Statistics and Its Interface Volume 0 (2021) 1–9

Study of automatic choice of parameters for forecasting in singular spectrum analysis

Safia ALMarhoobi, and Andrey Pepelyshev*

Singular spectrum analysis (SSA) is a popular tool for analysing and forecasting time series. The SSA forecasting algorithms have two parameters which should be chosen by the researcher or using the so-called automatic choice based on the root mean squared errors (RMSE) of retrospective forecasts. We study the sensitivity of the RMSE and investigate the reliability of the automatic choice of parameters for forecasting monthly temperature and humidity recorded at three meteorological stations in Oman.

AMS 2000 subject classifications: Primary 62M20.
Keywords and phrases: Recurrent SSA, Vector SSA, temperature, humidity.

1. INTRODUCTION

Singular Spectrum Analysis (SSA) is a non-parametric, powerful method for time series analysis and provides meaningful findings in many research areas without imposing any restrictive assumptions on processed data [5, 6, 8]. SSA can be used for parametric estimation, forecasting, and gap filling [4, 11, 16]. Moreover, SSA is a very useful tool for extracting various signals from noisy observations [9, 12]. A modification of SSA was used in [7] to find structures in short time series by extracting seasonality and simultaneously extracting cycles of small and long periods.

Basic SSA is a core version of SSA that involves embedding a time series to the space of Hankel matrices and then decomposing it into rank-one matrices using the standard singular value decomposition. By then inverting the embedding procedure, SSA yields a decomposition of the original time series into the sum of number of components such as a trend, oscillatory components, and noise [5, 8, 11].

The SSA algorithm has two parameters: the window length L and the number of singular values r. Choice of parameters L and r is depending on both the structure of the time series and the forecasting aims [5, 10]. The window length L defines the number of columns of the Hankel matrix and an inappropriate choice of L leads to a poor decomposition, incomplete reconstruction, and nonaccurate forecasting [3, 14, 19]. The parameter r should correspond to the rank of the signal. As noted in [3, 10], the authors argue that the value of r needs to be greater than what it should be. The increase of noise in the reconstructed series is likely to be small and therefore less detrimental than missing parts of the signal which can occur if r is too small.

There are two main SSA forecasting algorithms: recurrent SSA (SSA-R) forecasting and vector SSA (SSA-V) forecasting, which both depends on two parameters L and r. These parameters can be chosen by an expert by looking on the signal structure or using the automatic choice based on the root mean squared error (RMSE) of retrospective forecasts [5].

For assessing the quality of the automatic choice, we firstly analyze the sensitivity of the RMSE on parameters and then perform a study of reliability of the automatic choice for forecasting monthly temperature and humidity recorded at three meteorological stations in Oman.

The remainder of this paper is structured as follows. Section 2 provides a brief description of the methodology of SSA, forecasting algorithms and the automatic choice of parameters. Section 3 introduces time series of temperature and humidity at three meteorological stations in Oman. Section 4 reports the sensitivity of the RMSE. Section 5 studies the accuracy of forecasts with the automatic choice of parameters. Section 6 offers some concluding remarks.

2. SSA METHODOLOGY

2.1 Basic SSA

Following [6, Sec 2.1], Basic SSA can be described in brief as follows. Let $x_1, x_2, \ldots, x_N$ be a time series. For a given window length L $(1 < L < N)$, we construct the L-lagged vectors $X_{(i)} = (x_i, \cdots, x_{i+L-1})^T$, $i = 1, 2, \ldots, K = N - L + 1$, and compose these vectors into the Hankel $L \times K$-matrix

$$\mathbf{X} = (x_{i+j-1})_{i,j=1}^{L,K} = \left[X_{(1)}, \ldots, X_{(K)}\right],$$

which is called 'trajectory matrix'. The singular-value decomposition of the matrix $\mathbf{X}\mathbf{X}^T$ gives a set of L eigenvalues and eigenvectors. For a given integer r when $1 < r < L$ we create a group using r largest eigenvalues and corresponding eigenvectors of $\mathbf{X}\mathbf{X}^T$. The chosen eigenvectors determine an r-dimensional subspace in R^L which is denoted as S_r. The L-dimensional data $X_{(1)}, \ldots, X_{(K)}$ is then projected onto this r-dimensional subspace S_r and the subsequent averaging over the diagonals gives us some Hankel matrix $\tilde{\mathbf{X}}$, which

arXiv: 0000.0000
*Corresponding author.

