

FORECASTING THE NUMBER OF PASSENGER AT JENDERAL AHMAD YANI SEMARANG INTERNATIONAL AIRPORT USING HYBRID SINGULAR SPECTRUM ANALYSIS-NEURAL NETWORK (SSA-NN) METHOD

Tresiani Yunitasari¹, M. Al Haris^{2*}, Prizka Rismawati Arum³

^{1,2,3} Department of Statistics, Universitas Muhammadiyah Semarang, Semarang, Indonesia

Email: *fickriamalia@pnb.ac.id*

Article Info:

Received: April 10, 2023

Accepted: May 12, 2023

Available Online: May 31, 2023

Keywords:

Forecasting; Hybrid SSA-NN; Passenger; Transportation

Abstract: Transportation was an important sector of supporting the economic growth of a country. The impact of the Covid-19 2020 pandemic at Achmad Yani International Airport in Semarang resulted in the movement of the number of passengers decreasing quite drastically, but in mid-2020 the movement of the number of passengers had slowly increased. Forecasting was done to determine the flow of movement of the number of passengers in the future using the Hybrid Singular Spectrum Analysis (SSA)-Neural Network (NN) method. The SSA method was expected to be able to decompose various patterns in the data into trend, seasonality and noise. Furthermore, the NN method was used to analyze nonlinear patterns in the data. The results based on the number of domestic air passengers data at Jenderal Ahmad Yani International Airport Semarang from January 2006 to December 2021 showed that the best method was a combination of the SSA method with a window length of 40 and the NN method with a 6-8-1 network architecture (6 input neurons, 8 hidden neurons and 1 output neuron) for the trend component, 11-15-1 (11 neurons input, 15 hidden neurons and 1 output neuron) for the seasonal component, and 10-15-1 (10 input neurons, 15 hidden neurons and 1 output neuron) for the noise component. The model produces a prediction error based on a MAPE value of 0.54% or an accuracy rate of 99.46%.

1. INTRODUCTION

Indonesia as a country with a large enough population, where the increasing population growth was followed by the increasing need for transportation facilities [1]. Transportation facilities had a very important role for the community because they can support all aspects of life making it easier to carry out activities between regions. One of the transportation facilities that can be used is air transportation [2]. Today there were many human activities that used airplanes. This can be seen from the increasingly crowded community activities at Achmad Yani International Airport Semarang which is a part of PT Angkasa Pura I (Persero) which is engaged in air transportation flight services [1]. Aviation services were air transportation services that were in great demand by the public because they provided a feeling of comfort, safety, and had a fairly high speed with a relatively short travel time compared to other transportation [3].

Semarang Achmad Yani International Airport was the busiest airport in terms of flight activity with the highest number of passengers from several airports in Central Java [4].

However, the Covid-19 pandemic resulted in a drastic decreasing in the number of passengers. Based on data from the Badan Pusat Statistik (BPS) of Central Java Province, the movement of the number of domestic flight passengers at Jenderal Ahmad Yani International Airport in Semarang in 2020 decreased by 46.58% or 2.3 million passengers from the previous year [5]. This decreasing in the number of passengers occurred in line with the stipulation of Government Regulation No. 21 of 2020 concerning Large-Scale Social Restrictions in the context of accelerating the handling of Covid-19 [6]. In 2022, the movement of the number of passengers will slowly start to increase, but still not showing too big a change [6]. Therefore, forecasting was very important to use to determine the flow of movements of the number of domestic flight passengers at Achmad Yani International Airport, Semarang in the coming period as a material consideration in policy making [6], [7].

Forecasting was a method used to estimate a value in the future using data in the past [8]. The development of the number of passengers who have a certain pattern of movement, will provide a separate calculation in forecasting. So a forecasting method was needed that can determine the components of the data pattern separately by decomposing it into sub patterns so that better forecasting accuracy can be obtained. The method that can be used to decompose time series data patterns was the Singular Spectrum Analysis (SSA) method [9]. SSA was a forecasting method that combined elements of classical time series analysis, multivariate statistics, multivariate geometry, dynamic systems and signal processing. The main objective of this method was to decompose the original time series into a small number of components that can be identified by patterns such as trend, seasonality and noise [10].

In addition to the trend, seasonality and noise components, nonlinear patterns were often found in the time series data. Therefore, additional methods were needed that can capture nonlinear patterns in the data, one of which was the Neural Network (NN) method [9], [11]. This research will forecast the number of passengers on domestic flights at Achmad Yani International Airport in Semarang using the hybrid SSA-NN method. The SSA method was expected to be able to decompose various patterns in the number of passengers into trend, seasonality and noise, while the NN method was used to analyze data that has a nonlinear relationship pattern. The combination of methods carried out can increase the accuracy of the forecast results, because the combination of the two methods tends to produce better forecasts compared to using one method [9].

