RESEARCH ARTICLE • OPEN ACCESS

Implementation of Moving Average Filter in SARIMA-ANN and SARIMA-SVR Methods for Forecasting Pneumonia Incidence in Jakarta

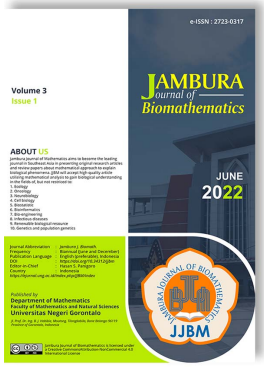
Muhammad Majid Rafi Musyaffa, Gatot Fatwanto Hertono, and Bevina Desjwiandra Handari

Volume 6, Issue 3, Pages 262–269, September 2025

Received 7 February 2025, Revised 26 March 2025, Accepted 19 September 2025, Published Online 30 September 2025

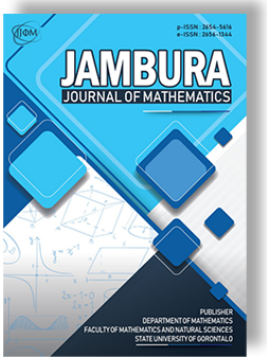
To Cite this Article : M. M. R. Musyaffa, G. F. Hertono, and B. D. Handari,"Implementation of Moving Average Filter in SARIMA-ANN and SARIMA-SVR Methods for Forecasting Pneumonia Incidence in Jakarta", *Jambura J. Biomath*, vol. 6, no. 3, pp. 262–269, 2025, *https://doi.org/10.37905/jjbm.v6i3.30558*

JOURNAL INFO • JAMBURA JOURNAL OF BIOMATHEMATICS



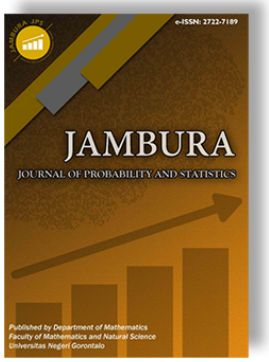
Homepage	:	http://ejurnal.ung.ac.id/index.php/JJBM/index
Journal Abbreviation	:	Jambura J. Biomath.
Frequency	:	Quarterly (March, June, September and December)
Publication Language	:	English
DOI	:	https://doi.org/10.37905/jjbm
Online ISSN	:	2723-0317
Editor-in-Chief	:	Hasan S. Panigoro
Publisher	:	Department of Mathematics, Universitas Negeri Gorontalo
Country	:	Indonesia
OAI Address	:	http://ejurnal.ung.ac.id/index.php/jjbm/oai
Google Scholar ID	:	XzYgeKQAAAAJ
Email	:	editorial.jjbm@ung.ac.id

JAMBURA JOURNAL • FIND OUR OTHER JOURNALS

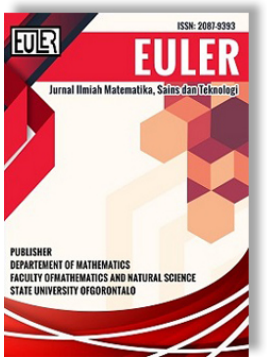


Jambura Journal of Mathematics

Jambura Journal of Mathematics Education



Jambura Journal of Probability and Statistics



EULER : Jurnal Ilmiah Matematika, Sains, dan Teknologi

Jambura Journal of Biomathematics, Volume 6, Issue 3, Pages 262–269, September 2025 https://doi.org/10.37905/jjbm.v6i3.30558

Research Article

Check for updates

Implementation of Moving Average Filter in SARIMA-ANN and SARIMA-SVR Methods for Forecasting Pneumonia Incidence in Jakarta

Muhammad Majid Rafi Musyaffa[1], **Gatot Fatwanto Hertono**[1], **and Bevina Desjwiandra Handari**[1,*]

[1]*Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Indonesia*

ARTICLE HISTORY
Received 7 February 2025
Revised 26 March 2025
Accepted 19 September 2025
Published 30 September 2025

KEYWORDS
Pneumonia
SARIMA
ANN
SVR
Hybrid model

ABSTRACT. *In this study, we implemented a moving average filter in SARIMA-ANN and SARIMA-SVR to predict Pneumonia incidence in Jakarta. Pneumonia is one of the highest causes of death in children throughout the world. Forecasting pneumonia incidence in the future can help to reduce the spread of cases, so that the number of deaths due to pneumonia can be reduced. In general, time series data consists of linear and nonlinear patterns, which cannot be properly modeled by linear or nonlinear models alone. One way to solve this issue is to use a hybrid model that combines several models to overcome the limitations of each component model and improve predicting performance. SARIMA-ANN and SARIMA-SVR methods combine a linear seasonal autoregressive integrated moving average (SARIMA) model and a nonlinear artificial neural network (ANN) or support vector regression (SVR) model to capture the linear and nonlinear characteristics of the data. Parameter estimation in SARIMA uses Gaussian Maximum Likelihood Estimation. Initially, the time series will be transformed by a moving average (MA) filter, so SARIMA can model the data well. Meanwhile, the remaining components separated from the transformation will be modeled with a nonlinear model such as ANN in the SARIMA-ANN method, or SVR in the SARIMA-SVR method. The simulation results show that the SARIMA-ANN method is superior to the SARIMA-SVR method in predicting incidences in West Jakarta and East Jakarta, with a MAPE difference ranging from 0.6% to 0.75%. Meanwhile, in North, South, and Central Jakarta, the SARIMA-SVR method is superior to the SARIMA-ANN method, with MAPE differences ranging from 1.6% to 3.99%. The SARIMA-SVR model achieves better results across the majority of municipalities, indicating that the SARIMA-SVR model generally provides better result for predicting Pneumonia incidence in Jakarta.*

1. Introduction

Pneumonia is a lung infection that is generally caused by bacteria, viruses, or fungi. The immune system's reaction to this infection causes the lung's air sacs, or alveoli, to fill with fluid or pus. Pneumonia is one of the leading causes of death for those over 65 and children under five throughout the world [1]. In 2019 alone, 740,180 children under five years old died from pneumonia [2]. Pneumonia is commonly seen in various health settings across the world and remains an intense area of study given its significance and the gaps in knowledge that persist [3]. Pneumonia survivors do not return to their pre-infection health trajectories but instead experience an accelerated health decline with an increased risk of cardiovascular disease [4]. Forecasting the number of pneumonia patients in the future can help reduce the number of deaths due to pneumonia. Medical practitioners can prepare the necessary medicines, adequate medical personnel, or provide education to people who are vulnerable to pneumonia before a spike in cases can occur. This problem causes great concern for the lives of many people. Therefore, adequate accuracy is required for forecasts [5].

