An Empirical Study of Error Evaluation in Trend and Seasonal Time Series Forecasting Based on SSA

Winita Sulandari
Department of Mathematics, Universitas Gadjah Mada, Indonesia
Study Program of Statistics, Universitas Sebelas Maret, Indonesia
winita@mipa.uns.ac.id; wnidha@yahoo.com

Subanar
Department of Mathematics
Universitas Gadjah Mada, Indonesia
subanar@ugm.ac.id

Suhartono
Department of Statistics
Institut Teknologi Sepuluh Nopember, Indonesia
suhartono@statistika.its.ac.id

Herni Utami
Department of Mathematics
Universitas Gadjah Mada, Indonesia
herni_utami@mipa.ugm.ac.id

Muhammad Hisyam Lee
Department of Mathematical Sciences
Universiti Teknologi Malaysia, Malaysia
mhl@utm.my

Abstract

SSA (Singular Spectrum Analysis) starts to become a popular method in decomposing time series into some separable and interpretable series. This study provides an error evaluation in the SSA-based model for trend and multiple seasonal time series forecasting. This error evaluation is obtained by means of a numerical study on the mean square error of the estimators and mean absolute percentage error of the forecast values. Four distinct types of data generating processes (DGP) with varying sample sizes are considered in this experimental study. The parameters are estimated from the component series of SSA. Each DGP is decomposed into trend, periodic and irregular components. All these components except the irregular one are fitted by appropriate deterministic function separately. Based on the numerical simulation results, the estimated parameters are closer to the true values as the sample size increases. As the illustrative example of the real data set implementation, we used the monthly atmospheric concentrations of CO2 from Moana Loa observatory for period January 1959 to June 1972. The proposed method produces better forecast values than the results of SSA-LRF (Linear Recurrent Formula) and TLSAR (Two Level Seasonal Autoregressive). The results encourage the improvement in the time series modeling on the more complex pattern.

Keywords: SSA; Trend; Seasonal; Periodic; Deterministic.

1. Introduction

SSA (singular spectrum analysis) is a method in time series analysis. The origination of this method is always associated with the publication of Broomhead and King (1986).

Though, in the same year, Fraedrich (1986) also developed the SSA and applied the method to the weather and climate system. In the following years, Vautard and Ghil (1989) extended and refined some aspects of the SSA such as the influence of window length as the parameter of SSA and sample size on the results of SSA. Since then, a number of studies on SSA and its application have been carried out (see, for example, Hassani et al. (2009), Hassani et al. (2010), Hassani and Zhigljavsky (2009), Vautard and Ghil (1991), and Vautardet al. (1992)).

SSA has become popular for analyzing and forecasting time series since it was introduced by Elsner (2002), Elsner and Tsonis (1996), and Golyandina et al. (2001). The capability of SSA in decomposing time series into some components opens up the possibility for a new time series modeling procedure. Some researchers such as Li et al. (2014), Vahabie et al. (2007) and Zhang et al. (2011) combined SSA with autoregressive (AR) or autoregressive integrated moving average (ARIMA) model.

It has previously been observed that the components decomposed by SSA are not always stationary (Sulandari et al., 2017). Thus, a stationary time series model such as AR model is not appropriate to directly apply and even if its difference is stationary (ARIMA model). In this case, the polynomial regression model can be used to estimate the trend series (Kitagawa, 2010) while sinusoidal function can be chosen as the alternate approach for the seasonal series (De Livera et al., 2011; Soares and Medeiros, 2008). Moreover, the sinusoidal function is more general and flexible in its use (Wei, 2006).

There are several studies on SSA application in trend and multiple seasonal time series forecasting, such as study in electric load demand by Afshar and Bigdeli (2011) and Briceño et al. (2013), atmospheric concentration of CO2 series by Golyandina and Korobeynikov (2014) and US tourist arrivals by Hassani et al. (2015). Those papers discussed recurrent and vector forecasting algorithm. However, none of them that discussed the combination of polynomial and sinusoidal deterministic function to model the component of SSA. As far we know, the way in which we specify the models in each component of SSA decomposition is not common in SSA as a tool of time series forecasting. Most of researchers who concern in SSA used and developed SSA with linear recurrent formula (SSA-LRF) as a forecasting method. While others such as Vahabie et al. (2007) and Zhang et al. (2011) modeled the component of SSA by using AR or ARIMA model. In our point of view, not all components of SSA decomposition can be modeled by AR or ARIMA.

By combining the deterministic and stochastic model as in Soares and Medeiros (2008), our approach becomes more flexible. We apply basic SSA to decompose the series into some components and estimate each of the components except the irregular component by the function of time and this make the model easier to interpret. This work presents the algorithm to estimate the parameters of the trend and multiple seasonal time series model. The parameters of the model are separately estimated from each component and then combined them. In order to show that the proposed algorithm is acceptable, a simulation study of mean square error (MSE) of the estimators and the mean absolute percentage error (MAPE) of the forecast value was done.

The remaining part of this paper is organized as follows. The detail of materials and methods are presented in Section **2**. This section begins by a brief overview of the data generation used in the simulation, SSA decomposition, and the error evaluation of the model. It then continues with the steps of proposed algorithm in estimating parameters and evaluating the errors. The experimental results and the illustrative example to the real data are reported in Section **3**. Finally, the conclusion is presented in Section 4.

2. Materials and Methods

This simulation study is limited to the data that are generated from deterministic function with two different error structures. To investigate whether the proposed algorithm is acceptable, the discussion would be considered in the four distinct data generating processes, those are trend linear with multiple seasonal time series (Model 1), trend quadratic with multiple seasonal time series (Model 2), trend linear with multiple seasonal time series in the presence of autocorrelated error (Model 3) and trend quadratic with multiple seasonal time series in the presence of autocorrelated error (Model 4).

2.1 Data Generation

This simulation study is set up from the model presented in (1),

$$X_t = \theta_0 + \theta_1 t + \theta_2 t^2 + \sum_{j=1}^{2} \alpha_j \cos(\omega_j t) + \beta_j \sin(\omega_j t) + u_t. \tag{1}$$

Two different error structures were considered for $\{u_t\}$. The first type of errors is Gaussian white noise with mean 0 and variance 1, while the second type is the first order autoregressive process. The four distinct types of data generating processes (DGP) that will be studied further in this simulation study are presented in Table 1.

