BAREKENG: Journal of Mathematics and Its Applications
June 2025 Volume 19 Issue 2 Page 1329-1340
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol19iss2pp1329-1340

COMPARISON FORECASTING BETWEEN SINGULAR SPECTRUM ANALYSIS AND LOCAL LINEAR METHOD FOR SHIP ACCIDENT SEARCH AND RESCUE OPERATIONS IN INDONESIA

Rien Recylia[1], Toha Saifudin[2*], Nur Chamidah[3], M. Fariz Fadillah Mardianto[4]

[1,2,3,4]*Departement of Mathematics, Faculty of Science and Technology, Universitas Airlangga*
Jln. Dr. Ir. H. Soekarno, Mulyorejo, Surabaya, 60115, Indonesia

*Corresponding author's e-mail: * **tohasaifudin@fst.unair.ac.id***

ABSTRACT

Article History:

Received: 22nd October 2024
Revised: 22nd January 2025
Accepted: 28th February 2025
Published: 1st April 2025

Keywords:

Forecasting;
Local linear;
Search and rescue;
SSA.

As a maritime country strategically located along the world's leading transportation routes, Indonesia often faces increased ship accidents. Based on the Basarnas Statistics Book, ship accidents handled by Basarnas from 2021 to 2023 increased by 3%. This condition requires an effective forecasting method to carry out SAR operations to predict ship accidents in the Indonesian region in the future and assess the readiness and needs of Basarnas resources. This study compares the forecasting results obtained using the Singular Spectrum Analysis (SSA) and the Local Linear methods. Both methods do not require parametric assumptions. The data used in this study are divided into training data and test data. This data is secondary data obtained from the Basarnas Statistics Book. The training data in this study is the number of SAR operations from January 2021 to December 2022, while the testing data is from January 2023 to December 2023. From the analysis results, it is known that the method with the smallest MAPE is the Local Linear method with a MAPE of test data of 18.67% (good forecasting category), optimal bandwidth (h) = 4.299, and CV (h) = 231.39 where bandwidth is used to determine the level of smoothness of the estimate, while the CV (h) value is used to select the optimal bandwidth that minimizes the estimation error. At the same time, the SSA method has a MAPE of 40.27% (fair forecasting category). This shows that the Local Linear method provides a more accurate forecast of the number of SAR operations related to ship accidents in Indonesia. This research contributes to the SDGs to make Basarnas an effective and accountable institution and improve the planning and decision-making process in SAR operations through accurate forecasting research is relevant to accurate forecasting.

How to cite this article:

R. Recylia, T. Saifudin, N. Chamidah and M. F. F. Mardianto., "COMPARISON FORECASTING BETWEEN SINGULAR SPECTRUM ANALYSIS AND LOCAL LINEAR METHOD FOR SHIP ACCIDENT SEARCH AND RESCUE OPERATIONS IN INDONESIA," *BAREKENG: J. Math. & App.*, vol. 19, iss. 2, pp. 1329-1340, June, 2025.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

Research Article · **Open Access**

1. INTRODUCTION

As the largest archipelagic country in the world, Indonesia has a high rate of ship accidents due to its dense sea traffic. The Central Statistics Agency (BPS) recorded that the number of islands owned by Indonesia reached 17,001 (seventeen thousand one) islands spread across 34 (thirty-four) provinces in 2022 **[1]**. As mandated in the 1945 Constitution of the Republic of Indonesia, the Unitary State of the Republic of Indonesia is responsible for protecting all Indonesian people and all of Indonesia's territory to protect their souls and bodies. One form of implementing state protection for its citizens is by carrying out search and rescue activities or what the public knows as Search and Rescue (SAR).

According to Law Number 29 of 2014 concerning Search and Rescue, Search and Rescue is all efforts and activities to search for, help, rescue, and evacuate humans facing emergencies and/or dangerous situations in accidents, disasters, or conditions that endanger humans. Basarnas can implement search and rescue operations for ship and aircraft accidents, accidents with special handling, disasters at the emergency response stage, and/or conditions that endanger humans. Based on the Basarnas statistics book, ship accidents handled by Basarnas increased from 811 incidents in 2021 to 846 incidents in 2023. In addition to the limited human resources owned by Basarnas and other operational challenges, forecasting the number of ship accident search and rescue operations is needed to assist planning and decision-making in implementing search and rescue operations in the future period.

Data on the number of ship accident search and rescue operations is time series data and requires an analysis method to produce accurate forecasts. However, the method selection depends on several aspects, namely time, data patterns, and the desired level of forecast accuracy. One method that is often used to analyze time series data is the Singular Spectrum Analysis (SSA) method. In their research, Hassani and Zigljavsky **[2]** once compared the SSA, SARIMA, ARAR, and Seasonal Holt-Winter methods by predicting accidental deaths in the USA in 1973 and showed that the SSA method was superior to other methods. Further research conducted by **[3]**, showed that the SSA method is a forecasting method with a non-parametric approach, which means that this method is flexible because it is free from its parametric assumptions.

