Multivariate Posterior Singular Spectrum Analysis

Ilkka Launonen[1] · **Lasse Holmström**[2]

Abstract A generalized, multivariate version of the Posterior Singular Spectrum Analysis (PSSA) method is described for the identification of credible features in multivariate time series. We combine Bayesian posterior modeling with multivariate SSA (MSSA) and infer the MSSA signal components with a credibility analysis of the posterior sample. The performance of multivariate PSSA (MPSSA) is compared to the single-variate PSSA with an artificial example and the potential of MPSSA is demonstrated with real data using NAO and SOI climate index series.

Ilkka Launonen
Department of Mathematical Sciences
University of Oulu, Finland
E-mail: ilkka.launonen@oulu.fi

Lasse Holmström
Department of Mathematical Sciences
University of Oulu, Finland

Multivariate Posterior Singular Spectrum Analysis

Abstract A generalized, multivariate version of the Posterior Singular Spectrum Analysis (PSSA) method is described for the identification of credible features in multivariate time series. We combine Bayesian posterior modeling with multivariate SSA (MSSA) and infer the MSSA signal components with a credibility analysis of the posterior sample. The performance of multivariate PSSA (MPSSA) is compared to the single-variate PSSA with an artificial example and the potential of MPSSA is demonstrated with real data using NAO and SOI climate index series.

Keywords Time series · SSA · Bayesian inference · multivariate · climate index

1 Introduction

Singular spectrum analysis (SSA) is a broad methodology of time series analysis whose applications include e.g. smoothing, noise reduction, extraction of trend and periodicities, missing data imputation and forecasting. SSA is an algebraic technique without any statistical model or stationarity assumptions. SSA is based on the singular value decomposition of an embedded time series and is thus related to principal component analysis. The original series, decomposed into 'eigentriples' of singular values and their associated eigenvectors, is regrouped to form various component series, which may be interpreted as a slowly-varying trend, periodic series and noise.

The origins of SSA are usually associated with the publications Broomhead and King (1986a) and Broomhead and King (1986b). Work on so-called 'Caterpillar' SSA was done independently among statisticians in St.Petersburg, where emphasis is placed on the separability of component series from each other. This is reflected in the monograph Golyandina et al. (2001) on SSA and its

––––––––––––––––––––

theoretical foundations. The recent book Golyandina and Zhigljavsky (2013) on the methodology of SSA is aimed at a wide scientific audience.

SSA can be generalized to multivariate time series and 2D-image analysis (Golyandina and Usevich, 2010). These generalizations differ from the basic single-variate SSA by using an embedding which reflects the different nature of the data matrix. In basic SSA, the embedding of a time series of length N produces an $L \times (N - L + 1)$ Hankel matrix whose columns are segments of the original series formed by sliding a window of length L along the series. In multi-channel SSA (MSSA), the embedding of the multivariate time series produces a stacked-Hankel matrix, and in 2D-SSA, a Hankel-block-Hankel matrix. SSA is a special case of MSSA, which itself is a special case of 2D-SSA with an $L \times 1$ rectangular window, if the multivariate time series are of equal length. These versions of SSA are special cases of shaped-SSA (Shlemov and Golyandina, 2014), where the data matrix can have omitted values and the embedded data is a quasi-Hankel matrix. The sums of the elementary matrices corresponding to the grouped eigentriples of these embedded Hankel-type matrices may not have a Hankel structure, and thus require a re-Hankelization prior to transforming them into time series or image components via an inverse embedding. The generalized SSA algorithms are procedurally the same as the basic SSA, except that the embedding and the re-Hankelization steps deal with generalized Hankel-type matrices. For more information, see Golyandina et al. (2015).

SSA is not well-suited to separation of autoregressive red noise. The Monte Carlo SSA (Allen and Smith, 1996) (MC-SSA) and its multivariate generalization (Allen and Robertson, 1996) remedy this with a statistical test of the eigentriples against a null hypothesis of colored noise by sampling 'surrogate' data and projecting them onto the eigentriples. As a Bayesian alternative to MC-SSA, Holmström and Launonen (2013) proposed the Posterior SSA (PSSA) and it was subsequently applied to paleoclimate reconstruction from varved lake sediment cores in Ojala et al. (2015). The PSSA method consists of three separate phases. First, the formulation of the Bayesian posterior model of the signal and producing a sample from the signal's posterior distribution. Second, the projection of the sample onto the one-dimensional subspaces defined by the left eigenvectors of the SSA eigentriples of, depending on the context, the original noisy time series data or the posterior mean of the sample. For each eigentriple, the projection phase thus gives a projected sample which is interpreted as a sample from the posterior distribution of the time series SSA component associated with the eigentriple. The final, third phase of PSSA makes inferences about the SSA components using the projected samples. In Holmström and Launonen (2013), this was done by computing a so-called credibility map for each SSA component and then visually inspecting those maps for credible features. The credibility maps consist of tapered pillars of black and white on a gray background at the time points, where the black and white features show credibly decreasing or increasing slope and gray means neither holds credibly. The SSA components whose maps are sufficiently colored are

inferred to be signal and those whose maps are mostly gray are inferred to be noise. For examples of such credibility maps, see the lower panel of Figure 3.

The PSSA is preferable over other methods, e.g. perturbation (cf. Hassani et al. (2011)), in the sense that its Bayesian paradigm allows taking into account the uncertainties involved in a principled and coherent manner while permitting a flexible posterior model without restrictive model assumptions. Bayesian modeling is also more flexible than straightforward smoothing of the data. In the present paper in particular, we use a smoothing prior for the signal that does not commit to a single fixed smooth. Uncertainty about the prior smoothing level is extended further by setting a hyperprior for the smoothing prior's smoothing parameter λ_0.

The modification of the projection phase of PSSA for the various generalizations of SSA is straightforward. In Holmström and Launonen (2013), the projecting operators contain the basic SSA embedding and re-Hankelization, which are replaced by their MSSA, 2D-SSA or shaped-SSA variants in the corresponding generalizations of PSSA. However, the inference phase requires a different scheme depending on the data type. For multivariate time series, the method of the credibility maps is similar to that of the single-variate PSSA. For image analysis, inference on the slope of the components is not possible. One option of inference for images was considered in Godtliebsen et al. (2004) in a frequentist scale-space context, where statistically significant gradients in the image pixels were used to identify significant features, e.g. peaks and ridges. This relies on a visual inspection of the gradients, which may be cumbersome for a large number of image components. A more straightforward approach was introduced in Holmström and Pasanen (2012), where the credibility of the image pixel values' sign was evaluated in the context of difference images. More generally, this inference is suitable for any zero-mean image, and thus could be applied to centered SSA image components. In this paper, we restrict ourselves to multivariate time series and present the multivariate generalization of PSSA (MPSSA).

The rest of the paper is organized as follows. In Section 2, we describe the MPSSA algorithm in greater detail. In Section 3, we demonstrate MPSSA with an artificial example and NAO/SOI climate index data, comparing the results to a previous study. In Section 4, we summarize the results and briefly discuss various aspects of the MPSSA method.

2 Multivariate PSSA

2.1 Basic SSA

We first describe the basic SSA algorithm following Golyandina et al. (2001) and Golyandina and Zhigljavsky (2013). Let $\mathbf{f} = [f_1, \ldots, f_N]^T$ be a time series of length N. SSA has four steps, where the first two are called the decomposition stage and the last two the reconstruction stage. Let $1 < L < N$ be an integer parameter called the window length. The first step of SSA is the em-

bedding, where the time series $\mathbf{f}$ is transformed to an $L \times K$ trajectory matrix $\mathbf{X}$, where $K = N - L + 1$ and the ith column of $\mathbf{X}$ is the lagged subseries $[f_i, \ldots, f_{i+L-1}]$ of length L, i.e.

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

The trajectory matrix $\mathbf{X}$ is a Hankel matrix, i.e., each of its skew-diagonals (antidiagonals) has a constant value. The embedding is a linear mapping from $\mathbb{R}^N$ to the space of $L \times K$ Hankel matrices $\mathcal{M}^{L \times K}$. We denote the embedding function by $\mathcal{T}$.

The second step is the singular value decomposition of the trajectory matrix $\mathbf{X}$. Let $\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_L} \geq 0$ be the singular values and $\mathbf{u}_i \in \mathbb{R}^L$ and $\mathbf{v}_i = \mathbf{X}^T \mathbf{u}_i / \sqrt{\lambda_i} \in \mathbb{R}^K$ the corresponding left and right singular vectors of $\mathbf{X}$. Then $\mathbf{X} = \mathbf{X}_1 + \cdots + \mathbf{X}_L$, where $\mathbf{X}_i = \sqrt{\lambda_i} \mathbf{u}_i \mathbf{v}_i$ are rank 1 elementary matrices. The triplet $(\sqrt{\lambda_i}, \mathbf{u}_i, \mathbf{v}_i)$ associated with $\mathbf{X}_i$ is called the ith eigentriple.

The third step, starting the reconstruction stage, is the grouping of the eigentriples. The index set $\{1, \ldots, L\}$ is partitioned into disjoint subsets $I_1, \ldots, I_m$ and, for each subset, the elementary matrices with the corresponding indices are summed to form the resultant matrices $\mathbf{X}_{I_j} = \sum_{i \in I_j} \mathbf{X}_i$. Then $\mathbf{X} = \mathbf{X}_{I_1} + \cdots + \mathbf{X}_{I_m}$. The resultant matrix $\mathbf{X}_{I_j}$ can be calculated by projecting the column vectors of $\mathbf{X}$ onto the subspace spanned by the eigenvectors $\{\mathbf{u}_i \,|\, i \in I_j\}$. We denote by $\mathbf{P}_{I_j} \in \mathbb{R}^{L \times L}$ the projection matrix corresponding to this basis.