Table 1: The four DGPs used in the simulation study

Seasonal	Trend	Noise	Model	Parameters
Double period	Linear	White Noise	1	$\varphi = [\theta_0\ \theta_1\ \alpha_1\ \beta_1\ \alpha_2\ \beta_2\ \omega_1\ \omega_2\ \sigma^2]$ $= [1204\ -0.0122\ -25\ 331\ 100\ -40\ 0.8976\ 0.2096\ 1]$
		AR(1)	3	$\varphi = [\theta_0\ \theta_1\ \alpha_1\ \beta_1\ \alpha_2\ \beta_2\ \omega_1\ \omega_2\ \rho_1\ \sigma^2]$ $= [1204\ -0.0122\ -25\ 331\ 100\ -40\ 0.8976\ 0.2096\ 0.7\ 1]$
	Quadratic	White Noise	2	$\varphi = [\theta_0\ \theta_1\ \theta_2\ \alpha_1\ \beta_1\ \alpha_2\ \beta_2\ \omega_1\ \omega_2\ \sigma^2]$ $= [1204\ -0.0122\ 0.0001\ -25\ 331\ 100\ -40\ 0.8976\ 0.2096\ 1]$
		AR(1)	4	$\varphi = [\theta_0\ \theta_1\ \theta_2\ \alpha_1\ \beta_1\ \alpha_2\ \beta_2\ \omega_1\ \omega_2\ \rho_1\ \sigma^2]$ $= [1204\ -0.0122\ 0.0001\ -25\ 331\ 100\ -40\ 0.8976\ 0.2096\ 0.7\ 1]$

These four DGP models are written as follows.

1) Model 1: u_t is independently identically distributed (i.i.d) N(0,1),

$$X_t = 1204 - 0.0122t - 25\cos\left(\frac{2\pi}{7}t\right) + 331\sin\left(\frac{2\pi}{7}t\right) + 100\cos\left(\frac{2\pi}{30}t\right) - 40\sin\left(\frac{2\pi}{30}t\right) + u_t.$$

2) Model 2: u_t is i.i.d N(0,1),

$$X_t = 1204 - 0.0122t + 0.0001t^2 - 25\cos\left(\frac{2\pi}{7}t\right) + 331\sin\left(\frac{2\pi}{7}t\right) + 100\cos\left(\frac{2\pi}{30}t\right) - 40\sin\left(\frac{2\pi}{30}t\right) + u_t.$$

3) Model 3: $u_t = \rho_1 u_{t-1} + e_t$, where ρ_1=0.7 and e_t is i.i.d N(0,1),

$$X_t = 1204 - 0.0122t - 25\cos\left(\frac{2\pi}{7}t\right) + 331\sin\left(\frac{2\pi}{7}t\right) + 100\cos\left(\frac{2\pi}{30}t\right) - 40\sin\left(\frac{2\pi}{30}t\right) + u_t.$$

4) Model 4: $u_t = \rho_1 u_{t-1} + e_t$, where ρ_1=0.7 and e_t is i.i.d N(0,1),

$$X_t = 1204 - 0.0122t + 0.0001t^2 - 25\cos\left(\frac{2\pi}{7}t\right) + 331\sin\left(\frac{2\pi}{7}t\right) + 100\cos\left(\frac{2\pi}{30}t\right) - 40\sin\left(\frac{2\pi}{30}t\right) + u_t.$$

The sample sizes considered in this simulation study are N= 300, 500, 700, 1000, and 1500. The last 30 points of each data series are taken as the testing data, while the $(N - 30)$ first points are the training data. The data generating process of each model is replicated $R = 100$ times. Thus, for each model, there are 100 independent time series.

2.2 Singular Spectrum Analysis (SSA)

SSA method consists of four steps. The first step is embedding process. The series $\{X_t, t = 1, 2, \dots, N\}$ is mapped into a trajectory matrix, **X**, with the size $L \times K$. Notation L, the number of the rows, denoted the parameter of SSA, named window length and $K = N - L + 1$, is the number of the columns of **X**. The l^{th} (for l = 1, 2,…, L) row and k^{th} (for k = 1, 2, …, K) column of **X** is X_{l+k-1} that is the observation value at time $(l + k - 1)$. In decomposing time series using SSA, the window length$L$ is a parameter that must be specified by the researcher. Its choice may influence the decomposition process and the components of the series. Khan and Poskitt (2011) presented the theoretical analysis of signal-noise separation and reconstruction in SSA to provide the optimal window length guidance. Meanwhile Golyandina and Zhigljavsky (2013)identified that the window length should be large enough.

In the second step, matrix **X** is decomposed into some elementary matrices using singular value decomposition (SVD). The matrices obtained from step two are then analyzed to find out which matrices were separable. The separability between components is measured by weighted correlation defined in Elsner and Tsonis (1996), Golyandina et al. (2001), and Golyandina and Zhigljavsky (2013).

The third step is grouping. The matrices are grouped into several separable matrices. Finally, in the last step, each group of matrices are transformed into a new series of length N by diagonal averaging algorithm (see Golyandina and Zhigljavsky (2013) for detail algorithm).

In SSA forecasting, Golyandina et al. (2001) proposed SSA-LRF (linear recurrent formula). The theoretical results on the properties of forecast determined by SSA can also be found in Khan and Poskitt (2014).

2.3 Error Evaluation

The proposed algorithm is acceptable if the estimator established from this procedure is close to the real value (parameter) of the model. One of the most common measures for evaluating the expected distance between the estimator and the parameter is mean square error (MSE), and the instrument to evaluate the forecast value is mean absolute percentage error (MAPE).

The MSE of the estimators for each R data sets of size N with d parameters, can be calculated by

$$\mathrm{MSE}_N(\hat{\varphi}) = E\|\hat{\varphi} - \varphi\|^2 = \frac{1}{R}\sum_{r=1}^{R}\sum_{i=1}^{d}(\hat{\varphi}_{ri} - \varphi_i)^2 \qquad (2)$$

where d is the number of parameters to be estimated, $\hat{\varphi}_{ri}$ is the estimator of the *i*-th parameter obtained from the r^{th} replication data set. Estimator $\hat{\varphi}$ is convengent in mean square to the parameter φ if $\|\hat{\varphi} - \varphi\| \to 0$ as $N \to \infty$(Brockwell and Davis, 1991). In this study, some training sample sizes are considered to examine the effect of sample size to the MSE of estimators. The larger number of observations should provide much more information about the unknown parameters, and therefore, the MSE of the estimators would be smaller.

