Hybrid Modeling of Singular Spectrum Analysis and Support Vector Regression for Rainfall Prediction

Ibnu Athoillah*, Aji Hamim Wigena, and Hari Wijayanto
Department of Statistics, IPB University, Bogor, 16680, Indonesia

*E-mail: ibnu1990ibnu@apps.ipb.ac.id

Abstract. Rainfall can provide many benefits such as for agriculture, water resource management, and electricity. However, it can also cause hydrometeorological disasters such as floods and droughts, so that an accurate rainfall prediction is needed to anticipate the risks and minimize the losses. Rainfall characteristics are diverse, complex, and uncertain. Thus rainfall data are usually nonlinear time series and difficult to predict using traditional methods such as ARIMA. In recent years, machine learning models such as support vector regression (SVR) hybrid model has been developed to improve prediction accuracy. The SVR hybrid model can be carried out with Singular Spectrum Analysis (SSA) for data pre-processing. The SSA method is used to decompose original time series data into trend, oscillatory, and noise components. The SVR model is then used to predict rainfall based on reconstruction series from SSA without noise components. The grid search algorithm using an optimization method is used to estimate the parameters of the SVR model. This research aims to apply an SSA-SVR hybrid model and compares it to the SVR model using monthly rainfall data at Kemayoran Station from 1980 to 2019. Based on the result showed that the hybrid model yielded more accurate than the single model.

1. Introduction
Rainfall has a very important meaning for life, so the accurate rainfall prediction is needed. It can be used to support agricultural production, water resource management, and hydrometeorological disaster mitigation. Agricultural cropping patterns, yields, and the threat of disaster depend on the seasonal pattern of rainfall. The Accurate rainfall prediction can be used to anticipate the risks and minimize the losses. However, in building the prediction model there are challenges due to its complexity and various factors that influence rainfall.

Jakarta is located in the lowlands and confluence of several rivers from upstream area which have high rainfall and empties into Jakarta Bay. This condition makes this region has the risk of flooding [1]. Therefore, predicting monthly rainfall is very important for flood control management and minimizing disaster risk in Jakarta.

Currently, machine learning remains developing nonparametric approaches that are widely used in many fields, including rainfall modeling. Machine learning becomes a strategy in rainfall modeling because of its ability to learn flexible data and build models without prior knowledge [2]. Support Vector Machine (SVM) and Artificial Neural Network (ANN) are two examples of the models based on machine learning.

Support vector regression (SVR) is a machine learning-based nonparametric method that does not require assumptions to be more flexible to apply in various fields. The concept of SVR is to maximize...
the hyperplane to get data that becomes a support vector. One advantage of using SVR is that it can resolve overfitting [3]. Several studies have shown the advantages of SVR in many fields, such as using SVR for a statistical downscaling model to predict rainfall in Indramayu [4] and using SVR for modeling the Indonesian Sharia Stock Index [5].

Rainfall has variability, complexity, and uncertainty, so that it usually has nonlinear characteristic series. Therefore, Pre-processing data is needed to eliminate noise to get a better result for modeling purposes. One method to remove noise in time series is Singular Spectrum Analysis (SSA) method. The advantage of the SSA method is that it can find the structure of data components and decompose series data into several components: trend, oscillatory, and noise components.

Previous studies are used SSA for pre-processing, including rainfall modeling with streamflow using a combination of ANN and SSA [6]. The results have shown that combining the ANN and SSA methods provide better predictive results than the ANN method only. They also reported that the application of SSA in pre-processing data for various models could improve the reliability of predictions. SSA-SVM modeling has also been used to predict the number of tourist arrivals in Bali and obtained better results than the single MA and SVM models [7]. SSA and SVR modeling is carried out for rainfall in China with dragonfly optimization [8]. The result has been obtained that integrated SSA and SVR better than the single model. SSA-ARIMA hybrid model conducted on forecasting water demand [9]. The results showed that the aggregate SSA hybrid model better than individuals SSA components hybrid models.