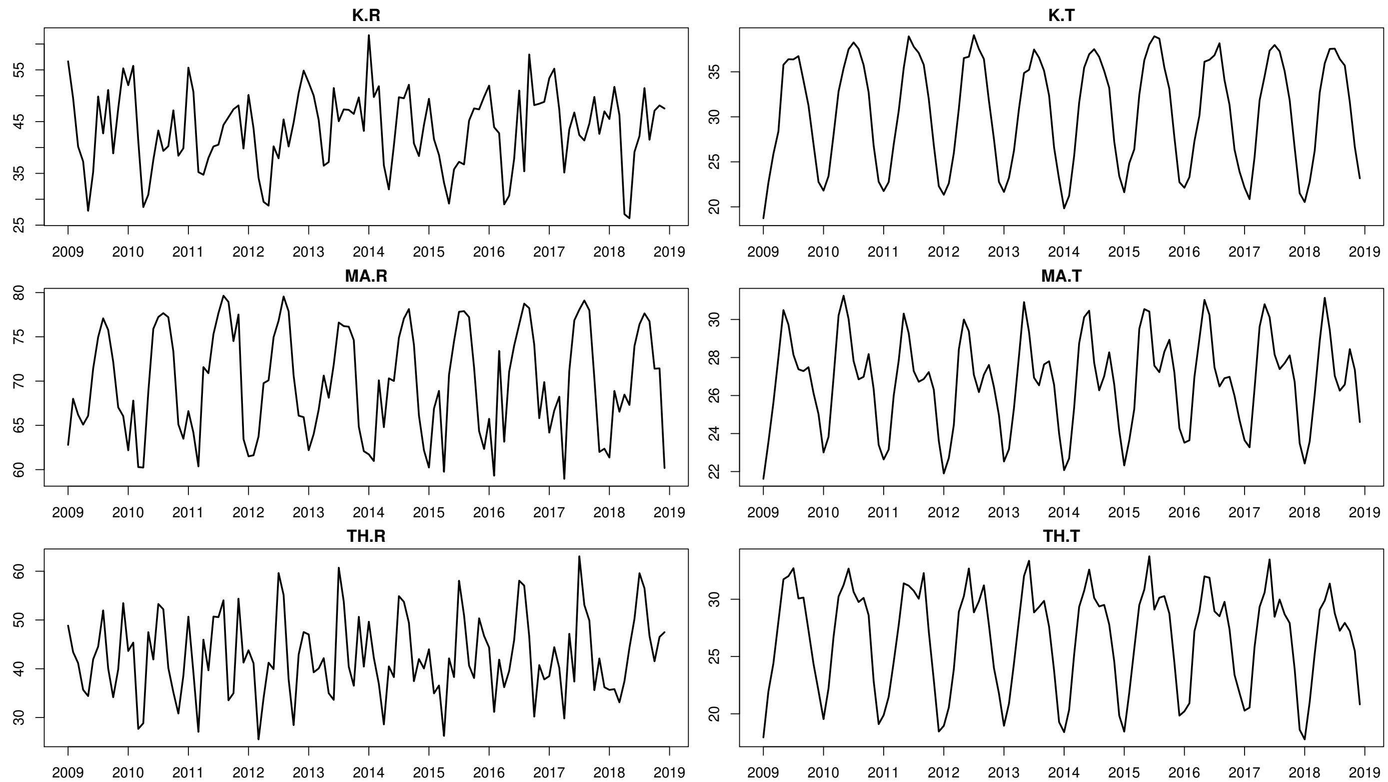

Figure 1. Monthly humidity (left) and temperature (right) at the stations K, MA and TH from 2009 to 2018.

we consider as an SSA approximation to $\mathbf{X}$. With a proper choice of r and L, the time series corresponding to $\tilde{\mathbf{X}}$ is called 'reconstructed series' and often used as an estimator of a signal or a trend.

The general guideline for selecting r and L is to take sufficiently large, say $L \approx \frac{N}{2}$ and, if we want to extract a periodic component with known period, to take the window length to be divisible by the period, while r is chosen on the base of relations between eigenvalues and the spectral properties of eigenvectors considered as time series, see [5].

2.2 Recurrent SSA forecasting

Recurrent SSA forecasting is based on the linear recurrence relation (LRR)

$$y_{i+d} = \sum_{k=1}^{d} a_k y_{i+d-k},$$

where $y_1, y_2, \ldots$ is a series and $a_1, \ldots, a_d$ are coefficients such that $a_d \neq 0$, see [5, Sec 3.1.1.1]. Applying the LRR iteratively, we can perform the natural recurrent continuation of any given series.

In the recurrent SSA forecasting algorithm, see e.g. [5, Sec 3.2.1.2], the coefficients $a_1, \ldots, a_d$ are computed using eigenvectors of the matrix $\mathbf{X}\mathbf{X}^T$ and the forecasted time series $(y_1, \ldots, y_{N+M})$ is defined by

$$y_i = \begin{cases} \tilde{x}_i & \text{for } i = 1, \ldots, N, \\ \sum_{j=1}^{d-1} a_j y_{i-j} & \text{for } i = 1+N, \ldots, N+M, \end{cases}$$

where $\tilde{x}_1, \ldots, \tilde{x}_N$ is the reconstructed time series.

2.3 Vector SSA forecasting

Following [5, Sec 3.2.1.3] the vector SSA forecasting algorithm is based on the sequential expansion of the set $Z_1, \ldots, Z_K$, where Z_i is the projection of the vector $X_{(i)}$ on the subspace S_r. The expanding vectors Z_k with $k > K$ are obtained from Z_{k-1} by shifting components and adding a special value into the last component. Hankelization of the matrix $(Z_1, \ldots, Z_k, Z_{k+1}, Z_{k+2}, \ldots)$ produces the vector forecast.