In general, time series data consist of linear and nonlinear patterns, which cannot be properly modeled by linear models or nonlinear models alone. One way to solve this issue is to use a hybrid model that combines several models to overcome the limitations of each component within the hybrid model and improve the forecasting performance [6]. Several previous studies that have been conducted have shown that methods that combine linear and nonlinear models will produce smaller forecast errors than using only one model [7–9].

In 2003, Zhang [10] proposed a hybrid Autoregressive Integrated Moving Average – Artificial Neural Network (ARIMA-ANN) model that combines the linear ARIMA model and the nonlinear ANN model to capture both linear and nonlinear patterns in the data. This method will input the data into ARIMA first, then the ARIMA forecast error from the training data, which is considered a nonlinear component, is modeled using ANN. The forecast results from ARIMA and ANN are then summed to obtain the final forecast. Zhang's ARIMA-ANN method produces better forecasts than using a single ARIMA or ANN on three different datasets, namely annual sunspot data, annual Canadian lynx

*Corresponding Author.

Email : *bevina@sci.ui.ac.id* (B. D. Handari)
Homepage : http://ejurnal.ung.ac.id/index.php/JJBM/index / E-ISSN : 2723-0317

data, and weekly exchange rates data. Over time, several studies began to try Zhang's method on data that had seasonal patterns, so instead of using ARIMA, seasonal ARIMA (SARIMA) was used. Research conducted by Moeeni and Bonakdari [11] shows that the SARIMA-ANN model displayed greater ability to forecast flood flows than an individual SARIMA or ANN model. Mukaram and Yusof [12] also shows that the SARIMA-ANN model is better than using SARIMA or ANN individually for forecasting monthly solar radiation. Several studies have also tried nonlinear models other than ANN such as support vector regression (SVR). Research conducted by Bouzerdoum, Mellit, and Pavan [13] shows that the SARIMA-SVR model outperforms SARIMA and SVR models for short-term power forecasting, and Xuemei et al. [14] shows that the SARIMA-SVR model outperforms SARIMA and SVR models in short-term cooling load forecasting. These studies stated the superior performance of the ARIMA-ANN, the SARIMA-ANN model and the SARIMA-SVR model compare to the performances of the corresponding individual models.

In 2014, Babu and Reddy [15] proposed a new hybrid ARIMA-ANN model that does not feed data into either ARIMA or ANN directly. Instead, the data is decomposed into two components, namely the low-volatile component and the high-volatile component, which can be considered as linear and nonlinear sequence, respectively. This decomposition is done by using a moving average filter, which is a technique commonly used in time series analysis to smooth out short-term fluctuations and highlight long-term trends or cycles [16]. The low-volatile component will be fitted to ARIMA, while the high-volatile component will be fitted to ANN. The forecast results from ARIMA and ANN are then summed to obtain the final forecast. This method produces better forecasts than Zhang's ARIMA-ANN model, single ARIMA, and single ANN, using the annual sunspot, hourly electricity price, and daily stock price data.

Several studies have been conducted on pneumonia forecasting in Indonesia using statistical and machine learning methods. For instance, Cahyati and Sari [17] employed the SARIMA model to forecast monthly cases of pneumonia among children in Semarang, while Tricahya and Rustam [5] compared the performance of a modified Fuzzy Time Series (FTS) against ARIMA and Holt's Exponential Smoothing for forecasting monthly pneumonia cases in Jakarta city. However, no research has explored the use of SARIMA-ANN and SARIMA-SVR models for pneumonia forecasting, particularly in Jakarta, where the data used is suspected to contain seasonal patterns.

In this study, the SARIMA-ANN and SARIMA-SVR methods, with a moving average filter, were used to forecast the weekly number of pneumonia incidence in Jakarta. The data that we use in this study is the weekly number of pneumonia incidence from every hospital in each municipality of Jakarta (except the data from the Thousand Islands) from 2017 until 2023. Through the forecasting performance of models SARIMA-ANN and SARIMA-SVR derived from this study, it is expected to serve as a tool to improve disease surveillance and early warning systems.

2. Theoretical Background

There are three models used, namely Seasonal Autoregressive Integrated Moving Average (SARIMA), Artificial Neural Network (ANN), and Support Vector Regression (SVR). A brief explanation of each model is as follows:

2.1. Seasonal Autoregressive Integrated Moving Average

Seasonal Autoregressive Integrated Moving Average or SARIMA is a forecasting model that predicts a future value in a time series using a linear combination of past values and past forecast errors. Over the past three decades, SARIMA has demonstrated significant success in both academic research and industrial applications [14]. Suppose y_t is the value of time series data at time point t and ϵ_t is the forecast error at time point t. The mathematical representation of a SARIMA(p,d,q) $(P,D,Q)_s$ can be written as follows:

$$\phi(B)\phi(B^s)(1-B)^d(1-B^s)^D y_t = \theta(B)\theta(B^s)\epsilon_t, \qquad (1)$$

where

$$\phi(B) = \left(1 - \sum_{i=1}^{p} \phi_i B^i\right), \quad \Phi(B^s) = \left(1 - \sum_{i=1}^{p} \phi_i B^{is}\right), \quad (2)$$

$$\theta(B) = \left(1 - \sum_{i=1}^{q} \theta_i B^i\right), \quad \Theta(B^s) = \left(1 - \sum_{i=1}^{Q} \Theta_i B^{is}\right). \quad (3)$$

In these equations, B is the backward operator with properties: $By_t = y_{t-1}, B^n y_t = y_{t-n}, B(ay_t + c) = ay_{t-1} + c$. where a and c are constants. The components of the SARIMA model are as follows:

- $\phi(B)$ and $\theta(B)$ represent the autoregressive (AR) and moving average (MA) parts, respectively, consisting of linear combinations of past values and past forecast errors. The order of AR and MA are respectively denoted as p and q.
- $\Phi(B^s)$ and $\Theta(B^s)$ denote the seasonal autoregressive and seasonal moving average parts, incorporating linear combinations of past seasonal values and seasonal forecast errors. The order of seasonal AR and seasonal MA are respectively denoted as P and Q, and s denotes the seasonal period of the data.
- $(1-B)^d$ and $(1-B^s)^D$ are the differencing and seasonal differencing operators, respectively, transforming the data (y_t) into its differences and seasonal differences. The orders of differencing and seasonal differencing are respectively denoted as d and D.

