ISSN 2089-8673 (Print) | ISSN 2548-4265 (Online)
Volume 13, Issue 1, March 2024

SARIMA WITH SLIDING WINDOW IMPLEMENTATION FOR FORECASTING SEASONAL DEMAND DATA

Made Rama Pradipta[1], Arya Sasmita[2], Hary Susila[3]

[1,2,3]Information Technology, Faculty of Engineering, Udayana University

email: rama16181@gmail.com[1], aryasasmita@unud.ac.id[2], harysusila@unud.ac.id[3]

Abstract

Demand forecasting is an essential part of business process management. A comparison of methods is needed to get the best model to provide good forecasting results. Difficulties in meeting consumer demands and predicting these requests using demand data at companies CV. ABCD is the main problem in this research. The SARIMA and decomposition methods are used for comparison and search for the best model before forecasting. SARIMA $(1,1,1)(1,1,1)^{12}$ with a windowing size of 56, indicating the smallest MAPE value of 3,91%. The value <10%, so it can be said to produce an excellent forecasting value. Forecasting results with SARIMA $(1,1,1)(1,1,1)^{12}$ show a meeting between actual and forecasting data in 2022. Therefore, it can be said the forecasting results for 2023 and 2024 can be used as a reference for the company CV. ABCD to meet customer demand and avoid stock shortages.
.

Keywords: Forecasting, SARIMA, Decomposition, Windowing, Seasonal, Demand

Received: 28-06-2023 | **Revised:** 27-09-2023 | **Accepted:** 02-03-2024
DOI: https://doi.org/10.23887/janapati.v13i1.59971

INTRODUCTION

Demand forecasting is an essential part of business process management. According to [1], obtaining reasonably accurate forecasts on future demand for a product based on historical data and current environmental conditions (e.g., political, social, economic) is helpful for companies to be able to plan and manage their business properly. The existence of demand forecasting can help companies to match sources, production, transportation, operating activities, and actions with customer needs [1].

CV. ABCD is a company engaged in fruit processing into fresh juice products. Hotel is the target market of CV. ABCD because its effects are often used at breakfast and in the bar section. CV. ABCD has 1000 customer data, 700 active customers, and 400 customer order periods daily. CV. ABCD also has ten types of products that are traded. However, with so many customers owned, consumer demand also often experiences a significant increase, so companies find it difficult to meet consumer demand due to insufficient stock.

Based on these problems, it is necessary to forecast future consumer demand. Thus, the company can have value recommendations for safety, minimum, and maximum stock. However, to produce good forecasts, a comparison of forecasting methods is needed [2]. That was done to find the best model suitable for forecasting the data demand of the company CV. ABCD. Several techniques can be used, including the SARIMA (Seasonal Autoregressive Integrated Moving Average) and decomposition methods.

The SARIMA (Seasonal Autoregressive Integrated Moving Average) method is a method that is often used for forecasting purposes. This method is a model developed from the ARIMA model (Box Jenkins) for time series data that has a seasonal pattern [3], [4]. Meanwhile, in the decomposition method, there is a process that divides the four elements of the time series data that are owned, including factors, trends, seasonality, and cycles. The division of 4 parts is carried out to produce a more accurate forecast. There are also two types of decomposition in the decomposition itself, namely multiplicative and additive decomposition [2], [5].

Suryani's research [6] proves that the ARIMA method can be used to make predictions. The ARIMA model $(1,1,1)$ is the best model that can be used to predict death in children under five due to pneumonia in that study. ARIMA model $(1,1,1)$ fulfills the parameter significance testing stage with an overall variable value of 0.05 and is the model that has the lowest forecasting accuracy value [6].

Prianda's research [3] uses seasonal data input, showing that the Seasonal Autoregressive Integrated Moving Average (SARIMA) method can predict by having a better MAPE level than the ELM method. In this study, SARIMA had a MAPE value of 4.97%, while ELM had a MAPE value of 7.62% [3]. Whereas, Lubis's research [7] uses seasonal data input and shows that decomposition methods, one of which is the Classical Multiplicative Decomposition (CMD) method, make predictions time series with a MAPE value of 10% which means it has good forecasting results [7].

The research by Fitriastutik and Anityasari [8] uses seasonal data input. It shows the results of comparing the decomposition method with the Seasonal Autoregressive Integrated Moving Average (SARIMA). Based on the data used in this study, the Seasonal ARIMA method is a model with parameters $(0,1,1)(0,1,0)^{12}$ be the best model for forecasting the increase in the weight of waste generated by the City of Surabaya for 2020. Therefore, it can be said that the SARIMA and decomposition methods can be used as a comparison method to find the best model [8]. Another study by Konarasinghe K [9] shows the results of a comparison of Decomposition, Holt's Winter, and SARIMA. Based on the data, SARIMA is suitable for forecasting the occupancy of guest nights in Hill Country of Sri Lanka [9].

The research of Rizki et al. [10] showed that the multiplicative decomposition (seasonal) method is the best method for forecasting supply and demand analysis in one of the motorcycle tire manufacturing industries in Indonesia with MAD (Mean Absolute Deviation) of 303.577, MSE (Mean Square Error) of 157,938,700,000 and MAPE of 14.15% [10].

The research of Kromkowski et al. [11] said that a sliding window combination is used to produce better forecasting accuracy. In the study of Dong et al. [12] using the sliding window implementation of the ARIMA method. The results shown in the model's performance, the ARIMA model with the most extended sample window (five years in the study), provides the best forecast accuracy [12]. Based on this research, implementing a sliding window in the forecasting method can be carried out, and it is possible to increase the results of forecasting accuracy.

Based on previous studies, it has been shown that the SARIMA and decomposition methods can be used to forecast seasonal data. In addition, previous research has shown that the two approaches can be compared to get the best model before forecasting. Several studies also show that SARIMA is the best forecasting model. However, several other studies have shown that decomposition is the best forecasting model. Therefore, in this research, the SARIMA and decomposition methods are used to get the best model for forecasting data demand for CV. ABCD. The difference between this study and previous research is the application of a sliding window used to improve forecasting accuracy. Output on this research the company CV. ABCD can have a recommendation value on the recommended value of safety stock, minimum, and maximum stock, and it has the best model for forecasting demand data.

METHOD

To get a good forecasting model, it is necessary to compare one model with another [2]. Therefore, this study uses the SARIMA method and decomposition as a comparison. This is because the data demand is owned by the company CV. ABCD has a seasonal pattern. The research method used can be seen in Figure 1.

