Weighted Linear Recurrent Forecasting in Singular Spectrum Analysis

Abstract

Singular Spectrum Analysis (SSA) is an increasingly popular time series filtering and forecasting technique. Owing to its widespread applications in a variety of fields, there is a growing interest towards improving its forecasting capabilities. As such, this paper takes into consideration the Recurrent forecasting approach in SSA (SSA-R) and presents a new mechanism for improving the accuracy of forecasts attainable via this method. The proposed Recurrent SSA-R approach is referred to as Weighted SSA-R (W:SSA-R), and we propose using a weighting algorithm for weighting the coefficients of the Linear Recurrent Relation (LRR). The performance of forecasts from the W:SSA-R approach are compared with forecasts from the established SSA-R approach. We exploit real data and various simulated time series for the comparison, so as to provide the reader with more conclusive findings. Our results confirm that the W:SSA-R approach can provide comparatively more accurate forecasts and is indeed a viable solution for improving forecasts by SSA.
Keywords: Time Series; Forecasting; Singular Spectrum Analysis; Recurrent forecasting.

1 Introduction

Recently, the Singular Spectrum Analysis (SSA) technique has gained considerable popularity as a powerful and effective time series filtering and forecasting tool. This popularity stems through its widespread and lucrative applications in a variety fields such as finance, economics, medicine, meteorology and genetics. As a result, there are continuous attempts at developing the underlying theory of SSA and improving its forecasting methods. Whilst the review of all applications of SSA are beyond the scope of this paper, those interested are referred to [1–15]. Nowadays, there is a bulk of research on SSA to develop its theory and applications and few such examples can be found in [16–22]. A detailed account of the theory and applications of SSA can be found in [23–25].

The SSA technique has two forecasting variations known as Recurrent and Vector forecasting which are referred to as SSA-R and SSA-V, respectively. Improving the performance of these forecasting methods constitutes an indispensable part of SSA development. For example, by drawing upon the general similarity between genetics Colonial Theory (CT), a nature inspired algorithm, and SSA; the performance of forecasting in subspace-based methods have been improved in [26] using CT. The SSA-V approach in particular has been widely adopted in applications as it has been declared in [23] that SSA-V is more robust than SSA-R when faced with time series which have unit root problems. Moreover, the SSA-V approach has shown better performance in the presence of outliers [27]. Recently, a comprehensive investigation was conducted in [28] to compare the forecasting capabilities of SSA-R and SSA-V forecasting algorithms via a simulation study and an application to 100 real data sets with varying structures from different fields. Statistically reliable results in [28] indicate that on average, SSA-V

forecasts are better in comparison to SSA-R as reported in [23, 29]. However, it is noteworthy that sample sizes and forecasting horizons were found to have an influence on which algorithm is more appropriate for forecasting with SSA.

Given that SSA-V outperforms SSA-R in most cases, there has been some interest in improving the forecasting capability of SSA-R. For example, in [30] the SSA-R algorithm has been promoted through generating the coefficients of the Linear Recurrent Relations (LRR) by filtered time series. In addition, a new parsimonious recurrent forecasting model has been proposed in [31]. Recently, a novel signal extraction approach for filtering and forecasting noisy exponential series has been introduced in [32]. Accordingly, this paper seeks to contribute further to the growing interest in improving SSA-R forecasts. However, in contrast to the aforementioned papers, here we seek to improve upon the forecasting capability of the SSA-R technique and develop a novel recurrent forecasting algorithm which can provide comparatively more accurate forecasts. In the novel approach, which we refer to as Weighted SSA-R, the coefficients of the LRR are weighed via a weighting algorithm. The performance of the new recurrent forecasting algorithm is evaluated and compared with the established SSA-R algorithm by utilizing various real and simulated time series. We obtain promising results which confirms that exploiting the Weighted SSA-R approach can lead to more accurate forecasts.

The remainder of this paper is organized as follows. Section 2 presents a review of SSA and the recurrent forecasting algorithm. The novel Weighted SSA-R forecasting approach is proposed in Section 3 and Section 4 is dedicated towards comparing the performance of Weighted SSA-R with the established SSA-R approach via a simulation study and an application to real time series data. The conclusions are presented in Section 5.

2 Singular Spectrum Analysis

The SSA method consists of two complementary stages: *Decomposition* and *Reconstruction*. Each of these stages includes two separate steps. At the first stage, the series is decomposed into several components in order to enable signal extraction and noise reduction. At the second stage, a less noisy series is reconstructed to be used to forecast new data points. More detailed information on the theory of basic SSA can be found in [23]. The basic SSA procedure is concisely presented below and in doing so we mainly follow [25, 38].

Stage 1: Decomposition (Embedding & Singular Value Decomposition)

In the embedding step, the multi-dimensional series $X_1, \ldots, X_K$ are built based on the original one dimensional time series $Y_N = \{y_1, \ldots, y_N\}$, where $X_i = (y_i, \ldots, y_{i+L-1})^T \in \mathbf{R}^L$ and $K = N - L + 1$. The vectors X_i are called *L-lagged vectors*. The Window Length L is the single choice of this step and is an integer such that $2 \leq L \leq N/2$. The output of the embedding step is the trajectory matrix $\mathbf{X} = [X_1 : \cdots : X_K]$, which is also a Hankel matrix.

In the Singular Value Decomposition (SVD) step, the trajectory matrix $\mathbf{X}$ is decomposed as a sum of rank-one elementary matrices. The eigenvalues of $\mathbf{X}\mathbf{X}^T$ are denoted by $\lambda_1, \ldots, \lambda_L$ in decreasing order of magnitude ($\lambda_1 \geq \cdots \geq \lambda_L \geq 0$) and by $U_1, \ldots, U_L$, the eigenvectors of the matrix $\mathbf{X}\mathbf{X}^T$ corresponding to these eigenvalues. If $d = \max\{i, \text{such that } \lambda_i > 0\} = rank(\mathbf{X})$ then the SVD of the trajectory matrix can be written as $\mathbf{X} = \mathbf{X}_1 + \cdots + \mathbf{X}_d$, where $\mathbf{X}_i = \sqrt{\lambda_i} U_i V_i^T$ and $V_i = \mathbf{X}^T U_i / \sqrt{\lambda_i}$ ($i = 1, \ldots, d$). The collection $(\sqrt{\lambda_i}, U_i, V_i)$ is called ith *eigentriple* of SVD.