Furthermore, the method that is often used in forecasting is the local polynomial method. Local polynomials have several advantages, including reducing asymptotic bias and producing reasonable estimates **[4]**. In local polynomial regression, the smoothness level of the function is determined by its bandwidth. The optimal bandwidth can be determined using the GCV (Generalized cross-validation) method **[5]**. In a study conducted by **[6]**, showed that modeling using the local polynomial method produced an optimum model and a MAPE value of 7.47%.

Based on previous research, Jing Wang's research **[2]** which combines singular spectrum analysis (SSA) time series techniques with multivariate forecasting parameterized autoregressive integrated moving average (ARIMA). The study results showed that SSA produces more accurate predictions than other methods. However, this study focuses on forecasting emergency ambulance demand (EAD) time series prediction in its non-stationary nature. His research predicts deaths in a general context, namely better EAD forecasting. Meanwhile, this study focuses on forecasting search and rescue (SAR) operations for ship accidents in Indonesia, which are more specific and directly related to the maritime context. The data used has a more varied time pattern and unique Indonesian maritime traffic conditions.

In addition, previous studies are theoretical and aim to prove the superiority of SSA compared to other methods in general accident forecasting. At the same time, this study is oriented towards practical solutions to improve SAR operations by Basarnas in Indonesia. This focus aims to help make better decisions to save lives in maritime accident situations. In addition, it was also found that the Local Linear method has better results. Therefore, the author is interested in conducting a forecasting study on the number of search and rescue operations in ship accidents in Indonesia based on the Singular Spectrum Analysis (SSA) and the Local Linear Method. It is hoped that Basarnas can improve planning and decision-making in search and rescue operations, add literature in the field of forecasting, and increase public awareness of search and rescue operation trends in Indonesia.

2. RESEARCH METHODS

The methods used in this study are Singular Spectrum Analysis (SSA) and the local linear approach. The data used is secondary data obtained from the Basarnas Statistics Book. The training data is the number of search and rescue operations from January 2021 to December 2022 (24 observations), while the test data is the number of search and rescue operations from January 2023 to December 2023 (12 observations). Before the analysis was carried out, the data had been prepared to ensure that there were no missing values, outliers, or data inconsistencies. In addition, only one scale was used in this data, namely the continuous scale. The SSA and Local Linear methods are suitable for predicting search and rescue operation data compared to other methods because this forecast uses time as a predictor variable, and the data has fluctuating characteristics.

2.1 Singular Spectrum Analysis (SSA) Method

Singular Spectrum Analysis (SSA) is a time series method that is quite powerful for forecasting. Based on the general structure of the algorithm underlying SSA, there are two essential things and vital parameters, namely the window length L and the number of eigentriples in the SSA procedure. The right choice of L and eigentriples can produce effective time series decomposition **[8]**. There are 2 (two) stages in the Singular Spectrum Analysis (SSA) method, namely:

a. Decomposition is the initial stage in the SSA method, which involves trend, seasonal, cyclical, and error components in a forecast **[9]**. Decomposition consists of two stages: Embedding and Singular Value Decomposition (SVD).

   i. Embedding is a stage where one-dimensional data is converted into multidimensional data and produces a Trajectory Matrix. Suppose the time series data of length N, without missing data is expressed by $X_N = \{x_1, x_2, \cdots, x_N\}$, the data is transformed into a matrix of size $L \times K$. With L being the windows length where $1 < L < N$. Based on the research of Marques et al. **[10]**, the determination of the value of L is chosen through trial and error process and $K = N - L + 1$. In the embedding stage, the time series X_N will be mapped into a row of lag vectors of size L, with the i-th lag vector expressed as follows:

$$X_i = (x_1, x_2, \cdots, x_{i+L-1})^T \tag{1}$$

   Or it can be expressed as the following trajectory matrix:

$$X = [X_1, X_2, \dots, X_K] = \begin{bmatrix} x_1 & x_2 & \cdots & x_K \\ x_2 & x_3 & \cdots & x_{K+1} \\ \vdots & \cdots & \ddots & \vdots \\ x_L & x_{L+1} & \cdots & x_N \end{bmatrix} \tag{2}$$

   ii. Singular Value Decomposition (SVD) is the stage of separating components that have different characteristics based on eigentriple. The second step in decomposition is to make Singular Value Decomposition (SVD) of the Trajectory Matrix. Consider **[11]**:

$$S = XX^T \tag{3}$$

   If $\lambda_1 \geq \cdots \geq \lambda_L \geq 0$ is the eigenvalue of matrix S, $d = rank(X) = max\{i, \lambda_i > 0\}$, $U_1, \cdots, U_d$ is the eigenvector of each eigenvalue, and $V_i = \frac{X^T U_i}{\sqrt{\lambda_i}}$ with $i = 1, \cdots, d$ is the factor vector. $X_i = \sqrt{\lambda_i} U_i V_i^T$, then the SVD of matrix X can be written as follows:

$$X = X_1 + X_2 + \cdots + X_d \tag{4}$$

   $(\sqrt{\lambda_i}, U_i, V_i)$ is called the i-th eigentriple, which contains the singular value $\sigma_i = \sqrt{\lambda_i}$, the left singular vector U_i, and the right singular vector V_i of matrix X.

