

# Change points detection for nonstationary multivariate time series

Yeonjoo Park<sup>a</sup>, Hyeongjun Im<sup>1,b</sup>, Yaeji Lim<sup>2,b</sup>

<sup>a</sup>Department of Management Science and Statistics, University of Texas at San Antonio, USA;

<sup>b</sup>Department of Applied Statistics, Chung-Ang University, Korea

---

## Abstract

In this paper, we develop the two-step procedure that detects and estimates the position of structural changes for multivariate nonstationary time series, either on mean parameters or second-order structures. We first investigate the presence of mean structural change by monitoring data through the aggregated cumulative sum (CUSUM) type statistic, a sequential procedure identifying the likely position of the change point on its trend. If no mean change point is detected, the proposed method proceeds to scan the second-order structural change by modeling the multivariate nonstationary time series with a multivariate locally stationary Wavelet process, allowing the time-localized auto-correlation and cross-dependence. Under this framework, the estimated dynamic spectral matrices derived from the local wavelet periodogram capture the time-evolving scale-specific auto- and cross-dependence features of data. We then monitor the change point from the lower-dimensional approximated space of the spectral matrices over time by applying the dynamic principal component analysis. Different from existing methods requiring prior information on the type of changes between mean and covariance structures as an input for the implementation, the proposed algorithm provides the output indicating the type of change and the estimated location of its occurrence. The performance of the proposed method is demonstrated in simulations and the analysis of two real finance datasets.

**Keywords:** change point detection, cumulative sum type statistics, locally stationary data, local wavelet periodogram, dynamic principal component analysis

---

## 1. Introduction

Detection of structural change points in univariate time series is an established field with a significant amount of work developed under parametric or nonparametric approaches; see, for example, Pitarakis (2004), Bücher *et al.* (2019), Wang *et al.* (2020), and Cho and Kirch (2021). Meanwhile, as applications have emerged that deal with multivariate or high-dimensional time series exhibiting nonstationary behaviors, recent works have begun addressing the change point estimation in nonstationary multivariate time series from many fields, including climatology (Heo and Manuel, 2022), finance (Banerjee and Guhathakurta, 2020), and neuroscience (Kirch *et al.*, 2015; Schröder and Om-bao, 2019).

The main statistical challenge under a multivariate setting is in proper consideration of its statistical properties varying in time along with potentially dynamic correlation among multiple time series to increase the detection power. There have been several works on the change point estimation by

---

<sup>1</sup> Yeonjoo Park and Hyeongjun Im are co-first authors.

<sup>2</sup> Corresponding author: Department of Applied Statistics, Chung-Ang University, Seoul 06974, Korea. E-mail: [yaeji.lim@gmail.com](mailto:yaeji.lim@gmail.com)

incorporating the nonstationarity properties through parametric modeling. Among many others, Aue *et al.* (2009) studied the volatility and co-volatility structures of multivariate dimensional random vectors, which allow both linear and nonlinear specifications, for testing the existence of second-order structural changes. Recently, Maciak *et al.* (2020) proposed self-normalized test statistics based on the CUSUM of partial residuals to detect a common breakpoint in dependent and nonstationary panels.

Besides the parametric approach, the representation of time series in the frequency domain is a widely used nonparametric technique in the change point detection problem, and various methods have been developed under this framework. For univariate time series, Wang (1995) studied the detection of jump and sharp cusp points in the white noise model using the wavelet transformation, and Qi *et al.* (2014) proposed a detection method based on Haar Wavelet and Kolmogorov-Smirnov statistic. In the multivariate case, Preuss *et al.* (2015) developed a nonparametric procedure to detect and estimate multiple structural breaks in the autocovariance function under the multivariate piecewise stationary process. Then a moving sum type test statistic involving the infinity norm is employed to measure differences in spectral matrices in adjacent segments of the observed process. Sundararajan and Pourahmadi (2018) also assumed the multivariate piecewise stationary process and focused on the detection of sudden changes in the covariance structure through the measure quantifying the difference of the spectral density matrices. Later Schröder and Ombao (2019) introduced a frequency-specific change-point detection method for epileptic seizure multi-channel electroencephalograms data again under the piecewise stationary process.

A locally stationary Wavelet (LSW) process, introduced by (Nason *et al.*, 2000), is also known to characterize the piecewise constant second-order structure of a time series through time-varying wavelet spectrum. For the univariate time series, Cho and Fryzlewicz (2012) applied the binary segmentation method on the wavelet periodograms under the assumption of the LSW process to detect change points in the second-order structure. Cho and Fryzlewicz (2015) further extended their previous works to the multivariate setting by assuming the multivariate LSW process and proposed a CUSUM-based binary segmentation algorithm, termed ‘sparsified binary segmentation’ (SBS), to detect the change point from local- and cross- periodograms of multivariate data.

In this article, we develop the two-step procedure that monitors the presence of structural changes, either on mean parameters or second-order structures, and estimates its position under the framework of the multivariate LSW process. The proposed procedure first scans the data by computing CUSUM-type statistic (Aue *et al.*, 2009; Horváth and Hušková, 2012; Cho and Fryzlewicz, 2012) over time to examine the presence of mean changes, such as mean shift. The CUSUM-based monitoring is a widely used sequential method to identify the likely position of change point on its trend, often as where its maximum is attained. In our procedure, we check two practical conditions derived from the properties of CUSUM-type statistics to decide the presence or absence of the trend change. When it is concluded that such change exists, its position is estimated accordingly.

If no mean change point is detected, we proceed to the second procedure in the detection of second-order structural change by modeling the multivariate time series as the multivariate LSW process (Park *et al.*, 2014; Cho and Fryzlewicz, 2015), allowing localized auto-correlation and cross-dependence. Based on the local wavelet periodogram defined by Park *et al.* (2014), spectral matrices can be estimated at time  $t$  and given scale  $j$ , providing a measure of the local contribution to both the auto-covariance and cross-covariance. We then monitor the change point in the lower-dimensional approximated space for the spectral matrices over time at the finest scale by applying dynamic principal component analysis (DPCA). The DPCA, first introduced by Brillinger (2001), enables extracting essential components from the multivariate temporal data by considering the serial dependence on them, as studied in Peña and Yohai (2016) and Peña *et al.* (2019).

PCA is one of the popular methods for multivariate data analysis but assumes independence among observations. To further consider serial dependence in multivariate time series data, dynamic PCA is developed, and in this paper, we apply wavelet transformation before using dynamic PCA. Cho and Fryzlewicz (2015) showed that the change points in the second-order structure of the multivariate time series data are detectable from the wavelet periodograms. Also, the local wavelet periodogram we used in this paper well captures the time-evolving scale-specific auto- and cross-dependence features of data. Therefore, we expect that the dynamic PCA applied to the wavelet periodogram reflects the complex structure of the given data.

We specifically identify the point displaying the abnormal change in the first component of dynamic PC scores using the CUSUM-type statistic. Unlike existing methods requiring prior information on the type of changes between mean and covariance structures as an input for the implementation, our proposed algorithm investigates the type of changes and provides the estimated location of occurrence.

The rest of the paper is organized as follows. We propose the two-step procedure for the change-point detection in Section 2 with brief reviews of the multivariate LSW process and DPCA in Section 2.2, and with the algorithm presented in 2.3. In Section 3, we compare the detection performances between the proposed and comparison methods through simulation studies, and the extension of the proposed approach to the multiple change points detection is further illustrated. Section 4 presents two applications to the finance and stock data. Finally, concluding remarks are provided in Section 5.

## 2. Methodology

Suppose we observe multivariate time series  $\mathbf{X}_{t,T} = (X_{t,T}^{(1)}, \dots, X_{t,T}^{(p)})'$ , at time points  $t = 0, \dots, T - 1$ , where  $T$  denotes the sample size. For each time series, we assume the following model,

$$X_{t,T}^{(d)} = \sigma^{(d)}\left(\frac{t}{T}\right) \left[ Y_{t,T}^{(d)} \right]^2, \quad (2.1)$$

where  $Y_{t,T}^{(d)}$  is a sequence of auto-correlated and non-stationary standard normal variables such that  $X_{t,T}^{(d)}$  is scaled  $\chi^2(1)$  distributed, as considered in Cho and Fryzlewicz (2015). Each  $\sigma^{(d)}(t/T)$  is a piecewise constant function, and our goal is to detect a change point in  $\sigma^{(d)}(t/T)$ .

We propose the two-step procedure that detects change points on nonstationary time series, either on location (mean) parameters or second-order structures. The nonstationary time series is specifically modeled through a multivariate LSW process (Park *et al.*, 2014). Cho and Fryzlewicz (2015) showed that the change points in the second-order structure were detectable from the wavelet periodograms under (2.1); thus, we apply multivariate LSW process for detecting the second-order structural change by allowing time-varying auto-correlation as well as cross-dependence among  $p$  time series with detailed procedures specified in Section 2.2. We note that although our proposed detection method is developed under (2.1) based on the part of derived asymptotic properties in Cho and Fryzlewicz (2015), it empirically shows good performance under other general model settings, as we shall see from examples in Section 2.3 and simulation studies in Section 3.

We assume that there is a subset of time series among  $X_{t,T}^{(d)}$ ,  $d = 1, \dots, p$ , displaying a change at one time point after the initial period  $t = 0$  of a structurally stable process. We especially focus on the presence of less sparse change points; in other words, it is assumed that the moderate proportion of  $X_{t,T}^{(d)}$  among  $p$  time series is affected by the incidence of event leading to the structural change on data. It is a natural appeal in many practical cases owing to cross-dependence among multiple time series. The practical level of moderate rate is further discussed through simulation studies in Section 3. We

note that the proposed algorithm is designed to detect one common change point, and we can apply it sequentially over the domain to find the following change points. Examples of multiple change points detection are illustrated through simulation studies (Section 3.3) and real data application (Section 4.2).