2.4 RMSE and the automatic choice of parameters

The root mean square error (RMSE) is the most popular criterion to measure the error of forecasts, see e.g. [5, 18, 15, 10]. The RMSE of the $1, 2, \ldots, h$-step ahead forecasts with

several truncation points is given by
(1)

$$\mathrm{RMSE}_{T_1}^{T_2} = \left(\frac{1}{(T_2 - T_1 + 1)h} \sum_{T=T_1}^{T_2} \sum_{j=1}^{h} (\tilde{y}_{T,j} - y_{T+j})^2 \right)^{1/2},$$

where $\tilde{y}_{T,j}$ is the j-step ahead forecast of the truncated time series $y_1, \ldots, y_T$, the value y_{T+j} is the true value of the given time series at time $T+j$, T_1 and T_2 are the first and last truncation points, respectively.

For the task of forecasting a given time series, the automatic choice of parameters L and r relies on finding values of parameters minimising the RMSE of forecasts with desired forecasting horizons, see [5, Sec 3.5.7] and [2, 11, 13, 17]. Further, the forecasting algorithm with automatically chosen parameters is applied to obtain future forecasts.

3. REAL TIME SERIES

We study the accuracy of forecasting algorithms for monthly time series of temperature (measured in deg C) and humidity (measured in %), which were provided by the Directorate General of Meteorology of Oman. The data was collected from Jan 2009 to Dec 2018 at three meteorological stations in the Sultanate of Oman: the Khasab Airport (K), the Masirah (MA) and the Thumrait (TH) stations.

In Figure 1 we depict all time series which do not exhibit trends as shown in [1]. We can observe that the annual pattern of temperature is rather stable from year to year. In particular, temperature at the station K has a simple sinusoidal shape but the annual pattern of temperature at the stations MA and TH is more complicated. We can also see that humidity is very volatile and the annual pattern is clearly visible only at the station MA. Note that humidity at the station MA is much larger than humidity at stations K and TH.

4. DEPENDENCE OF THE RMSE ON PARAMETERS

Let us study how the RMSE of 1,2,...,12-step ahead forecasts across several truncation points depends on the parameters L and r. We take Jan 2017 as the first truncation point and Dec 2017 as the last truncation point.

In Tables 1–6 we show the RMSE for six time series using two SSA forecasting algorithms with the wide range of parameters L and r. The lowest values for the RMSE for each forecasting algorithm are highlighted. We can see that the RMSE does not monotonically depends on parameters but tendencies are rather clear. Specifically, the RMSE is larger for very small r because such small values of r are smaller than the signal rank. The RMSE also become larger for large r especially for noisy time series because such a large r is greater than the signal rank and forecasting becomes unstable. We can observe that the RMSE for small $L = 16$ is larger because the structure of time series is not captured well. For fixed L and r, the RMSE of two forecasting algorithms are close to each other with a slight dominance of SSA-V forecasting.

In Table 1 we show the RMSE for humidity at the station K. We see that the lowest RMSE is 5.343 for SSA-R forecasting and 5.365 for SSA-V forecasting and it is attained at $L = 36$ and $r = 5$ for both algorithms. However, for other values of L and r, SSA-V forecasting is usually slightly better. Overall, the RMSE is weakly depending on L and r when $L \geq 24$.

In Table 2 we display the RMSE for temperature at the station K and we see that the lowest RMSE is 0.853 attained at $L = 48$ and $r = 5$ for both SSA-R and SSA-V forecasting algorithms. We observe that the RMSE is quite small for $r = 5$ and any L because the temperature at station K has a simple sinusoidal shape. The RMSE for $r = 4$ is larger than the RMSE for $r = 6$ indicating the acurracy of forecasting is dropping faster when taking r to be smaller than the signal rank.

Table 3 contains the RMSE for humidity at the station MA. The smallest RMSE is 2.917 attained at $L = 36$ and $r = 4$ for SSA-R forecasting and 3.145 attained at $L = 48$ and $r = 6$ for SSA-V forecasting. Overall, the RMSE is quite small for $r = 4$ or $L = 48$ indicating that the signal rank is 4.

In Table 4 we present the RMSE for the temperature at the station MA. We see that the lowest RMSE is 0.561 attained at $L = 24$ and $r = 5$ for SSA-R forecasting and 0.564 attained at $L = 36$ and $r = 9$ for SSA-V forecasting. In general, the RMSE is quite small for $r = 5$ and any L showing the signal rank is likely to be 5.

In Table 5 we show the RMSE for humidity at the station TH. We observe that the lowest RMSE is 4.441 attained at $L = 36$ and $r = 15$ for SSA-R forecasting and 4.445 attained at $L = 36$ and $r = 11$ for SSA-V forecasting. We see that the RMSE is quite small for any $r \geq 5$ and $L \in \{24, 36\}$. Note that humidity at the station TH is the most volatile in our study.

