



HAL
open science

Block bootstrap for periodic characteristics of periodically correlated time series

Anna E. Dudek

► **To cite this version:**

Anna E. Dudek. Block bootstrap for periodic characteristics of periodically correlated time series. 2017. hal-01305810v2

HAL Id: hal-01305810

<https://hal.science/hal-01305810v2>

Preprint submitted on 6 Oct 2017 (v2), last revised 5 Feb 2018 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Block bootstrap for periodic characteristics of periodically correlated time series

Anna E. Dudek^{1,2}

¹ Institut de Recherche Mathématique de Rennes,
Université Rennes 2, Rennes, France,

² AGH University of Science and Technology,
al. Mickiewicza 30, 30-059 Krakow, Poland,
email: aedudek@agh.edu.pl

Abstract

This research is dedicated to the study of periodic characteristics of periodically correlated time series such as seasonal means, seasonal variances and autocovariance functions. Two bootstrap methods are used: the extension of the usual Moving Block Bootstrap (EMBB) and the Generalized Seasonal Block Bootstrap (GSBB). The first approach is proposed, because the usual MBB does not preserve the periodic structure contained in the data and cannot be applied for the considered problems. For the aforementioned periodic characteristics the bootstrap estimators are introduced and consistency of the EMBB in all cases is obtained. Moreover, the GSBB consistency results for seasonal variances and autocovariance function are presented. Additionally, the bootstrap consistency of both considered techniques for smooth functions of the parameters of interest is obtained. Finally, the simultaneous bootstrap confidence intervals are constructed. A simulation study to compare their actual coverage probabilities is provided. A real data example is presented.

Keywords: Periodically correlated time series, consistency, seasonal means, seasonal variances, periodic autocovariance function.

MCS 2010 codes: 62F40 , 62G09, 62G20, 62M10.

1 Introduction

Seasonality appears naturally in economics, vibroacoustics, mechanics, hydrology and many other fields. A wide range of examples can be found e.g. in Antoni (2009), Gardner et al. (2006), Hurd and Miamee (2007) and Napolitano (2012). Periodicity is often present not only in the mean but also in the covariance function (see e.g. Broszkiewicz-Suwaj et al. (2004)). Thus, to build statistical models periodically correlated (PC) processes are used. A time series is called PC with period d , when it has periodic mean and covariance functions, i.e.

$$E(X_t) = E(X_{t+d}) \quad \text{and} \quad \text{Cov}(X_t, X_s) = \text{Cov}(X_{t+d}, X_{s+d}).$$

For more details, properties and examples we refer the reader to Hurd and Miamee (2007).

Analysis of PC time series can be performed in the time or in the frequency domain. In the first case the standard characteristics of interest are the seasonal means, the seasonal variances and the autocovariance function. In the latter case these are Fourier coefficients of the mean and the autocovariance functions,

spectral correlation density functions and the Loève bifrequency spectrum. However, one should note that the Loève bifrequency spectrum is impulsive (see (2.12c) in Napolitano (2016)) and hence it is not suitable for estimation. In contrast, estimators of the Fourier coefficients have been widely considered in the literature (see e.g. Hurd and Leśkow (1992a), (1992b), Dehay and Hurd (1994), Ciblat et al. (2002)). The overview of the aforementioned and also higher-order characteristics can be found in Hurd and Miamee (2007) and Napolitano (2016).

Independent of the characteristic that is considered, not only do its estimates need to be calculated, but also the confidence intervals are essential. Moreover, usually parameters of interest are multidimensional, e.g. vector of seasonal means or vector of Fourier coefficients. In such cases simultaneous confidence intervals are used. The asymptotic simultaneous confidence intervals are difficult to obtain. There are two reasons for that. Construction of the simultaneous confidence intervals requires estimation of the asymptotic variance and the calculation of quantiles. To obtain the simultaneous confidence intervals one needs to determine the asymptotic distribution of a smooth function of the vector of estimators (see e.g. Dudek and Leśkow (2011)). Moreover, when Fourier coefficients of the mean and the autocovariance functions are considered, the asymptotic covariance matrix is very difficult to estimate (Lenart et al. (2008), Dudek et al. (2014b), Dudek (2015)). Hence, to obtain confidence intervals resampling methods are used.

Presently there are two bootstrap approaches that can be applied for PC time series. These are the Generalized Seasonal Block Bootstrap (GSBB) introduced by Dudek et al. (2014a) and the Moving Block Bootstrap (MBB) proposed independently by Künsch (1989) and Liu and Singh (1992). The GSBB was designed for periodic processes and hence it preserves the periodic structure of the original data. The GSBB is the most general bootstrap algorithm for periodic data. Before the GSBB was introduced two other bootstrap approaches were used for data with periodic structure: the Seasonal Block Bootstrap (SBB) proposed by Politis (2001) and the Periodic Block Bootstrap (PBB) of Chan et al. (2004). The SBB assumes that the block length is always an integer multiple of the period length, while in the PBB case the block length is always smaller than the period length. For particular block length choices the GSBB reduces to the SBB and the PBB, but in general the GSBB does not require any relation between the block and the period length. The MBB completely destroys periodic structure contained in the original data and hence its application to PC times is limited. The main advantage of the MBB is that it does not require the knowledge of the period length. Moreover, it works very well for some non-periodic characteristics of PC processes and can be applied when the period length is unknown. Additionally, the MBB is widely used for almost periodically correlated (APC) time series. Note that APC processes are the generalization of PC processes. Their mean and covariance functions are almost periodic.

In the last years many bootstrap and subsampling consistency results for PC and APC processes were published. Consistency of the resampling methods for the Fourier coefficients of the mean and the autocovariance functions of PC and APC time series was obtained in Lenart et al. (2008) (subsampling), Dudek (2015) (the MBB) and Dudek et al. (2014b) (the GSBB). Additionally, Dehay and Dudek (2015, 2016) considered continuous APC processes that are not fully observed and applied the MBB for the coefficients of the mean and the autocovariance functions. Moreover, Lenart (2011) showed the subsampling consistency for the spectral density and for the magnitude of the coherence for APC time series. In Lenart (2016) a new approach called the Generalized Resampling Scheme is applied to the spectral density matrix of APC time series. The number of bootstrap consistency results in the time domain is limited. The MBB for the overall mean of APC time series was considered by Synowiecki (2007). For PC time series consistency of the GSBB for the overall mean and the seasonal means was shown by Dudek et al. (2014a). Finally, Leśkow and Synowiecki (2010) and Dudek et al. (2014a) obtained consistency of the PBB and the GSBB for periodic random arrays with growing period. Recently, Dudek (2016) showed

applicability of the MBB and the GSBB to the Fourier coefficients of the mean and the autocovariance function of such periodic random arrays.

In this paper we show consistency of the GSBB for the seasonal variances and the autocovariance function of PC time series. Since the MBB is not suitable for periodic characteristics, we introduce its extension and apply it to get consistent estimators of the seasonal means, the seasonal variances and the autocovariance function. Let us indicate that we use the GSBB in its original form and only the MBB requires modification to provide consistent results. Finally, we present results that allow for the construction of the bootstrap simultaneous confidence intervals for the aforementioned parameters. Moreover, in the simulation study we compare performance of both bootstrap methods. It will turn out that for short samples the MBB outperforms the GSBB.

The paper is organized as follows. In Section 2 the considered problem is formulated and the necessary definitions and notation are presented. Moreover, the MBB and the GSBB algorithms are recalled and modifications of the MBB is proposed. The bootstrap estimators of the seasonal means, the seasonal variances and the autocovariance function of PC time series are introduced. In all considered cases the consistency of both bootstrap approaches is shown. Finally, Section 2.5 contains the consistency results for smooth functions of the aforementioned parameters. In Section 2.4 we discuss the alternative representation of a PC sequence in the form of multidimensional stationary series and some disadvantages of this approach. Section 3 is dedicated to the simulation study of the behaviour of the both proposed bootstrap approaches. The simultaneous equal-tailed bootstrap confidence intervals for seasonal means and seasonal variances of a few simulated data examples are constructed and their actual coverage probabilities are calculated. Finally, in Section 4 the real data application is presented.

2 Problem formulation and consistency results

Let $\{X_t, t \in \mathcal{Z}\}$ be a PC time series with the known period d . In the sequel we follow the notation proposed in Dudek et al. (2014a). We assume that X_t is α -mixing i.e. $\alpha_X(k) \rightarrow 0$ as $k \rightarrow \infty$, where

$$\alpha_X(k) = \sup_t \sup_{\substack{A \in \mathcal{F}_X(-\infty, t) \\ B \in \mathcal{F}_X(t+k, \infty)}} |P(A \cap B) - P(A)P(B)|$$

and $\mathcal{F}_X(-\infty, t) = \sigma(\{X_s : s \leq t\})$, $\mathcal{F}_X(t+k, \infty) = \sigma(\{X_s : s \geq t+k\})$. If $\alpha_X(k) = 0$ it means that observations that are at least k units time apart are independent. α -mixing is an example of a weak dependence measure. Examples of other measures of weak dependence and their properties can be found in Doukhan (1994).

Let (X_1, \dots, X_n) be a sample from the considered series and B_i be the block of observations from this sample of the length $b \in \mathcal{N}$ that starts with observation X_i , i.e

$$B_i = (X_i, \dots, X_{i+b-1}).$$

Moreover, let $l \in \mathcal{N}$ be the smallest number for which $b(l+1) > n$. Below we recall the MMB and the GSBB algorithms.

Moving Block Bootstrap

1. Choose the block length b ($b < n$).
2. From the set $\{B_1, \dots, B_{n-b+1}\}$ select randomly with replacement $l + 1$ blocks that will create the bootstrap sample $(B_1^{*MBB}, \dots, B_{l+1}^{*MBB})$. Thus, each block is chosen with probability $1/(n - b + 1)$. Finally, take from the bootstrap sample the first n observations to obtain the sample $(X_1^{*MBB}, \dots, X_n^{*MBB})$ of the same size as the original one.

Generalized Seasonal Block Bootstrap

1. Choose the block length b ($b < n$).
2. For $t = 1, b + 1, 2b + 1, \dots, lb + 1$, let

$$(X_t^{*GSBB}, X_{t+1}^{*GSBB}, \dots, X_{t+b-1}^{*GSBB}) = (X_{k_t}, X_{k_t+1}, \dots, X_{k_t+b-1}).$$

Random variables $k_t, t = 1, b + 1, 2b + 1, \dots, lb + 1$ are i.i.d. from a discrete uniform random variable taking values in the set

$$S_{t,n} = \{t - dR_{1,n}, t - d(R_{1,n} - 1), \dots, t - d, t, t + d, \dots, t + d(R_{2,n} - 1), t + dR_{2,n}\},$$

with $R_{1,n} = [(t - 1)/d]$ and $R_{2,n} = [(n - b - t)/d]$.

3. Join the $l + 1$ blocks to get $(B_1^{*GSBB}, \dots, B_{l+1}^{*GSBB})$. Finally, take from the bootstrap sample the first n observations to obtain the sample $(X_1^{*GSBB}, \dots, X_n^{*GSBB})$ of the same size as the original one.

The MBB method in contrary to the GSBB does not keep the periodic structure of the original sample and cannot be applied in our problem. Having the MBB sample one cannot identify which seasons the bootstrap observations come from. Thus, we propose an extension of the usual MBB. A similar idea that allowed to use the MBB for the coefficients of the autocovariance function can be found in Dudek (2015).

Extension of the MBB method

Let us define a bivariate series $Y_i = (X_i, i)$ and then do the MBB on the sample (Y_1, \dots, Y_n) to obtain $(Y_1^{*MBB}, \dots, Y_n^{*MBB})$. In the second coordinate of the series $Y_1^{*MBB}, \dots, Y_n^{*MBB}$ we preserve the information on the original indices. Below we show how this information can be included in the definitions of the bootstrap estimators of the periodic characteristics of X_t to provide the consistency of the extended MBB. From now on when we use the abbreviation EMBB we have in mind the extended MBB version. Let $\{i_1, \dots, i_{l+1}\}$ be selected block numbers in step 2 of the MBB algorithm performed on (Y_1, \dots, Y_n) . Moreover, let $n = lb + r$, where $r \in \{0, 1, \dots, b - 1\}$. Using the first and the second coordinates of $Y_1^{*MBB}, \dots, Y_n^{*MBB}$ we form a vector containing the selected observations

$$(X_1^{*EMBB}, \dots, X_n^{*EMBB}) = (X_{i_1}, \dots, X_{i_1+b-1}, X_{i_2}, \dots, X_{i_2+b-1}, \dots, X_{i_l}, \dots, X_{i_l+b-1}, \\ X_{i_{l+1}}, \dots, X_{i_{l+1}+r-1})$$

and the vector of the original time indices

$$TI = (i_1, \dots, i_1 + b - 1, i_2, \dots, i_2 + b - 1, \dots, i_l, \dots, i_l + b - 1, i_{l+1}, \dots, i_{l+1} + r - 1).$$

By v_s^* we denote the number of elements in the EMBB sample that are from season s i.e.

$$\begin{aligned} v_s^* &= \#TI_s = \#\{t : t \in TI, t \bmod d = s\} \quad \text{for } s = 1, \dots, d-1, \\ v_d^* &= \#TI_d = \#\{t : t \in TI, t \bmod d = 0\}. \end{aligned}$$

Note that $n = v_1^* + \dots + v_d^*$ and that TI_s contains those time indices from TI that represent season s .

The next three sections are dedicated to the seasonal means, the seasonal variances and the autocovariance function of a PC time series. We introduce their bootstrap estimators and show the EMBB and the GSBB consistency in each case. Let us recall that we use the modified version of the usual MBB and the usual GSBB. This means that we do the GSBB on the original sample (X_1, \dots, X_n) .

2.1 Seasonal means

Denote by μ_s ($s = 1, \dots, d$) the seasonal means of X_t . Their estimators are defined as following

$$\hat{\mu}_s = \frac{1}{w_s} \sum_{i=0}^{w_s-1} X_{s+id} \quad \text{for } s = 1, \dots, d, \quad (1)$$

where w_s is the number of occurrences in the sample observations from season s .

Let us recall the GSBB estimator of μ_s introduced in Dudek et al. (2014a):

$$\hat{\mu}_s^{*GSBB} = \frac{1}{w_s} \sum_{i=0}^{w_s-1} X_{s+id}^{*GSBB} \quad \text{for } s = 1, \dots, d. \quad (2)$$

Finally, we define the bootstrap counterparts of (1) obtained using the EMBB algorithm by

$$\hat{\mu}_s^{*EMBB} = \frac{1}{v_s^*} \sum_{i \in TI_s} X_i^{*EMBB} \quad \text{for } s = 1, \dots, d. \quad (3)$$

If $v_s^* = 0$ for some s then we take $\hat{\mu}_s^{*EMBB} = 0$.