The order of differencing and seasonal differencing can be determined by using some unit root tests, or by looking at the autocorrelation function (ACF) plot and the plot of the time series [18, 19]. After d and D have been determined, the order of autoregressive and seasonal autoregressive are determined by examining at the significant values in the partial autocorrelation function (PACF) plot, while the orders of the moving average and seasonal moving average are examining by looking at the significant values in the ACF plot. Another way to determine $p, q, P,$ and Q is by evaluating the Corrected Akaike Information Criteria (AICc) value of the model. A model with a minimum value of AICc is often the best model for forecasting [18].

2.2. Artificial Neural Network

Artificial Neural Network (ANN) is a versatile and powerful model inspired by the structure and functioning of the human brain's neurons. ANN is particularly flexible in terms of architecture, allowing for various configurations to suit different types

of data and tasks. For time series forecasting, a three-layer ANN, comprised of an input layer, a hidden layer, and an output layer, is commonly used [15]. Each layer consists of one or more neurons, with every neuron in a layer connected to every neuron in the subsequent layer. In this study, since the objective is to forecast a single variable, the weekly pneumonia incident, a single neuron will be employed in the output layer. The mathematical representation of a three-layer ANN with one neuron in the output layer can be written as follows:

$$\hat{y}_t = f_2\left(\sum_{j=1}^{n_2} w_{j,1}^{(2)} \cdot f_1\left(\sum_{j=1}^{n_1} w_{i,j}^{(1)} y_{t-i} + b_j^{(1)}\right) + b_1^{(2)}\right), \quad (4)$$

where n_1 and n_2 respectively, are the number of neurons in the input layer and hidden layer, f_1 is the activation function in the hidden layer, f_2 is the activation function in the output layer, $w_{i,j}^{(1)}$ is the weight that connects neuron i in the input layer to neuron j in the hidden layer, $w_{j,1}^{(2)}$ is the weight that connects neuron j in the hidden layer to neuron 1 in the output layer, $b_j^{(1)}$ and $b_1^{(2)}$ are the biases for each neuron in the hidden and output layers, respectively.

To construct a nonlinear ANN model, a nonlinear activation function is required in the hidden layer. This study uses the sigmoid activation function, which is defined as follows:

$$f_1(x) = \frac{1}{1+e^{-x}}, \quad (5)$$

where x is an input vector. In the regression problems, the commonly activation function used in the output layer is the linear activation function, which is defined as follows:

$$f_2(x) = x. \quad (6)$$

The weights and biases in an ANN must be adjusted to minimize the prediction error. One method for optimizing these weights and biases is backpropagation, a training method in ANN used to reduce error by updating the weights and biases based on the error generated by the network.

2.3. Support Vector Regression

Support Vector Regression (SVR) is a type of Support Vector Machine (SVM) used for regression problems. Some of the main advantages of SVR is that it has an excellent generalization capability, and its computational complexity does not depend on the dimensionality [20]. The fundamental concept of the SVR model is to train and learn from all research data samples, distributing them between two parallel lines in such a way that the total deviation of all points is minimized [21]. Suppose x is a vector containing $y_{t-1}, y_{t-2}, \ldots, y_{t-m}$. The prediction formula for y_t is as follows:

$$\hat{y}_t = f(x) = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) K(x, x_i) + b, \quad (7)$$

where x_i is the training data that plays a direct role in determining the regression hyperplane, α_i and α_i^* are the Lagrange multipliers associated with the support vectors on the upper and lower margins, respectively, $K(x, x_i) = \varphi(x)\varphi(x_i)$ is a kernel function that measures the similarity between the input vector and the support vector in a high-dimensional space, b is a bias or constant, and n is the number of observations in training data.

SVR uses an ϵ-insensitive loss function, which ignores forecasting errors if the difference is not greater than error ϵ, the function is determined as follows:

$$L_\epsilon(y_t, \hat{y}_t) = \max(0, |y_t - \hat{y}_t| - \epsilon). \quad (8)$$

To make SVR a non-linear model, a non-linear kernel function is required. One of them is the Gaussian kernel function/radial basis function (RBF):

$$K(x, x_i) = \exp(-\gamma \|x - x_i\|^2), \quad (9)$$

where γ is a constant with a positive value and its value can be set independently.

To find the Lagrange coefficient α_i and α_i^*, the following dual problem needs to be solved:

$$\max_{\alpha,\alpha^*} L(\alpha, \alpha^*) = -\varepsilon \sum_{i=1}^{n} (\alpha_i + \alpha_i^*) + \sum_{i=1}^{n} y_i (\alpha_i - \alpha_i^*) - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) K(x_i, x_j),$$

subject to the constraints:

$$0 \le \alpha_i, \alpha_i^* \le C, \quad \forall i = 1, \ldots, n,$$

$$\sum_{i=1}^{m} (\alpha_i - \alpha_i^*) = 0,$$

where x_i are the vector containing past values in training data and C is a cost value that can be set independently.

3. Method

In general, time series data can exhibit both linear and nonlinear characteristics. Data that have linear characteristics can be modeled well by linear models, while data that have nonlinear characteristics can be effectively modeled by nonlinear models. To model a time series that has linear and nonlinear patterns, Babu and Reddy [15] tried to separate the time series into two components, namely the smoothed trend component $l_(t)$, which shows low volatility, and the residual component h_t, which exhibits high volatility. This separation can be achieved using a moving average filter.

