JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI e–ISSN: 2614–8811
Published by Departement of Mathematics, Hasanuddin University, Indonesia
https://journal.unhas.ac.id/index.php/jmsk/index
Vol. 21, No. 3, May 2025, pp. 897-908
DOI: 10.20956/j.v21i3.43304

Implementation of Singular Spectrum Analysis Method for Prediction of Average Sunshine Duration

Shifa Amelia Rachman[1], Putroue Keumala Intan*[2], Wika Dianita Utami[3], Adhim Salamiani[4]

[1,2,3]Department Mathematics, Sunan Ampel State Islamic University, Surabaya
[(4)] Meteorology, Climatology and Geophysics Agency
***Email**: [1]shifaamelia2507@gmail.com , [2]putroue@uinsa.ac.id , [3]wikadianita@uinsby.ac.id , [4]adhim.salamiani@bmkg.go.id*
**Corresponding author*
Received: 23 February 2025, revised: 26 April 2025, accepted: 28 April 2025

Abstract

Solar irradiance is the process by which radiant energy from the sun reaches the earth. BMKG states that solar irradiance reaches *100%* when the sun shines for 8 hours a day. Less than 8 hours of solar irradiance a day can affect local and global climate systems. This research aims to analyze and predict of average sunshine duration in Pasuruan with the Singular Spectrum Analysis (SSA) method. Based on the SSA model for optimal solar irradiation with $L = 41$ and Grouping Effect $r = 8$, this study analyzes the prediction of average sunshine duration in Pasuruan which produces a Mean Absolute Percentage Error (MAPE) value of 19.53%. The results indicate that the predictions are effectively categorized for estimating the average solar irradiance. The highest average was in July at 60.1% and the lowest average was in November at 12.82%

Keywords: BMKG, Climate, Prediction, Solar Irradiance, SSA

1. INTRODUCTION AND PRELIMINARIES

Weather is an essential aspect of everyday life, influencing various human activities and decisions. Weather has many elements, one of which is solar irradiance [20]. Solar irradiance is the process by which radiant energy from the sun reaches the earth. Solar irradiance plays an important role in various natural processes, such asp hotosynthesis by plants, global warming, rainfall and climate formation [13]. Based on information from the Meteorology, Climatology, and Geophysics Agency (BMKG), solar irradiance is said to be 100% when the sun shines for 8 hours a day. Less than 8 hours of solar irradiance a day can affect local and global climate systems [15]. Disturbances to the local and global climate system can have serious impacts on both the natural environment and human life, one of which is changes in extreme weather patterns. In Pasuruan, the maximum temperature ranges from 33 to 35 degrees Celsius during the day. Another factor to be aware of is the exposure to ultraviolet (UV) radiation, which is highest between 12:00 PM and 3:00 PM WIB [8]. The impact of this was felt in 2023, when Pasuruan city experienced suboptimal crop yields due to hot weather. This is the impact of extreme weather patterns, namely a prolonged dry season. This greatly affects the agricultural sector, where Pasuruan is one of the largest agricultural producing cities in East Java [6]. In a city like Pasuruan, an accurate understanding of average solar irradiance is important in optimizing various agricultural and other activities. Therefore, forecasting the sunshine duration is something that needs to be considered so that the next step can be taken in minimizing the impact of losses caused by solar irradiance [3]. This research aims to determine the accuracy of predicting the future average duration of solar irradiance. Forecasting sunshine duration is an important solution to minimize the impact of losses caused by solar solar irradiance. By using forecasting methods, we can obtain more accurate predictions regarding the intensity and duration of sunlight in the future, enabling more effective mitigation measures, both in the agricultural sector and other activities that depend on weather conditions.

Forecasting is the science of estimating future outcomes based on historical data. This analysis uses observations from previously collected data called time series [12]. Forecasting is used in decision making to obtain a certain goal by predicting future conditions [23]. There are many methods that can be applied to forecast future events, but one of the technologies developed by modern scientists to forecast future events today is Singular Spectrum Analysis (SSA). SSA decomposes a dataset into a sum of a few independent components, which can be interpreted as trend, seasonality, and noise. [21]. The advantage of SSA is that it can be used on multiple time series as no stationarity assumptions or logarithmic transformations are needed.This method is suitable for solar irradiance data. This is because solar irradiance data is not always stationary data. Solar irradiance data tends to have seasonal patterns and long-term variances that can make it non-stationary [22].