### 2.1. Mean change detection

In this section, we outline the first step for detecting changes in mean parameters, such as mean shift. Given  $\mathbf{X}_{t,T} = (X_{t,T}^{(1)}, \dots, X_{t,T}^{(p)})'$ , we identify the likely position of mean change point by computing CUSUM-type statistic (Cho and Fryzlewicz, 2015), for  $d = 1, \dots, p$ , over  $t = 0, \dots, T - 1$  as,

$$v_t^{(d)} := \left( \frac{1}{T} \sum_{u=0}^{T-1} X_{u,T}^{(d)} \right)^{-1} \left| \sqrt{\frac{T-t}{Tt}} \sum_{u=0}^{t-1} X_{u,T}^{(d)} - \sqrt{\frac{t}{T(T-t)}} \sum_{u=t}^{T-1} X_{u,T}^{(d)} \right|. \quad (2.2)$$

Note that the normalizing term is necessary to make the results independent of the level of  $\sigma^{(d)}(t/T)$  in the model (2.1). A large value of  $v_t^{(d)}$  indicates the potential presence of changes in mean trend at time point  $t$  of  $d^{\text{th}}$  time series, and the further discussion on the properties of  $v_t^{(d)}$  can be found in Cho and Fryzlewicz (2012, 2015). We next aggregate CUSUM-type statistics  $v_t^{(d)}$ ,  $d = 1, \dots, p$ , by using two methods discussed in Groen *et al.* (2013); pointwise maximum and pointwise average, defined respectively by,

$$v_t^{\max} = \max_{1 \leq d \leq p} v_t^{(d)}, \quad v_t^{\text{avg}} = \frac{1}{p} \sum_{d=1}^p v_t^{(d)}.$$

The location of mean change point is then estimated as  $b_{\max} := \arg \max_t v_t^{\max}$  or  $b_{\text{avg}} := \arg \max_t v_t^{\text{avg}}$ . We note that Cho and Fryzlewicz (2012) illustrated the ineffective performance of pointwise aggregation approaches under the high dimensional situation with extremely sparse change points, such as large  $p = 100$  with only one time series possessing one change point. However, we empirically examine through experiments that if the proportion of time series displaying changes at the common point is moderate or high, roughly above 20%, the pointwise-based estimator  $b_{\max}$  and  $b_{\text{avg}}$  still work superior. The detailed experimental settings and results are provided in Section 3. Under the moderate change points assumption, we propose to utilize both  $v_t^{\max}$  and  $v_t^{\text{avg}}$  in the mean change detection procedure based on their theoretical properties and remarkable finite sample performance as demonstrated in Groen *et al.* (2013).

Theoretically, if there exists a mean change at time point  $b$ ,  $v_t^{\max}$  and  $v_t^{\text{avg}}$  have the global maximum value at  $b = b_{\max} = b_{\text{avg}}$  without local maxima. In practice, however, due to finite samples and unexpected noise, we sometimes observe multimodal  $v_t^{\max}$  or  $v_t^{\text{avg}}$  involving several local maxima. We thus propose to check two conditions to confirm the presence of mean change; (i) either  $v_t^{\max}$  or  $v_t^{\text{avg}}$  features unimodal distribution; and (ii)  $|b_{\max} - b_{\text{avg}}| < \delta$ , for reasonably small  $\delta > 0$ . If both  $v_t^{\max}$  and  $v_t^{\text{avg}}$  display multimodal distributions, we conclude that mean change does not exist and terminate the detection step. Utilizing both aggregated statistics enables increasing detection power and filtering out false detection efficiently.

In detail, we check the unimodality of  $v_t^{\max}$  (or  $v_t^{\text{avg}}$ ) as follows. We first define two sets,  $\mathcal{D}_1 = \{\tau_1^1, \dots, \tau_k^1\}$  and  $\mathcal{D}_2 = \{\tau_1^2, \dots, \tau_k^2\}$ , consisting of  $k$  time points that divides  $[0, b_{\max}]$  and  $[b_{\max}, T - 1]$  with  $(k + 1)$  equally spaced intervals, respectively. If there is no exact  $t$  matching to  $\tau_i^1$  or  $\tau_i^2$ ,  $i =$

1, \dots, k, we choose the nearest  $t$  to define  $\mathcal{D}_1$  and  $\mathcal{D}_2$ . Then we calculate  $A^{\max}$  statistic as,

$$A^{\max} := \sum_{i=2}^k \mathbf{I} \left\{ v_{\tau_{i-1}^1}^{\max} > v_{\tau_i^1}^{\max} \right\} + \sum_{j=1}^{k-1} \mathbf{I} \left\{ v_{\tau_j^2}^{\max} < v_{\tau_{j+1}^2}^{\max} \right\}, \quad (2.3)$$

where an indicator function  $\mathbf{I}\{B\} = 1$ , when  $B$  is satisfied. Here  $A^{\max}$  counts the number of local maxima by examining the number of increases that occurred at the time before and after  $b_{\max}$ . In practice, obtained  $v_t^{\max}$  is often noisy with white noise local fluctuations due to finite sample circumstances, which may lead to over-counting local maxima. To avoid it, we suggest using the smoothed statistic. By applying the local polynomial smoothing to  $v_t^{\max}$ , we obtain  $\tilde{v}_t^{\max}$  and use it in calculating (2.3) as a replacement of  $v_t^{\max}$ . Here, we exclude boundary points in calculating  $A^{\max}$  to prevent any potential issues resulting from the boundary bias on  $\tilde{v}_t^{\max}$ .

Detailed calculation steps, including smoothing, are presented in the algorithm of Section 2.3. The same procedure is applied to obtain  $A^{\text{avg}}$  using  $v_t^{\text{avg}}$  or  $\tilde{v}_t^{\text{avg}}$ . We then have  $A^{\max} \times A^{\text{avg}} = 0$ , if either of the two aggregated CUSUM statistics is unimodal.

Here is the mean change point detection procedure and the stopping rule. If both conditions, (i)  $A^{\max} \times A^{\text{avg}} = 0$  and (ii)  $|b^{\max} - b^{\text{avg}}| < \delta$  are satisfied, we estimate the change point as  $b = (b_{\max} + b_{\text{avg}})/2$ . In our study, the threshold is set as  $\delta = \sqrt{T}/2$ , achieving the powerful detection from our experiments as well as from simulation studies in Cho and Fryzlewicz (2015). If at least one of two conditions is not satisfied, we conclude no mean change point over the domain and move on to the second procedure described in the next section.

## 2.2. Second-Order structural change detection

We now present the procedure for detecting the second-order structural change on multivariate time series by modeling  $\mathbf{X}_{t,T}$  with the  $p$ -variate LSW process. We briefly review the LSW model in Section 2.2.1 and propose to estimate the change point by implementing the DPCA on the local wavelet spectral (LWS) matrices in Section 2.2.2.

### 2.2.1. The multivariate LSW model and spectral matrix

The multivariate LSW process (Cho and Fryzlewicz, 2012; Park *et al.*, 2014) models  $p$ -variate nonstationary time series under rigorously defined time-varying auto-correlation property, especially allowing the individual-specific localized nonstationary behavior for each  $X_{t,T}^{(d)}$ , and the locally stationary cross-dependence between  $p$  series. In Park *et al.* (2014), individual processes are constructed using discrete wavelets  $\psi_j = \{\psi_{j,0}, \psi_{j,1}, \dots, \psi_{j,N_j-1}\}$  of length  $N_j$  for scales  $j$ , founded on  $\{h_k\}_{k=0}^{T-1}$  and  $\{g_k\}_{k=0}^{T-1}$  representing the low and high-pass quadrature mirror filters, respectively, commonly used in the construction of the Daubechies (Daubechies, 1992). Then  $\psi_j$ 's are specifically obtained as  $\psi_{1,n} = \sum_k g_{n-2k} \delta_{0,k} = g_n$  for  $n = 0, \dots, N_1 - 1$ , and  $\psi_{j+1,n} = \sum_k h_{n-2k} \psi_{j,k}$  for  $n = 0, \dots, N_{j+1} - 1$ . Here  $\delta_{0,k}$  is the Kronecker-delta function, and  $N_j = (2^j - 1)(N_h - 1) + 1$ , where  $N_h$  is the number of non-zero elements within filter  $\{h_k\}$ . More details on these wavelets can be found in Nason *et al.* (2000).