Research on the number of passengers was carried out by [1] to predict the number of airplane passengers at Ahmad Yani International Airport in Semarang using the Holt Winter's Exponential Smoothing method and the Exponential Smoothing Event Based method. The results of this study obtained that the Holt Winter's Exponential Smoothing method was the best method with the smallest accuracy error based on the Mean Absolute Percentage Error (MAPE) value of 5.644139%. Forecasting the number of passengers has also been carried out by (Larissa et al., 2021) [2] to predict the number of passengers at Soekarno-Hatta Airport. The results showed that the Holt's Winter Additive method was the best method based on the smallest MAPE value of 17.465%.

Forecasting using the Hybrid SSA-NN method was carried out by (Suhartono et al., 2019) [9] to predict the value of inflow and outflow fractions of currency in Indonesia by comparing the ARIMAX method with Hybrid SSA-NN. The results obtained where the Hybrid SSA-NN method provides a better forecasting value. The application of the other hybrid model with SSA has also been carried out by [11]. The results showed that the hybrid model yielded more accurate than the single model. Therefore, in this study it is hoped that the results of

forecasting the number of domestic flight passengers at Achmad Yani International Airport in Semarang using the Hybrid SSA-NN method can provide more accurate forecasting results.

2. LITERATURE REVIEW

2.1. Forecasting

Forecasting was a way to estimate a value in the future by taking into account past and current data [10]. Forecasting was the basis for long-term planning for a company or agency. The accuracy of forecasting results will increase the chances of achieving a profitable investment. Forecasting played an important role in making a policy, whether it was effective or not can be seen from the time the policy was taken [6], [7].

2.2. Singular Spectrum Analysis (SSA)

SSA was a technique for decomposing time series data into several pattern components (trend, seasonal and noise), so that they were easier to interpret. The SSA method did not require special assumptions, so its used becomes wider. The SSA method was divided into 2 main stages, namely decomposition and reconstruction [12].

2.3. Decomposition

The parameter used in the decomposition stage was the window length (L). Parameter L had a function to determine the number of dimensions of the path matrix. The determination of the value of L was done by checking through trial and error. This decomposition stage consists of two stages, namely the Embedding stage and the Singular Value Decomposition (SVD) stage [10].

The Embedding stage was carried out by converting a time series data into multidimensional data (matrix). Suppose the time series data with length N was expressed by $F = (f_1, f_2, \dots, f_N)$ where F was a time series data that was not zero or there was no missing data which was then converted into a path matrix with size $L \times K$. L parameter was specified by value $2 < L < \frac{N}{2}$ and then created a value $K = N - L + 1$. The path matrix of the F series can be described as follows (H. Utami et al., 2019):

$$F = X = [X_1: \dots: X_K] = (x_{ij})_{i,j=1}^{L,K} = \begin{pmatrix} f_1 & f_2 & f_3 & \dots & f_K \\ f_2 & f_3 & f_4 & \dots & f_{K+1} \\ f_3 & f_4 & f_5 & \dots & f_{K+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ f_L & f_{L+1} & f_{L+2} & \dots & f_N \end{pmatrix}$$

The next stage was the formation of a Singular Value Decomposition (SVD) from the X path matrix. On the X matrix, the eigenvalues were determined by $(\lambda_1, \lambda_2, \dots, \lambda_L)$ from a symmetric matrix $S = XX^T$ where $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_L > 0$ with the formula $|S - \lambda I| = 0$, eigenvector (U_1, U_2, \dots, U_L) from the S matrix corresponding to the eigenvalues [11].

Defined $d = \max \{i, \text{ so that } \lambda_i > 0\}$ was the rank of the X matrix. If the principal component is denoted by $V_i = \frac{x^T U_i}{\sqrt{\lambda_i}}$, for $i = 1, 2, \dots, d$, then the SVD of the X path matrix could be written as shown in Equation (1),

$$X = X_1 + X_2 \dots + X_d \quad (1)$$

with $X_i = \sqrt{\lambda_i}U_iV_i^T$ and matrix X_i had rank 1. Therefore the X_i matrix was an elementary matrix and a collection $(\sqrt{\lambda_i}, U_i, V_i)$ was called eigentriple ke- i dari SVD [13].

