Vector and Recurrent Singular Spectrum Analysis: Which is Better at Forecasting?

Mansi Ghodsi[a,*], Hossein Hassani[b], Donya Rahmani[a] and Emmanuel Sirimal Silva[c,1]

[a]The Business School, Bournemouth University, 89 Holdenhurst Road, Bournemouth. BH8 8EB, UK

[b]Research Institute of Energy Management and Planning, University of Tehran, No. 13, Ghods St., Enghelab Ave., Tehran, Iran

[c]Fashion Business School, London College of Fashion, University of the Arts London, 272 High Holborn, London. WC1V 7EY

[1]Corresponding author: e.silva@fashion.arts.ac.uk

[*]*Author names are appearing in alphabetical order.*

Vector and Recurrent Singular Spectrum Analysis: Which is Better at Forecasting?

15th November 2017

Abstract

Singular Spectrum Analysis (SSA) is an increasingly popular and widely adopted filtering and forecasting technique which is currently exploited in a variety of fields. Given its increasing application and superior performance in comparison to other methods, it is pertinent to study and distinguish between the two forecasting variations of SSA. These are referred to as Vector SSA (SSA-V) and Recurrent SSA (SSA-R). The general notion is that SSA-V is more robust and provides better forecasts than SSA-R. This is especially true when faced with time series which are non-stationary and asymmetric, or affected by unit root problems, outliers or structural breaks. However, currently there exists no empirical evidence for proving the above notions or suggesting that SSA-V is better than SSA-R. In this paper, we evaluate out-of-sample forecasting capabilities of the optimised SSA-V and SSA-R forecasting algorithms via a simulation study and an application to 100 real data sets with varying structures, to provide a statistically reliable answer to the question of which SSA algorithm is best for forecasting at both short and long run horizons based on several important criteria.

Keywords: Singular Spectrum Analysis; Vector SSA; Recurrent SSA; Forecasting.

1 Introduction

The Singular Spectrum Analysis (SSA) technique is a nonparametric time series analysis and forecasting technique which is transforming into an increasingly popular method for noise reduction and forecasting. Whilst it is not the objective of this paper to review all applications of SSA, we cite few of the recent articles as evidence of the increasing popularity of SSA (see for example, [1–14]). In brief, the SSA technique seeks to decompose a time series to identify the trend, signal, harmonic components and noise, and thereafter reconstructs a new, filtered time series which can be used for forecasting future data points [15]. In comparison to classical time series models, the SSA technique has the advantage of not been bound by the parametric assumptions of stationarity or normality [15] which are highly unlikely to hold in the real world.

The interest of this paper lies in the evaluation and comparison between the two SSA forecasting algorithms, with a view to identifying if one approach is strictly better than the other, or whether the best approach can be selected based on the structure of the time series in question. The two forecasting variations in SSA are referred to as Vector SSA (SSA-V) and Recurrent SSA (SSA-R). According to a suggestion by Golyandina et al. [16], the SSA-V algorithm is more robust than the SSA-R algorithm when faced with time series which have unit root problems. This was later confirmed in [17] where the author agreed with the conclusion in [16] which was based on a single application. Given the lack of statistically reliable experiments behind the aforementioned conclusions, one is unable to conclude with absolute confidence as to which of the two approaches are best for forecasting, or whether the best approach for a certain situation can be selected based on the structure of a given time series. Moreover, the SSA algorithms used in both [16, 17] were not optimal in terms of the selection of SSA choices, where the term choices refers to the parameters of a given SSA model [18].

In order to provide a more reliable comparison between the SSA forecasting algorithms, this paper adopts the basic SSA-V and SSA-R models with optimal choices [10, 20], along with an application into forecasting 100 real time series. These real time series include both stationary, and non-stationary data sets with varying fluctuations and seasonal components. Also considered is a simulation based on the Henon series. Given the significant increase in applications of SSA over the last decade, we believe this paper can provide enlightening insights to forecasters on the selection of the most suitable SSA forecasting approach based on the nature of the data being analysed.

The remainder of this paper is organized as follows. Section 2 describes the methodology underlying the SSA-R and SSA-V optimal forecasting algorithms whilst Section 3 is dedicated towards introducing the real data sets used in this paper. Section 4 reports the empirical results which includes the outcome from the simulation study and results following the application to real data, with the paper concluding in Section 5.

2 Methodology

In this section we present the SSA-R and SSA-V optimal forecasting algorithms. In doing so we mainly follow [10, 20].

1. Consider a real-valued nonzero time series $Y_N = (y_1, \ldots, y_N)$ of length N.

2. Divide the time series into two parts; $\frac{2}{3}^{rd}$ of observations for model training and testing, and the last $\frac{1}{3}^{rd}$ for validating the selected model.

3. Use the training data to construct the trajectory matrix $\mathbf{X} = (x_{ij})_{i,j=1}^{L,K} = [X_1, ..., X_K]$, where $X_j = (y_j, ..., y_{L+j-1})^T$ and $K = N - L + 1$. Initially, we begin with $L = 2$ ($2 \leq L \leq \frac{N}{2}$) and in the process, evaluate all possible values of L for Y_N.

4. Obtain the SVD of $\mathbf{X}$ by calculating $\mathbf{X}\mathbf{X}^T$ for which $\lambda_1, \ldots, \lambda_L$ denotes the eigenvalues in decreasing order ($\lambda_1 \geq \ldots \lambda_L \geq 0$) and by $U_1, \ldots, U_L$ the corresponding eigenvectors. The output of this stage is $\mathbf{X} = \mathbf{X}_1 + \ldots + \mathbf{X}_L$ where $\mathbf{X}_i = \sqrt{\lambda_i} U_i V_i^T$ and $V_i = \mathbf{X}^T U_i / \sqrt{\lambda_i}$.

5. Evaluate all possible combinations of r ($1 \leq r \leq L - 1$) singular values (step by step) for the selected L and split the elementary matrices $\mathbf{X}_i$ ($i = 1, \ldots, L$) into several groups and sum the matrices within each group.

6. Perform diagonal averaging to transform the matrix with the selected r singular values into a Hankel matrix which can then be converted into a time series (the steps up to this stage filters the noisy series). The output is a filtered series that can be used for forecasting.

7. Depending on the forecasting approach one wishes to use, select the SSA-R approach or SSA-V approach which are explained below in Sections 2.1 and 2.2 respectively.

8. Define a loss function $\mathcal{L}$.

9. When forecasting a series Y_N h-step ahead, the forecast error is minimised by setting $\mathcal{L}(X_{K+h} - \hat{X}_{K+h})$ where the vector $\hat{X}_{K+h}$ contains the h-step ahead forecasts obtained using the SSA-V or SSA-R forecasting algorithm.

10. Find the combination of L and r which minimises $\mathcal{L}$ and thus represents the optimal SSA choices.

11. Finally use the optimal L to decompose the series comprising of the validation set and select r singular values for reconstructing the less noisy time series. Thereafter, use this newly reconstructed series for forecasting the remaining $\frac{1}{3}^{rd}$ observations.

2.1 SSA-R

Let $v^2 = \pi_1^2 + \ldots + \pi_r^2$, where π_i is the last component of the eigenvector U_i ($i = 1, \ldots, r$). Moreover, suppose for any vector $U \in \mathbf{R}^L$ denoted by $U^\triangledown \in \mathbf{R}^{L-1}$ the vector consisting of the first $L-1$ components of the vector U. Let $y_{N+1}, \ldots, y_{N+h}$ show the h terms of the SSA recurrent forecast. Then, the h-step ahead forecasting procedure can be obtained by the following formula

$$y_i = \begin{cases} \widetilde{y}_i & \text{for } i = 1, \ldots, N \\ \sum_{j=1}^{L-1} \alpha_j y_{i-j} & \text{for } i = N+1, \ldots, N+h \end{cases} \tag{1}$$

where $\widetilde{y}_i$ ($i = 1, \ldots, N$) creates the reconstructed series (noise reduced series) and vector $A = (\alpha_{L-1}, \ldots, \alpha_1)$ is computed by:

$$A = \frac{1}{1 - v^2} \sum_{i=1}^{r} \pi_i U_i^\triangledown. \tag{2}$$

2.2 SSA-V

Consider the following matrix

$$\Pi = \mathbf{V}^\triangledown (\mathbf{V}^\triangledown)^T + (1 - v^2) A A^T \tag{3}$$

where $\mathbf{V}^\triangledown = [U_1^\triangledown, ..., U_r^\triangledown]$. Now consider the linear operator

$$\theta^{(v)} : \mathfrak{L}_r \mapsto \mathbf{R}^L \tag{4}$$

where

$$\theta^{(v)}U = \begin{pmatrix} \Pi U^{\triangledown} \\ A^T U^{\triangledown} \end{pmatrix}. \tag{5}$$

Define vector Z_i as follows:

$$Z_i = \begin{cases} \widetilde{X}_i & \text{for } i = 1, \ldots, K \\ \theta^{(v)} Z_{i-1} & \text{for } i = K+1, \ldots, K+h+L-1 \end{cases} \tag{6}$$

where, $\widetilde{X}_i$'s are the reconstructed columns of the trajectory matrix after grouping and eliminating noise components. Now, by constructing matrix $\mathbf{Z} = [Z_1, ..., Z_{K+h+L-1}]$ and performing diagonal averaging we obtain a new series $y_1, ..., y_{N+h+L-1}$, where $y_{N+1}, ..., y_{N+h}$ form the h terms of the SSA vector forecast.

Given that this paper is focussed entirely around SSA-V and SSA-R, we find it important to briefly comment on the computational complexity associated with the two SSA forecasting approaches. Also, this discussion could be useful for the cases when both approaches are equivalent. At the outset, it is noteworthy that both approaches are very similar in terms of computation as they both rely on the SSA choices of L and r for decomposition and reconstruction, and the linear recurrent formula for generating forecasts. As such, in terms of the computational complexity, there is no major distinguishable factor and both approaches will take a similar computation time to generate forecasts. However, SSA-V is known to provide a more robust analysis which is less sensitive to outliers [19], and even in its multivariate form there is evidence that SSA-V can provide better results than SSA-R [18].

3 Real Data

The real data used in this study have been obtained via the Data Market[1] and includes 100 data sets representing various fields and categories. A detailed account of the descriptives relating to the real data have been reported in Table 8 (see: Appendix). In order to provide a richer understanding on the nature of the real data, the mean, median, standard deviation (SD), coefficient of variation (CV), and skewness statistics, results from the normality (Shapiro-Wilk) and stationarity (Augmented Dickey-Fuller) tests have been reported via Table 8 in the Appendix. Below, we use Table 1 to present a concise summary on the nature of the 100 real data sets. Note that each time series used in this study has been given a code and the code is explained via Table 7 in the Appendix.

Table 1: Summary of the 100 real data.

	A	M	Q	W	D	H	+'ve Skew.	-'ve Skew.	Normal	Stationary	Non-stationary
Count	5	83	4	4	2	2	61	21	18	14	86

Note:A - Annual data, M - Monthly data, Q - Quarterly data, W - Weekly data, D - Daily data, H - Hourly data.

The first observation from Table 1 is that the study considers a variety of data with varying frequencies and distributions. Accordingly, we have considered data which represents annual, monthly, weekly, daily, and hourly frequencies with 18

[1]http://datamarket.com/

data sets which are normally distributed, and 14 data sets which are stationary. Moreover, there are 61 positively skewed data sets and 21 negatively skewed data sets. A majority of the time series used here are non-stationary and represents real life scenarios where non-stationarity is common. The nature of the selected data sets will enable an interesting comparison with regard to the impact of skewness, normality and stationarity of time series on SSA-V and SSA-R forecasting results.

It is also interesting to note that the 100 data sets evaluated in this study come from different fields. These include for example, crime, agriculture, economics, chemistry, ecology, energy, finance, health, tourism, housing market, and production. As such, we can ascertain the usefulness of SSA-V and SSA-R forecasts on a wide range of industries, which in turn improves the value of the output from this research. Figure 1 illustrates a selection of the 100 real time series used in this study. Prior to reporting the empirical results, we find it useful to describe certain characteristics of the time series shown in Figure 1 to give the reader a better understanding of the data used for real world applications.

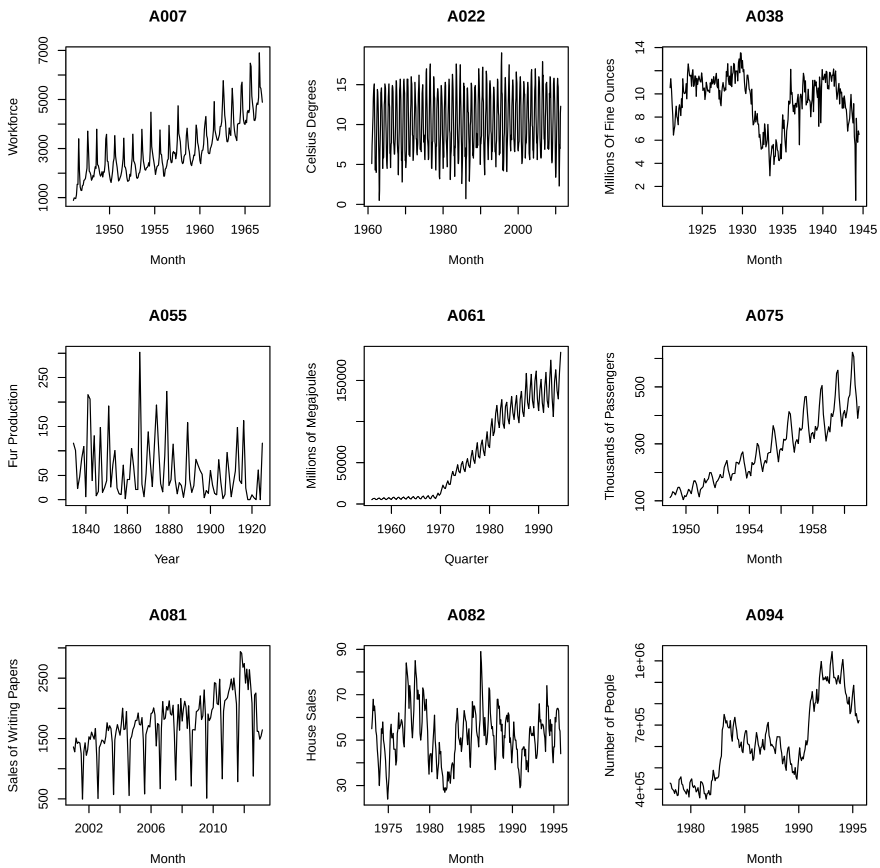

Figure 1: An example of 9 out of the 100 real time series.