With such wavelets as building blocks, the  $p$ -variate LSW process for  $\{\mathbf{X}_{t,T}\}_{t=0}^{T-1}$ , for  $T = 2^J$ ,  $J \in \mathbb{N}$ , is represented as (Park *et al.*, 2014),

$$\mathbf{X}_{t,T} = \sum_{j=1}^{\infty} \sum_{k=0}^{T-1} \mathbf{V}_j(k/T) \psi_{j,t-k} \mathbf{Z}_{j,k}, \quad (2.4)$$

where  $\mathbf{z}_{j,k}$  are  $(p \times 1)$  uncorrelated random vectors with zero mean and variance-covariance matrix equal to the identity matrix, and  $\mathbf{V}_j(\cdot)$  is the  $(p \times p)$  matrix of functions, named as *transfer function matrix*. The transfer functions, elements of *transfer function matrix*, provide a measure of the time-varying contribution to the variance at each particular scale  $j$  for each  $\{X_{t,T}^{(d)}\}_{t=0}^{T-1}$  and cross-dependence structure. See Park *et al.* (2014) for further remarks on  $\mathbf{V}_j(\cdot)$ . Then  $(p \times p)$  LWS matrix describing the time-scale decomposition of power under multivariate setting at scale  $j$  and rescaled time  $u \in (0, 1)$  is given as,

$$\mathbf{S}_j(u) = \mathbf{V}_j(u) \mathbf{V}_j'(u). \quad (2.5)$$

Under  $T = 2^J$ , we can estimate  $\mathbf{S}_j(u)$  up to scale  $J$ , by producing  $(p \times p)$  wavelet periodogram matrix  $\mathbf{I}_{j,t} = \mathbf{d}_{j,t} \mathbf{d}_{j,t}'$ , where  $\mathbf{d}_{j,t} = (d_{j,t}^{(1)}, \dots, d_{j,t}^{(p)})'$  whose elements are empirical wavelet coefficient vector for each  $X_{t,T}^{(d)}$ , calculated as  $\mathbf{d}_{j,k} = \sum_{t=0}^{T-1} \mathbf{X}_{t,T} \psi_{j,k}(t)$ . With  $D_{jl} = \langle \Psi_j, \Psi_l \rangle = \sum_{\tau} \Psi_j(\tau) \Psi_l(\tau)$ , where  $\Psi_j(\tau) := \sum_k \psi_{j,k} \psi_{j,k-\tau}$ , the bias-corrected estimator of the LWS matrix at time point  $t$  and scale  $j$  is given by

$$\hat{\mathbf{S}}_{j,t} = \sum_{l=1}^J D_{jl}^{-1} \tilde{\mathbf{I}}_{l,t}, \quad \text{for } j \in \{1, \dots, J\}, t \in \{0, \dots, T-1\}. \quad (2.6)$$

Here  $\tilde{\mathbf{I}}_{j,t}$  is the smoother version of  $\mathbf{I}_{j,t}$  for the bias reduction. In particular, Park *et al.* (2014) employed the smoothed estimator under a rectangular kernel smoother window of length  $2M + 1$ , written as  $\tilde{\mathbf{I}}_{j,t} = (2M + 1)^{-1} \sum_{m=-M}^M \mathbf{I}_{j,t+m}$ . A smoothing span  $M$  is selected by generalized cross-validated gamma deviance criterion (Ombao *et al.*, 2001). See Park *et al.* (2014) for further associated asymptotic properties. Consequently,  $\hat{\mathbf{S}}_{j,t}$  provides a measure of the local contribution to both the auto-covariance and cross-covariance at time  $t$  and given scale  $j$ .

We propose to examine the presence of the unusual behavior in  $\hat{\mathbf{S}}_{j,t}$  to detect the second-order structural change. In our study, the change at the finest scale  $j = J$  is particularly monitored, based on results from experimental studies presenting the most powerful detection performance. Korkas and Fryzlewicz (2017) similarly discussed that change points detected at the finest scales are likely to be more accurate, and they suggested moving to coarser only if necessary.

### 2.2.2. Dynamic PCA on LWS matrices

In this section, we describe the estimation of second-order structural change point using  $\hat{\mathbf{S}}_{J,t}$  obtained from (2.6). We note that  $\hat{\mathbf{S}}_{J,t}$  contains three-dimensional information, where the  $(p \times p)$  spectral matrix is estimated at each time points  $t = 0, \dots, T-1$ . To reduce the dimensionality of  $\hat{\mathbf{S}}_{J,t}$  as well as to reflect temporal dependence, we apply DPCA (Brillinger, 2001; Hörmann *et al.*, 2015; Peña *et al.*, 2019), which extracts essential components from the multivariate temporal data by considering the serial dependence on them. Since the dynamic lower-dimensional representation, i.e., dynamic PC score, is evaluated at  $t$ , the change point detection is fulfilled by monitoring PC scores over time.

The DPCA approximates the original multivariate time series data as close as possible under the least squares criterion of reconstruction error using  $q < p$  components. We specifically perform DPCA for each  $d^{\text{th}}$  row of  $\hat{\mathbf{S}}_{J,t}$ , denoted as  $\hat{\mathbf{S}}_{J,t}^{(d)}$ , over  $t = 0, \dots, T-1$ . Then the dynamic principal components are defined as the minimizers, over the matrix  $\mathbf{C}_t^{(d)} \in \mathbb{R}^{p \times q}$  and the vector  $\boldsymbol{\xi}_t^{(d)} \in \mathbb{R}^q$ , of

$$\sum_{t=0}^{T-1} \left\| \hat{\mathbf{S}}_{J,t}^{(d)} - \sum_u \mathbf{C}_{t-u}^{(d)} \boldsymbol{\xi}_u^{(d)} \right\|_2, \quad (2.7)$$

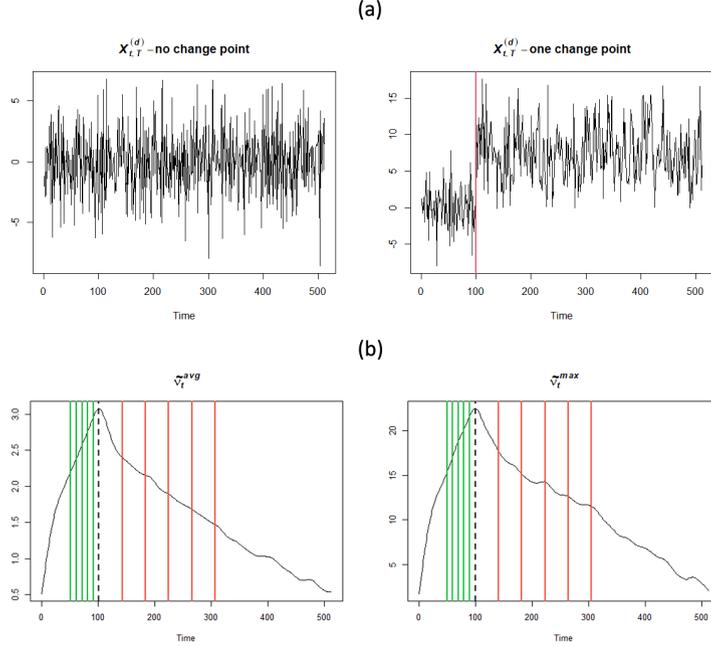


Figure 1: Example I – (a) Two randomly selected time series, and (b) smoothed CUSUM-type statistics,  $\tilde{v}_t^{\text{avg}}$  and  $\tilde{v}_t^{\text{max}}$ . The black dashed vertical line indicates the position where the maximum of distribution is achieved. Green and red lines indicate time points from  $\mathcal{D}_1$  and  $\mathcal{D}_2$ , respectively.

where the  $L_2$  norm of a matrix  $\|\mathbf{A}\|_2 := \sum_{i,j=1}^T A_{ij}^2$  for  $(T \times T)$  matrix  $\mathbf{A}$ . Based on the equivalent relationship between autocovariance time domain representation and spectral density frequency domain representation through Fourier transform, the minimum solution for dynamic PC scores  $\xi_t^{(d)}$  are computed as  $\xi_t^{(d)} = \sum_u \mathbf{b}_t^{(d)} \hat{\mathbf{S}}_{J,t}^{(d)}$  with  $q \times p$  filter  $\mathbf{b}_t^{(d)}$ , derived from

$$\mathbf{b}_t^{(d)} = \frac{1}{2\pi} \int_0^{2\pi} \mathbf{B}(\omega) e^{it\omega} d\omega,$$

where  $\mathbf{B}(\omega) = \overline{[\mathbf{U}_1(\omega), \dots, \mathbf{U}_q(\omega)]'}$ , complex conjugates of the transpose of the matrix with its columns consisting of  $\mathbf{U}_j(\omega)$ , an  $j^{\text{th}}$  eigenvector of  $\mathbf{f}_j^{(d)}(\omega)$ , the spectral density matrix of  $\hat{\mathbf{S}}_{J,t}^{(d)}$ . See Brillinger (2001) for detailed derivations and further reviews.

Among obtained PC scores  $\xi_t^{(d)} = (\xi_{1,t}^{(d)}, \dots, \xi_{q,t}^{(d)})'$ , for  $t = 0, \dots, T - 1$ , we especially examine the first dynamic PC score  $\xi_{1,t}^{(d)}$ , explaining major variations of data as well as presenting the superior empirical detection performance, although other orders can be used. Then the CUSUM-type statistic,  $\rho_t^{(d)}$ , for  $d = 1, \dots, p$ , is computed following (2.2) by replacing  $X_{u,T}^{(d)}$  with  $\xi_{1,t}^{(d)}$ . We further aggregate them using pointwise maximum,  $\rho_t^{\text{max}} = \max_{1 \leq d \leq p} \rho_t^{(d)}$ , and pointwise mean,  $\rho_t^{\text{avg}} = (1/p) \sum_{d=1}^p \rho_t^{(d)}$ , as discussed in Section 2.1. Then we set  $b_{\text{max}} = \arg \max_t \rho_t^{\text{max}}$  and  $b_{\text{avg}} = \arg \max_t \rho_t^{\text{avg}}$ . If  $|b_{\text{max}} - b_{\text{avg}}| < \delta$  is satisfied, we estimate the change point as  $b = (b_{\text{max}} + b_{\text{avg}})/2$ . We use the same threshold value  $\delta = \sqrt{T}/2$  as in Section 2.1.

