REVSTAT – Statistical Journal
Volume 11, Number 1, March 2013, 105–119

IMPROVING SSA PREDICTIONS BY INVERSE DISTANCE WEIGHTING

Authors: Richard O. Awichi
– Uganda Martyrs University, Nkozi, Uganda
ichbinrao@gmail.com

Werner G. Müller
– Johannes Kepler University, Linz, Austria
werner.mueller@jku.at

Abstract:

- This paper proposes a method of utilizing spatial information to improve predictions in one dimensional time series analysis using singular spectrum analysis (SSA). It employs inverse distance weighting for spatial averaging and subsequently multivariate singular spectrum analysis (MSSA) for enhanced forecasts. The technique is exemplified on a data set for rainfall recordings from Upper Austria.

Key-Words:

- *singular spectrum analysis; inverse distance weighting; spatio-temporal predictions.*

AMS Subject Classification:

- 49A05, 78B26.

1. INTRODUCTION

Singular spectrum analysis (SSA) is a recently popularized tool for time series analysis, cf. [10]. The origins of SSA can be traced to [2, 4, 6]. More information about the history of SSA can be found in [22]. It is a model free approach to time series analysis and literally any time series with a notable structure can be analysed using SSA. Indeed it has a wide area of applications ranging from mathematics and physics [10], to economics and financial mathematics [13, 14], environmental sciences [15], social sciences [12], and medicine [7]. It is now implemented under various software platforms, here we use Rssa, see [9] and a program called CaterpilarSSA as can be downloaded from `http://www.gistatgroup.com/cat/programs.html`. The aim of SSA is twofold:

i) To make a decomposition of the original series into a sum of a small number of independent and interpretable components such as a slowly varying trend, oscillatory components and a structure less noise;

ii) To reconstruct the decomposed series so as to make predictions without the noise component.

MSSA is an extension of SSA and takes advantage of the (delay) embedding procedure to obtain a similar formulation as SSA, albeit with larger matrices for multidimensional time series. It has previously been successfully applied to the study of climate fields, see [18]. Here we will employ it to jointly model an original time series with a spatial average of which we believe will improve predictions by pooling spatially dependent information.

One of the simplest but effective ways of generating spatial averages is inverse distance weighting, which was first introduced, incidentally also for the analysis of rainfall data in [11]. It was subsequently propagated in [20] and became thereafter one of the most popular spatial interpolation techniques (cf. eg. [16]).

Section 2 is devoted to reviewing the basics of SSA. Section 3 discusses forecasting, while Section 4 briefly presents MSSA, an extension of the SSA techniques to multivariate data and introduces a method of incorporating spatial dependence to improve forecasts. The application is presented in Section 5 and conclusions appear in Section 6.

2. SINGULAR SPECTRUM ANALYSIS

Most classical time series models devised for analysis and forecasting are based on restrictive assumptions of normality, linearity and stationarity, cf. [3].

A number of time series are deterministic, linear and dynamical systems thus allowing linear models to be used for modelling and forecasting. However, many time series exhibit nonlinear behaviour and therefore would require a method that works well for both linear and nonlinear, stationary and nonstationary data sets. SSA is one such technique.

2.1. A brief review of SSA

The Basic SSA, as it is commonly referred to, has two main stages: Decomposition and Reconstruction; each of which consists of two steps as described below. The main concept in SSA is the aspect of separability of the original time series into signal and noise so that the analysis and forecasting can be done on signal in the absence of noise. Separability will be mentioned again later. In the following discussion, we follow the approach in [10, Chapter 1].

Let $F_N = \{f_1, f_2, \ldots, f_N\}$ be a real valued, nonzero (at least one $f_i \neq 0$) time series data of sufficient length N without missing values.

