Annals of the Institute of Statistical Mathematics
https://doi.org/10.1007/s10463-026-00987-8

Confident grouping in singular spectrum analysis via multiple testing

Maryam Movahedifar[1] · **Friederike Preusse**[1] · **Anna Vesely**[2] · **Thorsten Dickhaus**[1]

Received: 25 August 2025 / Revised: 6 March 2026 / Accepted: 27 April 2026

Abstract
A key step in separating signal from noise in time series by means of singular spectrum analysis (SSA) is grouping. We present a multiple testing method for the grouping step in SSA. As separability criterion, we utilize the weighted correlation between the signal and the noise component of the (reconstructed) time series, and we test whether the population version of this weighted correlation is equal to zero. This test has to be performed for several possible groupings, resulting in a multiple test problem. The null distributions of the corresponding test statistics are approximated by a wild bootstrap procedure. The performance of our proposed method is assessed in a simulation study, and we illustrate its practical application with an analysis of real world data.

Keywords Family-wise error rate · Signal extraction · w-correlation · Wild bootstrap for dependent data

1 Introduction

Pre-processing data is an essential step for data analysis. In this regard, noise reduction and signal extraction are considered an important step of data analysis in all fields of study; see, among many others, Pyle (1999); Oliveri et al. (2019); Movahedifar et al. (2022). In the present work, we are concerned with the statistical analysis of

✉ Thorsten Dickhaus
dickhaus@uni-bremen.de

Maryam Movahedifar
movahedm@uni-bremen.de

Friederike Preusse
preusse@uni-bremen.de

Anna Vesely
anna.vesely2@unibo.it

[1] Institute for Statistics, University of Bremen, 28344 Bremen, Germany

[2] Department of Statistical Sciences, University of Bologna, 40126 Bologna, Italy

Published online: 25 June 2026

noisy time series. Singular Spectrum Analysis (SSA) is a non-parametric denoising method for analyzing time series. SSA is used in a wide range of application areas, including ecology and environmental research, medicine, economics, and finance (see, for example, Hassani et al. (2015); Hou et al. (2014); Liu et al. (2014); Le Bail et al. (2014); Chen et al. (2013); Muruganatham et al. (2013); Chao and Loh (2014); Aydın et al. (2011); Ghodsi and Yarmohammadi (2014); Sanei and Hassani (2016) and references therein). A detailed description of other popular denoising methods for one-dimensional time series is provided by Köhler and Lorenz (2005).

In the SSA method, the observed time series (modeled as signal plus noise) is transformed into a matrix $\mathbf{X}$, the columns of which consist of overlapping consecutive data points. This matrix $\mathbf{X}$ is called the trajectory matrix. The key step in SSA, when used for denoising, is grouping. The grouping step results in a decomposition of $\mathbf{X}$. Choosing the number of groups equal to two allows for applying a grouping criterion which is targeted towards separating signal components from noise components. We will provide more details in Sect. 2.1. General introductions to SSA have been provided by Golyandina et al. (2001, 2018), and Golyandina and Zhigljavsky (2020).

Traditional approaches for grouping in SSA are often subjective. These approaches typically involve selecting a group size based on expert knowledge or heuristic rules. A brief overview of some automatic grouping approaches is provided in Sect. 1.1. In the present work, we propose a new method for automatic grouping, which relies on multiple statistical hypotheses testing. Thus, the proposed method provides statistical guarantees for the resulting grouping, which is in contrast to all existing approaches that we are aware of. As separability criterion, we utilize the weighted correlation (w-correlation for short) between the signal and the noise component of the (reconstructed) time series. For optimal denoising, the population version of this w-correlation should be equal to zero, and we develop multiple tests for testing the null hypothesis of zero w-correlation for several possible groupings simultaneously; see Sect. 3 for details. A wild bootstrap procedure proposed by Hounyo (2023) is employed to approximate the null distribution of the empirical w-correlation.

1.1 Related work

Automatic grouping Generally, the sub-signals (or components) of a time series which correspond to the signal are the g leading components of the time series. In this, the term "leading" corresponds to the order of the eigenvalues of $\mathbf{X}\mathbf{X}^\top$, where $^\top$ denotes matrix transposition throughout; see the description around Equation (2.2) in Golyandina and Zhigljavsky (2020) and the explanations at the end of Section 2.1.2.3 in Golyandina and Zhigljavsky (2020). Therefore, determining the grouping in SSA corresponds to finding the group size g. If no prior information about g is available, methods for automatic grouping can be applied.

If there is no information about the rank of the signal, the w-correlation can be utilized. For example, Bilancia and Campobasso (2010) apply hierarchical clustering to the w-correlation matrix, while Alonso and Salgado (2008) utilize a k-means clustering procedure. If specific time series characteristics are of interest, one can evaluate different values of g by comparing the estimated characteristics of the corresponding

reconstructed signals with the observed characteristics. For example, if the aim of SSA is forecasting, one can apply SSA to historical data and compare the accuracy of the forecasts based on different values of g with already observed data. Then, g is chosen such that the corresponding forecast has the best accuracy; see Page 66 of Golyandina and Zhigljavsky (2020).

If the signal is assumed to have finite rank, subspace-based approaches can be used. In this case, determining g corresponds to estimating the rank of the signal. This can be done using information criteria, which was first proposed by Wax and Kailath (1985) (see Stoica and Selen, 2004, for an overview). If the signal is assumed to be a sum of complex exponentials, automatic estimation of the signal rank can be done based on the rotational invariance principle (Badeau et al. 2004; Papy et al. 2007; Roy and Kailath 1989) or the shift-invariance principle (Albert and Galarza, 2023). Estimation of the model rank using deep learning models has been proposed by Moon et al. (2021) for signals following an autoregressive moving-average (ARMA) process. An overview of the relation between subspace-based methods and SSA is given in Section 3.8 of Golyandina (2020).

Procedures based on w-correlation and subspace-based methods become unreliable if the signal and noise are not well separated (Golyandina, 2020). Furthermore, they offer no statistical guarantees that the determined group of signal components contain only true signal components of the time series. In some settings, it is important to be confident that the estimated signal does not contain any noise. Hence, in such settings statistical guarantees are of interest. The procedure to estimate g proposed in this work offers such guarantees. The proposed method is based on the w-correlation and is therefore applicable to a wide range of time series, which is in contrast to the subspace-based methods.

Note that variations of SSA exists which do not require to determine a grouping, see for example Yang and Buu (2025); Bógalo et al. (2021). However, in this work, we are interested in the classical framework of SSA which requires grouping.

Bootstrap procedures for time series When applying resampling procedures to time series, the temporal dependence of the series needs to be accounted for. One family of resampling procedures for dependent data are block-based bootstrap procedures, which resample blocks instead of single observations. Both Kunsch (1989) and Liu and Singh (1992) introduced the moving block bootstrap, a procedure which has been extended and adapted to different settings, see for example Carlstein (1986); Politis and Romano (1994); Paparoditis and Politis (2001); Andrews (2004); Paparoditis and Politis (2002). An alternative to block-based bootstrap procedures are wild bootstrap procedures for dependent data (e.g., Shao, 2010; Leucht and Neumann, 2013; Hounyo, 2023). In contrast to block-based methods, wild bootstrap methods for dependent data do not require the partitioning of the data into blocks. Instead, they utilize auxiliary variables that are dependent in a way that captures the (auto-)dependence in the original time series. Like the traditional wild bootstrap, introduced by Wu (1986) and Härdle and Mammen (1993), wild bootstrap procedures for dependent data are applicable to heterogeneously distributed data.

For overviews and comparisons of different bootstrap procedures for dependent data, we refer to Lahiri (2003) and Kreiss and Paparoditis (2011).

