A Hybrid Singular Spectrum Analysis and Neural Networks for Forecasting Inflow and Outflow Currency of Bank Indonesia

Suhartono[1](✉), Endah Setyowati[1], Novi Ajeng Salehah[1], Muhammad Hisyam Lee[2], Santi Puteri Rahayu[1], and Brodjol Sutijo Suprih Ulama[1]

[1] Department of Statistics, Institut Teknologi Sepuluh Nopember, Kampus ITS Sukolilo, Surabaya 60111, Indonesia
suhartono@statistika.its.ac.id

[2] Department of Mathematical Science, Universiti Teknologi Malaysia (UTM), 81310 Skudai, Johor, Malaysia

Abstract. This study proposes hybrid methods by combining Singular Spectrum Analysis and Neural Network (SSA-NN) to forecast the currency circulation in the community, i.e. inflow and outflow. The SSA technique is applied to decompose and reconstruct the time series factors which including trend, cyclic, and seasonal into several additive components, i.e. trend, oscillation and noise. This method will be combined with Neural Network as nonlinear forecasting method due to inflow and outflow data have non-linear pattern. This study also focuses on the effect of Eid ul-Fitr as calendar variation factor which allegedly affect inflow and outflow. Thus, the proposed hybrid SSA-NN is evaluated for forecasting time series that consist of trend, seasonal, and calendar variation patterns, by using two schemes of forecasting process, i.e. aggregate and individual forecasting. Two types of data are used in this study, i.e. simulation and real data about the monthly inflow and outflow of 12 currency denominations. The forecast accuracy of the proposed method is compared to ARIMAX model. The results of the simulation study showed that the hybrid SSA-NN with aggregate forecasting yielded more accurate forecast than individual forecasting. Moreover, the results at real data showed that the hybrid SSA-NN yielded as good as ARIMAX model for forecasting of 12 inflow and outflow denominations. It indicated that the hybrid SSA-NN could not successfully handle calendar variation pattern in all series. In general, these results in line with M3 competition conclusion, i.e. more complex methods do not always yield better forecast than the simpler one.

Keywords: Singular spectrum analysis · Neural network · Hybrid method Inflow · Outflow

1 Introduction

The currency has a very important role for the Indonesian economy. Although non-cash payment system has grown rapidly, currency or cash payment is still more efficient for individual payment for small nominal value. Forecasting inflow and outflow can be an

B. W. Yap et al. (Eds.): SCDS 2018, CCIS 937, pp. 3–18, 2019.
https://doi.org/10.1007/978-981-13-3441-2_1

option to maintain the stability of currency. The prediction of the amount of currency demand in Indonesia is often referred as the autonomous liquidity factor, so in predicting the demand for currency by society will be difficult [1]. The development of inflow and outflow of currency both nationally and regionally has certain movement patterns influenced by several factors, such as government money policy. Moreover, it is also influenced by trend, seasonal and calendar variation effect caused by Eid ul-Fitr that usually occurred at different date in each year [2]. Decomposition of time series data into sub patterns can ease the process of time series analysis [3]. Hence, a forecasting method that could capture and reconstruct each component pattern in the data was needed.

This study proposes a forecasting method that combining Singular Spectrum Analysis as decomposition method and Neural Network (known as SSA-NN) for forecasting inflow and outflow data in both scheme, i.e. individual and aggregate forecasting. The SSA method was applied to decompose and reconstruct the time series patterns in inflow and outflow data which including trend, cyclic, and seasonal into several additive components, while the NN method is used to handling non-linear pattern which contained in inflow and outflow data. In addition, this study also focuses to learn whether the SSA-NN could handle calendar variation effect in time series, particularly the effect of Eid ul-Fitr to inflow and outflow data.

As widely known, SSA is a forecasting technique that combines elements of classical time series analysis, multivariate statistics, multivariate geometry, dynamic systems, and signal processes [4]. SSA can decompose common patterns in time series data, trend, cycle, and seasonal factors into some additive components separated by trend, oscillatory, and noise components. SSA was first introduced by Broomhead and King [5] and followed by many studies that applied this method [6, 7]. SSA has a good ability in characterizing and prediction of time series [8]. Furthermore, it is also known that SSA method can be used for analyzing and forecasting short time series data with various types of non-stationary and produce more accurate forecast [9, 10].

In the past decades, many researchers have increasingly developed SSA by combining this method with other forecasting methods. Hybrid SSA model tends to be more significant and provides better performance than other methods [12]. A combination of SSA and NN could forecast more accurately and it could effectively reconstruct the data [13, 14]. A comparative study was done by Barba and Rodriguez [15] also showed that SSA-NN produced better accuracy for multi-step ahead forecasting of traffic accident data. The rapid research development about the combinations of SSA showed that this method can improve forecasting performance and could be a potential and competitive method for time series forecasting [16–18].

In this study, two types of data are used, i.e. simulation and real data about the monthly inflow and outflow data of 12 banknotes denomination from January 2003 to December 2016. These data are secondary data obtained from Bank Indonesia. The data are divided into two parts, i.e. training data (from January 2003 to December 2014) and testing data (from January 2015 to December 2016). The forecast accuracy of the proposed method is compared to ARIMAX model by using RMSE, MAE and MAPE criteria. The results of the simulation study showed that the hybrid SSA-NN with aggregate forecasting scheme yielded more accurate forecast than individual forecasting scheme. Moreover, the results at real data showed that the hybrid SSA-NN yielded as good as ARIMAX model for forecasting of 12 inflow and outflow denominations. It indicated that the hybrid SSA-NN could not handle calendar variation pattern in all data series. Generally, these results in line with M3 competition results which concluded that more complex methods do not always yield better forecast than the simpler one.