Stage 1: *Decomposition*
Step 1: Embedding

This (standard) time series procedure maps the one dimensional time series, F_N into multidimensional lagged vectors, $X_1 : \cdots : X_K$, where

$$X_i = (f_1, \ldots, f_{i+L-1})^T \in R^L\,, \qquad 1 \le i \le K \quad \text{and} \quad K = N - L + 1\ .$$

The single most important parameter of embedding is the window length, L, an integer such that $2 < L < N$. This parameter should always be large enough to permit reasonable separability. It should not be greater the $N/2$ for optimum results. See [8] for more on the choice of parameters for SSA. The vectors X_i, called the lagged vectors or L lagged vectors (to emphasize their dimension) form the K columns of the trajectory matrix X, i.e. $X = [X_1 : \cdots : X_K]$.

Specifically X is given as follows:

$$X = \begin{pmatrix} f_1 & f_2 & f_3 & \cdots & f_K \\ f_2 & f_3 & f_4 & \cdots & f_{K+1} \\ f_3 & f_4 & f_5 & \cdots & f_{K+2} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ f_L & f_{L+1} & f_{L+2} & \cdots & f_N \end{pmatrix}.$$

The $L \times K$ matrix X is a Hankel matrix, i.e. the elements along the anti-diagonal, $i + j =$ constant are equal, for the i^{th} row and j^{th} column.

Step 2: Singular Value Decomposition, SVD

This decomposes the trajectory matrix X and represents it as a sum of elementary matrices (rank-one bi-orthogonal). This is done by:

i) Calculating the matrix $S = XX^T$.

ii) Obtaining eigenvalues, λ_i of S such that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_L \geq 0$. Since S is positive definite, the eigenvalues are positive.

iii) For each λ_i, calculate U_i and V_i, the left and right singular vectors of X. The U_is are orthonormal system of eigenvectors corresponding to each λ_i such that $\langle U_i, U_j \rangle = 0$, $i \neq j$ (orthogonality) and $\|U_i\| = 1$ (unit norm property) and $V_i = X^T U_i / \sqrt{\lambda_i}$.

iv) Set $d = \max(i \colon \lambda_i > 0) = \operatorname{rank}(X)$. Then $X_i = \sqrt{\lambda_i}\, U_i V_i^T$ $(i = 1, \ldots, d)$, and the SVD of the trajectory matrix represents it as a sum of the X_i, i.e.:

$$X = \sum_{i=1}^{d} X_i = X_1 + X_2 + \cdots + X_d \; . \tag{2.1}$$

The collection $(\sqrt{\lambda_i}, U_i, V_i)$ is called the i^{th} eigentriple of X, $\sqrt{\lambda_i}$ are the singular values of X and the set $\{\sqrt{\lambda_i}\}_{i=1}^{d}$ is the spectrum of X.

The ratio $\lambda_i / \sum_{i=1}^{d} \lambda_i$ is the characteristic contribution (or its share) of X_i to (2.1). The first eigenvalue has the largest contribution and the last has the smallest.

If all the eigenvalues have multiplicity one, then (2.1) is uniquely determined.

Stage 2: *Reconstruction*

Step 3: Grouping

This corresponds to splitting the elementary matrices X_i into several groups and summing the matrices within each group. If $I = i_1, \ldots, i_p$ be one such group, then the matrix X_I corresponding to the group I is defined as:

$$X_I = X_{i_1} + \cdots + X_{i_p} \; .$$

For m such groups (disjoint), then X will be given as:

$$X = X_{I_1} + \cdots + X_{I_m} \; . \tag{2.2}$$

Matrices X_{I_i} are called *resultant matrices* and the procedure of choosing the sets $I_1, \ldots, I_m$ is called the *eigentriple grouping*.

The contribution of component X_I in (2.2) is measured by the share of the corresponding eigenvalues, i.e. $\sum_{i \in I} \lambda_i / \sum_{i=1}^{d} \lambda_i$.

Step 4: Diagonal Averaging

This (last) step transfers each resultant matrix into a time series, which is an additive component of the initial (original) series, F_N. If z_{ij} stands for an element of a matrix Z, then the k^{th} term of the resulting series is obtained by averaging z_{ij} over all i, j such that $i + j = k + 2$ ([10, page 17, 24], [12, page 242]). This is diagonal averaging or Hankelization of the matrix Z. The result of the Hankelization of a matrix Z is the matrix $\mathcal{H}Z$. Diagonal averaging is a linear operation and maps the trajectory matrix of the initial series into the original series itself, i.e. it transfers each matrix I into a time series which is an additive component of the initial series F_N.

