

Rainfall Forecasting Using the Singular Spectrum Analysis (SSA) Method

Nurhikmawati, Aswi Aswi, Ansari Saleh Ahmar

Universitas Negeri Makassar, Makassar, Indonesia

Article Info

Article history:

Received : 11-04-2024

Revised : 01-21-2025

Accepted : 07-10-2025

Keywords:

Forecasting;

Rainfall;

Singular Spectrum Analysis.

ABSTRACT

This study aims to evaluate the accuracy and performance of rainfall data forecasting in the city of Parepare using the Singular Spectrum Analysis (SSA) method. Situated in South Sulawesi Province, Parepare City is characterized by high rainfall intensity, which increases the likelihood of natural hazards such as flooding and landslides. These disasters have the potential to negatively impact key sectors, including economic activity, tourism, and transportation. Therefore, reliable rainfall prediction plays a crucial role in establishing a robust disaster early warning system. Monthly rainfall measurements from two stations, Bukit Harapan and Bulu Dua, are analyzed. The results reveal a Root Mean Square Error (RMSE) of 191.0566 for Bukit Harapan station and 346.023 for Bulu Dua station, underscoring the method's forecasting accuracy. A 12-month forecast predicts consistently high monthly rainfall in Parepare City, with the highest rainfall expected in December 2024 at Bukit Harapan station and in January 2024 at Bulu Dua station. Conversely, the lowest rainfall is anticipated at both stations in July 2024. Forecasts predicting increased rainfall during specific periods, particularly in December and January, offer critical insights for enhancing disaster preparedness and informing mitigation strategies. This information also plays a key role in minimizing adverse effects on the economic, transportation, and tourism sectors, while promoting more efficient and sustainable management of water resources.

Accredited by Kemenristekdikti, Decree No: 200/M/KPT/2020

DOI: <https://doi.org/10.30812/varian.v8i2.4571>



Corresponding Author:

Aswi Aswi,

Department of Statistics, Universitas Negeri Makassar,

Email: aswi@unm.ac.id

Copyright ©2025 The Authors.

This is an open access article under the [CC BY-SA](#) license.



How to Cite:

Nurhikmawati, N., Aswi, A., & Ahmar, A. S. (2025). Rainfall Forecasting Using the Singular Spectrum Analysis (SSA) Method. *Jurnal Varian*, 8(2), 233-248.

This is an open access article under the CC BY-SA license (<https://creativecommons.org/licenses/by-sa/4.0/>)

A. INTRODUCTION

A time series refers to a sequence of observational data recorded at regular time intervals in a specific order. Time series models aim to predict future conditions by analyzing historical data and extending observed patterns into future periods (Aswi & Sukarna, 2006). The purpose of time series analysis is to predict future trends, analyze relationships between variables, monitor processes, and determine whether they are operating within controlled parameters (Zulhijrah et al., 2023). Time series forecasting can be applied across a wide range of fields, including economics, finance, marketing, production, operations research, public administration, meteorology, geophysics, demography, and education (Setiawan et al., 2019).

Forecasting relies on pre-existing data and involves selecting appropriate forecasting methods to generate accurate and reliable

predictions (Sergio et al., 2023). A crucial step in selecting an appropriate forecasting model is to carefully consider the type of data pattern being analyzed (Hidayat et al., 2020). Data patterns can generally be categorized into four types: horizontal, seasonal, trend and cyclical. In many cases, time series data with seasonal patterns are commonly encountered during the forecasting modeling process (Hidayat et al., 2020). Several time series analysis methods can be used to analyze data with seasonal patterns, including the Seasonal Autoregressive Integrated Moving Average (SARIMA) and Holt-Winters methods (Aswi & Sukarna, 2006), and a more recent development in time series analysis is the Singular Spectrum Analysis (SSA) method.

SSA is a time series analysis method that integrates elements of classical time series analysis, multivariate statistics, geometry, dynamical systems, and signal processing (Purnama, 2022). According to Utami et al. (2021), SSA is a non-parametric method, recognized for its ability to forecast data that follows specific patterns, such as seasonality. It is more flexible than comparable forecasting models, as it does not rely on assumptions like stationarity or residual, normality and does not require logarithmic transformations (Hidayat et al., 2020). In SSA, there are two forecasting methods: the recurrent method (R-Forecasting) and the vector method (V-Forecasting). The recurrent method is the more commonly used approach due to its relative simplicity and ease of application. In SSA, there are two forecasting methods, the recurrent method (R-Forecasting) and the vector method (V-Forecasting). The recurrent method is the more commonly used approach due to its relative simplicity and ease of application (Golyandina & Zhigljavsky, 2020). The vector method is a modified version of the recurrent method.

Indonesia, as a country bisected by the equator, is particularly vulnerable to climate change and the associated shifts in rainfall patterns (Maulana & Yustiana, 2023). Indonesia's geography exhibits significant regional variability in rainfall, with rainfall patterns influenced by local factors such as topography, sea surface temperature, wind direction, and wind speed (Octavia et al., 2023). Parepare City, located in South Sulawesi Province, comprises four sub-districts: Soreang, Ujung, Bacukiki, and West Bacukiki. Parepare City is situated along a bay overlooking the Makassar Strait. It shares borders with Pinrang Regency to the north, Sidenreng Rappang Regency to the east, and Barru Regency to the south. Despite its coastal location, much of the city's terrain is characterized by hilly landscapes.

The city's diverse topography includes lowlands, coastal areas, and mountainous regions. Situated along the South Sulawesi coastline facing the Gulf of Bone, Parepare's proximity to the ocean results in a significant influx of water vapor from sea breezes, which increases the potential for higher rainfall. The city's mountainous terrain further influences rainfall patterns by forcing moist air masses to ascend leading to cooling, condensation, and subsequent precipitation. Regions with undulating topography often experience increased rainfall due to wind being forced to ascend or descend, depending on its direction. Parepare City is equipped with two rainfall observation stations: the Bukit Harapan rain gauge, located at the Agriculture, Marine and Fisheries Office, and the Bulu Dua rain gauge, situated at the BPP Bacucuki office.

Under extreme conditions, significant weather changes can result in severe consequences, including flooding, disease outbreaks, health issues, transportation disruptions, rising sea levels, and crop failures (Suhadi et al., 2023). In February 2023, three sub-districts in Parepare City—West Bacukiki, Soreang, and Ujung—experienced severe flooding and landslides. These natural disasters were triggered by the overflow of several rivers that pass through residential areas, following heavy rainfall accompanied by strong winds in the region. The disasters caused significant damage to hundreds of houses, public facilities, and places of worship. To mitigate the risks associated with extreme rainfall, it is crucial for the government to implement effective policy measures and adopt appropriate strategies to address this issue.