ISSN 2089-8673 (Print) | ISSN 2548-4265 (Online)
Volume 13, Issue 1, March 2024

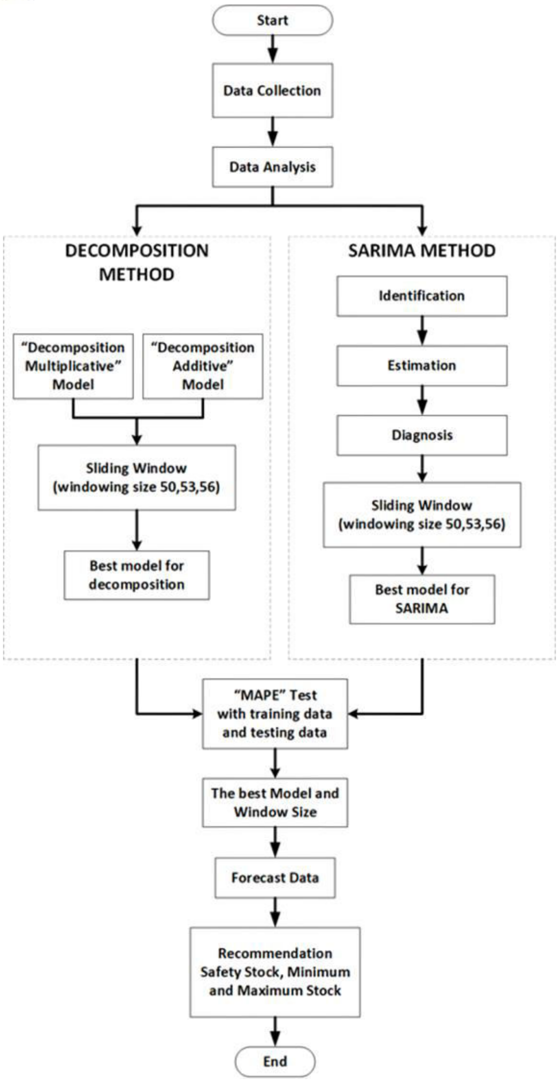

Figure 1. Research Flow

The comparison aims to get the best forecasting model from the two predetermined methods. In the SARIMA method, there will be a process of identifying, estimating, diagnosing, and applying a sliding window. Meanwhile, in the decomposition method, the model is divided into multiplicative and additive and the application of a sliding window. The best model will be determined based on the MAPE test. Then, the best model is used for forecasting to get the recommended value of safety stock, minimum, and maximum stock.

A. Data Collection

Data collection methods through observation and interviews with the company manager's CV. ABCD and make observations on how the forecasting process has previously been done. The material used in this study is data demand for the company's product CV. ABCD. Each product also has its demand data, so each product's demand data will be input into the research flow that will be carried out. The demand data obtained has a period starting from 2014 – 2019. Demand data is in the form of monthly periods. Table 1 is an example of demand data for a company CV. ABCD in 2014 on the first five data types (apple, guava, mango, orange, pineapple).

Table 1. Data Demand CV. ABCD in 2014

Month	Apple	Guava	Mango	Orange	Pineapple
1	28,457	26,587	25,457	93,732	45,886
2	28,627	26,747	25,609	94,294	46,161
3	28,318	26,458	25,332	93,274	45,661
4	29,783	27,827	26,643	98,100	48,024
5	29,314	27,389	26,224	96,556	47,268
6	30,051	28,077	26,882	98,983	48,456
7	29,539	27,600	26,425	97,299	47,632
8	29,914	27,949	26,760	98,532	48,236
9	30,052	28,079	26,884	98,988	48,458
10	29,704	27,753	26,572	97,841	47,898
11	29,487	27,550	26,378	97,125	47,547
12	26,021	24,312	23,278	85,709	41,958

B. SARIMA

Seasonal ARIMA $(p, d, q)(P, D, Q)^S$ It is a model developed from ARIMA that improves the performance of integrated autoregressive moving averages in seasonal series. The general form of the SARIMA method is stated in equation 1 below [13], [14].

$$\varphi_p(G)\varphi_p(G^{seasonal})(1-G)^d(1-G^{seasonal})^D X_t = \gamma_q(G)\omega_Q(G^{seasonal})e_t \quad (1)$$

In the above formula, the value $\varphi_p(G)$ and $\gamma_q(G)$ indicates the polynomial characteristic of the non-seasonal autoregressive (AR) and the non-seasonal component of the moving average (MA). Term from $\varphi_p(G^{seasonal})$ and $\omega_Q(G^{seasonal})$ indicates a seasonal autoregressive (SAR) and a seasonal polynomial moving average (SMA). $(1-G)$ and $(1-G^{seasonal})$ indicates time series differences, non-seasonal and seasonal. Term d and D indicate the value of non-seasonal differences in ARIMA and SARIMA. Additionally, on X_t shows the observed value at time t, e_t To indicate the predicted error value, s (seasonal) shows seasonal periods (e.g., s = 12, monthly period), and G is the backshift operator [13].

1. Identification

Before the research data can be ascertained that the data can be used in the SARIMA method, the identification stage is carried out. Not stationary data has three conditions: not stationary in the variance, not stationary in the mean, or both are not stationary. Data that is not stationary in the variant can be transformed to make the data stationary with box-cox transformation [15]. Meanwhile, the data is not stationary mean needs to be differentiated for stationary [16].

Table 2. Transformation Form

Value λ	Transformation
-1	$\frac{1}{Z_t}$
-0,5	$\frac{1}{\sqrt{Z_t}}$
0	$1n\, Z_t$
0,5	$\sqrt{Z_t}$
1	Z_t

In Table 2 above, the value Z_t is the initial data. If data is already stationary after transformation with optimal lambda (λ =1), then the data can be continued at the next stage [6], [15].

On average, data that is still non-stationary means that a differencing process is needed. Suppose at the zero order, the time series data is not stationary. In that case, the stationarity of the data can then be searched through the following order so that the stationarity level at the nth order is obtained (first difference, second difference, and so on). The differencing formula can be expressed in equation 2 [17].

$$\Delta Y_t = Y_t - Y_{t-1} \quad (2)$$

In the above formula value ΔY_t is a variable first difference at time t. Mark Y_t, Y_{t-1} is the value of the variable Y at time t and t-1 [17].