2.2. Separability

As mentioned earlier, the main concept in studying SSA properties is separability. This entails how well the components of the time series can be separated from each other to allow forecasting to be meaningfully done and also reliable construction of confidence bounds. Any time series may comprise trend (slowly varying component), periodic or quasi periodic components (like seasonal variations or harmonics generally) and noise. These may be generalized into signal and noise components. SSA decomposition of the series F_N can only be successful if the resulting additive components of the series are approximately separable from each other, [10, 12].

If a time series F_N can be split as $F_N = F_N^{(1)} + F_N^{(2)}$, then the matrix terms of the SVD step can be split into $X^{(1)}$ and $X^{(2)}$ respectively, i.e. $X = X^{(1)} + X^{(2)}$. This would imply that each row of $X^{(1)}$ is orthogonal to each row of $X^{(2)}$. Since rows (and columns) of the trajectory matrix X are themselves subseries of the initial series, the orthogonality condition of the rows of $X^{(1)}$ and $X^{(2)}$ is the condition of orthogonality of any subseries of length L and $K = N - L + 1$ of the series $F_N^{(1)}$ to any subseries of the same length, $F_N^{(2)}$. If this holds, then $F_N^{(1)}$ and $F_N^{(2)}$ are said to be weakly separable.

In geometrical terms, $F_N^{(1)}$ and $F_N^{(2)}$ are separable if and only if the subspace $\ell^{(L,1)}$ spanned by the columns of $X^{(1)}$ is orthogonal to the subspace $\ell^{(L,2)}$ spanned by the columns of $X^{(2)}$. One way to enhance separability of the series is auxiliary information about the series to help in choosing the window length, for example, if it is known that there is a seasonal component whose period is an integer, it is advisable to choose the window length which is a factor of the period, [10, page 44]. To choose eigentriples, one may use the graph of the logarithms of eigenvalues in which explicit plateau in the eigenvalue spectra prompts ordinal numbers of the eigentriple and a slowly decreasing sequence of singular values corresponds to noise components.

Another way to measure the separability between two series components, $F_N^{(1)}$ and $F_N^{(2)}$ (i.e. if $F_N = F_N^{(1)} + F_N^{(2)}$) is to calculate the weighted correlation or w-correlations between the two using the formula

$$\rho_{12}^w = \frac{\langle F_N^{(1)}, F_N^{(2)} \rangle_w}{\|F_N^{(1)}\|_w \, \|F_N^{(2)}\|_w} \ ,$$

where $\|F_N^{(i)}\|_w = \sqrt{\langle F_N^{(i)}, F_N^{(i)} \rangle_w}$, $i = 1, 2$, $\langle F_N^{(1)}, F_N^{(2)} \rangle_w = \sum_{i=0}^{N-1} w_i f_i^{(1)} f_i^{(2)}$, and the weights w_i defined as follows:

Let $L^\star = \min(L, K)$ and $K^\star = \max(L, K)$. Then,

$$w_i = \begin{cases} i+1 & \text{for } \ 0 \leqslant i \leqslant L^\star - 1 \ , \\ L^\star & \text{for } \ L^\star \leqslant i \leqslant K^\star \ , \\ N - i & \text{for } \ K^\star \leqslant i \leqslant N-1 \ . \end{cases}$$

A natural hint for grouping is the matrix of the absolute values of the w-correlations corresponding to a full decomposition. If the absolute value of the w-correlation is small then the corresponding series are almost w-orthogonal and is said to be weakly separable. The series $F^{(1)}$ and $F^{(2)}$ are w-orthogonal if $\langle F_N^{(1)}, F_N^{(2)} \rangle_w = 0$, [10, 12].

Separability is analogous to independence of random variables whence the covariance and correlation between such random variables are zero, [5, Section 4.5].

