Hindawi
Discrete Dynamics in Nature and Society
Volume 2023, Article ID 6285328, 12 pages
https://doi.org/10.1155/2023/6285328

Research Article

Forecasting the Effectiveness of COVID-19 Vaccination Using Vector Autoregressive with an Exogenous Variable: On the Cases of COVID-19 in Indonesia

Sukono Sukono,[1] **Riza Andrian Ibrahim**,[1] **Riaman Riaman**,[1] **Elis Hertini**,[1] **Yuyun Hidayat**,[1] **and Jumadil Saputra**[2]

[1]*Universitas Padjadjaran, Bandung, Indonesia*
[2]*Universiti Malaysia Terengganu, Kuala Terengganu, Terengganu, Malaysia*

Correspondence should be addressed to Jumadil Saputra; jumadil.saputra@umt.edu.my

Received 11 July 2022; Revised 16 December 2022; Accepted 17 February 2023; Published 14 March 2023

Academic Editor: Fabio Tramontana

This study aims to forecast the COVID-19 spread in Indonesia involving vaccination factors using vector autoregressive with exogenous variables (VARX). The COVID-19 spread represented by active, recovered, and death case rate indicators acts as endogenous variables, while the COVID-19 vaccination represented by second-dose vaccination rates acts as exogenous variables. Because the sum of three COVID-19 spread indicators in one day is one, only two indicators with the highest correlation rates are involved in VARX modelling. The other indicator is practically projected by subtracting one from the sum of two indicator projection results. Based on the analysis results, the active and recovered case rates are two indicators chosen in VARX modelling. Using Akaike information criterion, the most suitable VARX model to project the case and recovered case rates are VARX (7, 1). This model is expected to help the Indonesian government project the COVID-19 spread in Indonesia.

1. Introduction

COVID-19 has spread in Indonesia for over a year and a half since 11 March 2020. The spread of COVID-19 has continued to record terrible records. Based on data from the Ministry of Health of the Republic of Indonesia, on 23 June 2021, the number of daily confirmed positive cases of COVID-19 in Indonesia recorded a record high, namely 15,308 new cases. This record continued to be broken in the days that followed. The highest record recorded until 7 November 2021 was 56,757 cases, set on 15 July 2021. Then, the highest number of daily deaths due to COVID-19 recorded until 7 November 2021 was 2,069 cases, where the incident occurred on 21 July 2021. The Indonesian government has made various policies to reduce COVID-19 spread rates. One of the main policies is the National Vaccination Program [1–3]. In the first two months after the start of the policy on 13 January 2021, this policy showed positive developments toward the COVID-19 spread [4]. Active and death case rates of COVID-19 tend to decrease, while COVID-19 recovered case rates increase. As of 7 November 2021, active, recovered, and death case rates of COVID-19 are 0.2548%, 96.3662%, and 3.3790%, respectively. It further gives the Indonesian government a sense of optimism about reducing the COVID-19 spread through vaccination [5].

Knowledge of models that can assist this projection process is needed to project how many vaccination rates must be achieved every day to realize the desired COVID-19 spread rates. Therefore, this study aims to design a model for projecting the COVID-19 spread rates by involving the COVID-19 vaccination in Indonesia. The model used is vector autoregressive with exogenous variables (VARX). This model is chosen because it can explain the causal relationship between endogenous variables, where exogenous variables influence endogenous variables. The COVID-19 spread rates

can be measured using various indicators. This study uses three indicators: active, recovered, and death case rates. The sum of three indicators in one day is one. So, to simplify the projection, only two indicators are involved in VARX modelling. The projection of the other indicator is made by subtracting one from the projection results of the two indicators modelled by VARX. Then, the second-dose COVID-19 vaccination rates represent the COVID-19 vaccination factor against the Indonesian population. It is done because people in Indonesia have been wholly vaccinated if they have been vaccinated twice [6]. This research is expected to help the Indonesian government project vaccination rates the next day to control the COVID-19 spread in Indonesia.

2. Literature Review

Forecasting is closely related to time series analysis [7, 8]. Therefore, articles on modelling and forecasting the COVID-19 spread in Indonesia using the concept of time series analysis are discussed briefly in this section. The articles discussed are English articles indexed in the Scopus database. Kirana and Bhawiyuga [9] modelled and forecasted the COVID-19 spread using naive methods. The COVID-19 spread used is the number of monthly COVID-19 cases. Their model's mean absolute percentage error (MAPE) is 15.85%. On the basis of the accurate classification of the model, Wei [10] stated that this model is suitable for forecasting the number of monthly COVID-19 cases in the next month. Then, Djakaria and Saleh [11] modelled and forecasted the COVID-19 spread in Gorontalo Province, Indonesia, using the Holt-Winters smoothing method. The cumulative number of COVID-19 cases is used as the COVID-19 spread indicator. The MAPE of the model obtained is 6.14%, so the model is perfect for forecasting the cumulative number of COVID-19 cases on the next day based on the model's accurate classification by Wei [10].

Then, Aditya Satrio et al. [12] analyzed and forecasted the COVID-19 spread in Indonesia, where the COVID-19 spread indicators used are the cumulative number of cases, recovered, and death of COVID-19. Each indicator is modelled separately using the autoregressive integrated moving average (ARIMA) model. Using the same method, Fadly and Sari [13] modelled and forecasted the number of funerals carried out in DKI Jakarta Province, Indonesia, since COVID-19 hit. The articles described in the previous paragraph use univariate time series analysis, whereas multivariate time series analysis can also be used. Fitriani et al. [14] simultaneously model and forecast the number of daily COVID-19-positive cases in Indonesia and Singapore. The multivariate time series model used is vector autoregressive integrated (VARI). The MAPE obtained in each place was 32.74% and 37%, respectively. The MAPE values are enormous, so based on the model's accurate classification by Wei [10], the model is unsuitable for forecasting. Then, Meimela et al. [15] modelled and forecasted the COVID-19 spread in Indonesia using the vector autoregressive integrated moving average (VARIMA). The COVID-19 spread indicators used are the daily positive and death cases. VARIMA (1, 1, 1) is the best model got.

Also, Akbar et al. [16] analyzed the relationship between COVID-19 active case rates assumed to be influenced by population growth rates. Bandung City and Purwakarta Regency are the two areas considered, whereas the population growth rate in West Java Province represents the population growth rate. The vector autoregressive integrated with exogenous variables (VARX) describes the relationship. The MAPE obtained in each place is 1% and 11.8%, respectively. It aligns with Wei's accurate classification model (2006); the model is perfect for forecasting the COVID-19 spread in Bandung City, while the other is only good. As explained in the previous paragraphs, no articles in the Scopus database have explored modelling and forecasting active, recovered, and death case rates of COVID-19 in Indonesia, considering vaccination factors using the VARX model. Therefore, this is a good opportunity to perform such modelling and forecasting.

3. Materials and Methods

3.1. Materials. The data used in this study are (1) active, recovered, and death case rates of COVID-19 data and (2) second-dose COVID-19 vaccination rates data. The period considered is from 27 January 2021, when the second-dose vaccination program starts, until 7 November 2021. All data are obtained from the official website of the Science and Technology Index, National Research and Innovation Agency, Ministry of Research and Technology of the Republic of Indonesia.

3.2. Methods

3.2.1. Vector Autoregressive with Exogenous Variables. The general form of the vector autoregressive with exogenous variables (VARX) model with the order of the endogenous variables p and the order of the endogenous variables q, VARX (p, q), is expressed as follows [17–21]:

$$\mathbf{z}_t = \boldsymbol{\mu} + \sum_{i=1}^{p} \boldsymbol{\phi}_i \mathbf{z}_{t-i} + \sum_{j=0}^{q} \boldsymbol{\theta}_j \mathbf{x}_{t-j} + \mathbf{e}_t, \tag{1}$$

where $\mathbf{z}_t$ represents the n-dimensional vector containing the endogenous variables at time t, $\phi_i, i = 1, 2, \ldots, p$ express the $(n \times n)$-order coefficient matrix of $\mathbf{z}_{t-i}$, $\mathbf{x}_t$ represents the r-dimensional vector containing the exogenous variables at time t, $\theta_j, j = 0, 1, \ldots, q$ express the $(n \times r)$-order coefficient matrix of $\mathbf{x}_{t-j}$, and $\mathbf{e}_t$ represents the n-dimensional random error vector. Equation (1) can also be reformulated as follows [22]:

$$\phi(B_z)\mathbf{z}_t = \boldsymbol{\mu} + \theta(B_x)\mathbf{x}_t + \mathbf{e}_t, \tag{2}$$

where $\phi(B_z) = \mathbf{I}_n - \sum_{i=1}^{p} \phi_i B_z^i$, $\theta(B_x) = \sum_{j=0}^{p} \theta_j B_x^j$, $B_z^i \mathbf{z}_t = \mathbf{z}_{t-i}$, and $B_x^j \mathbf{x}_t = \mathbf{x}_{t-j}$.