1.2 Main contributions

The weighted correlation coefficient, or w-correlation, is a measure that can be used for judging how well the signal and noise have been separated in SSA. A value close to zero is an indication that the number of groups used is sufficient and hence that noise and signal are well separated. To determine the number g in the grouping step of SSA, we conduct inference on the population version of the weighted correlation coefficient utilizing the wild bootstrap procedure proposed by Hounyo (2023). This procedure allows us to approximate the (analytically intractable) sampling distribution of the empirical w-correlation. We utilize this approximation to make inference on the population version of the w-correlation. To be specific, we test all possible groupings obtained by assigning the first g SSA components to the signal subspace and the remaining components to noise, and assess for all such candidate values g whether the resulting w-correlation is significantly different from zero. To account for the multiplicity of the testing problem, we control the family-wise error rate (FWER). Therefore, the probability of making at least one false rejection is bounded from above by a pre-defined significance level, which we will denote by α. In other words, we control the probability that noise is included in the reconstructed signal. This is a novelty, since established automatic grouping procedures cannot give such guarantees. Since the proposed method is based on the w-correlation, it is applicable to time series with finite or infinite rank.

1.3 Overview of the rest of the material

The rest of this material is organized as follows. A description of the SSA methodology, including notations and assumptions, is given in Sect. 2. A description of the proposed procedure is given in Sect. 3. Simulation studies and real world data analysis that illustrate the application of the proposed methodology are given in Sects. 4 and 5, respectively. Finally, we conclude in Sect. 6 with a discussion of our findings and directions for future research. An overview of the parameters introduced in Sects. 2 and 3 is given in Table 1.

2 Singular spectrum analysis

The origins of SSA can be traced back to the work of Broomhead and King (1986). SSA, as a branch of time series analysis, is a nonparametric approach that can be used for denoising a time series. Furthermore, the SSA method is capable of filtering and forecasting time series. It can be carried out as a univariate or as a multivariate method. The major goal of the SSA method is to decompose the initial ordered series (e. g., a time series) into a sum of different components which can be considered as either a trend, a periodic, a quasi-periodic (perhaps amplitude-modulated), or a noise element. We refer to Movahedifar et al. (2018); Silva et al. (2019, 2018); Hassani et al. (2018, 2019) for more details. In the following subsections, descriptions of SSA and

Table 1 Overview of the notation introduced in Sects. 2–3. The abbreviation i.i.d. stands for independent and identically distributed

Symbol	Description
$\mathbf{Y}_N = (y_1, \ldots, y_N)^\top$	(Univariate) time series of length N
$t \in \{1, \ldots, N\}$	Indicates the considered time point
$\mathbf{S} = (S_1, \ldots, S_N)^\top$	Signal component of $\mathbf{Y}_N$
$\mathbf{Z} = (Z_1, \ldots, Z_N)^\top$	Noise component of $\mathbf{Y}_N$
Singular Spectrum Analysis (SSA)	
$L \in \{2, \ldots, N/2\}$	Window length in SSA (tuning parameter)
$\mathbf{X} = \left[\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(K)}\right]$	Trajectory matrix
$K = N - L - 1$	
$\mathbf{X}^{(i)} = \left(y_i, \ldots, y_{i+L-1}\right)^\top$	L-lagged vector
$\lambda_1, \ldots, \lambda_L$	Eigenvalues of $\mathbf{X}\mathbf{X}^T$
$\mathbf{U}_1, \ldots, \mathbf{U}_L$	Eigenvectors of $\mathbf{X}\mathbf{X}^T$ corresponding to $\lambda_1, \ldots, \lambda_L$
d	Rank of the trajectory matrix
$\mathbf{X}_1 + \cdots + \mathbf{X}_d$	Singular value decomposition of the trajectory matrix
$\mathbf{X}_i, i = 1, \ldots, d$	Rank-one elementary matrix
$\mathbf{X}_I = \mathbf{X}_{i_1} + \cdots + \mathbf{X}_{i_p}$	Matrix corresponding to group I, with $I \subset \{1, \ldots, d\}$
γ	Number of disjoint subsets $I_1, \ldots, I_\gamma$
$\tilde{\mathbf{X}}_I$	Hankelized Version of $\mathbf{X}_I$

Table 1 continued

Symbol	Description
$\rho^w(\mathbf{S}, \mathbf{Z})$	(Empirical) weighted correlation (w-correlation) of $\mathbf{S}$ and $\mathbf{Z}$
$w_t = \min\{t, L, N - t + 1\}$	Weights of the w-correlation
$g \in \{1, \ldots, d\}$	Grouping index
	defines groups associated with signal- and noise component
U_{gt}	Weighted product of signal- and noise component
	corresponding to grouping index g at time t
$\mathscr{H}_g$	Null hypothesis, if true: grouping index g leads to separability
Wild Bootstrap for dependent data (WBDD)	
$b = 1, \ldots, B$	Indicates the considered bootstrap sample
$\mathbf{R}_g \equiv \mathbf{R} \in \mathbb{R}^N$	Centered time series with $R_{gt} \equiv R_t :- U_{gt} - N^{-1} \sum_{t=1}^{N} U_{gt}$
$v_n(t) \in [0, 1]$	Data-tapering window (tuning parameter)
$v : \mathbb{R} \to [0, 1]$	Weight function used to compute $v_n(t)$
$\ell \in \{1, \ldots, N\}$	Block size of WBDD (tuning parameter)
$\bar{R}_{g\ell v} \equiv \bar{R}_{\ell v}$	Tapered moving block sample mean
$(a_j)_{j \in \{1, \ldots, N-\ell+1\}}$	Sequence of i.i.d. random variables (tuning parameter)
$R_g^\star \equiv R^\star$	WBDD pseudo observation
T_g	Test statistic corresponding to $\mathscr{H}_g$
p_g^{boot}	Bootstrap p-value corresponding to $\mathscr{H}_g$
p_g^{adj}	Adjusted p-value corresponding to $\mathscr{H}_g$ using the Šidák correction

of separability concepts are presented. For more information, see Sanei and Hassani (2016).

2.1 A brief description of basic SSA

Let $\mathbf{Y}_N = (y_1, \ldots, y_N)^\top$ denote a (univariate) time series of length N. Fix an integer L, which is called window length, where $2 \le L \le N/2$. Then, basic SSA consists of the following four steps:

1. *Embedding*: This step transforms the $\mathbf{Y}_N$ into the multi-dimensional series $\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(K)}$ with vectors $\mathbf{X}^{(i)} = (y_i, \ldots, y_{i+L-1})^\top \in \mathbb{R}^L$, where $K = N - L + 1$. The vectors $\mathbf{X}^{(i)}$ are called L-lagged vectors (or, simply, lagged vectors). The embedding step includes only one parameter, namely, the window length L. The purpose of this step is to form the trajectory matrix $\mathbf{X} = \left[\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(K)}\right] \in \mathbb{R}^{L \times K}$. Depending on several criteria like complexity of the data, the purpose of the analysis of data and, in the context of prediction, the forecasting horizon, the parameter L can be chosen. Sanei and Hassani (2016) recommend that L should be 'large', but not larger than $N/2$. In order to find a suitable value of L, the concept of separability can be considered; cf. Section 2.2. By definition, an $(L \times K)$-matrix $\mathbf{Z}$ is a Hankel matrix if its elements z_{ij} are constant on the "matrix diagonals", i. e., $z_{ij} \equiv \mathfrak{z}_k$ for all (i, j) such that $i + j = k$, for any $2 \le k \le L + K$. It is worthwhile to notice that any trajectory matrix is a Hankel matrix with $\mathfrak{z}_k = y_{k-1}$, $2 \le k \le L + K$.
2. *Singular Value Decomposition (SVD)*: In this step, the trajectory matrix $\mathbf{X}$ is decomposed into a sum of rank-one elementary matrices. Let $\lambda_1, \ldots, \lambda_L$ denote the eigenvalues of $\mathbf{X}\mathbf{X}^T$, which are ordered in decreasing magnitude such that $\lambda_1 \ge \cdots \ge \lambda_L \ge 0$, and let $\mathbf{U}_1, \ldots, \mathbf{U}_L$ denote the eigenvectors of $\mathbf{X}\mathbf{X}^\top$ corresponding to these eigenvalues. Letting $d = \max\{i : \lambda_i > 0\} = \operatorname{rank}(\mathbf{X})$, the SVD of the trajectory matrix can be written as $\mathbf{X} = \mathbf{X}_1 + \cdots + \mathbf{X}_d$, where $\mathbf{X}_i = \sqrt{\lambda_i}\mathbf{U}_i\mathbf{V}_i{}^\top$ and $\mathbf{V}_i = \mathbf{X}^\top\mathbf{U}_i/\sqrt{\lambda_i}$ for $i = 1, \ldots, d$. This SVD is optimal in the sense that among all the matrices $\mathbf{X}^{(g)}$ of rank $g < d$ the matrix $\sum_{i=1}^{g} \mathbf{X}_i$ minimizes the Frobenius norm $\|\mathbf{X} - \mathbf{X}^{(g)}\|_F$ of $\mathbf{X} - \mathbf{X}^{(g)}$, which is also sometimes referred to as the L_2-norm. Note that $\|\mathbf{X}\|_F^2 = \sum_{i=1}^{d} \lambda_i$ and $\|\mathbf{X}_i\|_F^2 = \lambda_i$ for $i = 1, \ldots, d$. Thus, we can consider the ratio $\lambda_i / \sum_{i=1}^{d} \lambda_i$ as a quantitative measure of the contribution of the matrix $\mathbf{X}_i$ to the expansion $\mathbf{X} = \mathbf{X}_1 + \cdots + \mathbf{X}_d$. Accordingly, $\sum_{i=1}^{g} \lambda_i / \sum_{i=1}^{d} \lambda_i$, the ratio of the sum of the first g eigenvalues over the sum of all d non-zero eigenvalues, is a quantitative measure of the optimal approximation of the trajectory matrix by matrices of rank g. Selecting an appropriate value of g is very important. On the one hand, by selecting g smaller than the true number of non-zero eigenvalues, some parts of the signal(s) will be lost. On the other hand, if one takes g greater than the value that it should be, then noise components will be included in the reconstructed series, which one may want to avoid in certain applications.
3. *Grouping*: In this step, the set of indices $\{1, \ldots, d\}$ is partitioned into γ disjoint subsets $I_1, \ldots, I_\gamma$. Let $I = \{i_1, \ldots, i_p\} \subset \{1, \ldots, d\}$ be such a subset (of size p). Then, the matrix $\mathbf{X}_I$ corresponding to the group I is given by $\mathbf{X}_I = \mathbf{X}_{i_1} + \cdots + \mathbf{X}_{i_p}$.

For example, if $I = \{2, 3, 5\}$, then $\mathbf{X}_I = \mathbf{X}_2 + \mathbf{X}_3 + \mathbf{X}_5$. Considering the SVD of $\mathbf{X}$, the split of the set of indices $\{1, \ldots, d\}$ into the disjoint subsets $I_1, \ldots, I_\gamma$ can be represented as

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

If the original series is modeled as signal plus noise, one considers $\gamma = 2$ groups of indices, namely $I_1 = \{1, ..., g\}$ and $I_2 = \{g + 1, ..., d\}$. The group I_1 is associated with the signal component and the group I_2 with noise. At the grouping step, diagnostic tools for differentiating between noise and signal are the periodogram, pairwise scatterplots of eigenvectors, or the eigenvalue spectrum, among others; cf. Section 3.3 of Hassani (2007). Once we have selected the eigenvalues corresponding to signal and noise, respectively, we can evaluate the effectiveness of this selection via the notion of separability; see Sect. 2.2.

4. *Diagonal Averaging*: The aim of this step is to transform every obtained matrix $\mathbf{X}_{I_j}$ from the grouping step into a Hankel matrix so that these can subsequently be converted back into a (reconstructed) time series. In basic SSA, Hankelization of $\mathbf{X}_{I_j}$ is achieved via diagonal averaging of its elements over all (i, j) such that $i + j = \text{const}$. By performing the diagonal averaging of all matrix components $\mathbf{X}_{I_j}$ in the expansion of $\mathbf{X}$ from (1), we obtain another expansion: $\mathbf{X} = \tilde{\mathbf{X}}_{I_1} + ... + \tilde{\mathbf{X}}_{I_\gamma}$, where $\tilde{\mathbf{X}}_{I_j}$ is the Hankelized version of the matrix $\mathbf{X}_{I_j}$. This is equivalent to the decomposition of the initial series $\mathbf{Y}_N = (y_1, \ldots, y_N)^\top$ into a sum of γ reconstructed series, namely $y_t = \sum_{j=1}^{\gamma} \tilde{\mathbf{y}}_t(j)$ for each $t \in \{1, \ldots, N\}$, where $\tilde{\mathbf{y}}_t(j)$ corresponds to the matrix $\tilde{\mathbf{X}}_{I_j}$ for each $t \in \{1, \ldots, N\}$. In this, the operation of back-transforming a Hankel matrix into an ordered univariate series is referred to as reconstruction.

2.2 Separability

In the context of SSA, the notion of separability is employed to assess the quality of the decomposition of $\mathbf{Y}_N$ into signal components and noise components. Here, we focus on the weighted scalar product and the weighted correlation, respectively, as separability criteria. Consider a given decomposition $\mathbf{Y}_N = \mathbf{S} + \mathbf{Z}$, where $\mathbf{S} = (S_1, \ldots, S_N)^\top$ corresponds to the signal components and $\mathbf{Z} = (Z_1, \ldots, Z_N)^\top$ represents the noise component. For each $t \in \{1, \ldots, N\}$, denote by $w_t = \min\{t, L, N - t + 1\}$ the number of times that the element y_t appears in the trajectory matrix of $\mathbf{Y}_N$. Then, the weighted scalar product of $\mathbf{S}$ and $\mathbf{Z}$ is defined as

$$\langle \mathbf{S}, \mathbf{Z} \rangle_w = \sum_{t=1}^{N} w_t S_t Z_t.$$

Correspondingly, the (empirical) weighted correlation (w-correlation) of $\mathbf{S}$ and $\mathbf{Z}$ is given by

$$\rho^w(\mathbf{S}, \mathbf{Z}) = \frac{\langle \mathbf{S}, \mathbf{Z} \rangle_w}{\|\mathbf{S}\|_w \cdot \|\mathbf{Z}\|_w}, \tag{2}$$

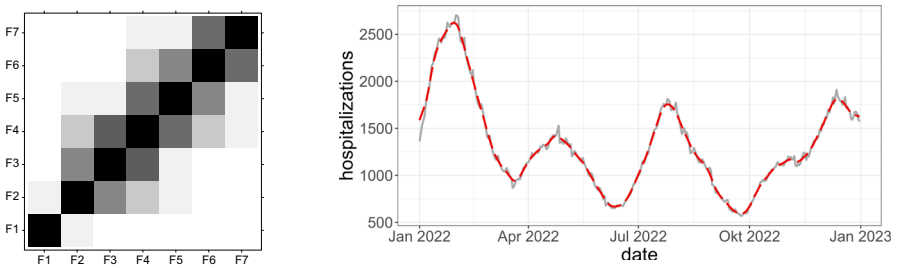

Fig. 1 Illustration of SSA results: Analysis of COVID-19 hospitalizations in Emilia Romagna during 2022. The left plot displays the matrix illustration of the w-correlations. The right plot shows the original time series (solid line) and the reconstructed signal based on the grouping according to the proposed procedure (dashed line)

where $\|\cdot\|_w$ denotes the norm induced by $\langle\cdot,\cdot\rangle_w$.

The absolute value of the w-correlation indicates the level of separability. Namely, a w-correlation close to zero means that the corresponding series are nearly w-orthogonal, and this is what one would expect in the case of a successful decomposition of $\mathbf{Y}_N$ into signal components and noise components. In contrast, a large w-correlation (in absolute value) suggests that the two series are far from being w-orthogonal. Visualizing the matrix of absolute w-correlations between series components in a grayscale format is a graphical tool to find a plausible value of g. In such a graph, low w-correlations appear as white, while w-correlations close to one are depicted in black; see the left graph in Fig. 1 for an example. This visual approach provides information about the relationships within the series components. For further details, we refer to Section 2.2.3.4 in Sanei and Hassani (2016).