The rest of paper is organized as follows: Sect. 2 reviews the methodology, i.e. ARIMAX, Singular Spectrum Analysis, and Neural Networks as forecasting method; Sect. 3 presents the results and analysis; and Sect. 4 presents the conclusion from this study.

2 Materials and Methods

2.1 ARIMAX

The ARIMAX is an ARIMA model with the addition of exogenous variables. This model has a similar form with linear regression that has additional variables such as trend, seasonal, and calendar variation factors, or other explanatory variables. The ARIMAX model which consists of linear trend (represented by t variable), additive seasonal (represented by $M_{i,t}$ variables), and calendar variation pattern (represented by $V_{j,t}$ variables) is written as follows:

$$Y_t = \beta_0 + \beta_1 t + \sum\nolimits_{i=1}^{I} \gamma_i M_{i,t} + \sum\nolimits_{j=1}^{J} \delta_j V_{j,t} + N_t \tag{1}$$

where $M_{i,t}$ is dummy variables for I seasonal effects, $V_{j,t}$ is dummy variables for J calendar variation effects, and N_t is noise variable that follows ARMA model. The identification of calendar variation effect, particularly about the duration effect, can be done graphically by using time series plot [19].

2.2 Singular Spectrum Analysis (SSA)

SSA is a forecasting method that combines elements of classical forecasting, multivariate statistics, multivariate geometry, dynamic systems, and signal processing. SSA method does not require the fulfillment of statistical assumptions such as stationary, and ergodicity. The main objective of the SSA method is to decompose the original time series into several additive components, such as trend, oscillatory, and noise components [4]. In general, SSA has two main stages as follows:

a. Decomposition (Embedding and Singular Value Decomposition)

The procedure in embedding is to map the original time series data into a multidimensional sequence of lagged vector. Let's assume L is an integer number represents window length with $1 < L < n$, the formation of lagged vectors where $K = n - L + 1$ is

$$Y_i = (f_i, f_{i+1}, \ldots, f_{i+L-1})^T, 1 \le i \le K \tag{2}$$

which has a dimension of L. If the dimensions of Y_i are emphasized, then Y_i is referred as L-lagged vectors. The path matrix of the F series is illustrated as follows:

$$\mathrm{Y} = [Y_i : \ldots : Y_K] = \begin{bmatrix} f_1 & f_2 & f_3 & \cdots & f_K \\ f_2 & f_3 & f_4 & \cdots & f_{K+1} \\ f_3 & f_4 & f_5 & \cdots & f_{K+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ f_L & f_{L+1} & f_{L+2} & \cdots & f_n \end{bmatrix} \tag{3}$$

Let $\mathbf{S} = \mathbf{Y}\mathbf{Y}^{\mathbf{T}}$ and $\lambda_1, \lambda_2, \ldots, \lambda_L$ be the eigenvalues of the matrix $\mathbf{S}$ Where $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_L \geq 0$ and $U_1, U_2, \ldots, U_L$ are eigenvectors of the matrix $\mathbf{S}$ corresponding to the eigenvalues. Note that $d = \max\{i\}$ so that $\lambda_i > 0$ is the rank of the matrix $\mathbf{Y}$. If $\mathbf{S} = V_i = \mathbf{Y}^{\mathbf{T}} U_i / \sqrt{\lambda_i}$ for $i = 1, 2, \ldots, d$, then SVD of the path matrix $\mathbf{Y}$ can be written as

$$\mathbf{Y} = \mathbf{Y}_1 + \mathbf{Y}_2 + \ldots + \mathbf{Y}_d \tag{4}$$

where $\mathbf{Y}_i = \sqrt{\lambda_i} U_i V_i^{\mathbf{T}}$. Matrix $\mathbf{Y}_i$ has rank 1 and often called as an elementary matrix. The set $\left(\sqrt{\lambda_i}, U_i, V_i\right)$ is called i-th eigentriple to SVD.

b. Reconstruction (Grouping and Diagonal Averaging)

After SVD equation is obtained, the grouping procedure will partition the set of indices $\{1, 2, \ldots, d\}$ into m subsets of mutually independent, $I_1, I_2, \ldots, I_m$. Let $I = \{i_1, i_2, \ldots, i_p\}$, the resulting $\mathbf{Y}_i$ matrix corresponds to group I defined as a matrix with $\mathbf{Y}_I = \mathbf{Y}_{i_1} + \mathbf{Y}_{i_2} + \ldots + \mathbf{Y}_{i_p}$. This matrix is calculated for groups $I = I_1, I_2, \ldots, I_m$ and this step will lead to decomposition form as follows:

$$\mathbf{Y} = \mathbf{Y}_{I_1} + \mathbf{Y}_{I_2} + \ldots + \mathbf{Y}_{I_m}. \tag{5}$$

Set m selection procedures, $I_1, I_2, \ldots, I_m$ are called eigentriple groupings. If $m = d$ and $I_j = \{j\}$, $j = 1, 2, \ldots, d$, then the corresponding grouping is called elementary.