Moving average (MA) filter is a statistical method commonly used in time series analysis to smooth short-term fluctuations and highlight long-term trends or cycles [16]. Note that the MA filter is totally different than moving average that is used in SARIMA model. The moving average filter takes m-sequential data and then looks for the average to be used as a new data point. Mathematically, the equation of the moving average filter can be seen in the following equation:

$$l_t = \frac{1}{m} \sum_{i=t-(m-1)}^{t} y_i, \quad (10)$$

where y_i is the data value at period i from the initial time series, l_t is the result of MA filter for the data at period t, and m is the

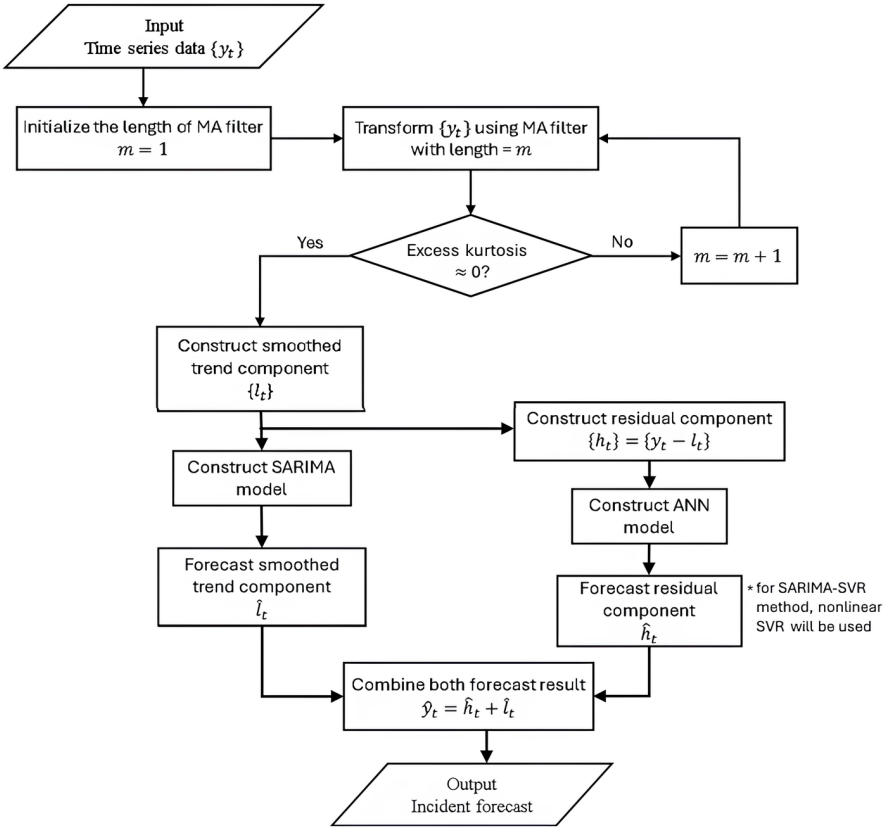


Figure 1. Flowchart of the Method (Derived from [15])

length of the MA filter. The value of m needs to be adjusted so that a clear long-term trend or cycle is visible.

The SARIMA model parameters are estimated using the Gaussian maximum likelihood, where a normally distributed time series provides more accurate parameter estimates. For this reason, the time series $\{y_t\}$ will be transformed using a MA filter, so that the kurtosis of $\{y_t\}$ approaches a normal distribution. This can be done by choosing a suitable MA filter length so that the sample excess kurtosis of the transformed time series is close enough to zero. One way to calculate the excess kurtosis is to use the following equation [22]:

$$K = \frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum_{i=1}^{n} \left(\frac{y_i - \bar{y}}{s} \right)^4 - \frac{3(n-1)^2}{(n-2)(n-3)}, \tag{11}$$

where $\bar{y}$ is the mean of $\{y_t\}$, s is the standard deviation of $\{y_t\}$, and n is the number of observations in $\{y_t\}$.

The transformation results from the moving average filter exhibit low volatility and predominantly consist of linear patterns. Therefore, SARIMA can model it more effectively than nonlinear models [15]. The results obtained from the MA filter will be called as the smoothed trend component (l_t). Assuming that the time series data is as a sum of smoothed trend component (l_t) and residual component (h_t), then

$$y_t = l_t + h_t. \tag{12}$$

As a result, the residual component can be obtained through subtraction of the initial data with its smoothed trend component,

$$h_t = y_t - l_t. \tag{13}$$

The residual component, which is highly volatile and consists of nonlinear patterns, will be fit into a nonlinear model such as a nonlinear ANN or a nonlinear SVR. The predicted results of smoothed trend component $(\hat{l}_t)$ and residual component $(\hat{h}_t)$ at period t will be combined to construct the forecast of the weekly incident at period t,

$$\hat{y}_t = \hat{l}_t + \hat{h}_t. \tag{14}$$

Summary of the used method in this study can be seen at the flowchart in Figure 1.

To evaluate the model's performance, the weekly forecasted incident results will be compared with actual values from the data using the mean absolute error (MAE) and the mean absolute percentage error (MAPE), as described by the following equations [23]:

$$MAE = \frac{1}{H} \sum_{t=n+1}^{n+H} |y_t - \hat{y}_t|,$$

$$MAPE = \frac{1}{H} \sum_{t=n+1}^{n+H} \frac{|y_t - \hat{y}_t|}{y_t} \times 100\%,$$

where H denotes the forecast horizon (number of observations in the testing data).