Some previous studies that applied the SSA method were conducted by [19] to forecast Rice Production in Southeast Sulawesi Province in the research, a MAPE value of 0.46% was obtained with a window length (L) of 25. which shows very good forecasting accuracy. Then the next research was conducted by [7] to forecast the United States Natural Gas Consumption based on this research obtained a MAPE value of 1.62%, with relatively constant forecasting results every year. The next research was conducted by [16] for monthly rainfall forecasting in Bogor city with comparison of SARIMA and SSA methods. The accuracy value obtained by applying the SSA method is greater than the SARIMA method, namely MAPE of 22%. And there are many more studies that implement the Singular Spectrum Analysis (SSA) approach.

2. Methods

The research methodology used in this study is quantitative. The data employed in this study is secondary data, namely monthly average data on the sunshine duration in Pasuruan for the time span of January 2010 to December 2023. There are obtained from the official institution of the Pasuruan Meteorology, Climatology and Geophysics Agency (BMKG) totaling 168 data samples. In this study, the dataset is divided into two subsets: 80% (134 data points) for training and 20% (24 data points) for testing. In this research, Singular Spectrum Analysis (SSA) is applied to predict the average solar irradiation. Singular Spectrum Analysis or SSA is a forecasting technique that offers greater flexibility compared to other methods due to its nonparametric approach, therefore, it does not rely on assumption tests like independence and normality of residuals, making it effective for both stationary and nonstationary data. There are two main stages in this method, namely decomposition and reconstruction.

A. Decomposition

The steps of the decomposition process consist of Embedding and Singular Value Decomposition (SVD).

1. Embedding

The embedding process involves transforming the initial time series data into a trajectory matrix or can be interpreted by transforming one dimensional data into a multidimensional matrix structure. [1]. The trajectory matrix is denoted as X has a dimension of $L \times K$. With L as the window length, indicating the total number of rows in the matrix while K will be the number of matrix columns. There is no definitive approach for identifying the exact value of L, so to determine the value of L is determined through trial and error until the MAPE reaches its minimum value is obtained. There is a range in determining the value of L, namely $2 \leq N \leq \frac{N}{2}$ with the assumption that L must be greater but not exceed the value of $\frac{N}{2}$ where the value of K is given as follows $N - L + 1$. This trajectory matrix can be identified as a Hankel matrix, which is defined as a matrix that whose antidiagonal elements have the same value and can be expressed as follows [4]

$$X = (X_i)_{L\times K} = \begin{bmatrix} x_1 & x_2 & \cdots & x_K \\ x_2 & x_3 & \cdots & x_{K+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{L+1} & x_{L+2} & \cdots & x_{L+K} \end{bmatrix} \quad 2.1$$

2. Singular Value Decomposition

SVD (Singular Value Decomposition) is a process of decomposing time series data into eigentriple components, including of Eigenvector U_i, Singular Value $\sqrt{\lambda_i}$ and Principal Component V_i^T. The first step to get the eigentriple value is to define the matrix S. In [10]:

$$S = XX^T \quad 2.2$$

where X is the matrix obtained in the embedding process. From the above calculation, the eigenvalues are obtained in the order according $\lambda_1 \geq \cdots \geq \ \lambda_1 \geq 0$ and $U_1 \ldots U_L$ is the eigenvector of each eigenvalue. The rank of matrix X can be represented by $d = \max\{i, \lambda_1, \geq 0\}$ if it is denoted as follows: [10].

$$\boldsymbol{V_1} = \frac{\boldsymbol{X^T U_i}}{\sqrt{\boldsymbol{\lambda_1}}} \text{ for } \boldsymbol{i} = \boldsymbol{1}, \ldots, \boldsymbol{d}, \quad 2.3$$

The SVD decomposition of the trajectory matrix is given as follows:

$$X = X_1 + X_2 + \ldots + X_d$$

$$X = U_1\sqrt{\lambda_1}V_1^T + U_2\sqrt{\lambda_2}V_2^T + U_3\sqrt{\lambda_3}V_3^T$$

$$\boldsymbol{X} = \sum_{\boldsymbol{i=1}}^{\boldsymbol{d}} \boldsymbol{U_i}\sqrt{\boldsymbol{\lambda_i}}\boldsymbol{V_i^T} \quad 2.4$$

The basic concept at this stage is to obtain a matrix row from the matrix S, where each matrix in the row contains the eigenvector U_i, *singular value* $\sqrt{\lambda_1}$ and principal component V_1^T which describes the characteristics of each row [5].