b. Reconstruction is the stage where data will be reconstructed into new time series data based on the values obtained in the previous stage through the grouping and diagonal averaging processes **[2]**.

   i. Grouping, the results of the decomposition of the $L \times K$ path matrix will be grouped by separating the SVD additive components into several subgroups, namely trend, seasonal, and noise [12]. For example, from d matrix indices grouped into m index groups, namely $I = I_1, I_2, \dots, I_m$, then the decomposition of the X matrix is as follows [11]:

$$X = X_{I1} + X_{I2} + \cdots + X_{Im} \tag{5}$$

with $X_I = \sum_{i \in I} X_i$.

ii. Diagonal Averaging, a transformation will be carried out from the grouping results into a new series with length N. At this stage, the X_{Ik} matrix obtained from the grouping stage will be rearranged into a new series with length N. For example, the reconstruction of the X_{Ik} matrix, namely [11]:

$$\widetilde{X_k} = \left\{\tilde{x}_1^{(k)}, \tilde{x}_2^{(k)}, \dots, \tilde{x}_N^{(k)}\right\} \tag{6}$$

To obtain the values of $\tilde{x}_p^{(k)}$ with $p = 1, \ldots$, the X_{Ik} matrix can be transformed into the Hankel matrix form. The elements of the i-th row and j-th column of the matrix $\tilde{x}_{ij}^{(k)}$, with $s = i + j$ and $N = L + K - 1$ are

$$\tilde{x}_{ij}^{(k)} = \begin{cases} \frac{1}{s-1}\sum_{l=1}^{s-1} x_{l,s-1} & for\ 2 \le s \le L-1 \\ \frac{1}{L^*}\sum_{l=1}^{L} x_{l,s-1} & for\ L \le s \le K+1 \\ \frac{1}{K+L-s+1}\sum_{l=s-K}^{L} x_{l,s-1} & for\ K+2 \le s \le K+N \end{cases} \tag{7}$$

Furthermore, from the elements $\widetilde{X_k} = \left\{\tilde{x}_1^{(k)}, \tilde{x}_2^{(k)}, \dots, \tilde{x}_N^{(k)}\right\}$ the diagonal average value of the Hankel matrix is obtained. It can be shown that the initial time series data $X_N = \{x_1, x_2, \cdots, x_N\}$ has been decomposed into the sum of m time series reconstructions $\tilde{X}^{(1)}, \tilde{X}^{(2)}, \dots, \tilde{X}^{(m)}$, with

$$x_p = \sum_{k=1}^{m} \tilde{x}_p^{(k)} \qquad for\ (p = 1,2, \dots, N) \tag{8}$$

2.2 W-Correlation

After the SSA stage is completed, it is continued with W-correlation. According to Golyandina et al. [11], SSA decomposition is said to be successful if the series of decomposition results do not correlate with each other. W-correlation measures the correlation between two components of the time series of SSA decomposition results. For example, the w-correlation can be measured $X_N^{(1)}$ and $X_N^{(2)}$:

$$\rho_{12}^{(w)} = \frac{\left(X_N^{(1)}, X_N^{(2)}\right)_w}{\left\|X_N^{(1)}\right\|_w \left\|X_N^{(2)}\right\|_w} \tag{9}$$

with

$$\left(X_N^{(1)}, X_N^{(2)}\right)_w = \sum_{i=1}^{N} w_i x_i^{(1)} x_i^{(2)} \tag{10}$$

w_i is the number of elements x_i that appear in the path matrix and

$$\left\| X_N^{(1)} \right\|_w = \sqrt{\left(X_N^{(i)}, X_N^{(i)} \right)_w} \tag{11}$$

with $i = 1, 2$.

2.3 Forecasting

The forecasting used in this study is SSA recurrent. The model is built using the linear recurrent formula (LRF). Suppose $u_i^{\ddot{v}}$ is the first $L - 1$ component vector of the eigenvector u_i and π_i is the last component of $u_i (i = 1, 2, \dots, l)$ with $v^2 = \sum_{i=1}^{l} \pi_i^2$, then R can be defined as a vector consisting of the LRF coefficients of a component [13]:

$$R = (a_{L-1}, \dots, a_1) = \frac{1}{1 - v^2} \sum_{i=1}^{l} \pi_i u_i^{\ddot{v}} \tag{12}$$

The time series used in this forecasting is the reconstructed series obtained from the diagonal averaging stage. The forecasting results are obtained through the following equation:

$$g_i = \begin{cases} \tilde{f}_i & , i = 1, \dots, N \\ \sum_{j=1}^{L-1} a_j g_{i-j} & , i = N + 1, \dots, N + M \end{cases} \tag{13}$$

where $\tilde{f}_i$ is the result of the reconstruction of the time series after diagonal averaging and $g_{N+1}, g_{N+2}, \dots, g_{N+M}$ are the forecasting results using the SSA method with a recursive formula, where a_j is the coefficient of LRF [16].