Previous research by Purnama (2022) applied the Singular Spectrum Analysis (SSA) method for rainfall forecasting in Gorontalo Province. The study found that using SSA with a window length of $L = 36$ resulted in a Mean Absolute Percentage Error (MAPE) of 0.029 (2.9%) for out-of-sample data, indicating a highly accurate forecast for rainfall in Gorontalo Province for the period of 2022-2023. The forecasting results suggest that rainfall levels remain relatively high, highlighting the need for government policies to mitigate potential negative impacts. Research conducted by Agustina et al. (2023) on rainfall forecasting in Medan using the Singular Spectrum Analysis (SSA) method achieved a MAPE accuracy value of 15.5%, with tracking signals remaining within the tolerance limits. Therefore, it can be concluded that the forecasting was performed effectively. Based on the literature review, no prior research has applied the Singular Spectrum Analysis (SSA) method for rainfall forecasting in Parepare City. Therefore, this study aims to forecast rainfall in Parepare City using the SSA method.

B. RESEARCH METHOD

1. Analysis of Time Series

Time series analysis was first introduced in 1970 by George E. P. Box and Gwilym M. Jenkins in their book *Time Series Analysis: Forecasting and Control*. Since then, the field has undergone significant development. Time series analysis is a

statistical procedure used to predict the probabilistic structure of future conditions to inform decision-making (Aswi & Sukarna, 2006). These models forecast future conditions by analyzing historical data and extrapolating patterns into the future. A crucial step in selecting an appropriate forecasting model is identifying the type of data pattern (Hidayat et al., 2020). Data patterns can be categorized into four types (Manuaba et al., 2022):

- Horizontal is data that is stationary with respect to its mean value or fluctuates constantly around the mean value.
- Seasonal is a data series that is affected by seasonal factors such as annual quarters, or days of the week.
- Trend is a pattern of data in which there is a long-term increase or decrease.
- Cyclical is a data pattern influenced by long-term fluctuations in economic values that have a relationship with the cycle of a business.

2. Singular Spectrum Analysis

Singular Spectrum Analysis (SSA) is a time series analysis method that integrates classical time series techniques, multivariate statistics, geometry, dynamical systems, and signal processing (Purnama, 2022). SSA is versatile and can address various time series problems, including forecasting, identifying simultaneous trends across multiple series, smoothing trend curves, extracting seasonal components, identifying cycles, uncovering short-term structures, and detecting change points (Marjuni, 2022). SSA decomposes time series data into interpretable components such as trends, seasonality, cycles, and noise (Sergio et al., 2023). The method involves two main stages: decomposition and reconstruction.

a. Decomposition

The decomposition process consists of two stages: Embedding and Singular Value Decomposition (SVD). An essential parameter in the decomposition stage is the Window Length (L), which plays a critical role in determining the effectiveness of the analysis (Darmawan, 2016).

Embedding

In this stage, the time series data is converted into a trajectory matrix by transforming the one-dimensional data (vector) into multidimensional data (matrix) (Utami et al., 2021). According Asrof et al. (2017), before proceeding step by step in SSA, it is crucial to determine the value of L . Given a time series dataset of length N with no missing data, expressed as $X = x_1, x_2, \dots, x_N$. X is transformed into an $L \times K$ matrix, where $K = N - L + 1$. L represents the window length, and it must satisfy the condition $2 < L < \frac{N}{2}$. The matrix form is expressed by the following Equation (1).

$$\mathbf{X} = [X_1, X_2, \dots, X_K] = \begin{bmatrix} x_1 & x_2 & \cdots & x_K \\ x_2 & x_3 & \cdots & x_{K+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_L & x_{L+1} & \cdots & x_N \end{bmatrix} \quad (1)$$

The matrix $\mathbf{X}$ is also known as the Hankel matrix, where all anti-diagonal elements have the same value. As a result, the output at this stage is a Hankel matrix of size $L \times K$ (Golyandina & Zhigljavsky, 2020).

Singular Value Decomposition

Singular Value Decomposition (SVD) aims to achieve component separation in the decomposition of time series data. It begins by determining the eigenvalues ($\lambda_1, \lambda_2, \dots, \lambda_L$) of the matrix $\mathbf{S}$ where (see Equation (2)),

$$\mathbf{S} = \mathbf{X} \mathbf{X}^T \quad (2)$$

with order according to $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_L \geq 0$ and $U_1, U_2, \dots, U_L$ is the eigenvector of each eigenvalue.

The rank of matrix $\mathbf{X}$ can be symbolized $d = \max i, \lambda_i \geq 0$ and $V_i = \frac{X^T U_i}{\sqrt{\lambda_i}}$ for $i = 1, \dots, d$. In this notation, the SVD of the trajectory matrix in Equation (3) (Golyandina & Zhigljavsky, 2020).

$$\begin{aligned} \mathbf{X} &= X_1 + X_2 + \dots + X_d \\ \mathbf{X} &= U_1 \sqrt{\lambda_1} V_1^T + U_2 \sqrt{\lambda_2} V_2^T + \dots + U_d \sqrt{\lambda_d} V_d^T \\ \mathbf{X} &= \sum_{i=1}^d U_i \sqrt{\lambda_i} V_i^T \end{aligned} \quad (3)$$

The basic concept at this stage is to obtain a matrix row from matrix S , where each matrix in the row contains an eigenvector (U_i), singular value ($\sqrt{\lambda_i}$) and principal component (V_i^T) that describes the characteristics of each matrix in the row.

b. Reconstruction

The reconstruction stage has two steps, namely, grouping and diagonal averaging. At this stage the parameters used are grouping effect (r).

Grouping

In this step, the grouping of the $L \times K$ trajectory matrix decomposition results will be carried out with the aim of separating the SVD additive components into several subgroups, namely trend, seasonality, and noise (Darmawan, 2016). Matrix X_i will be partitioned into m disjoint subsets denoted by $I = I_1 + \dots + I_m$, such that it yields the following Equation (4).

$$X_{Ix} = X_{I1} + X_{I2} + \dots + X_{Im} \quad (4)$$

The procedure of selecting sets $I_1, \dots, I_m$ is called eigentriples grouping. If $m = d$ and $I_j = j$, $j = 1, \dots, d$, then the corresponding grouping is called elementary (Golyandina & Zhigljavsky, 2020).