In Table 6 we present the RMSE for the temperature at the station TH and we observe that the lowest RMSE is 0.561 attained at $L = 24$ and $r = 5$ for the SSA-R forecasting and 0.564 attained at $L = 36$ and $r = 9$ for the SSA-V forecasting. Overall, the RMSE is small for $r = 7$ and $L \geq 24$.

In Figures 2 and 3 we depict humidity and temperature at three stations together with $1, 2, \ldots, 12$-month ahead SSA-R and SSA-V forecasts from 12 truncations points. We can observe that SSA-R and SSA-V forecasts are very stable and close each other. Note that humidity is quite volatile and therefore forecasts are not close to the observed values. In contrast, temperature has the clear annual pattern and consequently SSA forecasts are much more accurate.

Table 1. The RMSE of $1, 2, \ldots, 12$*-month ahead forecasts using SSA-R and SSA-V forecasting algorithms for humidity at the station K.*

	$L = 16$		$L = 24$		$L = 36$		$L = 48$	
r	SSA-R	SSA-V	SSA-R	SSA-V	SSA-R	SSA-V	SSA-R	SSA-V
4	6.382	7.013	6.808	6.780	6.132	5.950	6.354	6.177
5	6.395	7.071	6.210	6.284	**5.343**	**5.365**	5.590	5.386
6	9.346	9.819	7.800	8.003	6.859	6.415	5.861	5.898
7	10.057	7.987	6.026	5.831	6.642	6.652	6.651	6.334
8	11.924	9.461	5.890	5.934	5.629	6.032	6.986	6.460
9	12.086	10.538	6.145	6.396	6.683	6.142	6.687	6.276
10	11.874	10.375	7.532	6.313	6.996	6.126	6.662	6.098
11	15.015	11.339	6.503	5.980	7.115	6.153	6.622	6.092
12	14.935	10.494	6.888	6.085	6.987	6.114	6.721	6.019
13	12.106	10.279	7.293	5.920	6.641	6.142	6.741	5.983
14	37.023	20.195	7.753	6.390	6.762	6.822	6.772	6.631
15			8.228	6.324	6.671	6.722	6.669	6.549

Table 2. The RMSE of $1, 2, \ldots, 12$*-month ahead forecasts using SSA-R and SSA-V forecasting algorithms for temperature at the station K.*

	$L = 16$		$L = 24$		$L = 36$		$L = 48$	
r	SSA-R	SSA-V	SSA-R	SSA-V	SSA-R	SSA-V	SSA-R	SSA-V
4	1.026	1.240	1.250	1.273	1.206	1.297	1.170	1.243
5	0.882	0.972	0.948	0.991	0.867	0.963	**0.853**	**0.853**
6	1.064	1.082	0.931	1.016	0.875	0.939	0.864	0.950
7	1.198	1.131	1.263	1.079	1.145	1.126	1.069	1.076
8	1.378	1.154	1.373	1.066	1.179	1.142	1.103	1.069
9	2.095	1.206	1.337	1.136	1.240	1.154	1.028	1.065
10	2.869	1.161	1.399	1.132	1.315	1.159	1.037	1.066
11	3.225	1.449	1.482	1.214	1.359	1.151	1.024	1.000
12	4.106	1.564	1.500	1.225	1.416	1.154	1.024	1.003
13	4.916	1.998	1.544	1.221	1.434	1.150	1.077	1.042
14	7.291	10.167	1.722	1.144	1.477	1.175	1.104	1.029
15			1.804	1.218	1.557	1.190	1.133	1.009

Table 3. The RMSE of $1, 2, \ldots, 12$*-month ahead forecasts using SSA-R and SSA-V forecasting algorithms for humidity at the station MA.*

	$L = 16$		$L = 24$		$L = 36$		$L = 48$	
r	SSA-R	SSA-V	SSA-R	SSA-V	SSA-R	SSA-V	SSA-R	SSA-V
4	3.121	3.828	3.398	3.417	**2.917**	3.351	2.962	3.337
5	3.636	3.837	3.213	3.332	5.343	5.365	3.171	3.205
6	3.895	3.872	3.548	3.234	6.859	6.415	3.258	**3.145**
7	4.648	4.464	4.110	4.016	6.642	6.652	3.620	3.286
8	4.687	4.343	4.089	3.849	5.629	6.032	3.560	3.192
9	5.105	4.840	4.195	3.907	6.683	6.142	3.564	3.265
10	5.200	4.774	4.130	3.908	6.996	6.126	3.654	3.269
11	5.611	4.326	4.541	3.748	7.115	6.153	3.809	3.217
12	8.314	5.810	5.636	4.237	6.987	6.114	4.078	3.201
13	9.686	5.576	7.154	4.300	6.641	6.142	4.256	3.189
14	11.565	6.998	7.725	4.355	6.762	6.822	4.415	3.237
15			8.331	4.452	6.671	6.722	4.678	3.372