In previous studies in the reconstruction phase of grouping SSA, the determination of group members was carried out subjectively. Kalantari et al. (2019) proposed Auto SSA using hierarchical cluster method to define groups objectively [10]. The results have shown that the single linkage method is tend to be better than others. Therefore, this paper proposes a rainfall prediction model using SVR with pre-processing data based on SSA and Auto SSA. The SSA and Auto SSA remove noise from the original series and reconstruct the series without noise. The SVR uses various kernel functions and a grid search algorithm as an optimization method to estimate SVR model parameters. The result of the hybrid model SSA-SVR and Auto SSA-SVR will be compared to a single SVR model to assess the effectiveness of the proposed model.

2. Material
Data used in this study is the monthly rainfall in Kemayoran Jakarta. The rainfall pattern in Jakarta is a monsoonal type, which means there is a clear difference between the rainy season and the dry season. The rainy season's peak occurs in December-January-February, while the peak of the dry season occurs in July-August-September.

Rainfall data were collected from the Meteorological, Climatological, and Geophysical Agency (BMKG) at Kemayoran rainfall station, including the monthly rainfall data from 1980-2019. The data are divided into training and testing data, such as (1) the training data from 1980 to 2014 and the testing data from 2015 to 2019, (2) the training data from 1980-2015 and the testing data from 2016 to 2019, and (3) the training data from 1980-2016 and the testing data from 2017 to 2019. These three scenarios are conducted to know the consistency of the model predictions.

3. Method
This section discusses a literature review of the methods that be used in this study.

3.1 Singular Spectrum Analysis (SSA)
SSA consists of two stages, namely decomposition and reconstruction. In the decomposition stage, two steps must be taken to obtain eigentriple, namely embedding and singular value decomposition (SVD). For the reconstruction stage, two steps must be taken to get new series, namely grouping and diagonal averaging [11].

First, for the embedding process, transforming one-dimensional data into a series of multidimensional lag data. Suppose the initial time series \( X = (x_1, x_2, \cdots, x_n) \), transformed into a matrix
of size $L \times K$, where $L$ is window length with value is between $2 < L < N/2$. The appropriate $L$ value is gained from an optimization process and $K = N - L + 1$ lag data, $\mathbf{X} = (x_i, x_{i+1}, \ldots, x_{i+L-1})^T$. New series data can be changed into a matrix, which is known as a trajectory matrix, as follows:

$$\mathbf{x} = (x_{ij})_{i,j=1}^{LK} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_K] = \begin{bmatrix} x_1 & x_2 & \cdots & x_K \\ x_2 & x_3 & \cdots & x_{K+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_L & x_{L+1} & \cdots & x_N \end{bmatrix}$$

Then the trajectory matrix $\mathbf{X}$ will be changed into SVD,

$$\mathbf{X}_{lkk} = U_{lkk} \Sigma_{lkk} V_{lkk}^T$$

For instance, $\mathbf{S} = \mathbf{X} \mathbf{X}^T$, and $\lambda_1 \cdots \lambda_L$ is eigenvalue with decreasing order where $\lambda_1 \geq \cdots \geq \lambda_L \geq 0$. $U_1 \cdots U_L$ is eigenvector from each eigenvalue. Suppose rank of $\mathbf{X}$ is denoted by $d$ and $d = \max\{i, \lambda_i > 0\}$. If $V_i = \mathbf{X}^T U_i / \sqrt{\lambda_i}$ for $i = 1 \cdots d$. So that $\mathbf{X}$ can be formulated as follows:

$$\mathbf{X} = \mathbf{X}_1 + \cdots + \mathbf{X}_d,$$  \hspace{1cm} (1)

where $\mathbf{X}_i = \sqrt{\lambda_i} U_i V_i^T$, also called eigentriple.