2.4. Reconstruction

The reconstruction phase was the process of creating new time series data by grouping and diagonal averaging. The parameter used at this stage was the grouping effect (r) which functions to determine patterns in data plot. Grouping was a set of indices of $\{1, 2, \dots, d\}$ become m groups of mutually exclusive subsets denoted by I_1, I_2, \dots, I_m . Suppose $I = \{i_1, i_2, \dots, i_p\}$, then the resulting X_I matrix corresponds to group I which was defined as $X_I = X_{i_1} + \dots + X_{i_p}$. This matrix was calculated for $I = I_1, \dots, I_m$ and the expansion of Equation (1) causes the decomposition to become the Equation (2) [14],

$$X = X_{I_1} + X_{I_2} + \dots + X_{I_m}, \tag{2}$$

the process of selecting sets I_1, I_2, \dots, I_m was called eigentriple grouping. If $m = d$ and $I_j = \{j\}$, $j = 1, 2, \dots, d$, then the appropriate grouping was called elementary [13].

The last step in the SSA method was to change each matrix X_{I_m} from the grouped decomposition in Equation (2) into a new series with length N . Suppose Y was a matrix of size $L \times K$ with elements y_{ij} , $1 \leq i \leq L$, $1 \leq j \leq K$, for $L \leq K$. Given $L^* = \min(L, K)$, $K^* = \max(L, K)$, $N = L + K - 1$, $y_{ij}^* = y_{ji}$ for $L < K$ and $y_{ij}^* = y_{ij}$ for $L > K$. Based on these provisions, diagonal averaging was done by moving the Y matrix to the series g_0, g_1, \dots, g_{N-1} with the Equation (3) [9],

$$g_k \begin{cases} \frac{1}{k+1} \sum_{m=1}^{k+1} y_{m, k-m+2}^* & \text{untuk } 0 \leq k < L^* - 1, \\ \frac{1}{L^*} \sum_{m=1}^{L^*} y_{m, k-m+2}^* & \text{untuk } L^* - 1 \leq k < K^* \\ \frac{1}{N-k} \sum_{m=k-K^*+2}^{N-K^*+1} y_{m, k-m+2}^* & \text{untuk } K^* \leq k < N. \end{cases} \tag{3}$$

Equation (3) was related to the average of the matrix elements over the anti-diagonal $i + j = k + 2$. Choice $k = 0$ giving $g_0 = y_{1,1}$, choice $k = 1$ giving $g_1 = \frac{y_{1,2} + y_{2,1}}{2}$ and so on. Note that if the Y matrix was a path matrix of several series (h_1, \dots, h_N) , then $g_i = h_i$ for all i . If diagonal averaging equation (4) was applied to the resulting X_{I_k} matrix, a reconstructed series would be obtained $F^{(k)} = (f_1^{(k)}, \dots, f_N^{(k)})$. Therefore, the initial series f_1, \dots, f_N was decomposed into the sum of m reconstructed series as shown in Equation (4) [14]:

$$f_n = \sum_{k=1}^m f_n^{(k)}, \quad (n = 1, \dots, N) \tag{4}$$

2.5. Neural Network (NN)

The NN method had been widely developed and could be used in predicting past patterns because of its ability to remember and made generalizations from what has existed before, its ability to learn and was immune to errors, so it could create a system that was resistant to damage and consistently works well [9]. The NN architecture that was widely used and applied was MultiLayer Perceptrons (MLP) known as Feedforward Neural Networks (FFNN). FFNN in statistical modeling could be viewed as a flexible class of nonlinear functions. A special form of FFNN with one hidden layer consisting of q neuron units and an output layer consisting of only one neuron unit with response or output values \hat{y} was calculated by the Equation 5 [15]:

$$\hat{y}_{(k)} = f^0 \left[\sum_{j=1}^q \left[w_1^0 f_j^h \left(\sum_{i=1}^p w_{ji}^h x_{i(k)} + b_j^h \right) + b^0 \right] \right] \quad (5)$$

with

- $x_{i(k)}$: input variables $p, i=1, 2, \dots, p$;
- $\hat{y}_{(k)}$: The estimated of the output variable;
- k : Paired input-target data index $(x_{i(k)}, y_{(k)}), k = 1, 2, \dots, n$;
- w_{ji}^h : Input weights for i to j neuron of the hidden layer, $(j = 1, 2, \dots, q)$;
- b_j^h : Neuron bias for j of the hidden layer, $(j = 1, 2, \dots, q)$;
- f_j^h : Activation function in neurons to j of the hidden layer;
- w_1^0 : Neuron weights for j of the hidden layer to output layer ;
- b^0 : Neuron bias of the output layer;
- f^0 : Activation function in the output layer;

2.6. Hybrid Singular Spectrum Analysis-Neural Network (SSA-NN)

The application of the Hybrid SSA-NN method was carried out by decomposing a time series data into trend, seasonality and noise patterns. The results of the decomposition were forecasted using the FFNN method. Forecasting was done in aggregate, namely by summing the components that had the same pattern so that only three main patterns were formed, namely trend, seasonality and noise. If the resulting noise meets white noise, then forecasting was not necessary. The final forecasting result was the sum of all components into one time series data [9].