Let $\mathbf{Z}$ be $L \times K$ matrix with element z_{ij}, $1 \leq i \leq L$, $1 \leq j \leq K$ for $L \leq K$. Let's assume the values of $L^* = \min\{L, K\}$, $K^* = \max\{L, K\}$ and $n = L - K - 1$. If $L < K$ then $z_{ij}^* = z_{ij}$, and if $L > K$ then $z_{ij}^* = z_{ji}$. Diagonal averaging moves the $\mathbf{Z}$ matrix to the series $g_1, g_2, \ldots, g_n$ by following formula:

$$g_k = \begin{cases} \frac{1}{k}\sum_{m=1}^{k} z_{m,k-m+1}^* & \text{for } 1 \leq k < L^* \\ \frac{1}{L^*}\sum_{m=1}^{L^*} z_{m,k-m+1}^* & \text{for } L^* \leq k < K^* \\ \frac{1}{n-k+1}\sum_{m=k-K^*+1}^{n-K^*+1} z_{m,k-m+1}^* & \text{for } K^* \leq k < n \end{cases} \tag{6}$$

This equation corresponds to the average matrix element over the 'antidiagonals' $i + j = k + 1$. If the averaging diagonal is applied to the matrix Y_{I_k}, then this process will obtain a reconstructed series $F^{(k)} = \left(f_1^{(k)}, f_1^{(k)}, \ldots, f_1^{(k)}\right)$. Therefore, the initial series $f_1, f_2, \ldots, f_n$ are decomposed into a sum of the m reconstructed series, i.e.

$$f_j = \sum_{k=1}^{m} f_j^{(k)}, j = 1, 2, \ldots, n \tag{7}$$

2.3 Neural Networks

The most commonly used form of neural network architecture (NN) is Feedforward Neural Networks (FFNN). In statistical modeling, FFNN can be viewed as a flexible class of nonlinear functions. NNs has several unique characteristics features such as its adaptability, nonlinearity, arbitrary function mapping ability – make this method quite suitable and useful for forecasting tasks [20]. In general, this model works by accepting a vector from input x and then compute a response or output $\hat{y}(x)$ by processing (propagating) x through interrelated process elements. In each layer, the inputs are transformed into layers using a nonlinear form, and it will be processed forward to the next layer. Finally, the output values $\hat{y}$, which can be either scalar or vector values, are calculated on the output layer [21]. FFNN architecture with a hidden layer consisting of q unit neurons and output layer consisting only of one unit of neuron is shown as Fig. 1.

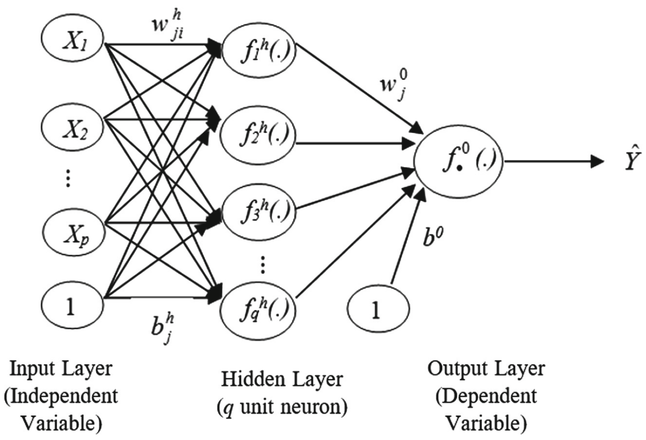

Fig. 1. FFNN architecture with one layer hidden, p input unit, q unit of neuron in hidden layer, and one output neuron unit

The response or output $\hat{y}$ values are calculated by:

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

where f_j^h is activation function in the j-th neuron in the hidden layer, and f^0 is the activation function of the neuron in the output layer.

2.4 Hybrid Singular Spectrum Analysis and Neural Network

In general, SSA method is able to decompose a data series into trend, seasonal, and noise patterns. From the decomposition of data patterns, the forecasting will be done using NN with inputs are the lags of component or known as Autoregressive Neural Networks (ARNN). Forecasting can be used either individual or aggregate scheme. Individual forecasting is done by forecasting every major component formed without combining as a trend and seasonal. Specifically, the noise components will be always modelled in aggregate scheme.

Aggregate forecasting is done by summing the components that have same pattern. Thus, the forecast value is calculated from three main patterns, i.e. trend, seasonal, and noise. Then, the results of forecasting with ARNN on individual patterns will be summed to get the forecast of main series (forecast aggregation). These procedure stages are shown in Figs. 2 and 3 for individual and aggregate scheme, respectively.