There are three assumptions of the VARX model presented as follows [23]:

(1) The VARX (p, q) model is stationary if the zero generators of the $\phi(v)$ and $\theta(w)$ determinants are outside the unit circle

(2) $\mathbf{e}_t$ is independent and identically multivariate-normally distributed with zero-vector mean and the constant covariance-matrix variance, and Σ. Σ is a positive semidefinite matrix expressed as follows: $\Sigma = [\sigma_{e_{k_1,t}, e_{k_2,t}}]$, where $k_1, k_2 = 1, 2, \ldots, n$

(3) $\mathbf{e}_t$ and $\mathbf{x}_t$ are independent

Many criteria can be used to determine the order of the VARX model. Lui et al. [22] stated that the order of the VARX model is determined through the Akaike information criterion (AIC). The model with the smallest AIC is the best. The AIC value of a model can be determined by the following equation [24]:

$$\text{AIC} = \ln\left(|\boldsymbol{\Sigma}|\right) + \frac{2L}{T}, \tag{3}$$

where T is the size of observation data, and L is many model parameters.

3.2.2. Augmented Dicky-Fuller Test. The unit root test equation for Augmented Dicky-Fuller (ADF) from data that have constant and trend elements is as follows [25–27]:

$$z'_{k,t} = \eta_k + \psi_k t + \beta_k z_{k,t-1} + \sum_{i=1}^{p-1} \zeta_{k,i} z'_{k,t-i} + e_{k,t}, \tag{4}$$

where $z_{k,t}$ represents observation data k at the time t, $z'_{k,t}$ represents the difference between $z_{k,t}$ and $z_{k,t-1}$, η_k is a constant, ψ_k represents the trend parameter, $\zeta_{k,i}$ is the parameter of $z'_{k,t-i}$, p represents the optimum lag, and $e_{k,t}$ is random error that is independent and identically normally distributed with zero mean and constant variance. The followings are the test hypotheses used: H_0: $\beta_k = 0$, and H_1: $\beta_k < 0$. The test statistic used is the t-ratio of the parameter estimator β_k estimated using the least square method stated as follows:

$$\tau = \frac{\widehat{\beta}_k}{s_{\widehat{\beta}_k}}, \tag{5}$$

where $\widehat{\beta}_k$ is an estimator β_k, and $s_{\widehat{\beta}_k}$ is the standard deviation of $\widehat{\beta}_k$. If τ less than the critical value, then H_0 rejected which means the data are stationary and vice versa. If H_0 is not rejected, then data must be differentiated until it is stationary [28, 29].

3.2.3. Granger Causality Test. Suppose that there are two stationary time series $z_{1,t}$ and $z_{2,t}$. In certain cases, $z_{1,t}$ can be affected by $z_{1,t-i}$ and $z_{2,t-i}$, and $z_{2,t}$ can also be affected by $z_{1,t-i}$ and $z_{2,t-i}$. This relationship is called a bi-directional causality relationship [30, 31]. The Granger causality test is a popular statistical test used to check the relationship.

The relationship between $z_{1,t}$ with $z_{1,t-i}$, which is also called a restricted relationship, is stated as follows [32]:

$$z_{1,t} = \mu_1 + \sum_{i=1}^{p} \phi_{1,i} z_{1,t-i} + e_{1,t}. \tag{6}$$

Then, the relationship between $z_{1,t}$ with $z_{1,t-i}$ and $z_{2,t-i}$, which is also called the unrestricted relationship, is stated as follows [33–35]:

$$z_{1,t} = \mu_1 + \sum_{i=1}^{p} \phi_{1,i} z_{1,t-i} + \sum_{i=1}^{p} \phi_{2,i} z_{2,t-i} + e_{1,t}. \tag{7}$$

The test hypotheses used are as follows: H_0: $\forall i, \phi_{2,i} = 0$, and H_1: $\exists i, \phi_{2,i} \neq 0$. Test statistics F determined by the following equation [36]:

$$F = \frac{\left(R^2_{ur} - R^2_r\right)(T - p - 1)}{\left(1 - R^2_{ur}\right)p}, \tag{8}$$

where R^2_{ur} and R^2_r represent the coefficient of determination in equations (6) and (7), respectively. If the F test statistic is greater than $F_{\alpha,p,T-p-1}$, then H_0 is rejected, which means $z_{1,t}$ influenced by $z_{1,t-i}$ and $z_{2,t-i}$ [37]. $z_{1,t}$ and $z_{2,t}$ have a bi-directional causality if $z_{1,t}$ influenced by $z_{1,t-i}$ and $z_{2,t-i}$, and $z_{2,t}$ influenced by $z_{1,t-i}$ and $z_{2,t-i}$ [38, 39].

3.2.4. VARX Parameter Estimation Using Maximum Likelihood (ML) Method. The VARX models with $n = 2, r = 1, p = 1,$ and $q = 0$ can be stated as follows:

$$\mathbf{y}_t = \mathbf{X}_t \boldsymbol{\beta} + \mathbf{e}_t, \tag{9}$$

where

$$\begin{aligned}
\mathbf{y}_t &= \begin{bmatrix} z_{1,t} \\ z_{2,t} \end{bmatrix}, \\
\mathbf{X}_t &= \begin{bmatrix} 1 & 0 & z_{1,t-1} & z_{2,t-1} & 0 & 0 & x_{1,t} & 0 \\ 0 & 1 & 0 & 0 & z_{1,t-1} & z_{2,t-1} & 0 & x_{1,t} \end{bmatrix}, \\
\boldsymbol{\beta} &= \begin{bmatrix} \mu_1 \\ \mu_2 \\ \phi^{(1)}_{11} \\ \phi^{(1)}_{12} \\ \phi^{(1)}_{21} \\ \phi^{(1)}_{22} \\ \theta^{(0)}_{11} \\ \theta^{(0)}_{21} \end{bmatrix}, \\
\mathbf{e}_t &= \begin{bmatrix} e_{1,2} \\ e_{2,2} \end{bmatrix}.
\end{aligned} \tag{10}$$

Based on equation (9), $\mathbf{e}_t$ can be stated as follows:

$$\mathbf{e}_t = \mathbf{y}_t - \mathbf{X}_t \boldsymbol{\beta}. \tag{11}$$

Because $\mathbf{e}_t \sim N(\mathbf{0}, \Sigma)$, then $(\mathbf{y}_t - \mathbf{X}_t\beta) \sim N(\mathbf{0}, \Sigma)$. Therefore, the distribution function of $\mathbf{e}_t$ can be expressed as follows [40, 41]:

$$f(\mathbf{e}_t|\boldsymbol{\beta}, \boldsymbol{\Sigma}) = \frac{1}{2\pi|\boldsymbol{\Sigma}|}\exp\left(-\frac{1}{2}(\mathbf{y}_t - \mathbf{X}_t\boldsymbol{\beta})^{'-1}\sum(\mathbf{y}_t - \mathbf{X}_t\boldsymbol{\beta})\right). \quad (12)$$

Vector estimator of β obtained by maximizing the likelihood function of $\mathbf{e}_t$ which is stated as follows:

$$l\left(\boldsymbol{\beta}, \boldsymbol{\Sigma}|\mathbf{e}_2, \ldots, \mathbf{e}_T\right) = \frac{1}{(2\pi|\boldsymbol{\Sigma}|)^{T-1}}\exp\left(-\frac{1}{2}\sum_{t=2}^{T}(\mathbf{y}_t - \mathbf{X}_t\boldsymbol{\beta})^{'-1}\sum(\mathbf{y}_t - \mathbf{X}_t\boldsymbol{\beta})\right). \quad (13)$$