2.7. Model Goodness Evaluation

The evaluation of model goodness is carried out to determine how well the model performs in predicting events for several future periods. The evaluation is done by looking at the accuracy of the forecasting results based on the Mean Absolute Percentage Error (MAPE) value [16]. MAPE is the average percentage error in forecasting compared to the actual value. The MAPE value can be obtained using the Equation (7) [9].

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{Y_t - \hat{Y}_t}{Y_t} \right| \times 100\% \quad (6)$$

with

- Y_t : The actual value for period t ;
 \hat{Y}_t : The predicted value for period t ;
 n : the number of sampel.

3. METHODOLOGY

The data used in this study was data on the number of passengers on domestic flights at Achmad Yani International Airport, Semarang. This data was monthly data collected from January 2006 to December 2021 obtained from the official website of the Badan Pusat Statistik (BPS) of Central Java Province [5].

The procedure to forecasting the number of passenger at Jenderal Ahmad Yani Semarang International Airport using Hybrid SSA-NN method is as follows:

1. Describing the data of the number of domestic flight passengers at Jenderal Ahmad Yani International Airport in Semarang from January 2006 to December 2021;
2. Decomposing data using the SSA model which includes the stages of embedding, singular value decomposition, grouping, and diagonal averaging;
3. Modeling the decomposed data (trend, seasonal, and noise components) using a neural network architecture;
4. Calculating the accuracy value of the SSA-NN model's prediction results;
5. Forecasting the number of domestic flight passengers at Jenderal Ahmad Yani International Airport in Semarang for the next 12 periods.

4. RESULTS AND DISCUSSION

The characteristics of data on the number of domestic flight passengers at Achmad Yani Semarang International Airport from January 2006 to December 2021 can be seen in Figure 1.

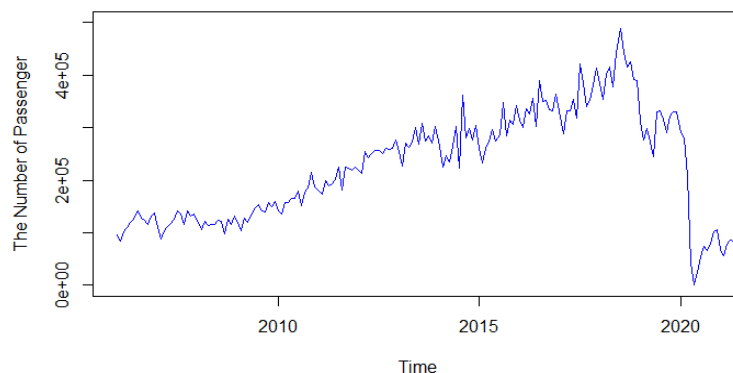


Fig 1. The Graph of the Number of Domestic Flight Passengers at Achmad Yani Semarang International Airport from January 2006 to December 2021

Figure 1 showed that the graph of the number of domestic flight passengers at Achmad Yani Semarang International Airport fluctuated every month and tends to increase from 2006

to 2018. It can be seen that the number of domestic flight passengers tends to increase during holiday seasons such as Eid al-Fitr, Christmas, and New Year's Eve. However, the situation changed in 2019 when the number of passengers began to decline. Angkasa Pura Corporation stated that the decrease in the number of passengers in 2019 was due to the issue of high ticket prices, which caused a decrease in public interest in purchasing them. In addition, the significant decrease in the number of flight passengers in 2020 was caused by the COVID-19 pandemic that entered Indonesia's territory.

4.1. Decomposition

The decomposition stage began with embedding. The embedding process was carried out by determining the value of the Windows Length (L) parameter with $2 < L < \frac{N}{2}$ through trial and error. The data used in this study consisted of 192 observations, so the value of L ranged from 2 to 96. Next, trial and error was carried out for the values of L = 10, 20, 30, 40, 50, 60, 70, 80, and 90, and then the value of L with the smallest MAPE was selected. The results of the trial and error of L values were presented in Table 1.