The final step is the (skew-)diagonal averaging of the resultant matrices $\mathbf{X}_{I_j}$, which transforms them back into time series components. For each resultant matrix, first the values x_{ij} of the kth skew-diagonal, where $i + j = k = 1$, are replaced with their average for $k = 1, \ldots, N$. This produces a matrix with constant skew-diagonals, i.e., a Hankel matrix, which is the trajectory matrix of some series of length N, since the embedding $\mathcal{T}$ is a one-to-one correspondence between $\mathbb{R}^N$ and $\mathcal{M}^{L \times K}$. The diagonal averaging of the resultant matrix $\mathbf{X}_{I_j}$ thus yields the reconstructed series component $\tilde{\mathbf{f}}_j$ and the original series is decomposed as $\mathbf{f} = \tilde{\mathbf{f}}_1 + \cdots + \tilde{\mathbf{f}}_m$. Formally, diagonal averaging is the orthogonal projection of the resultant matrix onto the space of Hankel matrices $\mathcal{M}^{L \times K}$, which we denote by the Hankelization operator $\mathcal{H}$. The SSA decomposition can thus be expressed as the collection of the composite linear operators $\mathcal{T}^{-1} \mathcal{H} \mathbf{P}_{I_j} \mathcal{T}$, $j = 1, \ldots, m$, where the reconstructed component $\tilde{\mathbf{f}}_j = \mathcal{T}^{-1} \mathcal{H} \mathbf{P}_{I_j} \mathcal{T} \mathbf{f}$.

2.2 MSSA

The generalizations of SSA (MSSA, 2D-SSA and shaped-SSA) share a common scheme with the basic algorithm (Golyandina et al., 2015). First, the data are

transformed via an embedding to a trajectory matrix, where the embedding depends on the generalization and the window shape parameter. Then, the singular value decomposition of the trajectory matrix and the grouping of the eigentriples are performed similar to the basic algorithm. Last, the multivariate series or image components are reconstructed by projecting, with respect to the Frobenius norm, the resultant matrices onto the space of Hankel-type matrices determined by the embedding, and finally taking the inverse embedding. The projection corresponds to modified diagonal averaging, depending on the form of the Hankel-type matrix.

Let $\mathbf{f} = \{\mathbf{f}^{(p)} = [f_1^{(p)}, \ldots, f_{N_p}^{(p)}]^T \mid p = 1, \ldots, s\}$ be a multivariate collection of s time series of lengths $N_1, \ldots, N_s$, let L be a common window length for all the series and denote $K = \sum_{p=1}^s K_p$, where $K_p = N_p - L + 1$. The embedding operator $\mathcal{T}_{\mathrm{MSSA}}$ of MSSA determines the trajectory matrix $\mathbf{X} \in \mathbb{R}^{L \times K}$ by

$$\mathbf{X} = \mathcal{T}_{\mathrm{MSSA}}\mathbf{f} = \begin{bmatrix} \mathbf{X}^{(1)} \, \mathbf{X}^{(2)} \, \cdots \, \mathbf{X}^{(s)} \end{bmatrix},$$

where $\mathcal{T}\mathbf{f}^{(p)} = \mathbf{X}^{(p)} \in \mathcal{M}^{L \times K_p}$ is the basic SSA embedding of the pth time series with window length L. The trajectory matrix $\mathbf{X}$ has thus a stacked Hankel structure determined by L and the time series lengths $N_1, \ldots, N_p$.

Let $\mathbf{X} = \mathbf{X}_1 + \cdots + \mathbf{X}_m$ be a decomposition of the trajectory matrix into resultant matrices, where $\mathbf{X}_j = \begin{bmatrix} \mathbf{X}_j^{(1)} \, \mathbf{X}_j^{(2)} \, \cdots \, \mathbf{X}_j^{(s)} \end{bmatrix}$ and $\mathbf{X}_j^{(p)} \in \mathbb{R}^{L \times K_p}$. The projection of $\mathbf{X}_j$ onto stacked Hankel matrices is then

$$\mathcal{H}_{\mathrm{MSSA}}\mathbf{X}_j = \begin{bmatrix} \mathcal{H}\mathbf{X}_j^{(1)} \, \mathcal{H}\mathbf{X}_j^{(2)} \, \cdots \, \mathcal{H}\mathbf{X}_j^{(s)} \end{bmatrix},$$

i.e., the submatrices $\mathbf{X}_j^{(p)}$ of $\mathbf{X}_j$ are diagonally averaged. We thus get the decomposition of $\mathbf{f}$ into multivariate time series components $\tilde{\mathbf{f}}_j$ by

$$\tilde{\mathbf{f}}_j = \mathcal{T}_{\mathrm{MSSA}}^{-1}\mathcal{H}_{\mathrm{MSSA}}\mathbf{P}_{I_j}\mathcal{T}_{\mathrm{MSSA}}\mathbf{f},$$

where $\mathbf{f}^{(p)} = \tilde{\mathbf{f}}_1^{(p)} + \cdots + \tilde{\mathbf{f}}_m^{(p)}$ for $p = 1, \ldots, s$.

In 2D-SSA, the window is rectangular with two parameters L_x and L_y and the embedding produces a Hankel-block-Hankel matrix where the trajectory matrix is formed by Hankel blocks which themselves form a Hankel matrix (Golyandina et al., 2015). In shaped-SSA, the window need not be rectangular but a two-dimensional grid shape, which slides along the image to form Hankel matrices from the pixels in its area. This embedding produces a quasi-Hankel trajectory matrix.

The MSSA has an advantage over SSA if the multivariate time series contain matching components. The MSSA reconstruction of the common signal is then more accurate, in the sense of mean squared error, than separate reconstructions with single-variate SSA (Golyandina and Stepanov, 2005; Golyandina et al., 2015). A time shift between the matching signals has no effect in MSSA. Thus, for example, sine waves with the same frequency but different phase strengthen each other.

2.3 Choice of parameters

The main parameter of SSA is the window length L whose choice is driven by the concept of separability. Two time series components are said to be strongly separable if, for any singular value decomposition of the trajectory matrix $\mathbf{X}$, it holds that $\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2$, where $\mathbf{X}_1$ and $\mathbf{X}_2$ are the trajectory matrices of the said components (cf. Golyandina et al. (2001)). I.e., the two components can be found by a correctly grouping the eigentriples into two sets within which the elementary matrices are summed to form $\mathbf{X}_1$ and $\mathbf{X}_2$. In practice with noisy series, only approximate separability is attainable, where the weighted so-called w-correlation is small between the series components. Given a fixed time series length N in basic SSA, the window length value which gives the best approximate separation of signal from noise is $L = N/2$ (Golyandina et al., 2001; Golyandina, 2010). More generally, the most detailed decomposition corresponds to a trajectory matrix $\mathbf{X}$ which has maximal rank (Golyandina et al., 2015). If the s series of a multivariate time series have equal length N, the choice of window length which maximizes the rank of $\mathbf{X}$ is thus approximately $[s/(s+1)] \cdot N$.

The grouping of the eigentriples may be considered the second parameter of SSA. This choice is dictated by the theory encompassing series of finite rank, such as polynomials and exponentially modulated harmonic series (cf. Golyandina et al. (2001)). For example, sine waves produce two matching eigentriples. Noise is described in this framework by not having finite rank, such that the number of eigentriples in its decomposition tends to infinity as N grows larger. In Posterior SSA, the eigentriple grouping is postponed and, instead, the elementary grouping $I_j = \{j\}$, $j = 1, \ldots, L$, is used. The orthogonal projection onto the ith eigentriple is thus handled by left multiplying $\mathbf{X}$ by the one-dimensional projection matrix $\mathbf{P}_i = \mathbf{u}_i \mathbf{u}_i^T$.

2.4 Posterior SSA and its generalization

The Posterior SSA method has three separate phases: posterior modeling and sampling, sample projection and posterior inference. Let $\{\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(n)}\}$ be a sample of n realizations from the posterior distribution of the signal time series of interest $\mathbf{g}$. PSSA distinguishes between two ways this sample can be obtained. In the first case, there is a single noisy observed time series $\mathbf{f}$ for which we build a Bayesian model and sample from the posterior $p(\mathbf{g}|\mathbf{f})$. In the second case, we have data only indirectly related to $\mathbf{g}$ which are then used to construct a Bayesian model and the resulting posterior $p(\mathbf{g}|\text{data})$ is sampled. The second case applies e.g. in the context of proxy data for climate time series reconstructions. The sample is then projected, in phase space, onto the left eigenvectors $\mathbf{u}_i$ of the SSA decomposition of either the noisy series $\mathbf{f}$ or the posterior mean $\mathbb{E}(\mathbf{g}|\text{data})$, which we denote in the following also by $\mathbf{f}$. In the case of directly observed data, either may be used, where the choice depends on the nature of the observed data and the quantity on which the modeler is
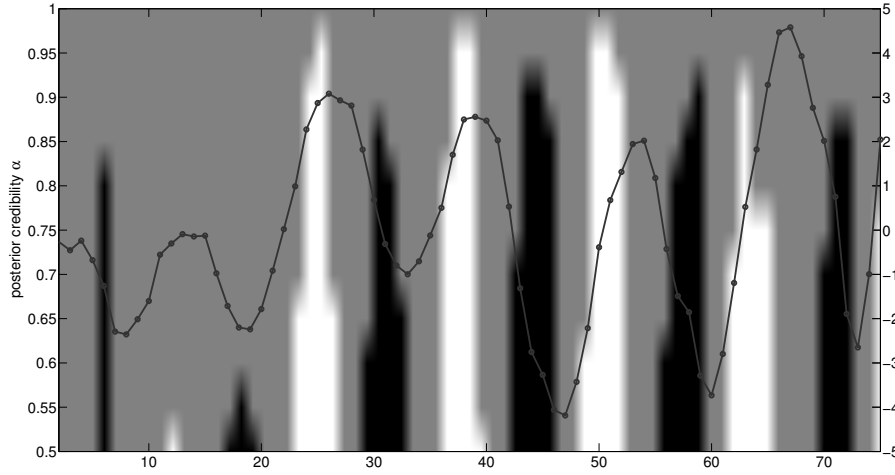
Fig. 1 An example of a credibility map its the associated SSA component (solid line, values at the time points shown by circles). See the text for further information.

interested in. Section 3.1 presents an example of using the noisy series, while Section 3.2 deals with the posterior mean. In the case of indirectly observed data, only the posterior mean is available. For either case, the PSSA method continues in the same vein and does not concern how the posterior sample was obtained or which series the SSA decomposition is based on.