Our objective in the next section is to choose a grouping with (minimal) grouping index $\hat{g} \in \{1, \ldots, d\}$ that leads to a separation of signal and noise in terms of a statistical criterion regarding the w-correlation. Namely, we propose to test for each candidate value of g the resulting w-correlation of reconstructed signal and reconstructed noise for being equal to zero. The null distribution of the tests will be approximated based on the wild bootstrap approach for dependent data (WBDD) by Hounyo (2023).

Since the w-correlation is not necessarily monotone in g, separability may alternate between poor and adequate levels as g increases. While a common strategy is to select the smallest g with acceptable separability, this may be suboptimal if early components do not yet form a reliable signal subspace. We therefore choose the first index at which separability occurs and remains stable for larger indices. We note that alternative selection criteria for $\hat{g}$ are possible once the non-separable groupings have been identified. Our proposed methodology can easily be adapted to such alternative criteria.

3 Identification of the grouping index

Suppose that we perform SSA on a time series $\mathbf{Y}_N = (y_1, \ldots, y_N)^\top$ with a certain window length $L \leq N/2$, obtaining $d \leq L$ non-null singular values. For any grouping

index $g \in \{1, \ldots, d\}$ we say that the main signal and noise are separable if the w-correlation given in Eq. (2) is close to zero, i.e., if the numerator

$$\sum_{t=1}^{N} U_{gt}, \qquad U_{gt} = w_t S_{gt} Z_{gt} \tag{3}$$

is close to zero, where (with slight abuse of notation) $\mathbf{S}_g$ and $\mathbf{Z}_g$ denote the reconstructed signal and noise components, respectively, for grouping index g. Our main goal is finding the best grouping, i.e., the smallest index $\hat{g}$ such that the main signal and noise are separable for all $g \in \{\hat{g}, \ldots, d\}$. Notice that we always have separability when $g = d$, as we obtain $\mathbf{Y} = \mathbf{S}_d$ and $\mathbf{Z}_d = \mathbf{0}$, and so

$$U_{dt} = 0 \text{ for all } t \in \{1, \ldots, N\} \quad \Longrightarrow \quad \sum_{t=1}^{N} U_{dt} = 0.$$

For any other grouping $g \in \{1, \ldots, d-1\}$, we are interested in testing

$$\mathscr{H}_g : \mathbb{E}\left[\sum_{t=1}^{N} U_{gt}\right] = 0, \tag{4}$$

against a two-sided alternative, where the expectation operator $\mathbb{E}$ refers to the true, but unknown data-generating process. If $\mathscr{H}_g$ is true, then the grouping with index g leads to separability (on the population level). Otherwise, the resulting main signal and noise are not separable, and we should use a larger grouping index. Therefore, we simultaneously test the null hypotheses $\left(\mathscr{H}_g : g \in \{1, \ldots, d-1\}\right)$ with control of the family-wise error rate (FWER). Then, we define the selected grouping index as

$$\hat{g} = \max\{g \in \{1 \ldots, d-1\} \ : \ \mathscr{H}_g \text{ is rejected}\} + 1 \tag{5}$$

if the maximum exists, and $\hat{g} = 1$ otherwise.

For given g, we conduct a bootstrap hypothesis test for $\mathscr{H}_g$ based on wild bootstrapping, accounting for the dependence among $U_{g1}, \ldots, U_{gN}$. To this end, we first give a short overview of general bootstrap hypothesis tests. We then present the procedure of generating wild bootstrap samples for testing $\mathscr{H}_g$, as well as the computation of the test statistic and corresponding p-value. Following this, we explain how to examine all null hypotheses $\mathscr{H}_1, \ldots, \mathscr{H}_d$ simultaneously in order to determine the selected grouping index $\hat{g}$.

3.1 Bootstrap hypothesis tests

To test a hypothesis using bootstrapping, B bootstrap samples, for which the null hypothesis is true and which consist of $\mathscr{N}$ pseudo-observations each, are generated. For each bootstrap sample $b = 1, \ldots, B$ the test statistic T^b is computed (Mackinnon,

2009). Then, the bootstrap p-value for a two-sided null hypothesis is defined as

$$p^{boot} = \frac{\sum_{b=1}^{B} \mathbf{1}(|T^b| \geq |T^{obs}|) + 1}{B+1}, \tag{6}$$

where T^{obs} corresponds to the test statistic based on the observed sample and $\mathbf{1}(\cdot)$ denotes the indicator function. Thus, the bootstrap p-value p^{boot} represents the proportion of test statistics that are at least as extreme as T^{obs} amongst all test statistics (Mackinnon, 2009). This accentuates the requirement that the bootstrap samples are obtained such that they follow the null hypothesis. It also indicates that we do not need to make any assumptions about the distribution of the test statistic T^{obs}. However, it is important that the bootstrap data generating process is valid, i.e., in our case accounts for the dependency and non-stationarity in the data. Thus, we use the wild bootstrapping procedure for dependent data (WBDD) introduced by Hounyo (2023).

3.2 Wild bootstrap for SSA

The framework of the WBDD permits general dependence conditions and data heterogeneity. WBDD combines the ideas of wild bootstraps and tapered block bootstrap.

In the context of SSA, we utilize WBDD to approximate the sampling distribution of $\mathbf{U}_g = (U_{g1}, \ldots, U_{gN})^\top$ as in (3). To this end, we consider the centered time series $\mathbf{R}$ with $R_t := U_{gt} - \bar{\mathbf{U}}_g$ for $\bar{\mathbf{U}}_g = N^{-1} \sum_{t=1}^{N} U_{gt}$. A sequence of data-tapering windows $v_n(t) \in [0, 1]$, $v_n(t) = 0$ for $t \notin \{1, \ldots, n\}$ is used as weights. Denote by $||v_n||_1$ and $||v_n||_2$ the L_1 and L_2-norm, respectively.

To generate a WBDD sample, we first compute the tapered moving block sample mean

$$\bar{R}_{\ell,v} = \frac{1}{Q} \sum_{j=1}^{Q} \sum_{i=1}^{\ell} \frac{v_\ell(i)}{||v_\ell||_1} R_{i+j-1},$$

with $\ell \in \{1, \ldots, N\}$ denoting the (fixed) block size and $Q = N - \ell + 1$. A WBDD pseudo observation $R^\star$ is is computed by

$$R_t^\star = (R_t - \bar{R}_{\ell,v})\eta_t + \bar{R},$$

with $\bar{R} = N^{-1} \sum_{t=1}^{N} R_t$ and

$$\eta_t := -\sum_{j=1}^{Q} \frac{v_\ell(t-j+1)}{||v_\ell||_2} \sqrt{\ell} a_j,$$

where $(a_j)_{j \in \{1,\ldots,Q\}}$ denotes a sequence of i.i.d. random variables.

As in Hounyo (2023), we consider

$$v_n(t) = v((t - 0.5)/n), \tag{7}$$

where $v(\cdot)$ is a single function $v : \mathbb{R} \to [0, 1]$. The function $v(\cdot)$ should meet three assumptions (Hounyo, 2023). To maintain a self-contained presentation, we repeat the assumptions here.

Assumption 1 We have $v(t) \in [0, 1]$ for all $t \in \mathbb{R}$, $v(t) = 0$ if $t \notin [0, 1]$, and $v(t) > 0$ for t in a neighborhood of $1/2$.

Assumption 2 The function $v(t)$ is symmetric around $t = 1/2$ and nondecreasing for $t \in [0, 1/2]$.

Assumption 3 The self-convolution is twice continuously differentiable at the point $t = 0$, where $v * v(t) = \int_{-1}^{1} v(x)v(x + |t|)dx$.

Additional assumptions about the sequence $(a_j)_{j\in\{1,\ldots,Q\}}$ are made (Hounyo, 2023). Once again, we repeat the assumptions here.