The decomposed $\mathbf{X}$ then will be grouped into $m$ disjoint subset $I_1 \cdots I_m$. Suppose $I = \{i_1, \ldots, i_p\}$, generated according to a group $I$ which defined as $\mathbf{X}_I = \mathbf{X}_{i_1} + \cdots + \mathbf{X}_{i_p}$, matrix $\mathbf{X}_I$ calculated for grouping $I_1 \cdots I_m$ so that equation (1) can be formulated as:

$$\mathbf{X} = \mathbf{X}_{I_1} + \cdots + \mathbf{X}_{I_m}.$$  \hspace{1cm} (2)

In general, the grouping process is manually identified by a plot of eigenvector and plot of w-correlation. In the one-dimensional plot of eigenvectors, all slowly varying components should be grouped into a trend. In the two-dimensional plot, pairs of eigenvectors that formed regular $P$-vertex polygon identified as a harmonic component. Those that do not include the above pattern are grouped into noise. Then the resulting of the grouping stage can be transformed into new $N$ series data. For instance, equation (1) generates $\mathbf{Y}$ matrix with size $L \times K$,

$$\mathbf{Y} = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1K} \\ y_{21} & y_{22} & \cdots & y_{2K+1} \\ \vdots & \vdots & \ddots & \vdots \\ y_{L1} & y_{L2} & \cdots & y_{LN} \end{bmatrix}$$

suppose $y_{ij}$ an element of $\mathbf{Y}$, and $1 \leq i \leq L, 1 \leq j \leq K$, then $L^* = \min(L, K)$, $K^* = \max(L, K)$, and $N = L + K - 1$. Then step of diagonal averaging is obtained by the formula,

$$y_k = \begin{cases} \frac{1}{K} \sum_{m=1}^{k} y_{m,k-m+1} & \text{for } 1 \leq k \leq L^* \\ \frac{1}{L^*} \sum_{m=1}^{L^*} y_{m,k-m+1} & \text{for } L^* \leq k \leq K^* \\ \frac{1}{N - k + 1} \sum_{m=1}^{N - k + 1} y_{m,k-m+1} & \text{for } K^* \leq k \leq N \end{cases}$$

### 3.2. Single linkage clustering

The grouping process in SSA with an identified by a plot of eigenvectors becomes subjective and then the clustering method proposes to overcome this. The hierarchical clustering method with a single linkage clustering algorithm can be measured from the distance or similarity between pairs of objects. Grouping is formed by the closest distance from the first group member to the other members.
The distance measured in this study is the w-correlation matrix, which shows how good the decomposition and how far the differences between the components in the time series. The formula to obtain w-correlation between two-time series \( Y_1 \) and \( Y_2 \) as follows,

\[
\rho^{(w)}_{12} = \frac{\sum_{i=1}^{N} w_i x^{(1)}_i y^{(2)}_i}{\left\| x^{(1)}_i \right\|_w \left\| y^{(2)}_i \right\|_w}
\]

\( w_i = \min(i, L, N - 1 + 1) \), \( \left\| y^{(k)}_i \right\|_w = \sqrt{\sum_{i=1}^{N} w_i y^{(k)}_i y^{(k)}_i}, k = 1, 2 \)

then the distance between the two components is defined as \( d_{ij} = 1 - \left| \rho^{(w)}_{ij} \right| \).

3.3. Support Vector Regression

Support vector machine initially applied to binary classification and pattern recognition. After being introduced Vapnik’s \( \epsilon \)-insensitive loss function, SVM has been extended to solve nonlinear regression estimation and applied prediction in many fields [12]. Smola et al. (2004) explain SVR by given training set data \( \{(x_i, y_i), ..., (x_n, y_n)\} \subset X \times \mathbb{R} \), where \( X \) is input space \( (X = R^d) \). Let \( f(x) \) is linear function (hyperplane) [13]:

\[
f(x) = w'^t x + b, \ w \in X, b \in \mathbb{R} \tag{3}
\]

where \( w \) is weight vector, \( b \) is bias, and \( x \) is a vector of training set data. Vectors are optimally separated by hyperplane if separated without error, and the distance between the closest vectors is maximal. Flatness for functions in equation (3) means searching for a minimum value of \( w \). It can be done by minimizing norm \( w \), \( \|w\|^2 = w'^t w \). This can be written in the optimization problem as follows,

\[
\min \frac{1}{2} \|w\|^2 \tag{4}
\]

with constraint

\[
y_i - w'^t x - b \leq \epsilon, i = 1, 2... n
\]

\[
w'^t x + b - y_i \leq \epsilon, i = 1, 2... n
\]