For the fast computation of dynamic PC scores, we can extract only the diagonal elements of  $\hat{\mathbf{S}}_{J,t}$  in (2.6), and denote it as  $\mathbf{w}_t$  of length- $p$  vector. Let  $\hat{\mathbf{S}}^{\text{diag}} := (\mathbf{w}_0, \dots, \mathbf{w}_{T-1})$ , which is a  $p \times T$

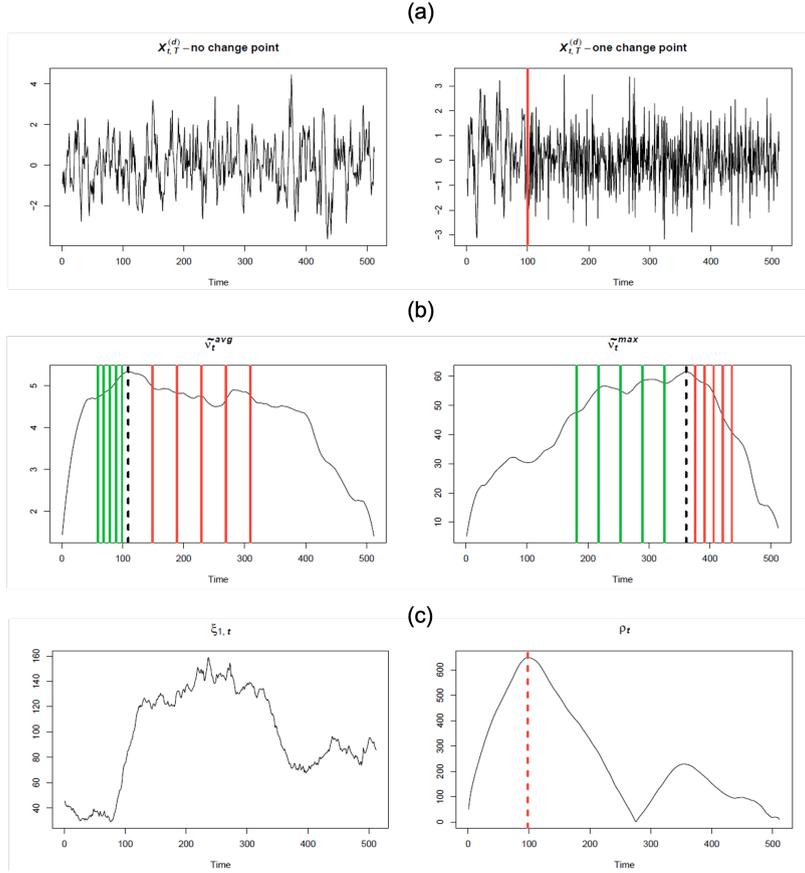


Figure 2: Example II – (a) Two randomly selected time series and (b) smoothed CUSUM-type statistics,  $\tilde{v}_t^{\text{avg}}$  and  $\tilde{v}_t^{\text{max}}$ . The black dashed vertical line indicates the position where the maximum of distribution is achieved. Green and red lines indicate time points from  $\mathcal{D}_1$  and  $\mathcal{D}_2$ , respectively. (c) The first dynamic PC score,  $\xi_{1,t}$ , obtained from DPCA (left) and its CUSUM-statistic  $\rho_t$  (right). The red dashed vertical line indicates  $b = 100$ , where the maximum of  $\rho_t$  is achieved.

matrix. Then, we apply DPCA to the  $\hat{\mathbf{S}}^{\text{diag}}$ , and obtain the first dynamic PC score,  $\xi_{1,t}$ . Note that, since we already aggregate the multivariate information in  $\mathbf{S}^{\text{diag}}$  by combining diagonal elements of the spectral matrix from each dimension, we obtain univariate CUSUM-statistic  $\rho_t$  over time without the pointwise average or max aggregation process. Then the position of the change point is estimated as  $b = \arg \max_t \rho_t$ .

### 2.3. Algorithm

The proposed procedures described in Section 2.1 and 2.2 are summarized in the following algorithm.

---

#### Algorithm 1 : Change point detection for nonstationary multivariate time series

---

- 1: Given  $p$ -variate time series,  $\mathbf{X}_{t,T} = (X_{t,T}^{(1)}, \dots, X_{t,T}^{(p)})'$ , compute CUSUM-type statistic  $v_t^{(d)}$  for each  $X_{t,T}^{(d)}$  as (2.2) and obtain the aggregated statistics,  $v_t^{\text{max}} = \max_{1 \leq d \leq p} v_t^{(d)}$  and  $v_t^{\text{avg}} = (1/p) \sum_{d=1}^p v_t^{(d)}$ .

- 2: Compute  $\tilde{v}_t^{\max}$  and  $\tilde{v}_t^{\text{avg}}$  by applying the local polynomial smoothing to  $v_t^{\max}$  and  $v_t^{\text{avg}}$ , respectively, and estimate the likely positions of change point,  $b_{\max} := \arg \max_t \tilde{v}_t^{\max}$  and  $b_{\text{avg}} := \arg \max_t \tilde{v}_t^{\text{avg}}$ . In practice, we use the relatively small value of span  $\alpha = 0.1$ , the size of the neighborhood in local smoothing, to prevent over-smoothing.
- 3: Check the unimodality of the distribution of  $\tilde{v}_t^{\max}$  and  $\tilde{v}_t^{\text{avg}}$  as follows.
  - i: Define two sets,  $\mathcal{D}_1 = \{\tau_1^1, \dots, \tau_k^1\}$  and  $\mathcal{D}_2 = \{\tau_1^2, \dots, \tau_k^2\}$ , consisting of  $k$  time points that divides  $[0, b_{\max}]$  and  $[b_{\max}, T - 1]$  with  $(k + 1)$  equally spaced intervals, respectively. In practice, we set  $k = 10$ .
  - ii: Calculate  $A^{\max}$  statistic as in (2.3). In our implementation, four boundary points at both ends are particularly excluded to alleviate potential boundary bias effects in smoothing,

$$A^{\max} = \sum_{i=6}^{10} \mathbf{I} \left\{ \tilde{v}_{\tau_{i-1}^1}^{\max} > \tilde{v}_{\tau_i^1}^{\max} \right\} + \sum_{j=1}^5 \mathbf{I} \left\{ \tilde{v}_{\tau_j^2}^{\max} < \tilde{v}_{\tau_{j+1}^2}^{\max} \right\}.$$

- iii: Similarly, use newly define  $\mathcal{D}_1$  and  $\mathcal{D}_2$  based on  $b_{\text{avg}}$ , and calculate  $A^{\text{avg}}$  statistic using  $\tilde{v}_t^{\text{avg}}$ .
- 4: If  $A^{\max} \times A^{\text{avg}} = 0$  and  $|b_{\max} - b_{\text{avg}}| < \sqrt{T}/2$ , make the conclusion on the presence of mean change and estimate the location of the change point as  $b = (b_{\max} + b_{\text{avg}})/2$ .
- 5: Else, move on to the second-order structural change detection procedure below.

- i: Compute an empirical wavelet coefficient ( $p \times 1$ ) vectors for each scale  $j = 1, \dots, J$ , for  $T = 2^J$  by  $\mathbf{d}_{j,k} = \sum_{t=0}^{T-1} \mathbf{X}_{t,T} \psi_{j,k}(t)$ , where  $k = 0, \dots, T - 1$ , and  $\psi_{j,k}(t)$  denotes discrete wavelets specified in Section 2.2.1.
- ii: Obtain the  $(p \times p)$  raw wavelet periodogram matrices,  $\mathbf{I}_{j,k} = \mathbf{d}_{j,k} \mathbf{d}_{j,k}'$ , and compute a smoothed and bias-corrected estimator of the LWS matrices at the finest scale  $j = J$ , i.e.,  $\hat{\mathbf{S}}_{J,t}$  for  $t \in \{0, \dots, T - 1\}$ , as in (2.6).
- iii: Perform DPCA for each  $d^{\text{th}}$  row of  $\hat{\mathbf{S}}_{J,t}$  over  $t = 0, \dots, T - 1$ , and obtain dynamic PC scores  $\xi_t^{(d)} = (\xi_{1,t}^{(d)}, \dots, \xi_{q,t}^{(d)})'$  over  $t$ .
- iv: For the first dynamic PC score,  $\xi_{1,t}^{(d)}$ , compute CUSUM-type statistic;

$$\rho_t^{(d)} = \left( \frac{1}{T} \sum_{k=0}^{T-1} \xi_{1,t}^{(d)} \right)^{-1} \left| \sqrt{\frac{T-t}{Tt}} \sum_{u=0}^{t-1} \xi_{1,u}^{(d)} - \sqrt{\frac{t}{T(T-t)}} \sum_{u=t}^{T-1} \xi_{1,u}^{(d)} \right|, \text{ for } d = 1, \dots, p.$$

- v: Aggregate them as  $\rho_t^{\max} = \max_{1 \leq d \leq p} \rho_t^{(d)}$  and  $\rho_t^{\text{avg}} = (1/p) \sum_{d=1}^p \rho_t^{(d)}$ , and obtain  $b_{\max} = \arg \max_t \rho_t^{\max}$  and  $b_{\text{avg}} = \arg \max_t \rho_t^{\text{avg}}$ .
- 6: If  $|b_{\max} - b_{\text{avg}}| < \sqrt{T}/2$  is satisfied, we conclude the presence of the second-order structural change and estimate its location as  $b = (b_{\max} + b_{\text{avg}})/2$ .

We illustrate the practical implementation of the proposed algorithm using two experimental datasets generated from Model A and Model B in simulation studies of Section 3.