B. Reconstruction

The reconstruction process consists of two steps that must be performed, that is Grouping and Diagonal Averaging

1. Grouping

In this step, eigentriple grouping is performed according to the characteristics of each component. Grouping is the first step of the reconstruction stage. The purpose of the grouping step is to separate the eigentriple components produced by the SVD process [14]. These components will be broken down into several subgroups including trend, seasonal, and noise groups. noise, grouping index sets $i = \{1, \dots, d\}$ into m mutually independent sets, $I_1, I_2 \dots I_m$ with $m = d$, where $d = max\{i, \lambda_i > 0\}$. Furthermore, X_i will be adjusted to group $I = \{i_{1,} i_2, \dots i_m\}$. Thus, $X_i = X_1 + X_2 + \cdots + X_d$ can be written as equation 2.5: [2]

$$\boldsymbol{X} = \boldsymbol{X_{I1}} + \boldsymbol{X_{I2}} + \dots + \boldsymbol{X_{Im}} \quad 2.5$$

2. Diagonal Averaging

As a result of this process, the grouping output is structured into a new sequence of length. At this step, the transformed of the X_I matrix grouping results reconstruction into a new sequence with a length of N will be carried out [11]. The purpose of this stage is to obtain the singular value of the separated components, which will subsequently be utilized for forecasting. The diagonal averaging step is performed by transforming each X_I matrix in the clustering stage into a new sequence with a length of N. Given Y a matrix of size $L \times K$ with elements y_{ij}, where $1 \le i \le L$ and $1 \le j \le K$ where $L \le K$ [9].

$$Y = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_K \\ y_{21} & y_{22} & \cdots & y_{K+1} \\ \vdots & \vdots & \ddots & \vdots \\ y_L & y_{L+1} & \cdots & y_N \end{bmatrix} \quad 2.6$$

The Y matrix through diagonal averaging will be Transformed into a series format $g_1, g_2, \dots, g_N$ with the following equation:

$$\boldsymbol{g_j} = \begin{Bmatrix} \frac{1}{j}\sum_{m=1}^{j} \boldsymbol{Y_{m,j-m+1}} & \boldsymbol{untuk\ 1 \le j \le L^*} \\ \frac{1}{L^*-1}\sum_{m=1}^{L^*-1} \boldsymbol{Y_{m,j-m+1}} & \boldsymbol{untuk\ L^* < j \le K^* + 1} \\ \frac{1}{N-j+1}\sum_{m=j-K^*+1}^{N-K^*-1} \boldsymbol{Y_{m,j-m+1}} & \boldsymbol{untuk\ K^* + 1 < j \le N} \end{Bmatrix} \quad 2.7$$

Where $L^* = (min\ L, K), K^* = (max\ L, K)\ and\ N = L + K - 1$ [9].

After obtaining the results of diagonal averaging, the next step is forecasting. Forecasting with the SSA method can be performed using two techniques, namely the recurrent method (R-Forecasting) and the vector method (V-Forecasting). RForecasting is a commonly used forecasting method because it is believed to be easier, this method performs continuation through the direct use of the Linear Recurrent Formula. The R-Forecasting method utilizes diagonal averaging for reconstruction, followed by processing with the Linear Recurrent Formula. The R-Forecasting method involves calculations based on the Linear Recurrent Formula, which is $a_1, \dots, a_d$ using the Eigenvector derived from the SVD process. Suppose U_i is the corresponding eigenvector $U_i = (u_1, u_2, \dots, u_{L-1}, u_L)^T$ and $U^{\bar{v}}$ is the first L - 1 coordinates of U_i defined $U^{\bar{v}} = (u_1, u_2, \dots, u_{L-1})^T$ while π_q corresponds to the last element of the eigenvector $U_i(\pi_q = u_L)$, the LRF coefficient can be calculated with the following formula: [17] $U^{\bar{v}} = (u_1, u_2, \dots, u_{L-1})^T$

$$\boldsymbol{R} = (\boldsymbol{a_{L-1}}, \dots, \boldsymbol{a_1}) = \frac{\boldsymbol{1}}{\boldsymbol{1-v^2}}\sum_{q=1}^{L-1} \boldsymbol{\pi_q U^{\bar{v}}} \quad 2.8$$

with $v^2 = \sum_{q=1}^{L-1} \pi_q^2$

In R-Forecasting, the time series data used is obtained through reconstruction via the diagonal averaging process. Forecasting results are obtained based on the following equation: [17]

$$\boldsymbol{g_i} = \begin{cases} \boldsymbol{\tilde{y}_\iota} & , \boldsymbol{i = 0, \dots, N} \\ \sum_{j=1}^{L-1} \boldsymbol{a_j g_{i-1}} & , \boldsymbol{i = N+1, \dots, N+M} \end{cases} \quad 2.9$$

Where $g_{N+1}, g_{N+2}, \dots, g_{N+M}$ are the outcomes of forecasting with the SSA method.