2.4 Local Linear Estimator

The local linear method is one of several nonparametric regression methods that can be used to estimate a regression curve. The advantage of the local linear method is that it can reduce asymptotic bias and produce good estimates [4]. Local linear estimation can use WLS (Weighted Least Square) by minimizing it [13]. In the local linear method, the smoothness of the function is determined by its bandwidth.

a. Kernel Weighting Function is one of the local linear estimation methods using WLS (Weighted Least Square) so that weighting is required with the Kernel function. The Kernel function K with bandwidth h is defined as follows:

$$K_h(x) = \frac{1}{h} K\left(\frac{x}{h}\right); \ -\infty < x < \infty \text{ and } h > 0 \tag{14}$$

Here are the properties of the kernel function:

i. $K(x) \geq 0$, for all values of x;

ii. $\int_{-\infty}^{\infty} K(x) \ dx = 1$

iii. $\int_{-\infty}^{\infty} xK(x) \ dx = 0$

iv. $\int_{-\infty}^{\infty} x^2 K(x) \ dx = \sigma^2 > 0$

Meanwhile, according to Hardle, there are several types of kernel functions in Table 1.

Table 1. Types of Kernel Functions

No	Kernel Types	Kernel Functions
1	Uniform Kernel	$K(x) = \frac{1}{2}; \ I(\lvert x \rvert < 1)$
2	Triangular Kernel	$K(x) = (1 - \lvert x \rvert); \ I(\lvert x \rvert < 1)$

No	Kernel Types	Kernel Functions
3	Epanechnikov Kernel	$K(x) = \frac{3}{4}(1-x^2);\ I(\lvert x\rvert < 1)$
4	Quadratic Kernel	$K(x) = \frac{15}{16}(1-x^2)^2;\ I(\lvert x\rvert < 1)$
5	Triweight Kernel	$K(x) = \frac{35}{32}(1-x^2)^3;\ I(\lvert x\rvert < 1)$
6	Cosine Kernel	$K(x) = \frac{\pi}{4}\cos\left(\frac{\pi}{2}x\right);\ I(\lvert x\rvert < 1)$
7	Gaussian Kernel	$K(x) = \frac{1}{\sqrt{2\pi}}\exp\left(\frac{1}{2}(-x^2)\right)$

b. Local Linear

The nonparametric regression model can be expressed as follows:

$$y_i = m(t_i) + \varepsilon_i, \qquad i = 1,2,\dots,n \tag{15}$$

where y is the response variable, $m(.)$ is the unknown regression function, and ε_i is the normally distributed error. To estimate the function $1m(t_i)$, local linear estimation can be used. The function $m(t_i)$, can be approximated by the 1st order Taylor expansion as follows [14]:

$$m(t_i) \approx \sum_{k=0}^{1} \frac{m^{(k)}(t_0)}{k!}(t-t_0)^k = \sum_{k=0}^{1} \beta_k(t_0)(t-t_0)^k \tag{16}$$

where $m^{(k)}(t_0)$ is the value of the k-th derivative function of $m(t_0), t \in (t_0 - h, t_0 + h)$ and $\frac{m^{(k)}(t_0)}{k!} = \beta_k(t_0)$.

Equation (16) can be written in matrix form as follows:

$$m(t_i) = t_{t_0}\beta(t_0) \tag{17}$$

$t_{t_0} = [1 \quad (t - t_0)], t \in (t_0 - h, t_0 + h))$ and $\beta(t_0) = [\beta_0(t_0) \quad \beta_1(t_0)]^T$

From Equation (15) and Equation (17) can be written as follows:

$$y_i = t_{t_0}\beta(t_0) + \varepsilon_i \tag{18}$$

The estimation of in $\beta(t_0)$ Equation (18) based on the local linear approach is obtained by taking n paired data samples $\{t_i, y_i\}_{i=1}^{n}$, and Equation (18) can be written as follows:

$$\left.\begin{aligned} y_1 &= \beta_0(t_0) + \beta_1(t_0)(t_1 - t_0) + \varepsilon_1 \\ y_2 &= \beta_0(t_0) + \beta_1(t_0)(t_2 - t_0) + \varepsilon_2 \\ &\vdots \\ y_n &= \beta_0(t_0) + \beta_1(t_0)(t_n - t_0) + \varepsilon_n \end{aligned}\right\} \tag{19}$$

Equation (19) can be expressed in the following matrix notation:

$$y^* = t_{t_0}^*\beta^*(t_0) + \varepsilon^* \tag{20}$$

where

$$t_{t_0}^* = \begin{bmatrix} 1 & (t_1 - t_0) \\ \vdots & \vdots \\ 1 & (t_n - t_0) \end{bmatrix}, y^* = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, \varepsilon^* = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix} \quad (21)$$