2. Estimation

After the data is stationary, the parameters are determined at the estimation stage. The parameter determination is based on displaying stationary Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots. Table 3 shows how ACF and PACF behave for seasonal and non-seasonal series [18].

Table 3. Characteristics of ACF and PACF

Model	Non – Seasonal ACF	PACF
AR (p)	Tail off at lag k	Cuts off at lag p
MA (q)	Cuts off at lag q	Tails off at lag k
ARMA (p, q)	Tails off	Tails off
Model	**Seasonal ACF**	**PACF**
AR (P)	Tail off at lag K	Cuts off at lag P
MA (Q)	Cuts off at lag Q	Tails off at lag K
ARMA (P, Q)	Tails off	Tails off

3. Diagnosis

After finding models that are suspected of being able to produce good forecasts, a diagnosis is made. At this stage, a diagnostic test uses the theory of White Noise. This process is carried out to determine a model's feasibility in selecting the best model. It is also necessary to carry out a white noise diagnostic test through the Ljung-Box test to determine whether the residue meets the requirements of the statistical hypothesis in Equation 3 and Equation 4 [19].

$$H_0 : \rho_1 = \rho_2 = \cdots = \rho_n = 0, \text{ White Noise} \quad (3)$$

$$H_1 : \rho_k \neq 0,\ k = 1,2,3,\ldots,n \text{ Not White Noise} \quad (4)$$

H_0 not accepted when $Q(Chi - Square) > X^2_{(1-\alpha),df}$ $df = K_{-p-q}$. When *P-Value* < α, then H_0 not accepted, which means it does not meet the residue process of White Noise. If value $Q(Chi - Square) < X^2_{(1-\alpha),df}$ or *P-Value* > α then H_0 acceptable means it meets the White Noise residue process with a provisioned value of α = 5% = 0, 05 [19].

Furthermore, testing is repeated to ensure the best model by conducting a normal distribution test. Figure 2 is a graphical example of the normal distribution test. The model is the best if the data is around the red line [20]. The closer the data is to the red line, the better the model is. Besides that, H_0 accepted if the P-value > α, with α as the significance level. Mark α which is often used α = 0.05 = 5%. H_0 rejected (H_1 accepted) if P-value < α [20], [21].

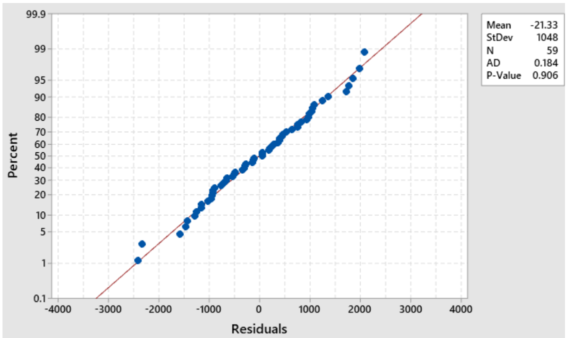

Figure 2. Normal Distribution Test

C. Decomposition

In the decomposition method, the type of decomposition is divided into two types: multiplicative and additive. The decomposition process assumes that data is influenced by four factors: Trend, Cyclical, Seasonal, and random error [5]. Equation 5 is the formula of decomposition multiplicative. Meanwhile, equation 6 is the formula for the decomposition additive [22].

$$Y = T.S.C.I \quad (5)$$

$$Y = T + S + C + I \quad (6)$$

In the equation above, the value of Y is the time series value (actual data), the T value is the trend component, the S value is the seasonality component, the C value is the cycle component, and the I value is the error or random element [22].

D. Sliding Window

Windowing, commonly called a sliding window, forms the available time series data structure. In the windowing process, changing a window's size produces the lowest error value [23]. The sliding window stage forms the segmentation of weekly or monthly pattern data to predict the next day's data [24].

Figure 3 shows the sliding window process with a window size of 50. The number in the box shows the monthly data for the company's CV. ABCD. The number 1 indicates the 1st month, 2 indicates the 2nd month, 3 indicates the 3rd month, and so on. Using a windowing size of 50, months 1 to 50 will be used to predict month 51. Then, later, with a windowing size of 50, data from months 2 to 51 are used to indicate the value in the 52nd month. And so on until all data is forecasted.

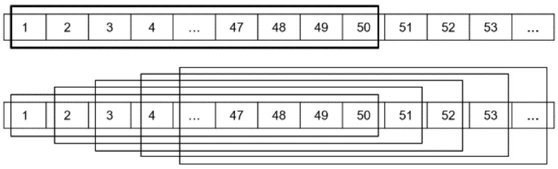

Figure 3. Sliding Window

E. MAPE Test

In order to know the value of forecasting errors that occur, a method is needed, Mean Absolute Percentage Error (MAPE), to see the error level between the actual data and forecasting results. The forecasting results will be better if the average percentage error value produces a smaller value. However, the forecasting results will worsen if the value obtained is more significant. The Mean Absolute Percentage Error (MAPE) is mathematically formulated in equation 7 [25].

$$MAPE = \frac{\sum \frac{|\varepsilon_i|}{X_i} 100\%}{n} = \frac{\sum \frac{|X_i - F_i|}{X_i} 100\%}{n} \quad (7)$$

In the above formula, the value X_i is actual data, value F_i forecasting value of value X_i And value n is the number of periods of forecasting which are involved. The MAPE value has an interpretation consisting of 4 parts described in Table 4 [25].

Table 4. MAPE Interpretation

Results	Interpretation
< 10%	Very Good Forecasting Results
(10 – 20) %	Good Forecasting Results
(20 – 50) %	Fairly Good Forecasting Results
> 50%	Bad Forecasting Results

If the results of the MAPE value are < 10%, then it can be said that the forecasting results are excellent. If the results of the MAPE value are (10 – 20) %, then it can be said that the forecasting results are promising. If the results of the MAPE value are (20 – 50) %, then it can be said that the forecasting results are pretty good. If the results of the MAPE value are > 50 %, then it can be said that the forecasting results are wrong.

RESULT AND DISCUSSION

The results and discussion will consist of the process of the SARIMA method in determining the model, the operation of the decomposition method, the MAPE test from a comparison of the SARIMA and decomposition methods, the application of sliding windows and the MAPE test for each window size, and the results of the recommendations given. The following describes each of these processes.