MAPE is selected to measure the error of *h*-steps ahead forecast values. In this case, the testing data sets in the simulation study are the last 30 points of each independent series with varying sample sizes that are in fact different from each other. The benefit of this scaled-independent MAPE is that it can be used to make comparisons between different data sets (Hyndman et al., 2008). The MAPE is defined by

$$\mathrm{MAPE}_{N_h} = \frac{1}{N_h}\sum_{h=1}^{N_h} 100\% \left|\frac{X_{t+h} - \hat{X}_{t+h}}{X_{t+h}}\right|, \qquad (3)$$

where X_{t+h} and $\hat{X}_{t+h}$ are the observation value at time $(t+h)$ and the forecast value at time $(t+h)$, respectively. N_h is the length of the forecast future period.

2.4 Experimental Design

Each generated series is divided into two data sets. The first data set is the training data and the second one is the testing data set. The length of the training data set is notated by N_{tr}. SSA is used to decompose each training data set into several separable components. In this simulation study, L is proportional to N_{tr}. Its value is set to be the biggest integer less than or equal to $N_{tr}/2$ as recommended by Golyandina (2010).

Each separable component obtained from SSA decomposition is then modeled and combined using the proposed algorithm. The process stages of the proposed algorithm are visualized in Figure 1 and presented as follows.

Step 1: Estimating the trend function.
Let $\{T_t, t = 1, 2, .., N_{tr}\}$ is the trend component obtained from SSA decomposition and $T_t = \sum_{i=0}^{N_{pol}} \theta_i t^{i-1} + \varepsilon_T$ is the polynomial model with the order N_{pol}. Find the best fit function using the following procedure.

(a) Estimate the order of the polynomial model N_{pol}, based on the behavior of $\{T_t, t = 1, 2, .., N_{tr}\}$ series.
(b) Estimate the parameter θ_i $(i = 0, \dots, N_{pol})$ using ordinary least square (OLS) method.
(c) Define $\hat{T}_t = \sum_{i=0}^{N_{pol}} \hat{\theta}_i t^i$ and $\varepsilon_t = T_t - \hat{T}_t$.

Step 2: Estimating sinusoidal function for the seasonal component.
Let $\{S_t^{(j)}, t = 1, \dots, N_{tr}\}$ is the j^{th} seasonal component series and the seasonal model is presented in equation (4) below,

$$S_t^{(j)} = \alpha_j \cos(\omega_j t) + \beta_j \sin(\omega_j t) + \eta_t^{(j)} \qquad (4)$$

for $j = 1, \dots, N_S$. $N_S$is the number of seasonal components that are not correlated one another. For every j, estimate α_j, β_j and ω_j using the following procedure.

(a) Express $S_t^{(j)}$ in the Fourier representation, $S_t^{(j)} = \sum_{p=0}^{[N_{tr}/2]} \alpha_{j_p} \cos(\omega_{j_p} t) + \beta_{j_p} \sin(\omega_{j_p} t)$, where $\omega_{j_p} = \frac{2\pi p}{N_{tr}}, p = 0, 1, \dots, [N_{tr}/2]$ are the Fourier frequencies and

$$\alpha_{j_p} = \begin{cases} \dfrac{1}{N_{tr}} \displaystyle\sum_{t=1}^{N_{tr}} S_t^{(j)} \cos\left(\omega_{j_p} t\right), & p = 0 \text{ and } p = \dfrac{N_{tr}}{2} \text{ if } N_{tr} \text{ is even,} \\ \dfrac{2}{N_{tr}} \displaystyle\sum_{t=1}^{N_{tr}} S_t^{(j)} \cos(\omega_{j_p} t), & p = 1,2, \dots, \left[\dfrac{N_{tr}-1}{2}\right], \end{cases}$$

$$\beta_{j_p} = \frac{2}{N_{tr}} \sum_{t=1}^{N_{tr}} S_t^{(j)} \sin(\omega_{j_p} t), \qquad p = 1,2, \dots, \left[\frac{N_{tr}-1}{2}\right],$$

are Fourier coefficients (Wei, 2006).

(b) Estimate the initial $\hat{\omega}_j$ using periodogram $I(\omega_{j_p})$,

$$I\left(\omega_{j_p}\right) = \begin{cases} N_{tr}\alpha_{j_0}^2, & p = 0 \\ \dfrac{N_{tr}}{2}\left(\alpha_{jp}^2 + \beta_{jp}^2\right), & p = 1, 2, \dots, \left[\dfrac{N_{tr}-1}{2}\right], \\ N_{tr}\alpha_{j_{\frac{N_{tr}}{2}}}^2, & p = \dfrac{N_{tr}}{2} \text{ if } N_{tr} \text{ is even.} \end{cases}$$

Initial $\hat{\omega}_j$ can be determined between the value $\omega_{j_{q-1}} = \frac{2\pi(q-1)}{N_{tr}}$ and $\omega_{j_{q+1}} = \frac{2\pi(q+1)}{N_{tr}}$ where ω_{j_q} is the maximizer of $I(\omega_{j_p})$. The procedure is presented in Figure 1 Step 2.

(c) Using a given value of $\hat{\omega}_j$ obtained from point (b), the parameters α_j and β_j are then can be estimated by OLS method. Intercept could be included if needed. Calculate $\hat{S}_t^{(j)} = \hat{\alpha}_j \cos(\omega_{j_k} t) + \hat{\beta}_j \sin(\omega_{j_k} t)$ and

$$\text{RMSE}\left(\eta^{(j)}\right) = \sqrt{\sum_{t=1}^{N_{tr}} (S_t^{(j)} - \hat{S}_t^{(j)})^2 / N_{tr}}. \qquad (5)$$

(d) Define the most appropriate frequency ω_j^* between $\omega_{j_{q-1}}$ and $\omega_{j_{q+1}}$. Repeat (c) for a given frequency value between $\omega_{j_{q-1}}$and $\omega_{j_{q+1}}$. The most appropriate estimated function $\hat{S}_t^{(j)}$ of the j^{th} sinusoidal component is $\hat{S}_t^{(j)}$ with the estimators $\hat{\omega}_j^*, \hat{\alpha}_j^*, \hat{\beta}_j^*$ so that minimizes the RMSE$\left(\eta^{(j)}\right)$.