The number v_s^* in (3) is equal to the number of observations from season s in the bootstrap sample. Note that when the GSBB is used then the periodic structure is kept and in the bootstrap sample each season is represented by exactly the same number of observations as in the original sample. This is not the case for the EMBB. When the block length is smaller than the period length ($b < d$), it may happen that in the bootstrap sample there are no observations from some seasons. In general some seasons may be underrepresented and others overrepresented.

Remark 2.1 *It is worth noticing that the EMBB estimator of the overall mean ($\bar{\mu} = 1/d \sum_{i=1}^d \mu_s$), i.e. $1/d \sum_{i=1}^d \hat{\mu}_s^{*EMBB}$, is equal to the mean from the bootstrap sample $1/n \sum_{i=1}^n X_i^{*EMBB}$. Thus, in this case we do not need to use information from which seasons the observations in the bootstrap sample are from and hence the standard MBB approach can be used (consistency of the MBB for $\bar{\mu}$ is shown in Synowiecki (2007)).*

Before we formulate consistency results, we need to introduce some additional notation. Let $\boldsymbol{\mu} = (\mu_1, \dots, \mu_d)'$ and $\hat{\boldsymbol{\mu}} = (\hat{\mu}_1, \dots, \hat{\mu}_d)'$ be the vectors of the seasonal means and their estimators, respectively. By $(\cdot)'$ we denote transposition. Moreover, let $\hat{\boldsymbol{\mu}}^{*EMBB}$ be the bootstrap counterparts of $\hat{\boldsymbol{\mu}}$ obtained using the EMBB.

Theorem 2.1 Let $\{X_t, t \in \mathcal{Z}\}$ be an α -mixing PC time series. Assume that for some $\delta > 0$,

- (i) $\sup_{t \in \mathcal{Z}} \mathbb{E}|X_t|^{4+\delta} < \infty$;
- (ii) $\sum_{\tau=1}^{\infty} \tau \alpha_X^{\delta/(4+\delta)}(\tau) < \infty$;

Then, if $b \rightarrow \infty$ as $n \rightarrow \infty$ such that $b = o(n)$, the EMBB for the seasonal means is consistent, i.e.

$$\rho_d(\mathcal{L}(\sqrt{w}(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu})), \mathcal{L}^*(\sqrt{w}(\hat{\boldsymbol{\mu}}^{*EMBB} - \hat{\boldsymbol{\mu}}))) \xrightarrow{p} 0 \quad (4)$$

unless asymptotically $\mathcal{L}(\sqrt{w}(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}))$ is degenerate. Moreover, $w = \lfloor n/d \rfloor$.

By ρ_d we denote a metric metrizing weak convergence of probability measures on \mathbb{R}^d . Moreover, $\mathcal{L}(\cdot)$ denotes a probability law and $\mathcal{L}^*(\cdot)$ its bootstrap counterpart conditional on (X_1, \dots, X_n) . Note that the GSBB consistency results presented in Dudek et al. (2014a) are obtained for Mallows' metric. Here it can also be considered. Theorem 2.1 states consistency of the EMBB under exactly the same assumptions as was obtained for the GSBB by Dudek et al. (2014a).

2.2 Seasonal variances

In this section we consider the other important characteristic of a periodic time series, e.g. seasonal variances. So far in the literature there is no bootstrap consistency result for them. Below we recall definition of the seasonal variances of a PC time series and introduce their bootstrap estimators. Since the GSBB technique keeps the periodic structure of the data, it provides a very elegant and intuitively natural estimator. The EMBB estimator requires the same kind of modifications that we proposed in the seasonal means case.

Let σ_s^2 ($s = 1, \dots, d$) denote the seasonal variances of the considered PC time series X_t with the period length d . Their estimators are defined as

$$\hat{\sigma}_s^2 = \frac{1}{w_s} \sum_{i=0}^{w_s-1} (X_{s+id} - \hat{\mu}_s)^2 \quad \text{for } s = 1, \dots, d.$$

Finally, the bootstrap versions of $\hat{\sigma}_s^2$ obtained using the GSBB and the EMBB are of the form

$$\begin{aligned} \hat{\sigma}_s^{*2 \text{ GSBB}} &= \frac{1}{w_s} \sum_{i=0}^{w_s-1} (X_i^{*GSBB} - \hat{\mu}_s)^2 \quad \text{for } s = 1, \dots, d, \\ \hat{\sigma}_s^{*2 \text{ EMBB}} &= \frac{1}{v_s^*} \sum_{i \in TI_s} (X_i^{*EMBB} - \hat{\mu}_s)^2 \quad \text{for } s = 1, \dots, d. \end{aligned}$$

Again if $v_s^* \equiv 0$ for some s then we define $\hat{\sigma}_s^{*2 \text{ EMBB}} = 0$.

Let us define three d-dimensional vectors containing all seasonal variances, their estimators and bootstrap estimators, i.e. $\boldsymbol{\sigma}^2 = (\sigma_1^2, \dots, \sigma_d^2)'$, $\hat{\boldsymbol{\sigma}}^2 = (\hat{\sigma}_1^2, \dots, \hat{\sigma}_d^2)'$ and $\hat{\boldsymbol{\sigma}}^{*2} = (\hat{\sigma}_1^{*2}, \dots, \hat{\sigma}_d^{*2})'$. Since the consistency for both bootstrap methods holds under exactly the same assumptions, for the sake of simplicity we present the results not using additional superscripts. We also keep this rule for the autocovariance function we introduce later.

Theorem 2.2 Assume that $\{X_t, t \in \mathcal{Z}\}$ is an α -mixing PC time series, WP(3) and WP(4) such that

$$(i) \sup_{t \in \mathcal{Z}} \mathbb{E}|X_t|^{8+2\delta} < \infty;$$

$$(ii) \sum_{\tau=1}^{\infty} \tau \alpha_X^{\delta/(4+\delta)}(\tau) < \infty;$$

If $b \rightarrow \infty$ as $n \rightarrow \infty$ such that $b = o(n)$, then

$$\rho_d(\mathcal{L}(\sqrt{w}(\hat{\sigma}^2 - \sigma^2)), \mathcal{L}^*(\sqrt{w}(\hat{\sigma}^{*2} - \sigma^2))) \xrightarrow{p} 0 \quad (5)$$

unless asymptotically $\mathcal{L}(\sqrt{w}(\hat{\sigma}^2 - \sigma^2))$ is degenerate.

By $WP(k)$ we denote a weakly periodic process of order k . Process X_t is $WP(k)$ if $\mathbb{E}|X_t|^k < \infty$ and for any $t, \tau_1, \dots, \tau_{k-1} \in \mathcal{Z}$

$$\mathbb{E}(X_t X_{t+\tau_1} \dots X_{t+\tau_{k-1}})$$

is periodic in the variable t . For more details on weakly periodic processes see e.g. Synowiecki (2007).

2.3 Autocovariance function

For the last one we consider the autocovariance function of a PC time series with known period of length d . By $B(t, \tau)$ we denote the autocovariance function of X_t . To be precise $B(t, \tau) = \text{Cov}(X_t, X_{t+\tau})$. Arguments t and τ represent time and shift, respectively. Note that $B(t, \tau)$ is a periodic function of the time argument. Moreover, the seasonal variances σ_s^2 ($s = 1, \dots, d$) that were considered in Section 2.2 are values of the autocovariance function at lag 0, i.e., $\sigma_s^2 = B(s + kd, 0)$ for $k \in \mathcal{N}$. The detailed description of $B(t, \tau)$ for PC time series, its properties and some examples can be found in the book of Hurd and Miamee (2007).

The estimator of $B(s, \tau)$ is of the form

$$\hat{B}(s, \tau) = \frac{1}{w} \sum_{k=0}^{w-1} (X_{s+kd} - \hat{\mu}_s) (X_{s+\tau+kd} - \hat{\mu}_{s+\tau}), \quad (6)$$

where $s = 1, \dots, d$ and w is the number of full periods contained in the sample ($n = wd + r_1, r_1 \in \{0, 1, \dots, d-1\}$). All elements of the sum (6) for which the time index $s + \tau + kd > n$ are replaced by 0. Moreover, $\hat{\mu}_{s+\tau}$ is the estimator of the seasonal mean represented by index $s + \tau$.

The bootstrap estimators of $B(t, \tau)$ obtained with the GSBB and the EMBB are of the form

$$\hat{B}^{*GSBB}(s, \tau) = \frac{1}{w} \sum_{k=0}^{w-1} (X_{s+kd}^{*GSBB} - \hat{\mu}_s) (X_{s+\tau+kd}^{*GSBB} - \hat{\mu}_{s+\tau}), \quad (7)$$

$$\hat{B}^{*EMBB}(s, \tau) = \frac{1}{v_s^*} \sum_{t \in TI_s} (X_t^{*EMBB} - \hat{\mu}_s) (X_{t+\tau}^{*EMBB} - \hat{\mu}_{s+\tau}). \quad (8)$$

If $v_s^* \equiv 0$ for some s then we take $\hat{B}^{*EMBB}(s, \tau) = 0$.

Notice that in the formula (6) each summand $(X_{s+kd} - \hat{\mu}_s) (X_{s+\tau+kd} - \hat{\mu}_{s+\tau})$ consists of the observation from season s and the observation from the season represented by index $s + \tau$. Below we investigate if this structure is preserved in formulas (7)-(8). For this purpose we decompose the bootstrap estimator

into two components. For the sake of simplicity we focus only on the EMBB estimator. However, one should remember that the corresponding formula holds also for the GSBB. We have

$$\begin{aligned}\widehat{B}^{*EMBB}(s, \tau) &= \frac{1}{v_s^*} \sum_{t \notin C_{s,b,\tau}^*} (X_t^{*EMBB} - \widehat{\mu}_s) (X_{t+\tau}^{*EMBB} - \widehat{\mu}_{s+\tau}) + \\ &+ \frac{1}{v_s^*} \sum_{t \in C_{s,b,\tau}^*} (X_t^{*EMBB} - \widehat{\mu}_s) (X_{t+\tau}^{*EMBB} - \widehat{\mu}_{s+\tau}),\end{aligned}\quad (9)$$

where

$$C_{s,b,\tau}^* = \{t : t \in TI_s, (X_t^{*EMBB}, t^*) \in B_i^{*EMBB}, (X_{t+\tau}^{*EMBB}, (t+\tau)^*) \in B_{i+1}^{*EMBB} \text{ for } i \in \{1, 2, \dots, l\}\}.$$

The set $C_{s,b,\tau}^*$ contains those time indices from the set TI_s for which X_t^{*EMBB} and $X_{t+\tau}^{*EMBB}$ belong to two consecutive blocks B_i^{*EMBB} and B_{i+1}^{*EMBB} . Thus, for each element of the first summand of (9) we observe exactly the same situation as for $\widehat{B}(s, \tau)$, i.e. it represents seasons s and $s + \tau$. This is no longer the case for the second summand of (9). Despite the fact that X_t^* is from season s the observation $X_{t+\tau}^*$ can be from any season. This is the consequence of using the EMBB method. Since the GSBB reproduces the periodic structure of the original sample it provides the corresponding bootstrap estimator such that X_t^{*GSBB} and $X_{t+\tau}^{*GSBB}$ are always representing seasons s and $s + \tau$. However, we show that the part of the bootstrap estimator based on the set $C_{s,b,\tau}^*$ is not asymptotically meaningful and hence consistency of the EMBB can be obtained. In fact a similar argument is essential to get the GSBB consistency. One should be aware that the real problem that this part of the estimator is introducing, is the dependence structure that was not present in the original sample. Block bootstrap algorithms are designed to preserve dependence structure only within each block.

Let r be any positive integer and let

$$\mathbf{s} = [s_1, \dots, s_r]', \quad \boldsymbol{\tau} = [\tau_1, \dots, \tau_r]' \quad \text{and} \quad B(\mathbf{s}, \boldsymbol{\tau}) = [B(s_1, \tau_1), \dots, B(s_r, \tau_r)]',$$

where $s_1, \dots, s_r \in \{1, \dots, d\}$ and $\tau_1, \dots, \tau_r \in \mathcal{Z}$.

Theorem 2.3 *Under assumptions of Theorem 2.2 for any $\mathbf{s}, \boldsymbol{\tau}$ for which $\mathcal{L}\left(\sqrt{w}\left(\widehat{B}(\mathbf{s}, \boldsymbol{\tau}) - B(\mathbf{s}, \boldsymbol{\tau})\right)\right)$ is asymptotically non-degenerate we have*

$$\rho_r\left(\mathcal{L}\left(\sqrt{w}\left(\widehat{B}(\mathbf{s}, \boldsymbol{\tau}) - B(\mathbf{s}, \boldsymbol{\tau})\right)\right), \mathcal{L}^*\left(\sqrt{w}\left(\widehat{B}^*(\mathbf{s}, \boldsymbol{\tau}) - \widehat{B}(\mathbf{s}, \boldsymbol{\tau})\right)\right)\right) \xrightarrow{p} 0, \quad (10)$$

where ρ_r is a metric metricizing weak convergence of probability measures on \mathbb{R}^r .

As for the seasonal variances, we omit the superscript in $\widehat{B}^*(\mathbf{s}, \boldsymbol{\tau})$. That means that (10) holds for the GSBB and the EMBB.

Remark 2.2 *One may note that in formulas (7) and (8) the bootstrap versions X_t^* of data are substituted to define the bootstrap statistic $\widehat{B}^*(\mathbf{s}, \boldsymbol{\tau})$ but the sample averages $\widehat{\mu}_s$ from the original data (not the bootstrap data) are used for centering in $\widehat{B}^*(\mathbf{s}, \boldsymbol{\tau})$. Such definitions of the bootstrap estimators is not accidental. In many applications analysis of time series with some seasonal pattern is done in two main steps. In the first one data are demeaned and in the latter the second-order analysis is performed. The form of our bootstrap estimators corresponds to this approach. Moreover, the bias of the bootstrap version of $\widehat{\mu}_s$ may significantly influence the performance of bootstrap methods, especially in the case of short samples.*

2.4 Discussion of the approach via multidimensional stationary time series

In our work we consider one-dimensional PC time series. However, an alternative approach is possible. X_t can be equivalently represented as \mathcal{R}^d -valued stationary time series \mathbf{X}_m (see Proposition 1.1 in Hurd and Miamee (2007)), where

$$\mathbf{X}_m = (X_{1+(m-1)d}, \dots, X_{d+(m-1)d})', \quad m = 1, 2, \dots$$

Thus, \mathbf{X}_m consists of observations from the m -th period.