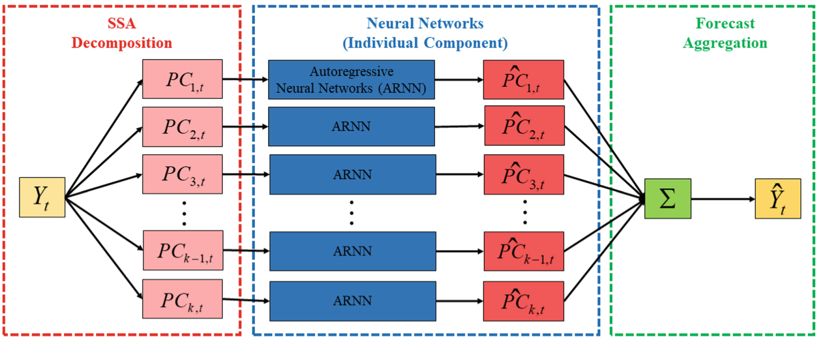

Fig. 2. SSA-NN forecasting using individual forecasting

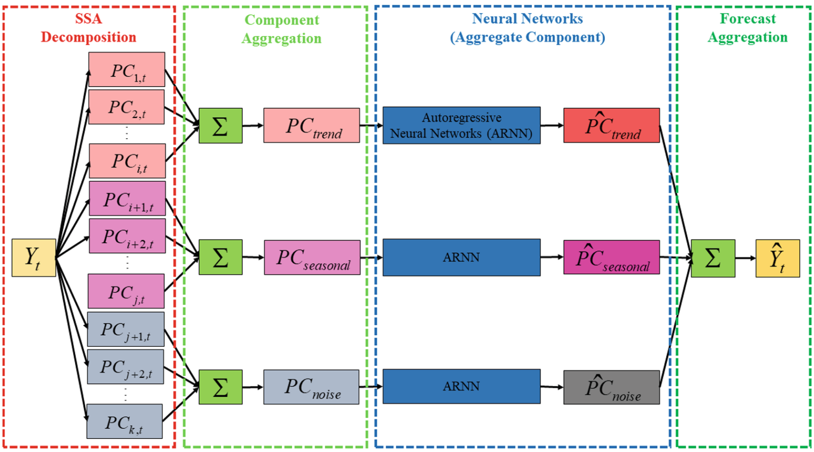

Fig. 3. SSA-NN forecasting using aggregate forecasting

The algorithm of SSA-NN has several steps as follows:

a. Data series decomposition with SSA
   i. Embedding
   ii. Singular Value Decompotition (SVD)
   iii. Grouping
   iv. Diagonal Averaging
b. Modeling the decomposition results using the NN method.
   i. Determine the input variables in NN based on the significant lags of the Partial Autocorrelation Function or PACF of stationary data [22].
   ii. Conduct a nonlinearity test using Terasvirta test.
   iii. Determine the number of units on hidden layer using cross validation method.
   iv. Estimate parameters/weights of NN by using backpropagation algorithm.
   v. Forecast the testing data.
c. Summarize the results of the forecast at each component to get the forecast of testing data.
d. Calculate the level of forecasting errors for testing data.
e. Forecast data by using the corresponding NN model for each data component.

2.5 Model Evaluation

Cross-validation is used for model evaluation, which focusing only on the forecast results for out-sample or testing data [23]. The model evaluation will be done based on the accuracy of the forecast by using RMSE, MAE and MAPE which shown in following equation [24], where C is the forecast period:

$$RMSE = \sqrt{\frac{1}{C}\sum_{c=1}^{C}\left(Y_{n+c} - \hat{Y}_n(c)\right)^2} \tag{2}$$

$$MAE = \frac{1}{C}\sum_{c=1}^{C}\left|Y_{n+c} - \hat{Y}_n(c)\right| \tag{3}$$

$$MAPE = \frac{1}{C}\sum_{c=1}^{C}\left|\frac{Y_{n+c} - \hat{Y}_n(c)}{Y_{n+c}}\right| \times 100\%. \tag{4}$$

3 Results

3.1 Simulation Study

Inflow and outflow currency data are suspected to contain trend, seasonal patterns and influenced by certain calendar variations. To gain better knowledge and understanding about the proposed SSA-NN method, a simulation study was conducted by assuming the data are observed on the period from January 2001 to December 2016 or have 192 observations. In this simulation study, data were generated for each component of trend, seasonal, calendar variation patterns as well as random and non-random noise (has nonlinear pattern) as follows:

a. Trend, $T_t = 0.2t$
b. Seasonal,

$$M_t = 20M_{1,t} + 23.7M_{2,t} + 25M_{3,t} + 23.7M_{4,t} + 20M_{5,t} + 15M_{6,t} + 10M_{7,t} + 6.3M_{8,t} + 5M_{9,t} + 6.3M_{10,t} + 10M_{11,t} + 15M_{12,t}$$

c. Calendar Variation,

$$V_t = 65V_{1,t} + 46V_{2,t} + 47V_{3,t} + 18V_{4,t} + 28V_{1,t+1} + 23V_{2,t+1} + 41V_{3,t+1} + 60V_{4,t+1}$$

d. Linear Noise Series (white noise assumption is fulfilled), $N_{1,t} = a_t$, where $a_t \sim IIDN(0,1)$
e. Nonlinear Noise Series which follow ESTAR(1) model, $N_{2,t} = 6.5N_{2,t-1}.\exp\left(-0.25N_{t-1}^2\right) + a_t$, where $a_t \sim IIDN(0,1)$.

There are two scenarios of simulation series that following equation,

$$Y_t = T_t + M_t + V_t + N_t$$

where the scenario 1 consisting of trend, seasonal, calendar variation and noise that fulfill white noise, and the scenario 2 containing of trend, seasonal, calendar variation and noise that follow nonlinear ESTAR model. Both scenarios are used to evaluate the performance of SSA in handling all these patterns, particularly calendar variation effect pattern. The time series plot of simulation data are shown in Fig. 4.

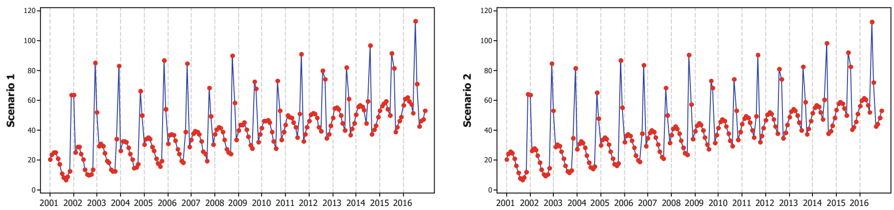

Fig. 4. Time series plot of the scenario 1 and 2 of simulation data

Decomposition results with SSA indicate that the effects of calendar variation could not be decomposed on their own. The results show that the effects of calendar variations on aggregate forecasting are captured into seasonal components and partly as noise components. Furthermore, both individual and aggregate data are modeled by using NN and it will be summed to obtain the forecast of data-testing. The model evaluation of individual and aggregate forecasting is shown in Table 1.