Step 3: Investigating the superposed irregular component.
The superposed irregular component is the sum of the irregular (I) component obtained from SSA decomposition and the residuals obtained from Step 1 and Step 2, that is

$$u_t = I_t + \varepsilon_t + \sum_{j=1}^{N_S} \eta_t^{(j)} \ , \quad t = 1, 2, \dots, N_{tr}. \tag{6}$$

(a) Find the superposed irregular component using equation (6), investigate whether it is white noise or not. If superposed irregular is white noise, the residuals of the model is $e_t = u_t$, for $t = 1, 2, \dots, N_{tr}$. If there were significant autocorrelation appeared in the superposed irregular, the process continues to (b).
(b) Apply first order AR model, $u_t = \rho_1 u_{t-1} + e_t$ to approximate the superposed irregular model. Estimate the parameter using OLS method and define the residuals of the model, $e_t = u_t - \hat{\rho}_1 u_{t-1}$.
(c) Estimate the variance, $\hat{\sigma}^2 = \sum_{t=2}^{N_{tr}} e_t^2 \,/(N_{tr} - 2)$.

Step 4: Evaluating the error.
(a) Calculate MSEs of estimators using formula (2) and MAPEs for the testing data using formula (3).
(b) Compare MSEs of estimators and MAPEs between the different sample sizes.

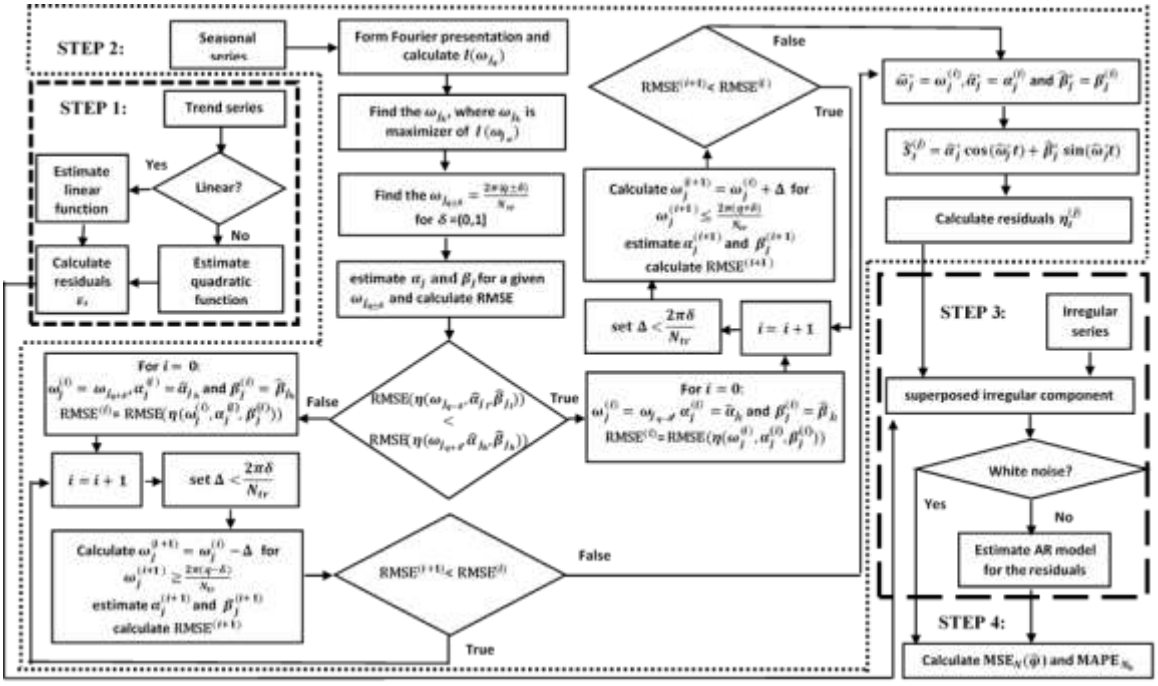

Figure 1: The flowchart of the proposed algorithm used in the simulation study

3. Results and Discussion

In this section, the proposed method is implemented to the four types of DGPs as a simulation study. The algorithm is then applied to CO2 time series CO2 from Mauna Loa to show how this algorithm works to the real data.

3.1 Simulation Study

The estimated parameters and the forecast values are then evaluated using MSE of estimators and MAPE for 1 to 30-steps ahead forecast values. All models are estimated on a notebook Intel Core i7 with 8 Gb of Ram memory and running Matlab R2015a. The computational time for all 2000 SSA decomposition (4 types × 5 sample sizes × 100 series) and estimating all 7000 models ((2 types × (3 + 4 components) × 5 sample sizes × 100 series) is negligible, even though its need much time.Figure 2 (left) presents one of w-correlation matrix result for the series generated from Model 1 with $N = 500$ and $L = 235$. The color in the Figure 2 (left) informs the degree of the correlation or separability between components. The stronger the correlation is shown by the darker the color and the weaker the correlation or the stronger the separability is shown by the lighter color.

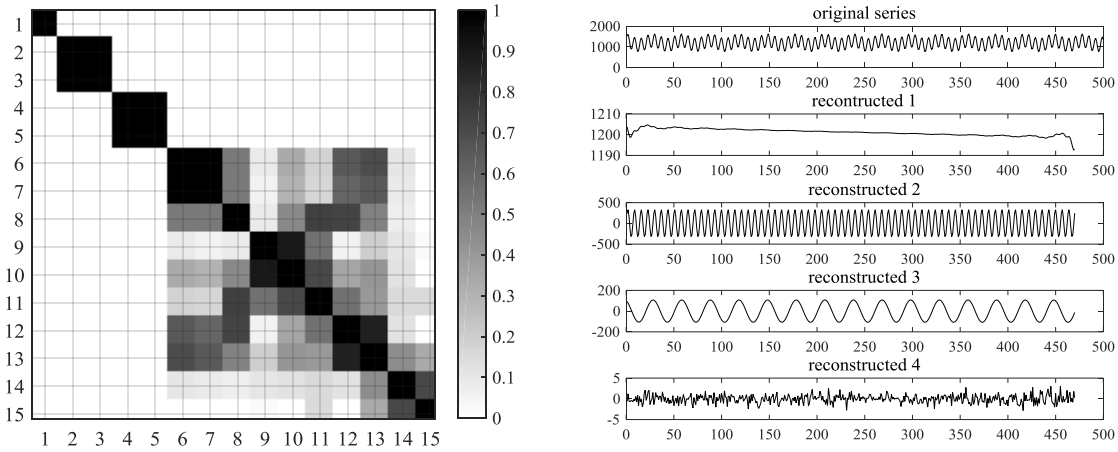

Figure 2: w-correlation matrix for Model 1 with N = 500 (left) and components of Model 1 with N = 500 (right)

In this case, the series can be decomposed into 4 component groups (see Figure 2 (right)), constructed from component 1, 2-3, 4-5 and 6-235. The first component is trend series, the next two group of components are the seasonal series, and the last group of component is the irregular series. The trend component of the series generated from Model 1 and Model 2 are estimated by linear model while those are generated from Model 3 and 4 are approximated by quadratic model. All seasonal components are estimated by sinusoidal function, and the residuals obtained from the polynomial and the sinusoidal model are then added to the irregular component, named superposed irregular.