A007 is an asymmetric non-stationary time series which represents the labour market in a U.S. county. It is clear that this monthly series is seasonal with a non-linear trend which appears to increase over time. On the other hand, A022 represents an asymmetric, yet stationary meteorological variable and appears to be highly seasonal right throughout with a high amplitude and possible sine wave pattern lying underneath. The time series in A038 is both asymmetric and non-stationary, and represents the production of silver. It has structural breaks of major magnitude visible through the entire series. The annual time series A055 is surprisingly stationary as per the ADF test and is also asymmetric. At first glance this time series appears to have no distinct underlying signal, however this series contains data on the production of coloured fox fur. A061 is an interesting quarterly series representing the energy sector and is non-stationary and asymmetric. This series has a non-linear trend along with an increasing seasonality over time. A075 is also asymmetric and non-stationary, and represents the airline industry. This time series is clearly seasonal along with a linear and increasing trend. A081 is representative of sales and whilst the trend suggests increasing seasonality over time, it is clear that there are major drops in the time series between each season. This series is also non-normal in distribution and non-stationary. A082 represents house sales and is a normally distributed, non-stationary time series. This particular series appears to have a slightly curved non-linear trend and a sine wave which is disrupted by noise. Finally, A094 once again represents the labour market, but in this case there are many structural breaks which makes the time series non-stationary, and this asymmetric series has seasonal periods visible with a non-linear trend.

In what follows, the empirical results are presented with a discussion on findings from both a simulation study and application to real data.

4 Empirical Results

4.1 Metrics

A key highlight of the simulation study is the consideration given to a variety of important factors in determining the true quality of a forecast from a given model. Firstly, the forecast error has been considered using both the Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) criteria. Secondly, the prediction of the correct direction of change has also been considered via a criterion referred to as Direction of Change (DC). Thirdly, consideration is also given to different forecasting horizons such that possible outcomes in both the short, medium and long term are taken into account. Below, we provide the formulae for calculating RMSE, MAE and DC prior to presenting the results from the Henon series simulation.

$$RMSE = \left(\frac{1}{n} \sum_{i=1}^{M} (Y_i - \hat{Y}_i)^2 \right)^{\frac{1}{2}} \tag{7}$$

$$MAE = \sum_{i=1}^{M} |Y_i - \hat{Y}_i| \tag{8}$$

The DC criterion is summarised below, and in doing so we mainly follow [5]. In the univariate case, for forecasts obtained using X_T, let D_{Xi} be equal to 1 if the forecast is able to correctly predict the actual direction of change and 0 otherwise. Then, $\tilde{D}_X = \sum_{i=1}^{n} D_{Xi}/n$ shows the proportion of forecasts that correctly identify the direction of change in the actual series.

4.2 Henon series simulation

We begin by presenting the results from a simulation study which considered the chaotic Henon series. In brief, it is a simple model which has the same essential properties of Lorenz system of differential equations [21]. The importance of considering this series for simulation purposes lies in the fact that the deterministic nature of these systems makes them unpredictable, and as such an experiment on predicting the chaotic time series from Henon map illustrates the performance of a given method, whilst the experimental result is also able to show the forecasting precision attainable via SSA-V or SSA-R when faced with such chaotic series. The results from this simulation are reported in Table 2.

We begin by comparing the forecasting results from SSA-R and SSA-V. The first observation is that the results from this simulation study are easily differentiable in this case based on the RMSE, RRMSE and MAE criteria. Based on the RMSE, SSA-V outperforms SSA-R at forecasting the Henon series across all four horizons at h = 1, 3, 6 and 12 steps-ahead. Across all four horizons, the SSA-V approach records an average RMSE of less than 1.00 whilst the average RMSE for SSA-R exceeds 1. If we consider the MAE criterion, SSA-V only outperforms SSA-R at $h = 3, 6$ and 12 steps-ahead, whilst SSA-R outperforms SSA-V at $h = 1$ step-ahead. However, this is only by 1 point and so it is difficult to conclude whether this is significant or not.

As such, in order to provide a better indication with regard to the performance of both approaches when forecasting the Henon series, we consider the RRMSE criterion. Based on the RRMSE, we are able to conclude that forecasts from SSA-V are 7%, 18%, 32% and 54% more accurate than the forecasts from SSA-R at $h = 1, 3, 6$ *and* 12 steps-ahead respectively. What is interesting is that as the horizon increases, SSA-V forecasts are seen performing comparatively better than SSA-R in providing the most accurate forecasts for the Henon series. In fact, the SSA-R forecasting performance is seen deteriorating heavily as the horizon increases beyond 1 step-ahead, whilst SSA-V shows comparatively more stable results. Finally, the average RRMSE result indicates that on average, across all four horizons evaluated here, forecasts from SSA-V are 28% better than forecasts from SSA-R. Accordingly, based on the loss functions, we are able to conclude that regardless of the horizon, SSA-V will provide a better forecast than SSA-R for the Henon series as proven by this simulation study.

Lastly we consider the DC criterion. The aim here is to ascertain whether the forecast is able to pick up the actual upwards or downwards trend in the real data. Across all four horizons SSA-V forecasts once again appears superior over SSA-R forecasts with a comparatively better accuracy in terms of the DC prediction. The average DC values makes it evident that, on average, when forecasting the Henon series, we can expect SSA-V to report a 74% accurate DC prediction in comparison to the SSA-V forecasts 69% average DC prediction. Accordingly, we are able to provide the following solid conclusion. When forecasting the Henon series, SSA-V is better than SSA-R in terms of the forecasting accuracy and the direction of change prediction in both the short and long run.

Table 2: Henon series forecasting results with SSA(3,1).

Horizon	SSA-V	SSA-R	RRMSE	SSA-V (MAE)	SSA-R (MAE)	SSA-V (DC)	SSA-R (DC)
1	0.87	0.92	0.93	0.66	0.65	0.73	0.73
3	0.84	1.01	0.82	0.69	0.78	0.73	0.71
6	0.83	1.21	0.68	0.73	0.97	0.75	0.69
12	0.91	2.29	0.46	0.78	1.64	0.75	0.64
Average	0.86	1.36	0.72	0.72	1.01	0.74	0.69

Note: RRMSE refers to the Ratio of the RMSE and here RRMSE = $\frac{VSSA}{RSSA}$. This means that when the RRMSE is less than 1, SSA-V outperforms SSA-R by 1-RRMSE percent and vice versa.

4.3 Application to Real Data

This section is dedicated towards reporting and analysing the out-of-sample forecasting results relating to the 100 real data sets that were introduced to the reader in Section 3. In analysing the application to real data, we rely on the RMSE, RRMSE nd DC criterions. Whilst a detailed account of the out-of-sample RMSE and RRMSE results can be found in Table 9 in the Appendix, we make use of a concise summary presented in Table 3 to draw our conclusions.

Analysis based on statistically significant outcomes.

In line with good practice, we have applied the modified Diebold-Mariano (DM) test in [22] to ascertain the statistically significant differences between SSA-V and SSA-R forecasts. However, it is pertinent to point out that if we rely on statistical significance as per the DM test, then we are unable to provide any form of differentiation between SSA-V and SSA-R as there are a very low number of statistically significant differences reported between these two approaches when applied to 100 data sets. In fact, if we were to present conclusions considering only these (very low) statistically significant outcomes we can infer that at $h = 1, 3$ *and* 12 steps-ahead, where SSA-V outperforms SSA-R based on the RMSE, the forecasts from SSA-V are likely to have a statistically significant difference in comparison to forecasts from SSA-R, whilst the outcomes are the exact opposite at $h = 6$ steps ahead. When drawing further conclusions, we do not rely on the statistically significant differences between the outcomes as the DM test hinders any further differentiation by suggesting that in majority of the cases there exists no statistically significant difference between the forecasts from SSA-V and SSA-R. We are of the view that it is factually incorrect to suggest there exists no statistically significant difference between the forecasts from these two approaches as such a conclusion does not appear to be justifiable given the empirical work previously carried out in [16, 17]. Moreover, there could be issues related to the Diebold-Mariano test statistic, and those interested are referred to [23] as the discussion of same is beyond the mandate of this paper. Yet, we do consider an alternate approach to determining and pointing out statistically reliable differentiations between the results obtained from this study and this has been explained in what follows.

Table 3: Summary of out-of-sample forecasts for 100 real data sets.

	SSA-V				SSA-R								
Score	*1*	*3*	*6*	*12*	*1*	*3*	*6*	*12*		*1*	*3*	*6*	*12*
General													
Sig.	**4**	**5**	1	**4**	3	2	**2**	3	$\mu\frac{SSA-V}{SSA-R}$	0.98	0.99	1.00	0.99
Overall	**62**	**46**	**41**	**57**	23	38	30	22	SSA-V=SSA-R	15	16	29	21
Data Type													
Normal	**14**	**10**	**7**	**15**	2	8	5	2	SSA-V=SSA-R	2	0	6	1
+'ve Skew.	**37**	**28**	22	**32**	19	24	**23**	15	SSA-V=SSA-R	5	9	16	14
-'ve Skew.	**11**	**8**	**10**	**11**	2	6	3	5	SSA-V=SSA-R	8	7	8	5
Station.	**10**	**10**	**7**	**9**	3	3	1	3	SSA-V=SSA-R	1	1	6	2
Non. Stat.	**52**	**36**	32	**47**	20	35	29	20	SSA-V=SSA-R	28	30	**39**	33
Frequency													
Annual	**4**	**5**	**3**	**4**	0	0	2	1	SSA-V=SSA-R	1	0	0	0
Monthly	**49**	34	30	**49**	27	**35**	**31**	24	SSA-V=SSA-R	7	14	22	10
Quarterly	**2**	1	1	1	0	**2**	**2**	**2**	SSA-V=SSA-R	**2**	1	1	1
Weekly	**4**	**2**	1	**2**	0	**2**	1	0	SSA-V=SSA-R	0	0	**2**	**2**
Daily	**2**	**2**	**2**	**1**	0	0	0	**1**	SSA-V=SSA-R	0	0	0	0
Hourly	**2**	**2**	**2**	**2**	0	0	0	0	SSA-V=SSA-R	0	0	0	0
Series Length													
$1<N\leq150$	**23**	**23**	**17**	**17**	8	10	12	12	SSA-V=SSA-R	2	0	4	4
$150<N\leq300$	**29**	**20**	**16**	**28**	9	14	13	10	SSA-V=SSA-R	5	9	14	5
$N>300$	10	6	6	**13**	**12**	**13**	**10**	7	SSA-V=SSA-R	2	5	8	4

Note: Except for $\mu\frac{SSA-V}{SSA-R}$, all other numbers appearing in this table represents the score. The score is defined as the umber of times SSA-V or SSA-R outperforms the alternate. Sig. represents the number of statistically significant scores. Shown in bold are the scores for the best performing model at the corresponding forecasting horizon.

The overall, general picture

Highlighted in bold are the instances where either SSA-V or SSA-R outperforms the alternate approach. It is evident from Table 3 that majority of the bold marks fall under SSA-V. As such it is clear that in general we can outline SSA-V as the better forecasting approach in comparison to SSA-R, regardless of the nature of the data. This conclusion is further supported by the 'Overall' results which shows that on majority of the instances, SSA-V outperforms SSA-R across all horizons evaluated in this study. Even though these results are supportive of the findings in [16, 17], they are not overly helpful to practitioners wishing to distinguish between SSA-R and SSA-V on a more micro level. As such, we analyse the results in more detail and present the following findings.

Inferences based on the RRMSE

The average RRMSE across 100 data sets show that in the short run ($h = 1$ step-ahead) SSA-V can provide forecasts which are on average 2% better than SSA-R, and that in the long run ($h = 12$ steps-ahead) SSA-V continues to provide forecasts which are on average 1% better than those provided by SSA-R. In the medium term ($h = 3$ and 6 steps-ahead), we find that on average SSA-V can provide a forecast which is 1% better than SSA-R at the horizon of three steps-ahead whilst there is on average no difference between the forecasts from SSA-V and SSA-R at $h = 6$ steps-ahead. Therefore, based on the average RRMSE we are able to recommend SSA-V as the better approach for forecasting in the short or long run (i.e. $h = 1$ or 12 steps-ahead), whilst for medium term forecasts we can recommend SSA-V to be the most appropriate for attaining $h = 3$ steps-ahead predictions whilst there is no difference between the two approaches at $h = 6$ steps-ahead.

Given the low statistically significant outcomes in this case, we believe it is important to consider the distribution of the RRMSE to provide further support to our claims. These distributions for each horizon are shown in Figure 2. It is clear from Figure 2 that in line with our conclusions there appears to be support for SSA-V being more likely to provide better forecasts than SSA-R at $h = 1, 3$ *and* 12 steps-ahead whilst the $h = 6$ steps-ahead distribution appears to be more less close to a normal distribution, confirming that there is on average likely to be no difference between SSA-V and SSA-R at this forecasting horizon.