Stage 2: Reconstruction (Grouping & Diagonal Averaging)

The grouping step splits the elementary matrices $\mathbf{X}_i$ into several groups and sums the matrices within each group. If a group of indices $i_1, \ldots, i_p$ is denoted by $I = \{i_1, \ldots, i_p\}$ then the matrix $\mathbf{X}_I$ corresponding to the group I is defined as $\mathbf{X}_I = \mathbf{X}_{i_1} + \cdots + \mathbf{X}_{i_p}$. Having the SVD of $\mathbf{X}$, the split of the set of indices $\{1, \ldots, d\}$ into the disjoint subsets $I_1, \ldots, I_m$ corresponds to the following representation:

$$\mathbf{X} = \mathbf{X}_{I_1} + \cdots + \mathbf{X}_{I_m}. \tag{1}$$

Diagonal averaging is a process which transforms each matrix $\mathbf{X}_{I_j}$ of the grouped decomposition (1) into a Hankel matrix so that these can subsequently be converted into a time series, which is an additive component of the initial series Y_N. Assume that z_{ij} stands for an element of a matrix $\mathbf{Z}$, then the k-th term of the resulting series is obtained by averaging z_{ij} over all i, j such that $i + j = k + 2$. This procedure is also known as Hankelization of the matrix $\mathbf{Z}$. The output of the Hankelization of a matrix $\mathbf{Z}$ is the Hankel matrix $\mathcal{H}\mathbf{Z}$, which is the trajectory matrix corresponding to the series obtained as a result of diagonal averaging. In its turn, the Hankel matrix $\mathcal{H}\mathbf{Z}$ uniquely defines the series by relating the value in the anti-diagonals to the values in the series. By applying the Hankelization procedure to all matrix components of (1), this expansion is obtained: $\mathbf{X} = \widetilde{\mathbf{X}}_{I_1} + \cdots + \widetilde{\mathbf{X}}_{I_m}$, where $\widetilde{\mathbf{X}}_{I_j} = \mathcal{H}\mathbf{X}_{I_j}$, $j = 1, \ldots, m$. This is equivalent to the decomposition of the initial series $Y_N = \{y_1, \ldots, y_N\}$ into a sum of m series: $y_t = \sum_{k=1}^{m} \widetilde{y}_t^{(k)} (t = 1, \ldots, N)$, where $\widetilde{Y}_N^{(k)} = \{\widetilde{y}_1^{(k)}, \ldots, \widetilde{y}_N^{(k)}\}$ corresponds to the matrix $\widetilde{\mathbf{X}}_{I_k}$.

The SSA technique has the capability of generating forecasts using the filtered time series following the decomposition and reconstruction stages which are crucial for signal extraction and noise filtering. In what follows, the SSA-R forecasting method is discussed at length.

2.1 Recurrent SSA Forecasting (SSA-R)

The main assumption to perform SSA forecasting is that the time series satisfies a Linear Recurrent Relation (LRR). The time series $Y_N = \{y_1, \ldots, y_N\}$ satisfies an LRR of order d if there exist the coefficients $a_1, \ldots, a_d$ such that:

$$y_{i+d} = \sum_{k=1}^{d} a_k y_{i+d-k}, \quad 1 \leq i \leq N - d, \quad a_d \neq 0, \quad d < N.$$

The coefficients $a_1, \ldots, a_d$ are called LRR-coefficients. The class of time series governed by LRRs which is rather wide and important for practical applications, contains the series that are linear combinations of products of exponential, polynomial and harmonic series [24].

Let I be the chosen set of eigentriples attained at the grouping step of SSA. For example in trend forecasting, I is the trend group and in harmonic components forecasting, I corresponds to harmonic groups. To obtain forecasts of reconstructed series (signal forecasting), the first r eigentriples can be selected.

Let $U_i \in \mathbf{R}^L, i \in I$ be the corresponding eigenvectors of chosen eigentriples, $\underline{U_i} \in \mathbf{R}^{L-1}$ be the vector consisting of the first $L-1$ components of the vector U_i, π_i be the last component of the vector U_i, $v^2 = \sum_{i \in I} \pi_i^2$, and $\widetilde{Y}_N = \{\widetilde{y}_1, \ldots, \widetilde{y}_N\}$ be the time series reconstructed by set I. Denote by $\mathcal{L} \subset \mathbf{R}^L$ the linear space spanned by the vectors $U_i, i \in I$; i.e., $\mathcal{L} = span(U_i, i \in I)$. Note that the set $\{U_i, i \in I\}$ forms an orthonormal basis in $\mathcal{L}$. It is assumed that $e_L \notin \mathcal{L}$, where $e_L = (0, 0, \ldots, 1)^T \in \mathbf{R}^L$; in other term, $\mathcal{L}$ is not a *vertical* space. Since $e_L \notin \mathcal{L}$, $v^2 < 1$. It can be proved that the last component z_L of any vector $Z = (z_1, \ldots, z_L)^T \in \mathcal{L}$ is a linear

combination of the first components $z_1, \ldots, z_{L-1}$, i.e., $z_L = a_1 z_{L-1} + \cdots + a_{L-1} z_1$, (see [24]) where the vector $R = (a_{L-1}, \ldots, a_1)^T$ is defined as:

$$R = \frac{1}{1 - v^2} \sum_{i \in I} \pi_i \underline{U_i}. \tag{2}$$

The SSA-R forecasting algorithm can then be summarized as follows.

1. The time series $Z_{N+h} = \{z_1, \ldots, z_{N+h}\}$ is defined by

$$z_i = \begin{cases} \widetilde{y}_i & \text{for } i = 1, \ldots, N \\ \sum_{j=1}^{L-1} a_j z_{i-j} & \text{for } i = N+1, \ldots, N+h \end{cases} \tag{3}$$

2. The numbers $z_{N+1}, \ldots, z_{N+h}$ are the h step-ahead recurrent forecasts.

It is clear that SSA-R forecasting is performed by the direct use of the LRR with coefficients $\{a_j, j = 1, \ldots, L-1\}$ as defined in (2).

Next, we introduce the newly proposed SSA-R forecasting algorithm.

3 Weighted SSA-R Forecasting Algorithm

It is well known that in real applications, the time series $Y_N = \{y_1, \ldots, y_N\}$ are not noise-free. Consequently, as shown in [30], the LRR-coefficients are contaminated with noise. It is noteworthy that these coefficients play a fundamental role in SSA-R forecasting as seen in (3), and using inadequate coefficients results in low accuracy predictions. As such, we seek to reduce the effect of noise level by weighting the LRR-coefficients in an appropriate manner. This simple, yet lucrative idea leads us to a new SSA-R forecasting algorithm whereby its LRR-coefficients are computed as follows:

$$R_w^{(h)} = w_h (a_{L-1}, \ldots, a_1)^T = \frac{w_h}{1 - v^2} \sum_{i \in I} \pi_i \underline{U_i}, \tag{4}$$

where w_h is the weight at forecast horizon h, defined as:

$$w_h = \frac{y_{N+h}}{z_{N+h}}. \tag{5}$$

Having the weighted LRR-coefficients from (4), the weighted SSA-R forecasting algorithm can be proposed as follows.

1. The time series $Z_{N+h} = \{z_1, \ldots, z_{N+h}\}$ is defined by

$$z_i = \begin{cases} \widetilde{y}_i & \text{for } i = 1, \ldots, N \\ \sum_{j=1}^{L-1} w_{i-N} a_j z_{i-j} & \text{for } i = N+1, \ldots, N+h \end{cases} \tag{6}$$

2. The numbers $z_{N+1}, \ldots, z_{N+h}$ are the h step-ahead weighted SSA-R forecasts.

Evidently, the weight w_h defined in (5) can only be used when y_{N+h} and z_{N+h} are known. However, in reality, y_{N+h} and z_{N+h} are not available because it is assumed that the observed time series $Y_N = \{y_1, \ldots, y_N\}$ has only N available data. To resolve this issue, we propose the following weighting scheme.

