APPLICATION OF THE HYBRID SINGULAR SPECTRUM ANALYSIS – ARIMA MODEL FOR INDONESIA'S INFLATION RATE (2018-2023)

Sri Rahayu1)
Aswi Aswi2)
Muhammad Fahmuddin Sudding3)
1,2,3)Statistics Study Program, Faculty of Mathematics and Natural Sciences, Universitas Negeri Makassar, Makassar, Indonesia
e-mail: aswi@unm.ac.id

ABSTRACT

Accurately forecasting inflation rates is crucial for maintaining economic stability and guiding macroeconomic policy decisions in Indonesia. However, the inherent volatility and complex patterns of economic data, which often include both trend and seasonal components, present significant challenges. This research aims to determine the results and accuracy of forecasting inflation rates in Indonesia using Hybrid Singular Spectrum Analysis (SSA) – Autoregressive Integrated Moving Average (ARIMA). Hybrid SSA-ARIMA combines two complementary time series methods to enhance forecasting accuracy, particularly for economic data characterized by trends and seasonality. The data used consists of national consumer price inflation rates (Y-on-Y) from January 2018 to December 2023. The forecast accuracy, as measured by Mean Absolute Percentage Error (MAPE), showed 56.26797% for Singular Spectrum Analysis and 18.88851% for Hybrid SSA-ARIMA. This demonstrates that Hybrid SSA-ARIMA has superior forecasting capabilities compared to Singular Spectrum Analysis in predicting inflation rates in Indonesia.

Keywords: ARIMA, Forecasting, Hybrid, Inflation, SSA.

INTRODUCTION

Inflation is one of the most critical macroeconomic indicators, as it significantly influences a nation's economic conditions. Ensuring its stability is essential to prevent adverse effects and financial vulnerabilities. Consequently, forward-looking inflation information is crucial for formulating effective macroeconomic policies and programs, which can be obtained through statistical forecasting techniques (Fajar & Rachmad, 2018). The development of these methods has accelerated and become more complex with advancements in computing technology. Notably, the emergence of hybrid time series forecasting methods, which combine two different forecasting approaches, exemplifies this progress (Fajar, 2019). Hybrid models leverage the strengths of multiple methods, leading to more accurate forecasting results (Darmawan et al., 2022).

One widely used hybrid time series forecasting method is Singular Spectrum Analysis-Autoregressive Integrated Moving Average (SSA-ARIMA). SSA is a non-parametric time series analysis technique that does not require stationarity assumptions or residual normality, making it suitable for both stationary and non-stationary data. SSA is particularly effective at identifying trends, seasonal patterns, and noise in time series data. A key concept in SSA is 'separability,' which refers to the ability to decompose the original time series into distinct components, such as trend, seasonality, and noise (Siti et al., 2019). This separability is crucial for isolating these components effectively, leading to more accurate forecasts. This approach is particularly well-suited for economic and business data, which often exhibit trends and seasonal patterns. Meanwhile, ARIMA is a parametric time series analysis technique applicable to both stationary and non-stationary data, as well as seasonal data. ARIMA involves systematic stages, including model identification, parameter estimation, and diagnostic testing, making it an ideal candidate for hybrid modeling (Arumsari et al., 2021).

Several studies have utilized the hybrid SSA-ARIMA approach (Fajar, 2019), which compared the performance of ARIMA, SSA, and hybrid SSA-ARIMA in forecasting Indonesian economic growth. The study demonstrated that the hybrid SSA-ARIMA method outperformed both ARIMA and SSA, as indicated by its lower Root Mean Square Error (RMSE) values. Additionally, a research on inflation forecasting for East Kalimantan Province using the Hybrid SSA-ARIMA has been conducted (Arumsari et al.,2021), This study found that the hybrid SSA-ARIMA(1,1,1) model improved forecasting accuracy for inflation in East Kalimantan Province in 2021. The highest inflation rate in December 2021 was 0.92%, with forecasting accuracy metrics showing a Root Mean Square Error (RMSE) of 0.069399 and a Mean Absolute Percentage Error (MAPE) of 32.61084%.

The use of the Hybrid SSA-ARIMA model for analyzing seasonal time series data was explored (Darmawan et al.,2022). In this context, additive and multiplicative seasonal patterns refer to the way seasonal fluctuations in the data behave relative to the overall trend. An additive seasonal pattern means that the seasonal effect remains constant regardless of the overall trend, while in a multiplicative seasonal pattern, the seasonal effect increases or decreases proportionally with the trend. The study found that for data exhibiting an additive seasonal pattern, the Hybrid SSA-ARIMA method with Alexandrov automatic grouping was more accurate (MAPE=0.13%), while for data with a multiplicative seasonal pattern, the Hybrid SSA-ARIMA method with alternative automatic grouping was more accurate (MAPE=3.63%). These findings demonstrate the flexibility and effectiveness of the hybrid method in handling different types of seasonal variations in time series data.

Based on previous research, the hybrid SSA-ARIMA method has proven to be effective and superior in improving the accuracy of economic data forecasting, particularly for data containing seasonal and trend components. By combining the strengths of SSA in separating data components and ARIMA in handling stationarity, the hybrid approach consistently delivers more accurate results compared to single methods. This method has successfully reduced forecasting errors in both economic growth and inflation predictions. These advantages make it a powerful tool for complex economic data analysis and a more reliable approach for achieving precise predictions in various economic contexts. The aim of this research is to evaluate the performance and accuracy of the Hybrid Singular Spectrum Analysis - ARIMA model in forecasting Indonesia's inflation rate for the period from January to December 2023.

METHOD

Type of Research

This study employs a quantitative research approach, utilizing secondary data on inflation rates in Indonesia. The data were obtained from publications by the Central Statistics Agency, covering the period from January 2018 to December 2023. The primary variable in this study is inflation, defined as the continuous rise in the prices of goods and services over a specific period.