Assumption 4 The sequence of random variables (a_j), $j = 1, \ldots, Q = N - \ell + 1$ is independent of the original observed sample $\mathbf{U}_g$. Moreover, it is independent and identically distributed and satisfies the following regularity conditions: $\mathbb{E}(a) = 0$, $Var(\sqrt{\ell}a) = \ell\mathbb{E}(a)^2 \to 1$ and for some $\delta > 0$, $\ell\mathbb{E}(|a|^{2+\delta}) \to C_\delta < \infty$ as $N \to \infty$, $\ell \to \infty$ such that $\ell/N = o(1)$, where C_δ is a nonrandom constant.

Examples for the function $v(\cdot)$ as well as examples for the generation of the sequence $(a_j)_{j\in\{1,\ldots,Q\}}$ are given in Hounyo (2023). Some of these examples will be utilized in the simulation study in Sect. 4.

In the next section, we present a wild bootstrap hypothesis test for $\mathscr{H}_g$ for a fixed grouping index g. Subsequently, we will study all null hypotheses $\mathscr{H}_1, \ldots, \mathscr{H}_{d-1}$ simultaneously to find the selected grouping index $\hat{g}$ defined in (5).

3.3 Fixed grouping

Fix a grouping $g \in \{1, \ldots, d-1\}$. To test the null hypothesis $\mathscr{H}_g$ given in (4) against its two-sided alternative at significance level $\alpha \in (0, 1)$, we define the (observed) test statistic as

$$T_g^{obs} = \sum_{t=1}^{N} U_{gt}.$$

Under the assumption that the null hypothesis is true for the bootstrap samples, the bootstrap test statistics $\left(T_g^b : 1 \le b \le B\right)$ can be computed analogously, using (WBDD) pseudo-observations $\mathbf{U}_g^b$, $b = 1, \ldots, B$ instead of the observations $\mathbf{U}_g$. To ensure that the bootstrap samples are obtained under the null (where the expected value of the sum of the observed values is equal to zero), we center the observations around zero before generating bootstrap pseudo-observations.

Finally, the bootstrap p-values are given by

$$p_g^{boot} = \frac{1}{B+1}\left(\sum_{b=1}^{B} \mathbf{1}\{|T_g^b| \ge |T_g^{obs}|\} + 1\right) \qquad (g = 1, \ldots, d-1),$$

in accordance with Eq. (6).

3.4 All groupings

To compute the selected grouping index $\hat{g}$ defined in Eq. (5), we need to test all hypotheses $\mathscr{H}_1, \ldots, \mathscr{H}_{d-1}$ simultaneously. In this, we aim to control the probability of rejecting at least one true null hypothesis, which is called the FWER. Different procedures can be used to control the FWER at level α. If we assume that the p-values given in Eq. (6) have positive dependency, we may use the Šidák correction. In this case, we adjust the observed p-values according to the Šidák correction, such that $p^{adj} = 1 - (1 - p^{boot})^{d-1}$. The hypothesis $\mathscr{H}_g$ is rejected if and only if the corresponding adjusted p-value is at most α. If we do not make any assumption about the dependency structure of the p-values, we may use the Bonferroni-Holm procedure. This procedure requires ordering the p-values in an ascending manner, such that $p_{(1)} \leq p_{(2)} \leq \ldots, \leq p_{(d-1)}$. The hypotheses corresponding to $p_{(1)}, \ldots, p_{(i^\star)}$ are rejected, where $i^\star = \max\{i \in \{1, \ldots, d-1\} : p_{(j)} \leq \alpha/(d-j) \text{ for all } j = 1, \ldots, i\}$. If $i^\star$ does not exist, no rejection is made. For more information, see Dickhaus (2014).

Controlling the FWER ensures that the probability of having one or more false rejections (where we falsely state that a certain grouping does not lead to separability on the population level) is at most α. In the following simulation study, we compare both multiple testing procedures and investigate the influence of the WBDD parameters on the performance of the proposed procedure.

4 Simulation study

In this section, we aim to assess the behavior of the proposed method described in Sect. 3 by applying it to the following three distinct time series.

Example 1 For simulation studies, we have generated three times series ($y \in \mathbb{R}^t : t = 1, \ldots, N = 50$) according to a "signal plus noise" model of the form $y(t) = f(t) + \varepsilon(t)$, $t = 1, \ldots, N$. For the signal part f, the following three choices have been considered.

1. $f_1(t) = \sin(2\pi t/3)$
2. $f_2(t) = \exp(0.2t)$
3. $f_3(t) = 0.7 \cdot \cos(\pi t/2) + 0.5 \cdot \cos(\pi t/3)$.

It is known that the true grouping indices are given by $g_1^\star = 2$, $g_2^\star = 1$, $g_3^\star = 4$ for f_1, f_2 and f_3 respectively (Golyandina et al. (2015, Example 3) Golyandina (2020, p.44)). In all three cases, the noise terms ($\varepsilon(t), t = 1, \ldots, N = 50$) are independent and identically distributed with $\varepsilon(1) \sim N(0, Var(f)/SNR)$, with signal to noise ratio $SNR \in \{0.5, 2, 5\}$.

The embedding step as well as the SVD step SSA pipeline have been carried out using the R package "Rssa" (Golyandina et al. 2018). We have set the window length to $L = N/2$. To compute the grouping parameter as described above, we have employed the wild bootstrap method with $B = 1000$ bootstrap replications. Implementation of

the WBDD requires fixing the block size ℓ, the weight function $v(t)$ as in Eq. (7) and the sequence of random variables $(a_j)_{j\in\{1,\ldots,Q\}}$. In accordance with Hounyo (2023), we have set $\ell = N^{1/5}$. We have considered the two weight functions

$$v^{[1]}(t) = \begin{cases} t & \text{if } t \in (0, 0.5] \\ 1-t & \text{if } t \in (0.5, 1) \\ 0 & \text{else,} \end{cases}$$

and

$$v^{[2]}(t) = \begin{cases} t/0.43 & \text{if } t \in [0, 0.43] \\ 1 & \text{if } t \in [0.43, 1-0.43] \\ (1-t/0.43) & \text{if } t \in [1-0.43, 1] \\ 0 & \text{if } t \notin [0, 1], \end{cases}$$

where $v^{[2]}(t)$ corresponds to the example given in Hounyo (2023) with $c = 0.43$. To investigate the influence of the external random sequence used in the WBDD on the estimated grouping index $\hat{g}$, we have utilized four different random sequences, namely those introduced in Examples 2.3−2.5 and 2.7 in Hounyo (2023). We denote the sequences by $a^{[i]}$, $i \in \{3, 4, 5, 7\}$, where i indicates the corresponding example in Hounyo (2023). The grouping index $\hat{g}$ has then been obtained as described in Eq. (5). To account for multiplicity, we have used the Šidák and Bonferroni-Holm correction with a significance level of $\alpha = 0.1$.

We have compared the proposed procedure with the method for automatic grouping described by Bilancia and Campobasso (2010) as implemented by Golyandina et al. (2018), Algorithm 2.15. This approach uses hierarchical clustering applied to the w-correlation matrix. The procedure groups the indices corresponding to the elementary matrices. We have set the grouping parameter returned by the hierarchical clustering approach to be the largest index in the first group. Simulation results are based on 500 Monte-Carlo iterations.

In the simulation study, the choice of the weight function did not influence the results of the proposed procedure but the choice of the random sequence did. Furthermore, no differences in the results between utilizing the Šidák or the Bonferroni-Holm correction have occurred.