1. Split the time series $Y_N = \{y_1, \ldots, y_N\}$ into two parts. Use the first $\frac{2}{3}$rd observations to construct many different training sets, each one containing one more observation than the previous one. Apply the remaining $\frac{1}{3}$rd of observations to build corresponding test sets, each one containing one fewer observation than the previous one (see Figure 1).

2. First, use the observations of the ith training set at times $1, 2, \ldots, M+i-1$ to forecast the observation $y_{M+h+i-1}$ in the ith test set, where M is the minimum size of the training set. Then, compute $w_h^{(i)} = \dfrac{y_{M+h+i-1}}{z_{M+h+i-1}}$ (see Figure 2).

3. Repeat the Step 2 for $i = 1, 2, \ldots, N-M-h+1$.

4. Compute the weight w_h as $w_h = \underset{1 \le i \le N-M-h+1}{\text{median}} \{w_h^{(i)}\}$.

The reason for applying *median* instead of *mean* in Step 4 of the above weighting algorithm is that the median is robust against outlier weights which may be obtained via Step 2.

It is noteworthy that the proposed weight in (5) is closely related to the relative error. If the relative forecasting error at horizon h is denoted by r_h then, $r_h = \dfrac{y_{N+h} - z_{N+h}}{y_{N+h}} = 1 - \dfrac{z_{N+h}}{y_{N+h}} = 1 - \dfrac{1}{w_h}$. Hence, it can be concluded that $w_h = \frac{1}{1-r_h}$ and weighted LRR-coefficients in (4) can be computed as follows:

$$R_w^{(h)} = \frac{1}{1-r_h}(a_{L-1}, \ldots, a_1)^T = \frac{1}{(1-r_h)(1-v^2)} \sum_{i \in I} \pi_i \underline{U_i}, \tag{7}$$

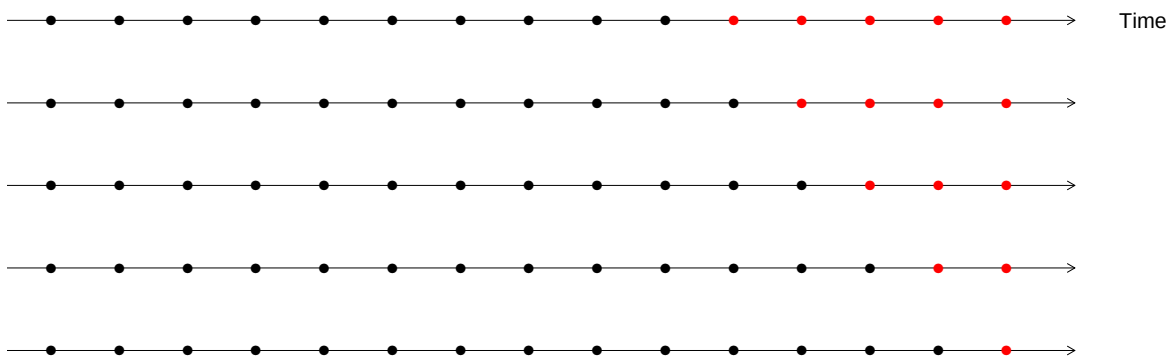

Figure 1: Training (black) and test (red) sets for $N = 15, M = 10$.

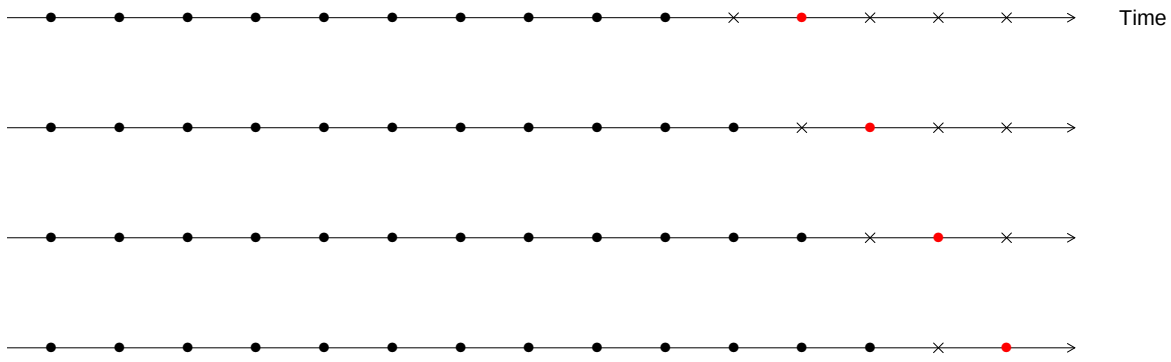

Figure 2: Computing the weights based on two-step forecasts for $N = 15, M = 10$. The black points are training sets, the red points are test sets and other points are ignored.

4 Empirical Results

In this section, the performance of SSA-R and Weighted SSA-R forecasting algorithms are evaluated by applying them to simulated time series and real data. The time series $Y_N = \{y_1, \ldots, y_N\}$ is divided into two parts to build training and test sets as explained in Step 1 of weighting algorithm in Section 3. The training sets are used to produce SSA forecasts. The accuracy of forecasting results are measured using the widely used metric, Root Mean Squared Error (RMSE). The following RMSE ratio is used to compare the SSA-R and newly proposed Weighted SSA-R forecasting algorithms:

$$RRMSE_h = \frac{RMSE_h(\text{weighted SSA-R})}{RMSE_h(\text{SSA-R})} = \frac{\left(\sum_{t=m}^{N-h}(y_{t+h} - \hat{y}_{t+h|t})^2\right)^{1/2}}{\left(\sum_{t=m}^{N-h}(y_{t+h} - \hat{\hat{y}}_{t+h|t})^2\right)^{1/2}}, \tag{8}$$

where, M is the minimum length of training sets, N is the length of the time series Y_N, h is the length of the forecast horizon, $\hat{y}_{t+h|t}$ and $\hat{\hat{y}}_{t+h|t}$ are the h step-ahead forecasts obtained from Weighted SSA-R and SSA-R, respectively. If the $RRMSE_h < 1$ then it can be concluded that the Weighted SSA-R forecast outperforms SSA-R forecasts at horizon h. Alternatively, when $RRMSE_h > 1$, it would indicate that the performance of SSA-R forecasting is better than that of the Weighted SSA-R forecasting approach. In order to enable better comparison, a dashed horizontal line $y = 1$ is added to all RRMSE figures.

4.1 Simulation Study

In the following two simulated series, 100 data points are generated and normally distributed noise is added to each point of the series. The minimum length of training sets is set to 67, i.e., $M = 67$. The number of eigentriples for reconstruction and forecasting (r) was selected according to the rank of the corresponding trajectory matrix. The simulation was repeated 1000 times and mean of RMSEs were calculated to compute $RRMSE_h$ as defined in (8). In order to assess the impact of noise levels on forecasting results, various signal to noise ratios (SNR) were employed as $SNR = 0.25, 0.5, 0.75, 1, 5, 10$.