Table 1. Model evaluation using individual and aggregate forecasting in simulation data

Method	Scenario 1			Scenario 2		
	RMSE	MAE	MAPE	RMSE	MAE	MAPE
Aggregate	8.17	6.63	10.5	8.64	7.23	11.8
Individual	8.60	7.18	11.4	8.73	7.52	12.4

Table 1 shows that in this simulation study, an aggregate forecasting has better results than an individual forecasting. It can be seen from the RMSE, MAE and MAPE of aggregate method are smaller than the individual method, both in scenario 1 and 2. Based on Table 1, it could be concluded that SSA-NN method yield better forecast on data containing random noise than data containing nonlinear noise.

3.2 Inflow and Outflow Data

The data that be used as case study are the monthly inflow and outflow data of banknotes per denomination from January 2003 to December 2016. These data are secondary data that be obtained from Bank Indonesia. The data are divided into training data (from January 2003 to December 2014) and testing data (from January 2015 to December 2016). The description of the data is shown at Table 2.

Table 2. Research variable (in billion IDR)

Inflow		Outflow	
Variable	Denomination	Variable	Denomination
$Y_{1,t}$	Rp2.000,00	$Y_{7,t}$	Rp2.000,00
$Y_{2,t}$	Rp5.000,00	$Y_{8,t}$	Rp5.000,00
$Y_{3,t}$	Rp10.000,00	$Y_{9,t}$	Rp10.000,00
$Y_{4,t}$	Rp20.000,00	$Y_{10,t}$	Rp20.000,00
$Y_{5,t}$	Rp50.000,00	$Y_{11,t}$	Rp50.000,00
$Y_{6,t}$	Rp100.000,00	$Y_{12,t}$	Rp100.000,00

The pattern of inflow and outflow of currency at Indonesia (National) from January 2003 until December 2016 shown in Fig. 5. The national inflow and outflow in Indonesia has generally fluctuated, although it declined in 2007 due to the implementation of Bank Indonesia's new policy on deposits and payments to banks. While starting in 2011, the inflow and outflow data increase due to imposition of deposits and withdrawals. In general, the increasing value of national inflow and outflow is high in certain months occurred as the effect of calendar variations, i.e. Eid ul-Fitr. The Eid ul-Fitr is suspected to affect certain months in both inflow and outflow data. In addition, Eid ul-Fitr that occur on different week will also give different impact on the increasing amount of inflow and outflow.

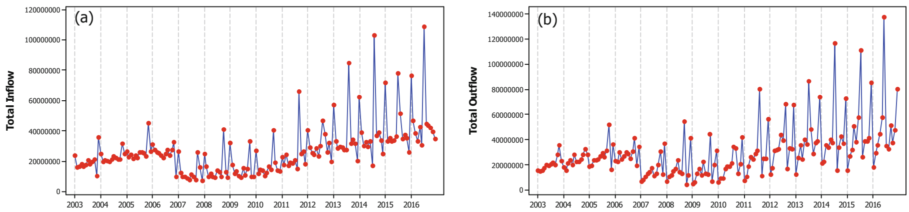

Fig. 5. Inflow (a) and outflow (b) in Indonesia (billion IDR)

The effect of Eid ul-Fitr influences the amount of inflow of Bank Indonesia in the one month after Eid ul-Fitr. This is related to people's habit to save money after carrying out Eid ul-Fitr holiday. In general, Eid ul-Fitr that occurs at the beginning of the month will result in a sharper increase in inflow. Otherwise, Eid ul-Fitr that occurs at the end of the month will yield the highest inflow in one month after this holiday. Additionally, the outflow was also affected by the occurrence of Eid ul-Fitr due to people tend to withdraw money to fulfill their needs during Eid ul-Fitr. In the month of Eid ul-Fitr, the highest outflow will happened when Eid ul-Fitr occurs at the end of the month. As for one month before Eid ul-Fitr, the highest outflow occurs when Eid ul-Fitr occurs at the beginning of the month.

3.3 Forecasting Inflow and Outflow Data Using ARIMAX

Forecasting inflow and outflow with ARIMAX method is using components as exogenous variables, which are trend, seasonal, and calendar variations component effects. These components are represented by the dummy variables as in Eq. (1).

The steps in ARIMAX method is to regress first the effects of trend, seasonal and calendar variation and then applying the ARIMA model on the residuals of this regression if these residuals not fulfill the white noise assumption. Based on model estimation at every denomination, the best ARIMA model for each residual time series regression is shown in Table 3. The ARIMAX equation model is obtained by combining time series regression model and the best ARIMA model from each of the inflow and outflow fractions. For example, the ARIMAX model for inflow Rp100.000,00 data can be written as:

$$
\begin{aligned}
Y_{7,t} = {} & 1.5t + 2956.4M_{1,t} - 1434.4M_{2,t} - 624.8M_{3,t} - 38.2M_{4,t} - 213.4M_{5,t} + \\
& 74.7M_{6,t} - 595.2M_{7,t} + 2941.3M_{8,t} - 3925.5M_{9,t} + 1433.0M_{10,t} - \\
& 11.8M_{11,t} + 428.9M_{12,t} + 5168.9V_{1,t} + 11591.4V_{2,t} + 9110.3V_{3,t} - \\
& 3826.1V_{4,t} - 6576.5V_{1,t+1} - 13521.5V_{2,t+1} - 13059.2V_{3,t+1} + \\
& 3102.7V_{4,t+1} + \frac{(1-0.80B)(1-0.17B^{12})}{(1-B)(1-B^{12})}a_t.
\end{aligned}
$$