Table 1. The Trial and Error Results of Selecting the L Value

| L | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 |
|------|-------|-------|-------|--------------|-------|-------|-------|-------|-------|
| MAPE | 31,58 | 38,79 | 48,37 | 12,39 | 22,02 | 23,51 | 27,06 | 24,69 | 25,31 |

Table 1 showed that the smallest value of L was 40 with the smallest MAPE value of 12.39%. Next, the value of $K = N - L + 1$ or $K = 192 - 40 + 1 = 153$ was used to form a trajectory matrix with the order of $L \times K$. Therefore, the trajectory matrix X (Hankel) can be arranged and was used to obtain eigen-triples by forming a symmetric matrix $S = XX^T$ as follows.

$$S = X_{(40 \times 153)} \cdot X_{(153 \times 40)}^T$$

$$S = \begin{bmatrix} 9.636933 \times 10^{12} & 9.653509 \times 10^{12} & 9.725311 \times 10^{12} & \dots & 4.084425 \times 10^{12} \\ 9.653509 \times 10^{12} & 9.807891 \times 10^{12} & 9.811341 \times 10^{12} & \dots & 4.126904 \times 10^{12} \\ 9.725311 \times 10^{12} & 9.811341 \times 10^{12} & 9.953584 \times 10^{12} & \dots & 4.191405 \times 10^{12} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 8.752223 \times 10^{12} & 8.820926 \times 10^{12} & 8.886778 \times 10^{12} & \dots & 1.100546 \times 10^{13} \end{bmatrix}$$

The symmetric matrix $S_{(40 \times 40)}$ that has been obtained was used to calculate the eigen-triples. the eigen-triples were used to calculate the principal component values that the result was shown in Table 2.

Table 2. Principal Component Values

| No | V_1 | V_2 | ... | V_{40} |
|-----|-------------|-----------------------------|-----|---------------|
| 1 | -0.03713652 | 6.699947×10^{-04} | ... | -0.080961445 |
| 2 | -0.03739637 | $-3.063632 \times 10^{-05}$ | ... | 0.041258635 |
| 3 | -0.03784994 | $-2.784025 \times 10^{-04}$ | ... | 0.059741302 |
| 4 | -0.03823682 | 8.958178×10^{-04} | ... | -0.033177156 |
| 5 | -0.03847604 | 1.743780×10^{-03} | ... | -0.038211754 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 153 | -0.05761631 | -0.188207694 | ... | -0.0121967577 |

4.2. Reconstruction

The reconstruction stage was carried out by grouping the eigentriples into trend, seasonal, and noise components based on the Effect grouping parameter (r). The value of r was determined based on the number of eigentriples that do not reflect noise in the singular value plot. The singular value plot based on 40 eigentriples was presented in Figure 2

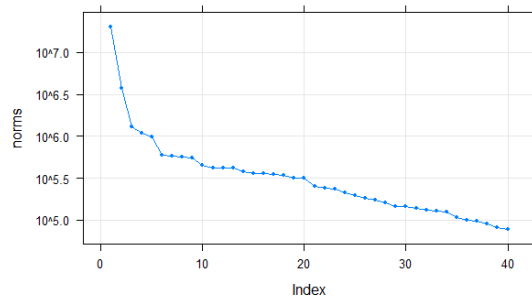


Fig 2. The Singular Value Plot

The singular values based on Figure 2 showed a slow decreasing pattern on eigenvectors 21 to 40 that indicating of 21 to 40 eigenvectors were grouped as noise components and affect the determination of $r = 20$. Furthermore, the eigenvectors that will be used to group trend and seasonal components were the first twenty eigenvectors presented in Figure 3.

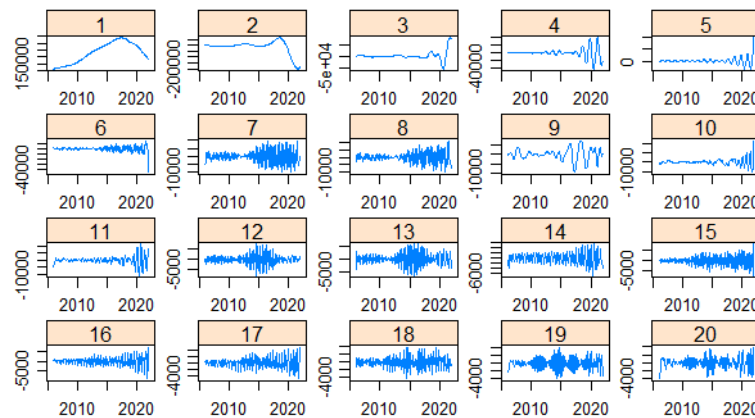


Fig 3. Reconstruction Plot of the First 20 Eigentriples

Based on the reconstructed series plot in Figure 3, it can be seen that the series reconstructed by eigentriples 1, 2, and 3 contain slowly varying components, so eigentriples 1, 2, and 3 were grouped into the trend component. The grouping of eigentriples related to seasonality was based on the similarity of the singular values of consecutive eigentriples. In the reconstructed series plot, the similarity of singular values results in the reconstructed series by an eigentriple having the same seasonal pattern and period as the series reconstructed by other eigentriples. Some pairs of consecutive eigentriples that have similar patterns were eigentriples 4 and 5, eigentriples 6, eigentriples 7 and 8, eigentriples 9, eigentriples 10 and 11,

eigentriples 12 and 13, eigentriples 14 and 15, eigentriples 16 and 17, eigentriples 18 and 19, and eigentriples 20.