Generally, equation (14) is transformed into a natural logarithm to facilitate maximization. This function is then referred to as the log-likelihood function, which is expressed as follows:

$$l\ln\left[l\left(\boldsymbol{\beta}, \boldsymbol{\Sigma}|\mathbf{e}_2, \ldots, \mathbf{e}_T\right)\right] = -(T-1)\ln(2\pi|\boldsymbol{\Sigma}|) - \frac{1}{2}\sum_{t=2}^{T}(\mathbf{y}_t - \mathbf{X}_t\boldsymbol{\beta})^{'-1}\sum(\mathbf{y}_t - \mathbf{X}_t\boldsymbol{\beta}). \quad (14)$$

Vector estimator of β is the zero-generator vector of the derivative of equation (15) on β. The following is the equation used to estimate the vector β:

$$\widehat{\boldsymbol{\beta}} = \left[\sum_{t=2}^{T}\mathbf{X}_t'\boldsymbol{\Sigma}^{-1}\mathbf{X}_t\right]^{-1}\sum_{t=2}^{T}\mathbf{X}_t'\boldsymbol{\Sigma}^{-1}\mathbf{y}_t. \quad (15)$$

3.2.5. Diagnostic Test. The diagnostic test is a feasibility test of the model for forecasting. This feasibility test includes checking the assumption of error and the model's mean absolute percentage error (MAPE). The model errors are tested to determine whether they are independent and identically normally distributed with zero mean and constant variance. This assumption is also known as the white noise assumption. One of the statistical tests commonly used to test whether the errors are independent or not is the Ljung-Box test. The null hypothesis (H_0) in this test is that the errors are independent, while this test's alternative hypothesis (H_1) is the opposite. The test statistic with the lag length M denoted by Q_M is determined using the following equation [42, 43]:

$$Q_M = T(T+2)\sum_{m=1}^{M}\frac{r_m^2}{T-m}, \quad (16)$$

where

$$r_m = \frac{\sum_{t=m+1}^{T}e_{k,t}e_{k,t-m}}{\sum_{t=1}^{T}e_{k,t}^2}. \quad (17)$$

Reject H_0 if $Q_M > \chi^2_{1-\alpha,df_{e_{k,t}}}$ [44, 45].

Meanwhile, the naked eye can check the assumption that the error is normally distributed with zero mean and constant variance using the quantile-quantile plot (Q-Q Plot) [46]. Suppose $F_{N_k}(e_{k,t})$ is the value of the normal cumulative distribution function with zero mean and constant variance of $e_{k,t}$. If the scatter of point pairs $(e_{k,t}, F^{-1}_{N_k}(e_{k,t}))$ is around a line with a gradient one, then $e_{k,t}$ is visually normally distributed with zero mean and constant variance. If the assumption of white noise in the model error is met and the MAPE of the model is less than 20%, then the model obtained is suitable for use in forecasting.

4. Results

The symbols used for each data are stated as follows: $z_{1,t}$ states daily COVID-19 active case rate data, $z_{2,t}$ states daily COVID-19 recovered case rate data, $z_{3,t}$ states daily COVID-19 death case rate data, and $x_{1,t}$ states second-dose COVID-19 vaccination rate data. The sum of the three indicators in one day is one, so to simplify the modelling and forecasting process, and only two indicators are used in VARX modelling. The remaining indicator is practically projected by subtracting one from the two indicators' projected results. The two indicators selected are the two indicators that have the highest correlation rates.

4.1. Data Descriptive. Descriptive statistics for each data obtained using Microsoft Excel software are presented in Table 1.

Table 1 shows the intensity of active, recovery, death, and vaccination rates from COVID-19 in Indonesia per day, each of which is different. The intensity values are 7.8119%, 89.2812%, 2.9069%, and 7.8703% per day, respectively. The intensity values of the active and death cases appear to be lower than that of vaccination. It indicates that the number of Indonesian citizens exposed and died from COVID-19 is lower than the number of people vaccinated daily. It shows that the Indonesian government is serious about handling COVID-19 via vaccination. Then, there are deviations from the active, recovery, death, and vaccination rates from

Table 1: Descriptive statistics of each data.

Descriptive statistics	$z_{1,t}$ (%)	$z_{2,t}$ (%)	$z_{3,t}$ (%)	$\mathbf{x_{1,t}}$ (%)
Average	7.8119	89.2812	2.9069	7.8703
Variance	0.2678	0.2470	0.0008	0.6484
Standard deviation	5.1750	4.9703	0.2807	8.0520

COVID-19 against each average daily. On average, the deviation values from the active, recovery, death, and vaccination rates from COVID-19 were 5.1750%, 4.9703%, 0.2807%, and 8.0520% per day, respectively. The daily deviations of the active and death rates are lower than the deviation of the vaccination rate. It indicates that Indonesian citizens are increasingly varied of COVID-19 with a high willingness to vaccinate. It prevents extreme value deviations from occurring. Finally, the variance of $z_{1,t}$, $z_{2,t}$, $z_{3,t}$, and $x_{1,t}$ is 0.2678%, 0.2470%, 0.0008%, and 0.6484%, respectively. The interpretation of the variance values is similar to the interpretation of the deviation value. Then, the correlation rates between $z_{1,t}$, $z_{2,t}$, and $z_{3,t}$ obtained using Microsoft Excel software are presented in Table 2.

Table 2 displays the correlation rate between $z_{1,t}$ and $z_{2,t}$, −0.9993, which is the highest correlation rate among the correlation rates of the other indicator data pairs. It is not surprising because the two are rationally the opposite of each other. If the active case rate increases, the recovered case rate decreases, and vice versa. This fact can be seen visually in Figure 1. Therefore, $z_{1,t}$ and $z_{2,t}$ are selected as endogenous variables. The high negative correlation between $z_{1,t}$ and $z_{2,t}$ is not a problem in VARX modelling. It is the same as regression modelling, where if the correlation rate between variables is high, the model formed can provide an excellent description of the data.

Table 3 captures the correlation rates between endogenous and exogenous variables above 0.6. It indicates that endogenous and exogenous variables are strong enough.

4.2. Data Stationary Test. In this section, $z_{1,t}$, $z_{2,t}$, and $x_{1,t}$ are checked for stationary visually and formally. Visually, the stationarity of the three data is seen from each graph, while formally, the stationary of the data is tested using the ADF test. Each data graph is observed whether it is around the average line or not. If the data graph is around the average line, then the data are stationary, and vice versa. The graphs of $z_{1,t}$, $z_{2,t}$, and $x_{1,t}$ with their respective average lines are presented in Figure 1.

Figure 1 displays that the graphs of the three data do not appear to be around the average line. Therefore, visually, the three data are not stationary. Then, the formal stationarity test of the three data is carried out using the ADF test. The significance level and optimal lag used are 0.05 and 6, respectively. The test statistics of each data are determined using equation (5). The formal stationary test result of the three data using the ADF test is presented in Table 4.

Based on Table 4, all data have a value of not less than the critical value. Therefore, H_0 on each data test is not rejected, which means that all the data are not stationary, so the differencing process using equation (3) is carried out. After

Table 2: The correlation rates between $z_{1,t}$, $z_{2,t}$, and $z_{3,t}$.

Data pairs	The correlation rates
$z_{1,t}$ and $x_{1,t}$	−0.6599
$z_{2,t}$ and $x_{1,t}$	0.6352

the data are differentiated, the next step is the three data's visual and formal stationarity tests. After the differencing process is carried out twice, all data are stationary. The graphs of $z''_{1,t}$, $z''_{2,t}$, and $x''_{1,t}$ with their respective average lines are presented in Figure 2.

Figure 2 shows the graphs of three data differentiated twice and appear to be around the average line. Therefore, visually, they are stationary. Next, the formal stationarity test is carried out using the ADF test. The significance level and optimal lag used are 0.05 and 6, respectively. The test statistics of each data are determined using equation (5). Furthermore, a formal summary of the three data stationary tests using the ADF test is presented in Table 5.

Table 5 captures the three data differentiated twice that have a τ value less than the critical value. Therefore, H_0 of each test is rejected, which means that all data are stationary.