A. SARIMA (Identification)

At this stage, the data identification process, which includes the process, is carried out time series plot, box–cox transformation, and differences. Figure 4 shows the data demand on the product Apple, which is still not stationary in the variance because the data fluctuations are still not constant. It can also be seen that the data demand product Apple is still not stationary in the average because the changes are not yet around a constant average value. Based on the series view plot in the data demand product Apple it can be assumed that the data is seasonal due to fluctuations in specific periods.

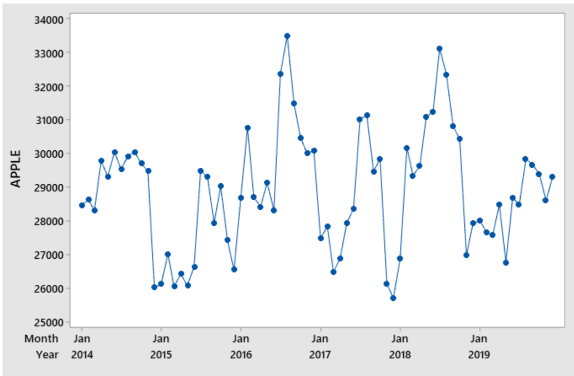

Figure 4. Apple's Time Series Plots

Figure 5 is a display time series plot multiple showing the overall comparison trend product. Based on the results of the time series plot multiple analysis, each product has a different quantity range. The difference in quantity illustrates the level of product interest based on the demanding consumer.

ISSN 2089-8673 (Print) | ISSN 2548-4265 (Online)
Volume 13, Issue 1, March 2024

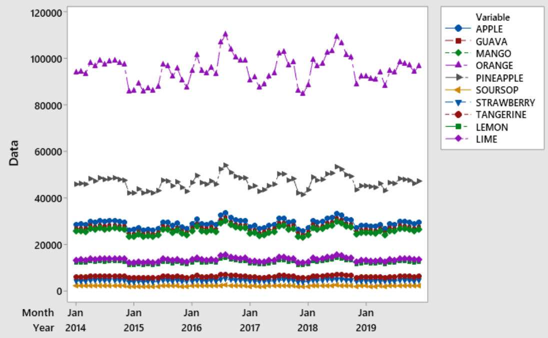

Figure 5. Time series Plot Display Multiple for all products

The next stage is the stationarity process for the variance by using the Box-Cox transformation. Table 5 shows the results of the Box-cox transformation on each product. It is obtained that each product needs to do Box-cox transformation three times to reach a rounded value = 1. Different values occur in transformation 2nd. However, at stage transformation, the 3rd round value of all products can get 1.00. So, it can be defined that the data demand for all products has reached stationary in variants.

Table 5. Box-cox Transformation Results for All Products

Item Name	Box-1	Box-2	Box-3
APPLE	0.00	1.69	1.00
GUAVA	0.00	1.67	1.00
MANGO	0.00	1.68	1.00
ORANGE	0.00	1.78	1.00
PINEAPPLE	0.00	1.72	1.00
SOURSOP	0.00	1.61	1.00
STRAWBERRY	0.00	1.44	1.00
TANGERINE	0.00	1.66	1.00
LEMON	0.00	1.66	1.00
LIME	0.00	1.67	1.00

At the level of differences, creating data derivatives time series as much as d times. Figure 6 displays the plot Autocorrelation Function (ACF), which shows that the data is still not stationary in the average because the data is still visibly dying down.

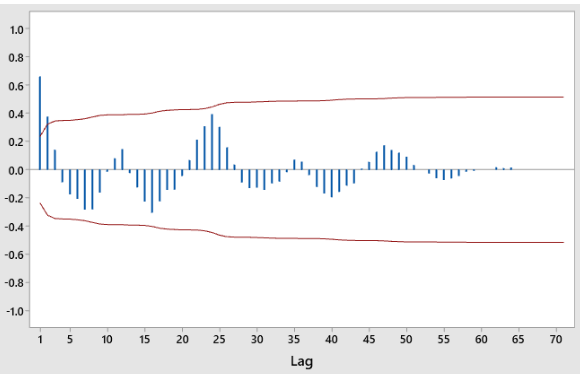

Figure 6. Display of ACF demand for Apple products

Figure 7 shows that in lag 12, the significant value is relatively high, namely T = 2.64, and in lag 24, the T = 1.75. With significant values at lag 12 and lag 24, the data can be categorized into seasonal data with a period of 12. Therefore, for the data to be stationary in the season, it is essential to differences with lag 12.

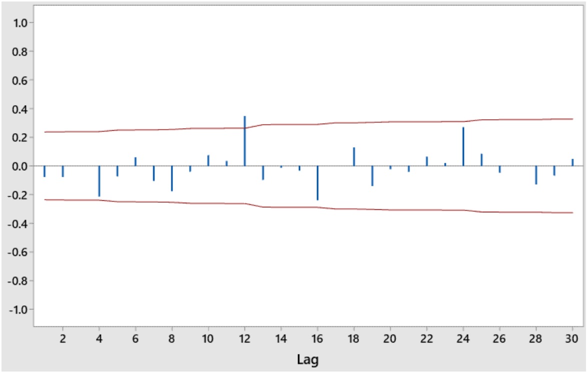

Figure 7. ACF Plot After First Apple Product Differences

Figure 8 is the display after making its differences with a lag of 12. Based on the ACF plot in Figure 8, it is stationary both in variance and average. The data is stationary because there is no pattern dying down or significant lag value at any given time.

Other products besides Apple have the same appearance as the ACF plot and the same PACF. This is because the data obtained has a similar trend on each product demand data. So, this can produce the best SARIMA model, which each data demand product can later use on applying the method Seasonal Autoregressive Integrated Moving Average (SARIMA).

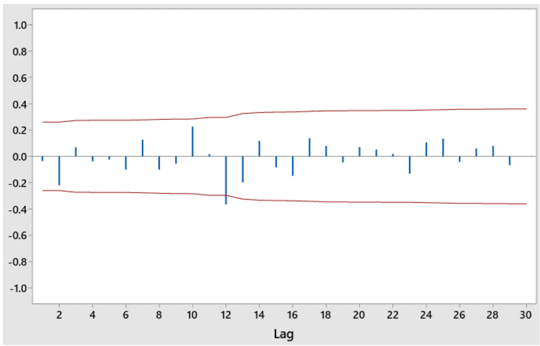

Figure 8. ACF Plot After Differences lag 12 Apple Products

B. SARIMA (Estimation)

In the process of estimating the SARIMA model, parameter values are determined $(p,d,q)(P,D,Q)^S$. The determination of these parameters is based on the display of the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots which are stationary in variance, average and seasonal effects. Figure 9 displays the ACF data plot demand for stationary Apple products.