However, it is not possible to be certain of the outcomes in this case by looking at the RRMSE value or the histograms relating to the distribution of RRMSE. Accordingly, we go a step further and study the cumulative distribution functions (c.d.f) of the RRMSE at each horizon. The resulting c.d.f's are presented in Figure 3. This approach enables us to quantify the findings further and give a more accurate picture in terms of a percentage. We analyse the c.d.f's to find out what percentage of RRMSE's lie below or above 1.00. If majority of the RRMSE's lie below 1, we can then conclude that on average, at a given horizon, SSA-V is better at providing out-of-sample forecasts in comparison to SSA-R and vice versa. Firstly, at $h = 1$ step-ahead, on average 60% of the RRMSE's are below 1.00, with approximately 20% equalling 1.00 and the remaining 20% exceeding 1.00. This makes it is clear that at $h = 1$ step-ahead, on average SSA-V will provide better forecasts than SSA-R. At $h = 3$ steps-ahead, on average, approximately 45% of the RRMSE's lie below 1.00, 15% equivalent to 1.00 and around 40% exceeds 1.00, thus providing weak evidence suggesting that on average, SSA-V can provide better forecasts than SSA-R in this case. At $h = 6$ steps-ahead approximately 40% of the RRMSE's are found to be below 1.00 whilst approximately 30% are seen being equivalent to and exceeding 1.00. Once again, there is weak evidence to conclude that SSA-V is on average better than SSA-R at this horizon. However, given the weak evidence in support of one particular forecasting approach at $h = 3$ and 6 steps-ahead, it is more appropriate to conclude that in the medium term, on average there is no significant difference between SSA-V and SSA-R forecasts. Finally, we consider the long run (i.e $h = 12$ steps-ahead). Once again, the results for the long run mirror the results at $h = 1$ step-ahead in terms of the approximate percentage values, suggesting that on average, in the long run SSA-V is likely to provide better forecasts than SSA-R.

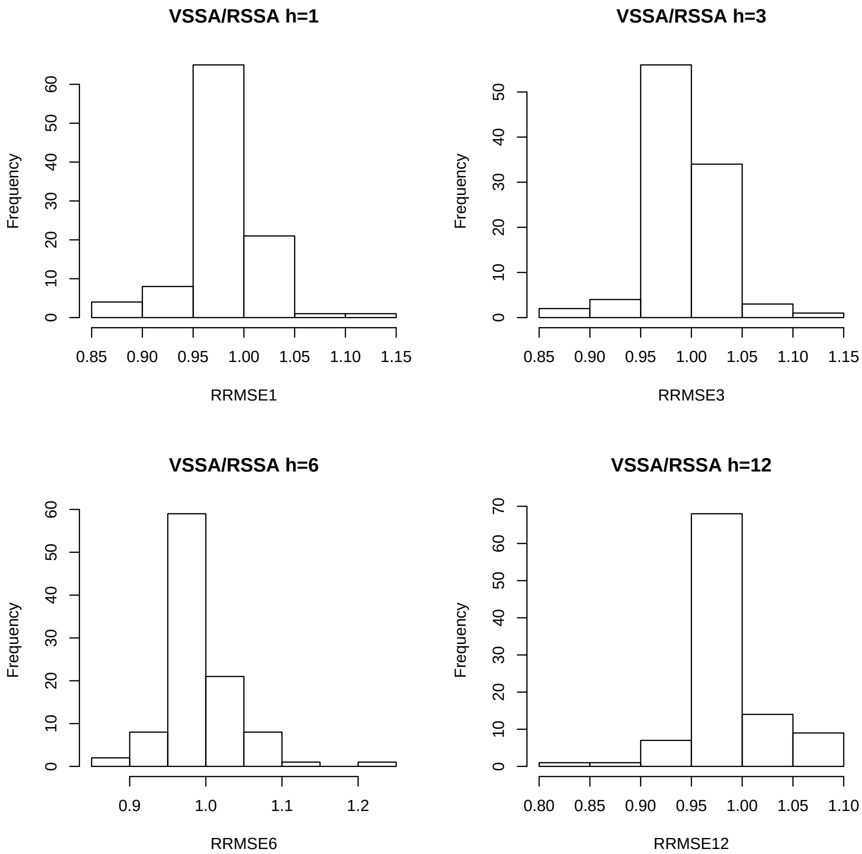

Figure 2: Distribution of RRMSE ($\frac{VSSA}{RSSA}$) for 100 data sets.

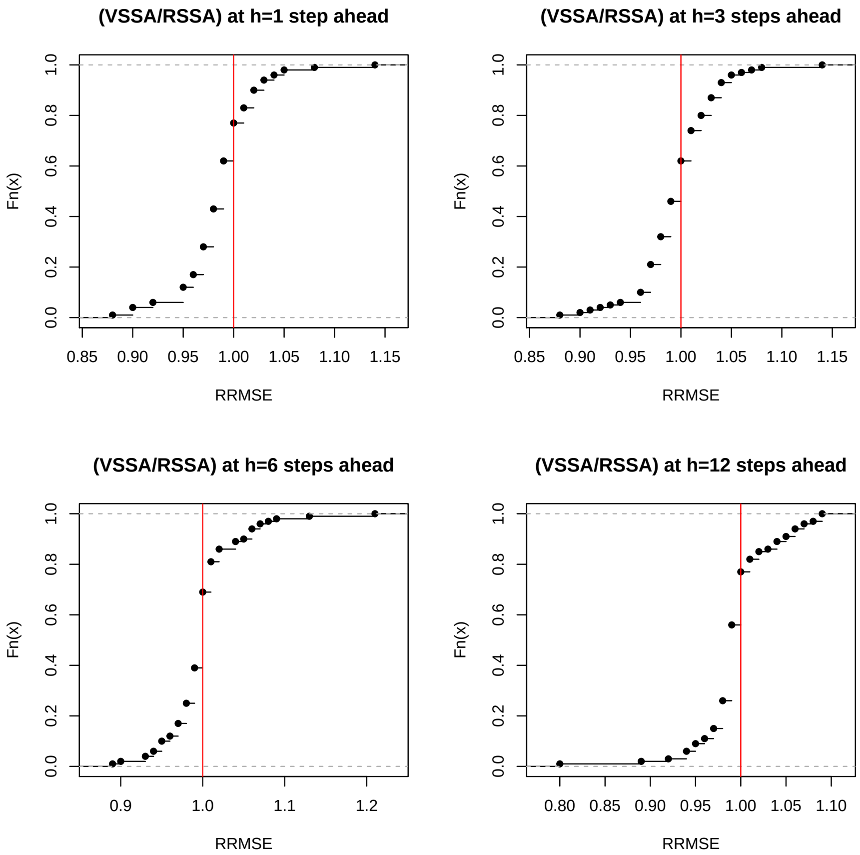

Figure 3: Cumulative distribution functions for RRMSE ($\frac{VSSA}{RSSA}$).

The distribution of data and its impact on SSA-V and SSA-R

Discussed herewith is the impact of the distribution of data (i.e. normal or skewed) on the out-of-sample forecasts attainable via both SSA-V and SSA-R. When the data is normally distributed, it is clear that SSA-V is most likely to provide a better forecast than SSA-R. This is evident as out of the 72 possible outcomes, SSA-V forecasts turn out to be better than SSA-R forecasts 64% of the time. As such, where data is normally distributed, SSA-V can be recommended to be the most appropriate approach for obtaining out-of-sample forecasts. Where the data is positively skewed, at horizons of 1, 3, and 12 steps-ahead, SSA-V is more likely to provide better forecasts than SSA-R, whilst at $h = 6$ steps-ahead SSA-R is seen outperforming SSA-V by 1 instance alone. Accordingly, it is safe to suggest that at $h = 6$ steps-ahead, there is no real difference between using SSA-V or SSA-R forecasts. As such when the data is positively skewed there is sufficient evidence to suggest that using SSA-V is likely to be more appropriate. In terms of the situation where data are negatively skewed the results are very clear that SSA-V forecasts are most likely to outperform SSA-R forecasts.

Stationarity and non-stationarity of data and its impact on SSA-V and SSA-R

In this section we consider the impact of stationary and non-stationary time series on SSA-V and SSA-R forecasts. This is important as in the real world we are faced with many non-stationary, and in some cases stationary time series. The ability to provide some insight in relation to the best approach to adopt under such circumstances would be helpful for many practitioners around the globe. As per the results in Table 3, it is clear that when the data is stationary SSA-V is once again most likely to provide better out-of-sample forecasts in comparison to SSA-R. In fact, where the data is stationary SSA-V forecasts outperformed SSA-R forecasts 64% of the time. When faced with non-stationary time series, in the short run and long run (i.e. $h = 1$ and 12 steps-ahead) there is a clear indication that SSA-V is most likely to provide better forecasts than SSA-R. However, when obtaining medium term forecasts with non-stationary data it appears that there is likely to be no difference between SSA-R and SSA-V.

Frequency of data and its impact on SSA-V and SSA-R

Based on the comments by an anonymous referee we included a summary of the results based on frequency of the time series. The results are summarised for the reader via Table 3. At the outset, it should be noted that given the disproportionate spread of frequencies in relation to those with monthly frequencies, we do not find it useful to comment on the other frequencies. However, with 83 time series representing the monthly frequency, we are able to draw some useful conclusions for the reader. Firstly, we find that when forecasting monthly data in the very short run ($h = 1$ step-ahead) and very long run ($h = 12$ steps-ahead), VSSA is more likely to provide a lower forecasting error than RSSA 59% of the time. Secondly, when forecasting monthly data in the medium term ($h = 3, 6$ steps-ahead), we do not find sufficient evidence to note that one approach is strictly better than the other.

Series length and its impact on SSA-V and SSA-R

The same anonymous referee suggested that we evaluate the impact of series length on SSA-V and SSA-R forecasts. Table 3 presents a summary of this analysis. As visible, it is clear that when the series length lies between 1-300, SSA-V is more likely to provide better forecasts than SSA-R across all horizons. In fact, the results show that SSA-V outperformed SSA-R 68% of the time at $h = 1$ step-ahead, 57% of the time at $h = 3$ steps-ahead, 43% of the time at $h = 6$ steps-ahead, and 59% of the time at $h = 12$ steps-ahead. However, interestingly, when the series length is beyond 300, then we notice that SSA-R forecasts outperform SSA-V at all horizons except $h = 12$ steps-ahead. As such, if the series length was the only criteria in question, then we can suggest that users rely on SSA-V for forecasting across all horizons when the series length falls between 1-300, and for long term forecasting at $h = 12$ steps-ahead when the series length is greater than 300. Where the series length is greater than 300, the most appropriate SSA forecasting approach for short and medium term forecasts would be SSA-R.

Analysis based on the DC metric

In this section we seek to identify as to which forecasting approach provides the best DC prediction under various scenarios. The detailed results are reported in Table 4 along with a concise summary at the bottom of the same table. Such an analysis is important as the results could provide further support to the conclusions made earlier and also provide practitioners with an idea in relation to the possible DC predictions one could expect from both SSA-V and SSA-R under varying conditions.

In general, based on Table 4 it is clear that both SSA-V and SSA-R are on average able to provide satisfactory DC predictions which exceeds beyond 50% across all horizons. Whilst there appears to be no major differences between the two approaches based on both the mean and median as measures of central tendency, the SSA-V approach has a slight advantage over SSA-R across all horizons. The minimum and maximum values, standard deviation (SD) and coefficient of variation (CV) are also reported. Based on the CV we can conclude that across all horizons there is likely to be less variation in the SSA-V DC results in comparison to SSA-R. This suggests that overall SSA-V produces comparatively more stable DC predictions around the reported mean across all horizons. Accordingly, based on the DC criterion our results indicate that on average both methods are able to provide good predictions of the actual direction of change and should one be interested in the method that is most likely to be best, then SSA-V would be the approach to select.

In summary, based on the analysis following applications to 100 real data sets, we can determine the superiority of SSA-V forecasts over SSA-R forecasts in majority of the instances (with the exception of where both approaches result in identical outcomes, and when forecasting series with lengths greater than 300 in the short and medium term). It is useful to briefly comment on the theoretical justifications for this superior performance of SSA-V over SSA-R forecasts. Even though both SSA forecasting approaches are based on the linear recurrent formula in Equations (1) and (6), when forecasting with SSA-V we rely on the entire vector for generating a forecast, whilst with SSA-R the forecast is based on a coefficient as opposed to a vector. The reliance of SSA-V on the vectorial form of the matrix, as indicated in Equation (5), means that this approach can capture more dynamical information about the whole structure of the underlying matrix in relation to SSA-R. As such, it is likely that the inclusion of more information aids SSA-V in developing comparatively more accurate forecasts than SSA-R.

4.3.1 SSA Choices

In order to enable replication of the results obtained in this study, and to provide an indication on the nature of SSA choices and how these differ between SSA-V and SSA-R, we report all SSA choices for all horizons in Table 5. Where r=1 has been selected as the optimal number of eigenvalues, this indicates that the SSA approach is relying on the trend alone to forecast the respective time series. In general, we can see that there is a significant difference between the SSA-V and SSA-R choices. However, interestingly, in certain cases we are able to notice that SSA-V and SSA-R relies on the same number of eigenvalues to forecast the same series across different horizons. The window length varies but r remains constant. See for example, A004 SSA-R and A005 SSA-V.