4.3. Causality Relationship Test. A check of the existence of a bidirectional causal relationship between $''z_{1,t}$ and $''z_{2,t}$ in this study is carried out using the Granger causality test. The significance levels, optimal lag, and residual degrees of freedom used are 0.05, 7, and 256, respectively. The effect of $z''_{1,t-1}$ and $z''_{2,t-i}$ on $z''_{1,t}$ is tested first. The F value obtained is 2.6931. This value is greater than the critical value, 2.0442. Therefore, H_0 is rejected, which means that $z''_{1,t}$ is affected by $z''_{1,t-1}$ and $z''_{2,t-i}$. Next, the effect of $z''_{1,t-1}$ and $z''_{2,t-i}$ on $z''_{2,t}$ is tested. The F value obtained is 2.7514. This value is also greater than the critical value, 2.0442. Therefore, H_0 is rejected which means that $z''_{2,t}$ is affected by $z''_{1,t-1}$ and $z''_{2,t-i}$. Because $z''_{1,t}$ is affected by $z''_{1,t-1}$ and $z''_{2,t-i}$, and $z''_{2,t}$ is affected by $z''_{1,t-1}$ and $z''_{2,t-i}$, then $z''_{1,t}$ and $z''_{2,t}$ have a bidirectional causality relationship.

4.4. VARX Order Identification. The AIC value, as described in Section 2, is used as a criterion to determine the best order of endogenous and exogenous variables of the VARX model. The model with the smallest AIC value is the most suitable. The AIC value for each model is determined by equation (3). The maximum order of the endogenous and exogenous variables considered is 10. The AIC value for each model is determined using the R software. Based on the results obtained from the software, the AIC value of the VARX (7, 1), −34.2194, is the smallest AIC value among other models. Therefore, the most suitable model chosen is the VARX (7, 1).

4.5. VARX (7, 1) Parameter Estimation Using ML Method. The VARX (7, 1) parameters are estimated using the ML method, equation (14), as described in Section 2. The estimation process is carried out using the help of Microsoft

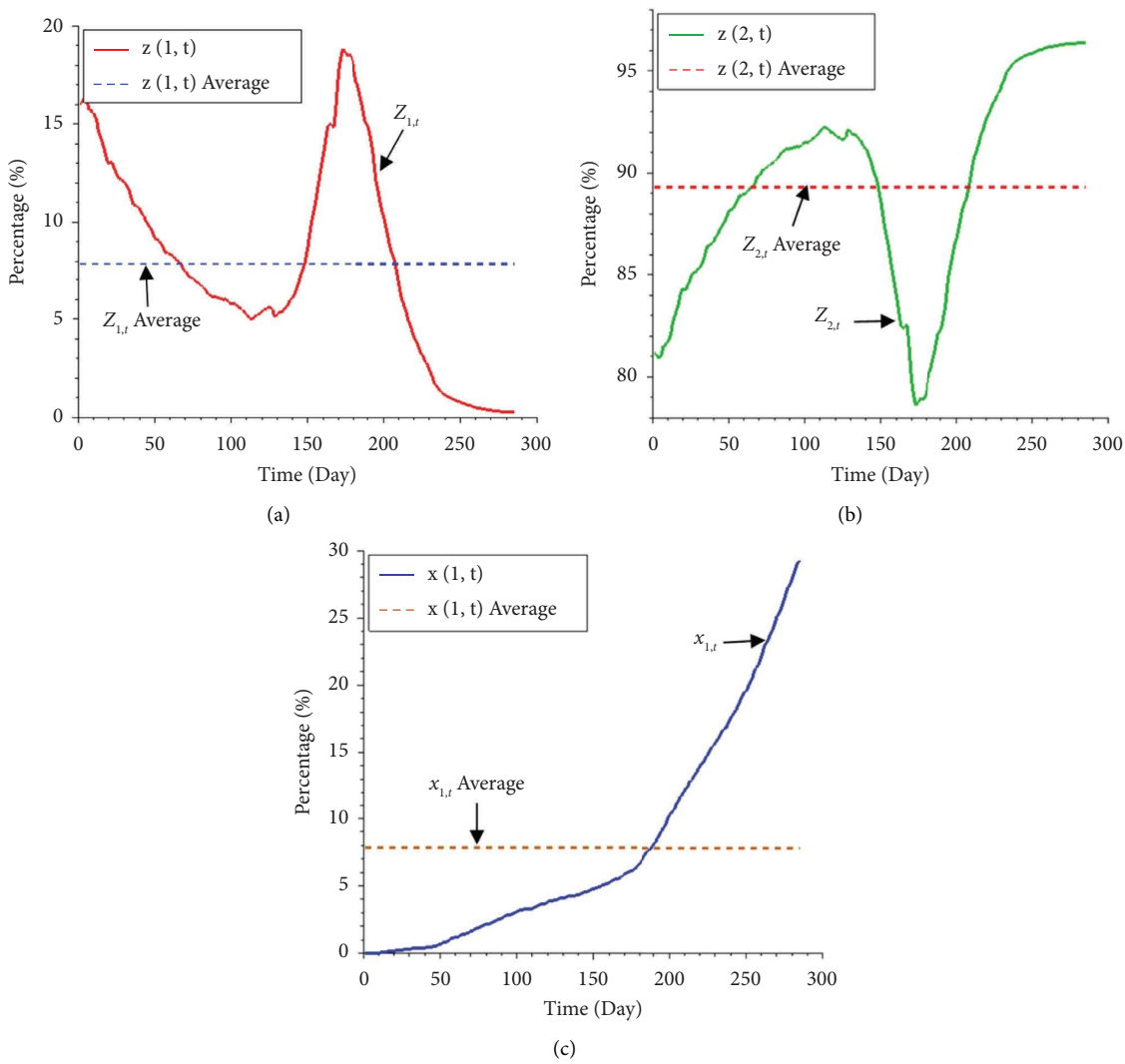

Figure 1: Graph of $z_{1,t}$ (a), $z_{2,t}$ (b), and $x_{1,t}$ (c).

Excel software. The parameter estimation results are presented in Table 6.

On the basis of the parameters obtained in Table 6, the VARX (7, 1) model is stated in equation (16).

$$\begin{aligned}\begin{bmatrix}\widehat{z}''_{1,t}\\ \widehat{z}''_{2,t}\end{bmatrix} &= \begin{bmatrix}1.7448\text{E}-05\\ -1.8174\text{E}-05\end{bmatrix} + \begin{bmatrix}0.5495 & 0.7931\\ -1.0376 & -1.2881\end{bmatrix}\begin{bmatrix}z''_{1,t-1}\\ z''_{2,t-1}\end{bmatrix} + \begin{bmatrix}1.2891 & 1.4981\\ -1.5079 & -1.7207\end{bmatrix}\begin{bmatrix}z''_{1,t-2}\\ z''_{2,t-2}\end{bmatrix}\\ &+ \begin{bmatrix}5.3873 & 5.6679\\ -5.5021 & -5.7843\end{bmatrix}\begin{bmatrix}z''_{1,t-3}\\ z''_{2,t-3}\end{bmatrix} + \begin{bmatrix}5.6900 & 6.0157\\ -5.8682 & -6.1934\end{bmatrix}\begin{bmatrix}z''_{1,t-4}\\ z''_{2,t-4}\end{bmatrix}\\ &+ \begin{bmatrix}8.9929 & 9.4176\\ -9.0363 & -9.4565\end{bmatrix}\begin{bmatrix}z''_{1,t-5}\\ z''_{2,t-5}\end{bmatrix} + \begin{bmatrix}3.2196 & 3.3846\\ -3.3553 & -3.5186\end{bmatrix}\begin{bmatrix}z''_{1,t-6}\\ z''_{2,t-6}\end{bmatrix}\\ &+ \begin{bmatrix}0.7175 & 0.6588\\ -0.8463 & -0.7909\end{bmatrix}\begin{bmatrix}z''_{1,t-7}\\ z''_{2,t-7}\end{bmatrix} + \begin{bmatrix}-0.0117\\ 0.0104\end{bmatrix}x''_{1,t} + \begin{bmatrix}0.2203\\ -0.2162\end{bmatrix}x''_{1,t-1}.\end{aligned} \tag{18}$$

Table 3: The correlation rates between endogenous and exogenous variables.