Diagonal Averaging

Following the grouping stage, the next step involves transforming the grouping results into a new series of length N . The objective of this step is to obtain the singular values of the separated components, which will subsequently be used for forecasting. The outcome of this stage is the F matrix as in Equation (5) (Darmawan, 2016).

$$F = \begin{bmatrix} f_{11} & f_{21} & \cdots & f_K \\ f_{12} & f_{22} & \cdots & f_{K+1} \\ \vdots & \vdots & \ddots & \vdots \\ f_L & f_{L+1} & \cdots & f_N \end{bmatrix} \quad (5)$$

Diagonal Average is formulated as in Equation (6).

$$g_k = \begin{cases} \frac{1}{k} \sum_{m=1}^k f_{m,k-m+1}^* & \text{for } 1 \leq k < L^* \\ \frac{1}{L^*-1} \sum_{m=1}^{L^*-1} f_{m,k-m+1}^* & \text{for } L^* \leq k \leq K^* + 1 \\ \frac{1}{N-k+1} \sum_{m=k-K^*+1}^{N-K^*+1} f_{m,k-m+1}^* & \text{for } K^* + 1 \leq k \leq N \end{cases} \quad (6)$$

where $L^* = \min(L, K)$ and $K^* = \max(L, K)$. From the above equation if collapsed into the resultant matrix X_{im} it will form a new series $\tilde{Y}^{(K)} = \tilde{y}_1^{(K)}, \tilde{y}_2^{(K)}, \dots, \tilde{y}_N^{(K)}$. Therefore, the result will be decomposed into the sum of m sequences as in Equation (7).

$$y_n = \sum_{k=1}^m \tilde{y}_n^{(k)} \quad (7)$$

3. Singular Spectrum Analysis (SSA) Forecasting

In Singular Spectrum Analysis (SSA), there are two primary forecasting methods: the R-forecasting method and the V-forecasting method. The recurrent method (R-forecasting) is the most commonly used approach due to its relative simplicity. The vector method (V-forecasting) is a modification of the recurrent method. In SSA forecasting, the model is constructed using a Linear Recurrent Formula (LRF) of polynomial form, as follows in Equation (8) (Golyandina & Zhigljavsky, 2020).

$$x_{i+d} = \sum_{k=1}^d r_k x_{i+d-k} \quad \text{for } 1 \leq i \leq N-d \quad (8)$$

The difference between R-forecasting and V-forecasting is that R-forecasting performs continuation directly (with the help of LRF), while V-forecasting deals with L-continuation. This causes the approximate continuation to usually give different results (Golyandina & Zhigljavsky, 2020).

In this research, the Recurrent Forecasting method will be used to perform forecasting. Recurrent forecasting is the most commonly used SSA forecasting technique. The main property of SSA decomposition is that the original series can fulfill the Linear Recurrent Formula (LRF) equation. The LRF equation is shown in Equation (9),

$$f_n = a_1 f_{n-1} + \dots + a_d f_{n-d} \quad (9)$$

where a_i is the LRF coefficient with $i = 1, 2, \dots, d$.

If the original series f_n satisfies Equation (10) then the series f_n can be expressed as the sum of exponential, polynomial, and harmonic components over a short period of time. This enables future forecasting using the same LRF values, as the continuation effect remains valid for short-term predictions.

The recurrent forecasting algorithm is as follows (Golyandina & Zhigljavsky, 2020):

1. Time series $Y_{N+M} = (y_1, \dots, y_{N+M})$ defined by

$$y_i = \begin{cases} \bar{x}_i & \text{untuk } i = 1, \dots, N \end{cases} \quad (10)$$

2. The numbers $y_{N+1}, \dots, y_{N+M}$ will create the M terms of recurrent forecasting.

Recurrent forecasting is done with Linear Recurrence Relation (LRR) using the coefficient of $\{a_i, j = 1, \dots, L-1\}$.

The following is the formula the linear operator $P_{Rec} : R^L R^L$, presented in Equation (11).

$$P_{Rec} = \begin{pmatrix} \bar{Z} \\ R^T \bar{Z} \end{pmatrix} \quad (11)$$

Where, $\bar{Z}$ consists of the last L-1 of Z, then

$$Y_i = \begin{cases} \bar{X}_i & \text{untuk } i = 1, \dots, K \\ P_{Rec} Y_{i-1} & \text{untuk } i = K+1, \dots, K+M \end{cases} \quad (12)$$

matrix $Y = [Y_1 : \dots : Y_{K+M}]$ is the trajectory matrix of Y_{N+M} . Therefore, Equation (12) is considered as the vector form of Equation (10).

4. Root Mean Square Error (RMSE)

The Root Mean Square Error (RMSE) is used to evaluate the accuracy of the forecasting model. RMSE quantifies how well the model predicts or estimates the true values of the data. It measures the average of the squared differences between the predicted and actual values. The RMSE is formulated as follows in Equation 13 (Aditya Pratama et al., 2024).

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (Y_t - \hat{Y}_t)^2}{n}} \quad (13)$$

Where,

Y_t = actual value in the period t

$\hat{Y}_t$ = forecasting value in the period t

n = number of data

5. Rainfall

Rain occurs when water vapor in the atmosphere condenses to form clouds. When the water droplets within the clouds grow large enough, gravity causes them to fall to the earth's surface as rain. Rainfall is measured as the height of rainwater collected in a rain gauge placed on a flat, non-absorbent surface that prevents seepage, runoff, and evaporation. The recorded water level indicates the intensity of the rainfall, with higher measurements corresponding to heavier rainfall in the area being monitored (Badri et al., 2023).

According to Tjasyono, Indonesia is generally divided into 3 main climate patterns by looking at rainfall patterns during the year, as follows (Kusumawardhani et al., 2015):

1. Regions with monsoonal rainfall patterns are characterized by unimodal precipitation, with a single peak during the rainy season. June, July, and August are typically dry months, while December, January, and February experience heavy rainfall. The remaining six months are transitional periods, with three months marking the shift from the dry season to the rainy season, and three months marking the transition from the rainy season to the dry season. Areas predominantly influenced by this monsoonal pattern include southern Sumatra, Central and South Kalimantan, Java, Bali, Nusa Tenggara, and parts of Papua.
2. Regions with an Equatorial rainfall pattern are characterized by bimodal precipitation, with two peak rainy seasons typically occurring around March and October, coinciding with the equinoxes. This pattern is observed in areas such as the central and northern parts of Sumatra, as well as the northern regions of Kalimantan.
3. Regions with a Localized rainfall pattern are characterized by a unimodal precipitation pattern (a single peak), but this pattern is opposite to that of monsoonal rainfall. This pattern is found in areas such as Maluku, Sulawesi, and parts of Papua.