Given that these SSA choices are the crucial determinants underlying the performance of the SSA forecasting algorithms, we find it pertinent to briefly comment on the differences between the historical approach and the relatively new, automated approach which has been considered in this paper. There are several historical approaches for determining L and r for a given time series. In [24] the authors suggest that selecting L as equal to a quarter of the length of a given series is common practice. However, previously in [16] it was noted that L should not exceed half of a given time series. One of the most brief and easy to understand explanations of the historical approach can be found in [25]. It begins with an analysis of the periodogram to find out any strong signals (e.g. seasonal fluctuations) in the data set. Thereafter, one selects L proportionate to the seasonal fluctuations after which an analysis of the scree plot or paired eigenvectors enables to differentiate between signal and noise. At this stage, one would select the appropriate number of eigenvalues r for reconstruction and consider the remainder as noise. Whilst this task would be simple in the case of a small time series, it becomes increasingly complicated and difficult when one has to analyse a huge number of paired eigenvectors for larger time series. Moreover, in the absence of seasonal fluctuations, the selection of L and r would be even more difficult, and in such cases the starting point is to select L such that it is less than half of the series length. The importance of the accurate selection of r is noted in [18] where it is stated that choosing r greater than the actual requirement will result in the incorporation of noise in the reconstructed signal. Two other approaches for the selection of SSA choices were presented in [26–28] where the authors consider the selection of L based on the concept of separability between signal and noise. In addition, when forecasting during a recession or immediately following the impact of a major structural break, in [29,32] it was shown that a small trajectory matrix approach whereby L is considered to be equal to 3 can provide sound forecasts. More recently, a Colonial Theory based approach for selecting SSA choices was introduced in [30].

In contrast, the automated approach which is documented in [10] and used in this paper enables one to overcome problems associated with the selection of SSA choices. In brief, the automated approach considers the training set of a given time series and evaluates the forecasting performance in relation to a loss function by considering every possible L and r. It then picks the L and r which minimises a given loss function and considers these to be the optimal SSA choices for forecasting the out-of-sample data. Whilst this approach saves time and effort, it should be remembered that the selection of L and r is optimized to obtain the best possible out-of-sample forecast such that one is able to prove the validity of the selected SSA choices from a statistical perspective. It has the drawback that, for example, if we were interested in capturing and analysing only the seasonal components or the trend of the series we would not be able to get the best possible decomposition and reconstruction for such purposes using this automated approach. If our objective is such, then it is more reliable to return to the historical approach which will enable one to analyze each paired eigenvector and select those representing the seasonal components which are of interest to us. However, in this paper we have focused on the behaviour of SSA-V and SSA-R algorithms when forecasting with the automated approach for several reasons. Firstly, this paper is focused on determining whether SSA-V or SSA-R is better for forecasting with SSA as opposed to signal extraction. Secondly, automated forecasting algorithms are rapidly gaining importance, especially since the introduction of the 'forecast' package in R [31]. Thirdly, automated SSA has been increasingly applied in the recent past for forecasting applications [10, 20].

The SSA choices reported in Table 5 are sensitive to the size of the training set. Selecting a larger or smaller training set will provide a different combination of L and r. Whilst in this paper we use the widely accepted two-thirds and one-third split between training and test sets, more research is needed in order to determine the optimal number of observations to include in a training set for a given time series in order to optimize the SSA choices selected via the automated approach. Another interesting point is that always when $L = 2$ and $r = 1$, provided that both SSA-V and SSA-R have chosen these as the optimal SSA choices for a particular horizon, the out-of-sample forecast from these two approaches should produce similar results.

4.3.2 Strengths and Weaknesses of SSA

Given that SSA is in the process of gaining popularity amongst time series analysts, we find it pertinent to discuss the strengths and weakness of SSA. In terms of the merits, firstly, being nonparametric means SSA can provide a more accurate representation of the real world scenario where parametric assumptions are unlikely to hold. As such, when using SSA there is no need for data transformations which leads to a loss of information [7]. Secondly, the noise reduction capabilities of SSA are not present in classical time series analysis and forecasting methods, and filtering enables SSA to provide a better fit to the data and obtain more accurate forecasts [20]. Also, the moving average component of ARIMA is known to be better at forecasting less volatile data whilst Single Exponential Smoothing cannot be used in the presence of seasonality [20], but SSA is not faced with any such restrictions. Thirdly, SSA enables one to obtain a richer understanding of the dynamics underlying time series by analysing the trend and seasonal fluctuations in isolation. Moreover, SSA can forecast a particular signal which is of interest, such as extracting and forecasting the trend alone, or 12 or 3-month seasonal fluctuations depending on the requirements. Fourthly, SSA can forecast with a minimum of 3 observations [32] whilst other time series analysis methods require larger historical data sets.

However, SSA is not without its limitations. It is well known that parametric models are preferred for certain scenarios because unlike with SSA, the parameters (e.g. regression parameters) allow interpretations on the exact effect of a given independent variable on the dependent variable [20]. Also, there exists a range of historical literature based on parametric models which enables users to easily compare and contrast between the findings. In addition, SSA is highly sensitive to the selection of L and r which leads to the argument that the decomposition process could result in a loss of some deterministic structures. It is noteworthy that the Colonial Theory based approach for selecting SSA choices helps overcome this issue to a certain extent by not relying on the historical binary approach to decomposition and reconstruction [30].

Here, it is worthwhile to draw the reader's attention to recent studies which have compared the application of SSA with other parametric and nonparametric time series analysis and forecasting techniques. In [3, 5] there is evidence of SSA outperforming Holt-Winters and ARIMA at forecasting industrial production. An application of SSA, ARIMA and Holt-Winters (HW) to eight UK economic time series before, during and after the recession, showed that SSA is least sensitive to the impact of the recession in relation to ARIMA and Holt-Winters as SSA produced comparatively superior forecasting results [29]. In [32] the authors evaluated the impact of the 2008 recession on forecasting US trade with SSA in relation to the optimal ARIMA and Exponential Smoothing (ETS) models, and Neural Networks, and found SSA to be superior. More recently, an application which compared ARIMA, ETS, Neural Networks (NN), TBATS, ARFIMA and SSA at forecasting European tourist arrivals resulted in SSA outperforming the other models on most instances [33].

Table 4: Out-of-sample DC prediction results for the 100 time series.

	SSA-V				SSA-R					SSA-V				SSA-R			
Code	*1*	*3*	*6*	*12*	*1*	*3*	*6*	*12*	Code	*1*	*3*	*6*	*12*	*1*	*3*	*6*	*12*
A001	0.70	0.66	0.57	0.60	0.68	0.71	0.67	0.62	**A002**	0.75	0.68	0.61	0.67	0.76	0.70	0.60	0.71
A003	0.90	0.57	0.58	0.53	0.93	0.57	0.58	0.50	**A004**	0.69	0.75	0.60	0.64	0.77	0.78	0.63	0.63
A005	0.71	0.52	0.81	0.75	0.68	0.52	0.62	0.80	**A006**	0.67	0.51	0.50	0.47	0.45	0.51	0.50	0.47
A007	0.80	0.74	0.62	0.53	0.82	0.73	0.61	0.51	**A008**	0.58	0.39	0.44	0.40	0.57	0.39	0.44	0.40
A009	0.47	0.41	0.35	0.38	0.47	0.41	0.35	0.38	**A010**	0.55	0.59	0.60	0.52	0.62	0.59	0.58	0.55
A011	0.79	0.71	0.71	0.59	0.67	0.66	0.73	0.61	**A012**	0.62	0.52	0.51	0.55	0.61	0.52	0.51	0.51
A013	0.71	0.77	0.63	0.58	0.73	0.68	0.61	0.67	**A014**	0.78	0.71	0.58	0.58	0.80	0.71	0.59	0.51
A015	0.78	0.63	0.63	0.53	0.85	0.65	0.63	0.53	**A016**	0.49	0.49	0.48	0.48	0.49	0.49	0.49	0.48
A017	0.53	0.58	0.62	0.60	0.60	0.62	0.58	0.43	**A018**	0.51	0.48	0.58	0.49	0.55	0.58	0.58	0.56
A019	0.89	0.82	0.61	0.58	0.90	0.81	0.61	0.62	**A020**	0.87	0.81	0.59	0.55	0.88	0.81	0.59	0.55
A021	0.86	0.82	0.56	0.61	0.87	0.82	0.56	0.61	**A022**	0.90	0.83	0.61	0.57	0.88	0.83	0.60	0.58
A023	0.95	0.85	0.60	0.62	0.95	0.84	0.60	0.69	**A024**	0.55	0.58	0.58	0.59	0.55	0.57	0.58	0.58
A025	0.56	0.58	0.58	0.63	0.55	0.55	0.61	0.60	**A026**	0.66	0.57	0.68	0.60	0.67	0.56	0.61	0.70
A027	0.58	0.64	0.46	0.54	0.63	0.40	0.46	0.61	**A028**	0.87	0.86	0.52	0.52	0.87	0.86	0.52	0.52
A029	0.78	0.54	0.65	0.50	0.80	0.58	0.63	0.49	**A030**	0.57	0.57	0.56	0.66	0.50	0.57	0.59	0.66
A031	0.90	0.71	0.66	0.72	0.93	0.71	0.66	0.69	**A032**	0.86	0.63	0.67	0.63	0.84	0.61	0.70	0.48
A033	0.68	0.64	0.61	0.59	0.66	0.69	0.67	0.59	**A034**	0.63	0.72	0.58	0.70	0.63	0.67	0.55	0.67
A035	0.75	0.74	0.70	0.67	0.75	0.73	0.72	0.70	**A036**	0.87	0.70	0.66	0.55	0.89	0.75	0.66	0.43
A037	0.71	0.65	0.61	0.62	0.81	0.67	0.65	0.44	**A038**	0.70	0.68	0.66	0.69	0.64	0.41	0.67	0.63
A039	0.78	0.72	0.70	0.66	0.79	0.48	0.56	0.63	**A040**	0.72	0.44	0.42	0.44	0.75	0.32	0.48	0.48
A041	0.70	0.78	0.77	0.76	0.70	0.75	0.74	0.76	**A042**	0.69	0.59	0.59	0.50	0.64	0.62	0.53	0.46
A043	0.54	0.54	0.53	0.51	0.56	0.54	0.53	0.51	**A044**	0.79	0.59	0.59	0.50	0.76	0.66	0.57	0.50
A045	0.90	0.82	0.82	0.58	0.58	0.63	0.15	0.55	**A046**	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
A047	0.66	0.68	0.68	0.68	0.64	0.65	0.66	0.65	**A048**	0.48	0.48	0.47	0.57	0.48	0.43	0.49	0.54
A049	0.88	0.67	0.71	0.53	0.89	0.65	0.65	0.51	**A050**	1.00	0.96	0.96	1.00	0.97	0.96	0.96	1.00
A051	0.55	0.68	0.68	0.59	0.55	0.61	0.64	0.55	**A052**	0.89	0.92	0.74	0.66	0.95	0.95	0.63	0.66
A053	0.77	0.57	0.76	0.68	0.77	0.50	0.76	0.53	**A054**	0.87	0.71	0.92	0.84	0.87	0.75	0.88	0.74
A055	0.70	0.64	0.72	0.74	0.67	0.61	0.76	0.74	**A056**	0.37	0.41	0.48	0.47	0.37	0.41	0.49	0.47
A057	0.97	0.61	0.47	0.54	0.91	0.61	0.37	0.54	**A058**	0.83	0.85	0.60	0.58	0.89	0.85	0.60	0.42
A059	0.76	0.74	0.78	0.76	0.78	0.79	0.80	0.79	**A060**	0.84	0.77	0.77	0.60	0.79	0.78	0.67	0.54
A061	0.50	0.50	0.50	0.51	0.50	0.50	0.53	0.51	**A062**	0.73	0.67	0.58	0.74	0.71	0.65	0.53	0.53
A063	0.91	0.53	0.67	0.64	0.91	0.56	0.67	0.53	**A064**	0.78	0.72	0.72	0.76	0.80	0.70	0.78	0.74
A065	0.98	0.74	0.65	0.47	0.95	0.76	0.65	0.47	**A066**	0.94	0.78	0.47	0.54	0.96	0.78	0.47	0.46
A067	0.71	0.72	0.69	0.65	0.65	0.69	0.71	0.60	**A068**	0.55	0.52	0.50	0.45	0.55	0.55	0.50	0.45
A069	0.70	0.73	0.56	0.53	0.70	0.74	0.56	0.51	**A070**	0.88	0.79	0.60	0.55	0.90	0.84	0.54	0.55
A071	0.91	0.67	0.51	0.58	0.91	0.69	0.51	0.56	**A072**	0.87	0.70	0.55	0.50	0.91	0.70	0.58	0.53

A073	0.84	0.56	0.42	0.48	0.71	0.58	0.52	0.37	**A074**	0.91	0.47	0.52	0.52	0.78	0.57	0.52	0.67
A075	0.96	0.70	0.49	0.59	0.90	0.67	0.44	0.59	**A076**	0.82	0.69	0.57	0.71	0.79	0.62	0.57	0.71
A077	0.91	0.70	0.67	0.67	0.80	0.88	0.67	0.67	**A078**	0.74	0.70	0.77	0.67	0.69	0.79	0.77	0.67
A079	0.63	0.52	0.70	0.63	0.69	0.67	0.70	0.63	**A080**	0.84	0.69	0.62	0.42	0.80	0.67	0.67	0.45
A081	0.73	0.68	0.64	0.63	0.73	0.51	0.59	0.53	**A082**	0.76	0.66	0.60	0.58	0.75	0.69	0.62	0.58
A083	0.86	0.75	0.57	0.50	0.88	0.77	0.57	0.54	**A084**	0.91	0.66	0.51	0.57	0.84	0.69	0.51	0.58
A085	0.75	0.74	0.69	0.60	0.75	0.68	0.67	0.67	**A086**	0.90	0.67	0.54	0.79	0.92	0.63	0.57	0.79
A087	0.86	0.65	0.49	0.59	0.86	0.60	0.52	0.60	**A088**	0.79	0.65	0.82	0.49	0.83	0.63	0.84	0.52
A089	0.54	0.43	0.47	0.49	0.56	0.43	0.47	0.56	**A090**	0.74	0.65	0.70	0.63	0.82	0.61	0.70	0.56
A091	0.86	0.60	0.66	0.63	0.89	0.66	0.66	0.63	**A092**	0.81	0.76	0.66	0.56	0.78	0.74	0.61	0.53
A093	0.84	0.74	0.61	0.45	0.83	0.77	0.56	0.41	**A094**	0.80	0.63	0.51	0.49	0.77	0.62	0.55	0.53
A095	0.75	0.60	0.50	0.51	0.75	0.58	0.50	0.52	**A096**	0.74	0.69	0.69	0.40	0.79	0.69	0.71	0.42
A097	0.89	0.71	0.62	0.48	0.88	0.70	0.63	0.52	**A098**	0.82	0.78	0.54	0.58	0.82	0.77	0.53	0.57
A099	0.58	0.60	0.59	0.58	0.56	0.56	0.59	0.58	**A100**	0.80	0.65	0.66	0.54	0.80	0.65	0.68	0.44
				SSA-V					SSA-R								
					1	3	6	12		1	3	6	12				
				Average	**0.76**	**0.66**	**0.62**	**0.59**		0.75	0.65	0.60	0.58				
				Median	**0.78**	**0.67**	**0.61**	**0.58**		0.78	0.66	0.60	0.55				
				Min	0.37	0.39	0.35	0.38		0.37	0.32	0.15	0.37				
				Max	1.00	1.00	1.00	1.00		1.00	1.00	1.00	1.00				
				SD	0.14	0.12	0.11	0.11		0.75	0.13	0.12	0.11				
				CV	**18.36**	**17.97**	**18.32**	**18.41**		18.76	19.85	19.36	19.77				

Table 5: SSA choices for the 100 time series.