Data pairs	The correlation rates
$z_{1,t}$ and $z_{2,t}$	−0.9993
$z_{1,t}$ and $z_{3,t}$	−0.7421
$z_{2,t}$ and $z_{3,t}$	0.7162

Table 4: Data stationary test $z_{1,t}$, $z_{2,t}$, and $x_{1,t}$.

Data	$z_{1,t}$	$z_{2,t}$	$x_{1,t}$
τ	−2.7926	−2.7815	2.4323
Critical value	−3.2000	−3.2000	−3.2000

In order to be used to forecast the actual data, equation (18) is transformed to the actual data as follows:

$$\begin{aligned}\begin{bmatrix}\widehat{z}_{1,t}\\ \widehat{z}_{2,t}\end{bmatrix} &= \begin{bmatrix}1.7448\text{E}-05\\ -1.8174\text{E}-05\end{bmatrix}\begin{bmatrix}2.5495 & 0.7931\\ -1.0376 & 0.7119\end{bmatrix}\begin{bmatrix}z_{1,t-1}\\ z_{2,t-1}\end{bmatrix}+\begin{bmatrix}-0.8099 & -0.0882\\ 0.5673 & -0.1446\end{bmatrix}\begin{bmatrix}z_{1,t-2}\\ z_{2,t-2}\end{bmatrix}\\ &+\begin{bmatrix}3.3586 & 3.4649\\ -3.5239 & -3.6309\end{bmatrix}\begin{bmatrix}z_{1,t-3}\\ z_{2,t-3}\end{bmatrix}+\begin{bmatrix}-3.7954 & -3.8220\\ 3.6281 & 3.6545\end{bmatrix}\begin{bmatrix}z_{1,t-4}\\ z_{2,t-4}\end{bmatrix}\\ &+\begin{bmatrix}3.000 & 3.0540\\ -2.8021 & -2.8541\end{bmatrix}\begin{bmatrix}z_{1,t-5}\\ z_{2,t-5}\end{bmatrix}+\begin{bmatrix}-9.0762 & -9.4349\\ 8.8491 & 9.2011\end{bmatrix}\begin{bmatrix}z_{1,t-6}\\ z_{2,t-6}\end{bmatrix}\\ &+\begin{bmatrix}3.2713 & 3.3072\\ -3.1719 & -3.2103\end{bmatrix}\begin{bmatrix}z_{1,t-7}\\ z_{2,t-7}\end{bmatrix}+\begin{bmatrix}1.7845 & 2.0670\\ -1.6627 & -1.9367\end{bmatrix}\begin{bmatrix}z_{1,t-8}\\ z_{2,t-8}\end{bmatrix}\\ &+\begin{bmatrix}0.7179 & 0.6588\\ -0.8463 & -0.7909\end{bmatrix}\begin{bmatrix}z_{1,t-9}\\ z_{2,t-9}\end{bmatrix}+\begin{bmatrix}-0.0117\\ 0.0104\end{bmatrix}x_{1,t}+\begin{bmatrix}0.2437\\ -0.2370\end{bmatrix}x_{1,t-1}+\begin{bmatrix}-0.4523\\ 0.4428\end{bmatrix}x_{1,t-2}\\ &+\begin{bmatrix}0.2203\\ -0.2162\end{bmatrix}x_{1,t-3}.\end{aligned} \tag{19}$$

Meanwhile, $z_{3,t}$ can be modelled as follows:

$$\widehat{z}_{3,t} = 1 - \widehat{z}_{1,t} - \widehat{z}_{2,t}. \tag{20}$$

4.6. Diagnostic Test of VARX (7, 1). Independence of both $e_{1,t}$ and $e_{2,t}$ to time is carried out using the Ljung-Box test described in section 2. The significance levels and the degrees of freedom used are 0.05 and 7, respectively. The test statistic is determined using equation (18). The test statistics of $e_{1,t}$ and $e_{2,t}$ obtained are 3.3610 and 3.4270, respectively. The test statistic values are smaller than the critical value, 14.0671. Therefore, H_0 of each independence test of $e_{1,t}$ and $e_{2,t}$ to times is not rejected, which means that $e_{1,t}$ and $e_{2,t}$ are independent of time. Next, the normality test of the VARX (7, 1) model error is visually carried out first through the *Q-Q* plot. *Q-Q* plots for $e_{1,t}$ and $e_{2,t}$ are presented in Figure 3.

Figure 3 indicates the scattering of points in each plot appears to be around a line with one gradient. Therefore, $e_{1,t}$ and $e_{2,t}$ visually follow a normal distribution with zero mean and constant variances, namely 0.00108 and 0.00107. Based on this, the model equation (20) errors fulfill the assumption that they are independent and identically normally distributed with zero mean and constant variance. The next step is checking the error size of the model. This study measures the error size using the mean absolute percentage error (MAPE). The MAPE value of $z_{1,t}$ and $z_{2,t}$ is 1.5751% and 0.0833%, respectively. The MAPE value is less than 10%, even close to zero. Therefore, according to Wei [10], the VARX (7, 1) model is very feasible to forecast $z_{1,t}$ and $z_{2,t}$ in the next day based on the feasibility criteria of the MAPE-based model.

The feasibility of the VARX (7, 1) model is also checked using the visual comparison between the forecast results and the actual data. The visual comparison between the forecast results and the actual data is presented in Figure 4.

Figure 4 shows that the forecast line of $z_{1,t}$ and $z_{2,t}$ appears very close to the actual data of $z_{1,t}$ and $z_{2,t}$. It means that the VARX (7, 1) model is visually feasible to forecast $z_{1,t}$ and $z_{2,t}$ in the next day. The VARX (7, 1) model is also practically forecast $z_{1,t}$ and $z_{2,t}$ in the next few days.

4.7. Forecasting. Forecasting $z_{1,t}$, $z_{2,t}$, and $z_{3,t}$ are carried out on six days, from 8 November 2021 until 13 November 2021. These forecasts are practically carried out using equation (19) for $z_{1,t}$ and $z_{2,t}$ and equation (20) for $z_{3,t}$. The second-dose COVID-19 vaccination rates used for this forecast are the same as the actual on that date, 29.5200%, 29.7817%, 30.1154%, 30.4491%, 30.7828%, and 31.0465%. The $z_{1,t}$, $z_{2,t}$, and $z_{3,t}$ forecasting results obtained are presented in Table 7.

Table 7 shows that the forecasting results of $z_{1,t}$, $z_{2,t}$, and $z_{3,t}$ from 8 November 2021 until 13 November 2021 tend to decrease. To determine the accuracy of the forecast results presented in Table 7, we will look at the mean absolute percentage error (MAPE) size. The MAPE of $z_{1,t}$, $z_{2,t}$, and