Data Analysis Technique

The data analysis in this study follows these steps:

1) Conduct a descriptive analysis of the inflation variable
2) Split the inflation data into in-sample and out-sample datasets

3) Perform time series analysis using SSA on the in-sample data, with the following steps:
   a. Embedding: Convert the Indonesia inflation rate data into a trajectory matrix.
   b. Singular Value Decomposition (SVD): Identify the eigentriples, including singular values $\sqrt{\lambda_i}$, eigenvector U_i, and principal components V_i^T.
   c. Grouping: Group the eigentriples obtained from the SVD step into components such as trend, seasonality, and noise.
   d. Diagonal Averaging: Calculate the average diagonal values of the separated components to obtain singular values for the forecasting stage.
   e. Conduct SSA forecasting using the Recurrent (*R-forecasting*) method with the trend and seasonal components
   f. Assess the accuracy of SSA forecasting using MAPE
4) Perform time series analysis with the ARIMA model on the noise component data obtained from the SSA reconstruction, including the following steps:
   a) Check the stationarity of the noise component data.
   b) Plot the Auto-Correlation Function (ACF) and Partial Auto-Correlation Function (PACF) to identify the appropriate ARIMA model.
   c) Estimate model parameters and assess their significance
   d) Perform diagnostic checks for white noise and normality.
   e) Estimate model parameters and assess their significance
   f) Conduct forecasting using the ARIMA model.
5) Implement hybrid SSA-ARIMA forecasting by combining the forecasts from SSA and ARIMA models.
6) Evaluate the forecasting accuracy of the hybrid SSA-ARIMA model using MAPE.

Singular Spectrum Analysis (SSA)

According to Golyandina et al. (2001), the basic SSA algorithm consists of two main stages, namely decomposition (embedding & singular value decomposition) and reconstruction (grouping & diagonal averaging), which are explained as follows:

1) Embedding

Embedding is the process of converting one-dimensional time series data into a trajectory matrix by transforming the data from a vector to a matrix. For example, time series data of length n, without missing data, is expressed by X = $\{x_1, x_2, \dots, x_n\}$. The data is transformed into a matrix of size L × K with L being the window length which becomes the matrix row where $2 < L < n/2$. Since there is no definitive method for determining the optimal value of L, it is typically selected through trial and error. The number of columns, K, is given by $K = n - L + 1$. The resulting matrix can be represented as follows:

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

Matrix **X** is also referred to as a Hankel matrix, characterized by having identical values along all its anti-diagonal elements. Therefore, at this stage, the output is a Hankel matrix of size L × K (Khaeri et al., 2017).

2) Singular Value Decomposition (SVD)

SVD starts by determining the eigenvalues $(\lambda_1, \lambda_2, \dots, \lambda_L)$ of the matrix **S** = **XX**^{T} where $\lambda_1 \geq \cdots \geq \lambda_L > 0$, and the eigenvector $(U_1, U_2, \dots, U_L)$ of the matrix. If notated $V_i = \frac{X^T U_i}{\sqrt{\lambda_i}}$ then the SVD of the path matrix is as follows (Irmawati et al., 2018) :

$$\boldsymbol{X} = X_1 + X_2 + \cdots + X_d$$
$$\boldsymbol{X} = U_1\sqrt{\lambda_1}V_1^T + U_2\sqrt{\lambda_2}V_2^T + \cdots + U_d\sqrt{\lambda_d}V_d^T \tag{2}$$
$$\boldsymbol{X} = \sum_{i=1}^{d} U_i\sqrt{\lambda_i}V_i^T$$

with:

i = 1, 2, ..., d and d = max $\{i\ ;\lambda_i > 0\}$

Basic concepts at this stage is to get $\boldsymbol{X_i}$, where each matrix in the sequence contains an eigentriple consisting of eigenvector U_i, singular value $\sqrt{\lambda_i}$, and principal components $\boldsymbol{V_i^T}$.

3) Grouping

Grouping is the stage of separating additive components such as trend, seasonality and noise contained in time series data. The grouping process is carried out by grouping sets of indices {1, 2, ..., d} into *m* subsets which can be denoted by $I = I_1, I_2, \dots, I_m$ which then forms a matrix based on the singular value decomposition $\boldsymbol{X_i}$ as follows:

$$\boldsymbol{X_I} = \boldsymbol{X_{I1}} + \boldsymbol{X_{I2}} + \cdots + \boldsymbol{X_{Im}} \tag{3}$$

The steps for looking at the set $I = I_1, I_2, \dots, I_m$ are called eigentriple grouping. Eigentriples with almost the same characteristics will be grouped into one group or component (Satriani et al., 2020).

4) Diagonal Averaging

The next stage is to reconstruct each matrix in $\boldsymbol{X_I}$ into new time series data with length n (Satriani dkk., 2020). Suppose the matrix $\boldsymbol{Y}^{(k)}$ is of size $L \times K$ with elements $y_{ij}^{(k)}$, $where\ 1 \leq i \leq L, 1 \leq j \leq K$. The matrix $\boldsymbol{Y}^{(k)}$ is converted into time series data $f_0, \dots, f_{n-1}$ through diagonal averaging as follows (Arumsari et al., 2021):

$$\hat{f}_t^{(k)} = \begin{cases} \frac{1}{t}\sum_{i=1}^{t} y_{i,t-i+1,}^{*(k)} & for\ 1 \le t < L^* \\ \frac{1}{L^*}\sum_{i=1}^{L^*} y_{i,t-i+1,}^{*(k)} & \text{for } L^* \le t \le K^* \\ \frac{1}{n-t+1}\sum_{i=t-K^*+1}^{n-K^*+1} y_{i,t-i+1,}^{*(k)} & \text{for } K^* \le t < n \end{cases} \tag{4}$$

The diagonal averaging stage aims to get the singular value of the components that have been separated by finding the average of the diagonals which is then used for forecasting.

5) SSA Forecasting

Reccurent forecasting (R-Forecasting) is one of the forecasting methods in SSA. In the R-forecasting method, diagonal averaging is used to obtain reconstruction and continuation using the Linear Reccurent Formula (LRF). LRF coefficient estimation uses eigenvectors that have been obtained at the SVD stage. Let u_i^{∇} be the first *L*-1 component vector of the eigenvector u_i dan π_i be the last component of u_i $(i = 1, \dots, I)$ with $v^2 = \sum_{i=1}^{I} \pi_i^2$. R can be defined as a vector consisting of the LRF coefficients of a component and can be calculated using the following equation (Hidayat et al., 2020):

$$R = (R_{L-1}, \dots, R_1) = \frac{1}{1 - v^2}\sum_{i=1}^{I} \pi_i\, u_i^{\nabla} \tag{5}$$

Forecasting results are obtained based on the following equation (Siringoringo et al., 2022) :

$$\hat{f}_t^{(k)} = \begin{cases} \hat{f}_t^{(k)} & \text{for t=1, 2, …, } n \\ \sum_{j=1}^{L-1} R_{j=1}^{(k)}\, \hat{f}_{t-j}^{(k)} & \text{for t=}n\text{+1, }n\text{+2,…, }n+m \end{cases} \tag{6}$$

The SSA model used to obtain estimated values can be written as follows (Ete et al., 2020):

$$\hat{f}_t = \hat{f}_t^{(1)} + \hat{f}_t^{(2)} \tag{7}$$

where $k = 1$ is the trend component and *k*=2 is the seasonal component (Siringoringo et al., 2022).