#### 1. Example I: Detection of the mean change.

We generate multivariate time series  $\mathbf{X}_{t,T} = (X_{t,T}^{(1)}, \dots, X_{t,T}^{(30)})'$  with  $T = 512$  as follows.

$$X_{t,T}^{(d)} = \alpha_d X_{t-1,T}^{(d)} + \sigma_d \epsilon_t^{(d)}, \text{ for } t = 0, \dots, 511, d = 1, \dots, 30,$$

where autoregressive (AR) coefficients  $\alpha_d \stackrel{\text{i.i.d.}}{\sim} U(-0.5, 0.5)$ ,  $\sigma_d \stackrel{\text{i.i.d.}}{\sim} U(0.5, 2)$ , and  $\epsilon_t^{(d)} \stackrel{\text{i.i.d.}}{\sim} N(0, 4)$ . For randomly selected 15 time series, new randomly generated coefficients  $\alpha_d$  and  $\sigma_d$  are used after  $t = 100$ . Also,  $\epsilon_t^{(d)}$  is regenerated at  $t = 100$  from  $N(\mu_d, 4)$ , where  $\mu_d \sim U(2, 5)$ . Therefore the data display the mean structural change at  $t = 100$ . Figure 1(a) displays two randomly selected time series; one without the change and the other one with the change. Then aggregated and smoothed CUSUM-type statistics,  $\tilde{v}_t^{\text{avg}}$  and  $\tilde{v}_t^{\text{max}}$ , are presented in Figure 1(b). Here, we observe  $b_{\text{avg}} = 102$  and  $b_{\text{max}} = 100$  (marked by dashed vertical lines), where green and red lines indicate time points in  $\mathcal{D}_1$ , and  $\mathcal{D}_2$ , derived from  $b_{\text{avg}}$  and  $b_{\text{max}}$ , respectively. We obtain  $A^{\text{max}} = 0$  and  $A^{\text{avg}} = 0$ , indicating  $A^{\text{max}} \times A^{\text{avg}} = 0$ , and also  $|b_{\text{max}} - b_{\text{avg}}| = 2$ , which is less than  $\sqrt{T}/2 = \sqrt{512}/2$ . Two conditions are satisfied, we thus conclude the presence of the mean change, and the estimated location is calculated as  $b = (b_{\text{max}} + b_{\text{avg}})/2 = (100 + 102)/2 = 101$ , close to the actual change point at  $t = 100$ .

## 2. Example II: Detection of the second-order structural change.

We generate multivariate time series  $\mathbf{X}_{t,T} = (X_{t,T}^{(1)}, \dots, X_{t,T}^{(50)})'$  with  $T = 512$  as follows.

$$X_{t,T}^{(d)} = \alpha_d X_{t-1,T}^{(d)} + \epsilon_t^{(d)}, \quad \text{for } t = 0, \dots, 511, \quad d = 1, \dots, 50,$$

where AR coefficient  $\alpha_d \stackrel{\text{i.i.d.}}{\sim} U(0.5, 1)$  and  $\epsilon_t^{(d)} \stackrel{\text{i.i.d.}}{\sim} N(0, 1)$ . For randomly selected 25 time series, AR coefficient  $\alpha_d$  is identically regenerated from  $U(-0.5, 0.3)$  and used after  $t = 100$ . These contaminated time series then exhibit second-order structural changes at  $t = 100$ . Two generated time series with and without change are illustrated in Figure 2(a), with  $\tilde{v}_t^{\text{max}}$  and  $\tilde{v}_t^{\text{avg}}$  presented in Figure 2(b). Since neither of two conditions are satisfied,  $A^{\text{max}} \times A^{\text{avg}} = 2 \times 2 \neq 0$  and  $|b_{\text{max}} - b_{\text{avg}}| = 273 > \sqrt{512}/2$ , we conclude the absence of mean structural change. Next, we move on to the second-order structural change detection procedure, and the first dynamic PC score  $\xi_{1,t}$ , computed from the fast computation approach described in Section 2.2.2, is shown in the left panel of Figure 2(c). The CUSUM-statistic  $\rho_t$  displays its maximum at  $b = 100$ , which is equal to the true change point  $t = 100$ , as seen in the right panel of Figure 2(c).

## 3. Simulation

In this section, we consider three models to investigate the performance of the proposed method under various scenarios and conduct comparison studies with existing methods. We further illustrate the sequential application of the proposed algorithm for multiple change points detection problems in Section 3.3.

### 3.1. Data generation

#### • Model A : AR time series with mean structural change.

We generate 50 dimensional multivariate time series  $\mathbf{X}_{t,T} = (X_{t,T}^{(1)}, \dots, X_{t,T}^{(50)})'$  from  $X_{t,T}^{(d)} = \alpha_d X_{t-1,T}^{(d)} + \sigma_d \epsilon_t^{(d)}$ , for  $t = 0, \dots, T - 1$ , and  $d = 1, \dots, 50$ , where  $\alpha_d \stackrel{\text{i.i.d.}}{\sim} U(-0.5, 0.5)$ ,  $\sigma_d \stackrel{\text{i.i.d.}}{\sim} U(0.5, 2)$ , and  $\epsilon_t^{(d)} \stackrel{\text{i.i.d.}}{\sim} N(0, 4)$ . Let  $\rho$  represent the proportion of time series among  $p$  affected by structural changes, i.e.,  $0 \leq \rho \leq 1$ . For randomly selected  $\lfloor \rho p \rfloor$  number of time series,  $\alpha_d$  and  $\sigma_d$  are independently regenerated from the same distributions  $U(-0.5, 0.5)$  and  $U(0.5, 2)$ , respectively, at  $t = t_0$ . Also,  $\epsilon_t^{(d)}$  is regenerated at  $t = t_0$  from  $N(\mu_d, 4)$ , where  $\mu_d \sim U(2, 5)$ . Therefore mean structural change occurs at  $t_0$ . Here we consider two levels for the sparsity of the change points

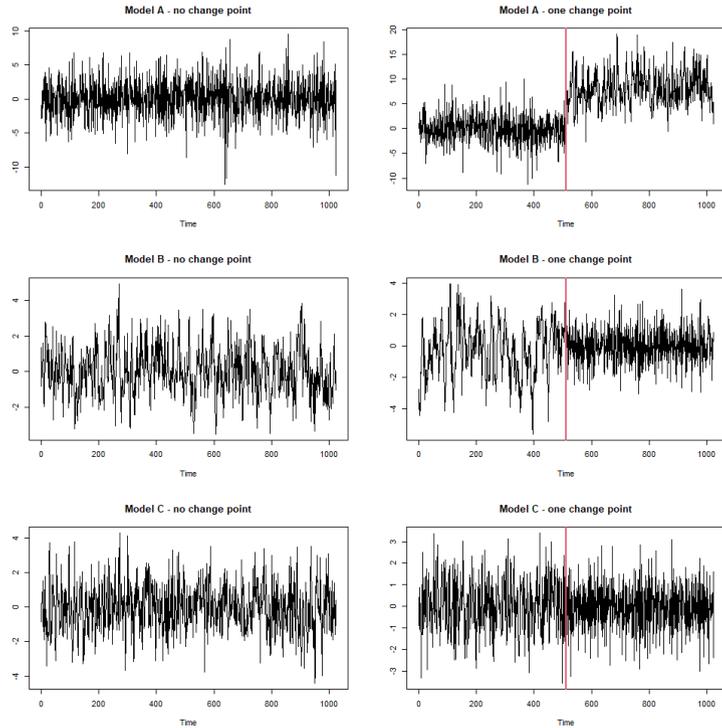


Figure 3: Randomly selected time series from Model A (top), Model B (middle), and Model C (bottom), under  $T = 1024$ , and three time series on the right panels contain structural changes at  $t_0 = 512$ , marked with red vertical lines.

across the time series; the moderate case ( $\rho = 0.7, 0.5, 0.3$ ) and the sparse case ( $\rho = 0.2, 0.1$ ). Then we set the length of each time series as  $T = 512$ , with the change point located at  $t_0 = 100$  or  $256$ . Additionally, we also consider  $T = 1024$  with change point located at  $t_0 = 100$  or  $512$ . The top left panel in Figure 3 shows the generated time series from Model A without change when  $T = 1024$ , and the top right plot represents the time series with one change in the mean structure at  $t_0 = 512$ .

- Model B : AR model with second-order structural change.

- (i) We generate 50 dimensional multivariate time series  $\mathbf{X}_{t,T} = (X_{t,T}^{(1)}, \dots, X_{t,T}^{(50)})'$  from  $X_{t,T}^{(d)} = \alpha_d X_{t-1,T}^{(d)} + \epsilon_t^{(d)}$ , for  $t = 0, \dots, T - 1$ , and  $d = 1, \dots, 50$ , where AR coefficients  $\alpha_d \stackrel{\text{i.i.d.}}{\sim} U(0.5, 1)$  and  $\epsilon_t^{(d)} \stackrel{\text{i.i.d.}}{\sim} N(0, 1)$ . For randomly selected  $\lfloor \rho p \rfloor$  number of time series, the same AR(1) model is employed but with different AR coefficients  $\alpha_d$ , regenerated from  $U(-0.5, 0.3)$  at the point  $t = t_0$ . Other settings have remained the same as in Model A. The left panel in the middle row of Figure 3 shows the generated time series with no change occurred when  $T = 1024$ , and the right plot illustrates a time series including one change point at  $t_0 = 512$ , where the second-order structural change has occurred.
- (ii) From Model B-(i),  $\epsilon_t = (\epsilon_t^{(1)}, \dots, \epsilon_t^{(50)})^T$  are generated from  $N_p(\mathbf{0}, \Sigma)$ . Here,  $(l, m)^{th}$  component of  $\Sigma$ ,  $\Sigma_{l,m} = 4(-0.95)^{|l-m|}$  for  $l, m = 1, \dots, 25$  and  $\Sigma_{l,l} = 4$  for  $l = 1, \dots, 50$ . Then for the

Table 1: Model A – Percentage of the correct identification of the change point at  $t_0$  over 100 simulated data

		Model A					
$T$	$t_0$	$\rho$	Proposed	Haar.AVG	Haar.MAX	SBS. mean	SBS.cov
512	100	0.5	<b>100</b>	92	23	<b>100</b>	31
		0.3	<b>100</b>	68	23	<b>100</b>	16
		0.2	93	53	19	<b>100</b>	12
		0.1	69	19	13	<b>100</b>	4
	256	0.5	<b>100</b>	90	35	<b>100</b>	78
		0.3	<b>100</b>	76	29	<b>100</b>	77
		0.2	<b>100</b>	53	27	<b>100</b>	64
		0.1	<b>99</b>	28	29	<b>100</b>	45
1024	100	0.5	<b>99</b>	94	39	<b>100</b>	26
		0.3	<b>97</b>	76	34	<b>100</b>	24
		0.2	82	62	24	<b>100</b>	13
		0.1	40	36	18	<b>100</b>	8
	512	0.5	<b>100</b>	<b>95</b>	66	<b>100</b>	86
		0.3	<b>100</b>	75	51	<b>100</b>	90
		0.2	<b>100</b>	56	53	<b>100</b>	82
		0.1	<b>98</b>	34	47	<b>100</b>	64

\*Boldfaces indicate higher than 95 (%) accuracy.

randomly chosen  $\lfloor \rho p \rfloor$  elements of  $\epsilon_t$  are swapped with those of other  $\lfloor \rho p \rfloor$  randomly chosen elements at the point  $t = t_0$ .