Example 4.1. Consider the Exponential series:

$$y_t = \exp(0.01t) + \varepsilon_t, \quad t = 1, 2, \ldots, 100,$$

where ε_t is the normally distributed noise series with zero mean. The first eigentriple was selected for reconstruction and forecasting ($r = 1$). Figure 3 shows the RRMSE of 1, 3, 6, and 12 steps-ahead forecast horizons for Exponential series. At each of forecast horizons h, different values of SNR's have been used. As can be seen in this figure, for all forecast horizons, the Weighted SSA-R forecasting method always outperforms SSA-R, especially for greater values of window length (L). It can also be seen that RRMSEs are very close together for various SNR's.

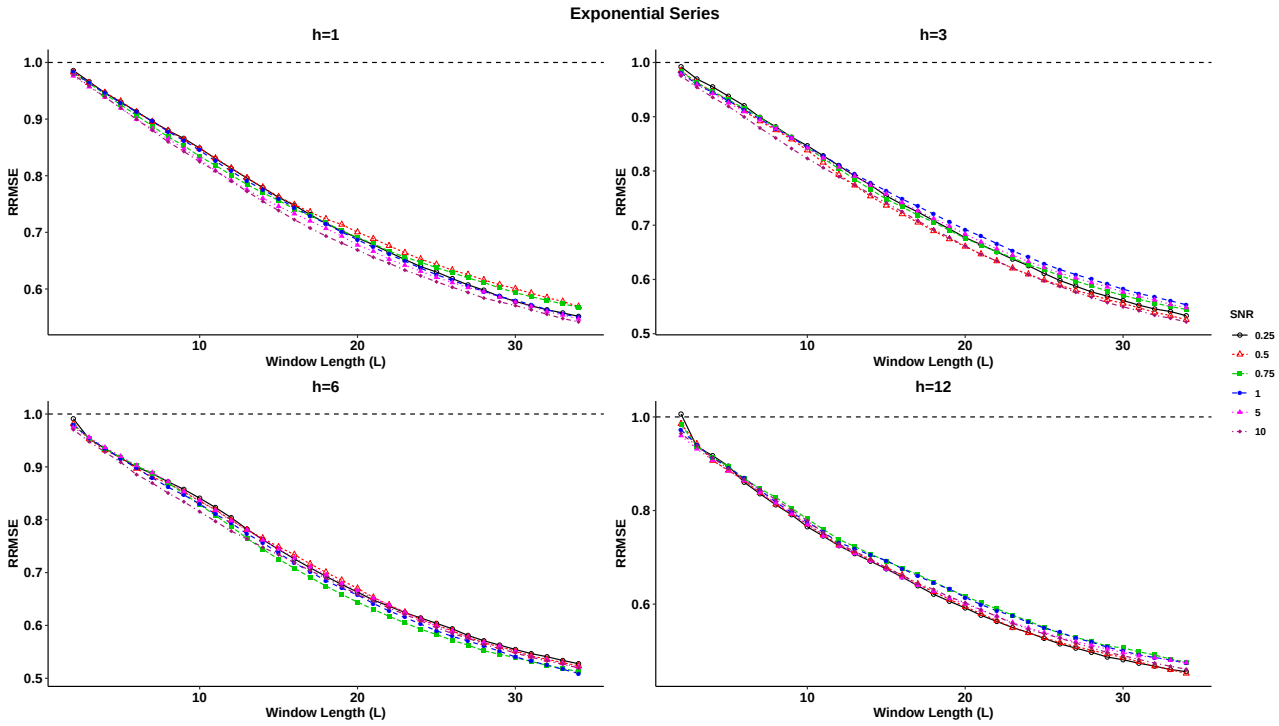

Figure 3: RRMSE for Exponential series.

Example 4.2. As a second example, consider the Sine series:

$$y_t = \sin(\pi t/3) + \varepsilon_t, \quad t = 1, 2, \ldots, 100,$$

where ε_t is the normally distributed noise series with zero mean. The first two eigentriples were selected for reconstruction and forecasting ($r = 2$). Figure 4 shows the RRMSE of 1, 3, 6, and 12 steps-ahead forecast horizons for the Sine series. Similar to Example 4.1, different values of SNR's have been used for each h and it can be concluded from this figure that for all forecast horizons, the Weighted SSA-R forecasting method always outperforms SSA-R forecasts. As can be seen in the RRMSE figure for this simulated series, unlike the simulated Exponential series, there are some differences between RRMSEs for various SNR's at lower values of window length (L).

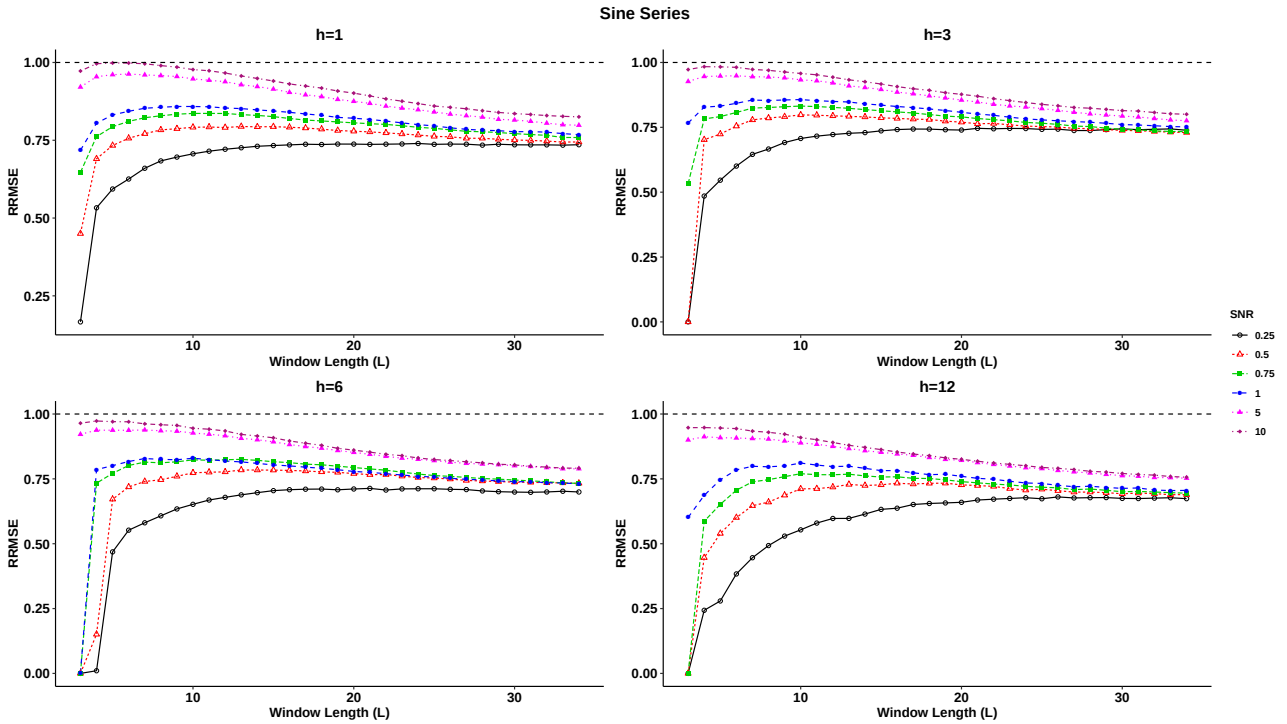

Figure 4: RRMSE for Sine series.