After obtaining the prediction results, the next step is model evaluation. in this study using the Mean Absolute Percentage Error. Mean Absolute Percentage Error (MAPE) method is most commonly used to measure the average forecasting error. MAPE calculates the mean percentage deviation between actual and predicted results. The smaller the prediction error rate, the better the forecast or prediction value. MAPE can be computed using the formula below [18]:

$$\mathbf{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{f_i - g_i}{f_i}\right| \times \mathbf{100}\% \tag{2.10}$$

Where N represents the total data, f_i denotes the observed data values, and g_i refers to the predicted values. The forecasting MAPE value is defined as follows:

Table 2.1 MAPE Value Terms

MAPE	Description
$< 10\%$	Highly Accurate
$10\% - 20\%$	Accurate
$> 50\%$	Inaccurate

The process of implementing the Singular Spectrum Analysis (SSA) method for average sunshine duration forecasting is shown in Figure 2.1.

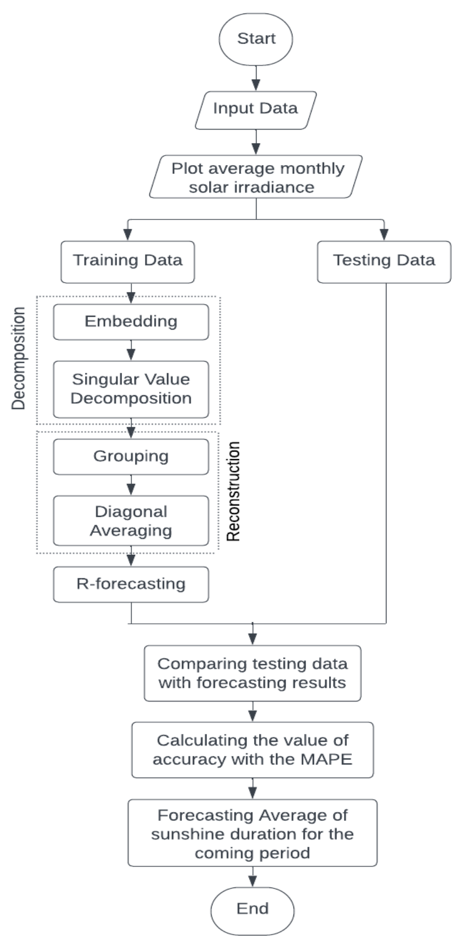

Figure 2. 1 Research flow chart

2. MAIN RESULTS

A. Data Pattern Analysis

The data employed in this research is consist of secondary data, namely monthly average data on the sunshine duration in Pasuruan for the period January 2010 to December 2023 with a total of 168 data. The data will be organized into time series data. Furthermore, a time series data plot is carried out to see how the data plot will be managed. The plot below represents the complete time series data

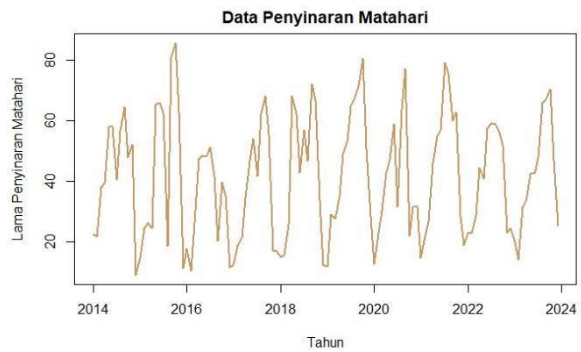

Figure 3. 1 Time Series Plot of Sunshine Duration

The graph of sunlight duration from 2014 to 2024 shows a clear seasonal pattern, with regular up-and-down oscillations indicating an annual cycle of sunlight exposure. Peaks generally occur during months with longer daylight periods, while troughs correspond to periods with less sunlight, such as during the rainy season. Periodicity is consistent across the years, although some anomalies are noticeable, particularly in 2015 and 2021, where the fluctuations are sharper and less regular. Significant deviations are also observed, such as an unusually high peak around late 2015 to early 2016 and deeper troughs around 2018 and 2022. These variations suggest that while the data generally follows a seasonal rhythm, external factors occasionally influence sunlight duration [9].

B. Singular Spectrum Analysis

Before entering the calculation process, the data is split into two sections namely 80% training and 20% testing data.