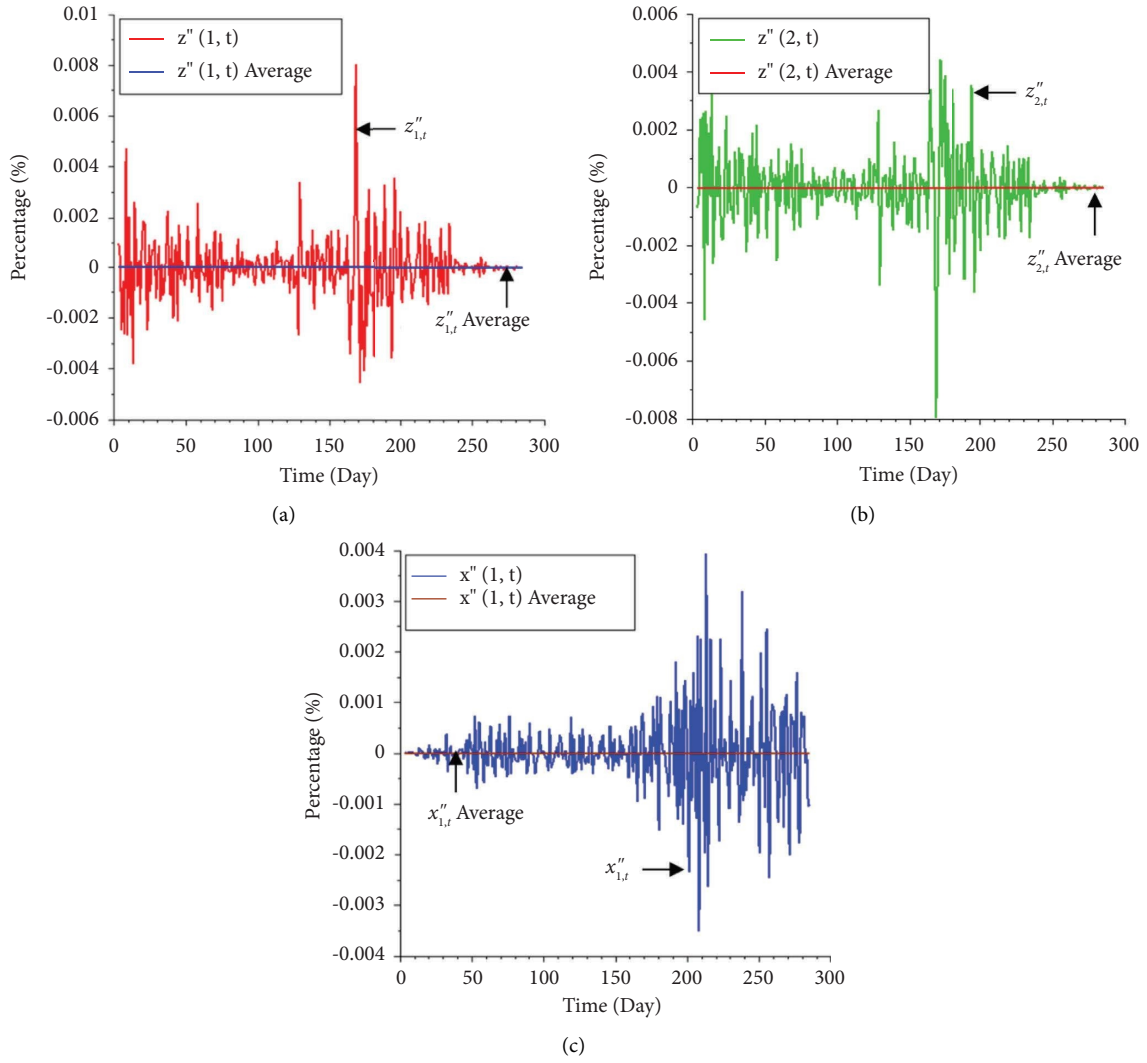

Figure 2: Graph of $z''_{1,t}$ (a), $z''_{2,t}$ (b), and $x''_{1,t}$ (c).

Table 5: Data stationary test $z''_{1,t}$, $z''_{2,t}$, and $x''_{1,t}$.

Data	$z''_{1,t}$	$z''_{2,t}$	$x''_{1,t}$
τ	−8.2813	−8.3078	−13.7290
Critical value	−3.2000	−3.2000	−3.2000

$z_{3,t}$ forecasting results on 8 November 2021 until 13 November 2021 is 2.2853%, 0.0051%, and 0.0082%, respectively. This MAPE of $z_{1,t}$, $z_{2,t}$, and $z_{3,t}$ forecasting results is tiny. Therefore, equations (19) and (20) are suitable for forecasting $z_{1,t}$, $z_{2,t}$, and $z_{3,t}$ for the next six days.

5. Discussions

First, an analysis of the influence of the second-dose COVID-19 vaccination rates on active and recovered case rates of COVID-19 is discussed. Generally, equation (19) shows that the second-dose COVID-19 vaccination rate in the previous three days influences active and recovered case rates of COVID-19 today. It does not always decrease active COVID-19 case rates but increases active COVID-19 case rates. It is also an increase and decrease in the COVID-19 recovered case rate. The active case rates of COVID-19 today are −1.17% of second-dose COVID-19 vaccination rate today, 24.37% of second-dose COVID-19 vaccination rate yesterday, −45.23% of second-dose COVID-19 vaccination rates the previous two days, and 22.03% of second-dose COVID-19 vaccination rate previous three days. A negative percentage means that the second-dose COVID-19 vaccination rate will decrease today's COVID-19 active case rate. In contrast, a positive percentage means that the second-dose COVID-19 vaccination rate will increase today's COVID-19 active case rate.

Meanwhile, the recovered case rates of COVID-19 today are 1.04% of the second-dose COVID-19 vaccination rate today, −23.70% of the second-dose COVID-19 vaccination rate yesterday, 44.28% of the second-dose COVID-19 vaccination rate previous two days, and −21.62% of second-dose COVID-19 vaccination rate previous three days. A positive percentage means that the second-dose COVID-19 vaccination rate will increase today's COVID-19 recovered case rate. In contrast, a negative percentage means that the

Table 6: Results of parameter estimates of VARX (7, 1).

Parameter	Estimator
μ_1	$1.745E-05$
μ_2	$-1.817E-05$
$\phi_{1,1}^{(1)}$	0.5495
$\phi_{1,2}^{(1)}$	0.7931
$\phi_{2,1}^{(1)}$	−1.0376
$\phi_{2,2}^{(1)}$	−1.2881
$\phi_{1,1}^{(2)}$	1.2891
$\phi_{1,2}^{(2)}$	1.4981
$\phi_{2,1}^{(2)}$	−1.5079
$\phi_{2,2}^{(2)}$	−1.7207
$\phi_{1,1}^{(3)}$	5.3873
$\phi_{1,2}^{(3)}$	5.6679
$\phi_{2,1}^{(3)}$	−5.5021
$\phi_{2,2}^{(3)}$	−5.7843
$\phi_{1,1}^{(4)}$	5.6900
$\phi_{1,2}^{(4)}$	6.0157
$\phi_{2,1}^{(4)}$	−5.8682
$\phi_{2,2}^{(4)}$	−6.1934
$\phi_{1,1}^{(5)}$	8.9929
$\phi_{1,2}^{(5)}$	9.4176
$\phi_{2,1}^{(5)}$	−9.0363
$\phi_{2,2}^{(5)}$	−9.4565
$\phi_{1,1}^{(6)}$	3.2196
$\phi_{1,2}^{(6)}$	3.3846
$\phi_{2,1}^{(6)}$	−3.3553
$\phi_{2,2}^{(6)}$	−3.5186
$\phi_{1,1}^{(7)}$	0.7175
$\phi_{1,2}^{(7)}$	0.6588
$\phi_{2,1}^{(7)}$	−0.8463
$\phi_{2,2}^{(7)}$	−0.7909
$\theta_{1,1}^{(0)}$	−0.0117
$\theta_{2,1}^{(0)}$	0.0104
$\theta_{1,1}^{(1)}$	0.2203
$\theta_{2,1}^{(1)}$	−0.2162

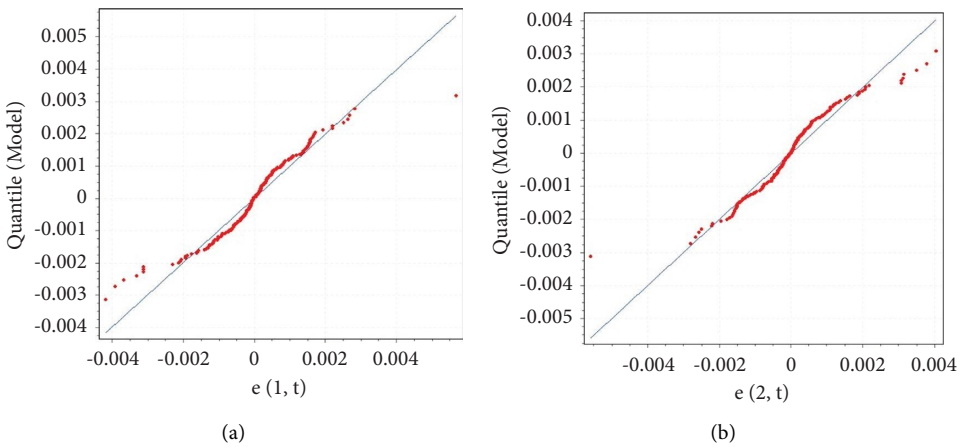

Figure 3: Q-Q plots of $e_{1,t}$ (a) and $e_{2,t}$ (b).