Autoregressive Integrated Moving Average (ARIMA)

The ARIMA model is a time series model that addresses non-stationarity by applying a differencing process of order *d* to achieve stationarity. In general, the ARIMA model is denoted as ARIMA(p, d, q) and can be formulated as follows (Aswi & Sukarna, 2006) :

$$\phi_p(B)(1-B)^d Z_t = \theta_q(B)a_t \tag{8}$$

with:

p, d, q = orders of *autoregressive* (AR), differencing *(d),* and moving average (MA) components

B = *Backshift operator*

$\phi_p(B)$ = non-seasonal operator AR = $(1-\phi_1 B-\phi_2 B-\cdots-\phi_p B^p)$

$(1-B)^d$ = non-seasonal differencing of orde *d*

$\theta_q(B)$ = non-seasonal MA operator = $(1-\theta_1 B-\theta_2 B-\cdots-\theta_q B^q)$

The first step in constructing an ARIMA model is determining whether the data is stationary, which is crucial for accurate time series modeling. If the data is non-stationary with respect to the mean, differencing is applied; if non-stationary with respect to variance, a Box-Cox transformation is used. Once stationarity is achieved, the ARIMA model is identified by analyzing the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots to suggest initial values for the model parameters (p, d, q). After selecting an initial ARIMA model, parameter estimation is conducted to determine the model coefficients. Model adequacy is then evaluated using diagnostic tests. The significance of the model parameters is tested with the t-test, assuming normally distributed errors. To confirm that the residuals behave as white noise, the Ljung-Box test is employed, testing for the absence of autocorrelation in the residuals. Additionally, the Shapiro-Wilk test is used to verify that the residuals follow a normal distribution. If multiple models pass the diagnostic tests, the Akaike Information Criterion (AIC) is compared, and the model with the lowest AIC is selected as the best fit. Once the optimal ARIMA model is chosen, forecasting is performed by applying the model's parameters to predict future values of the time series, taking into account past values and residuals to enhance accuracy.

Hybrid Autoregressive Integrated Moving Average – Singular Spectrum Analysis

The forecasting results of the SSA - ARIMA hybrid model ($\hat{X}_t$) are given by (Arumsari et al., 2021):

$$\hat{X}_t = \hat{f}_t + \hat{Z}_t \tag{9}$$

where $\hat{f}_t$ represents the forecast obtained using the SSA method, and $\hat{Z}_t$ represents the forecast obtained from the ARIMA method, applied to the noise component of the data.

Mean Absolute Percentage Error (MAPE)

The MAPE value can be calculated using the formula (Wulandari & Yurinanda, 2021) :

$$MAPE = \left(\frac{1}{n}\sum_{t=1}^{n}\left|\frac{X_t - \hat{X}_t}{X_t}\right|\right) \times 100\% \quad (10)$$

with :
X_t = actual data value
$\hat{X}_t$ = forecasted data value
n = Number of data points

Table 1 presents the MAPE value criteria (Wulandari & Yurinanda, 2021):

Table 1. MAPE value criteria

MAPE (X)	Interpretation
X < 10%	Forecasting ability is very good.
10% ≤ X < 20%	Forecasting ability is good.
20% ≤ X < 50%	Forecasting ability is sufficient.
X ≥ 50%	Forecasting ability is poor

RESULTS AND DISCUSSION

Descriptive Analysis

The descriptive statistics for Indonesia's inflation rate are provided in Table 2.

Table 2. Descriptive statistics of Indonesia's inflation rate

Variable	Mean	Standard Deviation	Min	Max
Inflation	2.95	1.19	1.32	5.95

Table 2 shows that the average inflation rate in Indonesia is 2.95 with a standard deviation of 1.19. The lowest inflation rate value of 1.32 occurred in August 2020 and the highest inflation rate value of 5.95 occurred in September 2022. Figure 1 shows the time series plot of inflation rate data in Indonesia.

Figure 1 presents the time series plot of the inflation rate in Indonesia from January 2018 to December 2022. The plot shows that the inflation rate exhibited a decreasing trend from January 2018 to December 2020, followed by an increasing trend from January 2021 to December 2022. The first step in this analysis is to divide the inflation rate data into two parts: in-sample data and out-sample data. The in-sample data, covering the period from January 2018 to December 2022 with a total of 60 data points, will be used for model analysis. Meanwhile, the out-sample data, consisting of 12 data points from January 2023 to December 2023, will be used for forecasting and evaluating the accuracy of the forecast.

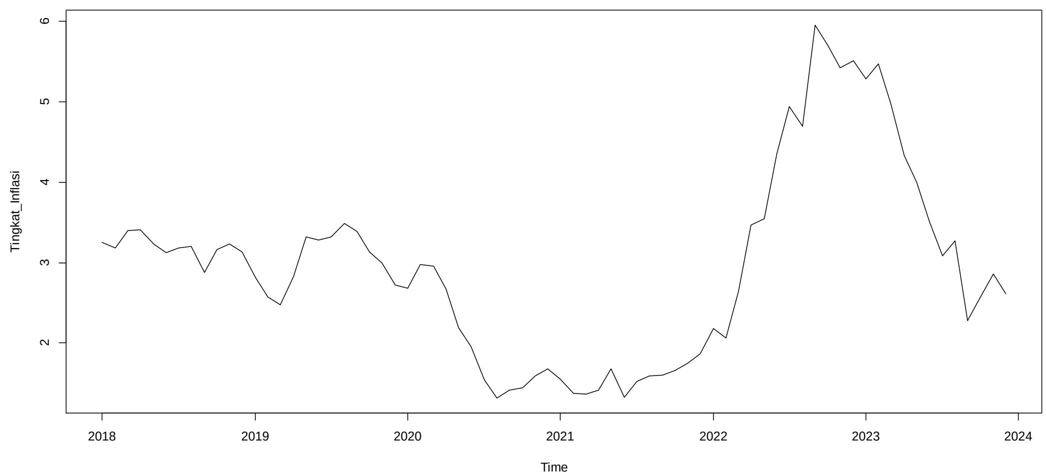


Figure 1. Time Series plot of inflation rate data in Indonesia

Singular Spectrum Analysis

1) Embedding

At this stage, the window length (L) value is determined using a trial and error process. The L value to be analyzed is 2 < L < n/2 or 2 < L < 30. Based on the smallest MAPE value, the L value used in this analysis is L = 25. Based on the L value obtained, K = n – L + 1 = 60 – 25 + 1 = 36. The trajectory matrix is arranged as follows:

$$\mathbf{X}_{25\times 36} = \begin{bmatrix} 3.25 & 3.18 & 3.40 & \cdots & 1.68 \\ 3.18 & 3.40 & 3.41 & \cdots & 1.55 \\ 3.40 & 3.41 & 3.23 & \cdots & 1.38 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 2.68 & 2.98 & 2.96 & \cdots & 5.51 \end{bmatrix}$$

2) Singular Value Decomposition

At this stage, calculations are carried out to obtain the triple eigenvalues based on the trajectory matrix $\mathbf{X}_{(25\times36)}$. The first step that needs to be done is to form a symmetric matrix $\mathbf{S} = \mathbf{XX}^T$. After getting the symmetric matrix $\mathbf{S}_{(25\times25)}$, the next step is to calculate the eigentriple. This process will produce 25 eigentriples consisting of singular value, eigenvector and principal component, each of which is in Tables 3, 4, and 5.

Table 3. Singular Values

No	Eigenvalues	Singular values
1	5337.246	73.0564567
2	402.9124	20.0726790
3	55.87322	7.4748394
⋮	⋮	⋮

No	Eigenvalues	Singular values
25	0.0474577	0.2178479

Table 4. Eigenvector

No	U_1	U_2	...	U_{25}
1	-0.2247623	-0.15872677	...	-0.06880761
2	-0.2217043	-0.16487471	...	0.22662370
3	-0.2183852	-0.17108019	...	-0.37164978
⋮	⋮	⋮	⋱	⋮
25	-0.2099606	0.42532626	...	0.06326486

Table 5. Principal component

No	V_1	V_2	...	V_{25}
1	-0.2118303	-0.005990249	...	-0.16055731
2	-0.2110611	-0.007168110	...	0.17410403
3	-0.2104319	-0.008564221	...	0.01860075
⋮	⋮	⋮	⋱	⋮
36	-0.1925521	0.4009263	...	-0.03531605

3) Grouping

In this step, eigentriple grouping obtained from the SVD stage is performed. Eigentriples with similar characteristics are grouped into a single component or cluster through graphical analysis, using singular value plots. Figure 2 represents a singular value plot displaying values for i = 1, 2, 3, ..., 25. Eigentriple grouping is essential as it allows for the identification and separation of dominant patterns or components, such as trends and seasonal effects, from noise. By grouping eigentriples with similar characteristics, we can concentrate on the primary signals that significantly contribute to the time series structure, thereby enhancing the accuracy of the analysis. This grouping also reduces data dimensionality, making subsequent modeling more efficient and interpretable. Practically, it leads to more precise forecasts by isolating the most informative components and filtering out irrelevant fluctuations, ultimately improving the model's robustness.

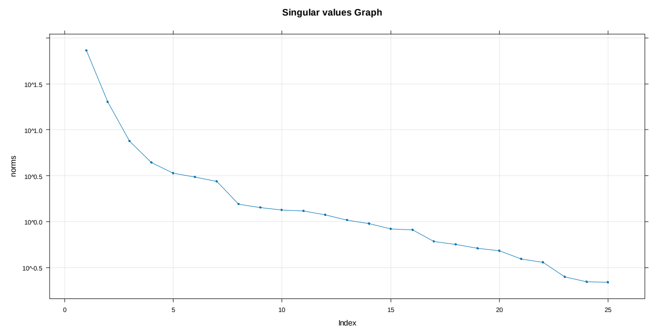


Figure 2. Singular values Graph

In Figure 2, singular value 1, which is 73.0564567, has the highest contribution compared to other singular values, indicating that it has the greatest influence on the time series component and data characteristics. The singular values begin to gradually decline from eigentriple 8 to eigentriple 25, which can be identified as the noise component. However, this grouping remains somewhat subjective, so it is necessary to re-examine all eigentriples. One approach to identifying eigentriples that contain elements of trend, seasonality, and noise is to analyze the eigenvector plot, as shown in Figure 3.

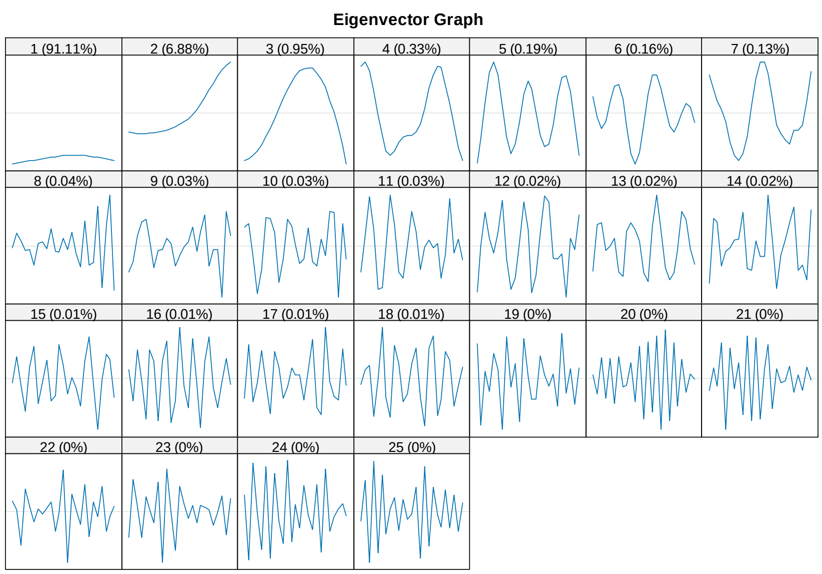


Figure 3. Eigenvector Graph

Based on Figure 3, the eigenvector 2 graph shows a slow and steady variation, indicating a persistent and long-term pattern. Specifically, the trend component exhibits a monotonic increase from the bottom left to the top right, without frequent fluctuations or cyclical patterns. This gradual upward movement suggests that eigenvector 2 captures the underlying trend of the data. Furthermore, for grouping seasonal components, one method that can be used is periodogram analysis according to Golyandina et al. (2018). The seasonal components, as identified through periodogram analysis, display regular and repetitive patterns. These eigenvectors capture cyclical variations that repeat at consistent intervals, which is indicative of seasonality. The periodic nature of these eigenvectors is a key indicator of seasonal behavior. The eigenvectors grouped into the noise category do not exhibit clear patterns of trend or seasonality. Instead, they are characterized by irregular, random fluctuations with no discernible repetitive structure. This randomness suggests that these eigenvectors represent noise, or the residual variation that cannot be attributed to trend or seasonality. The results of grouping the eigenvector into trend, seasonal and noise components are obtained as follows:

Table 6. Eigenvector Grouping

Group	Eigenvector
Tren	2
Seasonal	8, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25
Noise	1, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13

The w-correlation matrix is a tool used to evaluate the strength of correlations between different components in time series analysis. It provides a visual representation of how closely related various components are, based on their correlation coefficients. In this context, the w-correlation matrix is used to assess the separation of trend, seasonal, and noise components identified in the eigenvector analysis. According to Golyandina et al. (2018), examining the w-correlation matrix helps verify the accuracy of the component separation. In the matrix plot, darker color gradations indicate a strong correlation, suggesting that components with similar characteristics, such as trend or seasonality, are closely related. On the other hand, lighter color gradations reflect weaker correlations, which are typically associated with noise components that do not exhibit clear patterns of trend or seasonality. The following image presents the w-correlation matrix plot for the identified trend, seasonal, and noise components, demonstrating the correlation relationships between these groups.

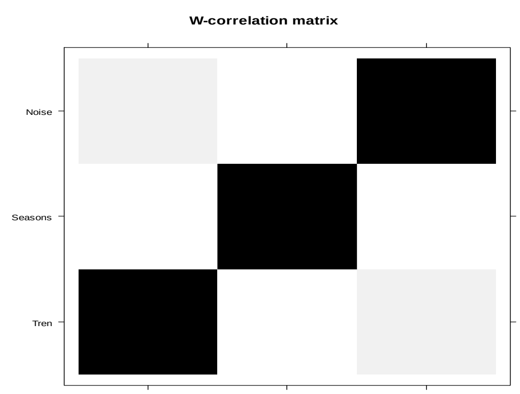
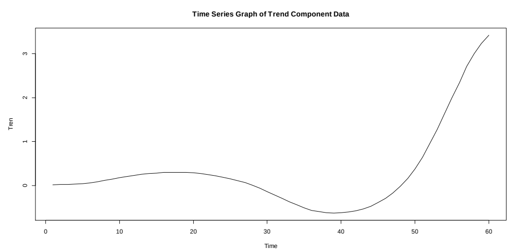
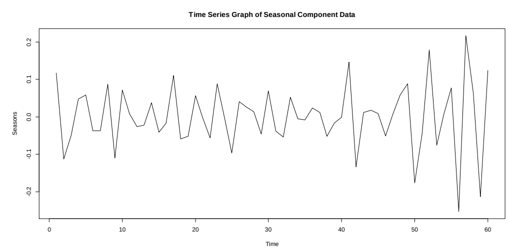


Figure 4. Plot W-Correlation Matrix

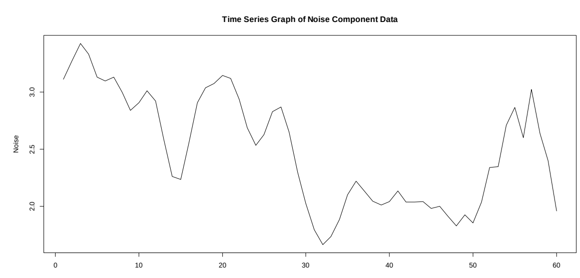
In Figure 4, the w-correlation matrix plot visually represents the correlation between different components (trend, seasonal, and noise). The color gradation in this matrix is determined by the correlation coefficients calculated between pairs of components. A dark color gradation between two components signifies a strong correlation. For instance, if the trend component and the seasonal component show no significant color gradation and have a correlation value of 0.002, it implies that there is minimal correlation between these components, indicating they can be effectively separated. Conversely, lighter color gradations indicate weak correlations. For example, a correlation value of 0.014 between seasonal and noise components, with no significant color gradation, suggests that these components are also well separated. When slight color gradation is observed, such as a correlation value of 0.088 between the trend and noise components, it implies a weak correlation. Although the correlation is not close to 1, the relatively minor gradation indicates that the components can still be separated, though there might be a very slight overlap. Based on these observations, the w-correlation matrix plot confirms that the trend, seasonal, and noise components are well separated, as indicated by the minimal color gradation and low correlation values between them.



(a)



(b)



(c)

Figure 5. (a) Trend (b) Seasonal (c) Noise

4) Diagonal Averaging

At this stage, each component is reconstructed using the corresponding eigentriple. Diagonal averaging is achieved by summing the trend and seasonal components from the reconstruction results. The series of reconstructed components and diagonal averaging results are presented in Table 7.

Table 7. Reconstruction Results and Diagonal Averaging

Nth Time	Reconstruction		Diagonal Averaging
	Trend	Seasonal	
1	0.019085360	0.1171898914	0.13627525
2	0.021331351	-0.1120307957	-0.09069944
3	0.023859886	-0.0483674363	-0.02450755
⋮	⋮	⋮	⋮
60	3.422883012	0.1247528010	3.54763581

5) SSA Forecasting

Following the diagonal averaging, the next step is to perform modeling for forecasting the subsequent 12 months. The SSA forecasting method utilized is the R-forecasting technique with LRF coefficient estimation. The forecasting model for the trend component is presented as follows:

$$\hat{f}_t^{(1)} = -0.082\hat{f}_{t-1}^{(1)} - 0.0867\hat{f}_{t-2}^{(1)} + \cdots + 0.206\hat{f}_{t-24}^{(1)} \tag{11}$$

Similarly, the forecasting model for the seasonal component is given below

$$\hat{f}_t^{(2)} = -0.119\hat{f}_{t-1}^{(2)} + 0.015\hat{f}_{t-2}^{(2)} + \cdots - 0.574\hat{f}_{t-24}^{(2)} \tag{12}$$