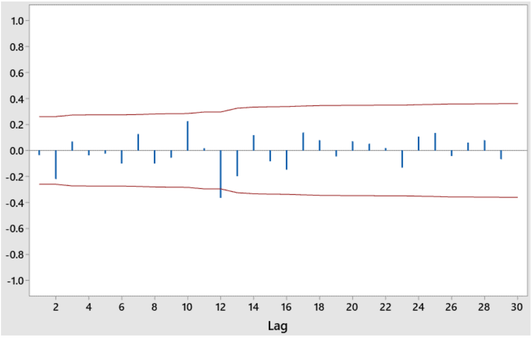

Figure 9. Display of the ACF data demand plot for Apple products

Figure 10 displays the PACF data plot demand for stationary Apple products. According to [18], the estimation of the SARIMA parameter (p,d,q)(P,D,Q) can be identified from the ACF and PACF plots. The ACF plot indicates the q and Q values in the SARIMA model, and the PACF plot indicates the p and P values in the SARIMA model. Based on Figure 9 and Figure 10, it is shown that there are cuts off after lag 2. Thus, it is suggested that the possible values of q are q=0, q=1, and q=2. Meanwhile, the possible p are p=0, p=1, and p=2. According to [26], if the data at lag 12 (seasonal lag) is significant in the ACF and PACF plots and other seasonal lags (lags 24, 36, etc.) die down to zero, then it is recommended to value P=1 and Q=1. Then, the values for d and D show the difference values. Based on the results of this study, it was shown that the values of d = 1 and D = 1 were due to differences twice in non-seasonal and seasonal differences. However, according to [27], based on parameter testing that has been carried out using the final estimates of parameters, the models that meet the P-Value < α (0,05) include the SARIMA model $(1,1,1)(1,1,1)^{12}$, $(0,1,1)(1,1,1)^{12}$, and $(0,1,0)(1,1,1)^{12}$.

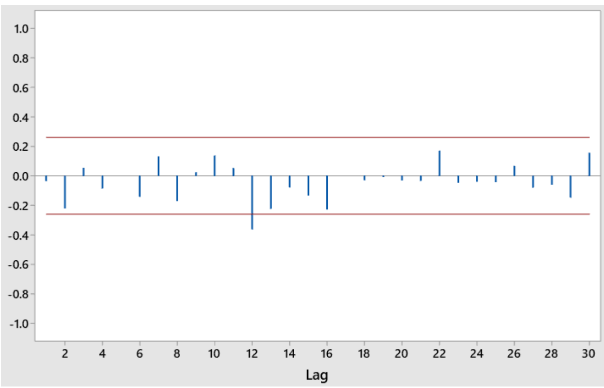

Figure 10. Display of the PACF plot on-demand data Apple products

C. SARIMA (Diagnosa)

At the diagnosis level, there is a diagnosis of white noise and normal distribution to determine the nature of the model white noise. Table 6 is the result Box–Pierce model $(1,1,1)(1,1,1)^{12}$ to perform a white noise test. The P-value for all lags > α (0,05) in the table. Therefore, it can be said that the model $(1,1,1)(1,1,1)^{12}$ accepted, which means the model meets the residue process of White Noise.

Table 6. Box – Pierce Model Results $(1,1,1)(1,1,1)^{12}$

Lag	12	24	36	48
Chi-Square	10.66	27.45	42.75	49.48
DF	8	20	32	44
P-Value	0.221	0.123	0.097	0.264

Table 7 is the result Box – Pierce model $(0,1,1)(1,1,1)^{12}$ to perform a white noise test. P-Value at lag 12,24, and 36 < α (0,05). Therefore, it can be said that the model $(0,1,1)(1,1,1)^{12}$ not accepted means the model does not meet the residue process White Noise.

Table 7. Box – Pierce Model Results $(0,1,1)(1,1,1)^{12}$

Lag	12	24	36	48
Chi-Square	17.17	39.42	54.82	60.33
DF	9	21	33	45
P-Value	0.046	0.009	0.010	0.063

Table 8 is the result Box – Pierce model $(0,1,0)(1,1,1)^{12}$ to perform a white noise test. The P-Value for all lags > α (0,05). Therefore, it can be said that the model $(0,1,0)(1,1,1)^{12}$ accepted, which means the model meets the residue process of White Noise.

Table 8. Box – Pierce Model Results $(0,1,0)(1,1,1)^{12}$

Lag	12	24	36	48
Chi-Square	14.94	25.18	42.70	50.45
DF	10	22	34	46
P-Value	0.134	0.288	0.146	0.302

Based on the white noise process that has been done, two models are obtained that fulfill the residual White Noise process. The model will be determined by comparing the residual white noise results with the normal distribution. All models meet the normal distribution because the overall value of p–value > α (0,05). However, by paying attention to the graph, it can be seen in Figure 11 that the SARIMA model $(1,1,1)(1,1,1)^{12}$ is the best model because the data is around the red line.

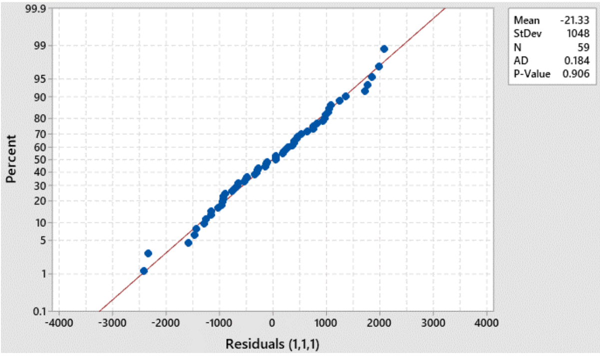

Figure 11. Display of Normal Distribution

Until the SARIMA model $(1,1,1)(1,1,1)^{12}$ selected as the best model in the SARIMA method for forecasting. That is because the model has fulfilled the stages of diagnosis, namely showing the SARIMA model $(1,1,1)(1,1,1)^{12}$ characteristic white noise.

D. Decomposition

In the process of the decomposition method, there are two types of model decomposition which will later be compared with the SARIMA method: method decomposition multiplicative and decomposition additive. Figure 12 is a graphical description produced by the method decomposition multiplicative. There are three graphs: actual, fits, and graphics trend. The blue color shows the basic graph, the red color indicates the chart fits, and the green color shows the graph trend. The measurement results include MAPE with a value of 4, MAD with a value of 1170, and MSD with a value of 1893145.