The estimation of $\beta^*(t_0)$ is obtained by performing least squares optimization or WLS (Weighted Least Square), namely:

$$Min\ Q(t_0) = Min\left(y^* - t_{t_0}^* \beta^*(t_0)\right)^T K_h(t_0)\left(y^* - t_{t_0}^* \beta^*(t_0)\right) \quad (22)$$

where,

$$K_h(t_0) = \begin{bmatrix} K_h(t_1 - t_0) & 0 & \cdots & 0 \\ 0 & K_h(t_2 - t_0) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & K_h(t_n - t_0) \end{bmatrix} \quad (23)$$

From **Equation (23)**, $K_h(.)$ is a kernel function with bandwidth h which can be defined as follows **[14]**: $K_h(t) = \frac{1}{h}K;\ -\infty < t < \infty$ and $h > 0$. The Gaussian Kernel can be used as follows: $K(t) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{1}{2}(-t^2)\right)$. Furthermore, from **Equation (22)**, the value of $\beta(t_0)$ can be derived by equating it to 0 and obtained:

$$\hat{\beta}(t_0) = \left(T_{t_0}{}^T K_h(t_0) T_{t_0}\right)^{-1} T_{t_0}{}^T K_h(t_0) y \quad (24)$$

From **Equations (17)** and **Equation (23)**, the local linear estimator of $\widehat{m}(t_i)$ namely:

$$\widehat{m}(t_i) = t_{t_0}\left(T_{t_0}{}^T K_h(t_0) T_{t_0}\right)^{-1} T_{t_0}{}^T K_h(t_0) y \quad (25)$$

c. Cross Validation (CV)

One way to determine the optimal bandwidth is by using the CV method. The CV function given is:

$$CV(h) = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \widehat{m}_{h,-i}(t_i)\right)^2 \quad (26)$$

2.5 Forecast Accuracy with Mean Absolute Percentage Error (MAPE)

If X_t is the actual data for the t-th period F_t and is the forecast result for the same period, then MAPE can be calculated using the following formula **[17]**:

$$MAPE = \frac{\sum_{t=1}^{N}\left|\left(\frac{X_t - F_t}{X_t}\right) \times 100\%\right|}{N} \quad (27)$$

with

X_t : t-th time series data

F_t : t-th time forecast result data

N : amount of data

The interpretation of the MAPE value can be seen in Table 2 below [9].

Table 2. Interpretation of MAPE Values

MAPE (%)	Interpretation
$MAPE < 10$	High accuracy forecast
$10 \leq MAPE < 20$	Good forecast
$20 \leq MAPE < 50$	Decent forecast
$MAPE \geq 50$	Inaccurate forecast

2.6 Procedure Analysis

a. Several stages were carried out with the SSA method in this study, namely:
   i. Inputting data into the R software.
   ii. Applying the SSA model to training data with 2 (two) stages: decomposition and reconstruction.
   iii. Modelling based on Linear Recurrent Formula (LRF) with SSA method.
   iv. Forecasting with the SSA model that has been built using testing data.
   v. Calculating the MAPE value from the testing data.

b. Several stages were carried out with the local linear method in this study, namely:
   i. Selecting the optimal bandwidth with a specific kernel weighting function for the local linear estimator from the training data using CV in Equation (24).
   ii. Estimating parameters with WLS for testing data.
   iii. Modelling based on WLS parameter estimates with the local linear method.
   iv. Calculating the MAPE value from the testing data.

c. In last stage, the MAPE value from the testing data based on the SSA method and the Local Linear method will be compared.

3. RESULTS AND DISCUSSION

The parameters obtained from SSA and the Linear Local Method will then be used to determine the predicted value for the training data. The following is a plot of the actual value versus the expected value for the training data using the SSA method, which can be seen in Figure 1.

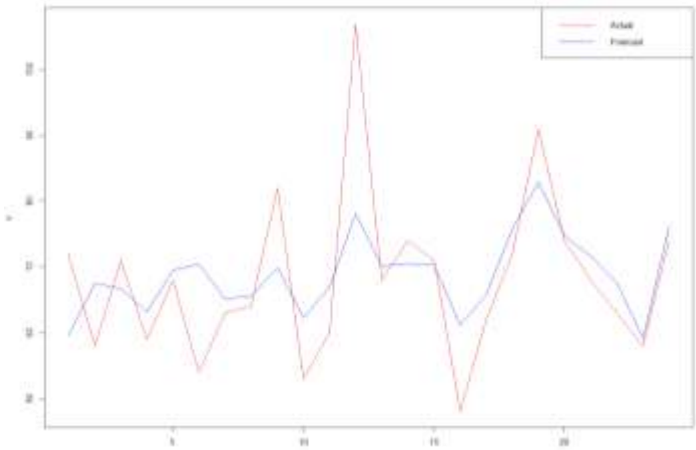

Figure 1. Plot Actual Value vs Predicted Value base on SSA Method

Figure 1 shows a fairly high error between the actual and the predicted data in some parts but not in all parts. Through the stages in the SSA method, the Windows length (L) obtained is 3 through the stages in the SSA method. Furthermore, the modeling will be built through the Linear Recurrent Formula (LRF) with the following equation model:

$$g_i = \sum_{j=1}^{2} a_j g_{i-j}$$

The following is an example of a model for $t = 25$:

$$g_{25} = -1.11(g_{24}) + 2.13(g_{23}) = 41.21$$

Next, in the same way, for the following periods, the predicted values are obtained, which are presented in **Table 3**, namely:

Table 4. MAPE Calculation Results for SSA Method

t	Current	Forecast	MAPE
25	69	41.21	
26	76	80.87	
27	62	62.42	
28	68	92.92	
29	84	56.31	40.27
30	70	51.39	
31	94	101.06	
32	93	44.52	
33	69	96.77	
34	55	121.35	
35	41	85.8	
36	65	71.55	

The SSA method has a more considerable MAPE value because the predictor variable used is only one variable, namely the time variable. This method may be better if, in the analysis, several predictor variables are sufficient to build the model.

The following plot is an estimate with training data based on the linear local method in **Figure 2**.

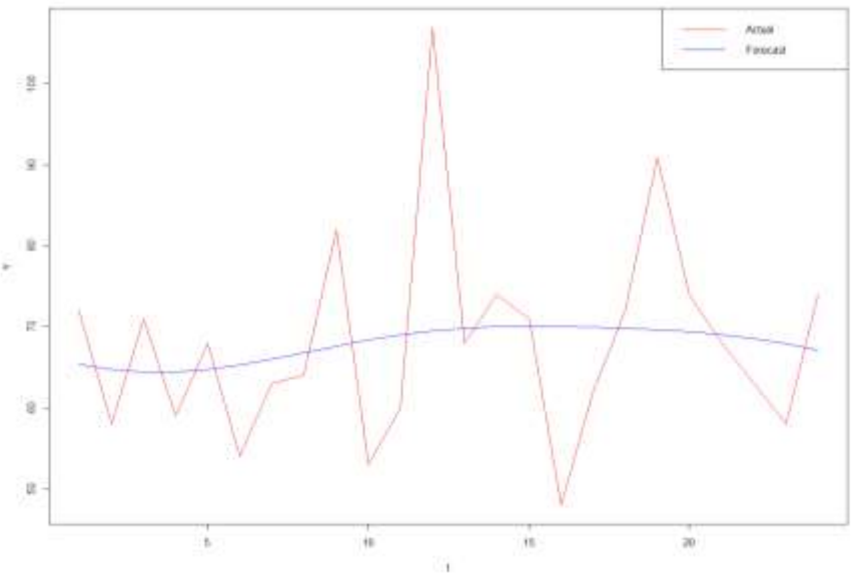

Figure 2. Plot Actual Value vs Predicted Value base on Local Linear Method

Based on **Figure 2**, compared with the forecasting results by SSA, the forecasting results using the local liner method can be seen that the forecasting model can resemble the trend of the original data plot. Meanwhile, for data forecasting using the local linear method with the Gaussian Kernel function, the tuning results for bandwidth obtained can be seen in $hCV(h)$ **Table 4**:

Table 5. Bandwidth and CV Values for Local Linear Method

Bandwidth (h)	CV
4	231.9547
4.15	231.5223
4.2	231.4485
4.25	231.406
4.29	231.393
4.291	231.393
4.292	231.393
4.293	231.3929
4.294	231.3929
4.295	231.3928
4.296	231.3928
4.297	231.3928
4.298	231.3927
4.29	231.3927
4.3	231.3927
4.35	231.4066
4.4	231.4457
4.45	231.508
4.5	231.5919
4.55	231.6955
4.6	231.8174
4.65	231.9559
4.7	232.1096

Based on Table 4, Figure 3 shows a Bandwidth versus CV plot so that it can be seen that the value given is the minimum CV. It can be seen that the plot results show that the maximum CV value occurs at the third bandwidth value of 4.299.