Slack variable \( \xi \) and \( \xi^* \) can be used to solve the observed value out of bounds \( \epsilon \). Then equation (4) can be formed as

\[
\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{\xi} (\xi_i + \xi^*_i) \tag{5}
\]

With constraint: \( y_i - w'^t x - b \leq \epsilon + \xi_i \)

\[
w'^t x + b - y_i \leq \epsilon + \xi^*_i
\]

\( \xi_i, \xi^*_i \geq 0 \)

C > 0 is the trade-off parameter between the flatness of the function \( f \) and limit \( \epsilon \). All outside periods will be multiplied by C. This is under the following \( \epsilon \)-insensitive loss function,

\[
L_\epsilon(y) = \begin{cases} 0, & f \text{ or } |f(x) - y| < \epsilon \\ |f(x) - y| - \epsilon, & \text{for others} \end{cases}
\]

Equation (5) is solved using the Lagrangian form

\[
L = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{\xi} (\xi_i + \xi^*_i) - \sum_{d=1}^{\eta} \eta_d (\xi_i + \eta^*_i) - \sum_{i=1}^{\xi} \alpha_i (\epsilon + \xi_i - y_i + w'^t x + b) - \sum_{i=1}^{\xi} \alpha^*_i (\epsilon + \xi^*_i + y_i - w'^t x - b) \tag{6}
\]
With the applied Karush-Kuhn-Tucker condition, equation 6 can be done as

\[
\text{Max} \ - \frac{1}{2} \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) k(x_i, x_j) - \varepsilon \sum_{i=1}^{n} (\alpha_i + \alpha_i^*) + \sum_{i=1}^{n} y_i (\alpha_i - \alpha_i^*)
\]

(7)

with condition,

\[\sum_{i=1}^{n} y_i (\alpha_i - \alpha_i^*) = 0 \text{ dan } \alpha_i, \alpha_i^* \in [0, C].\]

\(k(x_i, x_j)\) is the kernel function. Optimal weight vector obtained for the hyperplane is

\[w = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) x_i,\]

So that the regression model can be written:

\[f(x) = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) k(x_i, x_j) + b, \; i, j = 1, 2 \ldots n\]

SVR uses a kernel function to transform nonlinear input into a feature space which dimension is higher and then solved linearly. Performance of SVR determined by choice of kernel type and its parameters. The grid search algorithm is used to determine the optimal parameter kernel in the SVR model. Kernel functions that are used in this study as follows:

a. linear kernel

\[k(x_i, x_j) = x_i, x_j,\]

b. radial basis function (RBF) kernel

\[k(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \text{ where } \gamma = \frac{1}{2\sigma} \text{ is kernel parameter.}\]

### 3.4. Model Evaluation

To determine the best model, root mean square error of prediction (RMSEP) and Pearson correlation coefficient (r) is used.

\[\text{RMSEP} = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n}}\]

\[r = \frac{n \sum y_i \hat{y}_i - \sum y_i \sum \hat{y}_i}{\sqrt{(n \sum y_i^2 - (\sum y_i)^2)(n \sum \hat{y}_i^2 - (\sum \hat{y}_i)^2)}}\]

where \(y_i, \hat{y}_i\) are the observed and estimated values, respectively. The best model was chosen with a correlation coefficient value close to 1 and RMSE value close to 0.