Table 4. The RMSE of $1, 2, \ldots, 12$*-month ahead forecasts using SSA-R and SSA-V forecasting algorithms for temperature at the station MA.*

	$L = 16$		$L = 24$		$L = 36$		$L = 48$	
r	SSA-R	SSA-V	SSA-R	SSA-V	SSA-R	SSA-V	SSA-R	SSA-V
4	1.039	1.953	1.512	1.599	1.412	1.608	1.402	1.643
5	0.632	0.732	**0.561**	0.614	0.564	0.595	0.592	0.572
6	0.811	0.886	0.567	0.644	0.609	0.596	0.731	0.731
7	0.821	0.828	0.957	0.992	0.578	0.584	0.841	0.741
8	0.827	0.925	0.915	0.910	0.563	0.584	0.814	0.722
9	1.037	0.961	0.888	1.007	0.570	**0.564**	0.821	0.713
10	0.980	0.921	0.770	0.788	0.643	0.763	0.853	0.747
11	0.905	0.874	0.796	0.751	0.885	0.725	0.882	0.759
12	0.935	0.859	0.864	0.850	1.122	0.724	0.913	0.797
13	1.013	0.893	0.892	0.851	1.129	0.712	1.013	0.839
14	1.055	0.910	0.922	0.916	1.108	0.672	1.042	0.866
15	1.517	1.327	1.175	0.862	1.123	0.665	1.062	0.886

Table 5. The RMSE of $1, 2, \ldots, 12$*-month ahead forecasts using SSA-R and SSA-V forecasting algorithms for humidity at the station TH.*

	$L = 16$		$L = 24$		$L = 36$		$L = 48$	
r	SSA-R	SSA-V	SSA-R	SSA-V	SSA-R	SSA-V	SSA-R	SSA-V
4	6.979	9.040	6.383	8.319	6.744	8.214	7.077	8.217
5	6.099	6.671	4.529	4.871	4.582	4.486	4.979	5.009
6	6.223	6.490	4.871	4.937	4.875	4.583	5.638	5.256
7	5.972	6.715	4.897	4.854	4.767	4.531	5.383	5.365
8	5.996	6.236	4.943	5.203	4.818	4.888	5.372	5.415
9	5.799	6.247	4.881	5.091	4.810	4.715	5.422	5.278
10	5.873	6.527	4.742	4.886	4.537	4.442	5.256	5.165
11	6.244	6.323	4.769	4.863	4.553	**4.441**	5.215	5.123
12	6.680	7.439	4.797	4.791	4.603	4.498	5.301	5.265
13	8.299	6.245	4.869	4.792	4.495	4.555	5.331	5.318
14	8.953	8.393	5.189	4.842	4.484	4.562	5.357	5.344
15			5.216	4.891	**4.445**	4.529	5.333	5.255

Table 6. The RMSE of $1, 2, \ldots, 12$*-month ahead forecasts using SSA-R and SSA-V forecasting algorithms for temperature at the station TH.*

	$L = 16$		$L = 24$		$L = 36$		$L = 48$	
r	SSA-R	SSA-V	SSA-R	SSA-V	SSA-R	SSA-V	SSA-R	SSA-V
4	1.494	2.110	1.813	2.059	1.777	2.018	1.777	2.018
5	1.206	1.278	1.193	1.233	1.206	1.236	1.258	1.182
6	1.315	1.285	1.266	1.234	1.175	1.245	1.235	1.191
7	1.373	1.403	1.247	1.292	**1.149**	1.212	1.190	1.114
8	1.422	1.377	1.341	1.277	1.180	1.205	1.207	1.107
9	1.517	1.409	1.533	1.272	1.277	1.190	1.295	**1.092**
10	1.597	1.388	1.575	1.271	1.353	1.202	1.325	1.104
11	1.885	1.435	1.639	1.422	1.448	1.288	1.312	1.148
12	1.862	1.765	1.612	1.414	1.479	1.339	1.372	1.225
13	5.358	1.799	1.713	1.419	1.392	1.297	1.473	1.354
14	4.076	1.901	1.988	1.542	1.434	1.297	1.465	1.311
15	3.416	3.413	2.025	1.455	1.490	1.284	1.468	1.292