Rainfall is typically measured in millimeters or inches, but in Indonesia, the unit of measurement used is millimeters (mm). A rainfall measurement of 1 millimeter indicates that in an area of one square meter on a flat surface, 1 liter of water has been collected (Ajr & Dwirani, 2019). The limit values used to determine rain intensity according to BMKG are as follows:

- 0-100 mm = Low
- 101-300 mm = Medium
- 301-400 mm = High
- 401-500 mm = Very High

The chart in Figure 1 below presents an overview of data analysis techniques:

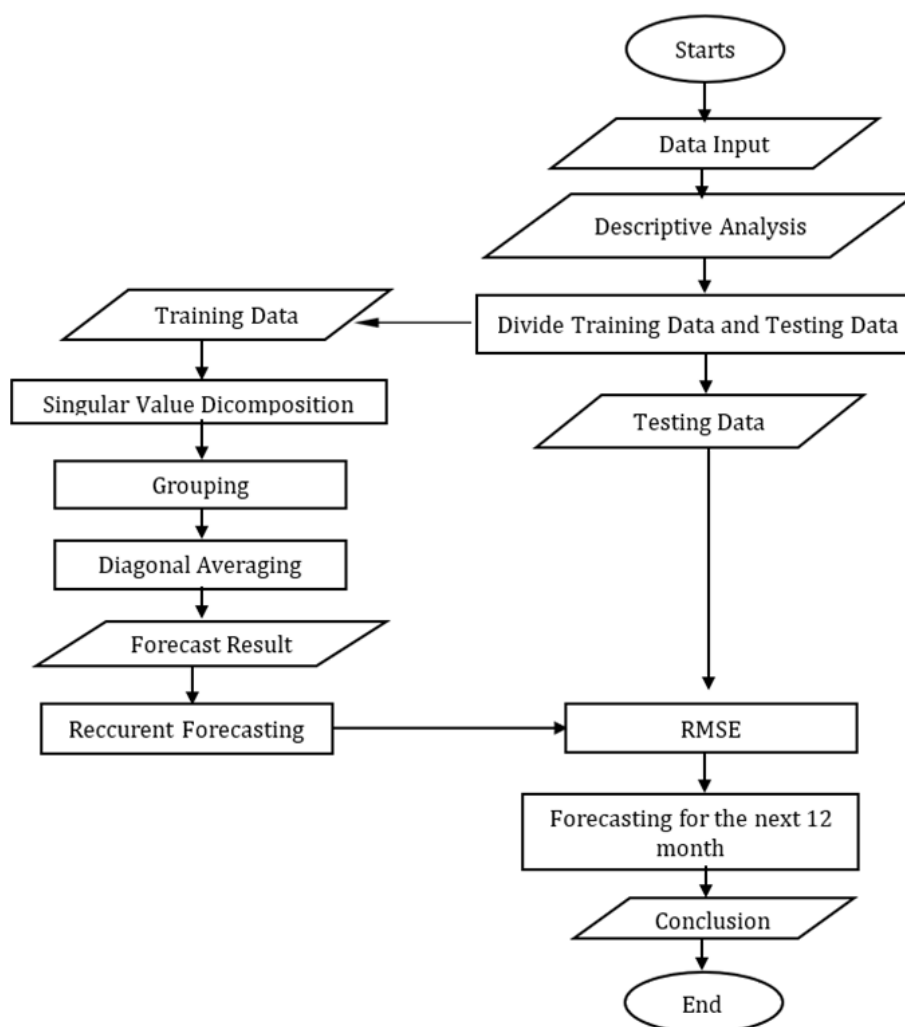


Figure 1. Data analysis techniques

C. RESULT AND DISCUSSION

1. Descriptive Analysis of Rainfall in Parepare City

The following is Figure 2, which presents rainfall data for Parepare City from January 2019 to December 2023.

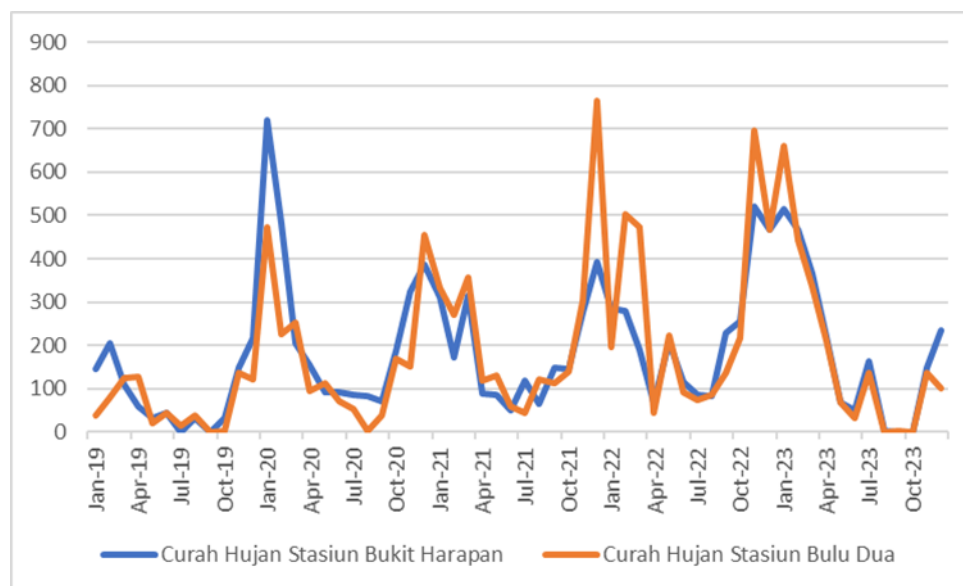


Figure 2. Rainfall in Parepare City from January 2019 to December 2023

Figure 1 illustrates that rainfall in Parepare City fluctuated over time. The highest rainfall intensity between January 2019 and December 2023 occurred during the periods of January to February 2020, November 2020 to March 2021, November 2021 to March 2022, and November 2022 to March 2023. Conversely, the lowest rainfall was recorded during May to October 2019, May to September 2020 and 2021, July to August 2022, and May to October 2023. A descriptive analysis of rainfall in Parepare City is presented in Table 1.

Table 1. Descriptive Analysis of Rainfall Data

Station	Minimum	Maximum	Mean	Median	Standard Deviation
Bukit Harapan	0.0	721.0	184.7	148.0	1.558.787
Bulu Dua	0.00	766.00	182.25	123.50	1.836.382

In Table 1, it can be seen that the average rainfall in Parepare City at Bukit Harapan Station is 184.7, and the distribution is 155.8787. The lowest amount of rainfall is 0 mm and the highest amount of rainfall is 721 mm. While the average rainfall in Parepare City at Bulu Dua station is 182.25 and the distribution is 183.6382. The lowest amount of rainfall is 0 mm and the highest amount of rainfall is 766 mm.