4.2 Real Data

Here, let us compare the efficiency of SSA-R and Weighted SSA-R forecasting algorithms using real data. Figures 5-7 show the time series plots of these data sets. These datasets were compiled through the INSEE (Institute National de la Statistique et des Etudes Economiques) for France, from Statistisches Bundesamt, Wesbaden for Germany, and from the Office for National Statistics (ONS) for the UK. It includes eight major components of real industrial production in France, Germany, and the UK. The series are seasonally unadjusted monthly indices for real output in Electricity/Gas, Chemicals, Fabricated Metals, Vehicles, Food Products, Basic Metals, Electrical Machinery and Machinery.

These same series were previously used in [39–41] and are considered to be important as they account for more than 50% of the total industrial production in each country. Those interested a summary of the data are referred to [39] instead of replicating this information here. In brief, the time series plots for these data clearly illustrate how they capture the effects of seasonality, non-stationarity and structural breaks, which adds further value to its choice as the real world examples in this study.

Tables 1–3 report the out-of-sample forecasting RRMSE's. Note that to help with replication, we also report the SSA choices used to generate each forecast within these same tables. We begin our analysis by considering the forecasts for industrial production in France as reported in Table 1. In this case, we notice the W:SSA-R model outperforms SSA-R across all horizons for all data. However, it is clear that for some series (e.g. Chemicals), the average accuracy gains reported by the W:SSA-R model is lower than for certain other series (e.g. Machinery).

Likewise, Table 2 reports the out-of-sample forecasts for industrial production in UK. Once again, for the three series reported here, we find the W:SSA-R model outperforming SSA-R.

In Table 3, the forecasting results for Germany are made available. Here too, we find the newly proposed W:SSA-R model outperforming the SSA-R model in terms of accuracy of

forecasts across all horizons.

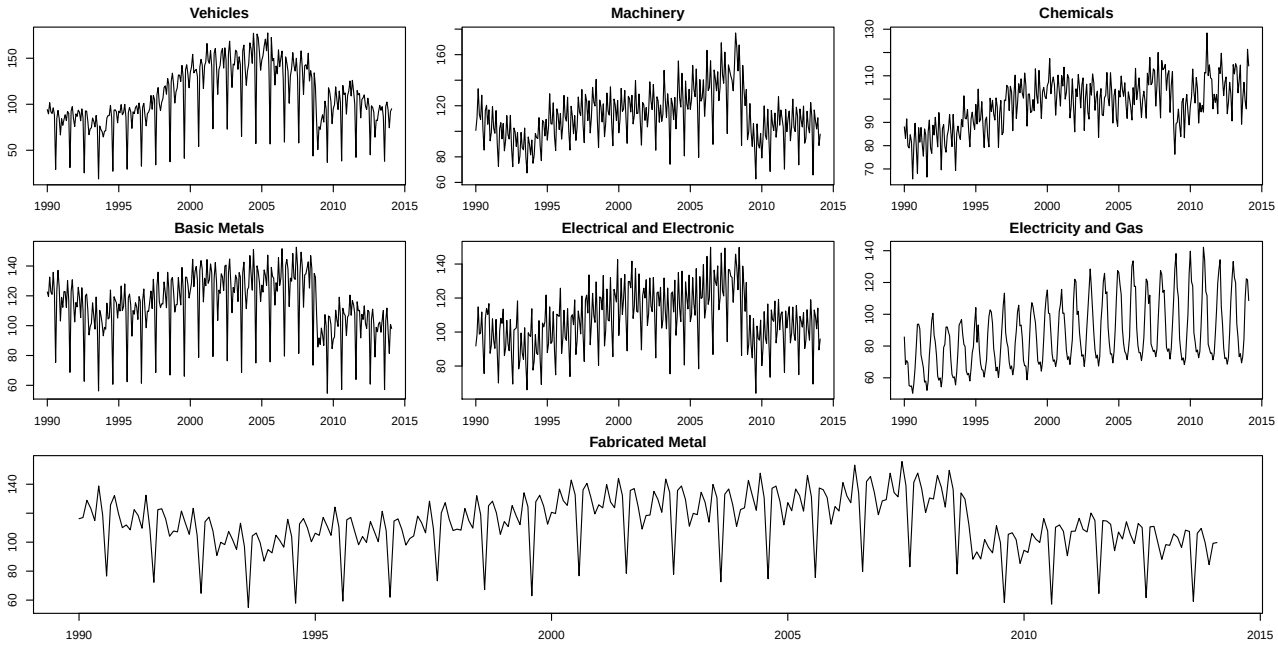

Figure 5: Time series plot of France industrial production data.

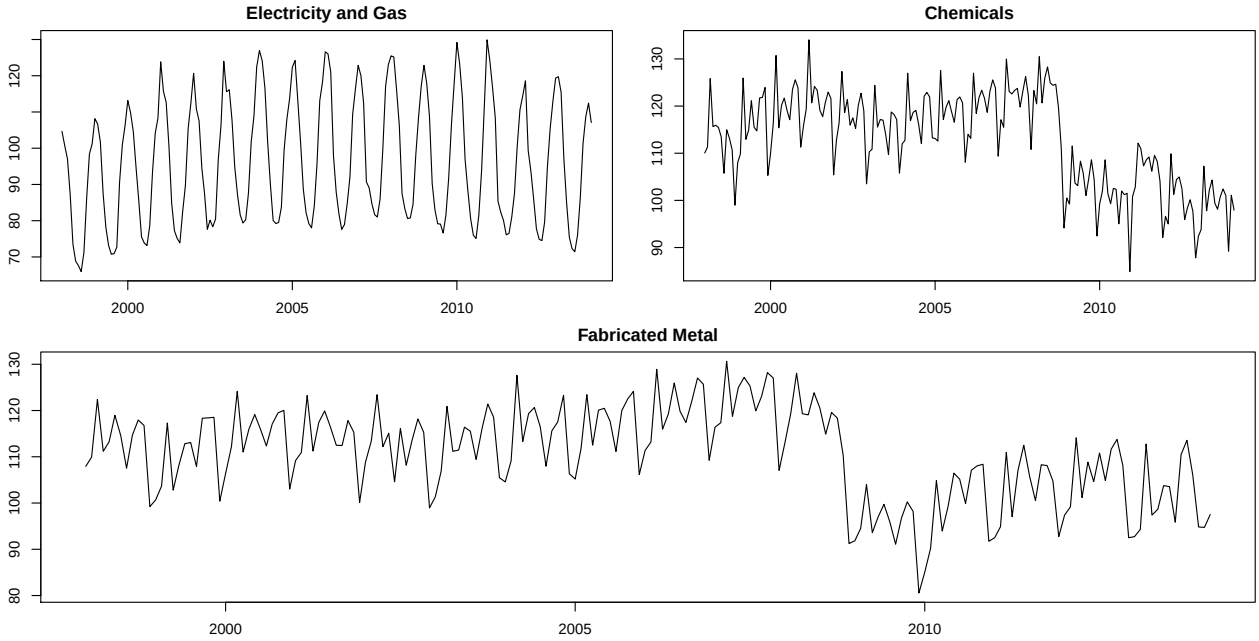

Figure 6: Time series plot of UK industrial production data.