For signal f_1 and f_3, the proportion of iterations for which $\hat{g} > g^\star$ has been below $\alpha = 0.1$ for all signal-to-noise ratios, both considered weight functions and all random sequences. Indeed, only for $SNR = 0.5$ did this proportion differ from zero. When comparing the observed FWER of the proposed procedure in terms of the four different random sequences, the observed FWER has been the largest for $a^{[3]}$ and $a^{[5]}$. For signal f_2 the proportion of iterations for which $\hat{g} > g^\star$ is greater than zero. In this case, the observed FWER has been lower for large SNR than for low SNR. For signal f_2, the observed FWER is below $\alpha = 0.1$ in case of $SNR \in \{2, 5\}$. However, it has been greater than $\alpha = 0.1$ for $SNR = 0.5$, with an observed maximum of 0.208 for $a^{[4]}$. Results for the grouping index returned by the hierarchical clustering approach are similar: The proportion of iterations with $\hat{g} > g^\star$ has differed from zero

Table 2 The average values of the selected grouping index $\hat{g}$ and the corresponding empirical standard deviations (in brackets) for different signals and varying SNR. The signals are defined in Example 1. The proposed procedure has been based on the wild bootstrap for dependent data by Hounyo (2023) using four different random sequences $a^{[i]}$, $i \in \{3, 4, 5, 7\}$, as given in Examples 2.3−2.7 in Hounyo (2023). Additionally, the average grouping parameters based on the hierarchical clustering approach (HC) by Bilancia and Campobasso (2010) and the corresponding standard deviations are reported. The true grouping index $g^\star$ corresponds to $g_1^\star = 2$ for signal f_1, $g_2^\star = 1$ for signal f_2 and $g_3^\star = 4$ for signal f_3. Results are based on 500 iterations

$a^{[i]}$	Signal f_1 $SNR = 0.5$	$SNR = 2$	$SNR = 5$	Signal f_2 $SNR = 0.5$	$SNR = 2$	$SNR = 5$
$i = 3$	1.244	1.942	2	1.622	1.180	1.036
	(± 0.6109)	(± 0.2340)	(± 0)	(± 1.5913)	(± 1.0798)	(± 0.5722)
$i = 4$	1.286	1.962	2	1.682	1.230	1.086
	(± 0.7247)	(± 0.1914)	(± 0)	(± 1.8318)	(± 1.2426)	(± 0.8878)
$i = 5$	1.210	1.914	2	1.528	1.174	1.038
	(± 0.5644)	(± 0.2806)	(± 0)	(± 1.5550)	(± 1.0854)	(± 0.5264)
$i = 7$	1.248	1.898	2	1.486	1.156	1.040
	(± 0.6027)	(± 0.3029)	(± 0)	(± 1.5014)	(± 1.0649)	(± 0.6318)
HC	2.09	2	2	1.774	1.008	1
	(± 0.5277)	(± 0)	(± 0)	(± 2.0812)	(± 0.0892)	(± 0)

$a^{[i]}$	Signal f_3 $SNR = 0.5$	$SNR = 2$	$SNR = 5$
$i = 3$	1.312	1.690	1.970
	(± 0.7615)	(± 0.5162)	(±0.3865)
$i = 4$	1.334	1.750	1.984
	(± 0.7235)	(± 0.5178)	(± 0.3795)
$i = 5$	1.288	1.652	1.908
	(± 0.9267)	(± 0.5248)	(± 0.4288)
$i = 7$	1.290	1.670	1.946
	(± 0.9185)	(± 0.5307)	(± 0.4092)
HC	2.358	2.212	2.124
	(± 1.0138)	(± 0.6163)	(± 0.4828)

for $SNR = 0.5$ for all signals and for $SNR = 2$ in the case of signal f_2. For f_2 and $SNR = 0.5$, the proportion has exceeded $\alpha = 0.1$, with a value of 0.272.

Next, we comment on the accuracy of the proposed procedure. We say that the procedure is accurate if $\hat{g}$ is close to the true grouping index $g^\star$. Table 2 displays the observed average and standard deviation of the values of the grouping index $\hat{g}$ based on the different random sequences $a^{[i]}$ as well as $\hat{g}$ based on the hierarchical clustering approach for the considered scenarios. We first focus on the results of the proposed procedure. Generally, accuracy has increased as the SNR increases. Furthermore, utilizing the random sequence $a^{[4]}$ in the WBDD has lead to the most accurate average grouping indices for signals f_1 and f_3. Regarding the standard deviation, utilizing $a^{[4]}$ has lead to the largest standard deviation for signal f_2, for the other examples, $a^{[7]}$ and

$a^{[5]}$ have lead to the largest standard deviations in most cases. Generally, the proposed procedure has been most accurate for signal f_1 and it has been least accurate for signal f_3. For signal f_3, the signal components predominantly correspond to seasonality. Thus, the proposed method appears to be somewhat insensitive to seasonal components of the signal. The automatic grouping approach based on hierarchical clustering has been either as accurate as or more accurate than the proposed procedure in most considered scenarios. The only exception has occurred for signal f_2 and $SNR = 0.5$. The proposed procedure tends to be slightly more conservative (i. e., tending to smaller values of the selected grouping index), as it provides the theoretical guarantee that the probability of including noise in the reconstructed signal remains below α, at least asymptotically as N tends to infinity. However, for small SNR (i.e. $SNR = 0.5$) and small $g^\star$, the proposed procedure overestimated $g^\star$ in more than $\alpha \cdot 100\%$ of the cases in our simulation study.

5 Real-world data analysis

This section demonstrates the application of the proposed procedure to disease surveillance data. A key task in this context is the extraction of underlying trends, as they provide a reliable picture of disease dynamics over time. However, raw data often contain substantial noise arising from reporting practices, calendar effects, and other short-term variations. Such noise can mask genuine patterns, lead to misleading interpretations, and affect forecasting. In contrast, isolating true signals provides a clear representation of disease progression and improves the capacity of the surveillance system to support public health decisions.

We have considered the daily series of Coronavirus Disease 2019 (COVID-19) hospitalizations recorded in the Emilia-Romagna region of Italy during 2022. The data were published by the Italian Civil Protection Department and are available at https://github.com/pcm-dpc/COVID-19. The time series consists of $N = 365$ observations, which exhibit strong day-of-week effects, largely due to reporting delays over weekends that are subsequently corrected on Mondays and Tuesdays. We have applied our proposed procedure to extract the underlying trend of hospitalizations, in order to describe the evolution of the epidemic.

For the procedure to be valid, the window length L must be fixed in advance. This parameter specifies the number of consecutive observations that are grouped in the embedding step of SSA and should approximate the local state of the underlying process (Hassani, 2007). Its value should be proportional to eventual periodicities of the data and, for time series with complex structure, not too large in order to avoid an undesirable decomposition of the components of interest (Golyandina et al. 2001). Moreover, larger values of L increase the number of tested hypotheses in our framework, producing generally higher adjusted p-values and thus a more conservative procedure, with the risk that the selected grouping $\hat{g}$ underestimates the optimal grouping index. Given the strong weekly structure of the series, we have set $L = 7$, consistently with previous applications of SSA to COVID-19 data (Alharbi, 2021). Furthermore, we have used the random sequence $a^{[4]}$, as given in Example 2.4 in Hounyo (2023), and the Bonferroni-Holm correction with $\alpha = 0.1$. In addition, we

have computed the grouping parameter using the hierarchical cluster approach by Bilancia and Campobasso (2010).

Both procedures consistently identify the same grouping index $\hat{g} = 1$. The w-correlation matrix (Fig. 1, left) further supports this choice, as it shows separability among the components. Finally, the reconstructed signal (Fig. 1, right) demonstrates a clear and accurate representation of the underlying structure, confirming the quality and reliability of the decomposition.