The objective of the sample projection phase is to get approximate samples from the elementary SSA components' posterior distributions. This is done in PSSA by applying the operator $\mathcal{T}^{-1}\mathcal{H}\mathbf{P}_i\mathcal{T}$ to the sample realizations $\mathbf{g}^{(k)}$. Thus, for each $\tilde{\mathbf{f}}_i$, we get a projected sample $\{\tilde{\mathbf{g}}_i^{(k)}\}_{k=1}^n$. The generalization of the projection phase of PSSA is straightforward as it only involves replacing the embedding and the Hankelization operators $\mathcal{T}$ and $\mathcal{H}$ of basic SSA with their counterparts in MSSA, 2D-SSA or shaped-SSA.

The posterior inference phase of PSSA makes inference about the credibility of the features of the SSA components. The credible features are indicated by a so-called credibility map which is computed from the sample $\{\tilde{\mathbf{g}}_i^{(k)}\}_{k=1}^n$. For an example of a detailed credibility map together with the associated SSA component, see Figure 1. The horizontal axis of a credibility map contains the time indices $t_2, \ldots, t_N$ and the vertical axis specifies a finite set of posterior credibility levels α contained in $]0.5, 1]$ (for our purposes, we use the credibility levels $0.5 + 1/n, 0.55, 0.6, 0.65, \ldots, 1$). At each pixel corresponding to a time index and a credibility level, the map has one of three colors white, black or gray. The credibility map is interpreted as follows. First, fix a credibility level α on the vertical axis and look at the colors on that line along the time indices. A white or black pixel indicates an increasing or decreasing slope, respectively, at the time index with a joint posterior probability of at least α. If a pixel is gray, the slope is neither increasing nor decreasing at the credibility level α. A feature such as a 'credible' local extremum located in a band of gray time

---

Algorithm 1 MPSSA

We assume a posterior model for the signal $\mathbf{g}$ of the multivariate time series $\mathbf{f} = \{\mathbf{f}^{(p)}\}_{p=1}^s$ and a sample $\{\mathbf{g}^{(k)} = \{\mathbf{g}^{(p,k)}\}_{p=1}^s\}_{k=1}^n$ drawn from it. A separate posterior model can be used for each component $\mathbf{f}^{(p)}$. Depending on the context, the multivariate series from which the MSSA eigentriples are calculated can be either the noisy series $\mathbf{f}$ or the posterior mean $\mathbb{E}(\mathbf{g}|\text{data})$ of the model. We denote both by $\mathbf{f}$.

1: **Input** : Multivariate series $\mathbf{f}$, sample $\{\mathbf{g}^{(k)}\}_{k=1}^n$.
2: Find a suitable window length L and apply MSSA to $\mathbf{f}$. This produces the eigentriples $(\sqrt{\lambda_i}, \mathbf{u}_i, \mathbf{v}_i)$ and the one-dimensional projections $\mathbf{P}_i = \mathbf{u}_i\mathbf{u}_i^T$, $i = 1, \ldots, d$. Here d is the number of time series components considered in posterior analysis.
3: **for all** $i = 1, \ldots, d \leq L$ **do**
4:     Compute the projected sample $\{\tilde{\mathbf{g}}_i^{(k)} = \mathcal{T}_{\text{MSSA}}^{-1}\mathcal{H}_{\text{MSSA}}\mathbf{P}_i\mathcal{T}_{\text{MSSA}}\mathbf{g}^{(k)}\}_{k=1}^n$.
5: **end for**
6: **for all** $p = 1, \ldots, s$ **do**
7:     **for all** $i = 1, \ldots, d \leq L$ **do**
8:         Compute the component $\tilde{\mathbf{f}}_i = \mathcal{T}_{\text{MSSA}}^{-1}\mathcal{H}_{\text{MSSA}}\mathbf{P}_i\mathcal{T}_{\text{MSSA}}\mathbf{f}$ and plot $\tilde{\mathbf{f}}_i^{(p)}$.
9:         Compute and plot the credibility map of the component part $\tilde{\mathbf{f}}_i^{(p)}$ by using the sample $\{\tilde{\mathbf{g}}_i^{(p,k)}\}_{k=1}^n$ or all the samples for simultaneous inference (cf. Algorithm 2).
10:    **end for**
11: **end for**
12: **Output** : An atlas of the MSSA components $\tilde{\mathbf{f}}_i^{(p)}$ and a corresponding credibility atlas for each $p = 1, \ldots, s$.

Use the plots of the components and their associated credibility maps to identify interesting credible features in the time series considered. Compare the credibility atlases $1, \ldots, s$ for matching credible components. Separate the credible components into groups according to SSA guidelines.

---

indices is thus indicated by white on one side and black on the other of the band.

The set of white and black pixels having at least the joint probability α, for each of the chosen credibility levels, was determined in Holmström and Launonen (2013) by a 'Highest Pointwise Probabilities' (HPW) algorithm, first proposed in Erästö and Holmström (2005). The credibility maps thus consist of a gray background filled by tapered white and black pillars, which cut off at the horizontal line where the credibility ceases to hold (specifically, at the mid-point between two credibility levels α_i and α_{i+1}, where the pixel is credible at the level α_i but not at α_{i+1}). We call this inference simultaneous within a credibility map, since HPW calculates a joint probability. The HPW can also be computed over the combined stacked time indices $1, \ldots, N-1$ of all or a subset of components $\tilde{\mathbf{f}}_i$. In this case, we calculate a single HPW credibility map, which is then segmented into separate maps corresponding to the individual components for easier visual representation. This type of inference is useful to suppress false positives, and we call it simultaneous over an atlas of credibility maps. On the condition that the posterior sample is of a sufficiently high quality, the credibility maps of the signal components are expected to show a plethora of white and and black pixels, while the maps of the noise artifacts are expected to be mostly gray. We represent the inference

visually by plotting the SSA components in a grid in descending order with respect to their singular value magnitude, and then plot the atlas of credibility maps underneath in a similar grid (cf. Figure 3).

The posterior inference of the multivariate time series is similar to the single-variate case. Now, for each multivariate SSA component $\tilde{\mathbf{f}}_i = \{\tilde{\mathbf{f}}_i^{(p)}\}_{p=1}^s$ there is a projected multivariate series sample $\{\{\tilde{\mathbf{g}}_i^{(p,k)}\}_{p=1}^s\}_{k=1}^n$. For each $p = 1, \ldots, s$, we pick the single-variate samples $\{\tilde{\mathbf{g}}_i^{(p,k)}\}_{k=1}^n$ for $i = 1, \ldots, L$ and calculate the atlas of the single-variate component parts $\tilde{\mathbf{f}}_i^{(p)}$ and their associated credibility maps computed from $\{\tilde{\mathbf{g}}_i^{(p,k)}\}_{k=1}^n$. The visual exposition of MPSSA is thus similar to PSSA, except that a complete atlas is shown for each $p = 1, \ldots, s$. Also, the HPW can be computed over the combined time indices of all the component parts, making inference simultaneous over an atlas and between atlases $1, \ldots, p$ of credibility maps. In Section 3, we use solely this fully simultaneous inference for multivariate series, and call the inference just simultaneous. The algorithmic structure of the MPSSA is described in Algorithm 1.

2.5 Construction of credibility maps

To describe map construction in more detail, consider a sample $\{\tilde{\mathbf{g}}_i^{(k)}\}_{k=1}^n$ from the posterior distribution of the ith elementary MSSA component. While our examples deal with time series observed on an equispaced time grid, the grid in general does not have to be uniform. In such a case the time indices j actually correspond to times t_j where the distances $t_{j+1} - t_j$ may vary with j. We formulate map construction in this slightly more general setting.

Let $\mathbf{D} \in \mathbb{R}^{(N-1) \times N}$ be the matrix that computes difference quotients, that is, for $\mathbf{z} = [z_1, z_2, \ldots, z_N]^T \in \mathbb{R}^N$, the slope of $\mathbf{z}$ is $\mathbf{Dz} = \mathbf{w}$ with $w_j = (z_{j+1} - z_j)/(t_{j+1} - t_j)$, $j = 1, \ldots, N-1$. Compute a sample of slopes of the ith elementary MSSA component time series parts $p = 1, \ldots, s$ at the time points $t_2, \ldots, t_{N_p}$ by the difference quotient vectors $\{\boldsymbol{\delta}_i^{(p,k)}\}_{p=1}^s = \{[\delta_1^{(i,p,k)}, \delta_2^{(i,p,k)}, \ldots, \delta_{N_p-1}^{(i,p,k)}]\}_{p=1}^s = \{\mathbf{D}\tilde{\mathbf{g}}_i^{(p,k)}\}_{p=1}^s$, $k = 1, \ldots, n$. Algorithm 2 describes how a simultaneous MPSSA credibility atlas is now constructed based on the posterior distribution of these slopes. The posterior probabilities in steps 3 and 4 of Algorithm 2 can be estimated using this sample.

3 Experimental results

3.1 Two noisy artificial series with a shared sinusoid

Let $N = 51$, $g_j = \sin(2\pi\omega t_j)$, $j = 1, \ldots, N$, where $\omega = 1/15$ and $t_j = 2(j-1)$, and let $\mathbf{f} = \mathbf{g} + \boldsymbol{\varepsilon}$, where $\boldsymbol{\varepsilon} \sim \mathrm{N}(\mathbf{0}, 1.4^2 \mathbf{I}_{51})$. Figure 2 shows the artificial time series $\mathbf{f}^{(1)}$ and $\mathbf{f}^{(2)}$, which contain two realizations from the multivariate Gaussian noise. Our aim is to demonstrate MPSSA in a case where the noise

---

Algorithm 2 Construction of simultaneous MPSSA Credibility Maps

---

1: Index the time series parts by $p = 1, \ldots, s$ and the elementary MSSA components by $i = 1, \ldots, d$. Select a set of credibilities $0.5 < \alpha_1 < \cdots < \alpha_M \leq 1$.