3. FORECASTING WITH SSA

Details of SSA Forecasting can be found in [10, Chapter 2, 5] and in [19]. We have three basic conditions:

1) Time series has structure.

2) A mechanism identifying this structure is found.

3) A method of time series continuation, based on the identified structure is available.

In SSA, forecasting is done through application of linear recurrent formulae (LRF) or equations. The class of series governed by LRF is rather wide; it contains harmonics, polynomials and exponential series and is closed under term-by-term addition and multiplication, [12]. An infinite series is governed by some LRF if and only if it can be represented as a linear combination of products of exponential, polynomial and harmonic series. (The signal component of a separable time series is always a linear combination of these series.)

An important property of SSA decomposition is that the original series satisfies an LRF of the form $f_n = a_1 f_{n-1} + \cdots + a_d f_{n-d}$ for some dimension d; $a_1, \ldots, a_d$ are constants.

Thus for any N and L, there are at most d nonzero singular values in the SVD of the trajectory matrix X and so even if L and $K = N - L + 1$ are larger than d, we need at most d matrices X_i to reconstruct the series. If f_n satisfies the LRF above, it will always be represented as a sum of products of exponentials, polynomials and harmonics, [10].

Alternatively put, if $r < L$, (r = number of terms in the SVD step), then the series satisfies some LRF of some dimension $d \leqslant r$. This result also implies that if $\dim(\ell_r) < L$, then the series satisfies a natural LRF of dimension $L - 1$. Any such series satisfying an LRF can then be forecast for an arbitrary number of steps using the LRF.

The selection of the resultant matrices in the third step of Basic SSA algorithm implies selection of the r-dimensional space $\ell_r \in R^L$ spanned by the corresponding left singular vectors and if ℓ_r is non-vertical, it produces an appropriate LRF which can be used in forecasting.

An LRF that governs a series with the help of SSA may be found as follows. Let d be minimal dimension of all LRFs governing a time series F_N. If the window length L is greater than d and N is large enough, then the trajectory space of F_N is d dimensional. The trajectory space determines an LRF of dimension $L - 1$ that governs the time series. If this LRF is applied to the last terms of the series, a forecast of the series is obtained. The same idea works for an additive component $F_N^{(1)}$ of F_N. The assumption here is that $F_N^{(1)}$ is (strongly) separable from the residual $F_N^{(2)} = F_N - F_N^{(1)}$ for the selected window length L. Normally (strong) separability of the components of a series implies that each component satisfies some LRF, [10, Chapter 6]. If $F_N^{(2)}$ is noise, then forecasting is done for $F_N^{(1)}$. Thus using a selected set of eigentriples, estimation can be performed on $F_N^{(1)}$ and its trajectory space. The basic inputs for the SSA LRF for a series F_N include the window length L, N, linear space ℓ_v which is not a vertical space and the number M of points to forecast. The linear space is used to obtain an orthonormal basis $P_1, \ldots, P_r$ used in the forecasting process.

Forecasting is also closely linked to separability of the series as mentioned above. If $F_N = F_N^{(1)} + F_N^{(2)}$, then forecasting is done for the signal $F_N^{(1)}$ in the presence of the noise component $F_N^{(2)}$ which is given as $F_N^{(2)} = F_N - F_N^{(1)}$.

[10, pages 95–107] gives an account of the forecasting algorithm and properties of LRFs.

Construction of confidence bounds can be done by either the empirical method or the bootstrap technique. The empirical confidence intervals are constructed for the entire series, which is assumed to have the same structure in the future. Bootstrap bounds are obtained for the continuation of the signal, [10, 12].

4. MSSA WITH INVERSE DISTANCE WEIGHTING

Data mining is an automated search for knowledge hidden in large collections of data set attributes. In environmental science and other areas where space-time behaviour is an important focus of investigation, it is not uncommon to have attributes whose values change with space and time and quite often, due to spillovers or unobservable variables or omitted factors. This leads to spatial dependence that subsequently influence data analysis.