Working with a stationary sequence may seem to be more convenient than analysis of a nonstationary process. However, there are several reasons why the approach introduced in this paper is important and worthy of consideration.

First of all let us recall arguments provided by Hurd and Miamee (2007). The period length is not always known, but it is required to transform a PC sequence to a vector stationary sequence. Analysis of PC time series in the time and in the frequency domain can provide its estimate. Moreover, one may test if periodicity is present only in the mean or also in the covariance function. Additionally, some problems are more easily understood for PC time series than for multidimensional sequences (see comment on innovation rank on p. 6 of Hurd and Miamee(2007)). Finally, methods developed for discrete PC time series, can be carried over to the APC case and continuous time. APC time series and PC/APC continuous time processes cannot be transformed to a stationary sequence.

At the end, let us compare resampling approaches for PC time series with those for stationary multidimensional time series. The overview of known consistency results for stationary processes can be found in Lahiri (2003) and Politis et al. (1999). Let (X_1, \dots, X_n) be a sample from a PC time series with the known period d and let $n = wd$, $w \in \mathcal{Z}$. If we do the usual MBB on our sample, we will select randomly with replacement blocks from the set of size $n - b + 1$ (see the description of the method at the beginning of Section 2). The block length b can be chosen as any integer number smaller than n . On the other hand, if we transform our sample into a stationary vector $(\mathbf{X}_1, \dots, \mathbf{X}_w)$ the number of possible block choices is $w - b + 1$ and $b < w$. Moreover, each block $(\mathbf{X}_i, \dots, \mathbf{X}_{i+b-1})$, $i = 1, \dots, w - b + 1$, consists of b full periods. If we transform it again into a one-dimensional vector

$$(X_{1+(i-1)d}, \dots, X_{d+(i-1)d}, \dots, X_{1+(i+b-2)d}, \dots, X_{d+(i+b-2)d})$$

one may notice that the first observation is always from the first season. It means that the usual MBB for \mathbf{X}_m is equivalent to the SBB for X_t . Using the usual MBB or its extension proposed by us or the GSBB, allows to choose blocks that do not contain an integer number of periods and do not necessarily start with observations from the first season. Moreover, when the PC sample is short there are very few possible block length choices for the SBB. Finally, the MBB and the GSBB techniques have higher overlap than the SBB among blocks to be bootstrapped, which leads to higher efficiency (see also comments in Dudek et al. (2014a)). Thus, in the bootstrap world, analysis of a PC time series is preferable.

Last but not least, one should be aware of the fact that analysis of time series can be done partially in the time and partially in the frequency domain. For example, some practitioners prefer to estimate seasonal means, then demean the data and perform analysis to find the second-order significant frequencies. Such an approach is preferable when detecting a fault of a machine. Approach via PC time series allows to use the same kinds of tools for the both steps of analysis.

2.5 Additional results

At the beginning we present the consistency results of both bootstrap approaches for the smooth functions of $\boldsymbol{\mu}$ and $B(\mathbf{s}, \boldsymbol{\tau})$. These results are essential to construct the bootstrap simultaneous confidence intervals

for seasonal means and seasonal variances. Their performance for some examples of PC time series is analyzed in Section 3.

Theorem 2.4 *Under the assumptions of Theorem 2.1 and taking function $H : \mathcal{R}^d \rightarrow \mathcal{R}^s$ such that it is:*

(i) *differentiable in a neighborhood of $\boldsymbol{\mu}$*

$$N_H = \left\{ x \in \mathcal{R}^d : \|x - \boldsymbol{\mu}\| < 2\eta \right\} \quad \text{for some } \eta > 0$$

(ii) $\nabla H(\boldsymbol{\mu}) \neq 0$

(iii) *the first-order partial derivatives of H satisfy a Lipschitz condition of order $\kappa > 0$ on N_H .*

If $b \rightarrow \infty$ as $n \rightarrow \infty$ such that $b = o(n)$, but $b^{-1} = o(\log^{-1} n)$, then

$$\rho_s \left(\mathcal{L} \left(\sqrt{w} (H(\hat{\boldsymbol{\mu}}) - H(\boldsymbol{\mu})) \right), \mathcal{L}^* \left(\sqrt{w} (H(\hat{\boldsymbol{\mu}}^{EMBB}) - H(\hat{\boldsymbol{\mu}})) \right) \right) \xrightarrow{p} 0,$$

where ∇ is the gradient operator.

The corresponding result for the GSBB can be found in Dudek et al. (2014a).

Theorem 2.5 *Under the assumptions of Theorem 2.3 and taking function $H : \mathcal{R}^r \rightarrow \mathcal{R}^s$ such that it is:*

(i) *differentiable in a neighborhood of $B(\mathbf{s}, \boldsymbol{\tau})$*

$$N_H = \left\{ x \in \mathcal{R}^{2r} : \|x - B(\mathbf{s}, \boldsymbol{\tau})\| < 2\eta \right\} \quad \text{for some } \eta > 0$$

(ii) $\nabla H(B(\mathbf{s}, \boldsymbol{\tau})) \neq 0$

(iii) *the first-order partial derivatives of H satisfy a Lipschitz condition of order $\kappa > 0$ on N_H .*

If $b \rightarrow \infty$ as $n \rightarrow \infty$ such that $b = o(n)$, but $b^{-1} = o(\log^{-1} n)$, then

$$\rho_s \left(\mathcal{L} \left(\sqrt{w} \left(H(\hat{B}(\mathbf{s}, \boldsymbol{\tau})) - H(B(\mathbf{s}, \boldsymbol{\tau})) \right) \right), \mathcal{L}^* \left(\sqrt{w} \left(H(\hat{B}^*(\mathbf{s}, \boldsymbol{\tau})) - H(\hat{B}(\mathbf{s}, \boldsymbol{\tau})) \right) \right) \right) \xrightarrow{p} 0.$$

Let us recall that $\hat{B}^*(\mathbf{s}, \boldsymbol{\tau})$ can be obtained using the GSBB or the EMBB.

As we mentioned before Theorems 2.4 and 2.5 can be used to construct bootstrap simultaneous consistent confidence intervals for $\boldsymbol{\mu}$ and $B(\mathbf{s}, \boldsymbol{\tau})$. Below we shortly present the idea of their construction for the example of the $(1-2\alpha)\%$ equal-tailed percentile simultaneous confidence intervals for $\boldsymbol{\mu}$. Quantiles of order α and $1 - \alpha$ are calculated using the maximum and the minimum statistics. As before we do not specify the bootstrap approach that is used, because the methodology does not depend on the chosen method. Define

$$\begin{aligned} K_{max}(x) &= P^* \left(\sqrt{w} \max_i (\hat{\mu}_i^* - \hat{\mu}_i) \leq x \right), \\ K_{min}(x) &= P^* \left(\sqrt{w} \min_i (\hat{\mu}_i^* - \hat{\mu}_i) \leq x \right). \end{aligned}$$

Then, the confidence region is of the form

$$\left(\hat{\mu}_i - \frac{K_{max}^{-1}(1-\alpha)}{\sqrt{w}}, \hat{\mu}_i - \frac{K_{min}^{-1}(\alpha)}{\sqrt{w}} \right), \quad (11)$$

where $i = 1, \dots, d$. More details and description of different confidence regions can be found e.g. in Dudek and Leśkow (2011).

In the second part of this section we would like to discuss some possible different bootstrap approaches, the modification of the assumptions of the presented theorems and problem of the block length choice. We start with the latter one. Strengthening the assumption (i) in Theorems 2.1-2.5, we are able to weaken (ii). If we replace (i) and (ii) by

(i)' the time series X_t is uniformly bounded;

(ii)' $\alpha_X(k) = O(k^{-2})$,

then the thesis of Theorems 2.1-2.5 remain valid. This is quite a standard result that appears in the papers concerning bootstrap for the first and the second order characteristics of PC time series. Thus, for proof we refer the reader e.g. to Dudek et al. (2014b).

One of the disadvantages of the usual MBB method is that the observations from the beginning and the end of the sample appear less often in the blocks than the other observations. For example X_1 is present only in B_1 . Politis and Romano (1992) proposed a modification of the MBB called the Circular Block Bootstrap (CBB). The idea of this approach is to treat data as wrapped on the circle. As a result instead of $n - b + 1$ blocks, n different blocks $\{B_1, \dots, B_n\}$ can be defined. For $t > n - b + 1$ block B_t is of the form $(X_t, \dots, X_n, X_1, \dots, X_{b-n+t-1})$. Thus, the missing part of the block is taken from the beginning of the considered sample. This idea can be also adapted to PC time series. However, to wrap the data we need to assume that the sample size is an integer multiple of the period length ($n = dw$).

In the stationary case the CBB often provides the unbiased bootstrap estimator, i.e. its expected value is equal to the considered estimator. The extension of CBB (ECBB) can be applied in our problem, but one should be aware that the obtained bootstrap estimator is only asymptotically unbiased. It is also worth noticing that this effect disappears in a case of the circular GSBB (cGSBB), because each bootstrap sample obtained with the cGSBB contains exactly w observations from each season.

Corollary 2.1 *If in Theorems 2.1-2.5 we replace the EMBB estimators by their ECBB or cGSBB counterparts, all the results remain valid.*

Finally, we would like to provide a comment on the block length choice. The problem of the block length choice in the case of stationary time series is well investigated (see e.g., Lahiri (2003)) and is based on the minimization of the mean squared error (MSE). In our nonstationary case independently on the parameter of interest, we do not know the form of the MSE. However, we believe that MSE expansions at least for the EMBB should be similar to those available for the usual MBB. For the EMBB, each resampled block basically gives about $\lfloor b/d \rfloor$ consecutive observations from the original data separated by the period d (for any seasonal starting point), and this effective bootstrap sample size should determine the bias/variance properties in estimating a seasonal mean in a manner analogous to the standard (usual) Moving Block Bootstrap. With this in mind, MSE expansions for the EMBB will probably be similar to those for the MBB, by replacing b with b/d and n with n/d in those expressions with the MBB. For example by this logic, considering variance estimation of a seasonal mean, the optimal block for the EMBB would seemingly act like $b/d = C(n/d)^{1/3}$ or $b = Cd^{2/3}n^{1/3}$ for a constant C that could be estimated with some plug-in approaches (see e.g. Politis and White (2004) in the stationary case).

3 Simulation study

Our study is divided into two parts. As first we compare performance of the ECBB and the cGSBB methods for the seasonal means μ_s ($s = 1, \dots, d$). The second part is dedicated to the seasonal variances σ_s^2 ($s = 1, \dots, d$). We consider the circular versions of the EMBB and the GSBB to reduce edge effects, caused by the fact that the observations from the beginning and the end of the sample appear in less blocks than the other observations. We restricted ourselves to μ_s and σ_s^2 because those two periodic characteristics appear the most often in applications.

For the seasonal means and the seasonal variations we calculated the actual coverage probabilities (ACPs) of the 95% equal-tailed simultaneous bootstrap confidence intervals. The number of algorithm iterations is 1000 and number of bootstrap samples is $B = 500$. Moreover, we decided to restrict our consideration to the short samples, i.e. $n \in \{60, 120, 240\}$. This decision is inspired by the work of Dudek and Potorski (2014) in which the authors compared the CBB and the cGSBB for the overall mean of PC time series. It seems that for the short samples the CBB overperforms the cGSBB, while for the longer ones this effect vanishes. Additionally, especially in the economic applications seasonal effects and short samples are present quite often and hence we believe that this kind of study can provide some suggestions regarding the advantages and disadvantages of the ECBB and the cGSBB in such situation. Moreover, we set the period length $d = 12$. Thus, in each sample we have 5, 10 or 20 periods. Finally, the block length $b \in \{3, 5, 8, 10, 13, 15, 18, 20\}$.

In the following we consider two kinds of models:

- (a) $X_t = f(t) + Y_t$,
- (b) $X_t = f(t) + Z_t$,

where $f(t)$ is a periodic or null function.

Series Y_t is an autoregressive sequence with periodic noise added:

$$Y_t = 0.5Y_{t-1} + \sin(2\pi t/12) a_t + \varepsilon_t,$$

where a_t are iid random variables from the standard normal distribution. Moreover, ε_t are iid white noise with continuous density function and they are independent on a_t .

Finally, Z_t is generated as a moving average series with periodic noise:

$$Z_t = 0.5 \cos(2\pi t/12) u_t + 0.5b_t, \quad \text{where} \quad b_t = 0.2b_{t-1} + 0.3b_{t-2} + c_t.$$

Sequences u_t and c_t are independent and are iid from the standard normal distribution.

Model (a) was chosen because it is a α -mixing (see Andrews (1984)). Mixing assumptions are very difficult to validate and we do not know if series (b) fulfills this condition. However, it is interesting to investigate how our bootstrap approaches perform for these kind of models.

Seasonal means case

In the first part of our study we consider 10 models

- M1: model (a) with $f(t) = 2 \cos(2\pi t/12)$ and ε_t that are Gaussian of mean 0 and variance 1/4
- M2: model (a) with $f(t) = 0.2 \cos(2\pi t/12)$ and ε_t that are Gaussian of mean 0 and variance 1/4
- M3: model (a) with $f(t) = 2 \cos(2\pi t/12)$ and ε_t that are iid from a mixture of normal distributions of means 1/2 and -1/2, variances 1 and weights 1/2

M4: model (a) with $f(t) = 0.2 \cos(2\pi t/12)$ and ε_t that are iid from a mixture of normal distributions of means $1/2$ and $-1/2$, variances 1 and weights $1/2$

M5: model (a) with $f(t) = 2 \cos(2\pi t/12)$ and ε_t that are iid from a Student-t distribution with location parameter 0, scale parameter $1/2$ and 4 degrees of freedom

M6: model (a) with $f(t) = 0.2 \cos(2\pi t/12)$ and ε_t that are iid from a Student-t distribution with location parameter 0, scale parameter $1/2$ and 4 degrees of freedom

M7: model (a) with $f(t) = 2 \cos(2\pi t/12)$ and ε_t that are iid from a Student-t distribution with location parameter 0, scale parameter $1/2$ and 2 degrees of freedom

M9: model (a) with $f(t) = 0.2 \cos(2\pi t/12)$ and ε_t that are iid from a Student-t distribution with location parameter 0, scale parameter $1/2$ and 2 degrees of freedom

M9: model (b) with $f(t) = 2 \cos(2\pi t/12)$

M10: model (b) with $f(t) = 0.2 \cos(2\pi t/12)$

We defined $M1-M10$ to analyze two situations: strong ($M1, M3, M5, M7, M9$) and weak ($M2, M4, M6, M8, M10$) periodic mean component when compared to the noise. In the first case values of the seasonal means are 10 times higher than in the latter one. Those examples were chosen to check the influence on ACPs of signal to noise ratio values, i.e. what effect on results is due to the power of noise. The examples of realizations of $M1-M10$ for the sample size $n = 240$ are presented in Figure 1. One can notice that in the left column (models $M1, M3, M5, M7, M9$), contrary to the right column (models $M2, M4, M6, M8, M10$), periodicity is quite strong and easy to observe. Finally, $M3-M10$ are modified versions of $M1-M2$, which allow us to investigate the influence of the different distributions of residuals on the bootstrap confidence intervals. Note that residuals in models $M5-M8$ are heavy-tailed and hence do not fulfill the assumption of existence of more than 4 or 8 moments (see Theorems 2.1 and 2.2, condition (i)). In fact residuals in $M7$ and $M8$ do not have even variance.