4. Results And Discussion

The data used in this study is the weekly number of pneumonia cases from every hospital in each of the five municipal-

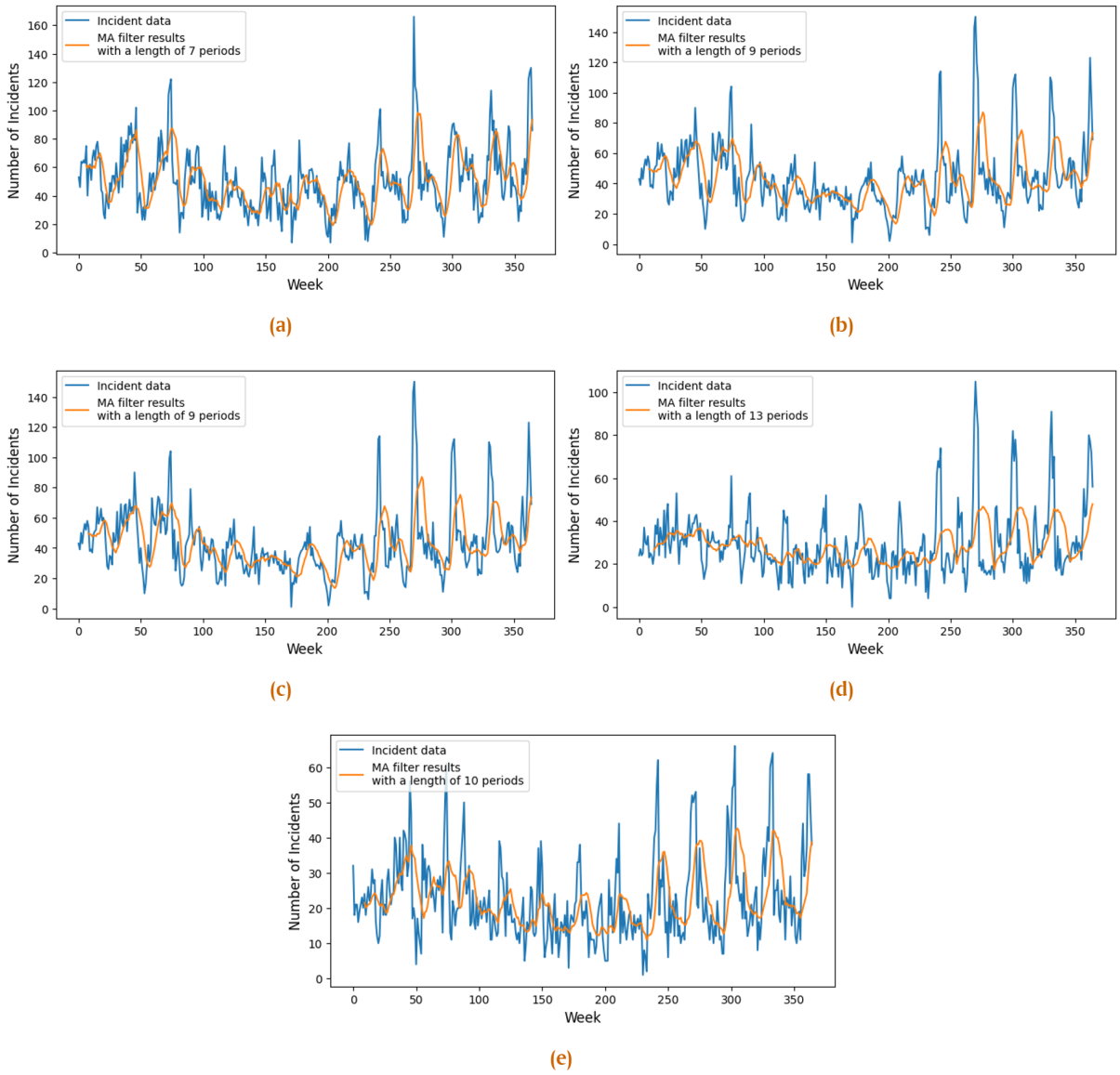


Figure 2. Weekly Pneumonia Cases and Its Smoothed Trend Component in (a) West Jakarta, (b) East Jakarta, (c) North Jakarta, (d) South Jakarta, (e) Central Jakarta.

ities of Jakarta: West Jakarta, East Jakarta, North Jakarta, South Jakarta, and Central Jakarta. The data consist of 365 weeks, ranging from January 2, 2017, to December 30, 2023. This data is collected from the official surveillance system of the Jakarta Health Department [24]. The weekly data can be seen in the blue curve in Figure 2. First, the value of an optimal length MA filter is determined by selecting the one that makes the excess kurtosis closest to zero after applying the MA filter transformation to the incident data for each municipality. These optimal lengths of MA filter are shown in Table 1.

Table 1. Optimal Length of MA Filter for the Pneumonia Data in Each Municipality

Municipality	MA filter length	Excess Kurtosis
West Jakarta	7	-0.023
East Jakarta	9	0.097
North Jakarta	10	-0.098
South Jakarta	13	0.057
Central Jakarta	10	-0.083

The smoothed trend component of the incident data in each municipality is then constructed using a moving average filter, with the length specified in Table 1. Note that the smoothed trend component is not a forecast, but rather the result of a decomposition that aims to remove short-term fluctuations. The visualization of the smoothed trend component in each municipality is presented in orange curves in Figure 2.

The smoothed trend component of the weekly data in each municipality was then tested using the Augmented Dickey-Fuller (ADF) test, and has been proven to be stationary by performing at most one differencing process in several municipalities (East Jakarta, South Jakarta, and Central Jakarta). The large number of significant values in the ACF and PACF plots, caused the differencing process performed twice for every smoothed trend component in each municipality. The order selection range of the AR