The final step in the reconstruction was diagonal averaging. Diagonal averaging was performed by summing up the reconstruction results for each component. The result of the diagonal averaging produces 3 time series data patterns consisting of trend, seasonal, and noise, as shown in Figure 4. The time series data patterns of trend, seasonal, and noise obtained from the SSA method reconstruction were then further processed using the NN method. The application of the NN method to the reconstruction results was expected to improve forecasting accuracy.

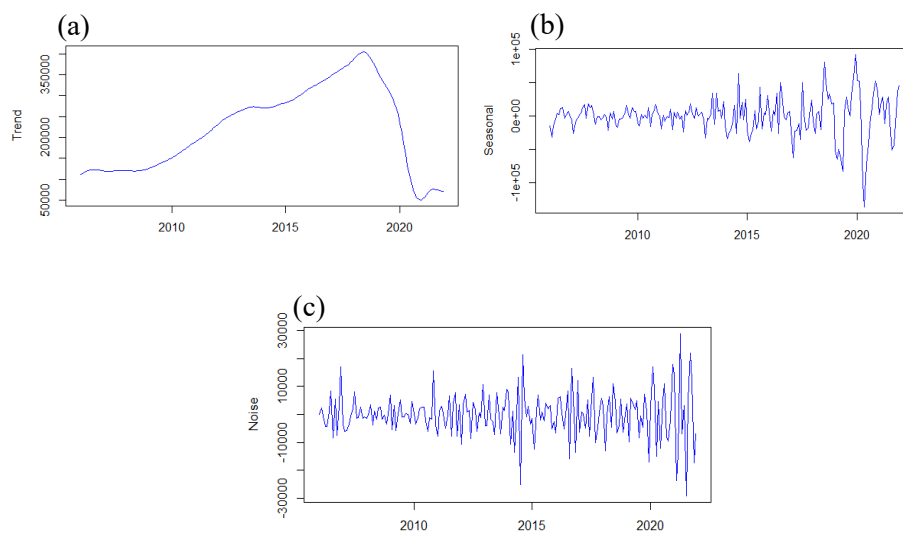


Fig 4. Diagonal Averaging Plot (a) Trend Patterns, (b) Seasonal Patterns and (c) Noise Patterns

4.3. Hybrid Singular Spectrum Analysis-Neural Network (SSA-NN)

Determining the input variables for the NN architecture in each component (trend, seasonality, noise) was done by observing the stationary ACF and PACF plots. The significant lags on the PACF plot will be used as input variables for The NN architecture. The NN architecture was formed using 1 hidden layer with the tanh activation function. The determination of the number of neurons in the hidden layer was done using the cross-validation method, where the number of hidden neurons tried includes 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15. The best NN architecture was determined based on the smallest MAPE value. The training results of the NN network in the three components were shown in Table 5.

Based on Table 5, the best architecture in the trend component was the architecture with input variables of 1, 2, 4, 5, 6, and 13 and the number of neurons in the hidden layer was 8, which produced the smallest MAPE value of 1.75%. The best architecture in the seasonal component was the architecture with input variables of 1, 5, 6, 9, 10, 12, 13, 14, 15, 16, and 18 and the number of neurons in the hidden layer was 15, which produced the smallest MAPE value of 3.91%. The best architecture in the noise component was the architecture with input variables of 1, 2, 3, 4, 6, 7, 8, 9, 10, and 12 and the number of neurons in the hidden layer was 15, which produced the smallest MAPE value of 5.44%.