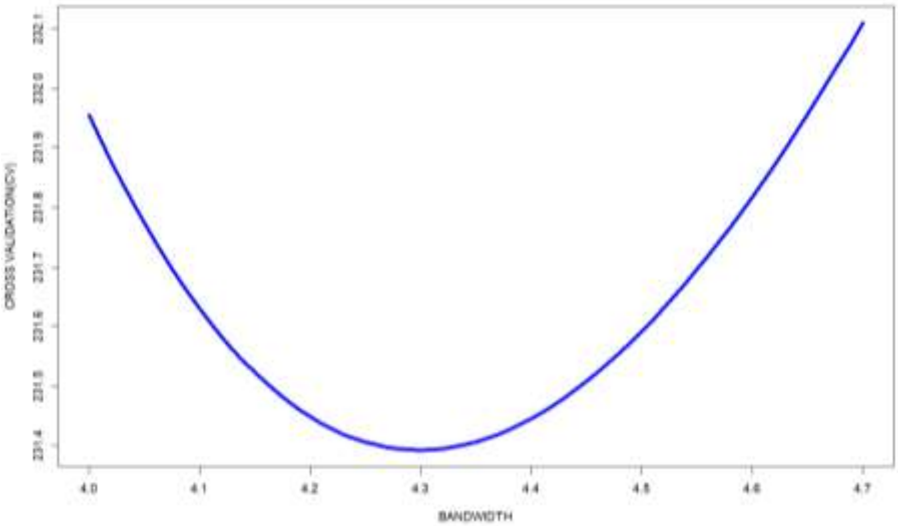

Figure 3. Plot Bandwidth versus CV

Based on **Table 4** and **Figure 3**, the optimal bandwidth is obtained with. Next, parameter estimation is carried out using a polynomial estimator and the values of and obtained for the testing data will be made into a model equation as follows: $h = 4.299$ $CV(h) = 231.39$. By obtaining the minimum CV, it can be ensured that the model using the bandwidth that has been selected based on the minimum CV has a minimal level of forecast error.

$$\hat{y} = \beta_0 + \beta_1(t - t_0)$$

$$\hat{y}_{25} = \beta_0 + \beta_1(t - t_0) = 70.22 + 0.84(t - t_0), \qquad t \in (20.7;\ 29.299)$$

$$\hat{y}_{26} = \beta_0 + \beta_1(t - t_0) = 7.61 + 0.6(t - t_0), \qquad t \in (21.7;\ 30.299)$$

$$\vdots$$

$$\hat{y}_{36} = \beta_0 + \beta_1(t - t_0) = 56.17 - 3.54(t - t_0), \qquad t \in (31.7;\ 40.299)$$

From the equation above, the forecasting results for the testing data are as follows at **Table 6**.

Table 7. Estimation Results for Testing Data Based on the Local Polynomial Method

t	Current	Forecast	APE (%) = $\left\|\frac{(\text{Actual}-\text{Forecast})}{\text{Actual}}\right\|$	MAPE
25	69	71.09	3	
26	76	71.33	6	
27	62	71.57	15	
28	68	71.82	5	
29	84	72.06	14	
30	70	72.30	3	
31	94	72.54	23	18.67
32	93	72.78	22	
33	69	73.02	6	
34	55	73.26	33	
35	41	73.50	79	
36	65	73.74	13	

Comparison of forecast results of the number of SAR operations for ship accidents in Indonesia. From the analysis of the results above, a comparison of forecasting results can be seen in the following **Table 6**.

Table 8. Comparison of MAPE Values Based on SSA and Local Linear Methods

Method	MAPE
Singular Spectrum Analysis (SSA)	40.27%
Local Linear	18.67%

From **Table 6**, forecasting the number of search and rescue operations for ship accidents in Indonesia using the local linear method has the smallest MAPE value of 18.67% (included in the good forecasting category) with optimal bandwidth $(h) = 4.299$ with $CV = 231.39$ compared to the SSA method of 40.27% (included in the decent forecasting category). Thus, we concluded that this study is based on previous studies but developed by focusing on some maritime issues and combining SSA with the local linear method to improve accuracy. In addition, this study has a more practical operational focus, which aims to support SAR operations in Indonesia, while previous studies were more theoretical and general in their application. In addition, we concluded that the local linear method is better for forecasting than the SSA method.

The results of this study allow SAR in Indonesia to mitigate potential disasters or search and rescue in the future. It is hoped that this study can reduce the accident rate. The results of this study allow SAR in Indonesia to mitigate potential disasters or search and rescue in the future. It is hoped that this study can reduce the accident rate. This study also has the potential to be developed, for example, by adding predictor variables or considering several other variables to predict in the hope of producing more accurate forecasts.

4. CONCLUSIONS

The overall study concluded that although the SSA method is flexible for non-parametric data, the results are less accurate in estimating the number of ship accident search and rescue operations in Indonesia, with a high MAPE value. This condition is likely due to the lack of predictor variables, which makes the model less able to represent the data pattern. On the other hand, the local linear method provides better accuracy than SSA, with a MAPE value of 18.67%, which is included in the good forecasting category. We concluded from the low MAPE value that the model based on the Local Linear method successfully represents the data pattern and has good predictive ability with a minimal error rate. Therefore, the local linear method is more suitable for estimating the number of ship accident search and rescue operations in Indonesia.