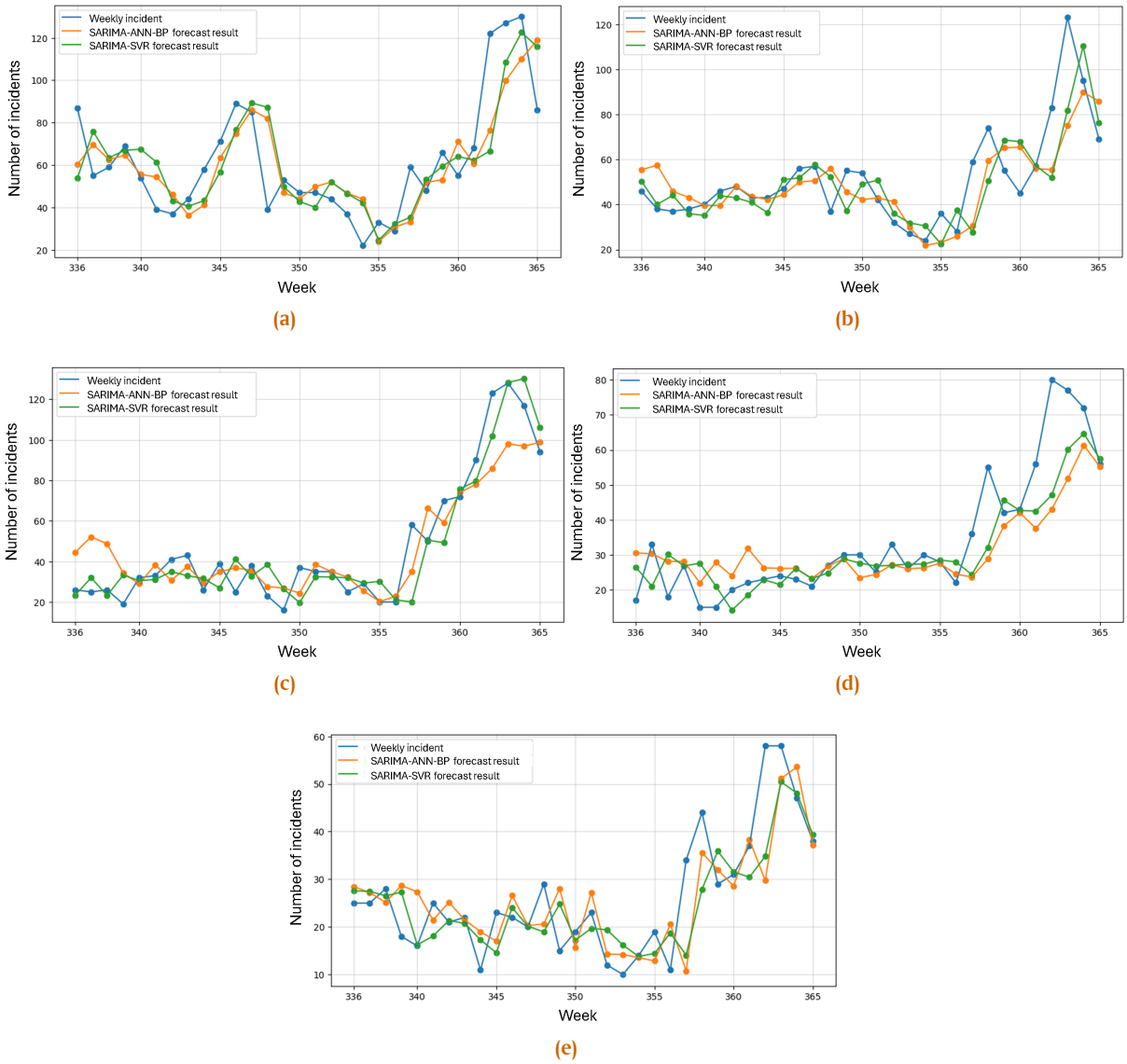


Figure 3. Forecast Results of SARIMA-ANN and SARIMA-SVR Models in (a) West Jakarta, (b) East Jakarta, (c) North Jakarta, (d) South Jakarta, (e) Central Jakarta.

and seasonal AR parameters was determined by observing the significant values in the PACF plot, meanwhile the order selection range of the MA and seasonal MA parameters was determined by considering at the significant values in the ACF plot. Table 2 presents the best SARIMA model for forecasting the smoothed trend component in each municipality by the lowest MAE and MAPE values in the final 30 weeks of the forecast period. For example, the best SARIMA model in West Jakarta has 2 differencing process, with 1 autoregressive and 1 moving average parameter, and 1 seasonal moving average parameter with 7 period of seasonality.

After hyperparameter tuning is performed using grid search, the best SARIMA model with the smallest error in each municipality can be seen in Table 2, with MAPE ranging from 1.83% to 3.65%. The Ljung-box test conducted on the forecast errors from each best SARIMA model also shows that there is no autocorrelation in the forecast errors. Afterwards, the residual component is constructed by subtracting the weekly incident data with its smoothed trend component as in eq. (11). Following that, the nonlinear ANN and SVR models are trained using the residual component.

Table 2. Best SARIMA Model for Forecasting Smoothed Trend Component in Each Municipality

Municipality	SARIMA Model	MAE	MAPE
West Jakarta	SARIMA$(1,2,1)(0,0,1)_7$	2.1510	3.65%
East Jakarta	SARIMA$(0,2,0)(0,0,1)_9$	1.2625	2.56%
North Jakarta	SARIMA$(0,2,0)(1,0,3)_{10}$	0.9767	2.59%
South Jakarta	SARIMA$(0,2,1)(2,0,1)_{13}$	0.6056	1.83%
Central Jakarta	SARIMA$(2,2,1)(1,0,3)_{10}$	0.9314	3.10%

Table 3 presents the best ANN model for forecasting residual component in each municipality by determining the lowest MAE in the final 30 weeks of the forecast period. In Table 3, the best ANN model in West Jakarta consists of 13 neurons in the input layer (uses the last 13 weeks of data), 11 neurons in the hidden layer, and 1 neuron in the output layer, using 92 epochs

Table 3. Best ANN Model to Forecast Residual Component in Each Municipality

Municipality	Architecture	Batch size	Learning rate	Epoch	MAE
West Jakarta	(13,11,1)	16	0.01	92	11.89
East Jakarta	(15,13,1)	32	0.1	18	9.40
North Jakarta	(13,11,1)	64	0.1	56	10.14
South Jakarta	(15,13,1)	64	0.1	36	7.34
Central Jakarta	(15,15,1)	32	0.1	26	5.74

Table 4. Best SVR Model to Forecast Residual Component in Each Municipality

Municipality	Length of input data	Cost	Epsilon	Gamma	MAE
West Jakarta	9	50	0.0001	0.0001	12.04
East Jakarta	15	100	0.1	0.0001	9.42
North Jakarta	7	50	0.1	0.0001	8.39
South Jakarta	15	10	0.1	0.0001	6.27
Central Jakarta	15	10	0.01	0.0001	5.24

(iterations) to train the model with a batch size of 16 (the model processes 16 training samples in each iteration to calculate the estimated error gradient) and a learning rate of 0.01 (a constant that determines how much influence the error gradient has on the weight and bias updates). The number of epochs used is determined by the early stopping technique, that can stop the learning process if the error of the test data does not improve sufficiently to avoid overfitting [25]. The MAE from the residual component forecast results using ANN in Table 3, which ranges from 5.74 to 11.89, is much larger than the MAE from the smoothed trend component forecast results, which range from 0.60 to 2.15 in Table 2. Table 4 presents the best SVR model for forecasting residual component in each municipality by selecting the lowest MAE and MAPE in the final 30 weeks of the forecast period.