2: **for all** $\alpha \in \{\alpha_1, \ldots, \alpha_M\}$ **do**

3:     Decompose the index set $\bigcup_{p=1}^{s} \{1, \ldots, N_p - 1\} \times \{1, \ldots, d\} \times \{p\}$ into three subsets,

$$I^+ = \{(j, i, p) \mid \mathbb{P}(\delta_j^{(i,p)} > 0 \mid \text{data}) \geq \alpha\},$$

$$I^- = \{(j, i, p) \mid \mathbb{P}(\delta_j^{(i,p)} < 0 \mid \text{data}) \geq \alpha\},$$

$$I^0 = \bigcup_{p=1}^{s} \{1, \ldots, N_p - 1\} \times \{1, \ldots, d\} \times \{p\} \setminus (I^+ \cup I^-).$$

4:     Let $E_{(j,i,p)}$ denote the event $\delta_j^{(i,p)} > 0 \mid \text{data}$, when $(j, i, p) \in I^+$ and $\delta_j^{(i,p)} < 0 \mid \text{data}$, when $(j, i, p) \in I^-$. Let $q_1, \ldots, q_S$ be an ordering of the triplets in $I^+ \cup I^-$ such that

$$\mathbb{P}(E_{q_1}) \geq \mathbb{P}(E_{q_2}) \cdots \geq \mathbb{P}(E_{q_S}) \geq \alpha$$

and let

$$k = \max\{\ell \mid \mathbb{P}(E_{q_1} \& \cdots \& E_{q_\ell}) \geq \alpha\}.$$

5:     Color pixels $(q_1, \alpha), \ldots, (q_k, \alpha)$ white or black depending on whether $q_\ell \in I^+$ or $q_\ell \in I^-$, and color the rest of the pixels (q, α) gray.

6: **end for**

7: The credibility map of the ith elementary MPSSA component for the time series part p is given by the colors of the pixels $((1, i, p), \alpha_1), \ldots, ((1, i, p), \alpha_M), \ldots, ((N_p - 1, i, p), \alpha_1), \ldots, ((N_p - 1, i, p), \alpha_M)$.

---

level is too high for the proper separation and credible identification of the signal $\mathbf{g}$ with single-variate PSSA.

In MPSSA, we must first define a posterior model of the signal. We use the additive model $\mathbf{f} = \mathbf{g} + \boldsymbol{\varepsilon}$ described in Erästö and Holmström (2005), where the likelihood is normal, $\mathbf{f} \sim N(\mathbf{g}, \sigma^2 \mathbf{I})$, and the variance σ^2 is unknown with a scaled inverse chi-squared prior, $\sigma^2 \sim \text{Scale-inv-}\chi^2(\nu_0, \sigma_0^2)$, where the degrees of freedom ν_0 determines how tight or loose the prior is. The signal $\mathbf{g}$ has a Gaussian smoothing prior conditional on the variance σ^2 and a prior smoothing parameter λ_0,

$$p(\mathbf{g}|\sigma^2, \lambda_0) \propto \left(\frac{\lambda_0}{\sigma^2}\right)^{\frac{N-2}{2}} \exp\left(-\frac{\lambda_0}{2\sigma^2}\|\mathbf{C}\mathbf{g}\|^2\right), \tag{1}$$

where $\mathbf{C} \in \mathbb{R}^{(N-2) \times N}$ is the second order differencing matrix for series of length N. For a fixed λ_0, the marginal posterior of $\mathbf{g}$ has a multivariate-t distribution, $\mathbf{g}|\mathbf{f} \sim t_{\nu_0 + N - 2}(\mathbf{S}_{\lambda_0}\mathbf{f}, \boldsymbol{\Sigma}_0)$, where $\mathbf{S}_{\lambda_0} = (\mathbf{I} + \lambda_0 \mathbf{C}^T \mathbf{C})^{-1}$ is a discrete spline smoothing matrix and the covariance is

$$\boldsymbol{\Sigma}_0 = \left(\frac{\|\mathbf{f}\|^2 - \mathbf{f}^T \mathbf{S}_{\lambda_0} \mathbf{f} + \nu_0 \sigma_0^2}{\nu_0 + N - 2}\right) \mathbf{S}_{\lambda_0}.$$

Before sampling from the posteriors $p(\mathbf{g}|\mathbf{f}^{(1)})$ and $p(\mathbf{g}|\mathbf{f}^{(2)})$ can take place, the hyperparameters of σ^2 and the fixed λ_0 need to be set. We used a maximum

likelihood algorithm, proposed in Appendix A of Holmström and Pasanen (2012) in the context of image analysis, to estimate the prior means of σ^2. The estimated values were about 1.5^2 and 1.45^2 for $\mathbf{f}^{(1)}$ and $\mathbf{f}^{(2)}$, respectively. We then chose a loose prior with degrees of freedom $\nu_0 = 10$ and chose σ_0^2 such that the prior means were the estimated variances. The algorithm also estimated λ_0, but these values gave extremely smooth posteriors due to the high level of noise. We thus settled to pick λ_0 on a subjective basis by inspecting the scale-derivative maps of $\mathbf{f}^{(1)}$ and $\mathbf{f}^{(2)}$. The scale-derivative map is a visual tool introduced by Pasanen et al. (2013) that helps to pick a value of λ_0 for which much of the noise is smoothed out by the smoother $\mathbf{S}_{\lambda_0}$ but the signal is left intact. The scale-derivative maps suggested a value of the order 10^0, and we chose $\lambda_0 = 5$ for both $\mathbf{f}^{(1)}$ and $\mathbf{f}^{(2)}$.

We drew a sample of 1000 realizations from the multivariate-t posteriors of $\mathbf{f}^{(1)}$ and $\mathbf{f}^{(2)}$ and projected the combined bivariate sample on the left eigenvectors of the bivariate time series' MSSA eigentriples, calculated with a window length $L = (2/3) \cdot 51 = 34$. The credibility maps for the MSSA component parts corresponding to $\mathbf{f}^{(1)}$ and $\mathbf{f}^{(2)}$ were then computed with the respective parts of the projected bivariate samples, using simultaneous inference over the first 25 maps of both series. The first 25 MSSA components of $\mathbf{f}^{(1)}$ and their credibility atlas are shown in Figure 3 and the components and credibility atlas of $\mathbf{f}^{(2)}$ are shown in Figure 4. The titles of the components show wavelength values which are determined as the maxima of the components' Lomb-Scargle periodograms. The first two MSSA components shown in the upper panels of Figures 3 and 4 reconstruct the true signal $\sin(2\pi(1/15)t)$, where the wavelength approximation is 15.38 for $\mathbf{f}^{(1)}$ and 14.81 for $\mathbf{f}^{(2)}$. Their respective credibility maps, shown in the lower panels of Figures 3 and 4, have high credibility values at the time points where the components' slope is non-zero. The other MSSA components pertaining to noise have only haphazard credible features.

For comparison, the single-variate SSA components and their simultaneous credibility atlas are shown in Figure 5 for $\mathbf{f}^{(1)}$ and in Figure 6 for $\mathbf{f}^{(2)}$. The window length with which the components were calculated was $L = 25$, half the length of the series rounded down. Compared to the MSSA components, the signal does separate from noise as well. The third SSA component of $\mathbf{f}^{(1)}$ in the upper panel of Figure 5 has approximately the same wavelength and the functional form as the signal, but no sinusoidal pair. Its credibility map shows credible features, but so does the false pair of the 4th and 5th components with a wavelength of 20. In the upper panel of Figure 6, the 1st and 5th SSA components of the series $\mathbf{f}^{(2)}$ have a wavelength of 15.38, but their functional form is not correct. The credibility atlas of $\mathbf{f}^{(2)}$ shows only some credible features. We conclude that in this artificial example, the superior performance of the MPSSA compared to single-variate PSSA is due to the better identification of the correct signal subspace with MSSA, since the relatively high noise level confounds the single-variate SSA.
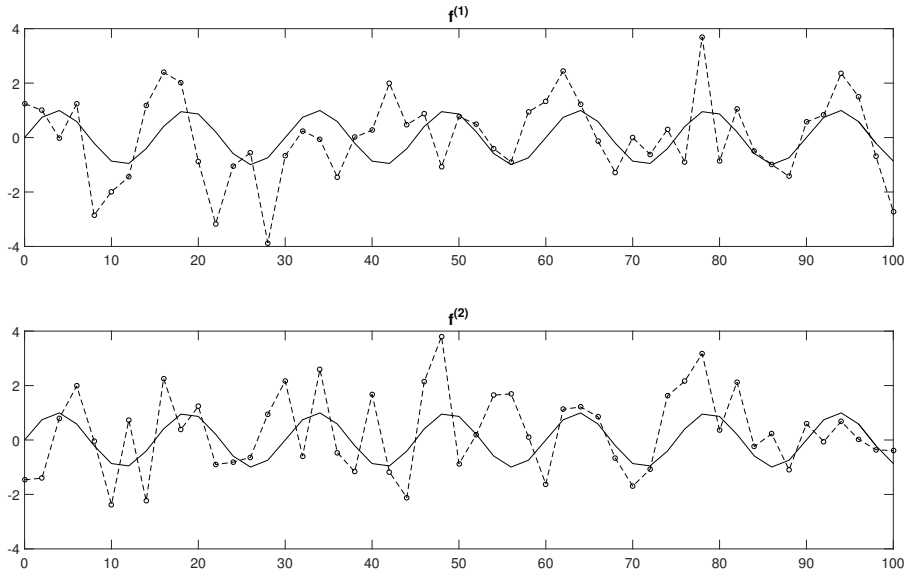
Fig. 2 The signal $\sin(2\pi(1/15)t)$ (solid line) and the artificial noisy series $\mathbf{f}^{(1)}$ (upper panel, dashed line) and $\mathbf{f}^{(2)}$ (lower panel, dashed line).

3.2 NAO and SOI

The NAO (North Atlantic Oscillation) index and SOI (Southern Oscillation Index) are two well-studied time series, defined as the normalized monthly mean sea level pressure difference between two locations in the North Atlantic and equatorial Pacific, respectively. For SOI, we use the data provided by the Australian Bureau of Meteorology, cf. `http://www.bom.gov.au/climate/current/soi2.shtml`, that consist of monthly pressure differences between Tahiti and Darwin beginning from January 1876. For NAO, we use the University of East Anglia data, cf. `http://www.cru.uea.ac.uk/cru/data/nao/`, where the pressure difference is calculated between Gibraltar and South-West Iceland. The NAO data extend to instrumental records from 1821 with some missing values, compiled by Jones in Jones et al. (1997) and updated from 2000 to May 2015 by Tim Osborn, cf. `http://www.cru.uea.ac.uk/~timo/datapages/naoi.htm`. The two series spanning the years 1876 to 2014 are shown in the bottom panels of Figure 7.