	SSA-V				SSA-R					SSA-V				SSA-R			
Code	*1*	*3*	*6*	*12*	*1*	*3*	*6*	*12*	Code	*1*	*3*	*6*	*12*	*1*	*3*	*6*	*12*
A001	(5,2)	(44,43)	(26,6)	(50,42)	(5,2)	(45,2)	(50,18)	(50,15)	**A002**	(36,5)	(34,5)	(35,27)	(35,18)	(22,4)	(22,5)	(36,10)	(13,6)
A003	(50,18)	(47,13)	(45,13)	(50,10)	(46,18)	(35,13)	(50,10)	(16,14)	**A004**	(25,29)	(46,9)	(45,9)	(44,10)	(14,6)	(47,9)	(47,9)	(49,9)
A005	(15,9)	(18,9)	(11,9)	(9,1)	(16,9)	(16,9)	(19,12)	(6,1)	**A006**	(19,5)	(2,1)	(2,1)	(2,1)	(15,8)	(2,1)	(2,1)	(2,1)
A007	(15,10)	(24,15)	(31,24)	(18,4)	(23,14)	(23,14)	(25,9)	(17,14)	**A008**	(6,3)	(2,1)	(2,1)	(2,1)	(6,3)	(2,1)	(2,1)	(2,1)
A009	(2,1)	(2,1)	(2,1)	(2,1)	(2,1)	(2,1)	(2,1)	(2,1)	**A010**	(15,6)	(2,1)	(9,1)	(5,3)	(15,6)	(2,1)	(9,1)	(6,3)
A011	(27,24)	(49,10)	(39,34)	(38,7)	(14,10)	(40,10)	(42,15)	(50,2)	**A012**	(21,11)	(2,1)	(37,1)	(33,1)	(3,2)	(2,1)	(31,1)	(28,1)
A013	(16,13)	(16,13)	(37,1)	(34,1)	(11,7)	(12,5)	(34,1)	(32,1)	**A014**	(14,7)	(13,8)	(20,5)	(14,11)	(14,7)	(16,6)	(18,5)	(21,14)
A015	(26,16)	(27,17)	(2,1)	(2,1)	(46,22)	(18,11)	(2,1)	(2,1)	**A016**	(31,17)	(32,16)	(2,1)	(7,4)	(32,16)	(32,16)	(2,1)	(2,1)
A017	(6,2)	(2,1)	(2,1)	(60,1)	(2,1)	(10,4)	(2,1)	(57,1)	**A018**	(17,6)	(2,1)	(2,1)	(14,9)	(5,2)	(5,2)	(2,1)	(2,1)
A019	(49,5)	(59,5)	(56,5)	(50,5)	(49,5)	(58,5)	(58,5)	(56,6)	**A020**	(42,5)	(60,5)	(57,5)	(54,5)	(50,5)	(60,6)	(55,4)	(60,7)
A021	(36,7)	(60,5)	(60,5)	(60,5)	(32,7)	(60,5)	(60,5)	(60,5)	**A022**	(29,7)	(60,3)	(57,3)	(52,5)	(50,5)	(51,5)	(51,5)	(52,5)
A023	(45,11)	(56,13)	(56,13)	(37,6)	(46,9)	(50,10)	(54,11)	(49,11)	**A024**	(2,1)	(4,1)	(16,6)	(41,39)	(2,1)	(3,1)	(36,5)	(49,3)
A025	(41,15)	(3,1)	(16,11)	(9,3)	(40,15)	(10,3)	(8,1)	(4,1)	**A026**	(15,5)	(20,8)	(2,1)	(40,1)	(2,1)	(19,8)	(19,8)	(35,1)
A027	(13,11)	(6,4)	(2,1)	(25,11)	(3,2)	(2,1)	(2,1)	(43,23)	**A028**	(24,11)	(22,11)	(13,11)	(24,11)	(22,10)	(14,11)	(19,10)	(24,11)
A029	(41,28)	(37,21)	(14,5)	(37,12)	(43,23)	(31,18)	(31,18)	(7,4)	**A030**	(4,1)	(4,1)	(28,1)	(2,1)	(3,1)	(4,1)	(7,1)	(2,1)
A031	(29,14)	(28,17)	(26,23)	(15,3)	(30,13)	(27,17)	(26,9)	(15,3)	**A032**	(16,14)	(18,14)	(23,17)	(20,11)	(17,13)	(19,13)	(23,15)	(23,17)
A033	(18,8)	(11,3)	(27,1)	(23,1)	(11,5)	(25,4)	(22,4)	(23,1)	**A034**	(10,6)	(8,6)	(7,1)	(25,1)	(10,6)	(7,5)	(5,1)	(25,1)
A035	(5,2)	(5,2)	(26,6)	(2,1)	(5,2)	(5,2)	(33,9)	(26,6)	**A036**	(36,26)	(45,19)	(39,32)	(39,32)	(40,22)	(42,22)	(47,17)	(47,21)
A037	(24,3)	(23,3)	(22,3)	(24,8)	(41,13)	(22,3)	(21,3)	(25,7)	**A038**	(26,2)	(24,2)	(32,2)	(4,1)	(4,1)	(23,2)	(29,2)	(3,1)
A039	(50,25)	(39,10)	(31,26)	(8,5)	(50,25)	(39,10)	(7,1)	(37,3)	**A040**	(18,12)	(9,4)	(25,1)	(25,1)	(23,15)	(6,3)	(25,1)	(25,1)
A041	(7,3)	(21,1)	(19,1)	(18,1)	(25,11)	(16,1)	(15,1)	(10,1)	**A042**	(29,4)	(26,4)	(12,1)	(6,1)	(30,4)	(29,4)	(29,4)	(23,3)
A043	(33,1)	(2,1)	(2,1)	(12,6)	(3,1)	(2,1)	(2,1)	(12,5)	**A044**	(21,15)	(26,8)	(21,14)	(22,14)	(24,7)	(25,5)	(21,9)	(28,6)
A045	(26,22)	(21,12)	(30,26)	(30,25)	(4,3)	(31,25)	(45,1)	(29,13)	**A046**	(12,6)	(41,33)	(22,16)	(43,23)	(12,6)	(24,43)	(9,5)	(43,23)
A047	(23,1)	(21,1)	(19,1)	(24,4)	(20,1)	(19,1)	(17,1)	(12,1)	**A048**	(3,1)	(8,3)	(8,3)	(10,1)	(3,1)	(8,3)	(8,3)	(7,1)
A049	(29,9)	(30,9)	(29,25)	(19,9)	(29,9)	(27,8)	(28,8)	(25,8)	**A050**	(17,3)	(15,3)	(13,3)	(8,3)	(6,2)	(13,3)	(13,3)	(9,5)
A051	(2,1)	(24,2)	(25,2)	(25,9)	(2,1)	(25,1)	(25,1)	(25,1)	**A052**	(25,13)	(23,13)	(25,14)	(25,12)	(25,13)	(25,13)	(23,14)	(19,11)
A053	(8,4)	(7,4)	(6,3)	(7,3)	(8,4)	(8,4)	(5,2)	(17,11)	**A054**	(20,3)	(19,8)	(16,8)	(20,8)	(18,6)	(18,7)	(18,7)	(16,9)
A055	(20,1)	(20,2)	(18,7)	(10,2)	(20,2)	(20,2)	(15,1)	(10,2)	**A056**	(2,1)	(2,1)	(2,1)	(2,1)	(2,1)	(2,1)	(2,1)	(2,1)
A057	(12,7)	(22,7)	(24,12)	(22,8)	(23,6)	((21,6)	((21,6)	(23,6)	**A058**	(11,5)	(23,7)	(23,9)	(13,9)	(11,5)	(11,5)	(23,6)	(16,7)
A059	(25,1)	(25,2)	(25,2)	(23,1)	(25,1)	(24,1)	(24,1)	(20,1)	**A060**	(38,12)	(38,13)	(37,13)	(21,11)	(22,8)	(22,8)	(37,13)	(20,9)
A061	(13,3)	(4,2)	(21,4)	(41,4)	(9,4)	(13,4)	(13,4)	(17,3)	**A062**	(12,3)	(10,3)	(21,4)	(18,2)	(11,3)	(11,3)	(11,4)	(24,11)
A063	(25,9)	(23,9)	(21,9)	(15,11)	(25,9)	(25,8)	(25,8)	(15,8)	**A064**	(14,4)	(11,8)	(8,2)	(11,8)	(12,3)	(9,4)	(8,2)	(7,2)
A065	(25,17)	(25,16)	(25,17)	(10,8)	(14,13)	(21,14)	(21,14)	(25,12)	**A066**	(24,18)	(22,14)	(22,14)	(22,14)	(14,13)	(23,13)	(21,14)	(24,9)
A067	(13,5)	(8,5)	(12,1)	(13,7)	(16,9)	(4,1)	(7,1)	(13,8)	**A068**	(14,1)	(12,1)	(10,1)	(2,1)	(2,1)	(12,1)	(10,1)	(2,1)
A069	(35,16)	(35,16)	(50,6)	(45,8)	(34,9)	(34,9)	(49,6)	(49,6)	**A070**	(25,21)	(14,8)	(12,10)	(11,9)	(16,9)	(15,6)	(11,7)	(11,6)
A071	(39,13)	(36,13)	(31,13)	(27,13)	(30,13)	(31,12)	(28,1)	(31,11)	**A072**	(29,12)	(29,12)	(19,11)	(18,11)	(29,14)	(29,12)	(23,8)	(13,4)

A073	(21,15)	(21,15)	(14,2)	(14,8)	(25,18)	(24,17)	(9,6)	(4,3)	**A074**	(27,6)	(23,5)	(26,7)	(28,5)	(19,4)	(23,4)	(20,4)	(27,4)
A075	(14,12)	(14,12)	(37,22)	(12,6)	(15,11)	(23,12)	(22,12)	(12,11)	**A076**	(18,12)	(18,12)	(19,12)	(13,12)	(19,12)	(20,12)	(17,12)	(13,12)
A077	(22,5)	(2,1)	(2,1)	(2,1)	(4,2)	(15,4)	(2,1)	(2,1)	**A078**	(22,5)	(23,5)	(23,5)	(2,1)	(4,2)	(23,5)	(23,5)	(2,1)
A079	(24,5)	(24,5)	(2,1)	(2,1)	(2,1)	(23,3)	(2,1)	(2,1)	**A080**	(14,7)	(17,9)	(16,8)	(11,4)	(14,7)	(14,7)	(18,6)	(9,3)
A081	(24,15)	(23,17)	(23,16)	(14,9)	(24,14)	(25,9)	(25,12)	(25,12)	**A082**	(22,7)	(13,8)	(22,12)	(13,7)	(34,15)	(22,7)	(25,11)	(2,1)
A083	(26,20)	(30,14)	(24,10)	(18,14)	(22,16)	(23,13)	(19,12)	(18,13)	**A084**	(35,12)	(28,13)	(33,11)	(40,10)	(40,11)	(40,11)	(40,11)	(38,11)
A085	(18,5)	(18,5)	(32,20)	(33,20)	(28,14)	(18,5)	(28,14)	(32,21)	**A086**	(26,20)	(30,14)	(18,16)	(18,16)	(25,19)	(30,14)	(29,14)	(18,16)
A087	(31,16)	(36,15)	(36,15)	(18,11)	(19,1)	(22,13)	(22,13)	(35,15)	**A088**	(25,15)	(39,24)	(28,20)	(35,21)	(40,19)	(38,19)	(28,18)	(28,18)
A089	(18,8)	(2,1)	(2,1)	(2,1)	(5,3)	(2,1)	(2,1)	(6,3)	**A090**	(38,17)	(27,8)	(38,7)	(14,6)	(24,7)	(30,8)	(33,8)	(35,9)
A091	(14,7)	(8,6)	(2,1)	(3,1)	(14,7)	(12,7)	(2,1)	(3,1)	**A092**	(15,9)	(31,9)	(20,12)	(35,11)	(16,8)	(24,8)	(24,9)	(28,11)
A093	(25,19)	(25,19)	(27,25)	(20,10)	(26,16)	(13,6)	(22,9)	(20,10)	**A094**	(38,16)	(35,14)	(40,9)	(40,9)	(37,16)	(34,14)	(39,9)	(38,9)
A095	(40,18)	(14,9)	(2,1)	(30,24)	(28,17)	(32,14)	(2,1)	(33,20)	**A096**	(37,14)	(29,13)	(31,13)	(14,12)	(22,13)	(29,13)	(33,17)	(33,20)
A097	(13,11)	(16,9)	(20,8)	(20,11)	(13,10)	(16,9)	(21,11)	(21,11)	**A098**	(37,22)	(40,23)	(40,23)	(40,20)	(39,23)	(34,17)	(40,20)	(40,21)
A099	(8,4)	(5,1)	(5,3)	(4,1)	(5,1)	(4,1)	(5,3)	(3,1)	**A100**	(39,17)	(35,13)	(35,13)	(35,10)	(39,17)	(34,16)	(33,16)	(35,13)

Note: Shown as (L, r) is the respective window length and number of eigenvalues.