### 3.5. Proposed hybrid model SSA-SVR and Auto SSA SVR

The hybrid model combines two different methods, such as SSA and SVR or Auto SSA and SVR. SSA or Auto SSA is used for pre-processing data to denoising the original series, and SVR is used for modeling prediction of the reconstruction data from SSA. In the SSA method, the grouping is identified manually, while in the Auto SSA, the step of grouping is done with single-linkage clustering to avoid subjectively. This model is designed to improve forecasting performance and reliability. The result of the hybrid model is compared to the single SVR model.

The steps of data analysis in this research:

**Step 1**: Data exploration

**Step 2**: Pre-processing using SSA and Auto SSA with single linkage clustering.

**Step 3**: Identification input variable for hybrid models using PACF from reconstruction series SSA and Auto SSA, and identification input variable for single model SVR using PACF from original series.
Step 4: Modeling hybrid SVR from the reconstruction series SSA and Auto SSA, and modeling SVR from the original series with two kernel types and optimization parameters using grid search algorithm.

Step 5: Compute RMSEP and correlation coefficient for each model to determine the best model.

4. Result and Discussion
Descriptive statistics were used to understand the characteristic of monthly rainfall in Jakarta. Based on data exploration, rainfall in Jakarta has a high variance indicated by a large value of standard deviation. Several global and local anomalies influence this condition. The skewness value was obtained 1.769, indicating that data is not normally distributed and skewed to the right. Monthly rainfall in Jakarta also suggested that data is not stationary in a variant.

The highest amount of rainfall in Kemayoran station Jakarta occurs in January-February. This period usually becomes the peak of the rainy season in Jakarta. During 1980-2019 there were several rainfall anomalies in Jakarta, which were indicated by outliers that produced extreme rainfall. It occurred in January 2014 and February 2015, reaching 925 mm/month and 939.5 mm/month. The smallest precipitation or peak of the dry season in Jakarta occurs in August-September, with the amount of rainfall is 0 mm/month.

Based on data exploration, the original rainfall data is pre-processed with SSA to decompose it into several components such as trend, seasonality, and noise by determining window length with several values to obtain the optimal value. The optimal window length $L$ used is 24 with RMSE 69.11. After the window length parameter is acquired, then trajectory matrix $X$ will be formed with size 24x457. The next step is singular value decomposition, which will decompose it into 24 eigentriples. The SSA grouping process is obtained from figure 4.1 with pairs of eigenvectors plots. Figure 4.1 shows 1st eigentriple (ET1) as a trend, then ET2-3, ET4-5, ET8-9, ET10-11 pairs of eigentriples as a harmonic component with periodicity 12, 6, 4, 3, respectively. Then diagonal averaging conduct to reconstruct new series from each group component. The result of reconstruction from each group component is then aggregated without noise to get a new series free from noise. Aggregated reconstruction series SSA without noise compared to the original series, as shown in figure 4.2.

The pre-processing step with Auto SSA is the same as the previous SSA but differs in group members determination. Window length parameter $L$ used is 24 with identification of group using single-linkage clustering. The plot of dendrogram in figure 4.3 shows that ET1 as a trend, ET2-3, ET4-5, ET6-7, ET8-9, ET12-13 as harmonic components, and later as noise. A plot of aggregated reconstructed series from Auto SSA without noise is shown in figure 4.4.

Compare the reconstruction series from SSA and Auto SSA obtained that the reconstruction series from Auto SSA closer to the original series pattern with RMSE 64.449 while from SSA has RMSE 69.11. The percentage of variance used in the reconstruction series from Auto SSA is 94.27%, while the reconstruction series from SSA is 86.84%. This reconstruction data series then used for modeling with SVR.

In SVR modeling, there are input and output. Where the output is the current rainfall, and the input is the previous rainfall data. Pacf plot is used to determine the input by looking at the significant lag from the reconstruction series. Pacf plot from SSA reconstruction in figure 4.5 show there are 16 significant lags that is lag 1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 16, 19, 23, 24, while pacf from Auto SSA reconstruction in figure 4.6 show there are 15 significant lags that is lag 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 16, 17. These significant lags will be used as input in SVR modeling.