Perhaps, the estimate parameters are affected by the choice of window length in SSA. By putting window length proportionally to the sample size, i.e. the biggest integer less than a half of training data, its influence to the decomposition results will be reduced.

The periodogram is only used for estimate the initial value of frequency of the sinusoidal function since it provides a crude estimate of frequency parameter (Quinn, 1994).Quinn (1994) presented the technique to find the frequency from three Fourier components that are the Fourier component at the maximizer of the periodogram and at the two adjacent Fourier frequencies. The estimate frequency parameters ω_j^* is obtained from the value $\omega_{j_{q-1}}$ to $\omega_{j_{q+1}}$ for a certain k such that ω_{j_q} is the maximizer of the periodogram. This paper presents the algorithm of finding estimate frequency in different way from Quinn (1994). In Quinn (1994), the estimate frequency is obtained by considering the real part ratio of exponential periodogram function at $\omega_{j_{q+1}}$ to ω_{j_q} and $\omega_{j_{q-1}}$to ω_{j_q}. While this study determines the estimate frequency iteratively between the value $\omega_{j_{q-1}}$ and $\omega_{j_{q+1}}$or perhaps in a narrower interval $(\omega_{j_{q-\delta}}, \omega_{j_{q+\delta}})$ for $0 < \delta < 1$ until the RMSE of equation (5) reaches the minimum value.Though the proposed algorithm probably needs more computation, the recent computational advances make it simpler and easier.

The results of estimated parameters are summarized in Table 2 to Table 5. As shown in Table 2 and Table 3, the estimated parameters for the sample of size $N = 300$ are far from the true values and consequently the MSEs of estimators become much larger than other sample sizes. As well as shown in Table 4 and Table 5. Generally, it is apparent from those tables that the average value of each estimated parameters tends to be closer to the true value as the sample size increases. It is also indicated by the value of MSEs of estimators that become smaller as the sample sizes increase, as can be seen from Table 2 to Table 5.

Table 2: The average of estimated parameters and MSEs of estimators for 100 replicates of series generated from Model 1

	Parameters of Model 1 ($\varphi = [\theta_0\ \theta_1\ \alpha_1\ \beta_1\ \alpha_2\ \beta_2\ \omega_1\ \omega_2\ \sigma^2]$)									
	θ_0	θ_1	α_1	β_1	α_2	β_2	ω_1	ω_2	σ^2	
	1204	−0.0122	−25	331	100	−40	0.8976	0.2094	1	
N	Estimated parameters of Model 1									$\mathrm{MSE}_N(\hat{\varphi})$
	$\hat{\theta}_0$	$\hat{\theta}_1$	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\alpha}_2$	$\hat{\beta}_2$	$\hat{\omega}_1$	$\hat{\omega}_2$	$\hat{\sigma}^2$	
300	1202.60	−0.0023	−23.00	330.59	99.15	−37.10	0.8976	0.2096	4.67	29.2396
500	1202.63	−0.0108	−24.47	330.97	100.12	−38.14	0.8976	0.2095	3.07	10.0029
700	1203.66	−0.0110	−24.39	331.35	99.06	−40.04	0.8976	0.2094	1.90	2.4194
1000	1204.47	−0.0126	−24.68	330.90	99.43	−39.89	0.8976	0.2094	1.68	1.3954
1500	1203.98	−0.0122	−25.01	330.99	99.92	−39.89	0.8976	0.2094	1.01	0.0459

Table 3: The average of estimated parameters and MSEs of estimators for 100 replicates of series generated from Model 2

Parameters of Model 2 ($\varphi = [\theta_0\ \theta_1\ \theta_2\ \alpha_1\ \beta_1\ \alpha_2\ \beta_2\ \omega_1\ \omega_2\ \sigma^2]$)										
θ_0	θ_1	θ_2	α_1	β_1	α_2	β_2	ω_1	ω_2	σ^2	
1204	−0.0122	0.0001	−25	331	100	−40	0.8976	0.209	1	
N	Estimated parameters of Model 2									$\text{MSE}_N(\hat{\varphi})$
	$\hat{\theta}_0$	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\alpha}_2$	$\hat{\beta}_2$	$\hat{\omega}_1$	$\hat{\omega}_2$	$\hat{\sigma}^2$	
300	1200.39	0.0466	0.0000	−24.83	330.46	99.15	−37.10	0.8976	0.2096	4.97	38.3990
500	1202.10	−0.0044	0.0001	−24.47	330.97	100.11	−38.11	0.8976	0.2095	3.18	12.2854
700	1203.98	−0.0103	0.0001	−24.39	331.35	99.05	−40.06	0.8976	0.2094	1.89	2.4115
1000	1204.52	−0.0122	0.0001	−24.69	330.90	99.42	−39.92	0.8976	0.2094	1.68	1.4491
1500	1203.73	−0.0112	0.0001	−25.01	330.99	99.92	−39.88	0.8976	0.2094	1.02	0.1220

Table 4: The average of estimated parameters and MSEs of estimators for 100 replicates of series generated from Model 3

Parameter of Model 3 ($\varphi = [\theta_0\ \theta_1\ \alpha_1\ \beta_1\ \alpha_2\ \beta_2\ \omega_1\ \omega_2\ \rho_1\ \sigma^2]$)										
θ_0	θ_1	α_1	β_1	α_2	β_2	ω_1	ω_2	ρ_1	σ^2	
1204	−0.0122	−25	331	100	−40	0.8976	0.2094	0.7	1	
N	Estimated parameters of Model 3									$\text{MSE}_N(\hat{\varphi})$
	$\hat{\theta}_0$	$\hat{\theta}_1$	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\alpha}_2$	$\hat{\beta}_2$	$\hat{\omega}_1$	$\hat{\omega}_2$	$\hat{\rho}_1$	$\hat{\sigma}^2$	
300	1204.94	−0.0023	−24.39	330.25	99.12	−37.13	0.8976	0.2096	0.92	3.26	17.1576
500	1204.20	−0.0109	−24.45	330.95	100.15	−38.10	0.8976	0.2095	0.83	1.25	6.3667
700	1204.11	−0.0111	−24.39	331.34	99.11	−40.01	0.8976	0.2094	0.76	1.17	1.7690
1000	1204.45	−0.0118	−24.61	330.91	99.48	−39.86	0.8976	0.2094	0.76	1.07	1.0482
1500	1203.95	−0.0121	−25.01	331.00	99.96	−39.85	0.8976	0.2094	0.70	1.04	0.1732