1. Decomposition

a. Embedding

Through this process, one-dimensional data is converted into multidimensional data, referred to as the trajectory matrix X which has$L \times K$. The value of L (window length) is established through a trial-and-error process, as the data employed in this study are 134 then the value of L that meets is $2 \leq L \leq 67$. To make the search for the optimum L more efficient, a trial of the value of $L = 30{,}40{,}50{,}60$ with the minimum MAPE value is carried out. The results obtained from the trial-and-error approach can be seen in the table below.

Table 3. 1 Trial-and-error results of MAPE value

L	30	40	50	60
MAPE	20,52%	20,04%	20,38%	22,2%

According to the table above, the L value that results in the lowest MAPE is $L = 40$, with a value of 20,04%. Additionally, further tracking is performed around $L = 40$ using the same method to find the L value with the lowest MAPE. The following table displays the results obtained from the trial-and-error approach.

Table 3.2 Trial-and-error results of MAPE value

L	41	42	43	44	45	46	47	48	49
MAPE	19,53%	20,06%	21,24%	21,71%	21,03%	21,48%	23,22%	21,49%	20,54%

Thus the L value used is $L = 41$ then the K value $= 134 - 41 + 1 = 94$. Thus, the trajectory matrix X (Hankel Matrix) can be structured as follows:

$$X_{(41\times 94)} = \begin{bmatrix} 20.30 & 26.48 & 26.34 & \cdots & 55.36 \\ 26.48 & 26.34 & 33.94 & \cdots & 17.16 \\ 26.34 & 33.94 & 28.72 & \cdots & 5.08 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 36.01 & 23.06 & 45.00 & \cdots & 21.60 \end{bmatrix}$$

b. Singular Value Decomposition

In this process, an SVD is established, consisting of three key elements: Singular Values ($\sqrt{\lambda_i}$), Eigenvectors (U_i), and Principal Components (V_i^T), known as eigentriple. According to the trajectory matrix $X_{(41x94)}$ obtained from the embedding process that has been carried out, it will then be used to obtain the eigentriple values needed at the Singular Value Decomposition stage. To obtain the eigentriple value, the first process will be formed symmetric matrix $S = XX^T$ as follows:

$$S = X_{(41\times 94)} . X^T_{(94\times 41)}$$

$$= \begin{bmatrix} 192085.8 & 175895.4 & 158954.6 & \cdots & 139447.0 \\ 175895.4 & 191968.1 & 175445.0 & \cdots & 149224.3 \\ 158954.6 & 175445.0 & 191292.7 & \cdots & 158315.5 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 139447.0 & 149224.3 & 158315.5 & \cdots & 138429.8 \end{bmatrix}$$

1) Singular Value ($\sqrt{\lambda_i}$)

From the symmetrical matrix $S_{(41x41)}$, the eigenvalue is first computed, followed by the singular value, which is derived as the square root of the corresponding eigenvalue. The results of these calculations are presented below.

Table 3.3 Eigenvalue and Singular Value Results

No	Eigen Value (λ_i)	Singular Value ($\sqrt{\lambda_i}$)
1	6338367.590	2517.61149
2	481845.038	694.15059
3	474387.302	688.75780
⋮	⋮	⋮
40	2182.120	46.71317
41	1625.646	40.31931

According to the table above, the eigenvalues can be observed have a decreasing order $\lambda_1 = 6338367.590 > \lambda_2 = 481845.038 > \lambda_3 = 474387.302 \cdots \lambda_{40} = 2182.120 > \lambda_{41} = 1625.646$

2) Eigenvector (U_i)

Once the eigenvalue is obtained, the eigenvector is then computed, yielding the following results:

$$U_1\begin{bmatrix}-0.1560\\-0.1569\\-0.1564\\\vdots\\-0.1565\end{bmatrix}, U_2\begin{bmatrix}-0.2056\\-0.1484\\-0.0435\\\vdots\\0.1750\end{bmatrix}, \cdots, U_{41}\begin{bmatrix}0.1450\\0.2088\\0.2147\\\vdots\\0.0732\end{bmatrix}$$

3) Principal Component (V_i^T)

Once the eigenvalue and eigenvector are obtained, the principal component is derived, with the results shown below

$$V_1\begin{bmatrix}-0.1005\\-0.1010\\-0.1025\\\vdots\\-0.1025\end{bmatrix}, V_2\begin{bmatrix}0.1034\\0.1350\\0.1398\\\vdots\\0.0835\end{bmatrix}, \cdots, V_{41}\begin{bmatrix}-0.0452\\0.0317\\0.0399\\\vdots\\0.1000\end{bmatrix}$$