Selecting the kernel function and determine the optimal parameter is vital for the SVR model; therefore, the grid search algorithm is used for optimization. The range of parameter values will be optimized for each kernel function as follow,

a. Linear kernel, range cost parameter (C): $2^2$, $2^4$, ..., $2^6$ and range epsilon parameter ($\varepsilon$): 0.001, 0.01, 0.1
b. Radial Basis Function (RBF) kernel, range cost parameter (C): $2^{-5}$, $2^{-4}$, $2^{-3}$, $2^{-2}$, $2^{-1}$, range gamma parameter: $2^{-8}$, $2^{-7}$, $2^{-6}$, range epsilon parameter ($\epsilon$): 0.001, 0.01, 0.1.

---

**Figure 4.1.** Plot pairs of eigenvector

**Figure 4.2.** A plot of original series (black line) and reconstructed series (red line) SSA

**Figure 4.3.** Dendrogram from single-linkage clustering

**Figure 4.4.** A plot of original series (black line) and reconstructed series (red line) Auto SSA

**Figure 4.5.** A plot of PACF from reconstructed series SSA

**Figure 4.6.** A plot of PACF from reconstructed series Auto SSA

Evaluation of the best models between hybrid model SSA-SVR, Auto SSA-SVR, and single model SVR are shown in table 4.1. It is determined by the smallest RMSEP and the greatest correlation. RMSEP from the hybrid model in various time scenarios is smaller than the single SVR model. The correlation coefficient of the hybrid model is greater than the single SVR model. It shows the advantages of a hybrid model with pre-processing data than a single model SVR without pre-processing. In the first scenario data sets, the best model obtained from the hybrid model Auto SSA-SVR with linear kernel function has RMSEP 79.286 and a correlation coefficient of 0.882. The second scenario data sets, the best model obtained from the hybrid model Auto SSA-SVR with a linear kernel function with RMSEP 68.146 and a correlation coefficient of 0.849. In the third scenario data sets, the best model obtained from hybrid model Auto SSA-SVR with linear kernel function with RMSEP 64.908 and correlation coefficient 0.864. these results show the advantage of the hybrid model Auto SSA-SVR with single
linkage clustering than the hybrid model SSA-SVR. A linear kernel function is superior in producing the best model than an RBF kernel. This result is relatively consistent and provides effective hybrid models compared to a single model.

A plot of comparison between hybrid model SSA-SVR, Auto SSA-SVR, and single models SVR with linear kernel during testing data 2017-2019 can be seen in figure 4.7. The plot of the Auto SSA-SVR model is closer to the original data pattern than others, while single SVR models tend to underestimate.

![Plot monthly rainfall 2017-2019](image)

**Figure 4.7. A compared plot of the hybrid model and single model SVR**

#### Table 4.1. Evaluation of the best models between SSA-SVR, Auto SSA-SVR, and SVR

| Data sets | kernel | Model | Parameter | RMSEP | r | Parameter | RMSEP | r | Parameter | RMSEP | r |
|-----------|--------|-------|-----------|-------|---|-----------|-------|---|-----------|-------|---|
| 1980-2014 (training) | Linear | SSA-SVR | C = 64 | 85.454 | 0.8599 | C = 4 | 79.286 | 0.882 | C = 0.5 | 122.755 | 0.715 |
| 2015-2019 (testing) | RBF | | C = 256 | 81.318 | 0.875 | C = 128 | 86.899 | 0.858 | C = 2 | 123.473 | 0.722 |
| | | | γ = 0.0039 | e = 0.01 | | γ = 0.0039 | e = 0.01 | | γ = 0.01 | |
| 1980-2015 (training) | Linear | Auto SSA-SVR | C = 16 | 78.891 | 0.7920 | C = 2 | 68.146 | 0.849 | C = 0.5 | 105.720 | 0.611 |
| 2016-2019 (testing) | RBF | | C = 256 | 75.023 | 0.822 | C = 256 | 71.955 | 0.835 | C = 8 | 94.080 | 0.7052 |
| | | | γ = 0.0039 | e = 0.01 | | γ = 0.0078 | e = 0.01 | | γ = 0.0156 | e = 0.001 | |
| 1980-2016 (training) | Linear | SVR | C = 16 | 68.158 | 0.862 | C = 2 | 64.908 | 0.864 | C = 8 | 97.711 | 0.652 |
| 2017-2019 (testing) | RBF | | C = 256 | 70.846 | 0.858 | C = 128 | 65.384 | 0.860 | C = 23 | 86.471 | 0.743 |
| | | | γ = 0.0039 | e = 0.01 | | γ = 0.0039 | e = 0.01 | | γ = 2.6 | e = 0.001 | |