NAO and SOI have several multiyear oscillations attributed to them. SOI is particularly thought to indicate El Niño and La Niña warming and cooling events. Feliks et al. (2013) examined NAO, SOI and Indian monsoon rainfall records for interactions between their oscillatory modes. They used single-variate Monte Carlo SSA on each time series and did pairwise comparisons of the series with MC-MSSA coupled with a varimax rotation of the eigenvectors for enhanced separation of oscillations. The monthly data were forward-backward filtered with a Chebyshev type 1 filter to cut off frequencies more
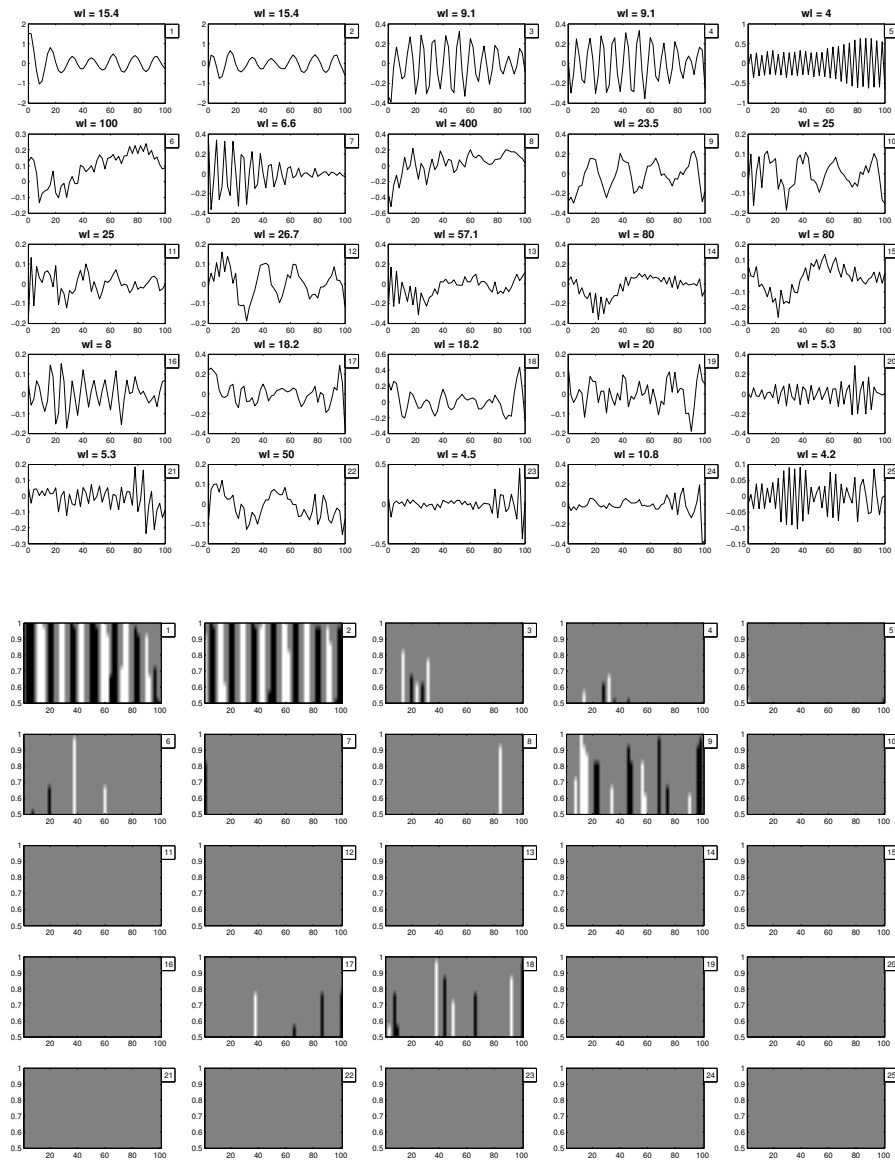
Fig. 3 MPSSA analysis of the artificial bivariate series. Upper panel: The $\mathbf{f}^{(1)}$ part of the bivariate MSSA components, calculated from the bivariate series $\{\mathbf{f}^{(1)}, \mathbf{f}^{(2)}\}$ with window length $L = 34$. Lower panel: The simultaneous credibility atlas for $\mathbf{f}^{(1)}$.
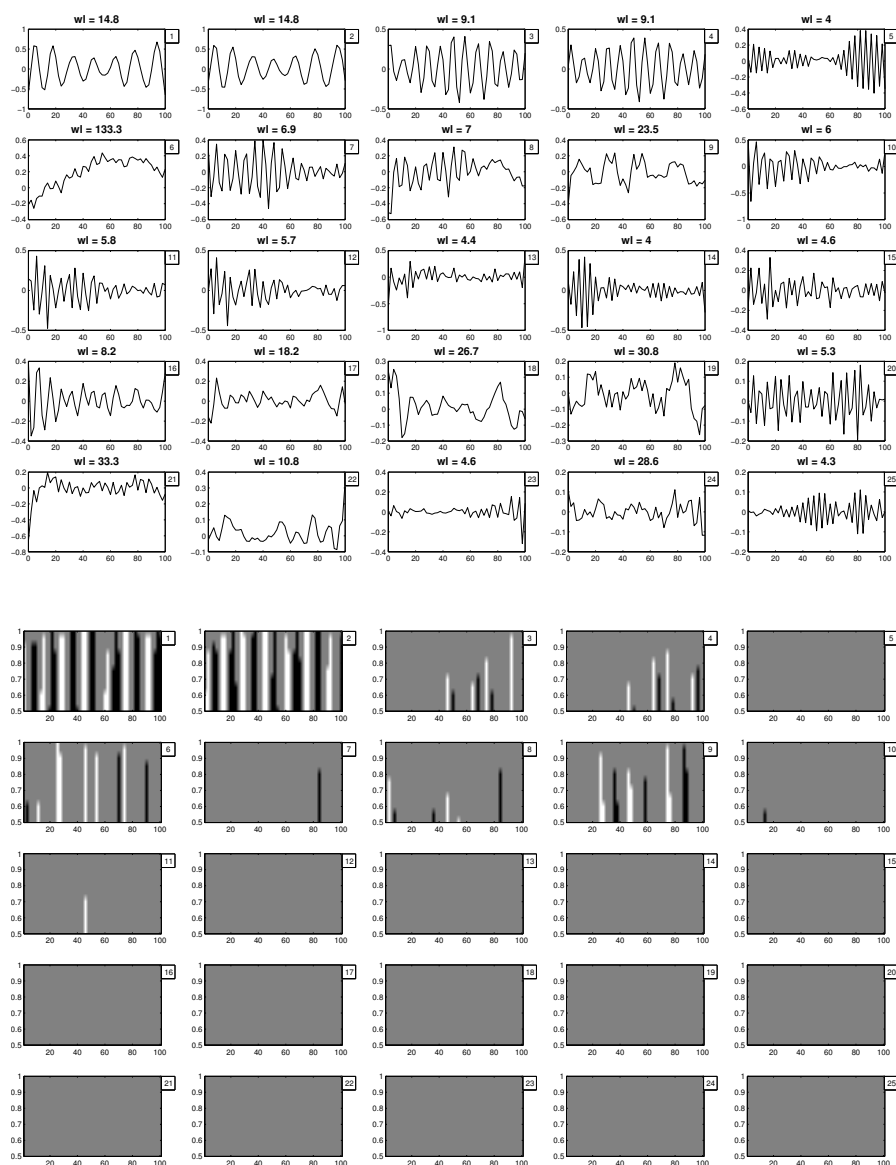
Fig. 4 MPSSA analysis of the artificial bivariate series. Upper panel: The $\mathbf{f}^{(2)}$ part of the bivariate MSSA components, calculated from the bivariate series $\{\mathbf{f}^{(1)}, \mathbf{f}^{(2)}\}$ with window length $L = 34$. Lower panel: The simultaneous credibility atlas for $\mathbf{f}^{(2)}$.
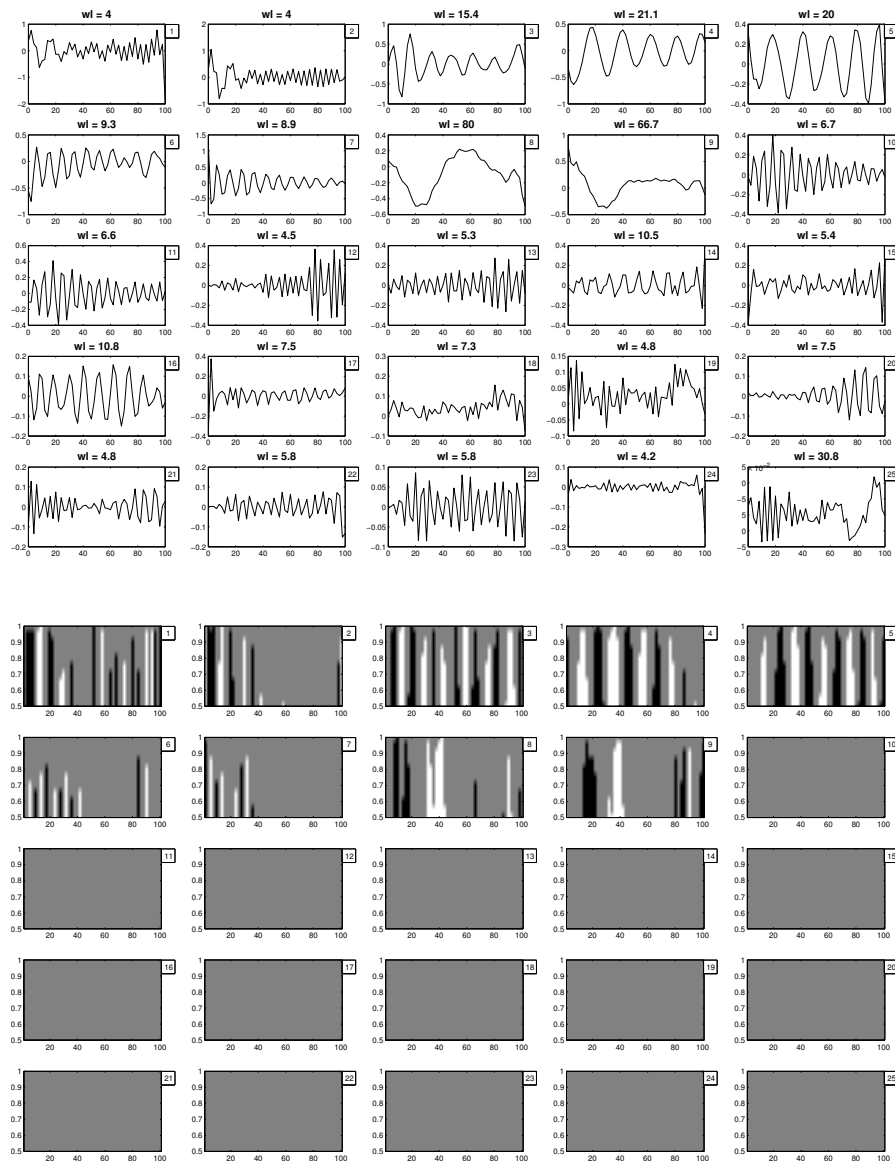
Fig. 5 PSSA analysis of the artificial series $\mathbf{f}^{(1)}$. Upper panel: The SSA components of $\mathbf{f}^{(1)}$, where $L = 25$. Lower panel: The simultaneous credibility atlas corresponding to the components.
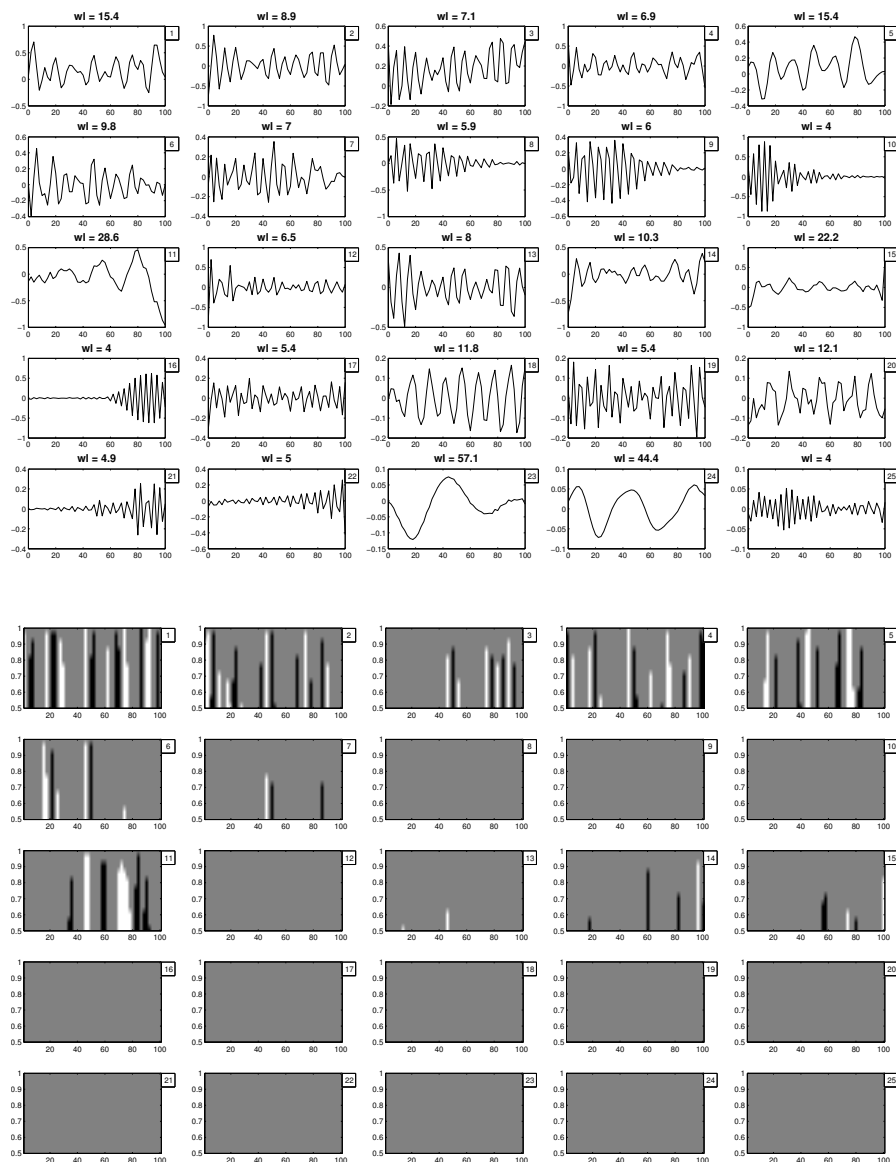
Fig. 6 PSSA analysis of the artificial series $\mathbf{f}^{(2)}$. Upper panel: The SSA components of $\mathbf{f}^{(2)}$, where $L = 25$. Lower panel: The simultaneous credibility atlas corresponding to the components.