Table 3. The best ARIMA model for each series

Data	ARIMA model	Data	ARIMA model
Y_1	ARIMA(0,1, [12])	Y_7	ARIMA(1,1,0)
Y_2	ARIMA(1,1, [1, 12])	Y_8	ARIMA(1,1, [1, 23])$(0,1,1)^{12}$
Y_3	ARIMA(0,1,1)	Y_9	ARIMA(1,0, [12, 23])
Y_4	ARIMA([12],1, [1, 23])	Y_{10}	ARIMA([1, 11, 12],1, [1, 12, 23])
Y_5	ARIMA(1,1, [12])	Y_{11}	ARIMA([12],0,2)
Y_6	ARIMA(0,1,1)$(0,1,1)^{12}$	Y_{12}	ARIMA([1, 10, 12],1, [1, 12, 23])

3.4 Forecasting Inflow and Outflow Data Using SSA-NN

In the previous study, it showed that a NN model could not capture the trend and seasonal patterns well [25]. To overcome it, the proposed SSA-NN firstly reconstruct the components of data using SSA. Each fraction of the inflow and outflow is decomposed by determining the L value of half of the data (L = 84). The SVD process obtained 50 eigentriples values. The grouping step is done by determining the value of effect grouping (r) to limit the number of eigentriples to grouping the trend and seasonal components. The r value is obtained from the sum of the singular values that have the graph show the noise component. Due to the simulation data show that aggregate forecasting gives better results than individual forecasting, then the forecasting inflow and outflow use aggregate forecasting. To do this, it is necessary to grouping the component pattern, i.e. trend, seasonal, and noise components.

Based on the results of principal component, there are 12 main components in inflow Rp 100.000,00. In Fig. 6, the components that tend to slowly increase or decrease are trend components, while components that follow periodic patterns and have corresponding seasonal periods are grouped into seasonal components, and other components are grouped into noise. Subsequent groupings are made to identify incoming inputs on the NN model. The result of reconstruction of each components of inflow Rp 100.000,00 is shown in Fig. 7.

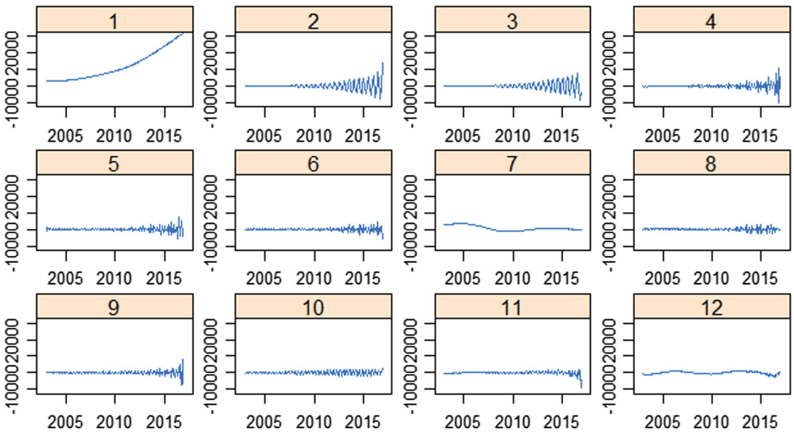

Fig. 6. Principal component plot of inflow Rp 100.000,00

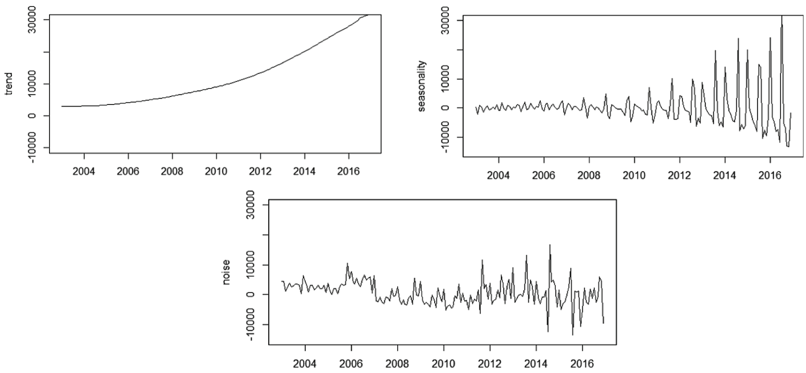

Fig. 7. Grouping trend, seasonal, and noise components of inflow Rp 100.000,00

Then, each component in Fig. 7 are modeled by NN. Based on the best model with the smallest value of goodness of fit criteria, the final NN model for each component can be written as

$$\hat{Y}_{7,t} = \hat{T}^*_{7,t} + \hat{S}^*_{7,t} + \hat{N}^*_{7,t}$$

where $\hat{T}^*_{7,t}$ is standardized value of T_t, $\hat{S}^*_{7,t}$ is standardized value of S_t, and $\hat{N}^*_{7,t}$ is standardized value of N_t. The architecture model of NN for denomination currency of Rp100.000,00 for trend and noise components are shown in Fig. 8.