Table 5: The average of estimated parameters and MSEs of estimators for 100 replicates of series generated from Model 4

Parameters of Model 4 ($\varphi = [\theta_0\ \theta_1\ \theta_2\ \alpha_1\ \beta_1\ \alpha_2\ \beta_2\ \omega_1\ \omega_2\ \rho_1\ \sigma^2]$)											
θ_0	θ_1	θ_2	α_1	β_1	α_2	β_2	ω_1	ω_2	ρ_1	σ^2	
1204	−0.0122	0.0001	−25	331	100	−40	0.8976	0.2094	0.7	1	
N	Estimated parameters of Model 4										$\text{MSE}_N(\hat{\varphi})$
	$\hat{\theta}_0$	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\alpha}_2$	$\hat{\beta}_2$	$\hat{\omega}_1$	$\hat{\omega}_2$	$\hat{\rho}_1$	$\hat{\sigma}^2$	
300	1201.97	0.0609	0.0001	−24.39	330.25	99.17	−37.10	0.8976	0.2096	0.93	1.63	15.5710
500	1202.97	−0.0000	0.0001	−24.44	330.95	99.94	−39.71	0.8976	0.2095	0.75	1.26	1.9405
700	1203.84	−0.0098	0.0001	−24.39	331.34	99.10	−40.03	0.8976	0.2094	0.77	1.18	1.8214
1000	1204.50	−0.0119	0.0001	−24.61	330.91	99.47	−39.89	0.8976	0.2094	0.76	1.06	1.0985
1500	1203.82	−0.0115	0.0001	−25.01	331.00	99.97	−39.84	0.8976	0.2094	0.70	1.00	0.2060

Figure 3 provides the MSEs of estimators obtained from the simulation study of Model 1 to Model 4 for the sample size of N = 300, 500, 700, 1000, and 1500 with 100 replications. In Figure 3 there is a clear trend of decreasing MSEs of estimators as the sample size increases. This means that the proposed algorithm yields consistent estimators.

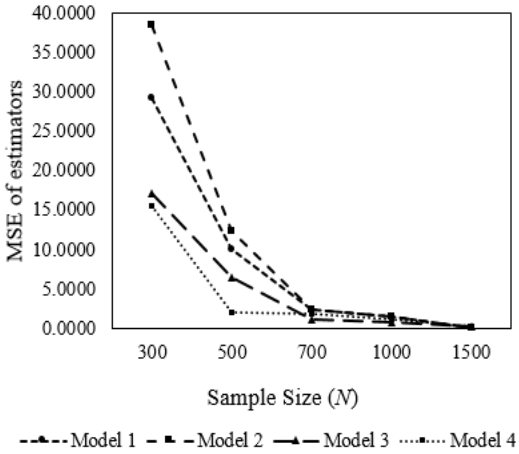

Figure 3: MSE of estimators for Model 1, 2, 3 and 4 with the sample size N = 300, 500, 700, 1000, and 1500.

Figure 4 shows the behavior of MAPEs for 1 to 30-steps ahead forecast values for Model 1 to Model 4. As shown in Figure 4, the models constructed from the largest sample size (in this case, $N = 1500$) produce the smallest MAPEs and otherwise the models constructed from the smallest sample size ($N = 300$) yield the larger and unstable MAPEs. It means that the model performance is getting better when the training sample size is increasing.

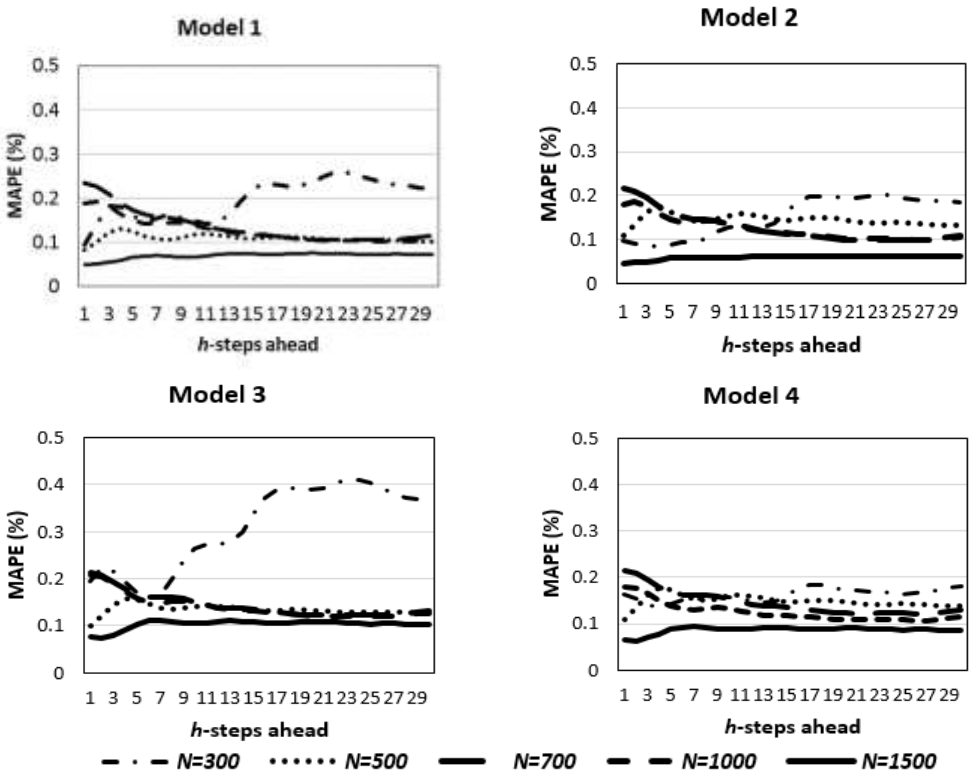

Figure 4: MAPEs *h*-steps ahead for Model 1, 2, 3 and 4 with the sample size N = 300, 500, 700, 1000, and 1500.

The discussion here is limited to the linear and quadratic trend and double stationary periodic function. After all, the empirical findings in this study provide a new

understanding of SSA based model procedure, especially for the time series with trend and multiple seasonal patterns. More research is required to determine the appropriate optimization algorithm for estimating the frequency of sinusoidal model. Further investigation and experimentation into more complex pattern, i.e. amplitude-modulated periodic series, is also needed to establish the more general procedure in modeling time series with trend and multiple seasonal patterns.