Based on the forecasting models developed for the trend and seasonal components, the next step is to combine the two components to obtain the SSA forecast. The results of the SSA forecasting are as follows:

Table 8. Forecasting Results of Inflation Rates in Indonesia Using SSA

Month	Forecasting		Inflation Rate
	Tren	Seasonal	
January 2023	3.01	-0.02	2.99
February 2023	3.03	-0.05	2.98
March 2023	2.98	0.09	3.07
April 2023	2.85	-0.06	2.79
May 2023	2.66	-0.03	2.63
June 2023	2.39	0.09	2.48
July 2023	2.04	-0.06	1.98

Month	Forecasting		Inflation Rate
	Tren	Seasonal	
August 2023	1.62	0.02	1.64
September 2023	1.12	0.01	1.13
October 2023	0.55	-0.04	0.51
November 2023	-0.09	0.05	-0.04
December 2023	-0.80	-0.03	-0.83

Based on Table 8, March 2023 exhibits the highest inflation rate (3.07), while December shows the lowest inflation rate (-0.83).

6) SSA Forecasting Accuracy

The forecasting accuracy was measured using the Mean Absolute Percentage Error (MAPE). Based on Equation 10, the MAPE for Indonesia's inflation rate data is 56.26797%. This relatively high value for the Singular Spectrum Analysis (SSA) model indicates limited forecasting accuracy. The high MAPE can be attributed to the inherent complexity and volatility of Indonesia's inflation data during the study period, reflecting the challenges of accurately modeling such complex data using the SSA method alone. The SSA MAPE value will be compared with the MAPE value of the Hybrid SSA-ARIMA model.

Hybrid Singular Spectrum Analysis – ARIMA

The data utilized are the noise components obtained through the Singular Spectrum Analysis (SSA) method. The following outlines the steps for the SSA-ARIMA hybrid analysis.\

1) Stasionarity

The stationarity of the noise component data can be assessed using several methods, one of which is a time series scatter diagram shown in Figure 6.

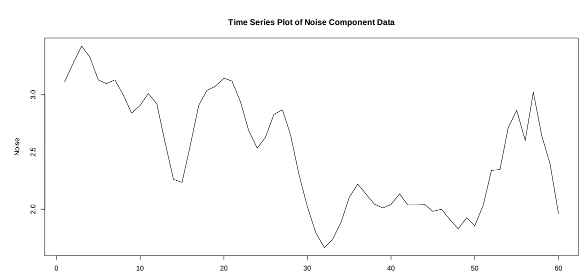
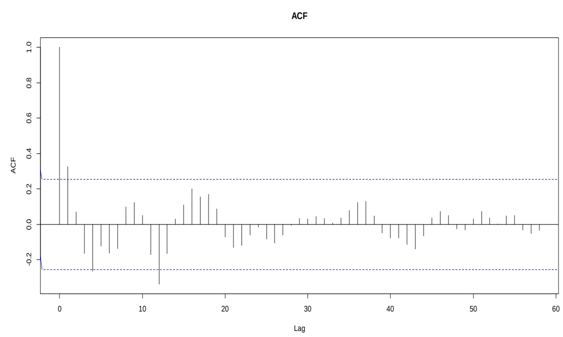


Figure 6. Time Series Plot of Noise Component Data

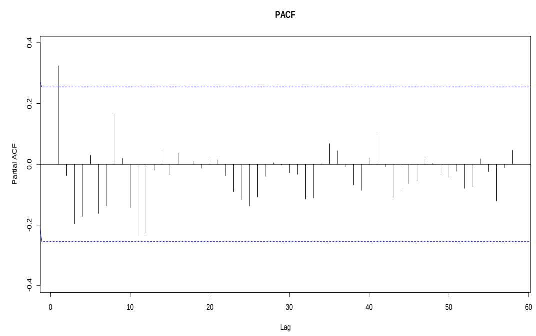
The stationarity of the data in terms of the mean was checked using the Augmented Dickey-Fuller (ADF) test. Initially, the data was non-stationary (Figure 6), as indicated by a p-value of 0.2726, which is greater than the significance level α = 0.05. To address this, differencing was applied once, which reduced the p-value to 0.01, making it less than α = 0.05. Therefore, after differencing once, the data became stationary with respect to the mean. To check for stationarity in the variance, the Box-Cox transformation was used. Data is considered stationary in variance if the λ (lambda) value is close to 1. The analysis resulted in a λ value of 1.334426, which suggests that the noise component data is stationary in terms of variance. In conclusion, the data is stationary in both the mean (after one round of differencing) and in variance (based on the Box-Cox transformation with λ = 1.334426).

2) Model Identification

Identifying the model in ARIMA starts with identifying the plot *Autocorrelation Function* (ACF) and plot *Partial Autocorrelation Function* (PACF).



(a)



(b)

Figure 7. (a) Plot ACF (b) Plot PACF

Based on Figure 7, the ACF plot shows a cut-off after lag 1 and 12, while the PACF plot shows a cut-off after lag 1. Therefore, it can be assumed that the preliminary models are ARIMA (1,1,0), ARIMA (0,1,1), ARIMA (1,1,0) $(0,0,1)^{12}$, ARIMA (0,1,1) $(0,0,1)^{12}$.

3) Parameter Estimation

After obtaining the preliminary model estimates, the next step is to examine the estimation results and parameter significance, as shown in Table 9.

Table 9. ARIMA Model Parameter Estimation

Model	Parameter	Estimation	*P-Value*	Parameter Significance
ARIMA (1,1,0)	ϕ_1	0.35769	0.004861 **	Significant
ARIMA (0,1,1)	θ_1	0.31815	0.007855 **	Significant
ARIMA	ϕ_1	0.324340	0.2405	Not Significant
(1,1,0) $(0,0,1)^{12}$	Θ_1	0.037783	0.8895	Not Significant
ARIMA	ϕ_1	0.16806	0.5613	Not Significant
(0,1,1) $(0,0,1)^{12}$	Θ_1	0.16806	0.5613	Not Significant

Based on Table 9, it can be seen that two models are significant: the ARIMA(1,1,0) model, with parameters having a p-value of 0.004861 (less than α = 0.05), and the ARIMA(0,1,1) model, with parameters having a p-value of 0.007855 (less than α = 0.05). This indicates that the parameters of both models are significant. After identifying these two best models, tests for white noise and normality of residuals were conducted.