2. Reconstruction

a. Grouping

This process is performed to group the eigentriples obtained from the SVD process. The grouping is done according to the characteristic properties of each component, including trends, seasonality, and noise. The grouping process uses Grouping Effect (r) as a parameter used to determine components related to trends and seasonality

1) Grouping of Noise Components

The noise components are grouped based on the parameter value r determined by looking at the number of eigentriples that do not describe the noise component in the following plot:

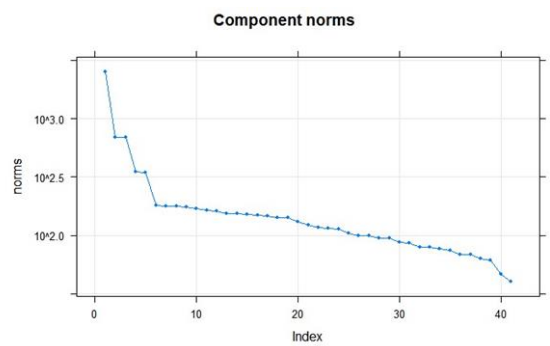

Figure 3. 2 Plot Eigentriple

It is said to be a noise component if the eigentriple plot decreases slowly. It can be observed in the plot above indicating that the eigentriple plot begin to decrease slowly at the 9th to 41st eigentriple so it can be said that the eigentriple is a noise component. Based on this, the 8 eigentriples that are not identified as noise components are the Grouping Effect (r) parameter values which will be utilized to recognize trend and seasonal components

2) Grouping of Trend and Seasonal components

After the noise component has been successfully grouped, the next step is to group the eigentriples that reflect the trend and seasonality components. The eigentriple numbers employed is 8 eigentriples consisting of eigentriples 1, 2, 3, 4, 5, 6, 7 and 8. To determine the trend and seasonal components, use the following plot of the reconstructed series:

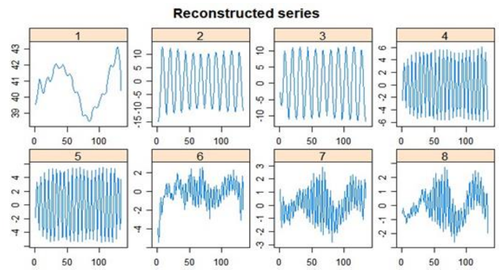

Figure 3.3 Plot Eigentriple 1-8

Trend components are defined as all elements on the plot that change slowly. It can be observed in the plot above that the series reconstructed by eigentriple 1 illustrates components that have slow variance, making eigentriple 1 identified as a trend component.

Eigentriples that define trend and seasonal components cannot be identified solely through the reconstructed plot and eigenvector pair scatterplot, but the identifier can also be seen by looking at the wcorrelation plot. The plot serves to find out how the correlation between eigentriples, if the plot color is solid black then the eigentriple has a strong correlation [2]. The w-correlation plot of 8 eigentriples is as shown below:

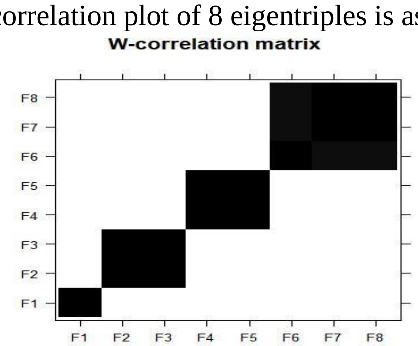

Figure 3.4 Plot w-correlation

The figure above describes that component F1 has no correlation with other components but only correlates with itself so it can be grouped into trends. Components F2 & F3, components F4 & F5 and components F7 & F8 have a strong correlation because the slices between these components have a dark color so they can be grouped into seasonal components.

b. Diagonal Averaging

The grouping outcomes are transformed into a new time series in this phase. This process is designed to derive the singular values of the decomposed components, which are then employed for forecasting.

Table 3.4 Results from Diagonal Averaging

No.	*Diagonal Averaging*
1.	18.335021
2.	20.055453
3.	25.235563
⋮	⋮
133	11.825159
134.	17.176906

The in-sample data produces a new time series with a length of 134, which is subsequently used for prediction through the R-forecasting method.

3. Applying R-Forecasting for Prediction

After obtaining the results from the calculation of diagonal averaging, forecasting is then conducted using the R-Forecasting technique, producing the following outcomes.