3.2 Application to CO2 Time Series

As an illustrative example of implementation of the proposed method on the real data, we used the monthly atmospheric concentrations of CO2 from Mauna Loa Observatory, Hawaii(Keeling and Whorf, 1997). Golyandina and Korobeynikov(2014) showed that the data have a trend and two seasonal components. The period of data used in this study is January 1959 to June 1972. The first 150 observations are used as the training data set and the next 12 observations are used as the testing data set.

The components obtained by SSA decomposition for the training data with L = 75 are depicted in Figure 5 and the most appropriate function for the reconstructed 1, 2 and 3 are presented in Table 6. The reconstructed 1 shows the trend series while the reconstructed 2 and 3 are the seasonal components.

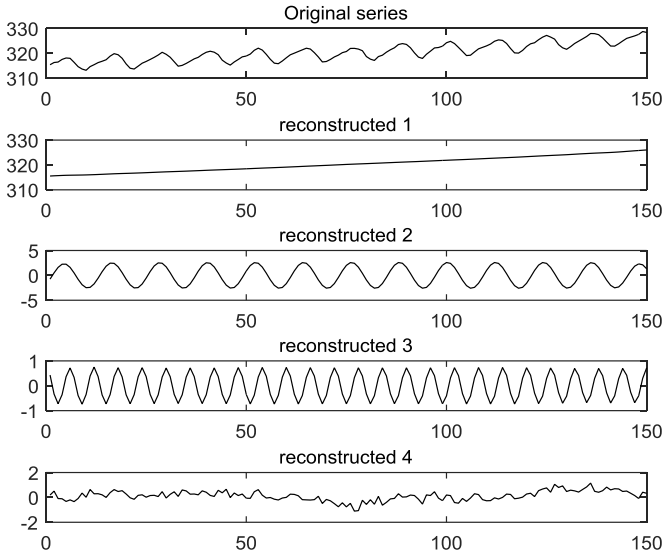

Figure 5: The original and components of CO2 series for period Jan 1959 – June 1971 (L = 75)

Table 6: Model for the trend and seasonal components of CO2

Reconstructed	Model
1	T_t = 315.4992 + 0.0547 t + 9.3161e-05 t^2 + ε_t
2	$S_t^{(1)}$ = – 0.0186 – 1.7543 cos (0.5244 t) + 1.8693 sin (05244 t) + $\eta_t^{(1)}$
3	$S_t^{(2)}$ = 0.0043 + 0.6902 cos (1.0430 t) – 0.1875 sin (1.0430 t) + $\eta_t^{(2)}$

The superposed residuals (u_t, t = 1, 2, …, 150) can then be calculated from ε_t, $\eta_t^{(1)}$, $\eta_t^{(2)}$, and the irregular (I) component (reconstructed 4). Since the superposed residuals are not stationary in mean, the series have to be made stationary by differencing so that the

autoregressive model can be implemented. In this case, the differenced first-order autoregressive model

$$\dot{u}_t = -0.3134\, \dot{u}_{t-1} + e_t,$$

where $\dot{u}_t = u_t - u_{t-1}$ is the best fit model to the superposed residuals. Further, the model CO2 can be represented as

X_t = 315.4849 + 0.0547t + 9.3161e-05 t^2 – 1.7543 cos (0.5244 t)
+ 1.8693 sin (0.5244 t) + 0.6902 cos (1.0430 t) – 0.1875 sin (1.0430 t)
+ 0.6866 u_{t-1}+0.3134 u_{t-2}+ e_t.

This model is then compared with SSA-LRF proposed by Golyandina et al. (2001) and two level seasonal autoregressive (TLSAR) model proposed by (Soares and Medeiros, 2008). The TLSAR model is constructed based on a two-step decomposition. The linear trend and harmonic of seasonal component based on Fourier series are involved in the first step and linear autoregressive model is involved in the second step.

Table 7: RMSE of the SSA-LRF, TLSAR, and the proposed method for the CO2 series

Method	RMSE	
	Training	Testing
SSA-LRF	0.6629	0.7789
TLSAR	0.2904	2.7534
The proposed method	0.3033	0.4391

Based on Table 7, TLSAR yields the smallest RMSE for the training data compared to others, but its value is very close to the result of the proposed method nevertheless. For the testing data, the proposed method produces the smallest RMSE and vice versa, the RMSE of TLSAR is the largest and even much larger than others.

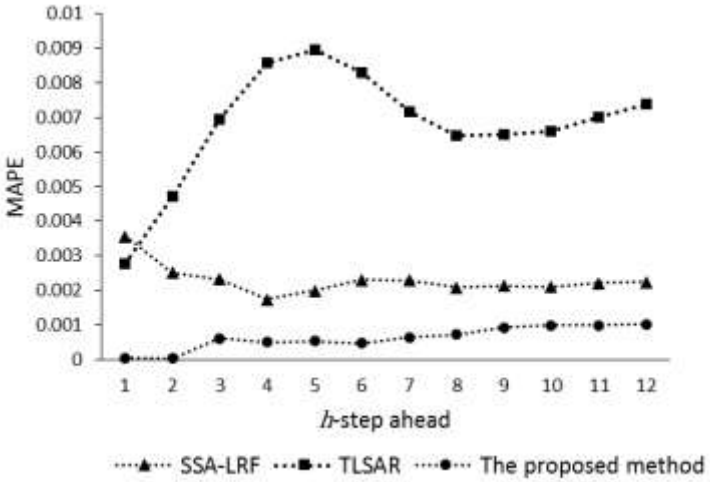

Figure 6: MAPE *h*-steps ahead for period July 1971 – June 1972

However, the MAPEs for *h*-steps ahead for period July 1971 to June 1972 obtained from the proposed method yields lowest values than those obtained from SSA-LRF and TLSAR (see Figure 6). For this case, the proposed method is slightly more efficient and gives better forecast values than the results of the SSA-LRF and TLSAR.

4. Conclusions

In this paper, the simulation study of error evaluation in trend and multiple seasonal time series forecasting based on SSA is presented. The error of estimators is measured by MSE of estimators while the error of forecast values is measured by MAPE. Based on the experimental result, MSEs of estimators tend to be smaller with the increasing of the sample sizes. It could be said that the experiment yields consistent estimators. And along with that, the forecasting values of the models constructed by the larger sample sizes produce the smaller MAPEs.