2. Singular Spectrum Analysis of Rainfall in Parepare City

Before the analysis stage using the Singular Spectrum Analysis (SSA) method, the data is divided into two subsets: training data (80%) and testing data (20%). Out of the 60 time series data points, 48 are designated as training data for model development, while the remaining 12 serve as testing data to evaluate the accuracy of the constructed forecasting model. The SSA method involves two key stages before forecasting: the decomposition stage and the reconstruction stage.

a. Decomposition

Embedding

In the embedding stage, the time series data is converted into a trajectory matrix or Hankel matrix. Determination of the L value is done by trial and error process using the smallest RMSE value. The data used in this study are rainfall data of Bukit Harapan station and Bulu Dua station, each with 48 data. So that the L value that meets is $3 < L < 24$. At Bukit Harapan station and Bulu Dua station, the best L values are 23 and 16, respectively. At Bukit Harapan station and Bulu Dua

station, the best L values are 23 and 16, respectively. Based on the L value, the dimensions for Bukit Harapan station are $K = N - L + 1 = 26$ and Bulu Dua station are $K = N - L + 1 = 33$. The X trajectory matrix of the two stations can be organized as equations X_1 and X_2 .

Trajectory Matrix X_1 for Bukit Harapan Station:

$$X_1 = (x_{i,j})_{23 \times 26} = \begin{bmatrix} 323 & 188 & \cdots & 144 \\ 386 & 323 & \cdots & 204 \\ \vdots & \vdots & \ddots & \vdots \\ 467 & 522 & \cdots & 172 \end{bmatrix}$$

Trajectory matrix X_2 for Bulu Dua Station:

$$X_2 = (x_{i,j})_{16 \times 33} = \begin{bmatrix} 96 & 251 & \cdots & 37 \\ 112 & 96 & \cdots & 80 \\ \vdots & \vdots & \ddots & \vdots \\ 468 & 696 & \cdots & 111 \end{bmatrix}$$

Singular Value Decomposition (SVD)

During the Singular Value Decomposition (SVD) stage, an S matrix is first constructed to determine the eigenvalues. This is achieved by multiplying the trajectory matrix by the transpose of itself, represented as ($S = X X^T$). Consequently, the S matrices for the Bukit Harapan and Bulu Dua stations are obtained in equations S_1 and S_2 .

Trajectory Matrix S_1 for Bukit Harapan Station:

$$S_1 = (x_{i,j})_{23 \times 23} = \begin{bmatrix} 1421029 & 1202041 & \cdots & 1218284 \\ 1202041 & 1499518 & \cdots & 1076443 \\ \vdots & \vdots & \ddots & \vdots \\ 1218284 & 1076443 & \cdots & 1668062 \end{bmatrix}$$

Trajectory matrix S_2 for Bulu Dua Station:

$$S_2 = (x_{i,j})_{16 \times 16} = \begin{bmatrix} 1080557 & 859494 & \cdots & 890182 \\ 859494 & 1098232 & \cdots & 1221034 \\ \vdots & \vdots & \ddots & \vdots \\ 890182 & 1221034 & \cdots & 2722547 \end{bmatrix}$$

After obtaining the S matrix, the next step in the Singular Value Decomposition (SVD) stage is to determine the eigentriples, which consist of singular values, eigenvectors, and principal components. The eigentriples are presented in Tables 2 to 4.

Table 2. Descriptive Analysis of Rainfall Data

Bukit Harapan Station			Bulu Dua Station		
i	Eigenvalues	Singular Values	i	Eigenvalues	Singular Values
1	20957485.644	4577.93465	1	19512007.22	4417.2398
2	3977750.448	1994.42986	2	4005586.68	2001.3962
3	3622305.355	1903.23550	3	3814429.44	1953.0564
4	706722.719	840.66802	4	1120872.94	1058.7129
5	685116.846	827.71785	5	923323.37	960.8972
...	...	...	...	...	...
23	5233.496	72.34291	16	28665.34	169.3084

Based on Table 2, the singular value with the largest value has the most significant influence on the characteristics of the time series data. While the singular value with the smallest value has the least influence on data characteristics.

Table 3. Eigenvector Value of Bukit Harapan Station and Bulu Dua Station

Bukit Harapan Station				Bulu Dua Station			
i	U_1	...	U_{23}	i	U_1	...	U_{16}
1	-0.1926803	...	0.115053012	...	-0.1875702	...	0.08836166
2	-0.2007515	...	-0.106710483	...	-0.1949765	...	-0.22148200
3	-0.1947815	...	0.151161878	...	-0.2059513	...	-0.51157283
4	-0.1916783	...	-0.307663437	...	-0.2411264	...	-0.43610172
5	-0.1891960	...	0.260618225	...	-0.2317182	...	-0.21429035
...	...	...	...	...	...	...	...
23	-0.2446074	...	-0.332860169	16	-0.3090067	...	0.30661264

Table 3 indicates that the Bukit Harapan eigenvector comprises 23 components, with values extending from -0.1926803 at U_1 to 0.115053012, and reaching -0.332860169 at U_{23} . In contrast, the Bulu Dua eigenvector contains 16 components, beginning at -0.1875702 at U_1 and ending at -0.30661264 at U_{16} . The component values for each eigenvector are visualized in Figures 3 and 4. The singular value and eigenvector obtained to calculate the principal component value using the formula $V_i = \frac{X^T U_i}{\sqrt{\lambda_i}}$. From the results of these calculations, the principal component components are obtained in Table 4.

Table 4. Eigenvector Value of Bukit Harapan Station and Bulu Dua Station

Bukit Harapan Station				Bulu Dua Station			
i	V_1	...	V_{23}	i	V_1	...	V_{16}
1	-0.1685723	...	-0.0753041969	1	-0.1062831	...	0.230927683
2	-0.4198156	...	0.2855971233	2	-0.2381608	...	-0.001928191
3	-0.4504203	...	-0.2085930058	3	-0.2411053	...	-0.003710098
4	-1.0092751	...	0.0009519360	4	-0.4239727	...	0.016469032
5	-1.0679353	...	-0.0087632984	5	-0.4341973	...	0.014280754
...	...	...	...	...	...	...	...
23	-0.293197	...	-0.2275390592	16	-0.2632067	...	-0.206137883

The obtained eigentriple values represent the entire decomposition process. Each component's value is utilized in the grouping stage.

b. Reconstruction

Grouping

In this step, the eigentriples are grouped into several components: trend, seasonality, and noise. The grouping effect determines the number of eigentriples used to identify the trend and seasonal components (r), which is based on the number of eigentriples that do not display noise in the singular value plot. The singular value plots for the Bukit Harapan and Bulu Dua stations are shown in Figures 3 and 4.