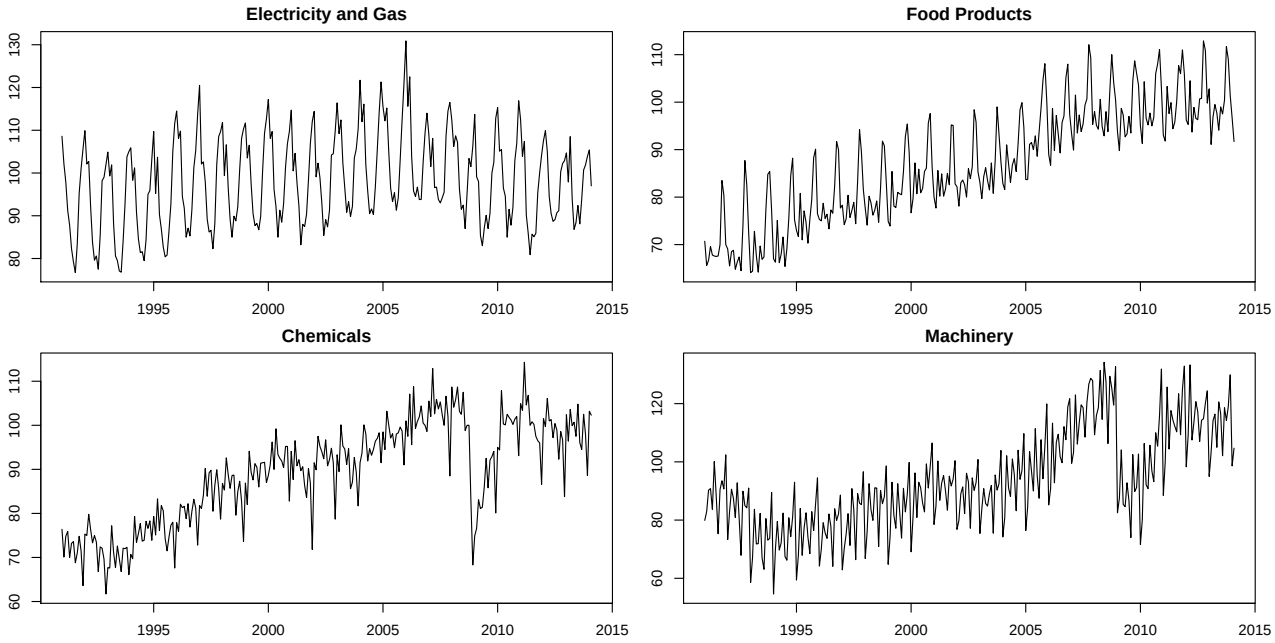

Figure 7: Time series plot of Germany industrial production data.

5 Conclusion

In this paper, we proposed a new forecasting algorithm for univariate time series by improving the capability of the recurrent forecasting approach within the SSA framework. In the proposed technique named Weighted SSA-R, the LRR coefficients were weighed via the weighting algorithm to obtain better forecasts.

The performance of the Weighted SSA-R algorithm was compared with the established SSA-R method with respect to the RMSE criteria using both simulated and real world data. The simulation in particular considered the impact of varying signal to noise ratios on the performance of the proposed algorithm. The simulation study indicated the superiority of the Weighted SSA-R forecasts in comparison to SSA-R forecasts across all horizons and for both exponential and sine series. Following the positive results from the simulation study, it was imperative to ascertain the performance of the proposed algorithm when faced with real data.

It is important to note that all real data sets considered in this study were previously adopted in other studies (as noted previously), thus validating its consideration here. Moreover, they related to annual, quarterly, monthly and daily frequencies, thereby providing the reader with a more comprehensive outlook on the performance of the proposed algorithm. The application to real data gave an indication of the algorithms sensitiveness to frequencies, forecasting horizons and choice of Window Length.

However, in general we found substantial evidence which supports the viability of the Weighted SSA-R algorithm. The findings of this paper opens up several research avenues relating to the improvement of the Weighted SSA-R algorithm further and calls for an increased application and comparison of the Weighted SSA-R algorithm not only with SSA-R forecast, but also with forecasts from SSA-V and other popular univariate forecasting models such as ARIMA, Holt-Winters, Exponential Smoothing and Neural Networks (to name a few).

	h	RRMSE (L, r)				
Vehicles	1	0.92 (13,9)	0.98 (21,14)	0.94 (50,14)	0.96 (73,21)	0.96 (90,22)
	3	0.90 (13,9)	0.96 (20,14)	0.96 (34,18)	0.95 (73,14)	0.92 (85,14)
	6	0.85 (15,8)	0.81 (21,8)	0.90 (29,12)	0.96 (80,14)	0.94 (96,14)
	12	0.86 (13,9)	0.80 (21,6)	0.70 (32,4)	0.98 (44,13)	0.92 (87,14)
Machinery	1	0.98 (23,14)	0.98 (34,16)	0.97 (47,16)	0.98 (81,18)	0.96 (92,20)
	3	0.97 (23,14)	0.98 (33,16)	0.96 (40,15)	0.97 (50,16)	0.97 (85,18)
	6	0.98 (20,3)	0.92 (35,3)	0.95 (55,15)	0.93 (76,16)	0.94 (85,16)
	12	0.95 (23,3)	0.87 (35,3)	0.82 (46,2)	0.80 (68,3)	0.81 (78,2)
Chemicals	1	0.98 (35,18)	0.98 (61,18)	0.99 (68,22)	0.99 (84,23)	0.99 (95,24)
	3	0.98 (34,5)	0.97 (46,4)	0.97 (65,3)	0.97 (81,8)	0.96 (96,8)
	6	0.97 (35,5)	0.96 (50,4)	0.96 (65,3)	0.97 (80,8)	0.96 (93,8)
	12	0.97 (43,5)	0.96 (57,3)	0.97 (66,3)	0.97 (81,8)	0.95 (95,8)
Basic Metals	1	0.97 (13,11)	0.98 (22,16)	0.97 (35,17)	0.96 (47,17)	0.97 (62,17)
	3	0.98 (9,7)	0.98 (20,15)	0.98 (33,14)	0.95 (47,16)	0.96 (59,17)
	6	0.95 (13,8)	0.93 (21,8)	0.90 (37,8)	0.93 (47,16)	0.93 (52,17)
	12	0.91 (20,6)	0.84 (34,5)	0.78 (42,5)	0.85 (55,8)	0.86 (91,8)
Electrical and Electronic	1	0.98 (23,14)	0.97 (58,20)	0.98 (70,21)	0.98 (81,21)	0.94 (88,22)
	3	0.98 (17,14)	0.97 (42,25)	0.95 (60,20)	0.95 (82,22)	0.96 (90,25)
	6	0.98 (18,5)	0.94 (27,5)	0.97 (38,5)	0.95 (67,17)	0.91 (83,5)
	12	0.94 (22,4)	0.94 (35,5)	0.82 (48,2)	0.91 (70,5)	0.81 (92,5)
Electricity and Gas	1	0.95 (14,6)	0.88 (26,6)	0.86 (33,8)	0.97 (39,9)	0.97 (52,14)
	3	0.95 (11,5)	0.87 (25,6)	0.84 (31,6)	0.97 (40,9)	0.98 (52,10)
	6	0.95 (10,7)	0.92 (20,6)	0.86 (28,6)	0.90 (37,9)	0.97 (49,10)
	12	0.90 (14,5)	0.83 (26,6)	0.88 (37,9)	0.98 (50,10)	0.97 (61,10)
Fabricated Metal	1	0.98 (13,11)	0.98 (53,20)	0.98 (62,22)	0.98 (77,23)	0.98 (92,24)
	3	0.97 (13,11)	0.97 (28,12)	0.97 (52,20)	0.98 (60,23)	0.97 (80,17)
	6	0.96 (13,11)	0.97 (29,12)	0.89 (40,12)	0.97 (58,16)	0.97 (75,16)
	12	0.87 (33,5)	0.78 (42,5)	0.87 (62,7)	0.87 (87,7)	0.91 (95,7)