In light of spatial dependence, an inverse distance weighting technique, see [1, 20], is proposed as a means of incorporating spatial information to improve the prediction. We construct an additional explanatory variable by taking spatially weighted averages

$$\bar{y}_t = \sum_{i=1}^{n} w_i \, y_{it} \qquad \text{where} \quad w_i = \frac{1/d_i}{\sum\limits_{i=1}^{n} (1/d_i)} \ ,$$

with d_i denoting distances between the target location and the i^{th} measurement site.

Multivariate (or multichannel) Singular Spectrum Analysis (MSSA) is an extension of SSA to multidimensional data.

Assume that $y_j = \left(y_j^{(1)}, \ldots, y_j^{(m)}\right)$ is an m-variate time series, L the window length, $X^{(i)}$ $(i = 1, \ldots, m)$ the trajectory matrices of the one dimensional time series $\{y_j^{(i)}\}$ $(i = 1, \ldots, m)$, the trajectory matrix X of the multivariate series is given as $X = (X^{(1)}, \ldots, X^{(r)}, \ldots, X^{(m)})$; [10, 17]. Note that X is now an $L \times mK$ block Hankel matrix (there are m blocks of $X^{(i)}$ matrices).

The aims and techniques of MSSA are straightforward extensions to those of SSA and so are the algorithms. Hence we refrain from any further discussion regarding the theory of MSSA. For more details see [21, 17]. The advantage of MSSA over SSA, however, is that it automatically utilizes dependencies among the time series in the analysis. Consequently, the quality of MSSA forecasts are typically improved when the series are more strongly correlated.

The above pooling of the spatial information by inverse distance weighting leads to a new time series $\bar{y}_t$ that can be used as a kind of covariable to the Linz rainfall series to improve the predictions. We thus now employ a MSSA with the original Linz series complemented by the pooled one, i.e. $m = 2$ in this case. Of course this can be performed for all the time series, not only the Linz one and even jointly, but we will refrain from this for the sake of expository simplicity.

5. THE APPLICATION

The complete data set consisted of $N = 192$ monthly recordings of rainfall at several locations in Upper Austria for the period 1994 Jan to 2009 Dec (see Figure 1 for a depiction of the measurement locations with the solid dot indicating Linz). The data is provided by the Zentralanstalt für Meteorologie in Austria and is described in more detail in [15].

Figure 1: Upper-Austrian rainfall measurement network.
Empty circle indicate measurement locations, solid circle Linz.

The time series graph Figure 2 shows the general behaviour of the logarithm of Linz rainfall series and the reconstructed series for the period above. Since it is annual data, it provides auxiliary information for the choice of the window length as a factor of the period, 12 monthlies, hence the choice of $L = 96 = N/2$ as the standard window length for the analysis. It can be inferred from the figure that the grouping employed yielded a reconstructed series fairly close to the original hence rather reliable in-sample predictions.

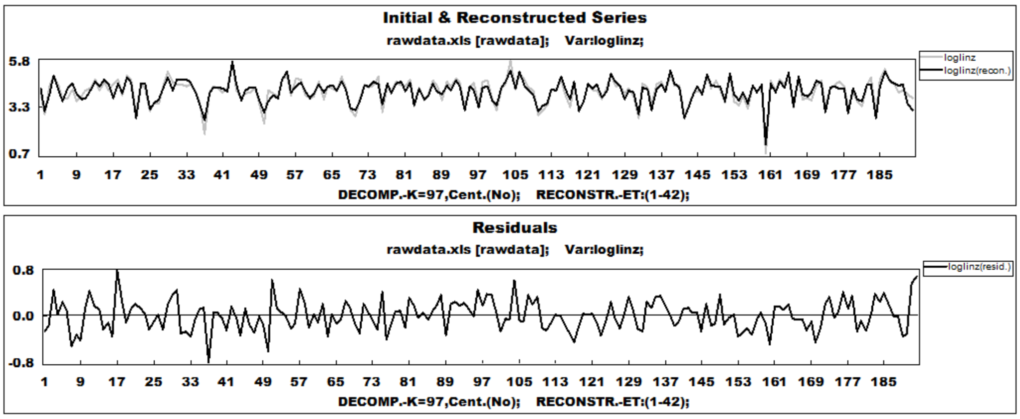

Figure 2: Initial (grey) and SSA reconstructed (black) series for Linz monthly rainfall data in logs; residuals below.