Based on the implementation of the proposed method on CO2 series, it can be seen that the proposed method is worthy to be considered in trend and multiple seasonal time series modeling. This conclusion is also reinforced by the comparative results of RMSE and MAPE *h*-steps ahead between the SSA-LRF, TLSAR, and the proposed method.

Acknowledgments

This research is supported by Doctoral Dissertation Research Grant from Non-Tax Revenue Universitas Sebelas Maret (PNBP UNS).We appreciate and thank to the editor and anonymous reviewers for all their valuable suggestions.

References

1. Afshar, K., and Bigdeli, N. (2011). Data analysis and short term load forecasting in Iran electricity market using singular spectral analysis (SSA). Energy, 36(5), 2620–2627.

2. Briceño, H., Rocco, C. M., and Zio, E. (2013). Singular spectrum analysis for forecasting of electric load demand. Chemical Engineering Transactions, 33, 919–924.

3. Brockwell, P. J., and Davis, R. A. (1991). Time Series: Theory and Methods (2nd ed.). Springer-Verlag.

4. Broomhead, D. S., Jones, R., and King, G. P. (1987). Topological dimension and local coordinates from time series data. Journal of Physics A: Mathematical and General, 20(9), L563.

5. Broomhead, D. S., and King, G. P. (1986). Extracting qualitative dynamics from experimental data. Physica D: Nonlinear Phenomena, 20(2-3), 217–236.

6. De Livera, A. M., Hyndman, R. J., and Snyder, R. D. (2011). Forecasting time series with complex seasonal patterns using exponential smoothing. Journal of the American Statistical Association, 106(496), 1513–1527.

7. Elsner, J. B. (2002). Analysis of Time Series Structure: SSA and Related Techniques. Journal of the American Statistical Association, 97(460), 1207–1208.

8. Elsner, J. B., and Tsonis, A. A. (1996). Singular Spectrum Analysis A New Tool in Time Series Analysis. Springer Science & Business Media.

9. Fraedrich, K. (1986). Estimating the dimensions of weather and climate attractors. Journal of the Atmospheric Sciences, 43(5), 419–432.

10. Golyandina, N. (2010). On the choice of parameters in singular spectrum analysis and related subspace-based methods. Stat Interface, 3(3), 259–279.

11. Golyandina, N., and Korobeynikov, A. (2014). Basic singular spectrum analysis and forecasting with R. Computational Statistics & Data Analysis, 71, 934–954.

12. Golyandina, N., Nekrutkin, V., and Zhigljavsky, A. (2001). Analysis of Time Series Structure: SSA and related techniques (Vol. 90). Chapman & Hall/CRC, Boca Raton, FL.

13. Golyandina, N., and Zhigljavsky, A. (2013). Singular Spectrum Analysis for time series. Springer Science & Business Media.

14. Hassani, H., Heravi, S., and Zhigljavsky, A. (2009). Forecasting European industrial production with singular spectrum analysis. International Journal of Forecasting, 25(1), 103–118.

15. Hassani, H., Soofi, A. S., and Zhigljavsky, A. (2010). Predicting Daily Exchange Rate with Singular Spectrum Analysis Data. Nonlinear Analysis: Real World Applications, 11(3), 2023–2034.

16. Hassani, H., Webster, A., Silva, E. S., and Heravi, S. (2015). Forecasting US tourist arrivals using optimal singular spectrum analysis. Tourism Management, 46, 322–335.

17. Hassani, H., and Zhigljavsky, A. (2009). Singular Spectrum Analysis: Methodology and Application to Economics Data. Journal of Systems Science and Complexity, 22(3), 372–394.

18. Hyndman, R., Koehler, A. B., Ord, J. K., and Snyder, R. D. (2008). Forecasting with exponential smoothing: the state space approach. Springer Science & Business Media.

19. Keeling, C. D., and Whorf, T. P. (1997). Atmospheric CO2 concentrations—Mauna Loa observatory, Hawaii, 1956-1997. Scripps Institution of Oceanography (SIO), University of California, La Jolla, California, USA, 92093-0220.

20. Khan, M. A. R., and Poskitt, D. S. (2011). Window Length Selection and Signal-Noise Separation and Reconstruction in Singular Spectrum Analysis. Monash Econometrics and Business Statistics Working Papers, 23(11), 2011–2023.

21. Khan, M. A. R., and Poskitt, D. S. (2014). On The Theory and Practice of Singular Spectrum Analysis Forecasting. Monash University, Department of Econometrics and Business Statistics.

22. Kitagawa, G. (2010). Introduction to time series modeling. Chapman & Hall, CRC press.

23. Li, H., Cui, L., and Guo, S. (2014). A hybrid short-term power load forecasting model based on the singular spectrum analysis and autoregressive model. Advances in Electrical Engineering, 2014.

24. Quinn, B. G. (1994). Estimating frequency by interpolation using Fourier coefficients. IEEE Transactions on Signal Processing, 42(5), 1264–1268.

25. Soares, L. J., and Medeiros, M. C. (2008). Modeling and forecasting short-term electricity load: A comparison of methods with an application to Brazilian data. International Journal of Forecasting, 24(4), 630–644.

26. Sulandari, W., Subanar, S., Suhartono, S., and Utami, H. (2017). Forecasting Time Series with Trend and Seasonal Patterns Based on SSA. In 2017 3rd International Conference on Science in Information Technology (ICSITech) "Theory and Applicattion of IT for Education, Industry and Society in Big Data Era" (pp. 694–699). Bandung, Indonesia: IEEE.

27. Vahabie, A. H., Yousefi, M. M. R., Araabi, B. N., Lucas, C., and Barghinia, S. (2007). Combination of singular spectrum analysis and autoregressive model for short term load forecasting. In Power Tech, 2007 IEEE Lausanne (pp. 1090–1093). IEEE.

28. Vautard, R., and Ghil, M. (1989). Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series. Physica D: Nonlinear Phenomena, 35(3), 395–424.

29. Vautard, R., and Ghil, M. (1991). Interdecadal oscillations and the warming trend in global temperature time series. Nature, 350(6316), 324.

30. Vautard, R., Yiou, P., and Ghil, M. (1992). Singular-spectrum analysis: A toolkit for short, noisy chaotic signals. Physica D: Nonlinear Phenomena, 58(1), 95–126.

31. Wei, W. W.S. (2006). Time series analysis: Univariate and Multivariate Methods (2nd ed.). Pearson Addison-Wesley.

32. Zhang, Q., Wang, B.-D., He, B., Peng, Y., and Ren, M.-L. (2011). Singular spectrum analysis and ARIMA hybrid model for annual runoff forecasting. Water Resources Management, 25(11), 2683–2703.