In Figures 2-11 we present the results of our study. In each figure in the left column are ACP curves for the ECBB and the cGSBB. The right column illustrates the differences between ACP values obtained for the ECBB and the cGSBB. Finally, rows correspond to different sample sizes ($n = 60, 120, 240$).

One can notice that performance of both bootstrap methods does not depend much on the signal to noise ratio. The ACPs obtained with the cGSBB and the ECBB for $M1, M3, M5, M7, M9$ are similar to corresponding ones for $M2, M4, M6, M8, M10$, especially for $n = 120$ and $n = 240$.

The strongest differences in performance between the bootstrap methods are observed for the shortest sample size. In this case ACPs obtained with the ECBB are too high for the short block lengths and they are decreasing when b is growing. On the other hand, the cGSBB usually provides too low ACPs for small values of b . Moreover, for $n = 60$ and all models except $M7-M8$ ACPs obtained with the cGSBB are smaller than 95% for all considered block lengths. Let us recall that residuals chosen for $M7-M8$ do not have variance and hence the moment assumption does not hold. For larger sample sizes the cGSBB curves are closer to 95% and some ACPs are greater than the nominal coverage level. For $n = 120$ and $n = 240$ the ECBB and the cGSBB ACP curves are closer to each other than for $n = 60$. In particular, for the largest considered sample size the distance between curves is not greater than 2% except for two cases, i.e. $M3$ with $b = 18$ and $M4$ with $b = 3$. For $n = 240$ ACP curves are quite flat and close to the nominal coverage probability. It is also worth noticing that the performance of both bootstrap techniques does not seem to be dependent on the distribution of the residuals. The shape of

the ACP curves for $M3 - M6$ is similar to those obtained for $M1 - M2$.

To understand better the behaviour of the ECBB and the cGSBB we compared the optimal block length choices (b_{opt}) for each case. By optimal for fixed model and sample size we mean such b from the considered set of block lengths that provides an ACP which is the closest to 95%. In Table 1 are presented the optimal block length choices b_{opt} together with ACPs for $M1 - M10$ and $n \in \{60, 120, 240\}$. Since the results for models $M7 - M8$ are slightly different than for other models we discuss them separately. For models $M1 - M6$ and $M9 - M10$ one may notice that for $n = 60$ the ECBB outperforms the cGSBB. It provides ACPs closer to 95% than the cGSBB. The absolute value of the difference between the nominal coverage probability and obtained ACP ranges from 0.2% to 1.8% for the ECBB and 1%-4.8% for the cGSBB. For $n = 120$ and $n = 240$ the corresponding differences are maximally equal to 0.4%. It is also worth noticing the values of b_{opt} for the ECBB and the cGSBB are usually different. Only in 4 cases are the optimal block lengths the same, i.e. for $M1 - M3$ with $n = 120$ and for $M6$ with $n = 240$. Finally, one may notice that for $n = 60$ and $n = 120$, b_{opt} is always smaller than the period length d . This means that the ECBB is providing ACPs close to the nominal one also for $b < d$. Let us recall that for $b < d$, in the ECBB sample some seasons may be not represented, while the cGSBB ensures that in the bootstrap sample there are exactly as many observations from each season as in the original one.

For models $M7 - M8$ results are a bit different than just described. One should remember that for $M7 - M8$ the moment assumption is strongly violated, because residuals have only the first moments. In this case the optimal block lengths for both methods are greater than for other cases. For the ECBB and $n = 60$ the chosen block length is equal to 13 and hence $b_{opt} > d$. Additionally, for $n = 60$ the ACPs obtained with the cGSBB are very close to the nominal coverage level and in contrary to the models $M1 - M6$ and $M9 - M10$ they are greater than 95%. For $n = 240$ independently of the considered bootstrap method to obtain ACPs close to 95% one needs to consider $b > 20$. In fact the obtained values of b_{opt} range from 40 to 53 and hence they are much higher than the corresponding values obtained for other models.

Seasonal variances case

In the second part of our study we constructed the 95% bootstrap equal-tailed simultaneous confidence intervals for the seasonal variances. All parameters set before remain unchanged. Only chosen PC time series are slightly modified. Obviously the shorter and more noisy the signal is then the quality of the seasonal means estimates will be lower. These estimates are essential to calculate the values of the seasonal variances estimators. Since the investigation of the influence of the seasonal means estimation quality on the variance estimator is outside the scope of this paper, we restrict our consideration to four models with the null mean function. To be more precise from $M1 - M6$ and $M9 - M10$ we removed the periodic mean component and we obtain the following models

V1: model $M1$ with $f(t) \equiv 0$,

V2: model $M3$ with $f(t) \equiv 0$,

V3: model $M5$ with $f(t) \equiv 0$,

V4: model $M9$ with $f(t) \equiv 0$.

The examples of their realizations for the sample size $n = 240$ can be found in Figure 12. In Figures 13-15 are presented ACP curves for both bootstrap methods together with the differences between the ECBB and the cGSBB. Since the main differences between the bootstrap methods are observed for the shortest sample and results for models $V2$ and $V3$ are similar to results for $V1$, to restrict the number

Table 1: Optimal block length choices (b_{opt}) for models $M1 - M10$ with sample sizes $n \in \{60, 120, 240\}$. ACPs obtained with b_{opt} for the ECBB are in column 4 and for the cGSBB in column 6.

		ECBB		cGSBB	
	n	b_{opt}	ACP	b_{opt}	ACP
$M1$	60	5	96.2%	5	92.0%
	120	8	95.4%	8	94.6%
	240	8	95.2%	3	95.6%
$M2$	60	3	96.4%	8	90.2%
	120	10	95.0%	10	95.4%
	240	8	95.2%	13	95.0%
$M3$	60	8	95.8%	3	92.8%
	120	8	95.2%	8	95.0%
	240	10	94.8%	8	95.4%
$M4$	60	8	95.2%	3	92.2%
	120	8	94.8%	5	95.0%
	240	8	95.2%	10	94.6%
$M5$	60	8	95.4%	13	93.0%
	120	10	95.4%	8	95.0%
	240	20	95.0%	8	95.4%
$M6$	60	5	96.6%	5	94.0%
	120	8	95.4%	10	95.4%
	240	20	95.2%	20	94.6%
$M7$	60	13	95.1%	8	95.2%
	120	26	94.8%	15	95.0%
	240	40	95.6%	40	95.0%
$M8$	60	13	95.0%	10	95.2%
	120	15	95.2%	18	95.6%
	240	53	95.2%	50	94.8%
$M9$	60	8	96.8%	3	92.4%
	120	10	94.6%	5	95.2%
	240	13	95.0%	10	94.8%
$M10$	60	8	96.0%	10	93.0%
	120	10	95.2%	3	95.0%
	240	15	94.8%	8	95.0%

of figures we presents ACP curves for $V2$ and $V3$ only for $n = 60$.

The main conclusions that were obtained for the seasonal means case are valid also for the seasonal variances. ACPs for the ECBB are usually higher than the corresponding ones for the cGSBB. Moreover, for $n = 120$ and $n = 240$ independently on the bootstrap method the ACP curves are quite flat. For $n = 60$ the cGSBB curves are below the nominal coverage level, while the ECBB ones are above 95% for the small values of b . The only exception of this rule is $V4$. In this case all ACP curves for $n = 120$ and $n = 240$ are entirely located above the nominal coverage level. This situation can be also observed for $n = 60$ and $b \leq 10$.

Table 2 contains the optimal block lengths. For model $V4$ we needed to perform additional simulations to find block lengths that provide an ACP close to 95%. One may notice that obtained values are much larger than corresponding ones for $V1$. Similarly to the seasonal means case, for $n = 60$ one may observe sometimes big differences in values of the ACPs obtained with the ECBB and the cGSBB. For $V1$ the optimal block length provides ACP equal to 95.2% (ECBB) and 91.8% (cGSBB). On the other hand for $V4$ ACPs obtained with b_{opt} are comparable for the both bootstrap techniques and close to the optimal coverage probability.

Table 2: Optimal block length choices (b_{opt}) for models $V1 - V4$ with sample sizes $n \in \{60, 120, 240\}$. ACPs obtained with b_{opt} for the ECBB in column 4 and for the cGSBB in column 6.

		ECBB		cGSBB	
	n	b_{opt}	ACP	b_{opt}	ACP
V1	60	5	95.2%	5	91.8%
	120	5	95.2%	13	95.0%
	240	8	95.4%	8	95.0%
V4	60	13	95.6%	10	95.2%
	120	28	95.0%	28	95.6%
	240	40	95.8%	40	94.4%
V2	60	8	94.8%	10	92.8%
V3	60	10	93.6%	5	94.0%

None of the bootstrap methods is uniformly better than the other. For shorter samples the ECBB is providing higher ACPs. This may be caused by the fact that the ECBB has a higher degree of overlap among blocks than the cGSBB (see Dudek et al. (2014a) and Lahiri (2003)). For longer samples the differences in performance between methods seem to vanish. However, if someone prefers to be conservative, we suggest using the ECBB, because it usually provides greater actual coverage probabilities than the cGSBB and simultaneously greater than the nominal coverage level.

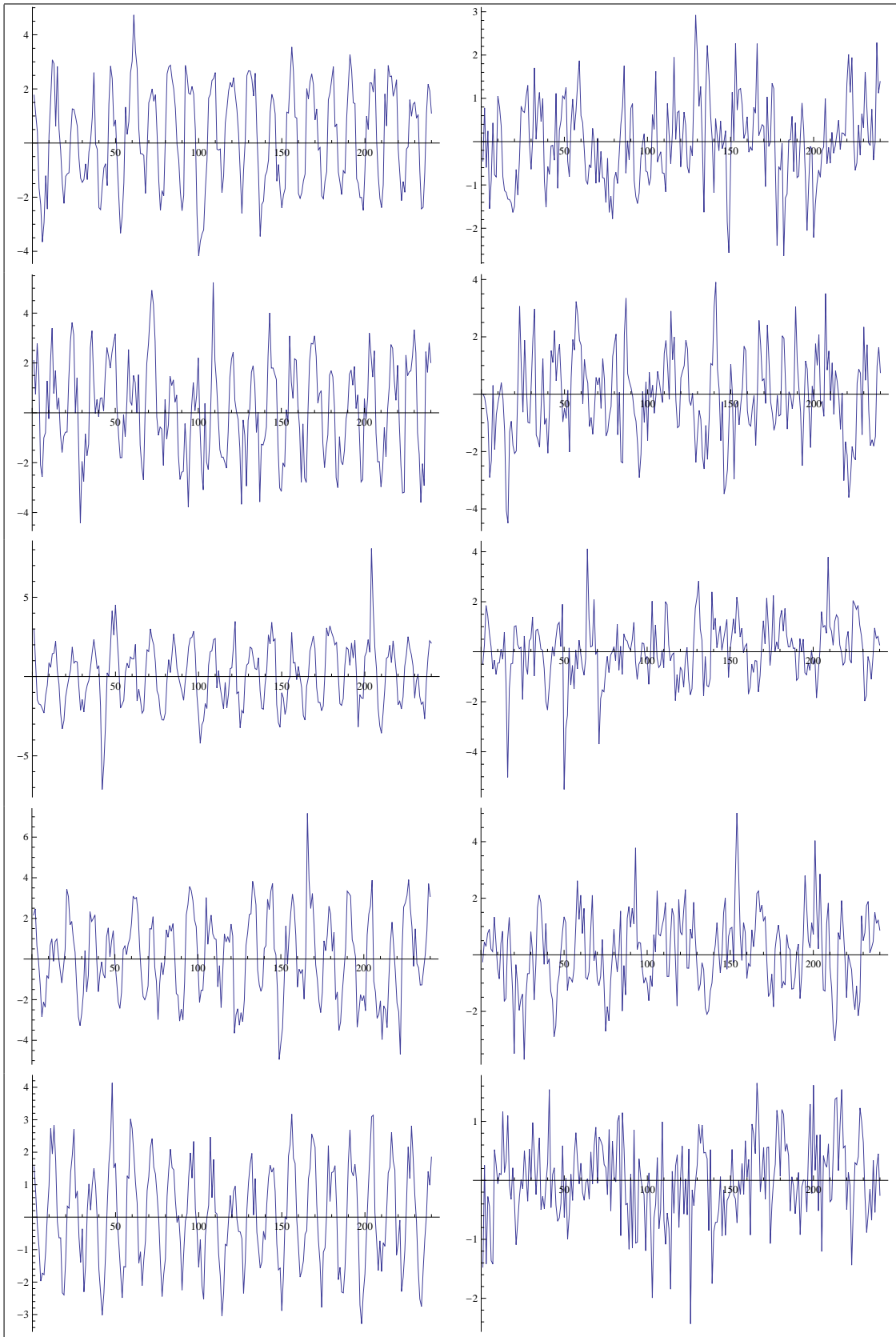


Figure 1: Examples of realizations: from top M1, M3, M5, M7, M9 (left column) and M2, M4, M6, M8, M10 (right column). Sample size $n = 240$.