5 Conclusions

This paper begins with the objective of providing a statistically reliable answer to the question, which SSA forecasting approach is best? Both a simulation study and an application to 100 real data sets have been used to determine the best approach between SSA-R and SSA-V forecasts. In addition, this paper considers an optimal SSA forecasting approach [10] to determine which SSA forecasting algorithm is best for a given situation. This study considers the effect of the distribution (i.e. normal or skewed) and stationarity of the data on SSA-V and SSA-R forecasts, in addition to relying on loss functions, the direction of change criterion, and cumulative distribution functions, to provide cogent conclusions.

The simulation study has clearly shown that when faced with chaotic time series like the Henon series for example, SSA-V has a higher forecasting precision than SSA-R based on the loss functions of RMSE and MAE, and that SSA-V also reports a better DC prediction in comparison to SSA-R. From the application to real data, we find evidence to conclude that in general one is more likely to find that SSA-V is the more suitable alternative to SSA-R with the following exceptions. Firstly, based on the RRMSE (as verified via the c.d.f.'s) we can conclude that SSA-V is on average better than SSA-R at forecasting in the short run ($h = 1$) and long run ($h = 12$). However, in the medium term ($h = 3$ and 6 steps-ahead) we find that there is likely to be no difference between the SSA-V and SSA-R forecasts. Secondly, where the data is normally distributed SSA-V forecasts are most likely to outperform SSA-R forecasts. When faced with positively skewed data, it is likely that SSA-V will continue to outperform SSA-R at $h = 1, 3$ and 12 steps-ahead, whilst at $h = 6$ steps-ahead both methods are unlikely to report a major difference in forecasts. Yet, when faced with negatively skewed data we have a clear winner in SSA-V which is most likely to provide better forecasts than SSA-R across all horizons. Thirdly, when the data is stationary we find that SSA-V is most likely to outperform SSA-R, but where the data is non-stationary this result only holds at $h = 1$ and 12 steps-ahead, whilst in the medium term there appears to be no distinguishable difference between the forecasts attainable via these two approaches. Finally, in terms of the DC criterion it is evident that both SSA-V and SSA-R are capable of providing sound direction of change predictions. However, we find evidence to support the notion that SSA-V is on average slightly better than SSA-R in terms of the reported DC predictions, and that the average SSA-V results for DC are more stable than the average SSA-R results as seen via the coefficient of variation statistic.

Table 6 summarises the findings of this study in tabular format to help the reader easily identify the conclusions. It is evident that our study has found overwhelming evidence in support of SSA-V forecasts as the better alternative in relation to SSA-R when it comes to forecasting with SSA. Where the results are inconclusive, which refers to cases when both approaches are equivalent, given that there is no computational complexity gains to be made between SSA-V and SSA-R, based on our previous discussions we can suggest the use of SSA-V to be more appropriate in general. However, under such scenarios it is advisable that users also evaluate the performance of SSA-R on their data for a complete picture. In contrast, if the series length was the only criteria, then we notice that SSA-R is a better contender than SSA-V for forecasting in the short and medium term when the series length exceeds 300.

Table 6: Suggested SSA forecasting models for different criteria and forecasting horizons following a detailed analysis.

Criterion	$h = 1$	$h = 3$	$h = 6$	$h = 12$
In general	SSA-V	SSA-V	SSA-V	SSA-V
RRMSE	SSA-V	Inconclusive	Inconclusive	SSA-V
Normally distributed data	SSA-V	SSA-V	SSA-V	SSA-V
Positively skewed data	SSA-V	SSA-V	Inconclusive	SSA-V
Negatively skewed data	SSA-V	SSA-V	SSA-V	SSA-V
Stationary data	SSA-V	SSA-V	SSA-V	SSA-V
Non-stationary data	SSA-V	Inconclusive	Inconclusive	SSA-V
Direction of change	SSA-V	SSA-V	SSA-V	SSA-V
Monthly frequencies	SSA-V	Inconclusive	Inconclusive	SSA-V
$1<N\leq300$	SSA-V	SSA-V	SSA-V	SSA-V
$N>300$	SSA-R	SSA-R	SSA-R	SSA-V

Note: 'In general' shows the forecasting approach reporting the highest score solely based on lowest RMSE, ignoring all other criteria. RRMSE looks at the average performance across all data sets taking into account the c.d.f related analysis. N is the length of the series.

In conclusion, we have successfully provided a statistically reliable answer to the question of which SSA forecasting approach is best. In brief, our results indicate that on average SSA-V forecasts are better in comparison to SSA-R as reported in [16,17]. For a single specific time series, both approaches must be evaluated. The consideration of various forecasting horizons, the distribution of data, stationarity and DC criterions, along with a simulation study and application to 100 real data sets has enabled this study to provide more insights and enlightenment in comparison to the conclusions previously derived in [16, 17]. We are of the view that our results presented in this paper would help practitioners and users of SSA to easily identify and distinguish between the two forecasting approaches when selecting this nonparametric technique for forecasting their given data sets depending on the horizon of interest.

References

[1] de 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.

[2] 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**, 183–191.

[3] Hassani, H., Heravi, H., and Zhigljavsky, A. (2009). Forecasting European industrial production with Singular Spectrum Analysis. *International Journal of Forecasting, 25*, 103–118.

[4] Gong, Y., Song, Z., He, M., Gong, W., and Ren, F. (2017). Precursory waves and eigenfrequencies identified from acoustic emission data based on Singular Spectrum Analysis and laboratory rock-burst experiments. *International Journal of Rock Mechanics and Mining Sciences*, **91**, 155–169.

[5] Hassani, H., Heravi, H., and Zhigljavsky, A. (2013a). Forecasting UK industrial production with multivariate singular spectrum analysis. *Journal of Forecasting, 32*, 395–408.

[6] Khan, M. A. R., and Poskitt, D. (2017). Forecasting stochastic processes using singular spectrum analysis: Aspects of the theory and application. *International Journal of Forecasting*, **33**(1), 199–213.

[7] Hassani, H., Soofi, A. S., and Zhigljavsky, A. (2013). Predicting inflation dynamics with singular spectrum analysis. *Journal of the Royal Statistical Society A, 176*, 743–760.

[8] Lai, L., and Guo, K. (2017). The performance of one belt and one road exchange rate: Based on improved singular spectrum analysis. *Physica A: Statistical Mechanics and its Applications*, **483**, 299–308.

[9] Mahmoudvand, R., and Rodrigues, P. C. (2017). A new parsimonious recurrent forecasting model in singular spectrum analysis. *Journal of Forecasting*, DOI: 10.1002/for.2484.

[10] 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.

[11] Merte, S. (2017). Estimating heat wave-related mortality in Europe using singular spectrum analysis. *Climatic Change*, **142**(3-4), 321–330.

[12] 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*, 708–727.

[13] Oropeza, V., and Sachchi, M. (2011). Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis. *Geophysics, 76*, V25–V32.

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

[15] Sanei, S., and Hassani, H. (2015). *Singular spectrum analysis of biomedical signals.* United States: CRC Press.

[16] Golyandina, N., Nekrutkin, V., and Zhigljavsky, A. (2001). Analysis of time series structure: SSA and related techniques. Chapman & Hall/CRC.

[17] Pepelyshev, A. (2010). *Comparison of Recurrent and Vector forecasting.* UK-China workshop on Singular Spectrum Analysis and its applications, Cardiff, September 20, 2010. Available via `http://ssa.cf.ac.uk/pepelyshev/pepelyshev_ssa_forecast.pdf`.

[18] Hassani, H., and Mahmoudvand, R. (2013). Multivariate Singular Spectrum Analysis: A general view and new vector forecasting approach. *International Journal of Energy and Statistics*, **1**(1), 55–83.

[19] 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**(4), 1450029.

[20] 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.

[21] Strogatz, S. H., (2009). Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry and Engineering (Studies in Nonlinearity), The Perseus Books Group.

[22] Harvey, D. I., Leybourne, S. J., and Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting, 13*, 281–291.

[23] Hassani, H., and Silva, E. S. (2015). A Kolmogorov-Smirnov Based Test for Comparing the Predictive Accuracy of Two Sets of Forecasts. *Econometrics*, **3**(3), 590–609.

[24] Elsner, J. B., and Tsonis, A. A. (1996). *Singular Spectrum Analysis: A new tool in time series analysis*. Springer.

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

[26] Hassani, H., Mahmoudvand, R., and Zokaei, M. (2011). Separability and window length in Singular Spectrum Analysis. *Comptes Rendus Mathematique*, **349**(17-18), 987–990.

[27] Hassani, H., Mahmoudvand, R., Zokaie, M., and Ghodsi, M. (2011). On the separability between signal and noise in Singular Spectrum Analysis. *Fluctuation and Noise Letters*, **11**(2), 1250014-1–1250014-11.

[28] Khan, M. A. R., and Poskitt, D. (2013). A note on window length selection in Singular Spectrum Analysis. *Australian and New Zealand Journal of Statistics*, **55**(2), 87–108.

[29] Hassani, H., Heravi, S., Brown, G., and Ayoubkhani, D. (2013). Forecasting before, during, and after recession with singular spectrum analysis. *Journal of Applied Statistics*, **40**(10), 2290–2302.

[30] 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.

[31] Hyndman, R. J., and Khandakar, Y. (2008) Automatic time series forecasting: the forecast package for R. *Journal of Statistical Software*, **27**, 1–22.

[32] 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.

[33] Hassani, H., Silva, E. S., Antonakakis, N., Filis, G., and Gupta, R. (2017). Forecasting accuracy evaluation of tourist arrivals. *Annals of Tourism Research*, **63**, 112–127.

Appendix

Table 7: Nomenclature.

Code	Name of Time Series
A001	US Economic Statistics: Capacity Utilization.
A002	Births by months 1853-2012.
A003	Electricity: electricity net generation: total (all sectors).
A004	Energy prices: average retail prices of electricity.
A005	Coloured fox fur returns, Hopedale, Labrador, 1834-1925.
A006	Alcohol demand (log spirits consumption per head), UK, 1870-1938.
A007	Monthly Sutter county workforce, Jan.1946-Dec.1966 priesema (1979).
A008	Exchange rates - monthly data: Japanese yen.
A009	Exchange rates - monthly data: Pound sterling.
A010	Exchange rates - monthly data: Romanian leu.
A011	HICP (2005 = 100) - monthly data (annual rate of change): European Union (27 countries).
A012	HICP (2005 = 100) - monthly data (annual rate of change): UK.
A013	HICP (2005 = 100) - monthly data (annual rate of change): US.
A014	New Homes Sold in the United States.
A015	Goods, Value of Exports for United States.
A016	Goods, Value of Imports for United States.
A017	Market capitalisation - monthly data: UK.
A018	Market capitalisation - monthly data: US.
A019	Average monthly temperatures across the world (1701-2011): Bournemouth.
A020	Average monthly temperatures across the world (1701-2011): Eskdalemuir.
A021	Average monthly temperatures across the world (1701-2011): Lerwick.
A022	Average monthly temperatures across the world (1701-2011): Valley.
A023	Average monthly temperatures across the world (1701-2011): Death Valley.
A024	US Economic Statistics: Personal Savings Rate.
A025	Economic Policy Uncertainty Index for United States (Monthly Data).
A026	Coal Production, Total for Germany.
A027	Coke, Beehive Production (by Statistical Area).
A028	Monthly champagne sales (in 1000's) (p.273: Montgomery: Fore. and T.S.).
A029	Domestic Auto Production.
A030	Index of Cotton Textile Production for France.