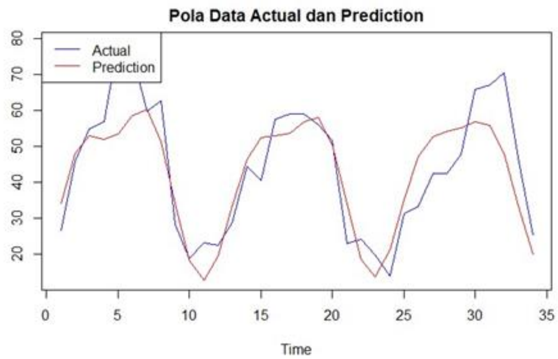

Figure 3. 5 Comparison of SSA-predicted and actual sunshine duration values

Figure 3.5 illustrates that the forecasting data forms a looping pattern comparable to the actual data, with predicted values closely following actual values. There are some data forecasting results that are very close to the actual data, but there are also some data that have a little distance difference.

Furthermore, the calculation of forecasting accuracy results in this research, MAPE method is utilized as follows.

$$\text{MAPE } = \frac{664.3197}{34} 100\% \ = \ 19,53\%$$

From the above calculations, it can be seen the obtained MAPE value is 19.53%, according to table 2.1 shows that the forecasting results have a fairly good level of accuracy using the Singular Spectrum Analysis method with a value of $L = 41$ and Grouping Effect $(r) = 8$ so that it can be used to forecast the average value of solar irradiation in the future. Furthermore, forecasting is carried out for the next 12 months and the following results are obtained:

Table 3.5 Predicted Results for 2024

No	Month	Average of Sunshine Duration
1	January	34,4%
2	February	48.14%
3	March	52,93%
4	April	52%
5	May	53.45%
6	June	58,6%
7	July	60,1%
8	August	51,23%
9	September	34,16%
10	October	18,42%
11	November	12,82%
12	December	19,51%

From Table 8, it is evident that the average duration of solar irradiation obtained in 2024 in Pasuruan is highest in July with a value of 60.1% while these results can provide more energy for activity and growth, they can also pose significant risks and impacts on the environment, human health, and life as a whole. While the lowest value acquisition is November with a value of 12.82%. If the average solar irradiation is low, some possible consequences are that it can slow down the photosynthesis process of plants, it can affect the availability of energy for ecosystems and human activities. The impact extends to energy production from solar panels and wind turbines in renewable energy facilities and related sectors.

3. CONCLUSION

The MSSA method successfully captured the trend, seasonality, and noise components in the sunlight duration data, achieving a minimum MAPE of 19.53%, indicating good predictive accuracy. The grouping process effectively separated meaningful patterns from noise, demonstrating the capability of MSSA to decompose and reconstruct time series data with complex structures. These results suggest that MSSA is a reliable approach for forecasting time series with seasonal characteristics and could be applied to similar environmental datasets for improved predictive performance.

REFERENCES

[1] Amalia Idrus, R., 2022. Penerapan Metode Singular Spectrum Analysis dalam Peramalan Jumlah Produksi Beras di Kabupaten Gowa. *Variansi J. Stat. Its Appl. Teach. Res*, Vol. 4, No. 2, 49–58.

[2] Andhika, G. B., Sumarjaya, I. W., & Srinadi, I. G. A. M., 2020. Peramalan Nilai Tukar Petani Menggunakan Metode Singular Spectrum Analysis. *E-Jurnal Mat*, Vol. 9, No. 3, 171.

[3] Arifin, J., 2024. Badan Riset dan Inovasi Nasional: Aktivitas Matahari Meningkat. *Radar Bromo*. https://radarbromo.jawapos.com/pasuruan/1001626607/badan-riset-dan-inovasi-nasional [14 Maret 2024].

[4] Basari, M. S. N., & Achmad, A. I., 2021. Metode Singular Spectrum Analysis untuk Meramalkan Indeks Harga Konsumen Indonesia Tahun 2019. *Prosiding Statistika*, Vol. 7, 484–491.

[5] Bounoua, W., & Aftab, M. F., 2024. Multivariate singular spectrum analysis and detrended fluctuation analysis for plant-wide oscillations denoising. *IFAC-PapersOnLine*, Vol. 58, No. 14, 676–681.

[6] Christienova, S. I., Pratiwi, E. W., & Darmawan, G., 2018. Perbandingan Model Peramalan Singular Spectrum Analysis (SSA) dan Fourier Series Analysis (FSA) pada Data Suhu Udara di Surabaya. *Berk. MIPA*, Vol. 25, No. 1, 94–106.

[7] Dwiadi, Q. A., Indriani, A., Nauli, T. S., 2023. Peramalan Konsumsi Gas Alam Amerika Serikat dengan Double Seasonality menggunakan Singular Spectrum Analysis (SSA). *Innovation Journal*, Vol. 3, 6405–6421.