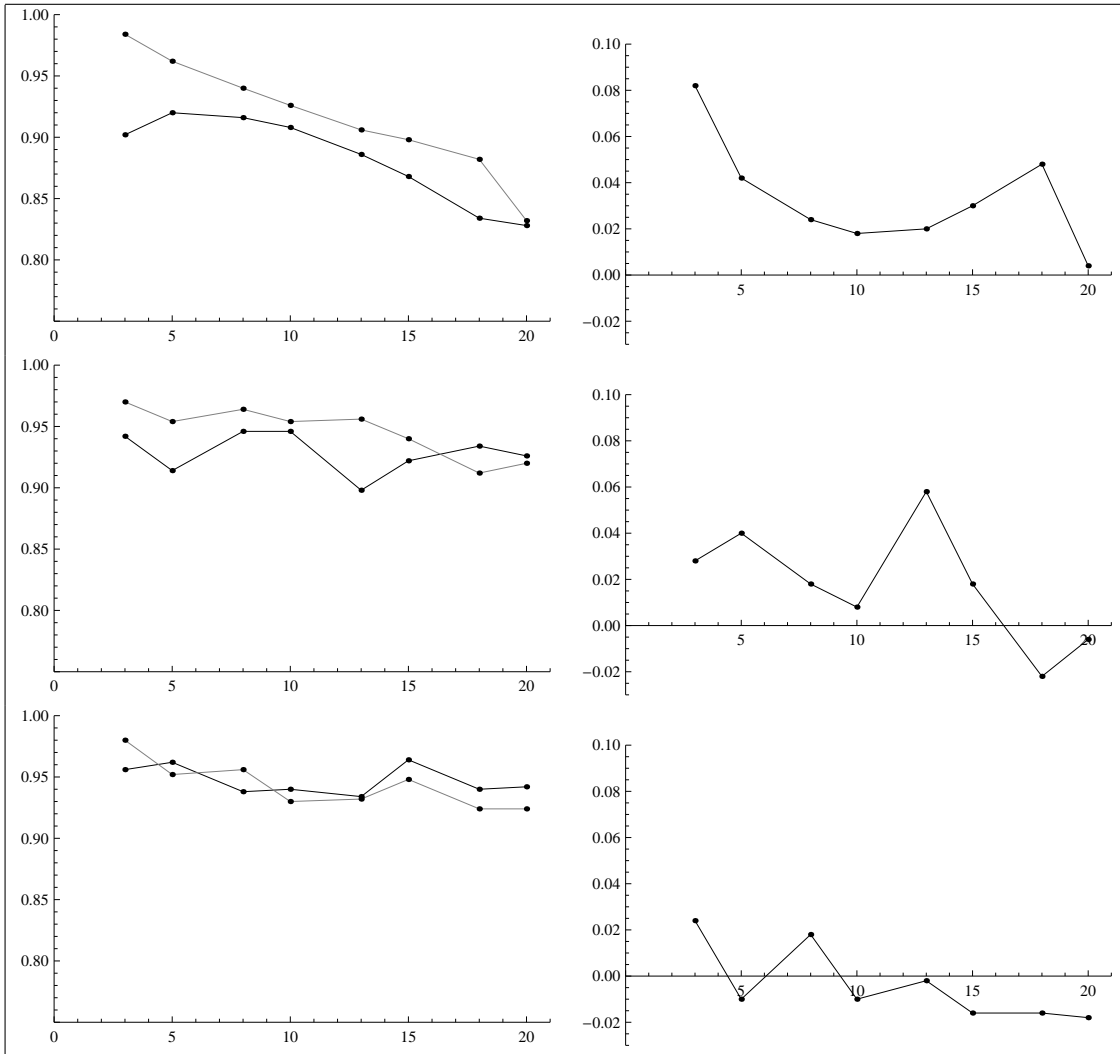


Figure 2: Model $M1$: ACPs of simultaneous equal-tailed percentile bootstrap confidence intervals for μ_i ($i = 1, \dots, d$) vs. block length b (left column). cGSBB method (black) and ECBB (gray). Right column: differences between ACPs obtained for ECBB and cGSBB. From top results for $n = 60, 120, 240$, respectively. Nominal coverage probability is 95%.

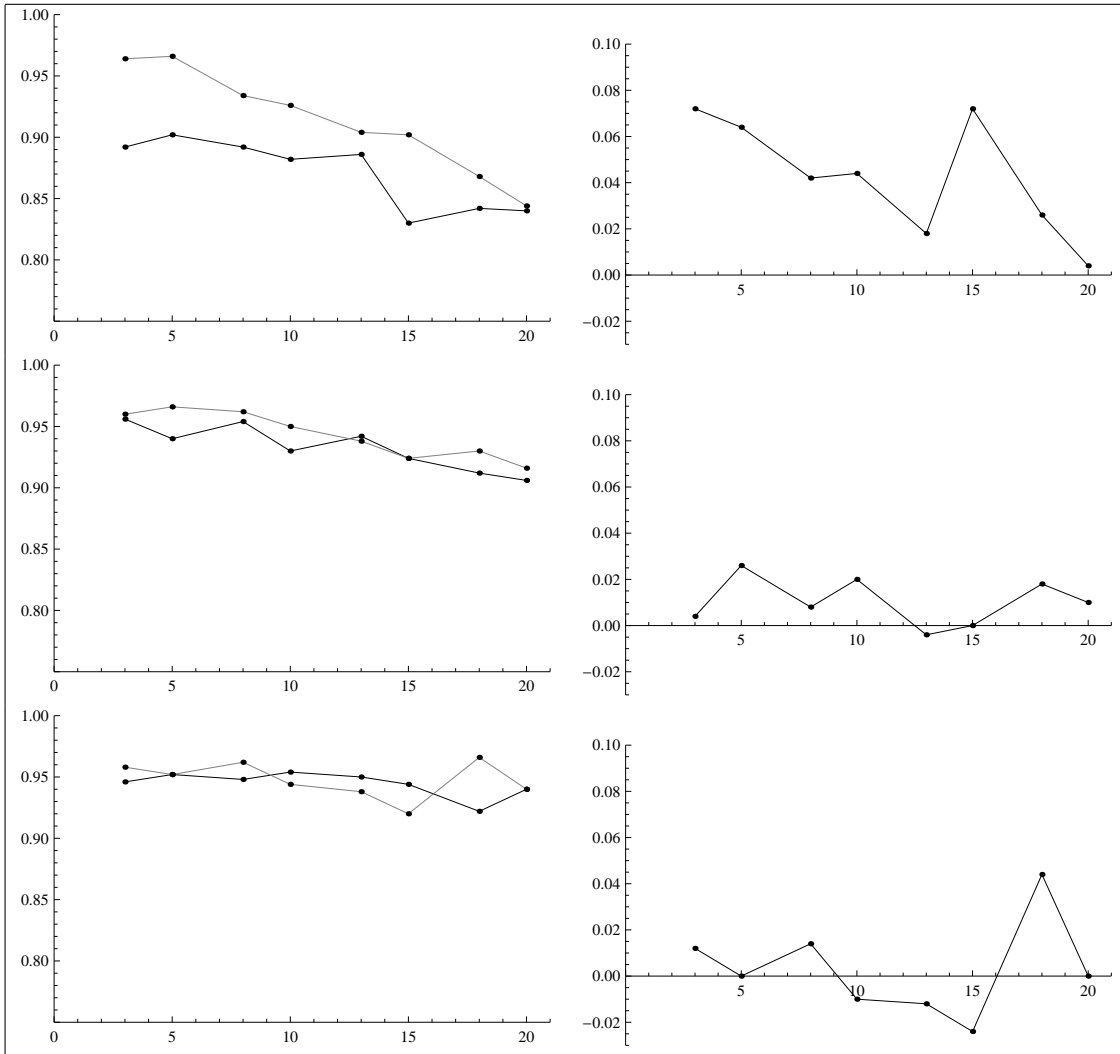


Figure 3: Model $M2$: ACPs of simultaneous equal-tailed percentile bootstrap confidence intervals for μ_i ($i = 1, \dots, d$) vs. block length b (left column). cGSBB method (black) and ECBB (gray). Right column: differences between ACPs obtained for ECBB and cGSBB. From top results for $n = 60, 120, 240$, respectively. Nominal coverage probability is 95%.

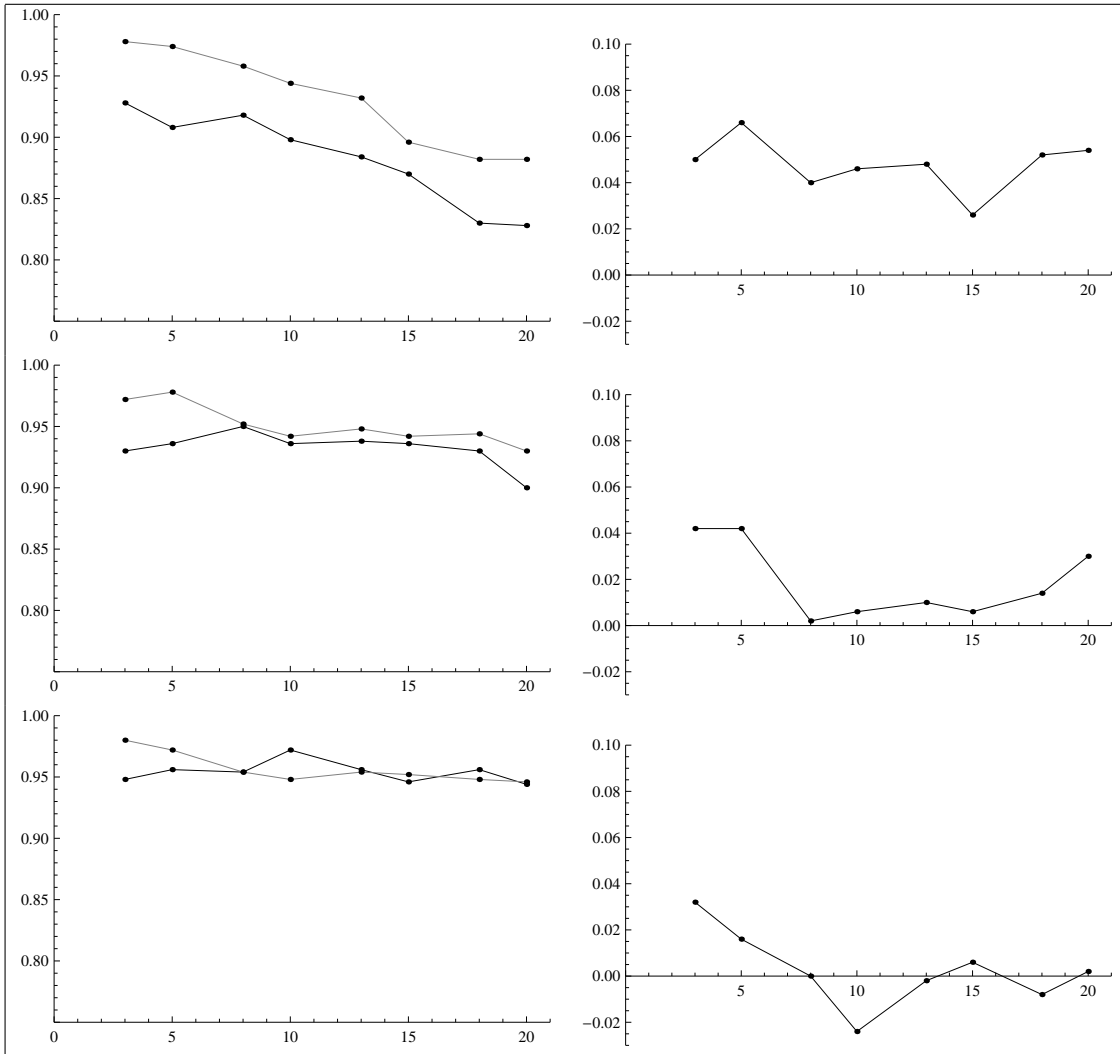


Figure 4: Model $M3$: ACPs of simultaneous equal-tailed percentile bootstrap confidence intervals for μ_i ($i = 1, \dots, d$) vs. block length b (left column). cGSBB method (black) and ECBB (gray). Right column: differences between ACPs obtained for ECBB and cGSBB. From top results for $n = 60, 120, 240$, respectively. Nominal coverage probability is 95%.

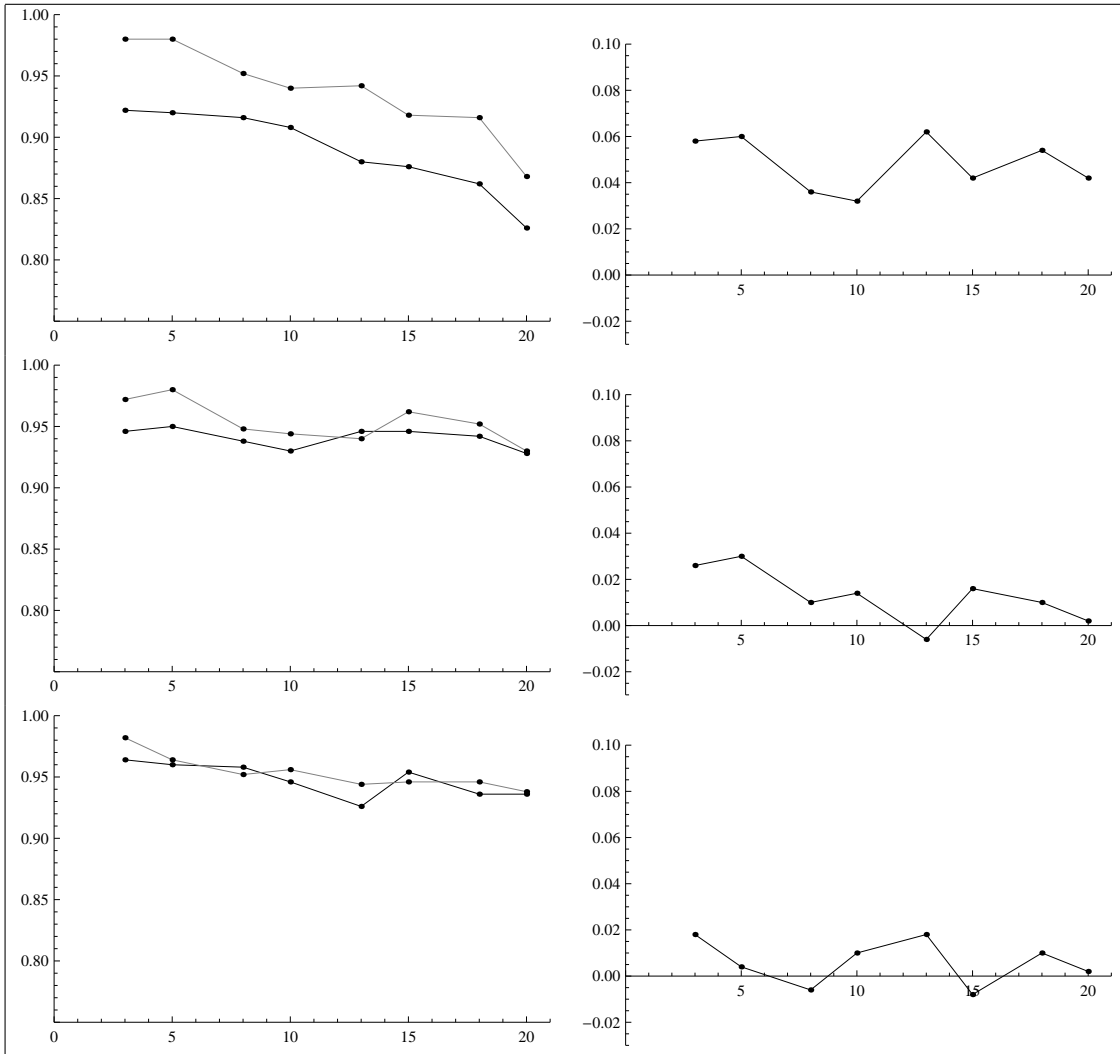


Figure 5: Model $M4$: ACPs of simultaneous equal-tailed percentile bootstrap confidence intervals for μ_i ($i = 1, \dots, d$) vs. block length b (left column). cGSBB method (black) and ECBB (gray). Right column: differences between ACPs obtained for ECBB and cGSBB. From top results for $n = 60, 120, 240$, respectively. Nominal coverage probability is 95%.