A031 Index of Production of Chemical Products (by Statistical Area).
A032 Index of Production of Leather Products (by Statistical Area).
A033 Index of Production of Metal Products (by Statistical Area).
A034 Index of Production of Mineral Fuels (by Statistical Area).
A035 Industrial Production Index.
A036 Knit Underwear Production (by Statistical Area).
A037 Lubricants Production for United States.
A038 Silver Production for United States.
A039 Slab Zinc Production (by Statistical Area).
A040 Annual domestic sales and advertising of Lydia E, Pinkham Medicine, 1907 to 1960.
A041 Chemical concentration readings.
A042 Monthly Boston armed robberies Jan.1966-Oct.1975 Deutsch and Alt (1977).
A043 Monthly Minneapolis public drunkenness intakes Jan.66-Jul78.
A044 Motor vehicles engines and parts/CPI, Canada, 1976-1991.
A045 Methane input into gas furnace: cu. ft/min. Sampling interval 9 seconds.
A046 Monthly civilian population of Australia: thousand persons. Feb 1978-Apr 1991.
A047 Daily total female births in California, 1959.
A048 Annual immigration into the United States: thousands. 1820-1962.
A049 Monthly New York City births: unknown scale. Jan 1946-Dec 1959.
A050 Estimated quarterly resident population of Australia: thousand persons.
A051 Annual Swedish population rates (1000's) 1750-1849 Thomas (1940).
A052 Industry sales for printing and writing paper (in Thousands of French francs).
A053 Coloured fox fur production, Hebron, Labrador, 1834-1925.
A054 Coloured fox fur production, Nain, Labrador, 1834-1925.
A055 Coloured fox fur production, oak, Labrador, 1834-1925.
A056 Monthly average daily calls to directory assistance Jan.62-Dec76.
A057 Monthly Av. residential electricity usage Iowa city 1971-1979.
A058 Montly av. residential gas usage Iowa (cubic feet)*100 71-79.
A059 Monthly precipitation (in mm), Jan 1983-April 1994. London, United Kingdom .
A060 Monthly water usage (ml/day), London Ontario, 1966-1988.
A061 Quarterly production of Gas in Australia: million megajoules. Includes natural gas from July 1989. Mar 1956-Sep 1994.
A062 Residential water consumption, Jan 1983-April 1994. London, United Kingdom.
A063 The total generation of electricity by the U.S. electric industry (monthly data for the period Jan. 1985-Oct. 1996).

A064 Total number of water consumers, Jan 1983-April 1994. London, United Kingdom.
A065 Monthly milk production: pounds per cow. Jan 62-Dec 75.
A066 Monthly milk production: pounds per cow. Jan 62-Dec 75, adjusted for month length.
A067 Monthly total number of pigs slaughtered in Victoria. Jan 1980-August 1995.
A068 Monthly demand repair parts large/heavy equip. Iowa 1972-1979.
A069 Number of deaths and serious injuries in UK road accidents each month. Jan 1969-Dec 1984.
A070 Passenger miles (Mil) flown domestic U.K. Jul. 62-May 72.
A071 Monthly hotel occupied room av. 63-76 B.L.Bowerman et al.
A072 Weekday bus ridership, Iowa city, Iowa (monthly averages).
A073 Portland Oregon average monthly bus ridership (/100).
A074 U.S. airlines: monthly aircraft miles flown (Millions) 1963-1970.
A075 International airline passengers: monthly totals in thousands. Jan 49-Dec 60.
A076 Sales: souvenir shop at a beach resort town in Queensland, Australia. Jan 1987-Dec 1993.
A077 Der Stern: Weekly sales of wholesalers A, 71-72.
A078 Der Stern: Weekly sales of wholesalers B, 71-72'
A079 Der Stern: Weekly sales of wholesalers 71-72.
A080 Monthly sales of U.S. houses (thousands) 1965-1975.
A081 CFE specialty writing papers monthly sales.
A082 Monthly sales of new one-family houses sold in USA since 1973.
A083 Wisconsin employment time series, food and kindred products, Jan. 1961-Oct. 1975.
A084 Monthly gasoline demand Ontario gallon millions 1960-1975.
A085 Wisconsin employment time series, fabricated metals, Jan. 1961-Oct. 1975.
A086 Monthly empolyees wholes./retail Wisconsin 61-75 R.B.Miller.
A087 US monthly sales of chemical related products. Jan 1971-Dec 1991.
A088 US monthly sales of coal related products. Jan 1971-Dec 1991.
A089 US monthly sales of petrol related products. Jan 1971-Dec 1991.
A090 US monthly sales of vehicle related products. Jan 1971-Dec 1991.
A091 Civilian labour force in Australia each month: thousands of persons. Feb 1978-Aug 1995.
A092 Numbers on Unemployment Benefits in Australia: monthly Jan 1956-Jul 1992.
A093 Monthly Canadian total unemployment figures (thousands) 1956-1975.
A094 Monthly number of unemployed persons in Australia: thousands. Feb 1978-Apr 1991.
A095 Monthly U.S. female (20 years and over) unemployment figures 1948-1981.
A096 Monthly U.S. female (16-19 years) unemployment figures (thousands) 1948-1981.

A097	Monthly unemployment figures in West Germany 1948-1980.
A098	Monthly U.S. male (20 years and over) unemployment figures 1948-1981.
A099	Wisconsin employment time series, transportation equipment, Jan. 1961-Oct. 1975.
A100	Monthly U.S. male (16-19 years) unemployment figures (thousands) 1948-1981.

Table 8: Descriptives for the 100 time series.

Code	F	N	Mean	Med.	SD	CV	Skew.	SW(p)	ADF	*Code*	F	N	Mean	Med.	SD	CV	Skew.	SW(p)	ADF
A001	M	539	80	80	5	6	-0.55	<0.01	-0.60†	**A002**	M	1920	271	249	88	33	0.16	<0.01	-1.82†
A003	M	484	2.59x10^5	2.61x10^5	6.88x10^5	27	0.15	<0.01	-0.90†	**A004**	M	310	7	7	2	28	-0.24	<0.01	0.56†
A005	D	92	47.63	31.00	47.33	99.36	2.27	<0.01	-3.16	**A006**	Q	207	1.95	1.98	0.25	12.78	-0.58	<0.01	0.46†
A007	M	252	2978	2741	1111	37.32	0.79	<0.01	-0.80†	**A008**	M	160	128	128	19	15	0.34	<0.01	-0.59†
A009	M	160	0.72	0.69	0.10	13	0.66	<0.01	0.53†	**A010**	M	160	3.41	3.61	0.83	24	-0.92	<0.01	1.58†
A011	M	201	4.7	2.6	5.0	106	2.24	<0.01	-2.66	**A012**	M	199	2.1	1.9	1.0	49	0.92	<0.01	-0.79†
A013	M	176	2.5	2.4	1.6	66	-0.52	<0.01	-2.27†	**A014**	M	606	55	53	20	35	0.79	<0.01	-1.41†
A015	M	672	3.39	1.89	3.48	103	1.09	<0.01	2.46†	**A016**	M	672	5.18	2.89	5.78	111	1.13	<0.01	1.91†
A017	M	249	130	130	24	19	0.35	<0.01	0.24†	**A018**	M	249	112	114	25	22	-0.01	0.01*	0.06†
A019	M	605	10.1	9.6	4.5	44	0.05	<0.01	-4.77	**A020**	M	605	7.3	6.9	4.3	59	0.04	<0.01	-6.07
A021	M	605	7.2	6.8	3.3	46	0.13	<0.01	-4.93	**A022**	M	605	10.3	9.9	3.8	37	0.04	<0.01	-4.19
A023	M	605	24	24	10	40	-0.02	<0.01	-7.15	**A024**	M	636	6.9	7.4	2.6	38	-0.29	<0.01	-1.18†
A025	M	343	108	100	33	30	0.99	<0.01	-1.23†	**A026**	M	277	11.7	11.9	2.3	20	-0.16	0.06*	-0.40†
A027	M	171	0.21	0.13	0.19	88	1.26	<0.01	-1.81†	**A028**	M	96	4801	4084	2640	54.99	1.55	<0.01	-1.66†
A029	M	248	391	385	116	30	-0.03	0.08*	-1.22†	**A030**	M	139	89	92	12	13	-0.82	<0.01	-0.28†
A031	M	121	134	138	27	20	0.05	<0.01	1.51†	**A032**	M	153	113	114	10	9	-0.29	0.45*	-0.52†
A033	M	115	117	118	17	15	-0.29	0.03*	-0.46†	**A034**	M	115	110	111	11	10	-0.53	0.02*	0.30†
A035	M	1137	40	34	31	78	0.56	<0.01	5.14†	**A036**	M	165	1.08	1.10	0.20	18.37	-1.15	<0.01	-0.59†
A037	M	479	3.04	2.83	1.02	33.60	0.46	<0.01	0.61†	**A038**	M	283	9.39	10.02	2.27	24.15	-0.80	<0.01	-1.01†
A039	M	452	54	52	19	36	-0.15	<0.01	0.08†	**A040**	Q	108	1382	1206	684	49.55	0.83	<0.01	-0.80†
A041	H	197	17.06	17.00	0.39	2.34	0.15	0.21*	0.09†	**A042**	M	118	196.3	166.0	128.0	65.2	0.45	<0.01	0.41†
A043	M	151	391.1	267.0	237.49	60.72	0.43	<0.01	-1.17†	**A044**	M	188	1344	1425	479.1	35.6	-0.41	<0.01	-1.28†
A045	H	296	-0.05	0.00	1.07	-1887	-0.05	0.55*	-7.66	**A046**	M	159	11890	11830	882.93	7.42	0.12	<0.01	5.71
A047	D	365	41.98	42.00	7.34	17.50	0.44	<0.01	-1.07†	**A048**	A	143	2.5x10^5	2.2x10^5	2.1x10^5	83.19	1.06	<0.01	-2.63
A049	M	168	25.05	24.95	2.31	9.25	-0.02	0.02*	0.07†	**A050**	Q	89	15274	15184	1358	8.89	0.19	<0.01	9.72†
A051	A	100	6.69	7.50	5.88	87.87	-2.45	<0.01	-3.06	**A052**	M	120	713	733	174	24.39	-1.09	<0.01	-0.78†
A053	A	91	81.58	46.00	102.07	125.11	2.80	<0.01	-3.44	**A054**	A	91	101.80	77.00	92.14	90.51	1.43	<0.01	-3.38
A055	A	91	59.45	39.00	60.42	101.63	1.56	<0.01	-3.99	**A056**	M	180	492.50	521.50	189.54	38.48	-0.17	<0.01	-0.65†
A057	M	106	489.73	465.00	93.34	19.06	0.92	<0.01	-1.21†	**A058**	M	106	124.71	94.50	84.15	67.48	0.52	<0.01	-3.88
A059	M	136	85.66	80.25	37.54	43.83	0.91	<0.01	-1.88†	**A060**	M	276	118.61	115.63	26.39	22.24	0.86	<0.01	-0.47†
A061	Q	155	61728	47976	53907	87.33	0.44	<0.01	0.06†	**A062**	M	136	5.72x10^7	5.53x10^7	1.2x10^7	21.51	1.13	<0.01	-0.84†
A063	M	142	231.09	226.73	24.37	10.55	0.52	0.01	-0.39†	**A064**	M	136	31388	31251	3232	10.30	0.25	0.22*	-0.16†
A065	M	156	754.71	761.00	102.20	13.54	0.01	0.04*	0.04†	**A066**	M	156	746.49	749.15	98.59	13.21	0.08	0.04*	-0.38†
A067	M	188	90640	91661	13926	15.36	-0.38	0.01*	-0.38†	**A068**	M	94	1540	1532	474.35	30.79	0.38	0.05*	0.54†
A069	M	192	1670	1631	289.61	17.34	0.53	<0.01	-0.74†	**A070**	M	119	91.09	86.20	32.80	36.01	0.34	<0.01	-1.93†
A071	M	168	722.30	709.50	142.66	19.75	0.72	<0.01	-0.52†	**A072**	W	136	5913	5500	1784	30.17	0.67	<0.01	-0.68†
A073	M	114	1120	1158	270.89	24.17	-0.37	<0.01	0.76†	**A074**	M	96	10385	10401	2202	21.21	0.33	0.18*	-0.13†

A075	M	144	280.30	265.50	119.97	42.80	0.57	<0.01	-0.35†	**A076**	M	84	14315	8771	15748	110	3.37	<0.01	-0.29†
A077	W	104	11909	11640	1231	10.34	0.60	<0.01	-0.16†	**A078**	W	104	74636	73600	4737	6.35	0.64	<0.01	-0.59†
A079	W	104	1020	1012	71.78	7.03	0.60	0.01*	-0.41†	**A080**	M	132	45.36	44.00	10.38	22.88	0.17	0.15*	-0.81†
A081	M	147	1745	1730	479.52	27.47	-0.39	<0.01	-1.15†	**A082**	M	275	52.29	53.00	11.94	22.83	0.18	0.13*	-1.30†
A083	M	178	58.79	55.80	6.68	11.36	0.93	<0.01	-0.92†	**A084**	M	192	1.62x10^{5}	1.57x10^{5}	41661	25.71	0.32	<0.01	0.25†
A085	M	178	40.97	41.50	5.11	12.47	-0.07	<0.01	1.45†	**A086**	M	178	307.56	308.35	46.76	15.20	0.17	<0.01	1.51†
A087	M	252	13.70	14.08	6.13	44.73	0.16	<0.01	1.13†	**A088**	M	252	65.67	68.20	14.25	21.70	-0.53	<0.01	-0.53†
A089	M	252	10.76	10.92	5.11	47.50	-0.19	<0.01	-0.05†	**A090**	M	252	11.74	11.05	5.11	43.54	0.38	<0.01	-0.88†
A091	M	211	7661	7621	819	10.70	0.03	<0.01	3.27†	**A092**	M	439	2.21x10^{5}	5.67x10^{4}	2.35x10^{5}	106.32	0.77	<0.01	1.61†
A093	M	240	413.28	396.50	152.84	36.98	0.36	<0.01	-1.60†	**A094**	M	211	6787	6528	604.62	8.91	0.56	<0.01	2.69†
A095	M	408	1373	1132	686.05	49.96	0.91	<0.01	0.60†	**A096**	M	408	422.38	342.00	252.86	59.87	0.65	<0.01	-1.95†
A097	M	396	7.14x10^{5}	5.57x10^{5}	5.64x10^{5}	78.97	0.79	<0.01	-2.51†	**A098**	M	408	1937	1825	794	41.04	0.64	<0.01	-1.15†
A099	M	178	40.60	40.50	4.95	12.19	-0.65	<0.01	-0.10†	**A100**	M	408	520.28	425.50	261.22	50.21	0.64	<0.01	-1.65†

Note:* indicates data is normally distributed based on a Shapiro-Wilk test at p=0.01.