than 0.45 cycles per year, after which annually subsampled July values were used for analysis. Using MC-SSA separately on each series with a window length of 40 years, they found statistically significant oscillatory modes of 6-7, 5, 3.6 and 2.5 years for SOI, and 7.8, 5.9, 5, 4.2 and 2.7 years for NAO. Using MC-MSSA with varimax rotation on NAO and SOI with a window length of 50 years, the statistically significant oscillatory modes for the multivariate system were 7.8, 5.8 and 3.6 years.

Our aim is to compare these previous results with Posterior MSSA of the NAO and SOI. First, we build a posterior model of the signal for the monthly NAO and SOI records. We use the additive model $\mathbf{f} = \mathbf{g} + \boldsymbol{\varepsilon}$, where $\boldsymbol{\varepsilon} \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma})$ and where the noise covariance matrix $\boldsymbol{\Sigma}$ is determined by the red noise AR(1) process $\varepsilon_t = \gamma \varepsilon_{t-1} + \sigma z_t$, $z_t \sim \mathrm{N}(0,1)$. The correlation parameter γ is assumed to have a uniform prior on the interval $[-1,1]$ and the variance σ^2 a scaled inverse chi-squared prior. As in the independent error model of Section 3.1, we assume a Gaussian smoothing prior for $\mathbf{g}$. However, where dividing the smoothing parameter λ_0 by σ^2 in (1) makes possible the simple closed form posterior, the dependent error structure complicates such an approach and the smoothing parameter is therefore left unscaled (cf. Erästö and Holmström (2007)). We also treat λ_0 as unknown and assign a gamma prior to it. Assuming prior independence of the model parameters, their joint posterior is

$$p(\mathbf{g}, \lambda_0, \gamma, \sigma^2 | \mathbf{f}) \propto p(\mathbf{f} | \mathbf{g}, \boldsymbol{\Sigma}) \cdot p(\mathbf{g} | \lambda_0) \cdot p(\lambda_0) \cdot p(\gamma) \cdot p(\sigma^2),$$

where

$$\mathbf{f} \mid \mathbf{g}, \boldsymbol{\Sigma} \sim \mathrm{N}(\mathbf{g}, \boldsymbol{\Sigma})$$
$$\mathbf{g} \mid \lambda_0 \sim \lambda_0^{\frac{N-2}{N}} \exp(-\tfrac{\lambda_0}{2} \|\mathbf{C}\mathbf{g}\|^2)$$
$$\lambda_0 \sim \mathrm{Gamma}(\eta_0, \beta_0)$$
$$\gamma \sim \mathrm{Unif}(-1, 1)$$
$$\sigma^2 \sim \text{Scale-inv-}\chi^2(\nu_0, \sigma_0^2).$$

The hyperparameters η_0 and β_0 are the shape and rate of the gamma prior of λ_0. For both NAO and SOI, we chose a loose prior for σ^2 with $\nu_0 = 3$. The choice of σ_0^2 was such that the prior mean was centered at an AR(1) noise variance value which was estimated from the residuals of the NAO and SOI series. The residuals were obtained by smoothing the data with the discrete spline smoother $\mathbf{S}_\kappa = (\mathbf{I} + \kappa \mathbf{C}^T \mathbf{C})^{-1}$, where the smoothing level κ was chosen so that the residual resembled AR(1) noise reasonably well in exploratory analyses, where we looked at the sample autocovariance and partial autocovariance plots of the data and the scatterplot and the Q-Q plot of the fitted residual. For the hyperparameters η_0 and β_0 of the gamma prior of λ_0, we chose $\eta_0 = 1$, making the prior an exponential distribution, and $\beta_0 = 1/\kappa$ so that the prior mean is κ.

We sampled from the posterior distribution with a Metropolis-Hastings within Gibbs algorithm, tuned separately for NAO and SOI, since the form of

the posterior does not permit direct sampling. The starting values for λ_0 and $\mathbf{g}$ were κ and $\mathbf{S}_\kappa \mathbf{f}$, and the starting values for γ and σ^2 were their estimates from the residual. We restricted the monthly NAO and SOI series to the years 1876-2014 and, after discarding burn-in period and thinning the sample by a factor of 1000, obtained a sample of 1000 realizations for both series. As in Feliks et al. (2013), we subsampled annual July values from the posterior realizations of $\mathbf{g}$ to get series of length 139. However, we didn't low-pass filter prior to subsampling, since the posterior realizations of $\mathbf{g}$ were already smooth. The posterior sample means of $\mathbf{g}_{\text{NAO}}$ and $\mathbf{g}_{\text{SOI}}$, shown in Figure 7, were then used as the multivariate series on which Posterior MSSA was based. We chose window length $L = 50$ and projected the combined NAO and SOI sample $\mathcal{S} = \{\{\mathcal{S}_{\text{NAO}}^{(1)}, \mathcal{S}_{\text{SOI}}^{(1)}\}, \ldots, \{\mathcal{S}_{\text{NAO}}^{(1000)}, \mathcal{S}_{\text{SOI}}^{(1000)}\}\}$ on the left eigenvectors of the bivariate posterior mean's 50 eigentriples. We then calculated the credibility atlases for the bivariate MSSA components, simultaneously over all the credibility maps. The NAO part of the MSSA components and their credibility atlas is shown in Figure 8 and the SOI part is shown in Figure 9. The titles of the component plots show the dominant wavelengths of their Lomb-Scargle periodograms.