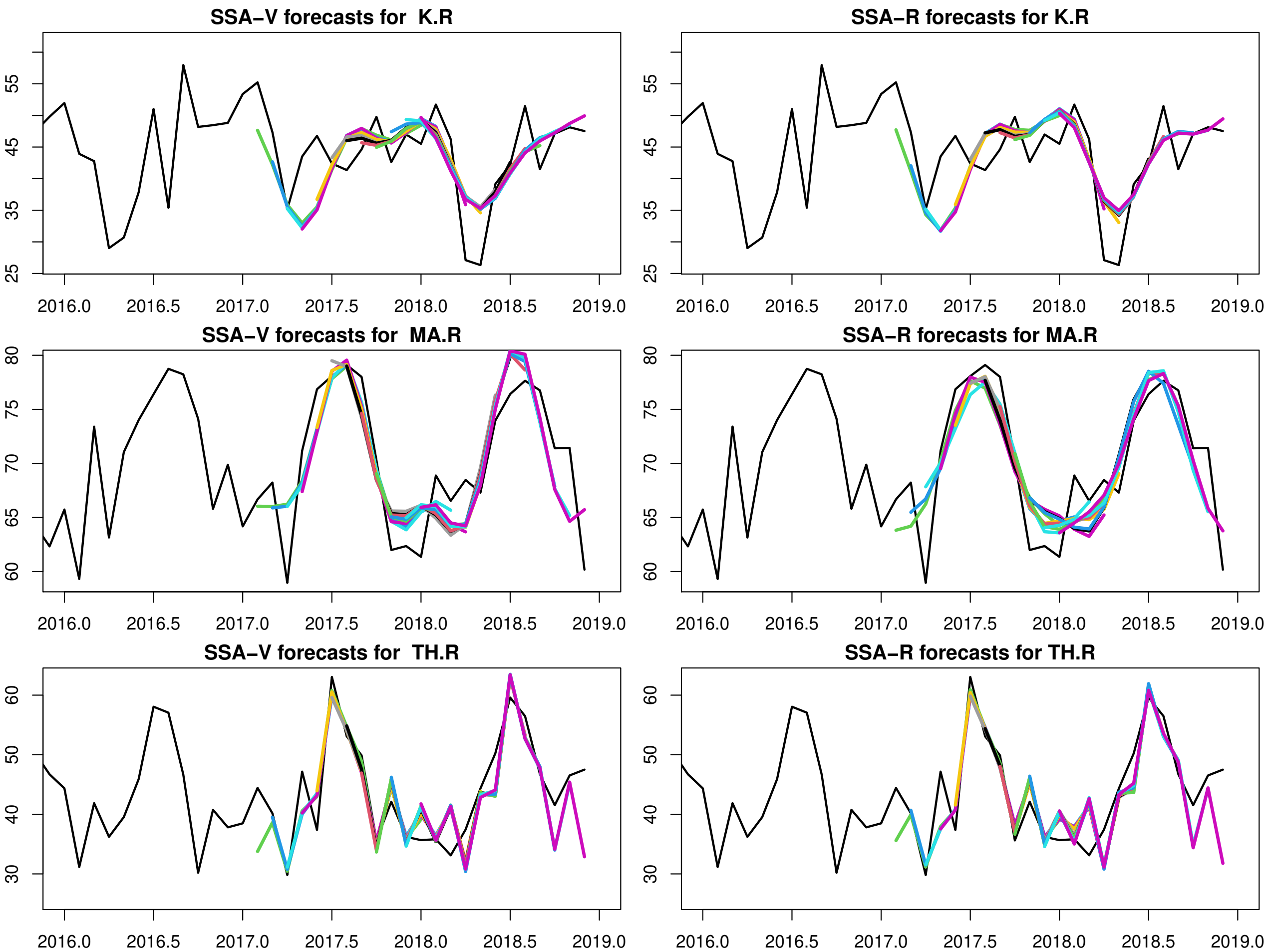

Figure 2. Humidity (black) at the stations K, MA and TH with $1, 2, \ldots, 12$-month ahead SSA-R and SSA-V forecasts (colored) with parameters L and r providing the smallest RMSE.

Table 7. The automatic choice of parameters L and r based on the $\mathrm{RMSE}_{Jan2016}^{Dec2016}$ for 1,2,. . . ,12-month ahead forecasts of humidity and temperature for the stations K, MA, and TH, the $\mathrm{RMSE}_{Jan2017}^{Dec2017}$ with chosen parameters and its efficiency. The last column contains the $\mathrm{RMSE}_{Jan2017}^{Dec2017}$ for ARIMA forecasting with automatic parameters.

		SSA-R					SSA-V					ARIMA
Series	Station	L	r	$\mathrm{RMSE}_{Jan2016}^{Dec2016}$	$\mathrm{RMSE}_{Jan2017}^{Dec2017}$	Eff	L	r	$\mathrm{RMSE}_{Jan2016}^{Dec2016}$	$\mathrm{RMSE}_{Jan2017}^{Dec2017}$	Eff	$\mathrm{RMSE}_{Jan2017}^{Dec2017}$
Hum.	K	36	5	6.441	5.479	1	36	13	5.906	6.192	0.875	5.783
Temp.	K	36	7	1.001	1.163	0.77	24	8	0.981	1.110	0.876	1.485
Hum.	MA	24	6	3.452	3.501	0.846	24	9	3.320	3.831	0.825	3.909
Temp.	MA	24	5	0.758	0.555	1	24	5	0.702	0.603	0.976	1.540
Hum.	TH	36	10	6.169	4.351	0.982	24	13	5.786	4.666	0.914	5.253
Temp.	TH	36	5	1.115	1.218	0.946	24	6	1.112	1.250	0.944	2.091