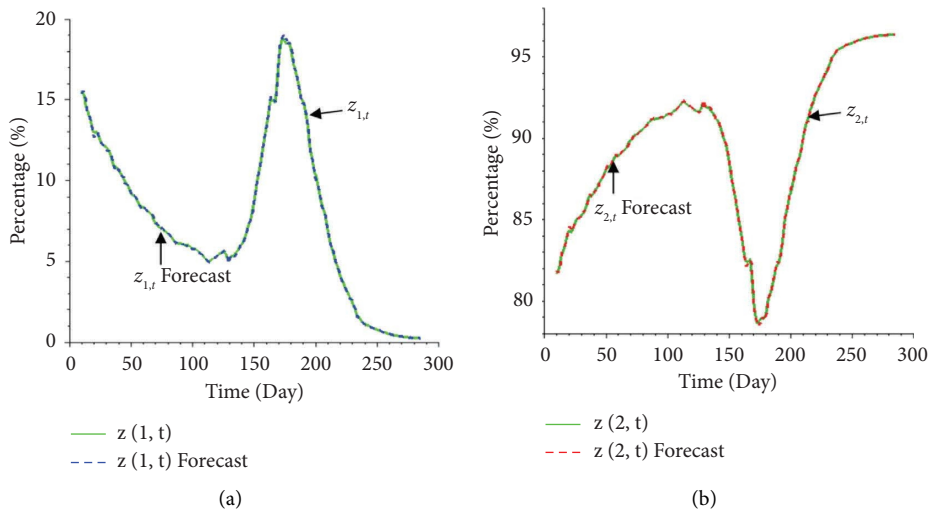

Figure 4: The visual comparison between the $z_{1,t}$ and $z_{1,t}$ forecast (a) and $z_{2,t}$ and $z_{2,t}$ forecast (b).

Table 7: The $z_{1,t}$, $z_{2,t}$, and $z_{3,t}$ forecasting results.

Date	$z_{1,t}$ forecast (%)	$z_{2,t}$ forecast (%)	$z_{3,t}$ forecast (%)
8 November 2021	0.2275	96.3930	3.3795
9 November 2021	0.2350	96.3862	3.3788
10 November 2021	0.2247	96.3960	3.3793
11 November 2021	0.2389	96.3822	3.3788
12 November 2021	0.2220	96.3988	3.3791
13 November 2021	0.2165	96.4042	3.3793

second-dose COVID-19 vaccination rate will reduce COVID-19 recovered case rate today. These facts do not mean that vaccination in Indonesia is not adequate. Suppose the percentage of the effect of today's vaccination up to the previous three days on the active and recovered case rates is added. In that case, the results have negative and positive values, respectively. Negative values indicate that vaccination reduces the daily active case rate, and positive values indicate that vaccination generally increases the daily recovery case rate. Next is an analysis of the influence of active and recovered case rates of COVID-19 on itself. In general, equation (19) shows that the active and recovered rates of COVID-19 today are influenced by themselves on the previous three days. The percentage of each day can be seen in equation (19).

The last is an analysis of the effect of second-dose COVID-19 vaccination rates on 8 November 2021 on active and recovered case rates of COVID-19. Suppose the second-dose COVID-19 vaccination rates on 8 November 2021 are assumed not to increase from the previous day. Thus, the forecasted active and recovered case rates of COVID-19 on 8 November 2021 are 0.2321% and 96.3886%, respectively. The forecast value of the COVID-19 active case rates is greater than the forecast value obtained in the forecasting section. The forecast value of the COVID-19 recovered case rates is also smaller than that obtained in the forecasting section. It shows that if the second-dose COVID-19 vaccination rates on 8 November 2021 are not increased, the decline in the COVID-19 spread rate in Indonesia will be slower. Therefore, although the second-dose COVID-19 vaccination rates do not always decrease the COVID-19 active case rates and increase the COVID-19 recovered case rates every day, the second-dose COVID-19 vaccination rates on 8 November 2021 must still be increased. It also applies from 9 November 2021 until 13 November 2021.

6. Conclusion

The VARX model (7, 1) is the best VARX model to describe a causal relationship between active and recovered case rates of COVID-19 in Indonesia influenced by the exogenous variable, second-dose COVID-19 vaccination rates. The assumption of white noise is fulfilled, and the MAPE of this model is tiny, so this model is very suitable to be used in the forecasting process the next day. Forecasting results in the six days also give very close results to the actual data (even if the absolute percentage error is almost zero). There is an oddity in this VARX (7.1) model. The second-dose COVID-19 vaccination rates carried out daily should always reduce COVID-19 active case rates and increase COVID-19 recovered case rates.

However, in this model, the impact of the second-dose COVID-19 vaccination rates carried out in the previous three days on COVID-19 active case rates can decrease and increase simultaneously. It also happened to the impact of the second-dose COVID-19 vaccination rates on the COVID-19 recovered case rates. The second-dose COVID-19 vaccination rates carried out in the previous three days on COVID-19 recovered case rates can increase

and decrease simultaneously. However, it does not indicate that vaccination does not affect daily activities and recovery case rates. The sum of the percentage effect of today's vaccination to the previous three days on the active and recovered case rates has negative and positive values, respectively. These negative and positive values indicate that vaccination generally decreases the daily active case rate and increases the daily recovery case rate.

Another analysis result obtained is that if the second-dose COVID-19 vaccination rates do not increase from the previous day, COVID-19 active case rates will only decrease slightly, and COVID-19 recovered case rates will only increase slightly. It has caused the COVID-19 spread in Indonesia to decline more slowly. Based on this, the vaccination rate must still be increased. With this model, the Indonesian government can project the second-dose COVID-19 vaccination rates that must be carried out today so that active and recovered case rates of COVID-19 can be achieved by that date.

Data Availability

All data are obtained from the official website of the Science and Technology Index, National Research and Innovation Agency, Ministry of Research and Technology of the Republic of Indonesia.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This project was funded by the Padjadjaran University Internal Research Grant, the "RISET DATA PUSTAKA DAN DARING (RDPD)" program under Prof. Dr. Sukono, MM., M.Sc., with Number: 2203/UN6.3.1/PT.00/2022.

References

[1] M. R. Aziz, M. A. Tavares, and C. J. Azhima, "COVID-19 vaccinations and the right to Health in Indonesia: social justice analysis," *Lentera Hukum*, vol. 8, no. 2, pp. 211–240, 2021.

[2] R. Sparrow, T. Dartanto, and R. Hartwig, "Indonesia under the new normal: challenges and the way ahead," *Bulletin of Indonesian Economic Studies*, vol. 56, no. 3, pp. 269–299, 2020.

[3] T. Toharudin, R. S. Pontoh, R. E. Caraka et al., "National vaccination and local intervention impacts on COVID-19 cases," *Sustainability*, vol. 13, no. 15, pp. 8282–8317, 2021.

[4] N. M. Nasir, R. N. Alkaff, D. Aristi, and J. F. Faiz, "A study of misinformation exposure of COVID-19 vaccine and the willingness to be vaccinated in Tangerang Selatan city, Indonesia," *Jurnal Kesehatan Reproduksi*, vol. 12, no. 1, pp. 1–13, 2021.

[5] C. Cucunawangsih, R. S. Wijaya, N. P. H. L. Lugito, and I. Suriapranata, "Post-vaccination cases of COVID-19 among healthcare workers at Siloam Teaching Hospital, Indonesia," *International Journal of Infectious Diseases*, vol. 107, no. 2021, pp. 268–270, 2021.

[6] M. Muhyiddin and H. Nugroho, "A year of COVID-19: a long road to recovery and acceleration of Indonesia's development," *Jurnal Perencanaan Pembangunan: The Indonesian Journal of Development Planning*, vol. 5, no. 1, pp. 1–19, 2021.

[7] S. G. Makridakis, S. C. Wheelwright, and R. J. Hyndman, *Forecasting: Methods and Applications*, John Wiley and Sons, New York, NY, USA, Third Edition, 1998.

[8] P. Das, *Econometrics in theory and practice*, Springer, Berlin, Germany, 2019.

[9] A. P. Kirana and A. Bhawiyuga, "Coronavirus (COVID-19) pandemic in Indonesia: cases overview and daily data time series using naïve forecast method," *Indonesian Journal of electronics, electromedical engineering, and medical informatics*, vol. 3, no. 1, pp. 1–8, 2021.