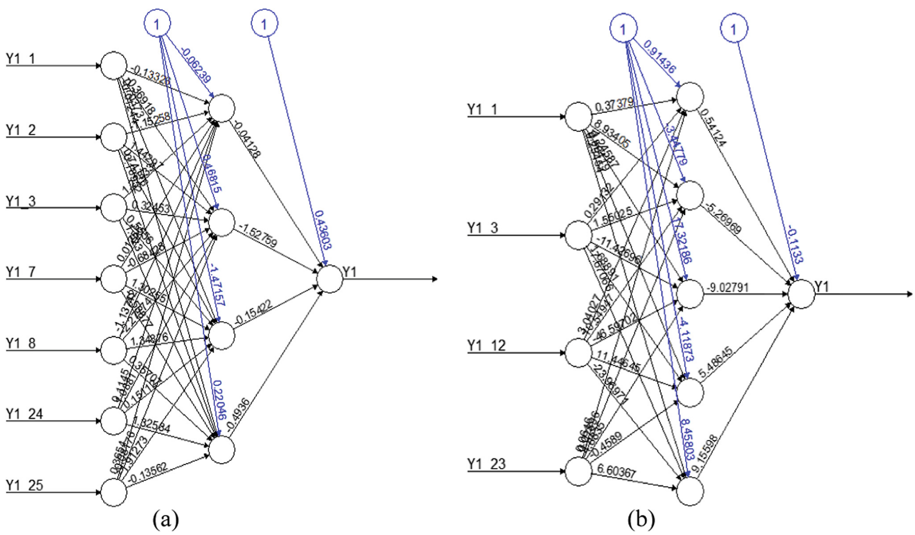

Fig. 8. NN architecture for trend (a) and noise (b) components of inflow Rp100.000,00

The SSA-NN modeling was performed on each denomination of inflow and outflow. In overall, hybrid SSA-NN model could capture well the pattern of calendar variation present in training data. However, in several denominations, the forecast value of testing data using SSA-NN model could not capture the calendar variation pattern (Fig. 9). The model evaluation using hybrid SSA-NN in each fraction presented in Table 4.

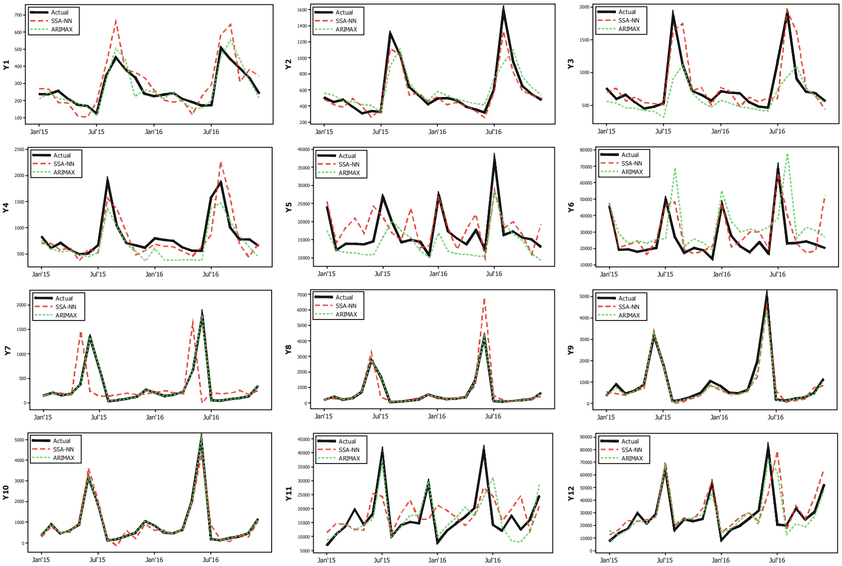

Fig. 9. Comparison of forecast value using SSA-NN and ARIMAX

Table 4. Model evaluation of hybrid SSA-NN model in each denomination

Variable	RMSE	MAE	MAPE	Variable	RMSE	MAE	MAPE
Y_1	84.2	67.9	26.8	Y_7	541.4	289.9	104.5
Y_2	87.3	63.9	11.2	Y_8	600.8	262.7	44.3
Y_3	245.8	154.5	19.3	Y_9	259.2	187.1	37.0
Y_4	254.5	190.3	20.9	Y_{10}	267.9	196.3	50.7
Y_5	4539	3701	23.2	Y_{11}	7591	6147	39.6
Y_6	9766	6368	29.7	Y_{12}	15138	8421	34.7

The forecast accuracy comparison between SSA-NN and ARIMAX methods for each denomination of inflow and outflow are shown in Fig. 9. Moreover, it is also necessary to analyze the reduced forecasting error for SSA-NN method compared to ARIMAX method. The comparison results of these methods are shown at Table 5.

The ratio value less than one indicates that SSA-NN with aggregate forecasting scheme is better and capable for reducing forecast error than ARIMAX based on RMSE criteria. In general, the results show that hybrid SSA-NN method give better results for predicting 6 out of 12 denominations of inflow and outflow. It is indicated by the RMSE ratio value that smaller than 1, which mean SSA-NN produces smaller forecast error than ARIMAX. Moreover, these results in line with M3 competition results, conclusion, and implication, i.e. more complex methods do not necessary yield better forecast than the simpler one [26].

Table 5. RMSE ratio between SSA-NN and ARIMAX methods

Data	RMSE Ratio	Data	RMSE Ratio
Y_1	1.92	Y_7	6.30
Y_2	0.51	Y_8	1.46
Y_3	0.78	Y_9	0.70
Y_4	1.16	Y_{10}	0.67
Y_5	0.93	Y_{11}	1.30
Y_6	0.55	Y_{12}	1.46

4 Conclusion

The results of simulation study showed that the proposed hybrid SSA-NN with aggregate forecasting scheme by grouping trends, seasonal, and noises yielded more accurate forecast than individual forecasting scheme. These results also showed that hybrid SSA-NN gave better performance in modeling series with random noise than nonlinear noise. Furthermore, the empirical study proved that Eid ul-Fitr had significant effect on the amount of inflow and outflow. The results for inflow and outflow data showed that hybrid SSA-NN could capture well the trend and seasonal pattern. Otherwise, this hybrid SSA-NN could not capture well the effects of calendar variations. Hence, it could be concluded that hybrid SSA-NN is a good forecasting method for time series which contain trends and seasonal only. Moreover, the results of forecast value comparison indicated that hybrid SSA-NN model performed as good as ARIMAX model, i.e. 6 of 12 denominations were better to be forecasted by the hybrid SSA-NN method, and the rests were more accurate to be forecasted by ARIMAX model. These results in line with the M3 competition conclusion, i.e. more complex methods do not necessary yield better forecast than the simpler one [26]. Hence, further research is needed to handle all patterns simultaneously, i.e. trend, seasonal, and calendar variation effects, by proposing new hybrid method as combination of SSA-NN and ARIMAX methods.