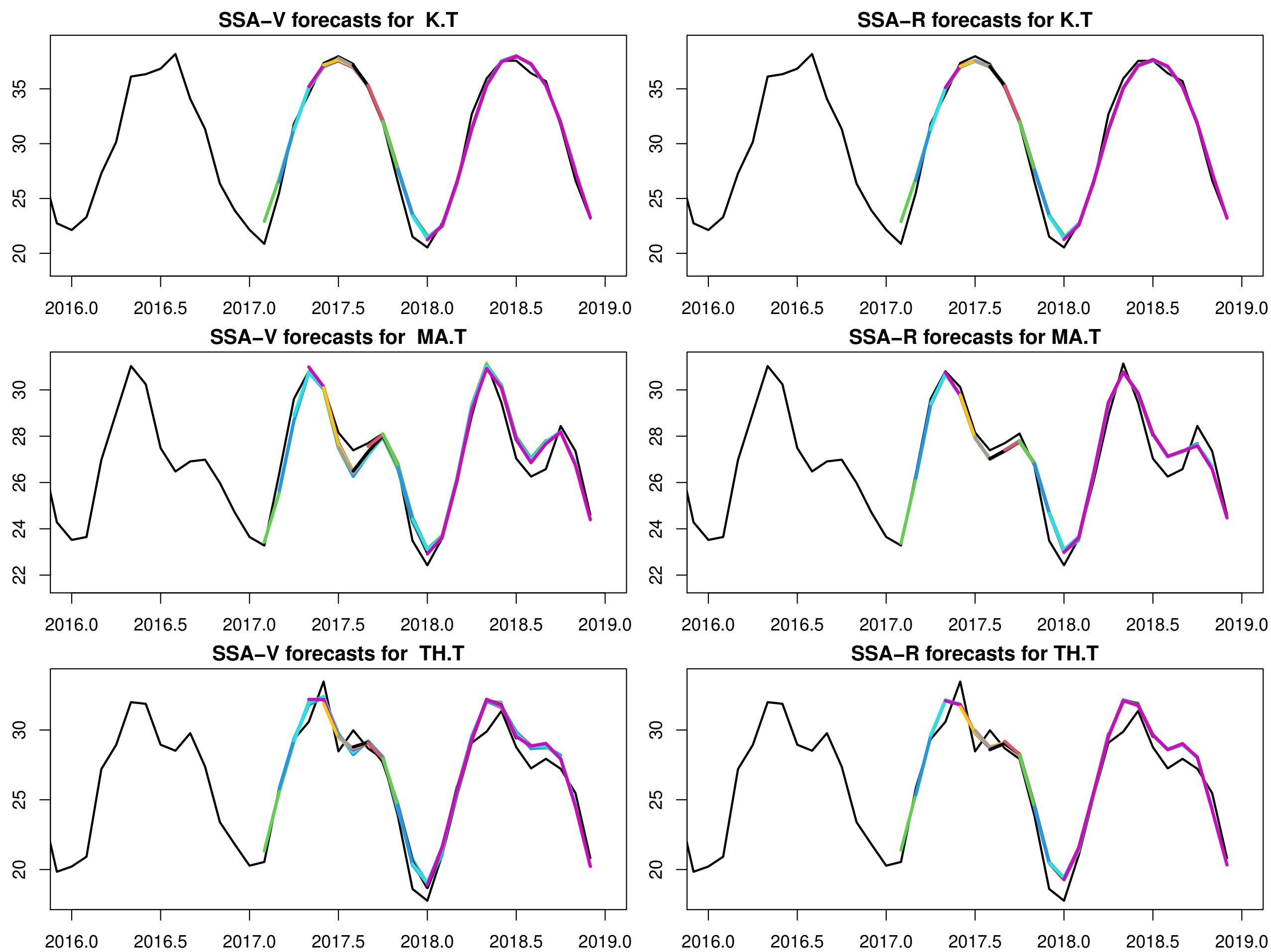

Figure 3. Temperature (black) at the stations K, MA and TH with $1, 2, \ldots, 12$-month ahead SSA-R and SSA-V forecasts (colored) with parameters L and r providing the smallest RMSE.

5. STUDY OF THE AUTOMATIC CHOICE OF PARAMETERS

For a proper assessment of the accuracy of forecasts with the automatic choice of parameters, we consider the time series from Jan 2018 to Dec 2018 as future values and perform the automatic choice of parameters for time series from Jan 2009 to Dec 2017. This automatic choice is based on minimizing the $\mathrm{RMSE}_{Jan2016}^{Dec2016}$ for $1, 2, \ldots, 12$-month ahead forecasts with truncation points from Jan 2016 to Dec 2016 and reported in Table 7. Also we compute the accuracy of future forecasts by means of the $\mathrm{RMSE}_{Jan2017}^{Dec2017}$ with chosen parameters and the efficiency of future forecasts as the ratio of the smallest $\mathrm{RMSE}_{Jan2017}^{Dec2017}$ across different parameters to the $\mathrm{RMSE}_{Jan2017}^{Dec2017}$ with chosen parameters.