† indicates a nonstationary time series based on the Augmented Dickey-Fuller test at p=0.01.

A indicates annual, M indicates monthly, Q indicates quarterly, W indicates weekly, D indicates daily and H indicates hourly.

N indicates series length.

Table 9: RMSE for out-of-sample forecasts.

	VSSA				RSSA				$\frac{VSSA}{RSSA}$			
Code	*1*	*3*	*6*	*12*	*1*	*3*	*6*	*12*	*1*	*3*	*6*	*12*
A001	0.56	1.12	2.03	3.47	0.57	1.08	2.01	3.48	0.98	1.04	1.01	0.99
A002	25.1	26.98	28.46	29.72	25.5	26.66	28.25	29.93	0.98	1.01	1.01	0.99
A003	10782	12128	12843	13937	11072	12476	13296	14174	0.97	0.97	0.97	0.98
A004	0.22	0.25	0.28	0.40	0.21	0.25	0.28	0.38	1.05	1.00	1.00	1.05
A005	47.06	47.06	43.71	53.57	49.22	49.22	45.96	52.82	0.96	0.96	0.95	1.01
A006	0.05	0.08	0.12	0.18	0.05	0.08	0.12	0.18	1.00	1.00	1.00	1.00
A007	316.13	367.98	391.86	404.19	313.11	357.37	386.55	393.41	1.01	1.03	1.01	1.03
A008	5.31	11.91	16.76	21.00	5.21	11.91	16.76	21.00	1.02*	1.00	1.00	1.00
A009	0.02	0.04	0.06	0.08	0.02	0.04	0.06	0.08	1.00	1.00	1.00	1.00
A010	0.08	0.17	0.23	0.28	0.08	0.17	0.23	0.27	1.00	1.00	1.00	1.04
A011	0.27	0.47	0.65	0.85	0.30	0.48	0.65	0.87	0.90*	0.98	1.00	0.97
A012	0.45	0.91	1.16	1.17	0.46	0.91	1.22	1.21	0.98	1.00	0.95	0.97
A013	0.68	1.88	2.61	2.58	0.77	2.09	2.51	2.42	0.88*	0.90	1.04	1.07
A014	5.37	7.35	8.50	10.71	5.19	7.35	8.42	10.85	1.03*	1.00	1.01	0.99
A015	$3.29\text{x}10^9$	$5.83\text{x}10^9$	$7.78\text{x}10^9$	$1.03\text{x}10^{10}$	$3.33\text{x}10^9$	$5.96\text{x}10^9$	$7.78\text{x}10^9$	$1.03\text{x}10^{10}$	0.99	0.98	1.00	1.00
A016	$6.18\text{x}10^9$	$1.05\text{x}10^{10}$	$1.68\text{x}10^{10}$	$2.01\text{x}10^{10}$	$6.27\text{x}10^9$	$1.07\text{x}10^{10}$	$1.68\text{x}10^{10}$	$2.02\text{x}10^{10}$	0.99	0.98	1.00	1.00
A017	7.34	11.50	17.59	22.06	7.35	11.34	17.59	22.20	0.95	1.01	1.00	0.99
A018	5.70	9.62	14.89	21.00	5.76	9.51	14.89	21.27	0.99	1.01	1.00	0.99
A019	1.34	1.31	1.30	1.30	1.32	1.30	1.30	1.30	1.02	1.01	1.00	1.00
A020	1.29	1.28	1.28	1.29	1.28	1.28	1.28	1.29	1.01	1.00	1.00	1.00
A021	1.01	1.06	1.05	1.06	1.04	1.05	1.05	1.05	0.97	1.01	1.00	1.01
A022	1.13	1.15	1.15	1.15	1.14	1.14	1.14	1.15	0.99	1.01	1.01	1.00
A023	1.79	1.91	1.95	1.94	1.82	1.89	1.95	2.09	0.98	1.01	1.00	0.92
A024	0.71	0.89	1.09	1.25	0.71	0.89	1.07	1.28	1.00	1.00	1.02	0.98
A025	21.50	27.06	29.38	33.57	21.2	25.54	30.12	33.78	1.01	1.06	0.98	0.99
A026	1.14	1.57	2.38	2.80	1.17	1.58	2.35	2.83	0.97	0.99	1.01	0.99
A027	0.05	0.11	0.16	0.17	0.05	0.12	0.16	0.16	1.00	0.92	1.00	1.06
A028	1355	1342	1325	1319	1307	1347	1317	1338	1.02	0.99	1.01	0.99

A029	37.86	50.24	64.62	73.16	38.04	48.83	67.59	76.70	0.99	1.03	0.96	0.95
A030	11.07	12.44	14.14	14.44	11.21	12.48	14.44	14.44	0.99	0.99	0.98	1.00
A031	1.90	3.02	5.65	8.67	1.95	3.06	5.73	8.86	0.97	0.99	0.99	0.98
A032	4.95	7.14	8.92	8.37	5.13	7.41	9.13	8.66	0.96	0.96	0.97	0.97
A033	8.27	11.22	14.36	14.60	8.36	10.66	11.88	14.97	0.99	1.05	1.21	0.98
A034	2.52	4.15	6.06	7.27	2.51	4.56	6.08	7.12	1.00	0.91*	1.00	1.02
A035	0.51	0.99	1.98	3.72	0.51	0.95	1.82	3.50	1.00	1.04*	1.09	1.06*
A036	0.09	0.15	0.14	0.17	0.10	0.15	0.15	0.19	0.90	1.00	0.93	0.89
A037	0.28	0.33	0.39	0.47	0.27	0.33	0.38	0.46	1.04	1.00	1.02	1.02
A038	1.33	1.42	1.54	1.76	1.36	1.45	1.54	1.76	0.98	0.97	1.00	1.00
A039	3.54	5.38	6.61	6.92	3.62	5.22	6.82	7.18	0.98	1.03	0.97	0.96
A040	261	656	934	1120	267	642	917	1039	0.98	1.02	1.02	1.08*
A041	1.90	3.02	5.65	8.67	1.95	3.06	5.73	8.86	0.97	0.99	0.99	1.00
A042	54.31	66.38	73.19	73.63	54.70	63.34	68.19	67.80	0.99	1.05	1.07	1.09
A043	33.17	37.63	48.23	73.29	33.73	37.63	48.23	73.90	0.98	1.00	1.00	0.99
A044	288.66	330.01	344.89	296.53	288.03	341.19	359.28	369.66	1.00	0.97	0.96*	0.80*
A045	0.18	0.64	0.93	0.98	0.19	0.66	0.99	1.00	0.95*	0.97*	0.94	0.98
A046	1.04	1.04	8.22	8.22	1.06	1.06	8.24	8.24	0.98	0.98	0.99	0.99
A047	7.35	7.47	7.60	7.64	7.51	7.57	7.63	7.66	0.98	0.99	0.99	0.99
A048	126025	164402	210423	274295	127864	166804	198390	251622	0.99	0.99	1.06*	1.09
A049	0.87	1.04	1.03	1.10	0.92	1.03	1.03	1.11	0.95	1.01	1.00	0.99
A050	27.64	34.73	62.19	94.28	27.61	36.34	63.08	94.12	1.00	0.96	0.99	1.00
A051	3.49	4.41	4.46	4.60	3.49	4.51	4.55	4.63	1.00	0.98	0.98	0.99
A052	53.66	55.16	55.92	63.02	53.64	54.58	56.23	62.37	1.00	1.01	0.99	1.01
A053	61.22	60.44	59.49	72.41	62.70	63.34	62.36	73.45	0.98	0.97	0.95	0.99
A054	84.24	82.08	80.20	88.85	86.57	83.73	82.02	92.90	0.97	0.98	0.98	0.96*
A055	41.95	42.11	42.86	41.78	42.06	42.32	42.66	41.96	0.99	0.97	1.00	0.99
A056	71.51	71.51	178.72	290.87	71.51	71.51	178.72	290.87	1.00	1.00	1.00	1.00
A057	44.07	44.91	45.85	46.74	42.97	44.90	45.31	43.67	1.03	1.00	1.01	1.07
A058	14.74	20.68	20.79	19.97	14.47	21.32	21.49	20.21	1.02	0.97	0.97	0.99
A059	37.96	38.13	38.55	38.41	38.20	38.64	39.03	38.88	0.99	0.99	0.99	0.99
A060	8.14	8.84	8.28	8.72	8.22	9.15	8.69	8.69	0.99	0.96	0.95	1.00
A061	7474.14	9996.28	17162.38	17162.38	7582.66	9869.33	16140.62	16140.62	0.99	1.01	1.06	1.06

A062	$6.60\text{x}10^6$	$7.32\text{x}10^6$	$7.12\text{x}10^6$	$7.85\text{x}10^6$	$6.67\text{x}10^6$	$7.08\text{x}10^6$	$7.12\text{x}10^6$	$7.48\text{x}10^6$	0.99	1.03	1.00	1.05
A063	6.92	6.95	7.11	6.84	6.99	7.04	7.07	6.98	0.99	0.99	1.01	0.98
A064	2914	2798	3031	3248	3009	2872	3025	3412	0.97	0.97*	1.00	0.95
A065	8.76	15.60	22.51	26.04	9.57	15.19	20.00	26.82	0.92	1.03	1.13	0.97*
A066	9.67	15.54	18.66	24.45	9.21	15.00	18.62	24.67	1.05	1.04	1.01	0.99
A067	9117	10323	11359	10944	9334	10390	11550	10970	0.98	0.99	0.98	0.99
A068	350.57	372.89	377.05	311.86	362.95	381.59	379.63	311.86	0.96	0.98*	0.99	1.00
A069	140.60	162.83	173.12	169.46	141.37	156.66	165.84	165.44	0.99	1.04	1.04*	1.02*
A070	6.05	6.60	6.69	6.44	6.23	6.75	6.81	6.83	0.97	0.98	0.98	0.94
A071	21.65	21.80	22.11	21.86	21.29	22.18	22.27	22.24	1.02	0.98	0.99	0.98
A072	544.61	626.01	681.04	707.54	550.32	640.10	695.69	723.97	0.99	0.98	0.98	0.98
A073	47.80	83.63	131.87	172.71	52.20	94.67	131.88	175.56	0.92	0.88	0.99	0.98
A074	1148	1165	1225	1146	1184	1200	1229	1148	0.97	0.97*	0.99	0.99
A075	15.64	21.63	25.16	25.06	15.91	20.76	24.28	26.27	0.98	1.04	1.04	0.95
A076	7376	7467	7356	7395	7313	7404	7369	7386	1.01	1.01	0.99	1.00
A077	284.40	584.83	940.23	1671.60	295.47	566.41	940.23	1671.60	0.96	1.03	1.00	1.00
A078	786.58	1470	2937	7008	836.09	1295	2258	7008	1.14	1.14	1.06*	1.00
A079	35.23	41.08	57.20	90.34	36.97	43.81	57.20	90.35	0.95	0.94	1.00	0.99
A080	4.40	7.79	10.38	11.54	4.27	7.48	10.22	11.71	1.03*	1.04	1.02	0.99
A081	359.25	428.19	402.64	435.72	359.01	460.02	446.26	434.48	1.00	0.93	0.90	1.00
A082	5.16	6.58	6.85	7.81	5.15	6.42	6.85	8.00	1.00	1.02	1.00	0.98
A083	1.26	1.67	1.68	1.86	1.22	1.62	1.66	1.87	1.03	1.03	0.89	0.99
A084	9187	8977	9362	9260	9244	9159	9218	9355	0.99	0.98	1.02	0.99
A085	0.57	1.18	1.88	2.96	0.60	1.16	2.00	3.14	0.95	1.02	0.94	0.94*
A086	1.73	3.69	4.56	5.70	1.78	3.43	4.24	5.73	0.97	1.08	1.08	0.99
A087	0.61	0.85	1.12	1.63	0.62	0.83	1.10	1.56	0.98	1.02	1.01	1.04
A088	4.02	4.14	4.22	4.15	3.73	4.04	4.21	4.23	1.08	1.02	1.00	0.98
A089	0.86	1.86	2.40	3.06	0.90	1.86	2.41	3.01	0.96	1.00	0.99	1.01
A090	1.90	2.13	2.32	2.38	1.95	2.10	2.17	2.18	0.97	1.01	1.07	1.09
A091	90.35	99.09	130.33	151.02	91.63	99.36	130.33	151.21	0.99	0.99	1.00	0.99
A092	12684	31798	61048	114449	12605	29818	57383	114332	1.01	1.07	1.06	1.00
A093	28.75	56.62	72.69	83.36	31.78	58.39	78.35	88.52	0.90*	0.97	0.93	0.94
A094	42.41	62.21	87.36	140.26	40.61	63.88	88.70	135.15	1.04	0.97	0.98	1.04

A095	125.38	212.82	280.64	387.12	123.12	213.13	280.64	386.79	1.02	0.99	1.00	1.00
A096	53.58	60.33	68.99	77.91	54.94	61.13	69.24	78.69	0.98	0.99	0.99	0.99
A097	32113	71065	118949	186205	31455	67931	117557	183615	1.02	1.05*	1.01	1.01
A098	148.99	320.59	489.00	643.24	147.85	320.58	503.24	644.74	1.01	1.00	0.97	0.99
A099	2.63	2.90	3.29	3.73	2.77	2.92	3.28	3.75	0.95	0.99	1.00	0.99
A100	65.59	82.26	98.41	119.21	65.44	80.83	93.73	120.72	1.00	1.02	1.05	0.99