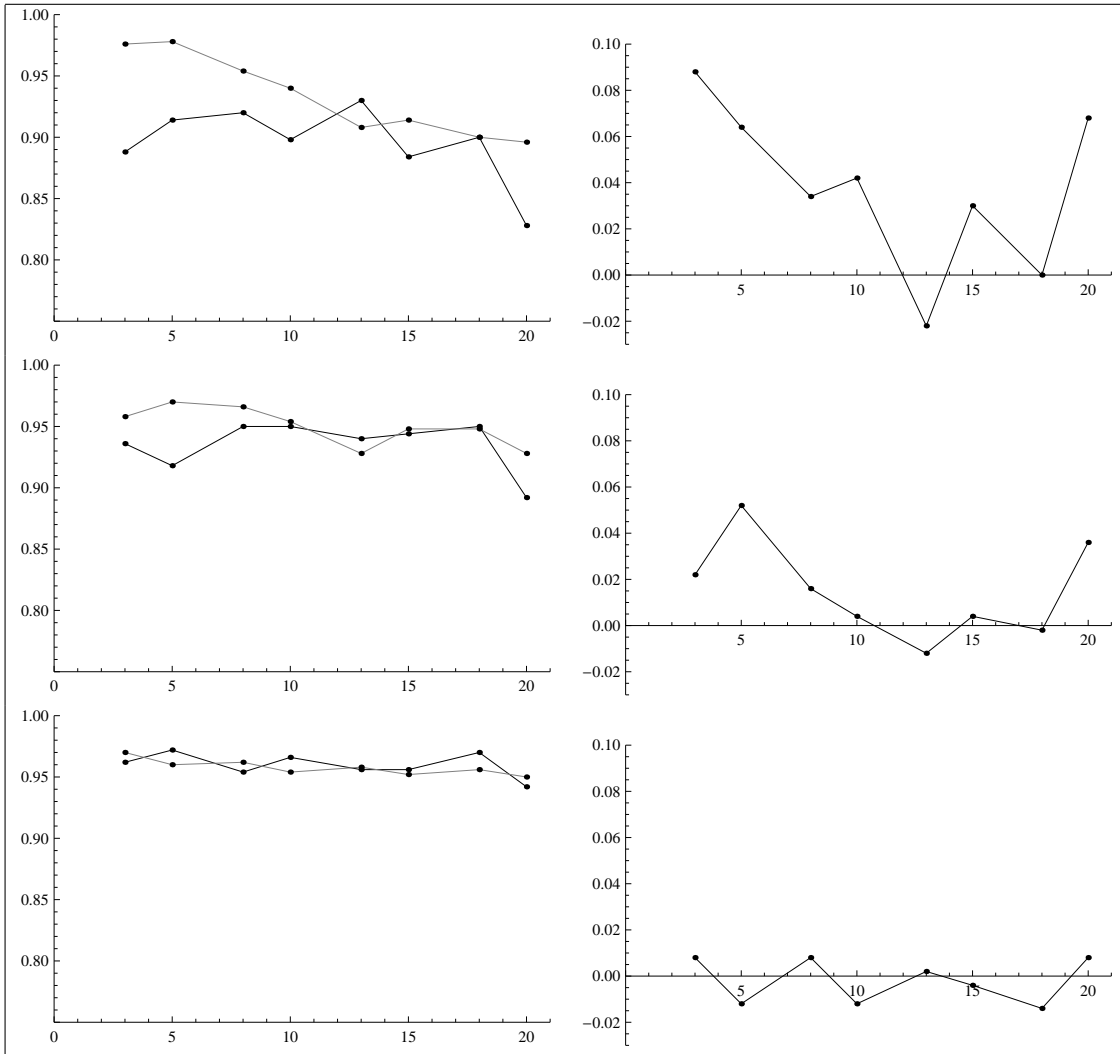


Figure 6: Model M_5 : ACPs of simultaneous equal-tailed percentile bootstrap confidence intervals for μ_i ($i = 1, \dots, d$) vs. block length b (left column). cGSBB method (black) and ECBB (gray). Right column: differences between ACPs obtained for ECBB and cGSBB. From top results for $n = 60, 120, 240$, respectively. Nominal coverage probability is 95%.

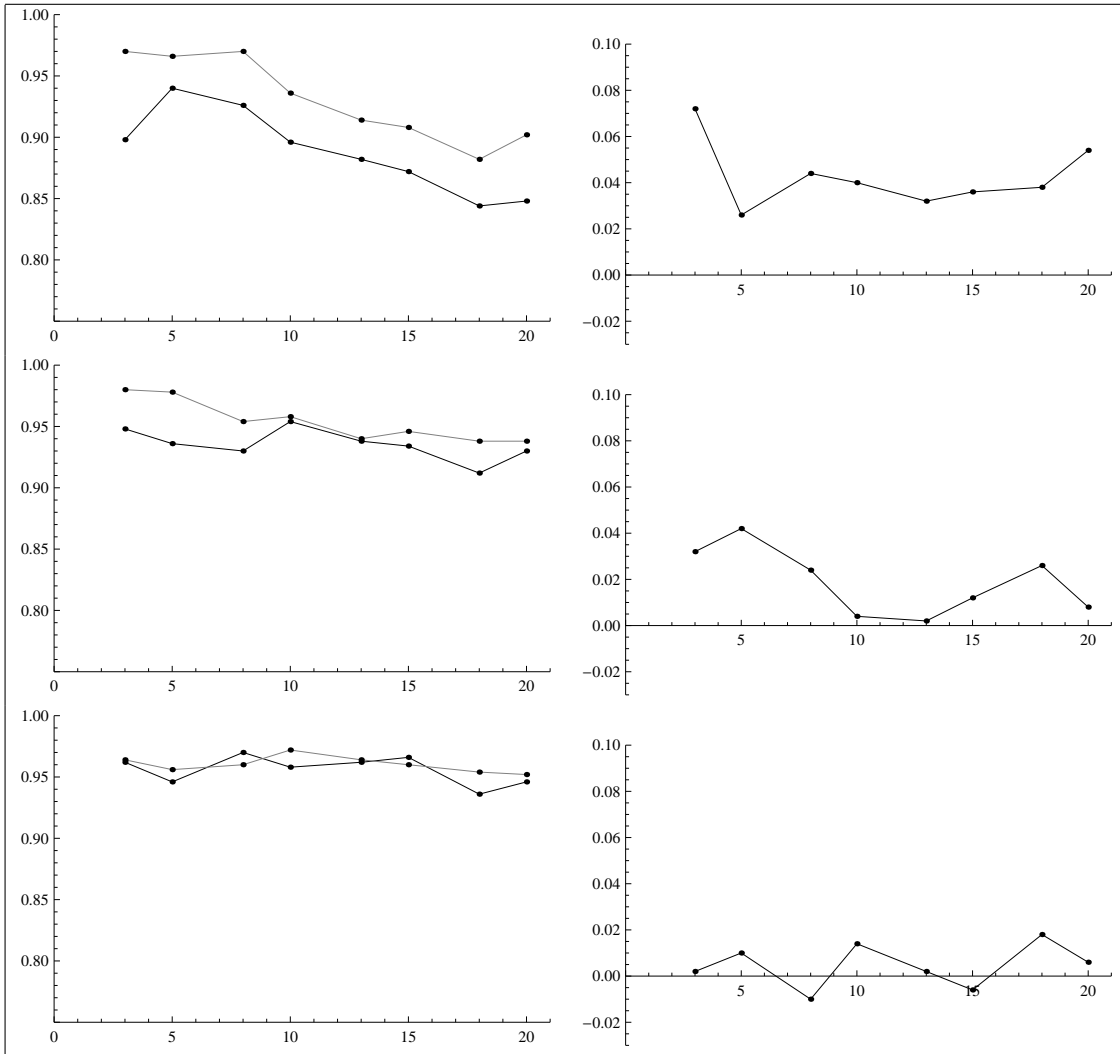


Figure 7: Model $M6$: ACPs of simultaneous equal-tailed percentile bootstrap confidence intervals for μ_i ($i = 1, \dots, d$) vs. block length b (left column). cGSBB method (black) and ECBB (gray). Right column: differences between ACPs obtained for ECBB and cGSBB. From top results for $n = 60, 120, 240$, respectively. Nominal coverage probability is 95%.

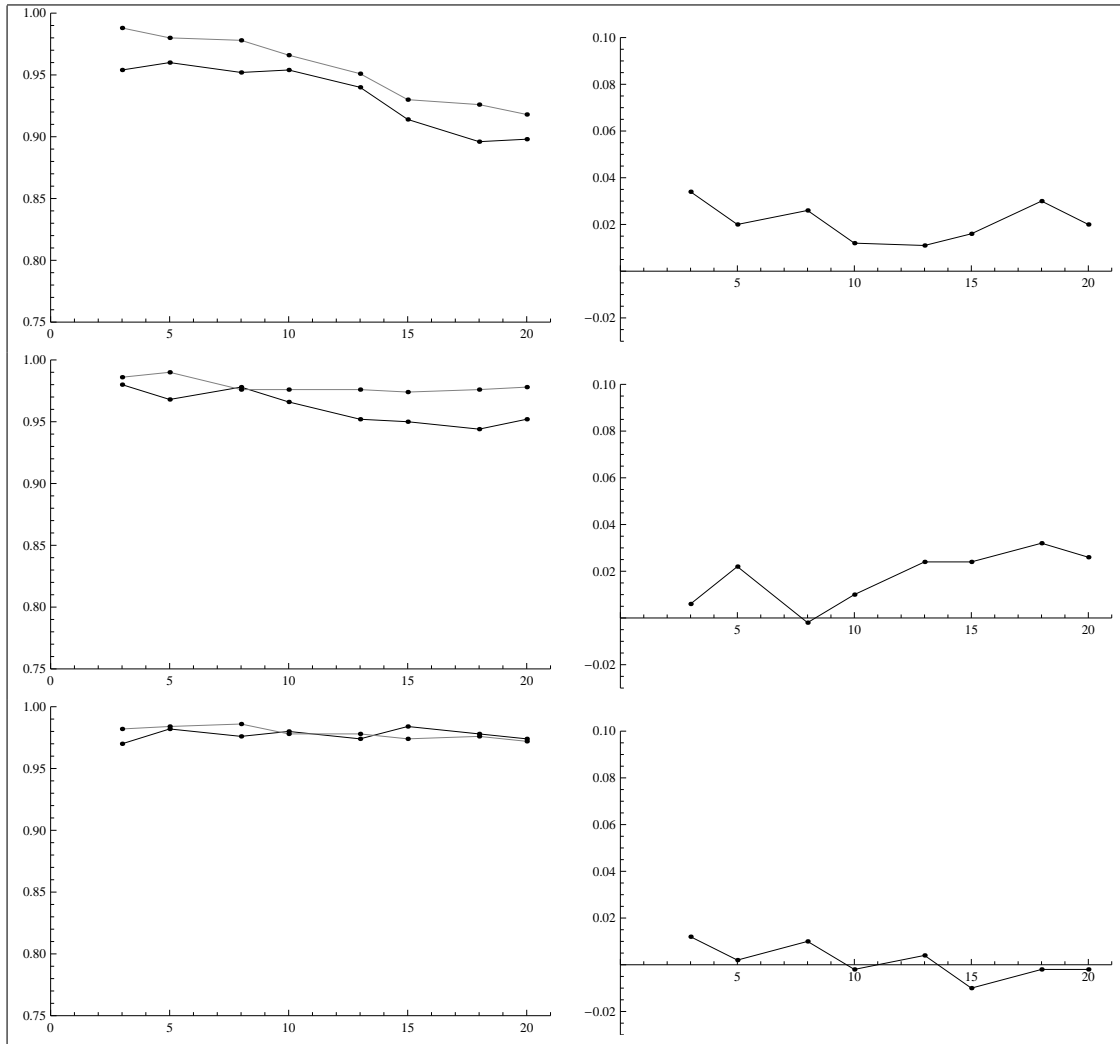


Figure 8: Model $M7$: ACPs of simultaneous equal-tailed percentile bootstrap confidence intervals for μ_i ($i = 1, \dots, d$) vs. block length b (left column). cGSBB method (black) and ECBB (gray). Right column: differences between ACPs obtained for ECBB and cGSBB. From top results for $n = 60, 120, 240$, respectively. Nominal coverage probability is 95%.

4 Real data example

In this section we present application of our results for energy market data. Energy market data is known to have periodic structures not only in the mean function but also in the autocovariance function (see e.g. Broszkiewicz-Suwaj et al. (2004)). We consider the series containing hourly observations from volumes of energy traded on the Nord Pool Spot exchange collected between 6 July and 31 August 2010. The detailed description of the series can be found in Dudek et al. (2015). Moreover, in the mentioned paper periodic ARMA model is fitted to these data.

In Figure 16 the considered data are presented. The length of the sample n is equal to 984, the period length d is 24 and hence the number of observed periods w is 41.