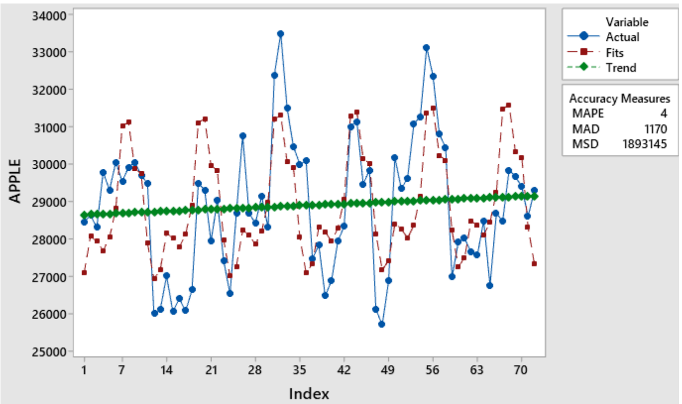

Figure 12. Display of multiplicative decomposition

Figure 13 is a graphical description produced by the method decomposition additive. The measurement results include MAPE with a value of 4, MAD with a value of 1168, and MSD with a value of 1890063. If the MAPE values are known to have the same value, then the two decomposition models will be tested directly by MAPE with the SARIMA model. Because the MAPE value produced by the two models has the same value, both models are tested for MAPE with the SARIMA model that has been determined. Based on Figures 12 and 13, it can be seen that the MAD value in the additive decomposition model is smaller than in the multiplicative decomposition model. Therefore, the additive decomposition model is the best model for the decomposition method.

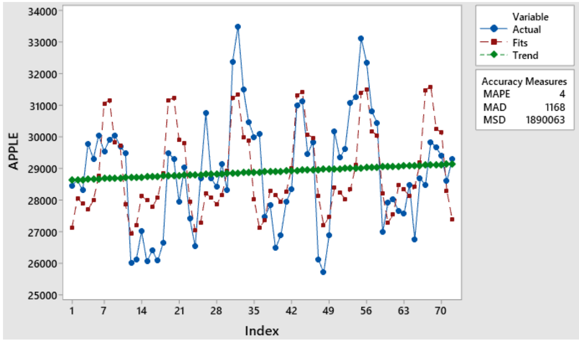

Figure 13. Display of additive decomposition

E. Sliding Window

After determining the best model in the SARIMA method, namely the SARIMA$(1,1,1)(1,1,1)^{12}$ model, and in the decomposition method, namely the additive decomposition model, the sliding window implementation was carried out on both models. In the sliding window process, the data used as training data is data from January 2014 to December 2018. Many possible sizes can be used in choosing the window size, but due to feasibility considerations, we select the candidate values for the time window through trial and error. The values tested ranged from 40 to 60. Based on the results of the experiments, we showed the three selected values for the time windows that represented the results of the experiments on both SARIMA and additive decomposition methods. Table 9 shows the forecasting value of the SARIMA model $(1,1,1)(1,1,1)^{12}$ in 2019.

Table 9. Forecasting results with SARIMA in 2019

Year	Month	W - 49	W - 50	W - 56
2019	Jan	27,257	27,242	26,834
2019	Feb	27,577	27,753	27,636
2019	Mar	25,891	26,088	26,359
2019	Apr	26,012	26,194	27,305
2019	May	26,440	26,631	28,450
2019	Jun	26,366	26,598	29,379
2019	Jul	29,998	30,226	30,981
2019	Aug	30,891	31,033	31,332
2019	Sep	29,103	29,234	29,634
2019	Oct	29,302	29,421	29,734
2019	Nov	27,600	27,705	27,786
2019	Des	26,964	26,799	27,123

Meanwhile, table 10 shows the value of the forecasting results from the additive decomposition model in 2019.

Table 10. Forecasting results with decomposition in 2019

Year	Month	W - 49	W - 50	W - 56
2019	Jan	28,235	28,131	28,031
2019	Feb	30,773	30,773	29,917
2019	Mar	29,518	29,519	28,986
2019	Apr	29,007	29,083	29,008
2019	May	29,545	29,607	29,373
2019	Jun	30,243	30,269	29,815
2019	Jul	33,157	33,160	32,788
2019	Aug	33,169	33,255	32,965
2019	Sep	31,232	31,326	31,292
2019	Oct	31,041	31,157	31,309
2019	Nov	27,259	27,383	28,500
2019	Des	27,999	28,080	28,409

F. MAPE Test

In the MAPE test, calculations are performed using data train and data testing 9:1 to determine the value of MAPE. The following is the calculation result of the MAPE test between the SARIMA and decomposition methods. Table 11 shows the results of the percentage of MAPE in the SARIMA method with the implementation of a sliding window. It can be seen in the table that the MAPE value of the SARIMA model $(1,1,1)(1,1,1)^{12}$ with windowing size 56 has the smallest MAPE value of 3.91%.

Meanwhile, table 12 shows the percentage results of MAPE in the additive decomposition method by implementing a sliding window. It can be seen in the table that the value of MAPE model decomposition additive with window size 56 has the smallest value of 5.82%.

Table 11. MAPE Test Results of the SARIMA method with Sliding Window

Year	Month	W-49	W-50	W-56
2019	Jan	0.03	0.03	0.04
2019	Feb	0.00	0.00	0.00
2019	Mar	0.06	0.05	0.04
2019	Apr	0.09	0.08	0.04
2019	May	0.01	0.00	0.06
2019	Jun	0.08	0.07	0.02
2019	Jul	0.05	0.06	0.09
2019	Aug	0.04	0.04	0.05
2019	Sep	0.02	0.01	0.00
2019	Oct	0.00	0.00	0.01
2019	Nov	0.04	0.03	0.03
2019	Des	0.08	0.09	0.07
Sum		0.50	0.48	0.47
MAPE (%)		4.14	3.97	3.91

Table 12. MAPE Test Results Additive Decomposition method with Sliding Window

Year	Month	W-49	W-50	W-56
2019	Jan	0.01	0.00	0.00
2019	Feb	0.11	0.11	0.08
2019	Mar	0.07	0.07	0.05
2019	Apr	0.02	0.02	0.02
2019	May	0.10	0.11	0.10
2019	Jun	0.05	0.06	0.04
2019	Jul	0.16	0.16	0.15
2019	Aug	0.11	0.11	0.10
2019	Sep	0.05	0.06	0.05
2019	Oct	0.06	0.06	0.07
2019	Nov	0.05	0.04	0.00
2019	Des	0.04	0.04	0.03
Sum		0.84	0.85	0.70
MAPE (%)		7.02	7.07	5.82

Therefore, the SARIMA model $(1,1,1)(1,1,1)^{12}$ is the best model for forecasting data demand for each product of the company CV. ABCD. This is due to the MAPE value in the SARIMA model $(1,1,1)(1,1,1)^{12}$ windowing size 56 is the smallest value of 3.91% compared to the additive windowing size 56 decomposition model of 5.82%.