Table 5. Determining the Best NN Architecture in the Three Components

| Trend Component Input Lags: 1, 2, 4, 5, 6, and 13 | | Seasonal Component Input Lags 1, 5, 6, 9, 10, 12, 13, 14, 15, 16 and 18 | | Noise Component Input Lags 1, 2, 3, 4, 6, 7, 8, 9, 10 and 12 | |
|--|-------------|--|-------------|--|-------------|
| Neurons in the Hidden Layer | MAPE | Neurons in the Hidden Layer | MAPE | Neurons in the Hidden Layer | MAPE |
| 1 | 2.94 | 1 | 226.24 | 1 | 550.99 |
| 2 | 4.07 | 2 | 177.41 | 2 | 197.33 |
| 3 | 2.80 | 3 | 184.94 | 3 | 316.98 |
| 4 | 4.07 | 4 | 125.17 | 4 | 534.35 |
| 5 | 1.97 | 5 | 133.78 | 5 | 297.76 |
| 6 | 5.60 | 6 | 126.93 | 6 | 195.61 |
| 7 | 4.17 | 7 | 75.76 | 7 | 270.68 |
| 8 | 1.75 | 8 | 58.20 | 8 | 158.00 |
| 9 | 2.89 | 9 | 27.63 | 9 | 174.17 |
| 10 | 2.86 | 10 | 23.96 | 10 | 86.59 |
| 11 | 3,11 | 11 | 14,44 | 11 | 37,04 |
| 12 | 2,28 | 12 | 9,35 | 12 | 44,71 |
| 13 | 1,9 | 13 | 7,15 | 13 | 16,68 |
| 14 | 5,61 | 14 | 4,89 | 14 | 8,60 |
| 15 | 1,8 | 15 | 3,91 | 15 | 5,44 |

The forecasting results for each component's architecture were then summed up to obtain the final Hybrid SSA-NN forecast. After obtaining the prediction results using the Hybrid SSA-NN method, the prediction results were evaluated on the data of the number of domestic flight passengers at Achmad Yani International Airport in Semarang, with the MAPE value as the criteria for the goodness of the model. The comparison of prediction accuracy results can be seen in Table 6.

Table 6. Determining the Best NN Architecture in the Three Components

| Method | MAPE |
|--|-------|
| Singular Spectrum Analysis | 12.39 |
| Hybrid Singular Spectrum Analysis-Neural Network | 0.54 |

Table 6 showed that the prediction using the Hybrid SSA-NN method resulted the MAPE value of 0.54%. This value was lower than the SSA method. Therefore, the application of the Hybrid SSA-NN method was better used to forecast the number of domestic airline passengers at Ahmad Yani International Airport Semarang. The forecasting results using both methods were also presented in Figure 7.

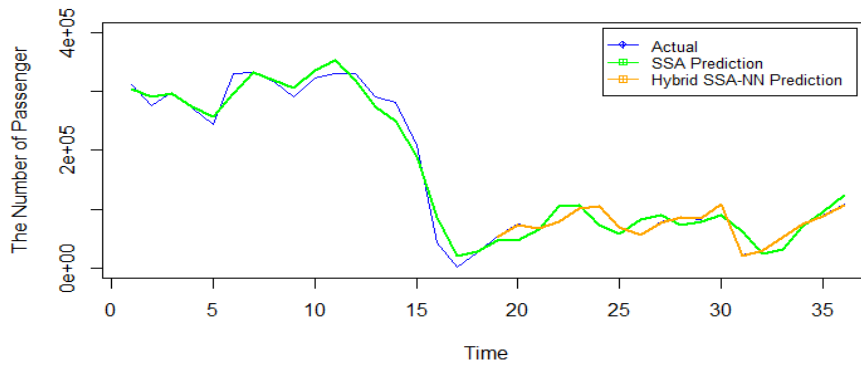


Fig 7. Comparison of Actual Data with the Prediction Results of Both Methods

5. CONCLUSION

The number of passengers at Achmad Yani Semarang International Airport has fluctuated due to the Covid-19 pandemic. The best model that can be used for forecasting is the Hybrid SSA-NN model with a window length of 40 and network architecture of 6-8-1 (6 input neurons, 8 hidden neurons, and 1 output neuron) for the trend component, 11-15-1 (11 input neurons, 15 hidden neurons, and 1 output neuron) for the seasonal component, and 10-15-1 (10 input neurons, 15 hidden neurons, and 1 output neuron) for the noise component. The model produces an MAPE value of 0.54% or an accuracy rate of 99.46%, which indicates that the prediction results are very good.