Table 1: Out-of-sample RRMSE for France industrial production.

	h	RRMSE (L, r)				
Electricity and Gas	1	0.97 (13,6)	0.90 (26,7)	0.97 (44,9)	0.98 (52,8)	0.98 (61,8)
	3	0.87 (15,3)	0.81 (22,3)	0.95 (30,5)	0.95 (41,8)	0.94 (50,8)
	6	0.87 (15,3)	0.80 (25,8)	0.95 (30,9)	0.92 (43,9)	0.92 (50,8)
	12	0.78 (15,3)	0.72 (25,4)	0.92 (38,12)	0.93 (45,5)	0.95 (56,11)
Chemicals	1	0.95 (12,6)	0.98 (18,3)	0.96 (46,5)	0.98 (55,9)	0.94 (64,8)
	3	0.96 (10,1)	0.98 (18,7)	0.96 (37,15)	0.93 (47,18)	0.84 (56,8)
	6	0.94 (9,1)	0.95 (15,2)	0.91 (35,21)	0.79 (44,8)	0.82 (64,21)
	12	0.91 (11,1)	0.84 (22,1)	0.75 (35,1)	0.88 (42,9)	0.81 (60,21)
Fabricated Metal	1	0.98 (8,6)	0.97 (12,8)	0.99 (22,15)	0.99 (47,21)	0.98 (61,22)
	3	0.98 (10,1)	0.99 (14,13)	0.98 (29,21)	0.98 (36,16)	0.97 (61,22)
	6	0.98 (14,6)	0.97 (29,21)	0.89 (38,1)	0.80 (53,1)	0.72 (64,1)
	12	0.96 (25,1)	0.90 (37,1)	0.86 (45,1)	0.78 (54,1)	0.71 (64,1)

Table 2: RRMSE for real data of UK.

	h	RRMSE (L, r)				
Electricity and Gas	1	0.98 (11,3)	0.98 (42,13)	0.96 (57,14)	0.96 (74,15)	0.90 (92,16)
	3	0.96 (11,3)	0.82 (27,3)	0.96 (61,5)	0.89 (78,8)	0.88 (90,16)
	6	0.95 (10,4)	0.89 (21,3)	0.79 (30,3)	0.98 (52,4)	0.86 (92,11)
	12	0.94 (11,3)	0.80 (25,3)	0.87 (74,11)	0.78 (82,10)	0.91 (91,5)
Food Products	1	0.90 (15,8)	0.98 (25,11)	0.96 (35,13)	0.97 (46,14)	0.97 (59,15)
	3	0.90 (13,8)	0.96 (26,11)	0.96 (33,13)	0.97 (42,14)	0.96 (61,15)
	6	0.88 (15,8)	0.94 (30,10)	0.94 (47,14)	0.94 (60,15)	0.95 (91,25)
	12	0.80 (14,9)	0.87 (33,13)	0.77 (50,10)	0.76 (61,10)	0.87 (92,9)
Chemicals	1	0.99 (12,4)	0.99 (30,18)	0.98 (60,24)	0.98 (73,24)	0.97 (85,17)
	3	0.96 (14,11)	0.98 (32,4)	0.88 (61,1)	0.84 (75,1)	0.84 (84,1)
	6	0.97 (16,1)	0.97 (32,1)	0.91 (50,1)	0.82 (73,1)	0.82 (91,1)
	12	0.95 (13,1)	0.95 (25,1)	0.93 (34,1)	0.84 (50,1)	0.76 (87,1)
Machinery	1	0.98 (36,18)	0.99 (62,24)	0.98 (71,18)	0.97 (82,18)	0.98 (90,18)
	3	0.99 (18,11)	0.97 (36,18)	0.96 (57,19)	0.98 (70,18)	0.98 (83,21)
	6	0.99 (19,18)	0.97 (38,18)	0.98 (44,18)	0.97 (71,18)	0.98 (78,18)
	12	0.85 (20,18)	0.88 (38,18)	0.94 (43,18)	0.88 (71,18)	0.98 (81,3)
Vehicles	1	0.99 (14,12)	0.99 (23,15)	0.97 (52,21)	0.97 (77,21)	0.99 (84,14)
	3	0.99 (15,12)	0.98 (22,21)	0.98 (53,13)	0.96 (69,11)	0.96 (78,21)
	6	0.99 (13,6)	0.98 (24,21)	0.99 (37,5)	0.96 (65,1)	0.91 (91,1)
	12	0.98 (19,3)	0.98 (33,5)	0.96 (61,1)	0.91 (76,1)	0.89 (90,1)

Table 3: RRMSE for real data of Germany.

References

[1] Aydin, S., Saraoglu, H. M., and Kara, S. (2011). Singular Spectrum Analysis of Sleep EEG in Insomnia, *Journal of Medical Systems*, **35**(4), 457–461.

[2] Muruganatham, B., Sanjith, M. A., Krishnakumar, B., and Satya Murty, S. A. V. (2013). Roller element bearing fault diagnosis using singular spectrum analysis, *Mechanical Systems and Signal Processing*, **35**(1-2), 150–166.

[3] Chen, Q., Dam, T. V., Sneeuw, N., Collilieux, X., Weigelt, M., and Rebischung, P. (2013). Singular spectrum analysis for modeling seasonal signals from GPS time series, *Journal of Geodynamics*, **72**, 25–35.

[4] Hou, Z., Wen, G., Tang, P., and Cheng, G. (2014). Periodicity of Carbon Element Distribution Along Casting Direction in Continuous-Casting Billet by Using Singular Spectrum Analysis, *Metallurgical and Materials Transactions B*, **45** (5), 1817–1826.

[5] Liu, K., Law, S. S., Xia, Y., and Zhu, X. Q. (2014). Singular spectrum analysis for enhancing the sensitivity in structural damage detection, *Journal of Sound and Vibration*, **333** (2), 392–417.

[6] Bail, K. L., Gipson, J. M., and MacMillan, D. S. (2014). Quantifying the Correlation Between the MEI and LOD Variations by Decomposing LOD with Singular Spectrum Analysis, *Earth on the Edge: Science for a Sustainable Planet International Association of Geodesy Symposia*, **139**, 473–477.