The plot in Figure 3 gives the eigenvalue graph. This graph is a plot of logarithms of the first 42 eigenvalues used in the reconstruction stage. As mentioned earlier, it shows the plateau for ordinal numbers in the eigentriple grouping. The remaining eigenvalues constitute the noise series and have not been included here. This graph shows a high percentage contribution of the first eigenvalue with a plateau for the second and third eigenvalues implying a particular type of signal. The other eigenvalues are gradually and slowly decreasing implying a strong tendency to noise after the 42^{nd} eigenvalue.

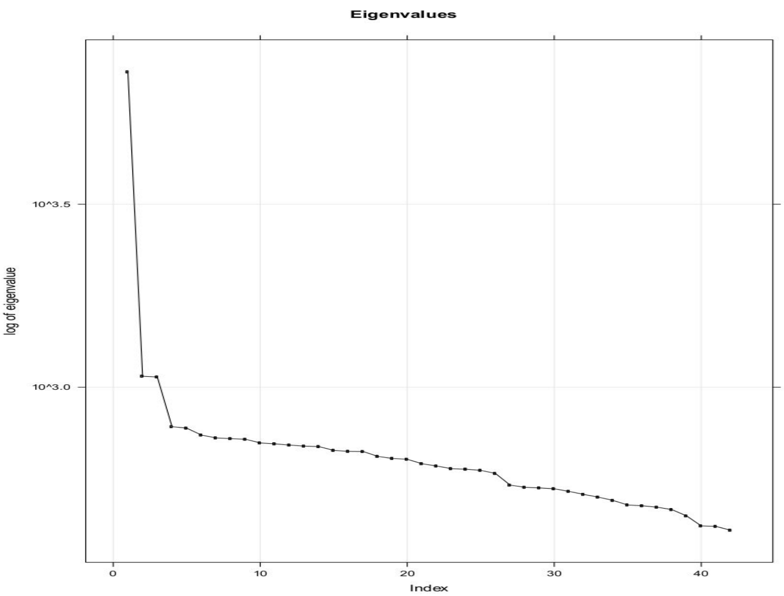

Figure 3: Eigenvalue graph for the first 42 eigenvalues used in the reconstruction stage of MSSA.

The graph in Figure 4 shows the w-correlations for the reconstructed components on a 20 grade grey scale from white to black corresponding to absolute values of correlations from 0 to 1, see [12, 10]. It shows the different eigenvalue groupings, even for the eigenvalues corresponding to the noise. This graph further illustrates the results of the grouping step and confirms the separation of signal from the noise for the original series as it clearly marks off the lags below 42. Furthermore some other possible eigentriple groupings were tried but the predictions were not better than for this particular grouping.

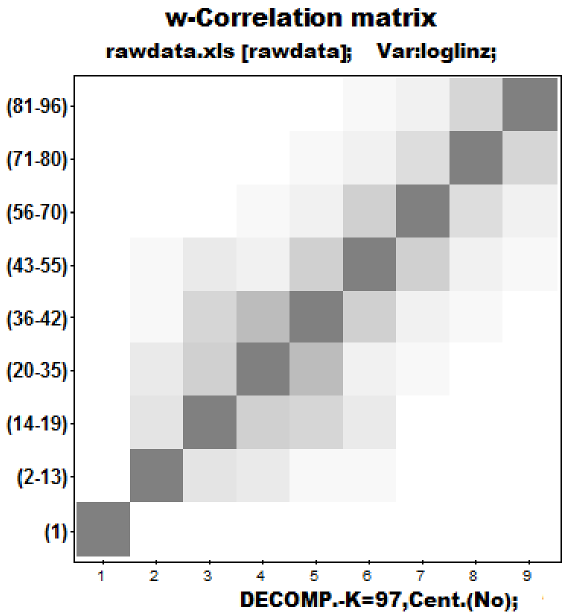

Figure 4: Matrix of w-correlations from the reconstructed components (1–42) and error (43–96) in the MSSA.