- Model C: Moving-Average (MA) time series with second-order structural change.

We generate 50 dimensional time series  $\mathbf{X}_{t,T} = (X_{t,T}^{(1)}, \dots, X_{t,T}^{(50)})'$  from  $X_{t,T}^{(d)} = \epsilon_t^{(d)} + \alpha_d \epsilon_{t-1}^{(d)}$ , for  $t = 0, \dots, T-1$ , and  $d = 1, \dots, 50$ , where MA coefficient  $\alpha_d \stackrel{\text{i.i.d.}}{\sim} U(0.5, 1)$  and  $\epsilon_t^{(d)} \stackrel{\text{i.i.d.}}{\sim} N(0, 1)$ . For randomly selected  $\lfloor \rho p \rfloor$  number of time series, MA coefficient  $\alpha_d$  is regenerated from  $U(-0.5, 0.3)$  at  $t = t_0$ . Other settings have remained the same as in Model A. The bottom panels in Figure 3 show the generated time series from Model C with and without a change when  $T = 1024$ . It is hard to detect the presence of change at  $t = t_0$  through the visual scanning of the data.

### 3.2. Evaluation of performance

The proposed algorithm is applied to each simulated set, and it is repeated 100 times. We also try the following three existing methods for comparison.

- Haar.MAX

Wavelet periodograms are computed using Haar wavelets at scale  $J$ , the finest scale, then compute CUSUM statistic for each time series, which is denoted as  $\zeta_t^{(d)}$ , for  $d = 1, \dots, p$ . Aggregate them as  $\zeta_t^{\max} := \max_{1 \leq d \leq p} \zeta_t^{(d)}$ , and estimate the location of the change point as  $b = \arg \max_t \zeta_t^{\max}$ .

- Haar.AVG

Aggregate multivariate CUSUM obtained from Haar.AVG as  $\zeta_t^{\text{avg}} := (1/p) \sum_{d=1}^p \zeta_t^{(d)}$ , and estimate the location of the change point as  $b = \arg \max_t \zeta_t^{\text{avg}}$ .

- SBS (Cho and Fryzlewicz, 2015)

Multiple change point detection algorithm based on the SBS of the second-order structure of a high dimensional time series by assuming the multivariate LSW process. There are two types of

Table 2: Model B – Percentage of the correct identification of the change point at  $t_0$  over 100 simulated data

			Model B - (i)				
$T$	$t_0$	$\rho$	Proposed	Haar.AVG	Haar.MAX	SBS.mean	SBS.cov
512	100	0.5	<b>98</b>	<b>100</b>	23	57	94
		0.3	92	<b>99</b>	28	43	71
		0.2	88	<b>97</b>	33	34	56
		0.1	78	41	31	25	29
	256	0.5	<b>97</b>	<b>100</b>	50	17	<b>100</b>
		0.3	<b>97</b>	<b>100</b>	46	13	<b>99</b>
		0.2	<b>95</b>	<b>98</b>	55	9	<b>98</b>
		0.1	86	76	53	13	<b>96</b>
1024	100	0.5	<b>98</b>	<b>100</b>	17	58	89
		0.3	<b>97</b>	<b>100</b>	29	49	70
		0.2	93	<b>99</b>	29	49	40
		0.1	88	62	36	28	21
	512	0.5	<b>98</b>	<b>100</b>	59	13	<b>100</b>
		0.3	<b>100</b>	<b>99</b>	69	8	<b>99</b>
		0.2	<b>98</b>	<b>100</b>	67	8	<b>100</b>
		0.1	93	91	62	3	<b>98</b>
			Model B - (ii)				
$T$	$t_0$	$\rho$	Proposed	Haar.AVG	Haar.MAX	SBS.mean	SBS.cov
512	100	0.7	50	44	54	12	6
		0.5	46	40	48	22	14
	256	0.7	60	58	62	6	68
		0.5	38	58	56	10	66
1024	100	0.7	56	52	52	28	6
		0.5	54	44	50	22	8
	512	0.7	68	54	64	6	67
		0.5	58	48	66	6	66

\*Boldfaces indicate higher than 95 (%) accuracy.

algorithms, named SBS.mean and SBS.cov, searching for change points in the mean and the second-order structures, respectively. Note that the type of changes should be provided as the algorithm input. The SBS may give multiple estimates of change points even with only one actual change point presence, owing to its capacity for simultaneous multiple change points detection. For this case, we evaluate the performance of the SBS by using the estimated point closest to the actual change point  $t_0$ .

For each method, we compute the percentage of accurate position identification at the true location  $t = t_0$  over 100 simulated data. As specified in Section 2.1, we set  $\delta = \sqrt{T}/2$ , and if the estimated point lies within the distance of  $\sqrt{T}/2$  from the true change point, the result is counted as the accurate one for all methods. The results are summarized in Tables 1–3. We highlight the percentages achieving an accuracy higher than 95% with boldfaces.

Table 1 shows the results from Model A, displaying the mean change in the data. Here SBS.mean works best and provides higher than 95% accuracy rates for all cases. However, SBS.cov, Haar.AVG, and Haar.MAX shows poor performance for all cases. The proposed method provides comparable results when the sparsity is moderate,  $\rho = 0.3, 0.5$ , and we observe the diminished detection power as  $\rho$  decreased, as discussed in Cho and Fryzlewicz (2015). However, it still shows improved performance than SBS.cov, Haar.AVG, and Haar.MAX display. The novelty of our algorithm is in not requiring the input about the types of change in advance, the mean or covariance, as the SBS algorithm does. The SBS works superior when the user correctly specifies the type of the change, but it collapses when the wrong information is entered. The proposed method displays higher than 70% accuracy levels when

Table 3: Model C – Percentage of the correct identification of the change point at  $t_0$  over 100 simulated data

$T$	$t_0$	$\rho$	Model C				
			Proposed	Haar.AVG	Haar.MAX	SBS.mean	SBS.cov
512	100	0.5	<b>98</b>	<b>99</b>	0	57	17
		0.3	<b>100</b>	91	22	43	7
		0.2	<b>95</b>	72	17	34	5
		0.1	73	47	21	25	3
	256	0.5	<b>98</b>	<b>100</b>	27	17	94
		0.3	<b>100</b>	<b>95</b>	37	13	88
		0.2	<b>99</b>	78	32	9	87
		0.1	86	48	35	13	54
1024	100	0.5	<b>98</b>	<b>100</b>	14	58	9
		0.3	<b>100</b>	93	17	49	4
		0.2	<b>97</b>	91	23	49	4
		0.1	81	49	24	28	1
	512	0.5	<b>99</b>	<b>99</b>	53	13	<b>98</b>
		0.3	<b>99</b>	94	49	9	87
		0.2	<b>98</b>	89	45	8	93
		0.1	85	60	56	4	84

\*Boldfaces indicate higher than 95(%) accuracy.

the sparsity is not extreme,  $\rho \geq 0.2$ . Under the sparse case  $\rho = 0.1$ , our approach based on both max and average aggregation displays improved detection power compared to Haar.AVG or Haar. MAX show.

Results from Model B are presented in Table 2. For the Model B - (i), it shows superior detection performances from the proposed method and Haar.AVG. The SBS.cov works good with the change that occurs at the middle location of the time series but shows ineffective performance when  $t_0 = 100$  with sparse cases,  $\rho = 0.2, 0.1$ . Since the coefficient of the AR model can have negative values at the change point  $t = t_0$  owing to the randomness in the data generation, it induces spuriously large values of CUSUM statistic, not from the actual location of the change point. Due to the same reason, Haar.max performs far behind for all cases. Since our method utilizes both aggregated CUSUM-type statistics, pointwise maximum and average statistics, the performance is not severely suffered from this phenomenon. For the Model B - (ii), all methods provide accuracy lower than 70%. However, proposed method works similar to the Haar.AVG and SBS.cov, and works better than SBS.cov when  $t_0 = 100$ .

Lastly, Table 3 shows that all comparison methods show low detection power in Model C under the MA(1) model. As seen at the bottom panel in Figure 3, the change point is not visually recognized over time. However, our proposed method still shows good detection power in almost all cases with high accuracy rates.

