., & Sunusi, N. (2018). Modified Exponential Weighted Moving Average (EWMA) control chart on autocorrelation data. *Journal of Physics: Conference Series*, *979*, 1–7. https://doi.org/10.1088/1742-6596/979/1/012097

Hieu, H. T., Chou, T. Y., Fang, Y. M., & Hoang, T. V. (2018). Statistical process control methods for detecting outliers in GPS time series data. *International Refereed Journal of Engineering and Science (IRJES), 7*(5), 8–15.

Jafarian-Namin, S., Fallahnezhad, M. S., Tavakkoli-Moghaddam, R., Salmasnia, A., & Fatemi Ghomi, S. M. T. (2021). An integrated quality, maintenance and production model based on the delayed monitoring under the ARMA control chart. *Journal of Statistical Computation and Simulation*, *91*(13), 2645–2669. https://doi.org/10.1080/00949655.2021.1904241

Katoch, R., & Sidhu, A. (2021). An application of ARIMA model to forecast the dynamics of COVID-19 epidemic in India. *Global Business Review*. https://doi.org/10.1177/0972150920988653

Kementerian Kesehatan Republik Indonesia. (2021). *Frequently Asked Questions (FAQ) seputar pelaksanaan vaksinasi COVID-19.* Retrieved from https://kesmas.kemkes.go.id/assets/uploads/contents/others/FAQ_VAKSINASI_COVID__call_center.pdf

Koutras, M. V., & Triantafyllou, I. S. (2020). Recent advances on univariate distribution-free Shewhart-type control charts. In M. Koutras & I. Triantafyllou (Eds.), *Distribution-free methods for statistical*

process monitoring and control (pp. 1–56). Springer. https://doi.org/10.1007/978-3-030-25081-2_1

Leonov, O. A., Shkaruba, N. Z., Vergazova, Y. G., Golinitskiy, P. V., & Antonova, U. Y. (2020). Quality control in the machining of cylinder liners at repair enterprises. *Russian Engineering Research*, *40*, 726–731. https://doi.org/10.3103/S1068798X20090105

Mandal, S., Roychowdhury, T., & Bhattacharya, A. (2021). Pattern of genomic variation in SARS-CoV-2 (COVID-19) suggests restricted nonrandom changes: Analysis using Shewhart control charts. *Journal of Biosciences*, *46*(1), 1–7. https://doi.org/10.1007/s12038-020-00131-5

Nguyen, H. D., Nguyen, Q. T., Tran, K. P., & Ho, D. P. (2019). On the performance of VSI Shewhart control chart for monitoring the coefficient of variation in the presence of measurement errors. *The International Journal of Advanced Manufacturing Technology*, *104*, 211–243. https://doi.org/10.1007/s00170-019-03352-7

Nguyen, H. D., Nadi, A. A., Tran, K. P., Castagliola, P., Celano, G., & Tran, K. D. (2021). *The effect of autocorrelation on the Shewhart-RZ control chart*. Retrieved from http://arxiv.org/abs/2108.05239

Oprime, P. C., Lizarelli, F. L., Pimenta, M. L., & Achcar, J. A. (2019). Acceptance X-bar chart considering the sample distribution of capability indices, $\hat{C}_p$ and $\hat{C}_{pk}$: A practical and economical approach. *International Journal of Quality & Reliability Management*, *36*(6), 875–894. https://doi.org/10.1108/IJQRM-11-2017-0239

Sahu, B. K. (2021). *Time series modelling in earth sciences*. CRC Press. https://doi.org/10.1201/9781003211280

Semenoglou, A. A., Spiliotis, E., Makridakis, S., & Assimakopoulos, V. (2021). Investigating the accuracy of cross-learning time series forecasting methods. *International Journal of Forecasting*, *37*(3), 1072–1084. https://doi.org/10.1016/j.ijforecast.2020.11.009

Srinivasa Rao, G., Raza, M. A., Aslam, M., AL-Marshadi, A. H., & Jun, C. H. (2019). A variable control chart based on process capability index under generalized multiple dependent state sampling. *IEEE Access*, *7*, 34031–34044. https://doi.org/10.1109/ACCESS.2019.2903892

Wang, Y., Yu, C., Hou, J., Chu, S., Zhang, Y., & Zhu, Y. (2022). ARIMA model and few-shot learning for vehicle speed time series analysis and prediction. *Computational Intelligence and Neuroscience*, *2022*, 1–9. https://doi.org/10.1155/2022/2526821

World Health Organization (WHO). (n.d.). *WHO Coronavirus (COVID-19) dashboard.* Retrieved from https://covid19.who.int/

Zhou, W., Cheng, C., & Zheng, Z. (2019). Optimal design of an attribute control chart for monitoring the mean of autocorrelated processes. *Computers & Industrial Engineering*, *137*, 1–15. https://doi.org/10.1016/j.cie.2019.106081