G. Recommendation

Based on the research results above, it is determined that the model that can be used with the lowest level of MAPE value is 3.91% in the ratio of 9:1 train testing, namely the SARIMA model $(1,1,1)(1,1,1)^{12}$ with sliding window technique using windowing size 56. This model forecasts the recommended value of safety stock, minimum and maximum stock based on data demand at the company CV. ABCD. Figure 14 is a graphical display of some products. The products in the picture include the color of apple blue, the color of guava orange, the color of mango gray, the color of the orange is yellow, and the color of pineapple green. The actual data in the graph is darker, while the yield values forecasting has a more faded color.

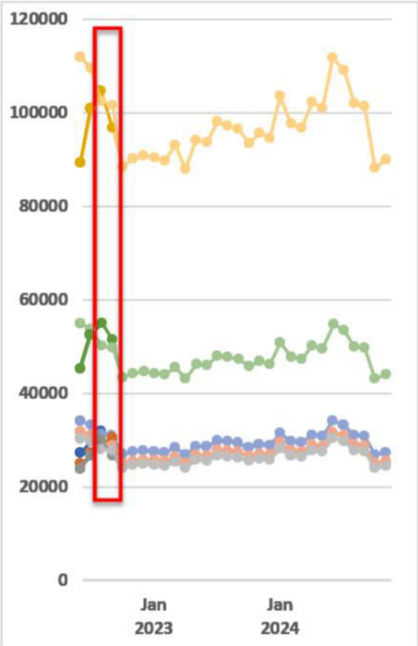

Figure 14. Graph of comparison of actual data and forecast results

In the picture above, the red box shows the meeting between the actual data and the results forecasting which has been done. The meeting between actual data and results forecasting is close enough. Therefore, the following forecasting results in 2023 and 2024 can be a reference for recommending safety, minimum, and maximum stock for companies CV. ABCD. Likewise, the chart movements show similar movements for other types of products. Table 13 shows the recommendations for each kind of fruit for the company's CV. ABCD. In the table, the recommendation data provided include minimum stock (Min), safety stock (SS), and maximum stock (Max) for 2023 and 2024.

Table 13. Display of recommendation results for each fruit

		2023			
-	**Apple**	**Mango**	**Pineapple**	**Strawberry**	**Lemon**
Min	26,684	23,870	43,027	3,970	11,236
SS	28,371	25,378	45,746	4,221	11,947
Max	29,709	26,575	47,904	4,420	12,510
-	**Guava**	**Orange**	**Soursop**	**Tangerine**	**Lime**
Min	24,931	87,890	1,817	5,555	12,263
SS	26,507	93,446	1,932	5,906	13,038
Max	27,758	97,853	2,023	6,185	13,653
		2024			
-	**Apple**	**Mango**	**Pineapple**	**Strawberry**	**Lemon**
Min	26,738	23,918	43,113	3,978	11,259
SS	30,661	27,427	49,439	4,562	12,911
Max	33,895	30,319	54,652	5,043	14,272
-	**Guava**	**Orange**	**Soursop**	**Tangerine**	**Lime**
Min	24,981	88,067	1,820	5,566	12,287
SS	28,647	100,988	2,087	6,383	14,090
Max	31,668	111,638	2,308	7,056	15,576

CONCLUSION

The best model for the comparison of the SARIMA and decomposition methods is the SARIMA model $(1,1,1)(1,1,1)^{12}$ with a windowing size of 56 with a MAPE value of 3.91%. The MAPE value is <10%, so the model can produce excellent forecasts. Based on the graph, when forecasting is done with the SARIMA model $(1,1,1)(1,1,1)^{12}$, there is a meeting between actual data and forecast data. Therefore, the best model for forecasting each type of fruit in CV. ABCD is the SARIMA model $(1,1,1)(1,1,1)^{12}$ with a windowing size of 56. Recommendation values in 2023 and 2024 for each fruit can be used as a reference for the company's CV. ABCD to meet customer demand and avoid stock shortages.

Based on the research, further analysis can also be suggested to use this method to forecast seasonal time series data other than comparing the SARIMA and decomposition methods. Different ways include double exponential smoothing, winter's method, and others. In addition, it is also recommended to use techniques of sliding windows with windowing size and utilize different models in application sliding windows. This is done to know the effect of sliding windows on other methods.

REFERENCES

[1] G. Merkuryeva, A. Valberga, and A. Smirnov, “Demand forecasting in pharmaceutical supply chains: A case study,” in *Procedia Computer Science*, Elsevier B.V., 2019, pp. 3–10. doi: 10.1016/j.procs.2019.01.100.

[2] D. Yuni, “Perbandingan Metode Exponential Smoothing dan Metode Decomposition Untuk Meramalkan Persediaan Beras (Studi Kasus Divre Bulog Lhokseumawe),” *Jurnal Visioner & Strategis*, vol. 10, no. 1, 2021.

[3] B. G. Prianda and E. Widodo, “PERBANDINGAN METODE SEASONAL ARIMA DAN EXTREME LEARNING MACHINE PADA PERAMALAN JUMLAH WISATAWAN MANCANEGARA KE BALI,” *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, vol. 15, no. 4, pp. 639–650, Dec. 2021, doi: 10.30598/barekengvol15iss4pp639-650.

[4] L. Parviz, “Comparative Evaluation of Hybrid SARIMA and Machine Learning Techniques Based on Time Varying and Decomposition of Precipitation Time Series,” 2020.

[5] G. Idfi, A. Yulistyorini, T. Rahayuningsih, V. A. K. Dewi, and E. Setyawan, “The forecasting model of discharge at Brantas sub-basin using autoregressive integrated moving average (ARIMA) and decomposition methods,” in *IOP Conference Series: Earth and Environmental Science*, IOP Publishing Ltd, Sep. 2021. doi: 10.1088/1755-1315/847/1/012029.