[10] W. W. S. Wei, "Time Series Analysis: Univariate and Multivariate Methods," *Technometrics*, vol. 33, no. 1, 1991.

[11] I. Djakaria and S. E. Saleh, "COVID-19 forecast using Holt-Winters exponential smoothing," *Journal of Physics: Conference Series*, vol. 1882, Article ID 12033, 12037 pages, 2021.

[12] C. B. Aditya Satrio, W. Darmawan, B. U. Nadia, and N. Hanafiah, "Time series analysis and forecasting of Coronavirus disease in Indonesia using ARIMA model and PROPHET," *Procedia Computer Science*, vol. 179, pp. 524–532, 2021.

[13] F. Fadly and E. Sari, "An an approach to measure the death impact of c-19 in Jakarta using autoregressive integrated moving average (ARIMA)," *Unnes Journal of Public Health*, vol. 9, no. 2, pp. 108–116, 2020.

[14] R. Fitriani, W. D. Revildy, E. Marhamah, T. Toharudin, and B. N. Ruchjana, "The autoregressive integrated vector model approach for COVID-19 data in Indonesia and Singapore," *Journal of Physics: Conference Series*, vol. 1722, pp. 012057-012058, 2021.

[15] A. Meimela, S. S. S. Lestari, I. F. Mahdy, T. Toharudin, and B. N. Ruchjana, "Modeling of COVID-19 in Indonesia using vector autoregressive integrated moving average," *Journal of Physics: Conference Series*, vol. 1722, Article ID 12079, 12087 pages, 2021.

[16] M. F. Akbar, E. Octaviany, T. Toharudin, and B. N. Ruchjana, "The application of vector autoregressive integrated with exogenous variable to model the relationship between COVID-19 positive numbers with population growth rate," *Journal of Physics: Conference Series*, vol. 1722, pp. 012081–012088, 2021.

[17] E. Muschilati and N. Irsalinda, "Forecasting tourist visit using the vector autoregressive exogenous method (*VARX*)," *Jurnal Ilmiah Matematika*, vol. 7, no. 2, pp. 81–87, 2020.

[18] A. R. Putri, M. Usman, W. Warsono, E. Virginia, and E. Virginia, "Application of vector autoregressive with exogenous variable: case study of closing stock price of PT INDF. Tbk and PT ICBP. Tbk," *Journal of Physics: Conference Series*, vol. 1751, Article ID 12012, 12110 pages, 2021.

[19] R. S. Tsay, *Multivariate Time Series Analysis with R and Financial Application*, John Wiley and Sons, Hoboken, NJ, USA, 2014.

[20] W. Warsono, E. Russel, W. Wamiliana, W. Widiarti, and M Usman, "Modeling and forecasting by the vector autoregressive moving average model for export of coal and oil data (case study from Indonesia over the years 2002-2017)," *International Journal of Energy Economics and Policy*, vol. 9, no. 4, pp. 240–247, 2019.

[21] W. W. S. Wei, *Multivariate Time Series Analysis and Applications*, John Wiley and Sons, Hoboken, NJ, USA, 2019.

[22] G. C. S. Lui, W. K. Li, K. M. Y. Leung, J. H. W. Lee, and A. W. Jayawardena, "Modelling algal blooms using vector autoregressive model with exogenous variables and long memory filter," *Ecological Modelling*, vol. 200, no. 1-2, pp. 130–138, 2007.

[23] W. Warsono, E. Russel, W. Wamiliana, and M. Usman, "Vector autoregressive with exogenous variable model and its application in modeling and forecasting energy data: case study of PTBA and HRUM energy," *International Journal of Energy Economics and Policy*, vol. 9, no. 2, pp. 390–398, 2018.

[24] F. Khan, A. Saeed, and S. Ali, "Modelling and forecasting of new cases, deaths and recover cases of COVID-19 by using vector autoregressive model in Pakistan," *Chaos, Solitons and Fractals*, vol. 140, Article ID 110189, 110195 pages, 2020.

[25] Y. W. Cheung and K. S. Lai, "Lag order and critical values of the augmented Dickey-Fuller test," *Journal of Business and Economic Statistics*, vol. 13, no. 3, pp. 277–280, 1995.

[26] G. Elliott, T. J. Rothenberg, and J. H. Stock, "Efficient tests for an autoregressive unit root," *Econometrica*, vol. 64, no. 4, pp. 813–836, 1996.

[27] R. I. D. Harris, "Testing for unit roots using the augmented Dickey-Fuller test," *Economics Letters*, vol. 38, no. 4, pp. 381–386, 1992.

[28] W. R. Bell, "An introduction to forecasting with time series models," *Insurance: Mathematics and Economics*, vol. 3, no. 4, pp. 241–255, 1984.

[29] Z. Hossain, A. Rahman, M. Hossain, and J. H. Karami, "Over-differencing and forecasting with non-stationary time series data," *Dhaka University Journal of Science*, vol. 67, no. 1, pp. 21–26, 2019.

[30] A. N. Ajmi, G. C. Aye, M. Balcilar, G. El Montasser, and R. Gupta, "Causality between US economic policy and equity market uncertainties: evidence from linear and nonlinear tests," *Journal of Applied Economics*, vol. 18, no. 2, pp. 225–246, 2015.

[31] W. N. Thurman and M. E. Fisher, "Chickens, eggs, and causality, or which came first?" *American Journal of Agricultural Economics*, vol. 70, no. 2, pp. 237-238, 1988.

[32] H. Lutkepohl, *New Introduction to Multiple Time Series Analysis*, Springer, Berlin, Germany, 2005.

[33] A. A. Anfofum, O. Joshua, and S. G. Joshua, "Measuring the import of financial sector development on Nigeria's economic growth," *International Journal of Educational Research*, vol. 2, no. 12, pp. 201–214, 2014.

[34] L. G. Barberia, M. L. Claro Oliveira, A. Junqueira, N. D. P. Moreira, and G. D. Whitten, "Should I stay or should I go? Embracing causal heterogeneity in the study of pandemic policy and citizen behaviors," *Social Science Quarterly*, vol. 102, no. 5, pp. 2055–2069, 2021.

[35] A. Khalik Salman and G. Shukur, "Testing for Granger causality between industrial output and CPI in the presence of regime shift: Swedish data," *Journal of Economics Studies*, vol. 31, no. 6, pp. 492–499, 2004.

[36] R. Chopra, C. R. Murthy, and G. Rangarajan, "Statistical tests for detecting Granger causality," *IEEE Transactions on Signal Processing*, vol. 66, no. 22, pp. 5803–5816, 2018.

[37] R. F. Engle and K. F. Kroner, "Multivariate simultaneous generalized ARCH," *Econometric Theory*, vol. 11, no. 1, pp. 122–150, 1995, https://www.jstor.org/stable/3532933.

[38] P. Foresti, "Testing for Granger causality between stock prices and economic growth," 2007, https://mpra.ub.uni-muenchen.de/id/eprint/2962MPRA Paper.

[39] J. D. Hamilton, *Time Series Analysis*, Princeton University Press, Princeton, NJ, USA, 1994.

[40] G. G. Hamedani and M. N. Tata, "On the determination of the bivariate normal distribution from distributions of linear combinations of the variables," *The American Mathematical Monthly*, vol. 82, no. 9, pp. 913–915, 1975.

[41] F. Sinz, S. Gerwinn, and M. Bethge, "Characterization of the p-generalized normal distribution," *Journal of Multivariate Analysis*, vol. 100, no. 5, pp. 817–820, 2009.

[42] P. J. Burns, "Robustness of the Ljung-Box test and its rank equivalent," 2002, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=443560.

[43] G. M. Ljung and G. E. P. Box, "On a measure of lack of fit in time series models," *Biometrika*, vol. 65, no. 2, pp. 297–303, 1978.

[44] W. H. Beyer, *Handbook of Tables for Probability and Statistics*, CRC Press, Boca Raton, FLa, USA, Second Edition, 2017.

[45] P. J. Brockwell and R. A. Davis, *Introduction to Time Series and Forecasting*, Springer, Berlin, Germany, 2002.

[46] A. J. Andersen and J. R. Dennison, "An introduction to quantile-quantile plots for the experimental physicist," *American Association of Physics Teachers*, vol. 87, no. 5, pp. 1–4, 2019.