REFERENCES

[1] S. Sadya, "INDONESIA HAS 17,001 ISLANDS IN 2022, WHERE ARE THE MOST PLACES?," Dataindonesia.id. [Online]. Available: https://dataindonesia.id/varia/detail/indonesia-punya-17001-pulau-pada-2022-paling-banyak-di-mana

[2] Jing Wang, Xuhong Peng, Jindong Wu, Youde Ding, Barkat Ali, Yizhou Luo, Yiting Hu, Keyao Zhang, SINGULAR SPECTRUM ANALYSIS (SSA) BASED HYBRID MODELS FOR EMERGENCY AMBULANCE DEMAND (EAD) TIME SERIES FORECASTING, *IMA Journal of Management Mathematics*, Volume 35, Issue 1, January 2024, Pages 45–64

[3] S. I. Christienova, E. W. Pratiwi, and G. Darmawan, "COMPARISON OF SINGULAR SPECTRUM ANALYSIS (SSA) AND FOURIER SERIES ANALYSIS (FSA) FORECASTING MODELS ON AIR TEMPERATURE DATA IN SURABAYA," *Berk. MIPA*, vol. 25, no. 1, pp. 94–106, 2018.

[4] Aizenbud, Yariv & Sober, Barak. (2021). CONVERGENCE RATES OF VECTOR-VALUED LOCAL POLYNOMIAL REGRESSION. 10.48550/arXiv.2107.05852.

[5] Emmenegger, Corinne & Bühlmann, Peter. (2021). DOUBLE MACHINE LEARNING FOR PARTIALLY LINEAR MIXED-EFFECTS MODELS WITH REPEATED MEASUREMENTS. 10.48550/arXiv.2108.13657.

[6] S. Suparti and A. Prahutama, "NONPARAMETRIC REGRESSION MODELING USING A LOCAL POLYNOMIAL APPROACH ON ELECTRICAL LOAD IN SEMARANG CITY," *Media Stat.*, vol. 9, no. 2, pp. 85–93, 2016.

[7] S. M. Sun MingDong and L. X. Li XuYong, "WINDOW LENGTH SELECTION OF SINGULAR SPECTRUM ANALYSIS AND APPLICATION TO PRECIPITATION TIME SERIES.," 2017.

[8] S. M. Sun MingDong and L. X. Li XuYong, "WINDOW LENGTH SELECTION OF SINGULAR SPECTRUM ANALYSIS AND APPLICATION TO PRECIPITATION TIME SERIES.," 2017.

[9] S. Satriani and R. Ibnas, "FORECASTING THE CONSUMER PRICE INDEX (CPI) IN SOUTH SULAWESI USING THE SINGULAR SPECTRUM ANALYSIS (SSA) METHOD," *J. MSA (Mathematics and Stat. and Apl.*, vol. 8, no. 1, pp. 82–89, 2020.

[10] Nguyen, Anh & Nguyen, Phi Le & Vu, Viet & Pham, Quoc & Nguyen, Viet & Nguyen, Minh & Nguyen, Hùng & Nguyen, Kien. (2022). ACCURATE DISCHARGE AND WATER LEVEL FORECASTING USING ENSEMBLE LEARNING WITH GENETIC ALGORITHM AND SINGULAR SPECTRUM ANALYSIS-BASED DENOISING. SCIENTIFIC REPORTS. 12. 10.1038/s41598-022-22057-8.

[11] S. Wadekar, A. Mahalkari, A. Ali and A. Gupta, "A REVIEW ON SINGULAR SPECTRUM ANALYSIS," 2022 IEEE INTERNATIONAL CONFERENCE ON CURRENT DEVELOPMENT IN ENGINEERING AND TECHNOLOGY (CCET), Bhopal, India, 2022, pp. 1-6, doi: 10.1109/CCET56606.2022.10080082.

[12] G. Darmawan, "IDENTIFICATION OF RAINFALL DATA PATTERNS IN THE GROUPING PROCESS IN THE SINGULAR SPECTRUM ANALYSIS METHOD," 2016.

[13] Nur Chamidah , Budi Lestari, *ANALISIS REGRESI NONPARAMETRIK DENGAN PERANGKAT LUNAK R*, Airlangga University Press, 2022

[14] V. Fibriyani and N. Chamidah, "PREDICTION OF INFLATION RATE IN INDONESIA USING LOCAL POLYNOMIAL ESTIMATOR FOR TIME SERIES DATA," in *Journal of Physics: Conference Series*, IOP Publishing, 2021, p. 12065.

[15] H. Hassani and R. Mahmoudvand, *SINGULAR SPECTRUM ANALYSIS: USING R*. Springer, 2018.

[16] Y. A. Jatmiko, R. L. Rahayu, and G. Darmawan, "COMPARISON OF THE ACCURACY OF THE FORECAST RESULTS OF SHALLOT PRODUCTION BY THE HOLT-WINTERS METHOD WITH SINGULAR SPECTRUM ANALYSIS (SSA)," *J. Mat. MANTIK*, vol. 3, no. 1, pp. 13–22, 2017.

[17] F. Fitri, R. Rahmat, and A. D. Pengestuti, "FORECASTING OF RAINFALL IN SUMATERA BARAT: SINGULAR SPECTRUM ANALYSIS (SSA) APPLICATION," in *Journal of Physics: Conference Series*, IOP Publishing, 2020, p. 12047.