[6] Ni Kadek Ary Indah Suryani, Oka Sudana, and Ayu Wirdiani, “Forecasting Pneumonia Toddler Mortality Using Comparative Model ARIMA and Multilayer Perceptron,” *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 6, no. 4, pp. 528–537, Aug. 2022, doi: 10.29207/resti.v6i4.4106.

[7] I. A. Lubis and P. Korespondensi, “USULAN PERENCANAAN SAFETY STOCK & FORECASTING DEMAND DENGAN METODE TIME SERIES PRODUKSI KERAN AIR DI PT KAYU PERKASA RAYA.” [Online]. Available: www.bps.go.id,

[8] E. Fitriastutik and M. Anityasari, “Forecasting Timbulan Sampah Kota Surabaya Menggunakan Time Series Analysis,” *Jurnal Teknik ITS*, vol. 9, no. 2, Feb. 2021, doi: 10.12962/j23373539.v9i2.56557.

[9] K. M. U. B. Konarasinghe, “Forecasting Foreign Guest Nights in Hill Country of Sri Lanka,” *Review of Integrative Business and Economics Research*, vol. 7, p. 41, 2018, [Online]. Available: https://www.researchgate.net/publication/342503004

[10] A. Rizki, G. Baskoro, and I. Mariza, “Supply and Demand Analysis by using Comparison of Forecasting Method in Motorcycles Tires Manufacturer,” *Conference on Management and Engineering in Industry (CMEI)*, 2021.

[11] P. Kromkowski, S. Li, W. Zhao, B. Abraham, A. Osborne, and D. E. Brown, “Evaluating Statistical Models for Network Traffic Anomaly Detection,” 2019.

[12] H. Dong, X. Guo, H. Reichgelt, and R. Hu, “Predictive power of ARIMA models in forecasting equity returns: a sliding window method,” *Journal of Asset Management*, vol. 21, no. 6, pp. 549–566, Oct. 2020, doi: 10.1057/s41260-020-00184-z.

[13] P. Manigandan *et al.*, “Forecasting natural gas production and consumption in united states-evidence from sarima and sarimax models,” *Energies (Basel)*, vol. 14, no. 19, Oct. 2021, doi: 10.3390/en14196021.

[14] L. Martínez-Acosta, J. P. Medrano-Barboza, Á. López-Ramos, J. F. R. López, and Á. A. López-Lambraño, “SARIMA approach to generating synthetic monthly rainfall in the Sinú river watershed in Colombia,” *Atmosphere (Basel)*, vol. 11, no. 6, Jun. 2020, doi: 10.3390/atmos11060602.

[15] A. Syarif, “Forecasting the Development of Islamic Bank in Indonesia: Adopting ARIMA Model,” *JTAM (Jurnal Teori dan Aplikasi Matematika)*, vol. 4, no. 2, p. 190, Oct. 2020, doi: 10.31764/jtam.v4i2.2790.

[16] L. Louisa, R. Fauzi, and E. Setiawan Nugraha, “Forecasting of Retirement Insurance Filled via Internet by ARIMA Models,” 2022.

[17] T. I. Aulia and A. Atiqi Rohmawati, “Prediksi Nilai Ekstrem Jumlah Kasus Harian Positif COVID-19 di Provinsi Jawa Timur dengan Model Vector Autoregressive Moving Average (VARMA).” [Online]. Available: https://kawalcovid19.id/.

[18] T. M. Wanjuki, A. Wagala, and D. K. Muriithi, “Evaluating the Predictive Ability of Seasonal Autoregressive Integrated Moving Average (SARIMA) Models using Food and Beverages Price Index in Kenya,” *European Journal of Mathematics and Statistics*, vol. 3, no. 2, pp. 28–38, Apr. 2022, doi: 10.24018/ejmath.2022.3.2.100.

[19] D. Wati, “PERAMALAN JUMLAH PENUMPANG KEBERANGKATAN BUS DI TERMINAL PURABAYA MENGGUNAKAN METODE SARIMA (SEASONAL AUTOREGRESSIVE INTEGRATED MOVING AVERAGE),” Surabaya, 2020.

[20] S. A. M. Tayib, S. R. M. Nor, and S. M. Norrulashikin, “Forecasting on the crude palm oil production in Malaysia using SARIMA Model,” in *Journal of Physics: Conference Series*, IOP Publishing Ltd, Aug. 2021. doi: 10.1088/1742-6596/1988/1/012106.

[21] R. A. Permana and D. Ikasari, “UJI NORMALITAS DATA MENGGUNAKAN METODE EMPIRICAL DISTRIBUTION FUNCTION DENGAN MEMANFAATKAN MATLAB DAN MINITAB 19,” 2023.

[22] W. G. S. Konarasinghe, “Comparison of Forecasting ability of Sama Circular

Model, ARIMA and Decomposition Techniques," 2020, [Online]. Available: www.imathm.edu.lk/publications

[23] R. E. Wahyuni, B. Provinsi, and J. Tengah, "OPTIMASI PREDIKSI INFLASI DENGAN NEURAL NETWORK PADA TAHAP WINDOWING: ADAKAH PENGARUH PERBEDAAN WINDOW SIZE?," 2021.

[24] Dwi Kartini, Friska Abadi, and Triando Hamonangan Saragih, "Prediksi Tinggi Permukaan Air Waduk Menggunakan Artificial Neural Network Berbasis Sliding Window," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 5, no. 1, pp. 39–44, Feb. 2021, doi: 10.29207/resti.v5i1.2602.

[25] N. Putu, R. Apriyanti, K. Gede, D. Putra, I. Made, and S. Putra, "Peramalan Jumlah Kecelakaan Lalu Lintas Menggunakan Metode Support Vector Regression," 2020.

[26] A. Otu, O. George A., O. Jude, M. Hope Ifeyinwa, and I. Andrew I., "Application of Sarima Models in Modelling and Forecasting Nigeria's Inflation Rates," *Am J Appl Math Stat*, vol. 2, no. 1, pp. 16–28, Jan. 2014, doi: 10.12691/ajams-2-1-4.

[27] A. A. Sidora and L. Handayani, *Autoregressive Integrated Moving Average (Arima) Modelling to Forecast Rainfall In Palu*, vol. 1, no. 1. 2016.