Acknowledgements. This research was supported by DRPM-DIKTI under scheme of “Penelitian Berbasis Kompetensi”, project No. 851/PKS/ITS/2018. The authors thank to the General Director of DIKTI for funding and to anonymous referees for their useful suggestions.

References

1. Sigalingging, H., Setiawan, E., Sihaloho, H.D.: Money Circulation Policy in Indonesia. Bank Indonesia, Jakarta (2004)
2. Apriliadara, M., Suhartono, A., Prastyo, D.D.: VARI-X model for currency inflow and outflow with Eid Fitr effect in Indonesia. In: AIP Conference Proceedings, vol. 1746, p. 020041 (2016)
3. Bowerman, B.L., O'Connell, R.T.: Forecasting and Time Series. Wadsworth Publishing Company, Belmont (1993)
4. Golyandina, N., Nekrutkin, V., Zhigljavsky, A.A.: Analysis of Time Series Structure: SSA and Related Techniques. Chapman & Hall, Florida (2001)
5. Broomhead, D.S., King, G.P.: Extracting qualitative dynamics from experimental data. Physica D **20**, 217–236 (1986)
6. Broomhead, D.S., King, G.P.: On the qualitative analysis of experimental dynamical systems. In: Sarkar S (ed.) Nonlinear Phenomena and Chaos, pp. 113–144. Adam Hilger, Bristol (1986)
7. Broomhead, D.S., Jones, R., King, G.P., Pike, E.R.: Singular Spectrum Analysis with Application to Dynamic Systems, pp. 15–27. IOP Publishing, Bristol (1987)
8. Afshar, K., Bigdeli, N.: Data analysis and short-term load forecasting in Iran electricity market using singular spectral analysis (SSA). Energy **36**(5), 2620–2627 (2011)
9. Hassani, H., Zhigljavsky, A.: Singular spectrum analysis: methodology and application to economic data. J. Syst. Sci. Complex. **22**, 372–394 (2008)
10. Zhigljavsky, A., Hassani, H., Heravi, S.: Forecasting European Industrial Production with Multivariate Singular Spectrum Analysis. Springer (2009)
11. Zhang, Q., Wang, B.D., He, B., Peng, Y.: Singular Spectrum Analysis and ARIMA Hybrid Model for Annual Runoff Forecasting. Springer, China (2011)
12. Li, H., Cui, L., Guo, S.: A hybrid short term power load forecasting model based on the singular spectrum analysis and autoregressive model. Adv. Electr. Eng. Artic. ID **424781**, 1–7 (2014)
13. Lopes, R., Costa, F.F., Lima, A.C.: Singular spectrum analysis and neural network to forecast demand in industry. In: Brazil: The 2nd World Congress on Mechanical, Chemical, and Material Engineering (2016)
14. Sun, M., Li, X., Kim, G.: Precipitation analysis and forecasting using singular spectrum analysis with artificial neural networks. Clust. Comput., 1–8 (2018, in press)
15. Barba, L., Rodriguez, N.: Hybrid models based on singular values and autoregressive methods for multistep ahead forecasting of traffic accidents. Math. Probl. Eng. **2016**, 1–14 (2016)
16. Zhang, X., Wang, J., Zhang, K.: Short-term electric load forecasting based on singular spectrum analysis and support vector machine optimized by cuckoo search algorithm. Electr. Power Syst. Res. **146**, 270–285 (2017)
17. Lahmiri, S.: Minute-ahead stock price forecasting based on singular spectrum analysis and support vector regression. Appl. Math. Comput. **320**, 444–451 (2018)
18. Khan, M.A.R., Poskitt, D.S.: Forecasting stochastic processes using singular spectrum analysis: aspects of the theory and application. Int. J. Forecast. **33**(1), 199–213 (2017)
19. Lee, M.H., Suhartono, A., Hamzah, N.A.: Calendar variation model based on ARIMAX for forecasting sales data with Ramadhan effect. In: Regional Conference on Statistical Sciences, pp. 349–361 (2010)
20. Zhang, P.G., Patuwo, E., Hu, M.Y.: Forecasting with artificial neural networks: the state of the art. Int. J. Forecast. **14**, 35–62 (1998)

21. Suhartono: New procedures for model selection in feedforward neural networks. Jurnal Ilmu Dasar **9**, 104–113 (2008)
22. Crone, S.F., Kourentzes, N.: Input-variable specification for neural networks - an analysis of forecasting low and high time series frequency. In: International Joint Conference on Neural Networks, pp. 14–19 (2009)
23. Anders, U., Korn, O.: Model selection in neural networks. Neural Netw. **12**, 309–323 (1999)
24. Wei, W.W.S.: Time Series Analysis: Univariate and Multivariate Methods, 2nd edn. Pearson Education, Inc., London (2006)
25. Zhang, G.P., Qi, M.: Neural network forecasting for seasonal and trend time series. Eur. J. Oper. Res. **160**(2), 501–514 (2005)
26. Makridakis, S., Hibon, M.: The M3-competition: results, conclusions and implications. Int. J. Forecast. **16**(4), 451–476 (2000)