4) Model Diagnostics

a. White Noise Test for Residual

Table 10. White Noise Test for Results

Model	Ljung-Box	*P-Value*
ARIMA (1,1,0)	0.0034784	0.953
ARIMA (0,1,1)	0.080207	0.777

Based on the results of the Ljung-Box test presented in Table 10, the p-values for both ARIMA(1,1,0) and ARIMA(0,1,1) models are greater than α (0.05). This indicates that the residuals of these models satisfy the white noise assumption.

b. Normally Distribution Test for Residual

Table 11. Shapiro Wilk Test Results

Model	Shapiro Wilk	*P-Value*
ARIMA (1,1,0)	0.98673	0.7595
ARIMA (0,1,1)	0.98445	0.6415

Based on the Shapiro-Wilk test results presented in Table 11, the p-values for all two models are greater than α (0.05). This indicates that the residuals of the ARIMA(1,1,0) and ARIMA(0,1,1) models are normally distributed.

5) Selection of The Best Model

The Table 12 represent the AIC values of the ARIMA (1,1,0) and ARIMA (0,1,1) models:

Table 12. ARIMA (1,1,0) and ARIMA (0,1,1) AIC values

Model	AIC Values
ARIMA (1,1,0)	-29.92
ARIMA (0,1,1)	-29.03

The ARIMA (1,1,0) model has the smallest AIC value, -29.92, indicating that it is the best model for forecasting

6) ARIMA and Hybrid SSA-ARIMA Forecasting

The next step is to model the forecasting for the next 12 months. The forecasting model for ARIMA(1,1,0) is as follows:

$$Z_t = Z_{t-1} + 0.36\, Z_{t-1} - 0.36\, Z_{t-2} + a_t \quad (13)$$

Table 13 presents the forecasting results for the next 12 months using the ARIMA(1,1,0) model, as shown below:

Table 13. ARIMA (1,1,0) Forecasting Results for Inflation Rate in Indonesia

Month	Forecasting
January 2023	1.80
February 2023	1.75
March 2023	1.73
April 2023	1.72
May 2023	1.72
June 2023	1.72
July 2023	1.72
August 2023	1.72
September 2023	1.72
October 2023	1.72
November 2023	1.72
December 2023	1.72

The next step is to forecast using the SSA-ARIMA hybrid model, which combines the forecasting results of the SSA and ARIMA(1,1,0) models. The following results present the forecast of the inflation rate in Indonesia for the next 12 months using the SSA-ARIMA(1,1,0) hybrid model.

Table 14. Hybrid SSA-ARIMA (1,1,0) Forecasting Results for Inflation Rate in Indonesia

Month	SSA	ARIMA (1,1,0)	SSA-ARIMA (1,1,0)
January 2023	2.99	1.80	4.79
February 2023	2.98	1.75	4.73
March 2023	3.07	1.73	4.80
April 2023	2.79	1.72	4.51
May 2023	2.63	1.72	4.35
June 2023	2.48	1.72	4.20
July 2023	1.98	1.72	3.70
August 2023	1.64	1.72	3.36
September 2023	1.13	1.72	2.85
October 2023	0.51	1.72	2.23
November 2023	-0.04	1.72	1.68
December 2023	-0.83	1.72	0.89

7) Hybrid SSA-ARIMA Forecasting Accuracy

The Table 15 presents the accuracy of SSA and Hybrid SSA-ARIMA forecasts, measured by their respective MAPE values: based on the MAPE values obtained, namely:

Table 15. Results of Forecasting Accuracy Measurements Based on MAPE

Model	MAPE
SSA	56.26797
***Hybrid* SSA - ARIMA**	18.88851

Based on Table 15, it is evident that the Hybrid SSA-ARIMA model (1,1,0) achieves a significantly lower MAPE value of 18.88851% compared to the SSA model's MAPE of 56.26797%. This substantial improvement indicates a more reliable forecast, particularly in terms of reducing error when predicting inflation rates in Indonesia. From a practical standpoint, the enhanced accuracy of the Hybrid SSA-ARIMA model can provide policymakers and economists with more precise predictions, enabling better decision-making and planning. This reduction in forecasting error could also contribute to minimizing risks associated with economic policy adjustments, ultimately leading to more stable economic outcomes. Additionally, as depicted in Figure 8, the forecasted values using the Hybrid SSA-ARIMA model more closely align with the actual values in the out-sample data. This further reinforces the model's practical applicability, as it suggests that the Hybrid SSA-ARIMA model can capture the underlying inflation trends more effectively than the SSA model alone.

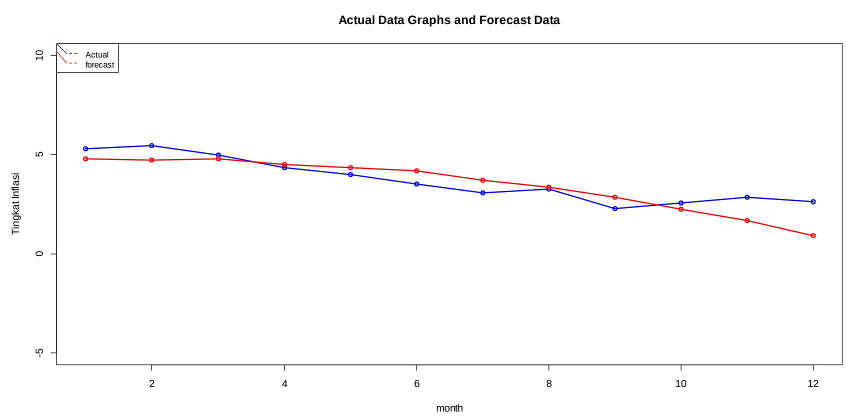


Figure 8. Actual Data Graphs and Forecast Data

CONCLUSION

Based on the analysis conducted, it can be concluded that the Hybrid SSA-ARIMA model demonstrates better forecasting accuracy in predicting the inflation rate in Indonesia compared to SSA. The forecasted inflation rates for the 12-month period from January 2023 to December 2023 indicate that the lowest inflation occurred in December at 0.89%, while the highest inflation occurred in March at 4.80%. The MAPE values further support this conclusion, with the SSA model yielding a MAPE of 56.26797%, while the Hybrid SSA-ARIMA model achieved a significantly lower MAPE of 18.88851%. It is evident that the forecasted values generated by the Hybrid SSA-ARIMA model closely follow the actual inflation values, especially in the out-sample data, indicating a superior fit and more accurate forecasting performance. In contrast, the SSA model shows greater deviation from the actual values, which is reflected in its higher MAPE. The key finding from this study is that incorporating the ARIMA model within the SSA framework significantly enhances forecasting accuracy. This improvement has practical implications, particularly for policymakers and economists who rely on accurate inflation forecasts to make informed decisions. By utilizing the Hybrid SSA-ARIMA model, stakeholders can achieve more reliable predictions, leading to better planning, reduced risks, and ultimately more effective economic policy interventions.