The pair of components 1 and 4 has a wavelength of about 13-14 years and their NAO maps show high credibility at nearly all time points. Their SOI maps have some non-credible gaps between the time points. In Feliks et al. (2013), the 7th and 8th varimax rotated eigenvectors of the MSSA of NAO and SOI had a wavelength of 13.4 and 13.5 years, respectively, but they were not statistically significant in the Monte Carlo analysis. The component pair 2 and 3 shows a wavelength of about 6.5 years and is credible for SOI but not for NAO. This is in line with Feliks et al. (2013), who found such a periodicity for SOI with a single-variate MC-SSA. The dominant wavelength of components 5 and 6 is about 7.5 years and has credible features for about the first half of the series, which compares to the 7.8 year cycle in Feliks et al. (2013). The 5.8 year cycle might correspond to the pair of components 10 and 11, which have a wavelength of about 5.6 years. Looking at their credibility maps, they seem credible for the most part for SOI but not for NAO. A credible wavelength corresponding to the 3.6 years cycle in Feliks et al. (2013) seems absent in the Posterior MSSA components.

4 Summary

In Section 2, we described MSSA and Posterior SSA, and showed how PSSA can be generalized to a multivariate version. The projection phase of PSSA was generalized by replacing the embedding and the Hankelization operators of the composite projection operator by their multivariate SSA variants, where the trajectory matrix had a stacked-Hankel structure. The inference of MPSSA was done similarly as in PSSA. We computed a credibility map for each single-variate SSA component part of the multivariate SSA components, using the corresponding single-variate part of the projected sample. We argued that the further generalization of the projection phase into 2D-SSA and shaped-SSA is
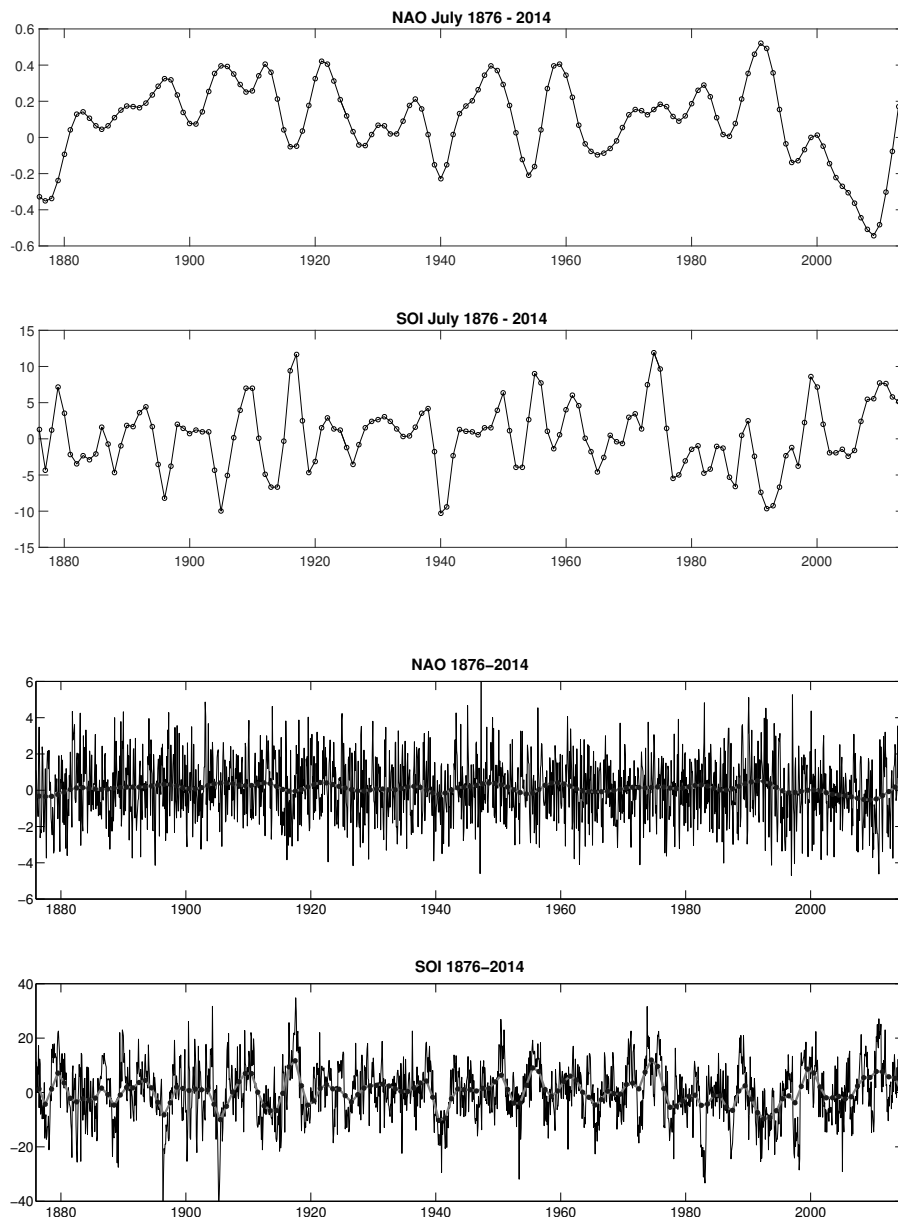
Fig. 7 The upper part shows the July values of the posterior sample means of NAO (first panel) and SOI (second panel). The lower part shows the raw NAO (third panel) and SOI (fourth panel) monthly time series used in posterior modeling for the years 1876 to 2014. The July values of the posterior means are superimposed over the data.
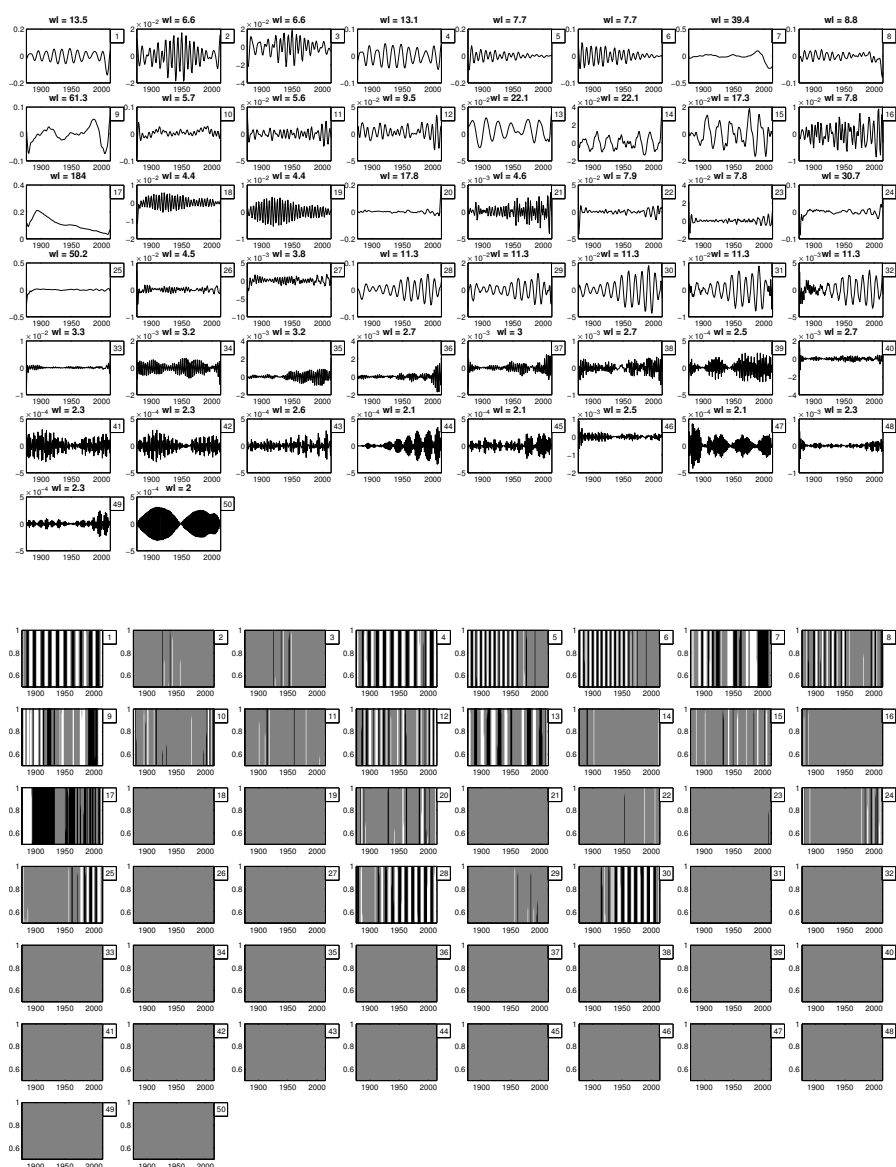
Fig. 8 MPSSA analysis of the NAO time series. Upper panel: The NAO part of the bivariate MSSA components, calculated from the bivariate posterior mean. Lower panel: The NAO components' simultaneous credibility atlas.
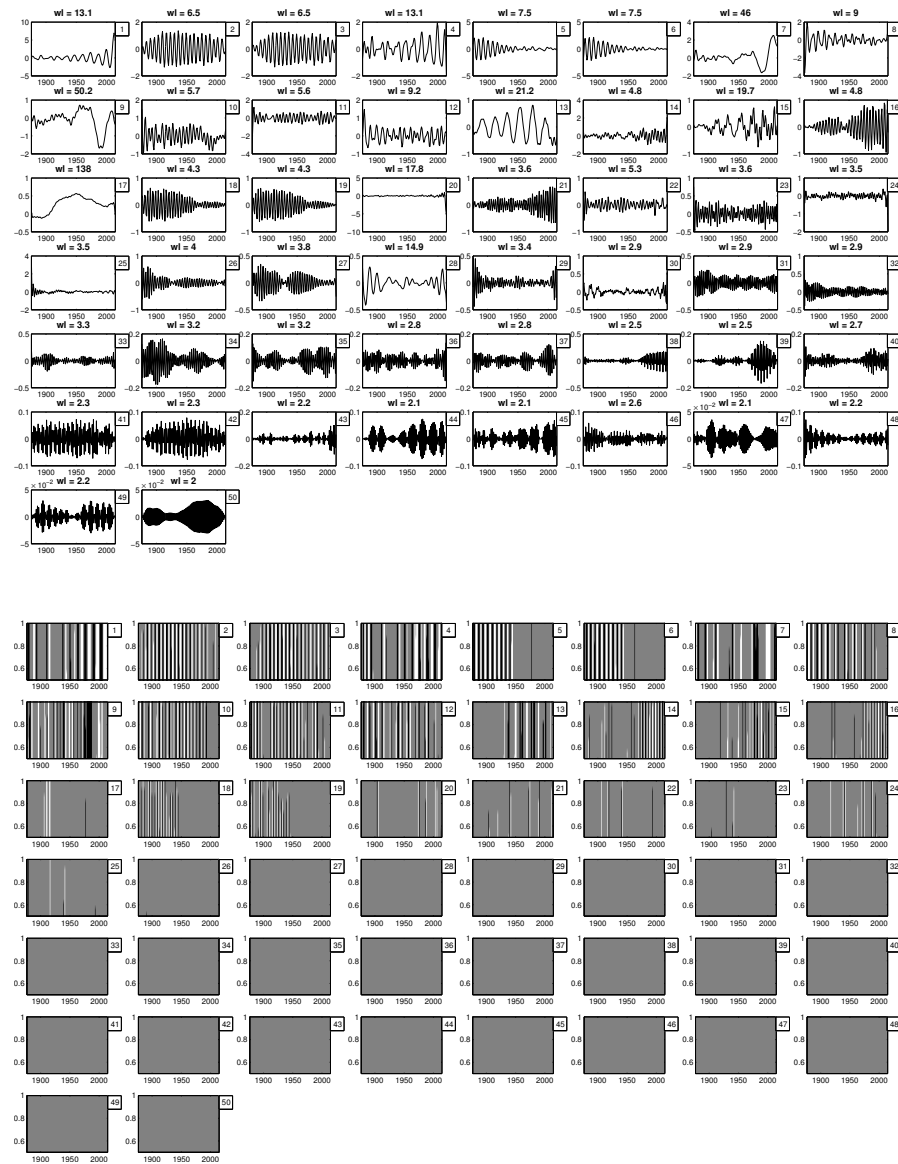
Fig. 9 MPSSA analysis of the SOI time series. Upper panel: The SOI part of the bivariate MSSA components, calculated from the bivariate posterior mean. Lower panel: The SOI components' simultaneous credibility atlas.