The following graphs in Figure 5 show the time series for the spatially pooled series and its effect on the Linz data series for the MSSA analysis. For the MSSA analysis we used two series, the Linz data and the inverse distance weighted average by employing the Euclidean distances d_i between Linz and each of the 36 other locations. Thus the nearer the locations to Linz, the stronger the weighting and vice versa. For missing values in the data, a new weight is calculated by excluding the corresponding distance measure from the w_is.

The in-sample SSA prediction was done with solely the Linz data to obtain the SSA prediction of Figure 1. Its root mean square error ($RMSE_{SSA}$) was found to be 0.247. The weighted average, using the inverse distance technique, was then included as a second series to study its effects, due to spatial spillover, on the Linz data for the MSSA prediction. This is shown in Figure 5. Its $RMSE_{MSSA}$ was found to be 0.245 and slightly less than $RMSE_{SSA}$. This indicates that the suggested technique of including spatial dependence in the SSA analysis may actually improve the forecasts. However, our results from other groupings show a less clear picture, particularly if not the standard window length of $L = N/2$ was used, and in further work, we want to investigate the capabilities of MSSA performing ensemble spatio-temporal predictions for the whole network of stations.

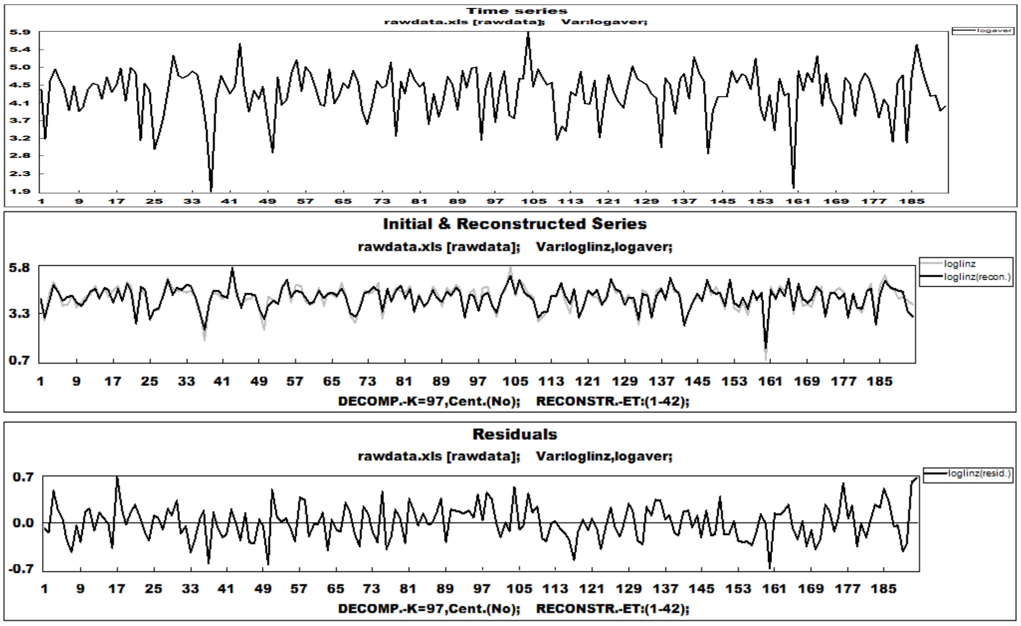

Figure 5: Time series for the inversely distance weighted data. Initial (black) and MSSA reconstructed (grey) series for Linz monthly rainfall data in logs; residuals below.