The best SVR models that show the smallest error for forecasting residual component in each municipality are shown in Table 4. We can see that the MAE of the residual component forecast results obtained using SVR in each municipality is not much different from the MAE of the residual component forecast results obtained using ANN (see Table 3).

The forecast results of the residual components obtained from the best ANN models in each municipality in Table 3, and the best SVR models in each municipality in Table 4, are then combined with the forecast results of the smoothed trend components obtained using the best SARIMA models in each municipality in Table 2, to form weekly forecasts for the SARIMA-ANN method and the SARIMA-SVR method. The weekly incident forecast from both methods for the last 30 weeks are shown in Figure 3.

From Figure 3, we can see that both methods often cannot predict well whether the weekly incident rate will rise or fall drastically. However, because the forecast error on the smoothed trend component is relatively small, the problem lies with the ANN and SVR, as they cannot predict whether the residual component will increase or decrease drastically. The weekly incident forecast error from both methods for the last 30 weeks is shown in Table 5.

As shown in Table 5, unlike the forecast results of the smoothed trend component in Figure 2, which have a quite small forecast errors, the weekly incident forecast results from the SARIMA-ANN and SARIMA-SVR methods have quite large forecast errors, with MAPE values ranging from 19.18% to 25.24%. The results indicate that the SARIMA-ANN model slightly outperforms

Table 5. Forecast Error Comparison Between SARIMA-ANN and SARIMA-SVR Methods in Each Municipality

Municipality	SARIMA-ANN		SARIMA-SVR		Error Difference	
	MAE	MAPE	MAE	MAPE	MAE	MAPE
West Jakarta	**13.98**	**24.87%**	14.15	25.63%	0.17	0.76%
East Jakarta	**10.29**	**19.18%**	10.53	19.78%	0.23	0.60%
North Jakarta	11.04	28.56%	**9.07**	**24.57%**	1.96	3.99%
South Jakarta	7.61	22.84%	**6.80**	**21.19%**	0.80	1.64%
Central Jakarta	6.34	28.05%	**5.85**	**25.24%**	0.48	2.81%

The bolded parts represent the smallest error for each municipality.

the SARIMA-SVR model in West Jakarta and East Jakarta, with MAPE differences of 0.6% and 0.76, respectively. Meanwhile, the SARIMA-SVR model slightly outperforms SARIMA-ANN in North Jakarta, South Jakarta, and Central Jakarta, with MAPE differences ranging from 1.64% to 3.99%. The SARIMA-SVR model achieves better results in terms of MAE and MAPE across the majority of municipalities.

Additionally, Table 5 also shows that SARIMA-ANN achieves its smallest MAPE in East Jakarta and its highest in North Jakarta, while SARIMA-SVR also records its smallest MAPE in East Jakarta but its highest in West Jakarta. Although the MAE values for South Jakarta and Central Jakarta are relatively smaller than those for East Jakarta, this is influenced by the smaller variance in the distribution of the incident data in South and Central Jakarta compared to the broader data distribution in East Jakarta (see Figure 3). Based on these findings, the best model's performance still depends on each characteristic of the municipality data, and the forecast results for East Jakarta are best suited to the SARIMA-ANN and SARIMA-SVR models compared to other municipalities.

5. Conclusion

This study utilized the SARIMA-ANN and SARIMA-SVR methods, combined with a moving average filter to separate the time series into two components, namely the smoothed trend component (which shows low volatility), and the residual component (which exhibits high volatility), to forecast the weekly number of pneumonia cases in each municipality of Jakarta. The results show that the SARIMA-ANN model slightly outperforms the SARIMA-SVR model for forecasting in West and East Jakarta, with MAPE differences are 0.6% and 0.76%, respectively. Meanwhile, the SARIMA-SVR model slightly outperforms the SARIMA-ANN model for forecasting in North, South, and Central Jakarta, with

MAPE differences ranging from 1.64% to 3.99%. The SARIMA-SVR model outperforms the SARIMA-ANN model in terms of MAE and MAPE across three municipalities, while the SARIMA-ANN performs better in two municipalities, indicating that the SARIMA-SVR model generally provides better results.

Despite these findings, both models exhibit relatively high forecast errors, with MAPE values ranging from 19.18% to 25.24%. These high errors are primarily due to significant inaccuracies in predicting the residual components using ANN and SVR. To address this issue, future research should focus on developing and integrating more accurate methods for forecasting residual components. This could involve exploring advanced machine learning techniques, such as deep learning models, or hybrid approaches that dynamically adjust to data characteristics. Further research in this area has the potential to provide healthcare providers with more reliable tools for predicting pneumonia incidence, ultimately aiding in better resource allocation and timely medical intervention.

Author Contributions. **Musyaffa, M. R.**: software, data curation, formal analysis, writing—original draft preparation. **Hertono, G. F.**: methodology, validation, writing—review, supervision, project administration, funding acquisition. **Handari, B. D.**: Conceptualization, methodology, validation, writing—review and editing, supervision.

Acknowledgement. The authors sincerely appreciate the valuable contribution of the editors, reviewers and colleagues for their support in the research process and in improving this manuscript.

Funding. This research is fully supported by The Ministry of Education, Culture, Research, and Technology No.: PKS-480/UN2.RST/HKP.05.00/2025.

Conflict of interest. The author declares that he has no conflicts of interest to report regarding the present study.