6 Conclusion and future work

In this work, we have derived a procedure to identify the grouping index for the separation of a time series into signal and noise in the context of SSA. In contrast to existing methods, the proposed procedure offers statistical guarantees in terms of a confidence statement regarding that no noise is included in the reconstructed signal. We assume separability if the w-correlation is close to zero and utilize the wild bootstrap procedure proposed by Hounyo (2023) to approximate the null distribution of the empirical w-correlation. The grouping index is determined by testing the null hypothesis of separability (on the population level) for all possible grouping indices and defining the selected grouping index $\hat{g}$ as the minimal index for which the null hypotheses cannot be rejected for all $g \in \{\hat{g}, \ldots, d-1\}$. To account for the multiplicity of this approach, the Bonferroni-Holm correction can be used to control the FWER. Our criterion (5) for selecting $\hat{g}$ is related to fixed-sequence multiple testing. Namely, we can start testing with $g = d - 1$ and iteratively decrease g until the maximum in (5) is found. However, our stopping rule is different from traditional fixed sequence multiple tests as discussed, e. g., in Section 3.1.2.2 of Dickhaus (2014). Namely, we stop as soon as a rejection takes place and not as soon as a non-rejection takes place. Therefore, it is (as far as we can see) not possible to re-cycle the significance level from one test step to the next. However, performing the individual tests in fixed sequence may reduce the computational efforts in comparison to "blindly" performing all tests at once, at least in cases where p-values do not need to be combined (in terms of their order statistics, for example).

We have investigated the performance of the proposed method in a simulation study. The results indicate that the proposed procedure returns accurate grouping indices if the signal is not complex, i.e., the true g is low. Specifically, the method appears to be insensitive to seasonal components. The accuracy of the proposed procedure might increase for complex signals if a different window length is chosen. As pointed out by Golyandina et al. (2018), shorter window lengths might capture complex time series structures better. For small signal to noise ratio in connection with a small samples size, the empirical FWER of the proposed procedure exceeded the target level α in our simulation study. While the main reason for this exceedance most likely is the asymptotic nature of the bootstrap approximation of the null distributions of the test statistics, it may be worthwhile to study the influence of tuning parameter selection (see Table 1 for an overview of these tuning parameters) on the realized FWER of the proposed procedure.

Furthermore, the proposed procedure has been applied to disease surveillance data, i.e., the daily count of COVID-19 hospitalizations. Results were consistent with established methodologies, including the grouping method of investigating the plot of the w-correlation. Since the practitioner does not need to choose between multiple possible groupings, our proposal might increase the reproducibility of SSA studies. In addition, it provides statistical guarantees that trends or periodicity observed in the reconstructed signal likely do not include any noise.

This study opens avenues for future investigations. Further exploration could delve into refining the parameterization of the SSA method, investigating the impact of different parameter choices beyond g, and expanding the evaluation across diverse datasets. Moreover, extending this method's application to practical scenarios and different types of time series data would be valuable. Additionally, exploring variations or enhancements of the SSA method could lead to improved accuracy and broader applicability in various domains.

Acknowledgements Friederike Preusse gratefully acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)- project number 281474342. Anna Vesely acknowledges partial financial support by the Deutsche Forschungsgemeinschaft (DFG) via Grant No. DI 1723/5-3, and from the Italian Complementary National Plan PNC-I.1 "Research initiatives for innovative technologies and pathways in the health and welfare sector" D.D. 931 of 06/06/2022, "DARE - DigitAl lifelong pReVEntion" initiative, code PNC0000002, CUP: B53C22006450001. We are grateful to Daniel Ochieng for his contributions to a previous version of this manuscript.

Author contributions MM and AV developed the methodology. TD recommended the usage of the wild bootstrap and wrote the introduction. AV and FP implemented the methodology and authored the corresponding sections. MM and FP performed the simulations. Application to real-world data was carried out by AV. All authors proofread and approved the final manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL.

Data availability The data used in this article, provided and published by the Italian Civil Protection Department, are available at https://github.com/pcm-dpc/COVID-19. The code for the simulation and its analysis is available upon request from the authors.

References

Albert, R., Galarza, C. (2023). A constraint optimization problem for model order estimation. *Signal Processing, 210*(109), 092. https://doi.org/10.1016/j.sigpro.2023.109092

Alharbi, N. (2021). Forecasting the COVID-19 Pandemic in Saudi Arabia Using a Modified Singular Spectrum Analysis Approach: Model Development and Data Analysis. *Journal of Medical Internet Research, 2*(1), e21,044. https://doi.org/10.2196/21044

Alonso, F., Salgado, D. (2008). Analysis of the structure of vibration signals for tool wear detection. *Mechanical Systems and Signal Processing, 22*(3), 735–748. https://doi.org/10.1016/j.ymssp.2007.09.012

Andrews, D. W. K. (2004). The Block-Block Bootstrap: Improved Asymptotic Refinements. *Econometrica, 72*(3), 673–700. https://doi.org/10.1111/j.1468-0262.2004.00509.x

Aydın, S., Saraoğlu, H. M., Kara, S. (2011). Singular Spectrum Analysis of Sleep EEG in Insomnia. *Journal of Medical Systems, 35*(4), 457–461. https://doi.org/10.1007/s10916-009-9381-7

Badeau, R., David, B., Richard, G. (2004). Selecting the modeling order for the ESPRIT high resolution method: an alternative approach. In: 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2, II–1025, https://doi.org/10.1109/ICASSP.2004.1326435

Bilancia, M., Campobasso, F. (2010). Airborne Particulate Matter and Adverse Health Events: Robust Estimation of Timescale Effects. In H. Locarek-Junge, C. Weihs (Eds.), *Classification as a Tool for Research* (pp. 481–489). Heidelberg: Springer.

Bógalo, J., Poncela, P., Senra, E. (2021). Circulant singular spectrum analysis: A new automated procedure for signal extraction. *Signal Processing, 179*(107), 824. https://doi.org/10.1016/j.sigpro.2020.107824

Broomhead, D., King, G. P. (1986). Extracting qualitative dynamics from experimental data. *Physica D: Nonlinear Phenomena, 20*(2), 217–236. https://doi.org/10.1016/0167-2789(86)90031-X

Carlstein, E. (1986). The Use of Subseries Values for Estimating the Variance of a General Statistic from a Stationary Sequence. *The Annals of Statistics, 14*(3), 1171–1179. https://doi.org/10.1214/aos/1176350057

Chao, S. H., Loh, C. H. (2014). Application of singular spectrum analysis to structural monitoring and damage diagnosis of bridges. *Structure and Infrastructure Engineering, 10*(6), 708–727. https://doi.org/10.1080/15732479.2012.758643

Chen, Q., van Dam, T., Sneeuw, N., Collilieux, X., Weigelt, M., Rebischung, P. (2013). Singular spectrum analysis for modeling seasonal signals from GPS time series. *Journal of Geodynamics, 72*, 25–35. https://doi.org/10.1016/j.jog.2013.05.005

Dickhaus, T. (2014). *Simultaneous statistical inference with applications in the Life Sciences*. Heidelberg: Springer.. https://doi.org/10.1007/978-3-642-45182-9

Ghodsi, M., Yarmohammadi, M. (2014). Exchange rate forecasting with optimum singular spectrum analysis. *Journal of Systems Science and Complexity, 27*(1), 47–55. https://doi.org/10.1007/s11424-014-3303-6

Golyandina, N. (2020). Particularities and commonalities of singular spectrum analysis as a method of time series analysis and signal processing. *WIREs Computational Statistics, 12*(4), e1487. https://doi.org/10.1002/wics.1487

Golyandina, N., Zhigljavsky, A. (2020). *Singular Spectrum Analysis for Time Series*. (2nd ed.). Heidelberg: Springer. https://doi.org/10.1007/978-3-662-62436-4

Golyandina, N., Nekrutkin, V., Zhigljavsky, A. (2001). *Analysis of Time Series Structure: SSA and Related Techniques,* (1st edn.). Boca Raton: Chapman & Hall/CRC. https://doi.org/10.1201/9780367801687

Golyandina, N., Korobeynikov, A., Shlemov, A., Usevich, K. (2015). Multivariate and 2D Extensions of Singular Spectrum Analysis with the Rssa Package. *Journal of Statistical Software*, 67(2), 1–78, https://doi.org/10.18637/jss.v067.i02

Golyandina, N., Korobeynikov, A., Zhigljavsky, A. (2018). *Singular Spectrum Analysis with R, 1st edn. Use R!*, Heidelberg: Springer. https://doi.org/10.1007/978-3-662-57380-8

Härdle, W., Mammen, E. (1993). Comparing Nonparametric Versus Parametric Regression Fits. *The Annals of Statistics, 21*(4), 1926–1947. https://doi.org/10.1214/aos/1176349403