### 3.3. Extension to the multiple change points detection

Although the proposed algorithm is designed to detect one common change point in the multivariate time series, we can extend it to detect multiple change points by combining the segmentation technique in Cho and Fryzlewicz (2012). To do this, we specifically apply the binary segmentation method of Cho and Fryzlewicz (2012) to our algorithm as follows: If the proposed algorithm detects a change point at  $t = b$ , we divide the time series into two parts:  $[0, b]$  and  $[b + 1, T - 1]$ , and the algorithm is applied to each segment. Once another change point is detected, we sequentially repeat the same process with further segmentation until no more change point is observed.

We illustrate the segmentation process and its performance with the following simulated data

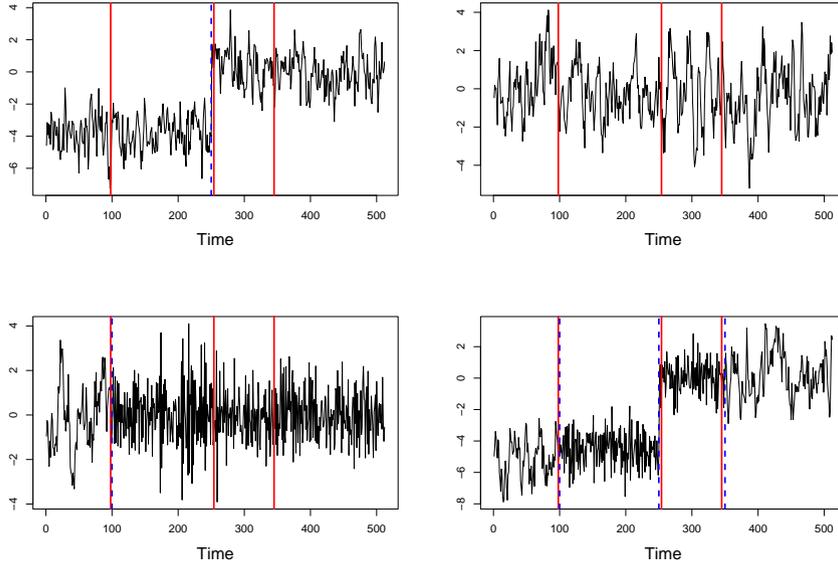


Figure 4: Randomly selected four time series with the presence of three change points at  $t = 100, 250,$  and  $350$ . Estimated locations of change points are marked with red solid vertical lines, and the actual locations of change points are marked with blue dashed vertical lines.

example. First, 30 dimensional multivariate time series  $\mathbf{X}_{t,T}^* = (X_{t,T}^{*(1)}, \dots, X_{t,T}^{*(30)})'$  is generated from  $X_{t,T}^{*(d)} = \alpha_d^* X_{t-1,T}^{*(d)} + \epsilon_t^{*(d)}$ , for  $t = 0, \dots, T - 1$ ,  $d = 1, \dots, 30$ , where  $T = 512$ ,  $\alpha_d^* \stackrel{\text{i.i.d.}}{\sim} U(0.5, 1)$  and  $\epsilon_t^{*(d)} \stackrel{\text{i.i.d.}}{\sim} N(0, 1)$ . Here we contaminate the data at point  $t = 100$  for 10 randomly selected time series, by using AR coefficients  $\alpha_d^*$  independently regenerated from  $U(-0.5, 0.3)$ . We then additionally generate the other 20 dimensional multivariate time series  $\mathbf{X}_{t,T}^\dagger = (X_{t,T}^{\dagger(1)}, \dots, X_{t,T}^{\dagger(20)})'$  from  $X_{t,T}^{\dagger(i)} = C_i \mathbf{I}(t \geq 250) + \alpha_i^\dagger X_{t-1,T}^{\dagger(i)} + \epsilon_t^\dagger(i)$  for  $t = 0, \dots, T - 1$ ,  $i = 1, \dots, 20$ , where  $T = 512$ , constant  $C_i \stackrel{\text{i.i.d.}}{\sim} U(0.5, 5)$ ,  $\mathbf{I}(\cdot)$  is an indicator function, and other parameters are set as the same in  $\mathbf{X}_{t,T}^*$ . For the randomly selected 10 time series,  $\alpha_i^\dagger$  is regenerated from  $U(-0.5, 0.3)$  at  $t = 100$ , and from  $U(0.5, 1)$  at  $t = 350$ . Consequently, 50-variate time series  $X_{t,T} = \{(\mathbf{X}_{t,T}^*)', (\mathbf{X}_{t,T}^\dagger)'\}'$  display one mean change at  $t = 250$  and two second-order structural changes at  $t = 100$  and  $350$ . Figure 4 illustrates the randomly selected four time series under this setting.

We apply the proposed algorithm sequentially with the binary segmentation technique, and the locations of change points are estimated as  $t = 98, 253,$  and  $345$  (red vertical lines in Figure 4), very close to the actual locations of change points at  $t = 100, 250,$  and  $350$ . A more precise extension for the multiple change points detection is left for future research.

#### 4. Real data application

In this section, we present the application of the proposed procedures to the two real data and conduct comparison studies.

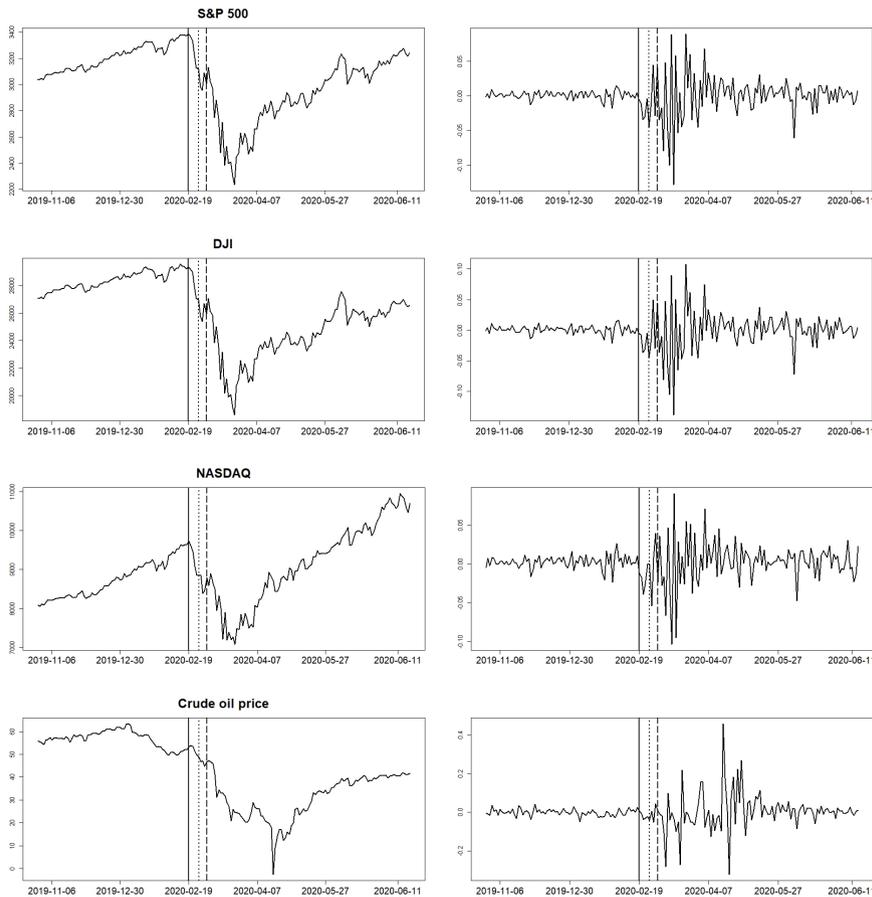


Figure 5: (Left) original financial variables and (Right) the corresponding log returns between October 28, 2019, and July 27, 2020. The solid, dotted, dashed vertical lines indicate the estimated locations of detected change points from the proposed method, SBS.cov, and Haar.AVG, respectively.

#### 4.1. Financial data

We consider four financial variables with 897 observations: S&P 500, Dow Jones industrial average (DJI), NASDAQ, and crude oil price from July 9, 2018 to July 27, 2020. Four plots in the left panel of Figure 5 show the raw collected data, and four plots in the right panel display their log returns. For a clearer presentation of the estimation results, time series plots are generated from October 28, 2019 to July 27, 2020, although we use all data in the detection process. The log return values are used in the detection procedures for all methods, and the vertical lines in each plot represent the estimated locations of change points from corresponding methods. The estimated locations from the proposed method, SBS.cov, and Haar.AVG are February 19, 2020, February 26, 2020, and March 3, 2020, respectively.