Using techniques described in Section 2 we constructed the 95% equal-tiled simultaneous bootstrap con-

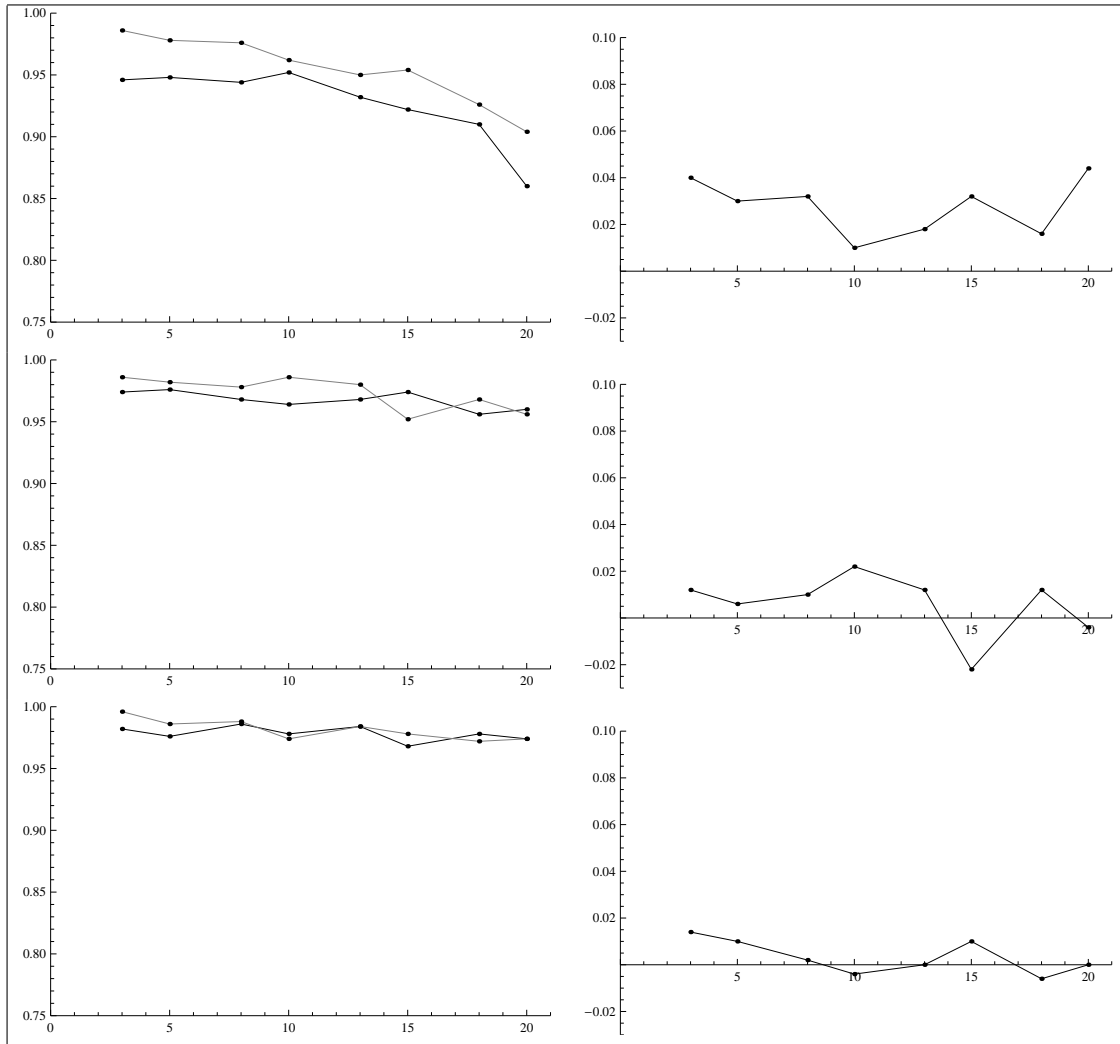


Figure 9: Model $M8$: ACPs of simultaneous equal-tailed percentile bootstrap confidence intervals for μ_i ($i = 1, \dots, d$) vs. block length b (left column). cGSBB method (black) and ECBB (gray). Right column: differences between ACPs obtained for ECBB and cGSBB. From top results for $n = 60, 120, 240$, respectively. Nominal coverage probability is 95%.

confidence intervals for seasonal means μ_s and seasonal variances σ_s^2 , $s = 1, \dots, 24$. To apply the cGSBB and the ECBB we set the block length $b = \lfloor \sqrt{n} \rfloor = 31$ and $b = \lfloor \sqrt[3]{n} \rfloor = 9$. Moreover, the number of bootstrap samples $B = 1000$. Problem of the block length choice is very difficult and so far in the literature for PC time series there is no result concerning this issue. However, this topic is well investigated for stationary processes (see e.g. Lahiri (2003)) and usually block lengths of order \sqrt{n} or $\sqrt[3]{n}$ are optimal. Since PC processes are periodically stationary we decided to use such kind of block lengths.

In Figures 17 and 18 the estimated values of μ_s and σ_s^2 together with obtained simultaneous confidence bands are presented. In the seasonal means case differences between the lower and the upper bounds obtained with the cGSBB and the ECBB are very small. The maximal observed difference is around 50, while the values of estimates range from 23 000 to 32 000. Moreover, all confidence bands are narrow. Independent of the chosen b and the bootstrap approach all 24 seasonal means are non-zero. None of the confidence bands are crossing the line $y = 0$.

The simultaneous confidence intervals for σ_s^2 , $s = 1, \dots, 24$ are quite wide. The lower bands are often

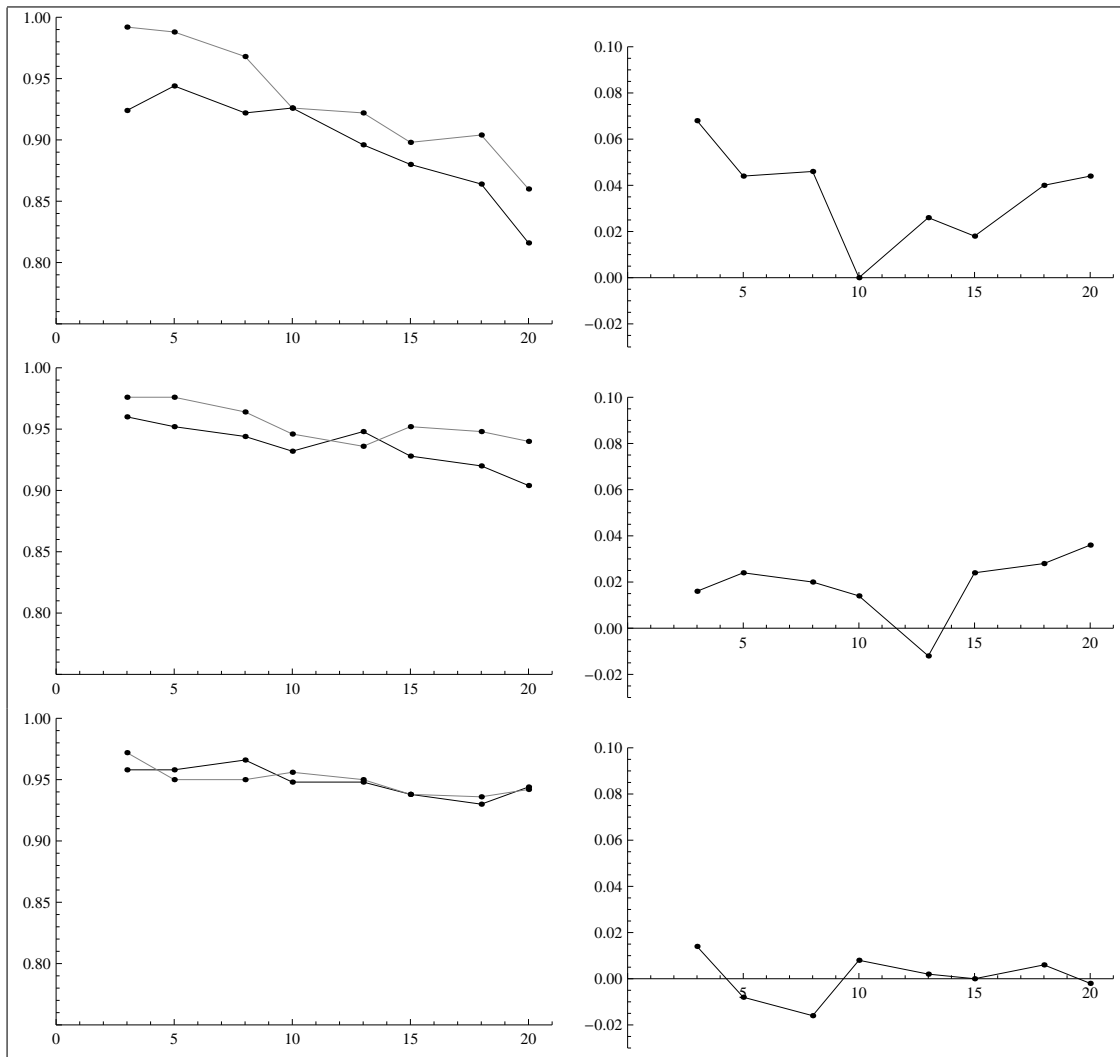


Figure 10: Model $M9$: ACPs of simultaneous equal-tailed percentile bootstrap confidence intervals for μ_i ($i = 1, \dots, d$) vs. block length b (left column). cGSBB method (black) and ECBB (gray). Right column: differences between ACPs obtained for ECBB and cGSBB. From top results for $n = 60, 120, 240$, respectively. Nominal coverage probability is 95%.

taking negative values. Since seasonal variances are always non-negative the lower bounds can be restricted from below by 0. However, for the sake of clarity we decided to present graphics in the current form. Independent of the block length and the bootstrap method significant values of σ_s^2 are observed for $s = 7, 8, 9, 21, 22$.

In the all considered cases independently of the considered parameter the ECBB confidence intervals are wider than the ones obtained with the cGSBB. As we discussed in Section 3 this property can often be observed (see also Dudek and Potorski (2014)). Since for now there is no method of the block choice for any bootstrap approach that can be applied for PC and APC time series, we recommend using the ECBB to those who prefer to be conservative.

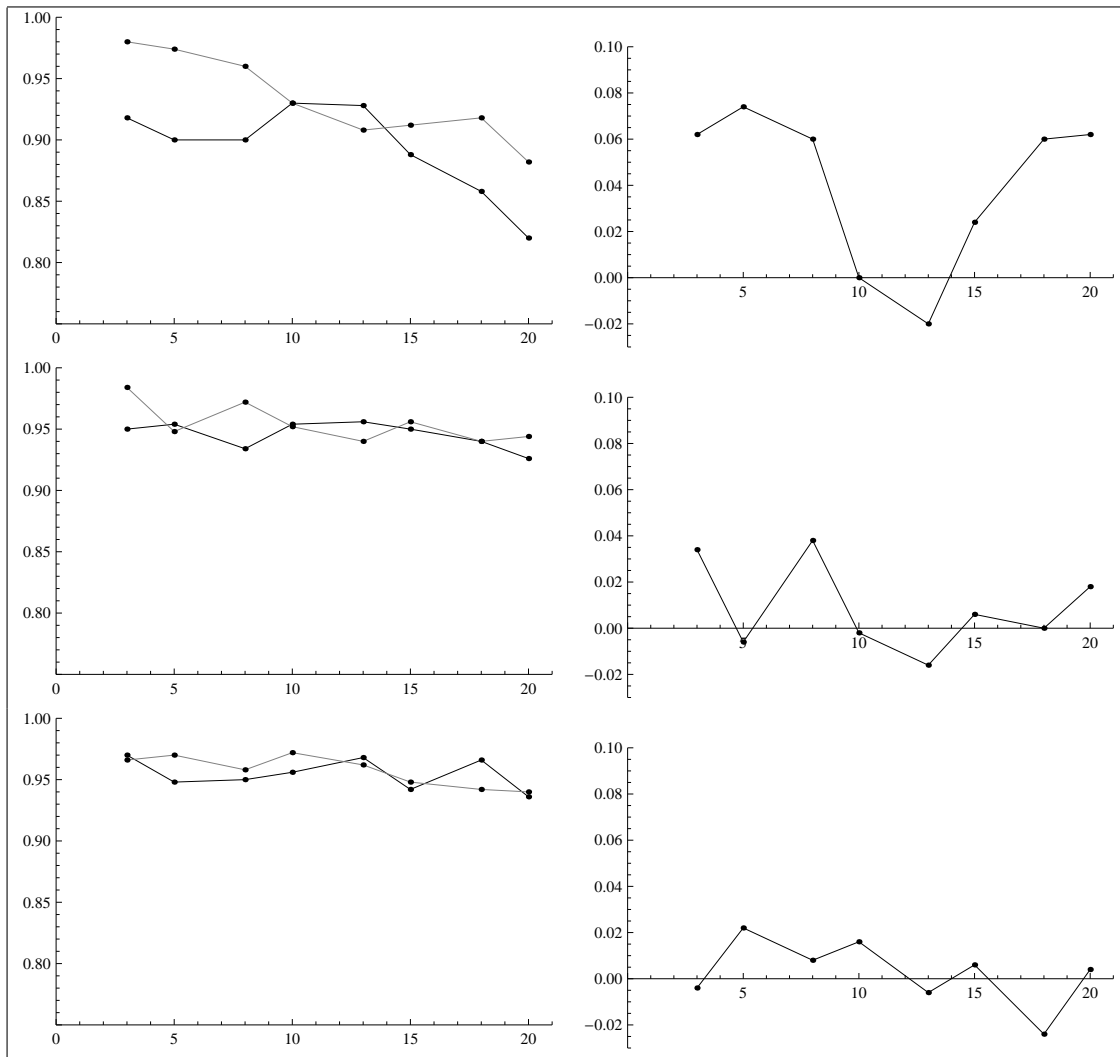


Figure 11: Model M_{10} : ACPs of simultaneous equal-tailed percentile bootstrap confidence intervals for μ_i ($i = 1, \dots, d$) vs. block length b (left column). cGSBB method (black) and ECBB (gray). Right column: differences between ACPs obtained for ECBB and cGSBB. From top results for $n = 60, 120, 240$, respectively. Nominal coverage probability is 95%.

Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 655394.