REFERENCES

- [1] Sofiana, Suparti, A. R. Hakim, and I. T. Utami, "Peramalan Jumlah Penumpang Pesawat di Bandara Internasional Ahmad Yani dengan Metode Holt Winter's Exponential Smoothing dan Metode Exponential Smoothing Event Based," *J. Gaussian*, vol. 9, no. 4, 2020, doi: 10.14710/j.gauss.v9i4.29448.
- [2] A. Larissa, F. C. Garini, W. Anbiya, R. A. Fatharani, F. Salsabila, and T. Toharudin, "Perbandingan Metode SES, Holt's Linear, Holt's Winter untuk Peramalan Jumlah Penumpang Pesawat: Penerbangan Domestik di Bandara Soekarno-Hatta," in *SEMINAR NASIONAL STATISTIKA X (2021)*, 2021.
- [3] W. M. Sialagan, "Jasa Transportasi Udara Indonesia," *J. Sos. Hum.*, vol. 2, no. 1, pp. 57–66, 2011.
- [4] M. U. Aidi, Y. Anas, B. Hario, and W. Kushardjoko, "Analisis Kapasitas Air side Rencana Pengembangan Bandar Udara Internasional Ahmad Yani Semarang," *J. Karya Tek. Sipil*, vol. 2, no. 4, pp. 127–136, 2013.
- [5] BPS Provinsi Jawa Tengah, *Jawa Tengah dalam Angka 2022*. 2022.
- [6] F. Rianda, "Pemodelan Intervensi untuk Menganalisis dan Meramalkan Jumlah Penumpang Pesawat di Bandara Soekarno-Hatta Akibat Pandemi Covid-19," in *Seminar Nasional Official Statistics 2021*, 2021, vol. 2021, no. 1, pp. 283–292, doi: 10.34123/semnasoffstat.v2021i1.857.
- [7] Y. Rizkiana, E. Zukhronah, and H. Pratiwi, "Peramalan Data Banyak Penumpang Bandara Adi Soemarmo Menggunakan Metode SARIMAX," *Pros. Sendika*, vol. 5, no. 2, pp. 183–188, 2019.

- [8] Suhartono, S. Isnawati, N. A. Salehah, D. D. Prastyo, H. Kuswanto, and M. H. Lee, "Hybrid SSA-TSR-ARIMA for water demand forecasting," *Int. J. Adv. Intell. Informatics*, vol. 4, no. 3, pp. 238–250, 2018, doi: 10.26555/ijain.v4i3.275.
- [9] Suhartono, E. Setyowati, N. A. Salehah, M. H. Lee, S. P. Rahayu, and B. S. S. Ulama, *A hybrid singular spectrum analysis and neural networks for forecasting inflow and outflow currency of bank Indonesia*, vol. 937. Springer Singapore, 2019.
- [10] H. Utami, Y. W. Sari, Subanar, Abdurakhman, and Gunardi, "Peramalan Beban Listrik Daerah Istimewa Yogyakarta dengan Metode Singular Spectrum Analysis (SSA)," *Media Stat.*, vol. 12, no. 2, p. 214, 2019, doi: 10.14710/medstat.12.2.214-225.
- [11] I. Athoillah, A. H. Wigena, and H. Wijayanto, "Hybrid Modeling of Singular Spectrum Analysis and Support Vector Regression for Rainfall Prediction," in *Journal of Physics: Conference Series*, 2021, vol. 1863, no. 1, doi: 10.1088/1742-6596/1863/1/012054.
- [12] D. Safitri, Subanar, H. Utami, and W. Sulandari, "Forecasting of jabodetabek train passengers using singular spectrum analysis and holt-winters methods," in *Journal of Physics: Conference Series*, 2020, vol. 1524, pp. 1–9, doi: 10.1088/1742-6596/1524/1/012100.
- [13] A. A. Ete, M. Fitriawanawati, and M. T. Arifin, "Forecasting the Number of Tourist Arrivals to Batam by applying the Singular Spectrum Analysis and the Arima Method," in *1st International Conference on Progressive Civil Society (IconProCS 2019)*, 2019, vol. 317, pp. 119–126, doi: 10.2991/iconprocs-19.2019.24.
- [14] Satriani, Nursalam, and R. Ibtnas, "Peramalan Indeks Harga Konsumen (IHK) di Sulawesi Selatan dengan Menggunakan Metode Singular Spektrum Analysis (SSA)," *J. MSA (Mat. dan Stat. serta Apl.)*, vol. 8, no. 1, p. 82, 2020, doi: 10.24252/msa.v8i1.17441.
- [15] A. Suharsono, M. Monica, and J. D. T. Purnomo, "Perbandingan Model Hybrid ARIMAX-FFNN-EGARCH dan Model Hybrid SETAR-EGARH untuk Peramalan (Studi Kasus: Data Cash Outflow dan Inflow Bank Indonesia Kota Kediri)," *Inferensi*, vol. 5, no. 1, p. 23, 2022, doi: 10.12962/j27213862.v5i1.12470.
- [16] S. W. Utami, I. M. Nur, and M. Al Haris, "Peramalan Nilai Ekspor Provinsi Jawa Tengah dengan Metode Fuzzy Time Series Berbasis Algoritma Novel," *J Stat. J. Ilm. Teor. dan Apl. Stat.*, vol. 15, no. 1, pp. 195–202, 2022.