Data availability. This data is collected from the official surveillance system of the Jakarta Health Department [24].

References

[1] A. O. Yunus, M. O. Olayiwola, and A. M. Ajileye, "A fractional mathematical model for controlling and understanding transmission dynamics in computer virus management systems," *Jambura Journal of Biomathematics (JJBM)*, vol. 5, no. 2, pp. 116–131, 2024. DOI:10.37905/jjbm.v5i2.25956

[2] WHO, "Pneumonia in children," https://www.who.int/news-room/fact-sheets/detail/pneumonia, 2022, Accessed on 10 June 2024.

[3] M. A. Khan, A. Bajwa, and S. T. Hussain, "Pneumonia: Recent updates on diagnosis and treatment," *Microorganisms*, vol. 13, no. 3, p. 522, 2025. DOI:10.3390/microorganisms13030522

[4] C. Stotts, V. F. Corrales-Medina, and K. J. Rayner, "Pneumonia-induced inflammation, resolution and cardiovascular disease: Causes, consequences and clinical opportunities," vol. 132, no. 6, pp. 751–774, 2023. DOI:10.1161/CIRCRESAHA.122.321636

[5] S. Tricahya and Z. Rustam, "Forecasting the amount of pneumonia patients in jakarta with weighted high order fuzzy time series," in *IOP Conference Series: Materials Science and Engineering*, vol. 546, no. 5, p. 052080, 2019. DOI:10.1088/1757-899X/546/5/052080.

[6] M. Khashei and M. Bijari, "A new class of hybrid models for time series forecasting," *Expert Systems with Applications*, vol. 39, no. 4, pp. 4344–4357, 2012. DOI:10.1016/j.eswa.2011.09.157

[7] L. B. Sina *et al.*, "Hybrid forecasting methods—a systematic review," *Electronics*, vol. 12, no. 9, p. 2019, 2023. DOI:10.3390/electronics12092019

[8] K. Y. Chen, "Combining linear and nonlinear model in forecasting tourism demand," *Expert Systems with Applications*, vol. 38, no. 8, pp. 10368–10376, 2011. DOI:10.1016/j.eswa.2011.02.049

[9] C. H. Aladag, E. Egrioglu, and C. Kadilar, "Forecasting nonlinear time series with a hybrid methodology," *Applied Mathematics Letters*, vol. 22, no. 9, pp. 1467–1470, 2009. DOI:10.1016/j.aml.2009.02.006

[10] P. G. Zhang, "Time series forecasting using a hybrid arima and neural network model," *Neurocomputing*, vol. 50, pp. 159–175, 2003. DOI:10.1016/S0925-2312(01)00702-0

[11] H. Moeeni and H. Bonakdari, "Forecasting monthly inflow with extreme seasonal variation using the hybrid sarima-ann model," *Stochastic Environmental Research and Risk Assessment*, vol. 31, pp. 1997–2010, 2017. DOI:10.1007/s00477-016-1273-z

[12] M. Z. Mukaram and F. Yusof, "Solar radiation forecast using hybrid sarima and ann model," *Malaysian Journal of Fundamental and Applied Sciences*, vol. 13, no. 4–1, pp. 346–350, 2017. DOI:10.11113/mjfas.v13n4-1.895

[13] M. Bouzerdoum, A. Mellit, and A. M. Pavan, "A hybrid model (sarima-svm) for short-term power forecasting of a small-scale grid-connected photovoltaic plant," *Solar Energy*, vol. 98, pp. 226–235, 2013. DOI:10.1016/j.solener.2013.10.002

[14] L. Xuemei *et al.*, "Hybrid support vector machine and arima model in building cooling prediction," *3CA 2010 - 2010 International Symposium on Computer, Communication, Control and Automation*, vol. 1, pp. 533–536, 2010. DOI:10.1109/3CA.2010.5533864

[15] C. N. Babu and B. E. Reddy, "A moving-average filter based hybrid arima-ann model for forecasting time series data," *Applied Soft Computing Journal*, vol. 23, pp. 27–38, 2014. DOI:10.1016/j.asoc.2014.05.028

[16] D. T. Thinh, N. B. H. Quan, and N. Maneetien, "Implementation of moving average filter on stm32f4 for vibration sensor application," in *Proceedings 2018 4th International Conference on Green Technology and Sustainable Development, GTSD 2018*, 2018. DOI:10.1109/GTSD.2018.8595630.

[17] W. Cahyati and M. Sari, "Forecasting of childhood pneumonia in semarang city," *Atlantis Press*, 2020. DOI:10.2991/icracos-19.2020.52

[18] R. J. Hyndman and G. Anthanasopoulos, "*Forecasting: Principles and Practice, 3rd ed.*." OTexts, 2021.

[19] Y. W. Chang and M. Y. Liao, "A seasonal arima model of tourism forecasting: The case of taiwan," *Asia Pacific Journal of Tourism Research*, vol. 15, pp. 215–221, 2010. DOI:10.1080/10941661003630001

[20] M. Awad and R. Khanna, "*Support Vector Regression*." Apress, pp. 67–80, 2015. DOI:10.1007/978-1-4302-5990-9_4

[21] D. Zhao and R. Zhang, "A new hybrid model sarima-ets-svr for seasonal influenza incidence prediction in mainland china," *Journal of Infection in Developing Countries*, vol. 17, pp. 1581–1590, 2023. DOI:10.3855/jidc.18037

[22] D. N. Joanes and C. A. Gill, "Comparing measures of sample skewness and kurtosis," *Journal of the Royal Statistical Society Series D: The Statistician*, vol. 47, pp. 183–189, 1998.

[23] X. Xian *et al.*, "Comparison of sarima model, holt-winters model and ets model in predicting the incidence of foodborne disease," *BMC Infectious Diseases*, vol. 23, 2023. DOI:10.1186/s12879-023-08799-4

[24] J. H. Department, "Weekly pneumonia cases surveillance data," https://surveilans-dinkes.jakarta.go.id/sarsbaru/rs_rekap.php, Accessed on 14 February 2024.

[25] I. Goodfellow, Y. Bengio, and A. Courville, "*Deep learning*." The MIT Press, 2017.