We can see that the efficiency of forecasting with the automatic choice is around 90%. Overall, SSA-V forecasts are slightly more accurate than SSA-R forecasts with the automatic choice. We can observe that the $\mathrm{RMSE}_{Jan2017}^{Dec2017}$ is close to the $\mathrm{RMSE}_{Jan2016}^{Dec2016}$ confirming that the structure of time series has not changed much.

In Table 7 we also presented the RMSE for ARIMA forecasting with automatic parameters implemented in the function `auto.arima` from the R package `forecast`. Specifically, we used the stationary seasonal ARIMA model for fitting and forecasting our monthly data. We can see that SSA-R and SSA-V forecasting is better than ARIMA forecasting of humidity at the station TH and temperature at all three stations.

6. CONCLUSION

In this paper, by using real data representing monthly temperature and humidity in Oman, we have provided a statistical framework for studying which SSA forecasting algorithm is best. We demonstrated that the sensitivity of the RMSE for retrospective forecasts is rather small to the parameters L and r. We shown that the efficiency of SSA forecasts with the automatic choice of parameters is rather high. We also found that SSA-R and SSA-V forecasts are similar to each other in terms of the RMSE with a slight dominance of SSA-V forecasts. We believe that the findings presented in this paper will increase the confidence of researchers to recognise and apply the SSA forecasting algorithms with the automatic choice of parameters.

ACKNOWLEDGMENTS

We thank the Directorate General of Meteorology in the Sultanate of Oman for providing raw meteorological data. The work of A. Pepelyshev was partially supported by the Russian Foundation for Basic Research (project no. 20-01-00096).

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

Received 3 January 2021

REFERENCES

[1] AL Marhoobi, S. and Pepelyshev, A. (2021). Analysis of temperature and humidity in Oman using singular spectrum analysis. *Communications in Statistics-Simulation and Computation* 1–14.

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

[3] Golyandina, N. (2010). On the choice of parameters in singular spectrum analysis and related subspace-based methods. *Statistics and Its Interface* **3** 259–279.

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

[5] Golyandina, N., Korobeynikov, A. and Zhigljavsky, A. (2018). *Singular spectrum analysis with R.* Springer.

[6] Golyandina, N., Nekrutkin, V. and Zhigljavsky, A. A. (2001). *Analysis of time series structure: SSA and related techniques.* CRC press.

[7] Golyandina, N. and Shlemov, A. (2013). Variations of singular spectrum analysis for separability improvement: non-orthogonal decompositions of time series. *arXiv preprint arXiv:1308.4022.*

[8] Golyandina, N. and Zhigljavsky, A. (2020). *Singular Spectrum Analysis for time series.* Springer.

[9] Harmouche, J., Fourer, D., Auger, F., Borgnat, P. and Flandrin, P. (2017). The sliding singular spectrum analysis: A data-driven nonstationary signal decomposition tool. *IEEE Transactions on Signal Processing* **66** 251–263.

[10] 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** 55–83.

[11] Hassani, H. and Mahmoudvand, R. (2018). *Singular Spectrum Analysis: Using R.* Springer.

[12] Hassani, H., Mahmoudvand, R., Zokaei, M. and Ghodsi, M. (2012). On the separability between signal and noise in singular spectrum analysis. *Fluctuation and Noise Letters* **11** 1250014.

[13] Hassani, H., Silva, E. S., Gupta, R. and Segnon, M. K. (2015). Forecasting the price of gold. *Applied Economics* **47** 4141–4152.

[14] Mahmoudvand, R., Najari, N. and Zokaei, M. (2013). On the optimal parameters for reconstruction and forecasting in singular spectrum analysis. *Communications in Statistics-Simulation and Computation* **42** 860–870.

[15] Mahmoudvand, R., Rodrigues, P. C. and Yarmohammadi, M. (2019). Forecasting daily exchange rates: A comparison between SSA and MSSA. *REVSTAT* **17** 599-616.

[16] Pepelyshev, A. and Zhigljavsky, A. (2017). SSA analysis and forecasting of records for Earth temperature and ice extents. *Statistics and Its Interface* **10** 151–163.

[17] Plaza, E. G. and López, P. N. (2017). Surface roughness monitoring by singular spectrum analysis of vibration signals. *Mechanical Systems and Signal Processing* **84** 516–530.

[18] Wang, W. and Lu, Y. (2018). Analysis of the mean absolute error (MAE) and the root mean square error (RMSE) in assessing rounding model. In *IOP Conference Series: Materials Science and Engineering* **324** 012049.

[19] 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.

Safia ALMarhoobi
Ministry of Higher Education, Research and Innovation
Oman
and
School of Mathematics
Cardiff University
Senghennydd Road
Cardiff, CF24 4AG, UK
E-mail address: safiaalmarhoobi@gmail.com

Andrey Pepelyshev
School of Mathematics
Cardiff University
Senghennydd Road
Cardiff, CF24 4AG, UK
E-mail address: pepelyshevan@cardiff.ac.uk