Hassani, H. (2007). Singular Spectrum Analysis: Methodology and Comparison. *Journal of Data Science, 5*(2), 239–257. https://doi.org/10.6339/JDS.2007.05(2).396

Hassani, H., Webster, A., Silva, E. S., Heravi, S. (2015). Forecasting U.S. Tourist arrivals using optimal Singular Spectrum Analysis. *Tourism Management, 46*, 322–335. https://doi.org/10.1016/j.tourman.2014.07.004

Hassani, H., Silva, E. S., Gupta, R., Das, S. (2018). Predicting global temperature anomaly: A definitive investigation using an ensemble of twelve competing forecasting models. *Physica A: Statistical Mechanics and its Applications, 509*, 121–139. https://doi.org/10.1016/j.physa.2018.05.147

Hassani, H., Rua, A., Silva, E. S., Thomakos, D. (2019). Monthly forecasting of GDP with mixed-frequency multivariate singular spectrum analysis. *International Journal of Forecasting, 35*(4), 1263–1272. https://doi.org/10.1016/j.ijforecast.2019.03.021

Hou, Z., Wen, G., Tang, P., Cheng, G. (2014). Periodicity of Carbon Element Distribution Along Casting Direction in Continuous-Casting Billet by Using Singular Spectrum Analysis. *Metallurgical and Materials Transactions B, 45*(5), 1817–1826. https://doi.org/10.1007/s11663-014-0103-2

Hounyo, U. (2023). A wild bootstrap for dependent data. *Econometric Theory, 39*(2), 264–289. https://doi.org/10.1017/S0266466621000487

Köhler, T., Lorenz, D. (2005). A comparison of denoising methods for one dimensional time series unpublished.

Kreiss, J. P., Paparoditis, E. (2011). Bootstrap methods for dependent data: A review. *Journal of the Korean Statistical Society, 40*(4), 357–378. https://doi.org/10.1016/j.jkss.2011.08.009

Kunsch, H. R. (1989). The Jackknife and the Bootstrap for General Stationary Observations. *The Annals of Statistics, 17*(3), 1217–1241. https://doi.org/10.1214/aos/1176347265

Lahiri, S. N. (2003). *Resampling Methods for Dependent Data*. *Springer Series in Statistics*, New York: Springer https://doi.org/10.1007/978-1-4757-3803-2

Le Bail, K., Gipson, J. M., MacMillan, D. S. (2014). Quantifying the Correlation Between the MEI and LOD Variations by Decomposing LOD with Singular Spectrum Analysis. In C. Rizos, P. Willis (Eds.), *Earth on the Edge: Science for a Sustainable Planet* (pp. 473–477). Heidelberg: Springer.

Leucht, A., Neumann, M. H. (2013). Dependent wild bootstrap for degenerate U- and V-statistics. *Journal of Multivariate Analysis, 117*, 257–280. https://doi.org/10.1016/j.jmva.2013.03.003

Liu, K., Law, S., Xia, Y., Zhu, X. (2014). Singular spectrum analysis for enhancing the sensitivity in structural damage detection. *Journal of Sound and Vibration, 333*(2), 392–417. https://doi.org/10.1016/j.jsv.2013.09.027

Liu, R. Y., Singh, K. (1992). Moving blocks Jackknife and bootstrap capture weak dependence. In R. L. L. Billard (Eds.), *Exploring the Limits of Bootstrap* (pp. 225–248). New York: Wiley Series in Probability and Statistics, Wiley.

MacKinnon, J. G. (2009). Bootstrap Hypothesis Testing. In: D. A. Belsley and E. J. Kontoghiorghes (Eds.) *Handbook of Computational Econometrics*, chap 6, 183–213, Chichester, UK: John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470748916.ch6

Moon, J., Hossain, M. B., Chon, K. H. (2021). AR and ARMA model order selection for time-series modeling with ImageNet classification. *Signal Processing, 183*(108), 026. https://doi.org/10.1016/j.sigpro.2021.108026

Movahedifar, M., Yarmohammadi, M., Hassani, H. (2018). Bicoid signal extraction: Another powerful approach. *Mathematical Biosciences, 303*, 52–61. https://doi.org/10.1016/j.mbs.2018.06.002

Movahedifar, M., Hassani, H., Yarmohammadi, M., Kalantari, M., RG. (2022). A robust approach for outlier imputation: Singular spectrum decomposition. *Communications in Statistics: Case Studies, Data Analysis and Applications,8*(2), 234–250. https://doi.org/10.1080/23737484.2021.2017810

Muruganatham, B., Sanjith, M., Krishnakumar, B., Satya Murty, S. (2013). Roller element bearing fault diagnosis using singular spectrum analysis. *Mechanical Systems and Signal Processing, 35*(1), 150–166. https://doi.org/10.1016/j.ymssp.2012.08.019

Oliveri, P., Malegori, C., Simonetti, R., Casale, M. (2019). The impact of signal pre-processing on the final interpretation of analytical outcomes – A tutorial. *Analytica Chimica Acta, 1058*, 9–17. https://doi.org/10.1016/j.aca.2018.10.055

Paparoditis, E., Politis, D. N. (2001). Tapered block bootstrap. *Biometrika, 88*(4), 1105–1119.

Paparoditis, E., Politis, D. N. (2002). Local block bootstrap. *Comptes Rendus Mathematique, 335*(11), 959–962. https://doi.org/10.1016/S1631-073X(02)02578-5

Papy, J. M., De Lathauwer, L., Van Huffel, S. (2007). A Shift Invariance-Based Order-Selection Technique for Exponential Data Modelling. *IEEE Signal Processing Letters, 14*(7), 473–476. https://doi.org/10.1109/LSP.2006.891324

Politis, D. N., Romano, J. P. (1994). The Stationary Bootstrap. *Journal of the American Statistical Association, 89*(428), 1303–1313. https://doi.org/10.2307/2290993

Pyle, D. (1999). *Data Preparation for Data Mining* (1st ed.). San Francisco: Morgan Kaufmann Publishers Inc.

Roy, R., Kailath, T. (1989). ESPRIT-estimation of signal parameters via rotational invariance techniques. *IEEE Transactions on Acoustics, Speech, and Signal Processing, 37*(7), 984–995. https://doi.org/10.1109/29.32276

Sanei, S., Hassani, H. (2016). *Singular Spectrum Analysis of Biomedical Signals*. Boca Raton, FL: CRC Press.

Shao, X. (2010). The Dependent Wild Bootstrap. *Journal of the American Statistical Association, 105*(489), 218–235. https://doi.org/10.1198/jasa.2009.tm08744

Silva, E. S., Hassani, H., Heravi, S. (2018). Modeling European industrial production with multivariate singular spectrum analysis: A cross-industry analysis. *Journal of Forecasting, 37*, 371–384. https://doi.org/10.1002/for.2508

Silva, E. S., Hassani, H., Heravi, S., Huang, X. (2019). Forecasting tourism demand with denoised neural networks. *Annals of Tourism Research, 74*, 134–154. https://doi.org/10.1016/j.annals.2018.11.006

Stoica, P., Selen, Y. (2004). Model-order selection: a review of information criterion rules. *IEEE Signal Processing Magazine, 21*(4), 36–47. https://doi.org/10.1109/MSP.2004.1311138

Wax, M., Kailath, T. (1985). Detection of signals by information theoretic criteria. *IEEE Transactions on Acoustics, Speech, and Signal Processing, 33*(2), 387–392. https://doi.org/10.1109/TASSP.1985.1164557

Wu, C. F. J. (1986). Jackknife, Bootstrap and Other Resampling Methods in Regression Analysis. *The Annals of Statistics, 14*(4), 1261–1295. https://doi.org/10.1214/aos/1176350142

Yang, J. J., Buu, A. (2025). Automated parameter selection in singular spectrum analysis for time series analysis. *Communications in Statistics - Simulation and Computation*. https://doi.org/10.1080/03610918.2025.2456575