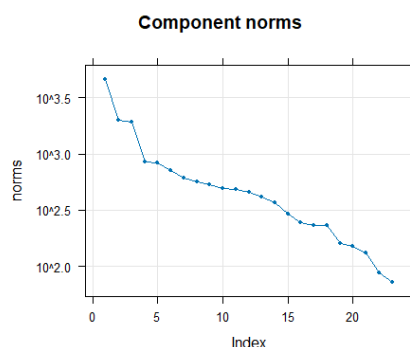


Figure 3. Singular Value Plot of Bukit Harapan Station

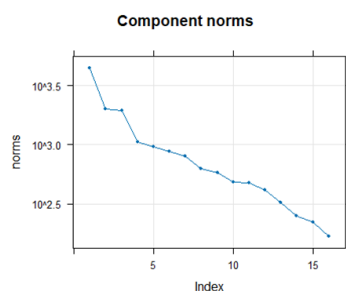


Figure 4. Singular Value Plot of Bulu Dua Station

It can be observed that the singular values begin to decline gradually from eigentriple 9 onward. Therefore, eigentriples 9 through 23 are classified as the noise component, with the value of r set to 8. It can be observed that the singular values begin to decrease from eigentriple 11 gradually. As a result, eigentriples 11 through 16 are classified as the noise component, with the value of r set to 10. The reconstructed sequence plots are shown in Figure 5 and Figure 6.

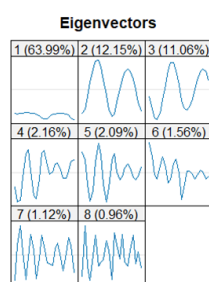


Figure 5. Eigenvector Graph of Bukit Harapan Station

Figure 5 shows that the series reconstructed by eigentriple 1 contains a slow-varying component, indicating that eigentriple 1 is the trend component. Grouping the eigentriples into seasonal components is based on the similarity of consecutive eigentriple patterns. In the figure, two pairs of eigentriples exhibit similar patterns: eigentriples 2 and 3, as well as eigentriples 4 and 5 (see Table 5).

Table 5. Eigenvector Value of Bukit Harapan Station and Bulu Dua Station

Component	Eigentriple
Trend	1
Seasonal 1	2 and 3
Seasonal 2	4 and 5
Noise	6, 7, and 8

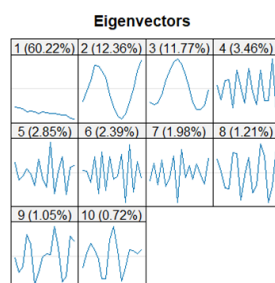


Figure 6. Eigenvector Graph of Bulu Dua Station

Figure 6 shows that the series reconstructed by eigentriple 1 contains a slow-varying component, indicating that eigentriple 1 is the trend component. Additionally, the grouping of eigentriples into seasonal components is based on the similarity of consecutive eigentriple patterns. In the figure, one pair of eigentriples, eigentriples 2 and 3, exhibits similar patterns (see Table 6).

Table 6. Eigentriple Couple Bulu Dua Station

Component	Eigentriple
Trend	1
Seasonal 1	2 and 3
Noise	4, 5, 6, 7, 8, 9, and 10

Based on the groups that have been formed, the weak separability between them will be evaluated to ensure there is no strong correlation. This is done using the w-correlation matrix (separability plot) (Lubis et al., 2017). Here's the separation plot of Bukit Harapan and Bulu Dua stations.

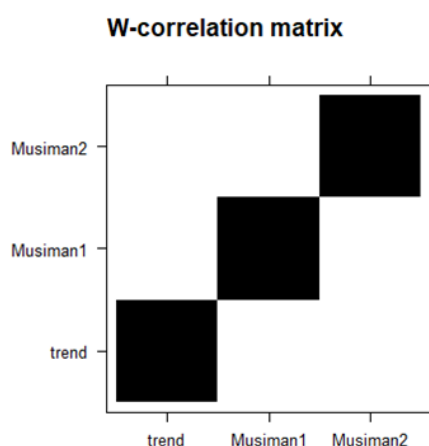


Figure 7. Separation Plot of Bukit Harapan Station

Figure 7 presents the separability plot for the trend component (eigentriple 1), seasonal component 1 (eigentriple pairs 2 and 3), and seasonal component 2 (eigentriple pairs 4 and 5). The plot indicates that these components are uncorrelated, as evidenced by the lack of colored areas that would suggest inter-group correlations. Figure 8 displays the separability plot for the trend component (eigentriple 1) and the first seasonal component (eigentriple pairs 2 and 3). The absence of any color gradients in the plot indicates no detectable correlation between these groups, confirming their clear separation.

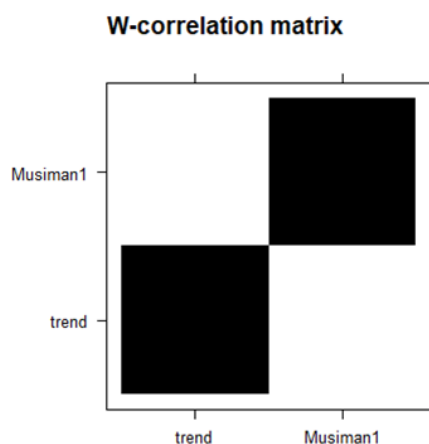


Figure 8. Separation Plot of Bulu Dua Station

Diagonal Averaging

Following the grouping stage, the next step is diagonal averaging, which transforms the grouped results into a time series of length 48. Each component is reconstructed using the corresponding eigentriple. In this study, for the Bukit Harapan station, the trend component is reconstructed using eigentriple 1. In contrast, the seasonal component is divided into two sub-components: the first season is reconstructed using eigentriples 2 and 3, and the second season using eigentriples 4 and 5. For the Bulu Dua station, the trend component is also reconstructed using eigentriple 1, while the first seasonal component is reconstructed using eigentriples 2 and 3. The results of diagonal averaging are presented in Tables 7 and 8.

Table 7. Diagonal Averaging Bukit Harapan Station

Time to-	Reconstruction			Diagonal Averaging
	Trend	Seasonal 1	Seasonal 2	
1	148.6939	139.42	37.837	325.9509
2	158.1262	148.2542	-10.6407	295.7396
3	161.1932	102.7354	-94.1463	169.7823
4	161.6469	17.5243	-128.3024	50.8688
5	162.8247	-78.2213	-62.6417	21.1968
...	...	...	...	...
45	211.0999	-1.921	-12.204	196.9749
46	220.1739	53.1452	1.01662	274.3357
47	234.7377	98.2888	16.9699	349.9965
48	246.4224	113.3126	17.9966	377.7316

Table 8. Diagonal Averaging Bulu Dua Station

Time to-	Reconstruction		Diagonal Averaging
	Trend	Seasonal 1	
1	88.06	64.9719	153.0319
2	90.4716	80.6565	171.1281
3	92.6504	71.4289	164.0794
4	96.8444	37.7316	134.576
5	97.2955	-15.7158	81.5797
...	...	...	...
45	282.3135	-88.2736	194.0399
46	301.3008	29.483	330.7839
47	339.023	175.185	514.2081
48	359.2658	245.7016	604.9673

The next step involves forecasting using a series of models derived from the previously processed time series data, utilizing the Recurrent Forecasting method within Singular Spectrum Analysis (SSA).