6. CONCLUSION

This short presentation illustrates the basic capabilities of SSA in separating the components of a time series and in forecasting without any assumptions about the time series data. It brings out the key advantage of the methodology of SSA in applied statistics: that of inference and prediction without specifying any particular model structure. Its extension to multidimensional data analysis, the MSSA is yet another elegant procedure to handle multidimensional data analysis without necessarily pre-specifying dependence structures. The suggested method of exploiting spatial dependence within the concept of MSSA is promising, particularly for the in-sample imputation of missing data. As mentioned earlier, we require further studies and refinements for assessing the capabilities of the technique for the out-of-sample predictions.

ACKNOWLEDGMENTS

We are grateful to the two anonymous referees for their attentive and detailed remarks which helped to considerably improve the paper.

REFERENCES

[1] Bakkali, S. and Amrani, M. (2008). About the use of spatial interpolation methods to denoising Moroccan resistivity data phosphate ”disturbances” map, *Acta Montanistica Slovaca*, **13**, 216–222.

[2] Basilevsky, A. and Hum, D.P.J. (1979). Karhunen–Loeve analysis of historical time series with an application to plantation births in Jamaica, *Journal of American Statistical Association*, **74**, 284–290.

[3] Brockwell, P.J. and Davis, R.A. (2002). *Introduction to Time Series and Forecasting*, Springer, New York.

[4] Broomhead, D.S. and King, G.P. (1986). Extracting qualitative dynamics from experimental data, *Physica D*, **20**, 217–236.

[5] Casella, G. and Berger, R.L. (2002). *Statistical Inference*, Second Edition, Duxbury / Thomson Learning, Australia, USA, Canada.

[6] Elsner, J.B. and Tsonis, A.A. (1996). *Singular Spectrum Analysis: A New Tool in Time Series Analysis*, Plenum.

[7] Ghodsi, M.; Hassani, H. and Sanei, S. (2010). Extracting fetal heart signal from noisy maternal ECG by multivariate singular spectrum analysis, *Statistics and Its Interface*, **3**, 399–411.

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

[9] Golyandina, N. and Korobeynikov, A. (2012). Basic singular spectrum analysis and forecasting with R, *arXiv:1206.6910.*

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

[11] Horton, R.E. (1923). Rainfall interpolation, *Monthly Weather Review*, **51**(6), 291–304.

[12] Hassani, H. (2007). Singular spectrum analysis: methodology and comparison, *Journal of Data Science*, **5**, 239–257.

[13] Hassani, H. and Thomakos, D. (2010). A Review on singular spectrum analysis for economic and financial time series, *Statistics and Its Interface*, **3**, 377–397.

[14] Kapl, M. and Müller, W.G. (2010). Prediction of steel prices: a comparison between a conventional regression model and MSSA, *Statistics and Its Interface*, **3**, 369–275.

[15] Mateu, J. and Müller, W.G. (Eds.) (2012). *Spatio-temporal Design: Advances in Efficient Data Acquisition; (Statistics in Practice)*, Wiley.

[16] Okabe, A. and Sugihara, K. (2012). *Spatial Analysis Along Networks: Statistical and Computational Methods; (Statistics in Practice)*, Wiley.

[17] Patterson, K.; Hassani, H.; Heravi, S. and Zhigljavsky, A. (2011). Multivariate singular spectrum analysis for forecasting revisions to real-time data, *Journal of Applied Statistics*, **38**(10), 2183–2211.

[18] Raynaud, S.; Yiou, P.; Kleeman, R. and Speich, S. (2005). Using MSSA to determine explicitly the oscillatory dynamics of weakly nonlinear climate systems, *Journal of Nonlinear Processes in Geophysics*, **12**, 807–815.

[19] Rodrigues, P.C. and de Carvalho, M. (2012). Spectral modeling of time series with missing data, *Applied Mathematical Modelling*, `http://dx.doi.org/10.1016/j.apm.2012.09.040`

[20] Shepard, D. (1968). A two-dimensional interpolation function for irregularly-spaced data, *Proceedings of the 1968 23$^{\text{rd}}$ ACM national conference*, ACM '68, ACM, New York, NY, USA, pp. 517–524.

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

[22] Zhigljavsky, A. (2010). Singular spectrum analysis for time series, *Statistics and Its Interface*, **3**, 255–258.