5. Conclusion
This research proposed a hybrid SSA-SVR and Auto SSA-SVR model for rainfall prediction. Based on the result showed that the hybrid model yielded more accurate than the single model. This means that pre-processing with SSA decomposition can improve the accuracy of the prediction. Comparing the
RMSEP value from each hybrid model in all scenarios obtained, the best model is the hybrid Auto SSA-SVR model using a linear kernel function. The best model is received in the third scenario with the smallest RMSEP of 64,908 and a correlation coefficient of 0.864.

For further research, it can be developed using another linking method or a non-hierarchical cluster method for grouping step of Auto SSA, also other nonparametric models can be used to be combined with SSA for rainfall prediction.

Acknowledgment
The correspondence author would like to thank the Ministry of Research and Technology/ National Research and Innovation Agency for the SAINTEK scholarship to study postgraduate in Statistics Department of IPB University.

References

[1] BPBD DKI Jakarta 2016 *Rencana Kontinjeni Banjir DKI Jakarta 2016* (Jakarta)
[2] Mehr D, Nourani V, Khosrowshahi K and Ghorbani M 2019 A hybrid support vector regression–firefly model for monthly rainfall forecasting *International Journal of Environmental Science and Technology* 16 (1) 335–346
[3] Yasin H, Prahutama A and Utami T 2014 Prediksi harga saham menggunakan support vector regression dengan algoritma grid search *Media Statistika* 7(1) 29–35
[4] Agmalaro M 2011 Statistical downscaling of gcm data using support vector regression to predict monthly rainfall in indramayu (Bogor: IPB) [thesis]
[5] Saputra G, Wigena A and Sartono B 2019 Penggunaan support vector regression dalam pemodelan indeks saham syariah indonesia dengan algoritme grid search *Indonesian Journal of Statistics and Its Applications* 3 (2) 148 – 160
[6] Wu C and Chau K 2011 Rainfall-runoff modeling using artificial neural network coupled with singular spectrum analysis *Journal of Hydrology* 399 394–409
[7] Sitohang Y, Andriyana Y and Chadidjah A 2018 The forecasting technique using SSA-SVM applied to foreign tourist arrivals to bali *Telkomnika* 16 (4) 1679–87
[8] Li W and Zhang J 2019 An innovated integrated model using singular spectrum analysis and support vector regression optimized by intelligent algorithm for rainfall forecasting *Journal of Autonomous Intelligence* 2 (1) 46
[9] Suhartono, Isnawati S, Salehah N, Prastyo D, Kuswanto H and Lee M 2018 Hybrid SSA-TSR-ARIMA for water demand forecasting *International Journal of Advance in Intelligent Informatics* 4 (3) 238-250
[10] Kalantari M and Hassani H 2019 Automatic grouping in singular spectrum analysis *Forecasting* 1 (1) 189–204
[11] Golyandina N and Zhigliavskii A 2013 *Singular Spectrum Analysis for Time Series* (London: Springer)
[12] Chapelle O, Vapnik V, Bousquet O and Mukherjee S 2002 Choosing multiple parameters for support vector machines *Machine Learning* 46 131–59.
[13] Smola A and Schölkopf B 2004 A tutorial on support vector regression *Statistics and Computing* 14 199-222