3. Recurrent Forecasting

Following the diagonal averaging process, the next step is to forecast the testing data using the recurrent forecasting method based on the previously constructed model. The overall forecast for the testing data is obtained by summing the forecasted values of both the trend and seasonal components. The forecasting results for the testing data at Bukit Harapan and Bulu Dua stations are presented in Tables 9 and 10.

Table 9. R-Forecasting Results on Bukit Harapan Station Testing Data

Month	Forecasting			Forecast Result	Actual Data	Percentage Error
	Trend	Seasonal 1	Seasonal 2			
January	236.784	81.7913	-1.7745	316.8	514	38887.682
February	239.143	35.5163	-7.9779	266.682	466	39727.825
March	241.679	-19.2426	-4.9227	217.514	366	22048.211
April	244.372	-66.5978	2.6082	180.383	219	1491.2881
May	247.229	-93.3007	6.0612	159.989	67	8646.9913
June	250.218	-92.4703	2.3978	160.146	50	12132.075

Month	Forecasting			Forecast Result	Actual Data	Percentage Error
	Trend	Seasonal 1	Seasonal 2			
July	253.348	-65.1968	-3.353	184.799	162	519.77616
August	256.634	-19.9427	-4.3215	232.37	2	53070.337
September	260.125	30.1029	0.6436	290.872	0	84606.404
October	263.728	70.5956	5.9028	340.227	0	115754.28
November	267.429	90.2331	5.354	363.016	147	46662.739
December	271.304	83.7782	-0.7323	354.35	234	14484.002
Total Percentage						438031.607
RMSE Value						191.0566

Table 10. R-Forecasting Results on Bulu Dua Station Testing Data

Month	Forecasting		Forecast Result	Actual Data	Percentage Error
	Trend	Seasonal 1			
January	342.448	200.0662	542.5141	662	14276.88
February	351.868	118.2493	470.1173	443	735.34796
March	361.504	0.8932	362.3975	337	645.03301
April	371.122	-117.146	253.976	211	1846.9366
May	381.438	-199.7866	181.6518	69	12690.428
June	392.348	-220.9149	171.4331	32	19441.589
July	404.111	-173.1964	230.9147	135	9199.6297
August	417.200	-70.0882	347.1122	0	120486.88
September	431.147	56.5226	487.6695	1	236847.2
October	445.730	165.8732	611.6032	0	374058.47
November	461.176	224.4385	685.614	135	303175.78
December	476.916	210.069	686.9853	101	343378.77
Total Percentage					119731.912
RMSE Value					346.023

4. Forecasting Accuracy with RMSE

The accuracy of the forecasting results for the testing data is evaluated using the RMSE (Root Mean Squared Error), which represents the average percentage error over several periods. Based on the analysis results shown in Tables 9 and 10, the RMSE for Bukit Harapan station is 191.0566, while for Bulu Dua station, it is 346.023. These values indicate that the recurrent forecasting method within SSA is effective in forecasting rainfall in Parepare City.

5. Forecasting

Rainfall forecasting in Parepare City, specifically at Bukit Harapan and Bulu Dua stations, using the SSA method with window lengths of 23 and 16, respectively, demonstrates a high degree of accuracy and can be effectively applied to future rainfall predictions. The dataset used in this study comprises 60 data points, consisting of both training and testing data. The rainfall forecast for Parepare City over the next 12 months is presented in Table 11 below.

Table 11. Rainfall Forecasting Results in Parepare City

Month	Bukit Harapan Station	Bulu Dua Station
January 2024	344.645	483.6746
February 2024	345.689	451.0712
March 2024	303.672	369.2584
April 2024	235.013	260.5752
May 2024	164.727	153.2882
June 2024	114.412	74.7664
July 2024	100.403	43.9931
August 2024	126.848	67.6511
September 2024	191.080	137.6940
October 2024	280.558	234.8837
November 2024	376.529	333.2218

Month	Bukit Harapan Station	Bulu Dua Station
December 2024	458.787	406.7334

Applying the previously validated window length and grouping parameters, Table 11 presents the monthly rainfall forecasts for Parepare City for the year 2024. Results indicate persistently high rainfall levels, with the maximum predicted at Bukit Harapan station in December and at Bulu Dua station in January. Conversely, both stations are expected to record their lowest rainfall in July 2024. The visualization of rainfall forecasting results for Parepare City is shown in Figures 9 and 10.

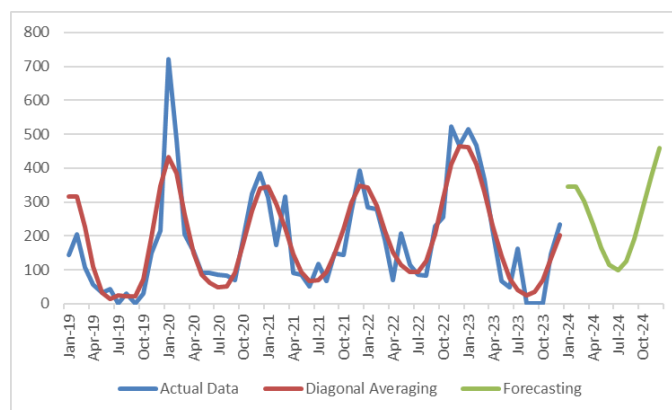


Figure 9. Rainfall Forecasting Result Chart of Bukit Harapan Station

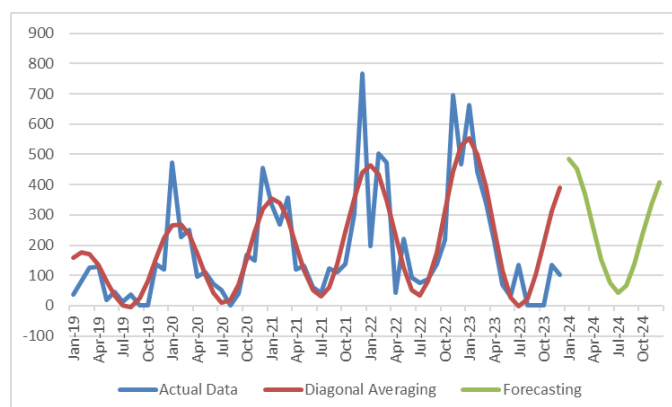


Figure 10. Rainfall Forecasting Result Chart of Bulu Dua Station

D. CONCLUSION AND SUGGESTION

The forecast results using the Singular Spectrum Analysis (SSA) method indicate that monthly rainfall in Parepare City remains relatively high. The highest rainfall is projected for December 2024 at Bukit Harapan station, while at Bulu Dua station, it is expected in January 2024. Conversely, the lowest rainfall is forecasted for both Bukit Harapan and Bulu Dua stations in July 2024. These projections can serve as a valuable reference for the government and local communities in preparing for and mitigating the potential negative impacts of high rainfall in Parepare City. Rainfall forecasting for Parepare City using the Singular Spectrum Analysis (SSA) method yielded an RMSE accuracy of 191.0566 for Bukit Harapan station and 346.023 for Bulu Dua station. Therefore, it can be concluded that this method performs well in forecasting rainfall in Parepare City and can be reliably used for future rainfall predictions.