similarly straightforward, but the inference requires a different scheme suited for image data, as discussed in Section 1.

In Section 3, we demonstrated the MPSSA algorithm with an artificial example and NAO and SOI climate index series. We compared MPSSA with the single-variate PSSA in the artificial example, where the performance of MPSSA was better. The reason for this was that the single-variate SSA, unlike MSSA, could not locate the signal subspace properly due to the high level of noise compared to the length of the series. The results of the NAO and SOI example were compared with those by Feliks et al. (2013). We found evidence for some harmonic features with a similar wavelength as those deemed statistically significant in Feliks et al. (2013), who analyzed the data with Monte Carlo MSSA and SSA combined with varimax rotation and Chebyshev type 1 pre-filtering and subsampling of the index series.

The posterior modeling and sampling of the signal plays a crucial role in MPSSA. If the model is not satisfactory, the credibility maps may show white and black false positives or gray false negatives even though the signal subspace is found with MSSA. In the artificial example of Section 3.1, the posterior model, which is based on smoothing, breaks down faster than MSSA if the noise is too high. This is because the sampling is done individually for both series and cannot take advantage of the common signal, which benefits MSSA. One topic for future work is to consider a joint Bayesian model for the component time series. Such a model could also explicitly take into account the spatial dependence between the time series.

The hierarchical model in Section 3.2 based on an AR(1)-noise assumption contains the random parameters λ_0, γ and σ^2 for the prior smoothing, autocorrelation and the noise variance, respectively. Furthermore, λ_0 and σ^2 have a gamma and a scaled inverse chi-squared prior, respectively, while γ has a uniform prior. The values of λ_0 and σ^2 can be estimated and set as the prior means, such that the modeler is left only to decide the vagueness of the priors by choosing the shape η_0 for the gamma prior and the degrees of freedom ν_0 for the scaled inverse chi-squared prior. Concerning the occurrence of false positives in the credibility maps, the parameter λ_0 is paramount. We tested the model with pure AR(1)-noise for various autocorrelations and found that the posterior values of λ_0, with the shape hyperparameter η_0 ranging from 1 to 2, were about an order of magnitude less than the ML-estimate of λ_0 (10^3 versus 10^4). Some low-frequency components may thus be suspect when inference is based on the model in Section 3.2, producing false positive features in the credibility maps. However, using a fixed ML-estimate of λ_0, the low-frequency features were eliminated in the resulting t-posterior of the signal and the false positives disappeared.

Comparing MPSSA to the Monte Carlo MSSA, MPSSA is pronouncedly a time-domain method which gives information about the credibility of time series component features at multiple time points, where the component may have credible features at some time intervals but not at others. Unlike Monte Carlo MSSA, MPSSA can point out components which may be common for only a proper subset of the multivariate series, when the credibility maps are

gray for some of the single-variate component parts with the same singular value magnitude ordinal number. However, the MPSSA as such is not suitable for the analysis of a large number of time series, since the computational burden and the manual burden of the visual inspection of the credibility maps may grow too large. The computational burden grows in particular if the inference is simultaneous between credibility atlases. The computational complexity of the Hankelization, up to a quasi-Hankel matrix, can however be brought down to $\mathcal{O}(N \log N)$ with a fast Fourier transform implementation (Korobeynikov, 2010; Golyandina et al., 2015). The analysis of a large set of data with MPSSA would likely require an inference limited to the L eigentriples of the trajectory matrix instead of the $s \cdot L$ multivariate time series components.

References

Allen M, Robertson A (1996) Distinguishing modulated oscillations from coloured noise in multivariate datasets. Clim Dynam 12:775–784

Allen M, Smith L (1996) Monte Carlo SSA: Detecting Irregular Oscillations in the Presence of Colored Noise. J Climate 9(12):3373–3401

Broomhead DS, King G (1986a) Extracting qualitative dynamics from experimental data. Physica D 20:217–236

Broomhead DS, King G (1986b) On the qualitative analysis of experimental dynamical systems. In: Sarkar S (ed.) Nonlinear Phenomena and Chaos, pp. 113–144, Adam Hilger, Bristol, Boston

Erästö P, Holmström L (2005) Bayesian Multiscale Smoothing for Making Inferences about Features in Scatter Plots. J Comput Graph Stat 14(3):569–589

Erästö P, Holmström L (2007) Bayesian analysis of features in a scatter plot with dependent observations and errors in predictors. J Stat Comput Sim 77(5):421–431

Feliks Y, Groth A, Robertson A, Ghil M (2013) Oscillatory Climate Modes in the Indian Monsoon, North Atlantic, and Tropical Pacific. J Climate 26(23):9528–9544

Godtliebsen F, Marron JS, Chaudhuri P (2004) Statistical significance of features in digital images. Image Vision Comput 22(13):1094–1104

Golyandina N (2010) On the choice of parameters in Singular Spectrum Analysis and related subspace-based methods. Stat Interface 3:259–279

Golyandina N, Korobeynikov A, Shlemov A, Usevich K (2015) Multivariate and 2D Extension of Singular Spectrum Analysis with the Rssa Package. J Stat Softw 67(2):1–78

Golyandina N, Nekrutkin V, Zhigljavsky A (2001) Analysis of Time Series Structure, SSA and related techniques. Monographs on Statistics and Applied Probability, Chapman & Hall/CRC

Golyandina N, Stepanov D (2005) SSA-based approaches to analysis and forecast of multidimensional time series. In: Proceedings of the 5th St.Petersburg Workshop on Simulation, pp. 293–298, St. Petersburg, Russia

Golyandina N, Usevich K (2010) 2d-extension of singular spectrum analysis: algorithm and elements of theory. In: Olshevsky V, Tyrtyshnikov E (eds.) Matrix Methods: Theory, Algorithms and Applications, pp. 449–473, World Scientific Publishing

Golyandina N, Zhigljavsky A (2013) Singular Spectrum Analysis for Time Series. SpringerBriefs in Statistics, Springer

Hassani H, Xu Z, Zhigljavsky A (2011) Singular spectrum analysis based on perturbation theory. Nonlinear Anal-Real 12(5):2752–2766

Holmström L, Launonen I (2013) Posterior Singular Spectrum Analysis. Stat Anal Data Min 6(5):387–402

Holmström L, Pasanen L (2012) Baysian Scale Space Analysis of Differences in Images. Technometrics 54(1):16–29

Jones PD, Jonsson T, Wheeler D (1997) Extension to the North Atlantic Oscillation using early instrumental pressure observations from Gibraltar and South-West Iceland. Int J Climatol 17:1433–1450

Korobeynikov A (2010) Computation- and Space-Efficient Implementation of SSA. Stat Interface 3(3):357–368

Ojala AEK, Launonen I, Holmström L, Tiljander M (2015) Effects of solar forcing and North Atlantic oscillation on the climate of continental Scandinavia during the Holocene. Quaternary Sci Rev 112:153–171, doi: http://dx.doi.org/10.1016/j.quascirev.2015.01.021

Pasanen L, Launonen I, Holmström L (2013) A scale space multiresolution method for extraction of time series features. Stat 2(1):273–291

Shlemov A, Golyandina N (2014) Shaped extensions of singular spectrum analysis. In: 21st International Symposium on Mathematical Theory of Networks and Systems, pp. 1813–1820, Groningen, The Netherlands