REFERENCE

Ahadi, G. D., Nur, N., & Ersela, L. (2023). The Simulation Study of Normality Test Using Kolmogorov-Smirnov, Anderson-Darling, and Shapiro-Wilk. *Eigen Mathematics Journal*, *6*(1).

Aktivani, S. (2020). Uji Stasioneritas Data Inflasi Kota Padang Periode 2014-2019. *Statistika*, *20*(2), 83–90.

Arumsari, M., Wahyuningsih, S., & Siringoringo, M. (2021). Inflation Forecasting for East Kalimantan Province Using Hybrid Singular Spectrum Analysis- Autoregressive Integrated Moving Average Model. *Jurnal Matematika, Statistika Dan Komputasi*, *18*(1), 78–92. https://doi.org/10.20956/j.v18i1.14284

Aswi, & Sukarna. (2006). *Analisis Deret Waktu Teori & Aplikasi*. Andira Publisher, Makassar.

BI. (2020). *Inflasi*. Bank Indonesia.

Choirunisa, P., & Kariyam. (2019). Perbandingan Metode Triple Exponential Smoothing Dan Metode Seasonal Arima Untuk Peramalan Inflasi Di Kota Tanjung Pandan. *Prosiding Sendika*, *5*(2), 76. http://eproceedings.umpwr.ac.id/index.php/sendika/article/view/844

Darmawan, G., Rosadi, D., & Ruchjana, B. N. (2022). Hybrid Model of Singular Spectrum Analysis and ARIMA for Seasonal Time Series Data. *CAUCHY: Jurnal Matematika Murni Dan Aplikasi*, *7*(2), 302–315. https://doi.org/10.18860/ca.v7i2.14136

Ete, A. A., Suhartno, & Atok, R. M. (2020). SSA and ARIMA for Forecasting Number of Foreign Visitor Arrivals to Indonesia. *INFERENSI*, *3(1)*, 55–63.

Fajar, M. (2019). Perbandingan Kinerja Peramalan Pertumbuhan Ekonomi Indonesia dengan ARIMA , SSA dan Hybrid ARIMA-SSA. *ResearchGate*, 1–7. https://doi.org/10.13140/RG.2.2.28620.26244

Fajar, M., & Rachmad, S. H. (2018). Inflation Forecasting By Hybrid Singular Spectrum Analysis-Multilayer Perceptrons Neural Network Method. *ResearchGate*, 1–10. https://doi.org/10.13140/RG.2.2.27432.47368

Golyandina, N., Korobeynikov, A., & Zhigljavsky, A. (2018). *Singular Spectrum Analysis*. https://doi.org/10.1007/978-3-030-26050-7_294-1

Golyandina, N., Nekrutkin, V., & Zhigljavsky, A. (2001). *Analysis of Time Series Structure : SSA and Related Techniques*. Chapman & Hall/CRC.

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

Idrus, R. A., Ruliana, & Aswi. (2022). Penerapan Metode Singular Spectrum Analysis dalam Peramalan Jumlah Produksi Beras di Kabupaten Gowa. *VARIANSI: Journal of Statistics and Its Application on Teaching and Research*, *4*(2), 49–58. https://doi.org/10.35580/variansiunm40

Irmawati, D. R., Atok, R. M., & Suhartono. (2018). Hybrid singular spectrum analysis-ARIMA modelling for direct and indirect forecasting of Farmer's Term of Trade in East Java. *2018 International Conference on Information and Communications Technology, ICOIACT 2018*, *2018-Janua*, 889–894. https://doi.org/10.1109/ICOIACT.2018.8350823

Khaeri, H., Yulian, E., & Darmawan, G. (2018). Penerapan Metode Singular Spectrum Analysis (Ssa) Pada Peramalan Jumlah Penumpang Kereta Api Di Indonesia Tahun 2017. *Euclid*, *5*(1), 8. https://doi.org/10.33603/e.v5i1.496

Makridakis S, Wheelwright SC, H. R. (1997). 1 / the Forecasting Perspective. *Forecasting Methods and Applications*, 1–632.

Rizki, M. I., & Taqiyyuddin, T. A. (2021). Penerapan Model SARIMA untuk Memprediksi Tingkat Inflasi di Indonesia. *Jurnal Sains Matematika Dan Statistika*, *7*(2). https://doi.org/10.24014/jsms.v7i2.13168

Satriani, Nursalam, & Ibnas, R. (2020). Peramalan Indeks Harga Konsumen (IHK) di Sulawesi Selatan dengan Menggunakan Metode Singular Spektrum Analysis (SSA). *Jurnal MSA (Matematika Dan Statistika Serta Aplikasinya)*, *8*(1), 82. https://doi.org/10.24252/msa.v8i1.17441

Siringoringo, M., Wahyuningsih, S., Purnamasari, I., & Arumsari, M. (2022). Peramalan Jumlah Produksi Kelapa Sawit Provinsi Kalimantan Timur Menggunakan Metode Singular Spectrum Analysis. *VARIANSI: Journal of Statistics and Its Application on Teaching and Research*, *4*(3), 162–172. https://doi.org/10.35580/variansiunm46

Siti, M., Basari, N., & Achmad, A. I. (2019). Metode Singular Spectrum Analysis untuk Meramalkan Indeks Harga Konsumen Indonesia Tahun 2019. *Prosiding Statistika*, 484–491.

Wulandari, S. S., & Yurinanda, S. (2021). *Penerapan Metode ARIMA Dalam Memprediksi Fluktuasi Harga Saham PT Bank Central Asia Tbk*. *11*, 53–68.

Yunita, T. (2019). Peramalan Jumlah Penggunaan Kuota Internet Menggunakan Metode Autoregressive Integrated Moving Average (ARIMA). *JOMTA Journal of Mathematics: Theory and Applications*, *1*(2), 16–22.