[7] Chao, H-S., and Loh, C-H. (2014). Application of singular spectrum analysis to structural monitoring and damage diagnosis of bridges, *Structure and Infrastructure Engineering: Maintenance, Management, Life-Cycle Design and Performance*, **10** (6), 708–727.

[8] Hassani, H., Webster, A., Silva, E. S. and Heravi, S. (2015). Forecasting U.S. Tourist arrivals using optimal Singular Spectrum Analysis, *Tourism Management*, **46**, 322–335.

[9] Silva, E. S. and Hassani, H. (2015). On the use of singular spectrum analysis for forecasting U.S. trade before, during and after the 2008 recession, *International Economics*, **141**, 34–49.

[10] Ghodsi, Z., Silva, E. S. and Hassani, H. (2015). Bicoid Signal Extraction with a Selection of Parametric and Nonparametric Signal Processing Techniques, *Genomics Proteomics Bioinformatics*, **13**(3), 183—191.

[11] Mahmoudvand, R. and Rodrigues, P. C. (2016). Missing value imputation in time series using singular spectrum analysis, *International Journal of Energy and Statistics*, **4**(1), 1650005.

[12] Carvalho, M. and Rua, A. (2017). Real-time nowcasting the US output gap: Singular spectrum analysis at work, *International Journal of Forecasting*, **33**(1), 185-198.

[13] Hassani, H., Silva, E. S. and Ghodsi, Z. (2017). Optimizing bicoid signal extraction, *Mathematical Biosciences*, **294**, 46–56.

[14] Silva, E. S., Ghodsi, Z., Ghodsi, M., Heravi, S. and Hassani, H. (2017). Cross country relations in European tourist arrivals, *Annals of Tourism Research*, **63**, 151–168.

[15] Silva, E. S., Hassani, H., Heravi, S., and Huang, X. (2019). Forecasting tourism demand with denoised neural networks, *Annals of Tourism Research*, **74**, 134–154.

[16] Hassani, H. (2010). Singular Spectrum Analysis Based on the Minimum Variance Estimator, *Nonlinear Analysis: Real World Applications*, **11**, 2065–2077.

[17] Hassani, H., Xu, Z. and Zhigljavsky, A. (2011). Singular spectrum analysis based on the perturbation theory, *Nonlinear Analysis: Real World Applications*, **12**, 2752–2766.

[18] Wang, R., Ma, H. G., Liu, G. Q. and Zuo, D. G. (2015). Selection of window length for singular spectrum analysis, *Journal of the Franklin Institute*, **352**, 1541–1560.

[19] Alharbi, N. and Hassani, H. (2016). A new Approach for Selecting the Number of the Eigen Values in Singular Spectrum Analysis, *Journal of the Franklin Institute*, **353**, 1–16.

[20] Kalantari, M., Yarmohammadi, M. and Hassani, H. (2016). Singular Spectrum Analysis Based on L_1-norm, *Fluctuation and Noise Letters*, **15**(1), 1650009.

[21] Papailias, F. and Thomakos, D. (2017). EXSSA: SSA-based reconstruction of time series via exponential smoothing of covariance eigenvalues, *International Journal of Forecasting*, **33**(1), 214–229.

[22] Rodrigues, P. C., Tuy, P. D. S. E., and Mahmoudvand, R. (2018). Randomized singular spectrum analysis for long time series, *Journal of Statistical Computation and Simulation*, DOI: 10.1080/00949655.2018.1462810.

[23] Golyandina, N., Nekrutkin, V. and Zhigljavsky, A. (2001). *Analysis of Time Series Structure: SSA and Related Techniques*, Chapman & Hall/CRC.

[24] Golyandina, N. and Zhigljavsky, A. (2013). *Singular Spectrum Analysis for Time Series*, Springer Briefs in Statistics, Springer.

[25] Sanei, S. and Hassani, H. (2016). *Singular Spectrum Analysis of Biomedical Signals*, Taylor & Francis/CRC.

[26] Hassani, H., Ghodsi, Z., Silva, E. S. and Heravi, S. (2016). From nature to maths: Improving forecasting performance in subspace-based methods using genetics Colonial Theory, *Digital Signal Processing*, **51**, 101–109.

[27] Hassani, H., Mahmoudvand, R., Omer, H. N. and Silva, E. S. (2014). A Preliminary Investigation into the Effect of Outlier(s) on Singular Spectrum Analysis, *Fluctuation and Noise Letters*, **13**(14), 1450029.

[28] Ghodsi, M., Hassani, H., Rahmania, D., and Silva, E. S. (2018). Vector and recurrent singular spectrum analysis: which is better at forecasting?, *Journal of Applied Statistics*, **45**(10), 1872–1899.

[29] Pepelyshev, A., Comparison of recurrent and vector forecasting. UK-China workshop on Singular Spectrum Analysis and its Applications, Cardiff, September 20, 2010. Available at http://ssa.cf.ac.uk/pepelyshev/pepelyshev-ssa-forecast.pdf.

[30] Hassani, H., Kalantari, M. and Yarmohammadi, M. (2017). An improved SSA forecasting result based on a filtered recurrent forecasting algorithm, *C.R. Acad. Sci. Paris, Ser.I*, **355**, 1026–1036.

[31] Mahmoudvand, R. and Rodrigues, P. C. (2017). A new parsimonious recurrent forecasting model in singular spectrum analysis, *Journal of Forecasting*, 1–10. https://doi.org/10.1002/for.2484.

[32] Hassani, H. and Kalantari, M. (2018). A novel signal extraction approach for filtering and forecasting noisy exponential series, *C.R. Acad. Sci. Paris, Ser.I*, **356**, 563–570.

[33] Hyndman, R. (2018). *Australian annual exports of goods and services in current prices: 1950-1990*, Available from Time Series Data Library (TSDL) Web site: https://datamarket.com/data/list/?q=cat:ehl provider:tsdl.

[34] Hyndman, R. (2018). *Quarterly gross fixed capital expenditure of Australia: 1959-1995*, Available from Time Series Data Library (TSDL) Web site: https://datamarket.com/data/list/?q=cat:ehl provider:tsdl.

[35] Cryer, J. D. and Chan, K. S. (2008). *Time Series Analysis: With Applications in R*, Second ed., Springer.

[36] Hyndman, R., Athanasopoulos, G., Bergmeir, C., Caceres, G., Chhay, L., O'Hara-Wild, M., Petropoulos, F., Razbash, S., Wang, E. and Yasmeen, F. (2018). forecast: Forecasting functions for time series and linear models, R package version 8.4, http://pkg.robjhyndman.com/forecast.

[37] Hydman, R. and Khandakar, Y. (2008). Automatic Time Series Forecasting: The forecast Package for R, *Journal of Statistical Software*, **27** (3), 1–22.

[38] Hassani, H. (2007). Singular Spectrum Analysis: Methodology and Comparison, *Journal of Data Science*, **5**(2), 239–257.

[39] Silva, E. S., Hassani, H., and Heravi, S. (2018). Modeling European industrial production with multivariate singular spectrum analysis: A cross-industry analysis, *Journal of Forecasting*, **37**(3), 371–384.

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

[41] Heravi, S., Osborn, D. R., and Birchenhall, C. R. (2004). Linear versus neural network forecasts for European industrial production series, *International Journal of Forecasting,*, **20**(3), 435–446.