ACKNOWLEDGEMENT

The title for the thank you to the institution or the person who has contributed during the research and references is not numbered. The Acknowledgments section is optional. Research sources can be included in this section.

DECLARATIONS

AUTHOR CONTRIBUTION

All authors contributed to this manuscript, from exploring ideas to writing this article.

FUNDING STATEMENT

This research received no specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article.

REFERENCES

- Aditya Pratama, M., Munawaroh, M., Joko Pranoto, W., Studi Teknik Informatika, P., Sains dan Teknologi, F., & Muhammadiyah Kalimantan Timur, U. (2024). Perbandingan Performa Algoritma Linear Regresi dan Random Forest untuk Prediksi Harga Bawang Merah di Kota Samarinda. *Jurnal Ilmu Teknik*, 1(2), 172–182.
- Agustina, S., Fitri, F., Vionanda, D., & Salma, A. (2023). Rainfall Forecasting in Medan City Using Singular Spectrum Analysis (SSA). *UNP Journal of Statistics and Data Science*, 1(3), 149–156. <https://doi.org/10.24036/ujsds/vol1-iss3/52>
- Ajr, E. Q., & Dwirani, F. (2019). Menentukan Stasiun Hujan dan Curah Hujan dengan Metode Polygon Thiessen Daerah Kabupaten Lebak. *Jurnal Lingkungan dan Sumber Daya Alam (JURNALIS)*, 2(2), 139–146.
- Asrof, A., Ischak, R., & Darmawan, G. (2017). Peramalan Produksi Cabai Merah di Jawa Barat Menggunakan Metode Singular Spectrum Analysis (SSA). *STATISTIKA: Journal of Theoretical Statistics and Its Applications*, 17(2), 77–87. <https://doi.org/10.29313/jstat.v17i2.2839>
- Aswi & Sukarna. (2006). *Analisis Deret Waktu: Teori dan Aplikasi* (Edisi Pert). Andira Publisher.
- Badri, A. A. A., Nandarie, A. C. A., & Haryanto, Y. D. (2023). Optimasi Model ARIMA dalam Prakiraan Curah Hujan di Jambi. *Gographia : Jurnal Pendidikan dan Penelitian Geografi*, 4(1), 39–43. <https://doi.org/10.53682/gjppg.v4i1.5776>
- Darmawan, G. (2016). Identifikasi Pola Data Curah Hujan pada Proses Grouping dalam Metode Singular Spectrum Analysis. *Seminar Nasional Pendidikan Matematika 2016*, 1–9.
- Golyandina, N., & Zhigljavsky, A. (2020). *Singular Spectrum Analysis for Time Series* (Second). Springer Berlin, Heidelberg. <https://doi.org/https://doi.org/10.1007/978-3-662-62436-4>
- Hidayat, K. W., Wahyuningsih, S., & Nasution, Y. N. (2020). Pemodelan Jumlah Titik Panas di Provinsi Kalimantan Timur dengan Metode Singular Spectrum Analysis. *Jambura Journal of Probability and Statistics*, 1(2), 78–88. <https://doi.org/10.34312/jjps.v1i2.7287>
- Kusumawardhani, Ismi, D., & Gernowo, R. (2015). Analisis Perubahan Iklim berbagai Variabilitas Curah Hujan dan Emisi Gas Metana (CH₄) dengan Metode Grid Analysis and Display System (GrADS) di Kabupaten Semarang. *Youngster Physic Journal*, 4(1), 49–54.
- Lubis, D. A., Johra, M. B., & Darmawan, G. (2017). Peramalan Indeks Harga Konsumen dengan Metode Singular Spectral Analysis (SSA) dan Seasonal Autoregressive Integrated Moving Average (SARIMA). *Jurnal Matematika "MANTIK"*, 3(2), 74–82. <https://doi.org/10.15642/mantik.2017.3.2.74-82>
- Manuaba, I. D. N. A., Manuaba, I. B. G., & Sudarma, M. (2022). Komparasi Metode Peramalan Grey dan Grey-Markov untuk mengetahui Peramalan PNBPN di Universitas Udayana. *Majalah Ilmiah Teknologi Elektro*, 21(1), 83–88. <https://doi.org/10.24843/mite.2022.v21i01.p12>
- Marjuni, A. (2022). Peramalan Harga Saham Serentak Menggunakan Model Multivariate Singular Spectrum Analysis. *Jurnal Sistem Informasi Bisnis*, 12(1), 17–25. <https://doi.org/10.21456/vol12iss1pp17-25>
- Maulana, M. I., & Yustiana, F. (2023). Analisis Kuantitatif dan Variabilitas Curah Hujan dengan Klasifikasi Iklim Mohr di Kota Padang. *FTSP Series: Seminar Nasional dan Diseminasi Tugas Akhir 2023*, 157–162.

- Purnama, E. (2022). Aplikasi Metode Singular Spectrum Analysis (SSA) pada Peramalan Curah Hujan di Provinsi Gorontalo. *Jambura Journal of Probability and Statistics*, 3(2), 161–170.
- Sergio, A., Wahyuningsih, S., & Siringoringo, M. (2023). Peramalan Inflasi Kota Balikpapan Menggunakan Metode Singular Spectrum Analysis. *Jurnal EKSPONENSIAL*, 14(1), 21–30.
- Setiawan, D. A., Wahyuningsih, S., & Goejantoro, R. (2019). Peramalan Produksi Kelapa Sawit Menggunakan Winter's dan Pegel's Exponential Smoothing dengan Pemantauan Tracking Signal. *Jambura Journal of Mathematics*, 2(1), 1–14. <https://doi.org/10.34312/jjom.v2i1.2320>
- Suhadi, Mabruroh, F., Wiyanto, A., & Ikra. (2023). Analisis Fenomena Perubahan Iklim Terhadap Curah Hujan Ekstrim. *OPTIKA: Jurnal Pendidikan Fisika*, 7(1), 94–100.
- Utami, N. A. G., Sulandari, W., & Handajani, S. S. (2021). Peramalan Curah Hujan Bulanan Di Pos Hujan Jatisrono Dengan Metode Singular Spectrum Analysis (SSA). *Prosiding Seminar Nasional Aplikasi Sains & Teknologi (SNAST)*.
- Zulhijrah, Isnaini, M., Sulastri, S., Zalza, & Aswi Aswi. (2023). Dampak Covid-19 terhadap Tingkat Inflasi di Indonesia. *Jurnal Matematika dan Statistika serta Aplikasinya*, 11(2), 57–63.