[8] Fahrizal Firmani, R. S., 2023. Suhu Panas di Jawa Timur hingga Akhir Oktober Imbas El Nino. *Radar Bromo*. https://radarbromo.jawapos.com/pasuruan/1003051735/suhu-panas-di-jawa-timur-hingga-akhir-oktober-imbas-el-nino-ini-penjelasan-bmkg-pasuruan [12 Mei 2025]

[9] Hassani, H., & Thomakos, D., 2010. A review on singular spectrum analysis for economic and financial time series. *Stat. Interface*, Vol. 3, No. 3, 377–397.

[10] Hidayat, K. W., Wahyuningsih, S., & Nasution, Y. N., 2020. Pemodelan Jumlah Titik Panas di Provinsi Kalimantan Timur dengan Metode Singular Spectrum Analysis. *Jambura J. Probab. Stat*, Vol. 1, No. 2, 78–88.
[11] Huang, W. L., Zhou, Y. X., Zhou, Y., Liu, W. J., & Li, J. D., 2024. A reweighted damped singular spectrum analysis method for robust seismic noise suppression. *Petroleum Science*, Vol. 21, No. 3, 1671–1682.
[12] Nugroho, P. A., 2023. Implementasi Jaringan Syaraf Tiruan Multi-Layer Perceptron Untuk Prediksi Penyinaran Matahari Kota Bandung. *Komputa J. Ilm. Komput. dan Inform*, Vol. 12, No. 1, 83–90.
[13] Pagano, P. P., Ciaccio, E. J., & Garan, H., 2022. Separation of cardiogenic oscillations from airflow waveforms using singular spectrum analysis. *Comput. Methods Programs Biomed*, Vol. 220, 106803.
[14] Pujiastuti, A., & Harjoko, A., 2016. Sistem Perhitungan Lama Penyinaran Matahari dengan Metode Otsu Threshold (Studi Kasus: St. Klimatologi Barongan). *Compiler*, Vol. 5, No. 2.
[15] Puspitasari, W., Rustiana, S., Suparman, Y., & Purwandari, T., 2019. Perbandingan Hasil Peramalan Curah Hujan Bulanan Kota Bogor dengan Seasonal Autoregressive Integrated Moving Average (SARIMA) dan Singular Spectrum Analysis (SSA). *Prosiding Sendik*, Vol. 5, No. 2, 206–217.
[16] Rahmani, D., & Fay, D., 2022. A state-dependent linear recurrent formula with application to time series with structural breaks. *J. Forecast*, Vol. 41, No. 1, 43–63.
[17] Ridla, M. A., Azise, N., & Rahman, M., 2023. Perbandingan Model Time Series Forecasting Dalam Memprediksi Jumlah Kedatangan Wisatawan Dan Penumpang Airport. *Simkom*, Vol. 8, No. 1, 1–14.
[18] Salsalina, S., & Widyasari, R., 2023. Prediksi Jumlah Pemakaian Air Bersih Menggunakan Metode Hybrid Singular Spectrum Analysis (SSA) dan SARIMA di PDAM Tirtanadi Sibolangit. *J. Lebesgue J. Ilm. Pendidik. Mat. Mat. dan Stat*, Vol. 4, No. 2, 1301–1310.
[19] Sugiarto, 2016. Aplikasi Metode Singular Spectrum Analysis (SSA) dalam Peramalan Produksi Padi di Provinsi Sulawesi Tenggara. *JPNM (Jurnal Pustaka Nusantara Multidisplin)*, Vol. 4, No. 1, 1–23.
[20] Sunarmi, N., 2022. Analisis Faktor Unsur Cuaca terhadap Perubahan Iklim di Kabupaten Pasuruan pada Tahun 2021. *Newton-Maxwell J. Phys*, Vol. 3, No. 2, 56–64.
[21] Syah, M. A., et al., 2025. Assessing Coffee Farmers' Awareness of Climate Change Through the Climate Change Awareness Index (CCAI) Approach in Tutur, Pasuruan, Indonesia. *Agriverse*, Vol. 1, No. 1, 23–33.
[22] Varadi, F., & Pap, J. M., 1996. Application of Singular Spectrum Analysis to Solar Irradiance Variability. *Sol. Drivers Interplanet. Terr. Disturbances*, Vol. 95, 576.
[23] Yustiana, F., & Ibrahim, N., 2023. Analisis Deret Waktu Curah Hujan dan Karakteristik Iklim di Kota Majalengka. *J. Tek. Sipil*, Vol. 12, No. 2.