The proposed method captures the second-order structural change with its estimated location close to February 20, 2020, corresponding to the date of the stock market plunge due to the COVID-19 outbreak. It is well-known that the 2020 stock price crash, beginning on February 20, 2020, lasted

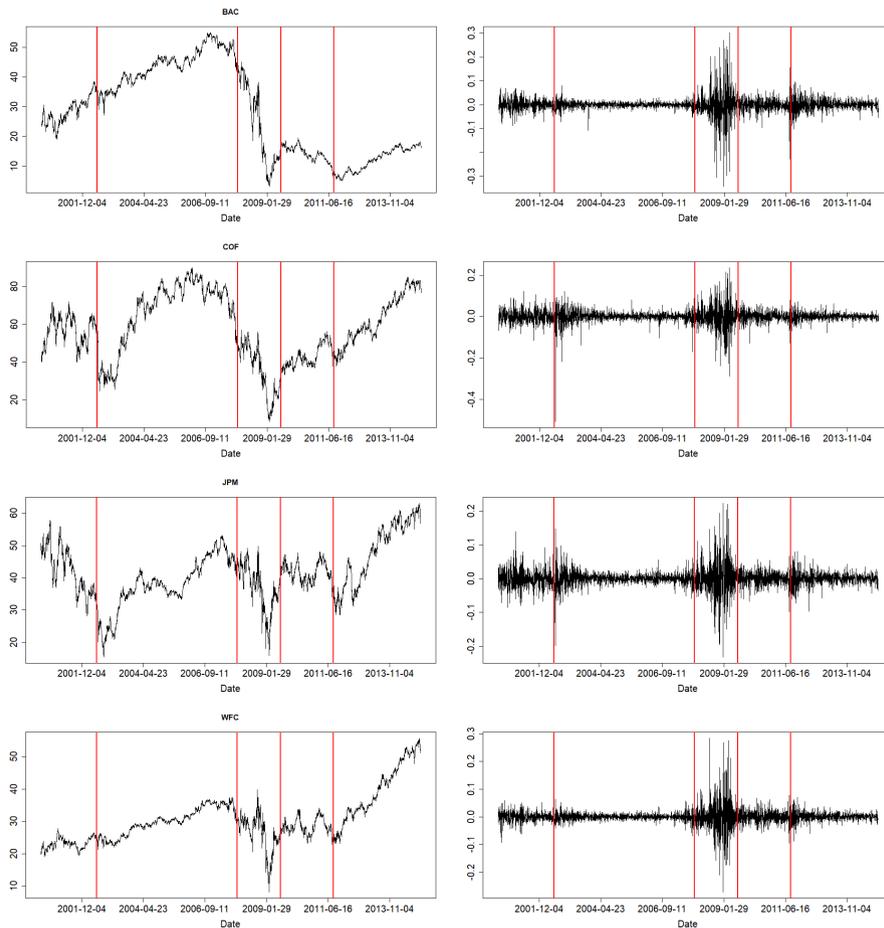


Figure 6: (Left) original stocks and (Right) the corresponding log-returns from May 1, 2000, to January 15, 2015. The red vertical lines represent the estimated location of change points from the proposed method.

approximately a month until April 7, 2020. We observe the relatively accurate detection of the start date of such sharp plunge from our method.

## 4.2. Stock data

We apply the proposed method to 12 stocks of the S&P100 with 4097 observations: AXP (American Express), BAC (Bank of America), BK (BNY Mellon), BLK (BlackRock), COF (Capital One), GS (Goldman Sachs), JPM (JPMorgan Chase), MET (MetLife), MS (Morgan Stanley), SCHW (Charles Schwab), USB (U.S. Bank), and WFC (Wells Fargo) from May 1, 2000 to August 11, 2016. Figure 6 shows the four raw stock variables and their log returns among 12 stocks. Again, for a clearer visualization of the estimation results, we generate the times series plot only from May 1, 2000 to January 15, 2015. The detection procedures are implemented on log return values, and we especially apply the segmentation method in Section 3.3 for the multiple change points detection.

The locations of estimated change points obtained from the proposed method are represented by the red vertical lines in Figure 6. Note that they are detected as second-order structural changes. One of estimated points at July 9, 2002 is quite close to the day of the stock market plunge due to WorldCom's bankruptcy on July 21, 2002. Indeed it was the largest scale of bankruptcy filing in United States history at that time. Two other detected change points at December 11, 2007, and August 14, 2009, capture the European financial crisis that occurred between 2007 and 2009. Finally, the estimated point at August 26, 2011, corresponds to the 2011 European sovereign debt crisis. We empirically confirm the superior detection performance of our proposed method even with the presence of multiple change points.

## 5. Concluding remarks

In this paper, we consider a change point detection problem for nonstationary multivariate time series data. We provide the two-step procedure that detects and estimates the location of mean change or second-order structural changes in data. The mean change detection procedure is based on the CUSUM-type statistics. If no mean change is detected, we move on to the second-order structural change point detection procedure, built on the novel combination of the LWS modeling and DPCA. The time localized property of the local wavelet spectrum enables modeling the time-varying auto- and cross-correlations of multivariate time series, and further, by applying DPCA to the spectrum, we could extract lower dimensional information of second-order structure over time. Simulation studies show that the proposed algorithm efficiently detects both types of changes, either in mean or second-order structures. Note that we assume a specific underlying model in (2.1) where wavelet transformation works well, but the proposed method shows superiority under various settings. Also, although the proposed algorithm is designed to detect one change point, we could extend it to the multiple change points problem using the segmentation technique with empirically confirmed remarkable performance through simulation examples and real data examples.

We have several future research directions. One direction is in developing theoretical properties of the proposed estimator based on the part of asymptotic properties derived in Cho and Fryzlewicz (2015). Also, we can improve some ad-hoc procedures in the algorithm. For example, in the second-order change detection procedure, examining the LSW matrices over all scales rather than the finest scale may increase the detection power by using richer information. The dimension reduction of LSW matrices from all scales should be considered in future research. Furthermore, we can use the second or third dynamic PC scores to scan data. Those PC scores may detect minor change points in the multivariate time series that the first PC scores could overlook.

## Acknowledgments

This research was supported by the Chung-Ang University Graduate Research Scholarship in 2021, the National Research Foundation of Korea (NRF) funded by the Korea government (2022R1F1A1074134), and Korea Institute of Energy Technology Evaluation and Planning (KETEP) and the Ministry of Trade, Industry and Energy (MOTIE) of the Republic of Korea (No.20199710100060).

## References

- Aue A, Hörmann S, Horváth L, and Reimherr M (2009). Break detection in the covariance structure of multivariate time series models, *The Annals of Statistics*, **37**, 4046–4087.
- Banerjee S and Guhathakurta K (2020). Change-Point analysis in financial networks, *Stat*, **9**, e269.

- Brillinger, D. R. (2001). *Time Series: Data Analysis and Theory*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia.
- Bücher A, Fermanian JD, and Kojadinovic I (2019). Combining cumulative sum change-point detection tests for assessing the stationarity of univariate time series, *Journal of Time Series Analysis*, **40**, 124–150.
- Cho H and Fryzlewicz P (2012). Multiscale and multilevel technique for consistent segmentation of nonstationary time series, *Statistica Sinica*, **22**, 207–229.
- Cho H and Fryzlewicz P (2015). Multiple-Change-Point detection for high dimensional time series via sparsified binary segmentation, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **77**, 475–507.
- Cho H and Kirch C (2021). Data segmentation algorithms: Univariate mean change and beyond, *Econometrics and Statistics*.
- Daubechies I (1992). *Ten Lectures on Wavelets*, Society for Industrial and Applied Mathematics, Philadelphia, PA.
- Groen JJ, Kapetanios G, and Price S (2013). Multivariate methods for monitoring structural change, *Journal of Applied Econometrics*, **28**, 250–274.
- Heo T and Manuel L (2022). Greedy copula segmentation of multivariate non-stationary time series for climate change adaptation, *Progress in Disaster Science*, **14**, 100221.
- Hörmann S, Kidziński Ł, and Hallin M (2015). Dynamic functional principal components, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **77**, 319–348.
- Horváth L and Hušková M (2012). change-point detection in panel data, *Journal of Time Series Analysis*, **33**, 631–648.
- Kirch C, Muhsal B, and Ombao H (2015). Detection of changes in multivariate time series with application to eeg data, *Journal of the American Statistical Association*, **110**, 1197–1216.
- Korkas KK and Fryzlewicz P (2017). Multiple change-point detection for non-stationary time series using wild binary segmentation, *Statistica Sinica*, **27**, 287–311.
- Maciak M, Pešta M, and Peštová B (2020). Changepoint in dependent and non-stationary panels, *Statistical Papers*, **61**, 1385–1407.
- Nason GP, Von Sachs R, and Kroisandt G (2000). Wavelet processes and adaptive estimation of the evolutionary wavelet spectrum, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **62**, 271–292.
- Ombao HC, Raz JA, Strawderman RL, and Von Sachs R (2001). A simple generalised crossvalidation method of span selection for periodogram smoothing, *Biometrika*, **88**, 1186–1192.
- Park T, Eckley IA, and Ombao HC (2014). Estimating time-evolving partial coherence between signals via multivariate locally stationary Wavelet processes, *IEEE Transactions on Signal Processing*, **62**, 5240–5250.
- Peña D, Smucler E, and Yohai VJ (2019). Forecasting multiple time series with one-sided dynamic principal components, *Journal of the American Statistical Association*, **114**, 1683–1694.
- Peña D and Yohai VJ (2016). Generalized dynamic principal components, *Journal of the American Statistical Association*, **111**, 1121–1131.
- Pitarakis JY (2004). Least squares estimation and tests of breaks in mean and variance under misspecification, *The Econometrics Journal*, **7**, 32–54.
- Preuss P, Puchstein R, and Dette H (2015). Detection of multiple structural breaks in multivariate time series, *Journal of the American Statistical Association*, **110**, 654–668.
- Qi JP, Zhang Q, Zhu Y, and Qi J (2014). A novel method for fast change-point detection on simulated time series and electrocardiogram data, *PloS one*, **9**, e93365.

- Schröder AL and Ombao H (2019). Fresped: Frequency-specific change-point detection in epileptic seizure multi-channel eeg data, *Journal of the American Statistical Association*, **114**, 115–128.
- Sundararajan RR and Pourahmadi M (2018). Nonparametric change point detection in multivariate piecewise stationary time series, *Journal of Nonparametric Statistics*, **30**, 926–956.
- Wang D, Yu Y, and Rinaldo A (2020). Univariate mean change point detection: Penalization, cusum and optimality, *Electronic Journal of Statistics*, **14**, 1917–1961.
- Wang Y (1995). Jump and sharp cusp detection by wavelets, *Biometrika*, **82**, 385–397.

*Received November 8, 2022; Revised November 30, 2022; Accepted December 1, 2022*