5 Appendix

For the sake of clarity and to simplify the notation we present all the proofs for the ECBB and the cGSBB. To get the corresponding results for the EMBB and the GSBB it is enough to change number of blocks that can be selected. For example for the ECBB we have n different blocks and for the EMBB $n - b + 1$. Moreover, we assume that the sample size is an integer multiple of the block length ($n = bl$) and that each block contains v full periods, i.e. $b = vd + r_2$, where $r_2 \in \{0, 1, \dots, d - 1\}$. Thus, in each

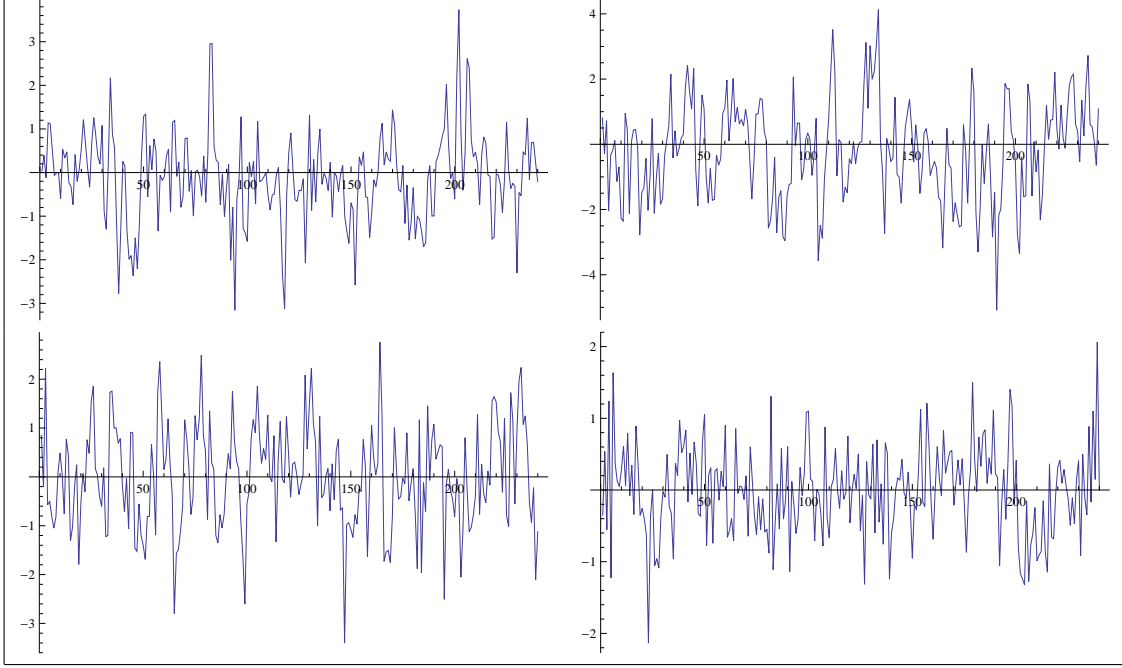


Figure 12: Examples of realizations: from top V1, V3 (left column) and V2, V4 (right column). Sample size $n = 240$.

selected block there are v or $v + 1$ observations from each season.

Proof of Theorem 2.1:

Below we present the asymptotic normality result for the seasonal means case under α -mixing assumption. Since the proof of the theorem is the simplified version of the proof of Theorem 5.2, we omit the details.

Theorem 5.1 *Assume that $\{X_t, t \in \mathcal{Z}\}$ is a PC, and α -mixing time series such that*

$$(i) \sup_{t \in \mathcal{Z}} \mathbb{E}|X_t|^{2+\delta} < \infty \text{ for some } \delta > 0;$$

$$(ii) \sum_{\tau=1}^{\infty} \alpha_X^{\delta/(2+\delta)}(\tau) < \infty;$$

Then

$$\sqrt{w}(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}) \xrightarrow{d} N_d(0, \boldsymbol{\Sigma}(\mathbf{s})),$$

where $\boldsymbol{\Sigma}(\mathbf{s}) = [\sigma_{ij}]_{i,j=1,\dots,d}$, $\sigma_{ii} = \sigma^2(s_i)$ and

$$\sigma_{ij} = \sum_{k=-\infty}^{\infty} \text{Cov}(X_{s_i}, X_{s_j+kd}).$$

One may notice that under the assumptions of Theorem 2.1, conditions (i) and (ii) of Theorem 5.1 are fulfilled. First we show the consistency in the one dimensional case. Without loss of generality we consider only the first season. We will show that

$$\rho_1(\mathcal{L}(\sqrt{w}(\hat{\mu}_1 - \mu_1)), \mathcal{L}^*(\sqrt{w}(\hat{\mu}_1^{*ECBB} - \hat{\mu}_1))) \xrightarrow{p} 0 \quad \text{as } n \rightarrow \infty. \quad (12)$$

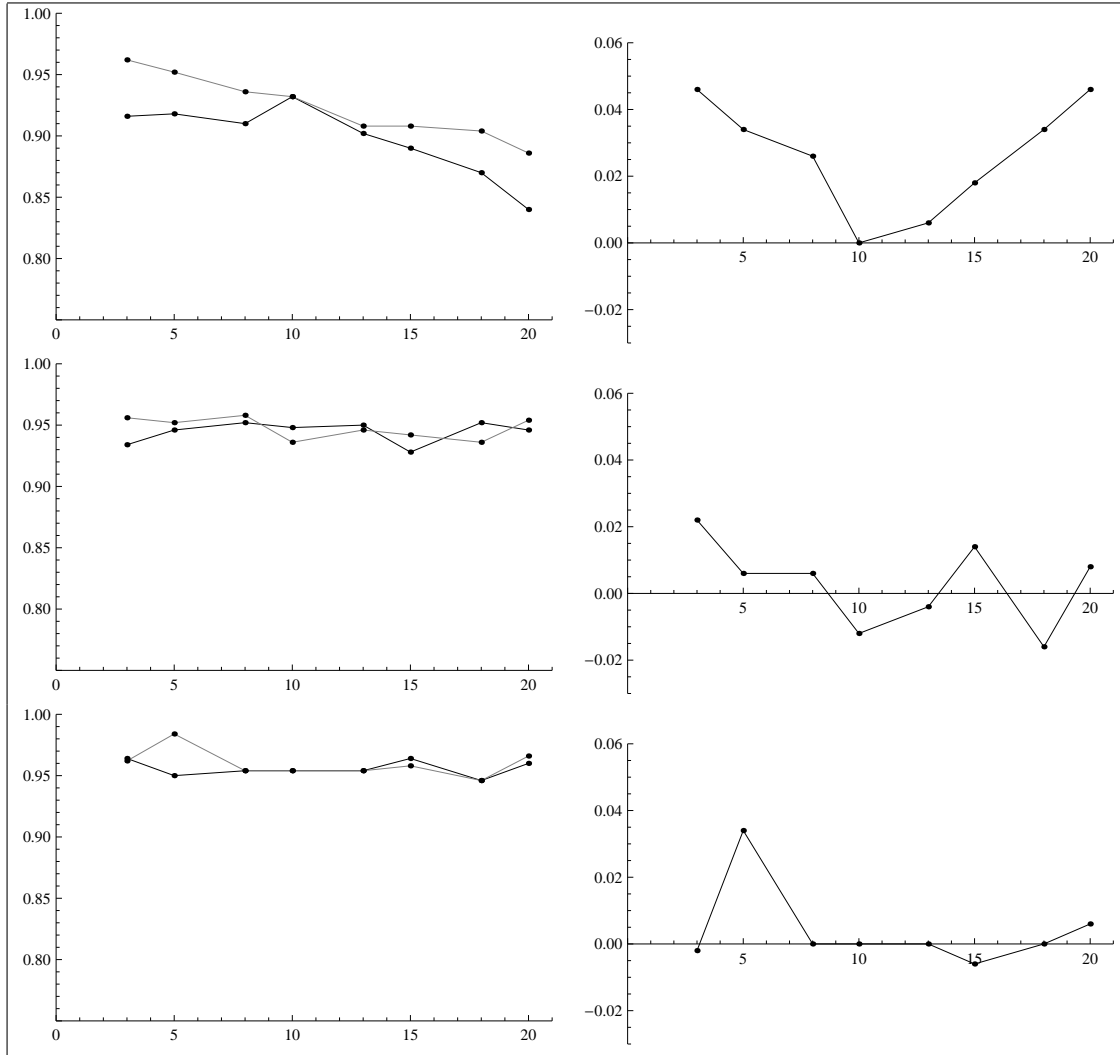


Figure 13: Model V1: ACPs of simultaneous equal-tailed percentile bootstrap confidence intervals for σ_i^2 ($i = 1, \dots, d$) vs. block length b (left column). cGSBB method (black) and ECBB (gray). Right column: differences between ACPs obtained for ECBB and cGSBB. From top results for $n = 60, 120, 240$, respectively. Nominal coverage probability is 95%.

In our consideration we follow the notation introduced in Dudek et al. (2014a).

Our sample of size n is divided into l disjoint blocks of the length b ($n = lb$). Let \tilde{U}_i and \tilde{U}_i^* , $i = 1, \dots, l$ be the sum of all observations from the first season contained in the i -th block of the original sample and the bootstrap sample, respectively. Then, the estimator (1) and its bootstrap counterpart (3) can be equivalently rewritten as

$$\hat{\mu}_1 = \frac{1}{w_1} \sum_{i=1}^l \tilde{U}_i \quad \text{and} \quad \hat{\mu}_1^{*ECBB} = \frac{1}{v_1^*} \sum_{i=1}^l \tilde{U}_i^*.$$

To get desired consistency of $\hat{\mu}_1^{*ECBB}$ we introduce another bootstrap estimator of μ_1 , which is of the form

$$\tilde{\mu}_1^{*ECBB} = \frac{1}{w_1} \sum_{i=1}^l \tilde{U}_i^*.$$

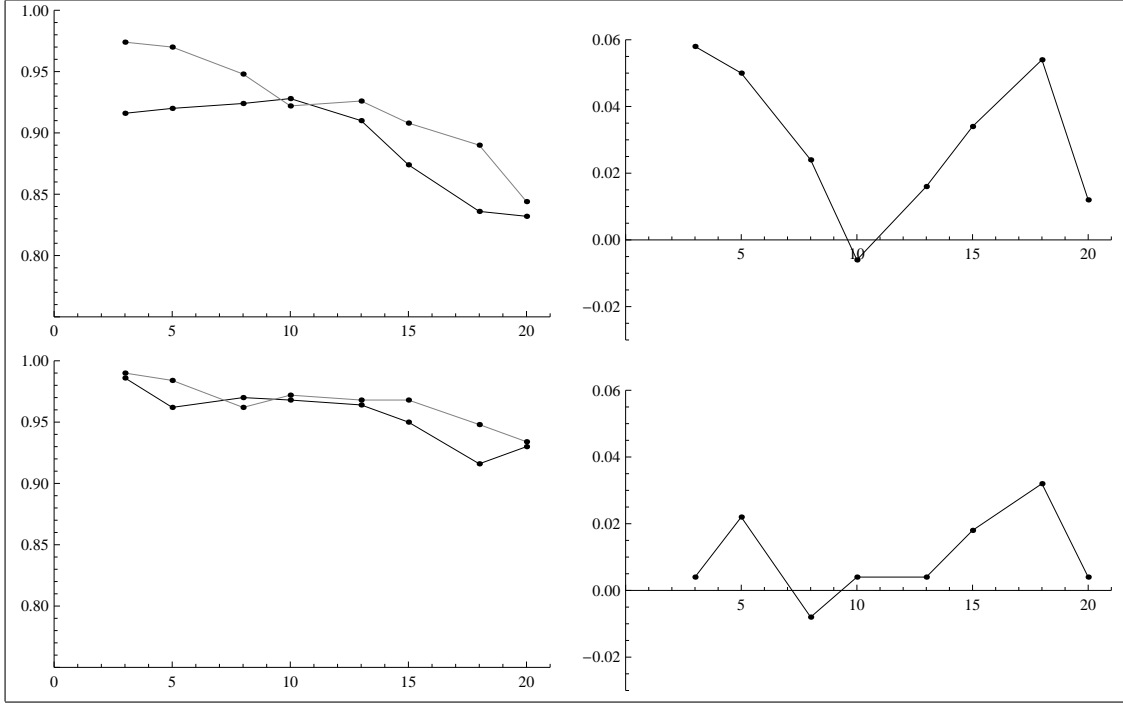


Figure 14: Model V2, V3: ACPs of simultaneous equal-tailed percentile bootstrap confidence intervals for σ_i^2 ($i = 1, \dots, d$) vs. block length b (left column). cGSBB method (black) and ECBB (gray). Right column: differences between ACPs obtained for ECBB and cGSBB. In first and second row results for V2 and V3, respectively. Sample size $n = 60$. Nominal coverage probability is 95%.

Let $U_i = \tilde{U}_i - E\tilde{U}_i$ and $U_i^* = \tilde{U}_i^* - E^*\tilde{U}_i^*$ for $i = 1, \dots, l$. Using Corollary 2.4.8 in Araujo and Giné (1980) and following the reasoning proposed in Dudek et al. (2014a) (see Theorem 3.1), one may easily obtain the consistency of $\tilde{\mu}_1^{*ECBB}$, i.e.

$$\rho_1(\mathcal{L}(\sqrt{w}(\hat{\mu}_1 - \mu_1)), \mathcal{L}^*(\sqrt{w}(\tilde{\mu}_1^{*ECBB} - \hat{\mu}_1))) \xrightarrow{P} 0 \quad \text{as } n \rightarrow \infty. \quad (13)$$

To obtain consistency of $\tilde{\mu}_1^{*ECBB}$ we use the Conditional Slutsky's Theorem (see Lemma 4.1 in Lahiri (2003)). Multiplying $\tilde{\mu}_1^{*ECBB}$ by w_1/v_1^* , one gets $\tilde{\mu}_1^{*ECBB}$. Note that $P^*(vl \leq v_1^* \leq (v+1)l) = 1$. Let us recall that $n = wd + r_1$. Since $l/w \rightarrow 0$ as $n \rightarrow \infty$, we get that

$$\frac{v_1^*}{vl} \xrightarrow{P^*} 1 \quad \text{and} \quad \frac{v_1^*}{w} \xrightarrow{P^*} 1 \quad \text{as } n \rightarrow \infty.$$

Moreover, $w/w_1 \rightarrow 1$ as $n \rightarrow \infty$ and

$$E^*\left(\sum_{i=1}^l \tilde{U}_i^*\right) = \sum_{i=1}^l E^*(\tilde{U}_i^*) = \sum_{i=1}^l \frac{1}{n} \sum_{j=1}^n \tilde{U}_j = \frac{bl}{n} w \hat{\mu}_1 = w \hat{\mu}_1.$$

Thus,

$$\sqrt{w} \left| \frac{1}{v_1^*} E^*\left(\sum_{i=1}^l \tilde{U}_i^*\right) - \hat{\mu}_1 \right| = \sqrt{w} \left| \frac{1}{v_1^*} w \hat{\mu}_1 - \hat{\mu}_1 \right| = \sqrt{w} |\hat{\mu}_1| \left| \frac{w}{v_1^*} - 1 \right|.$$

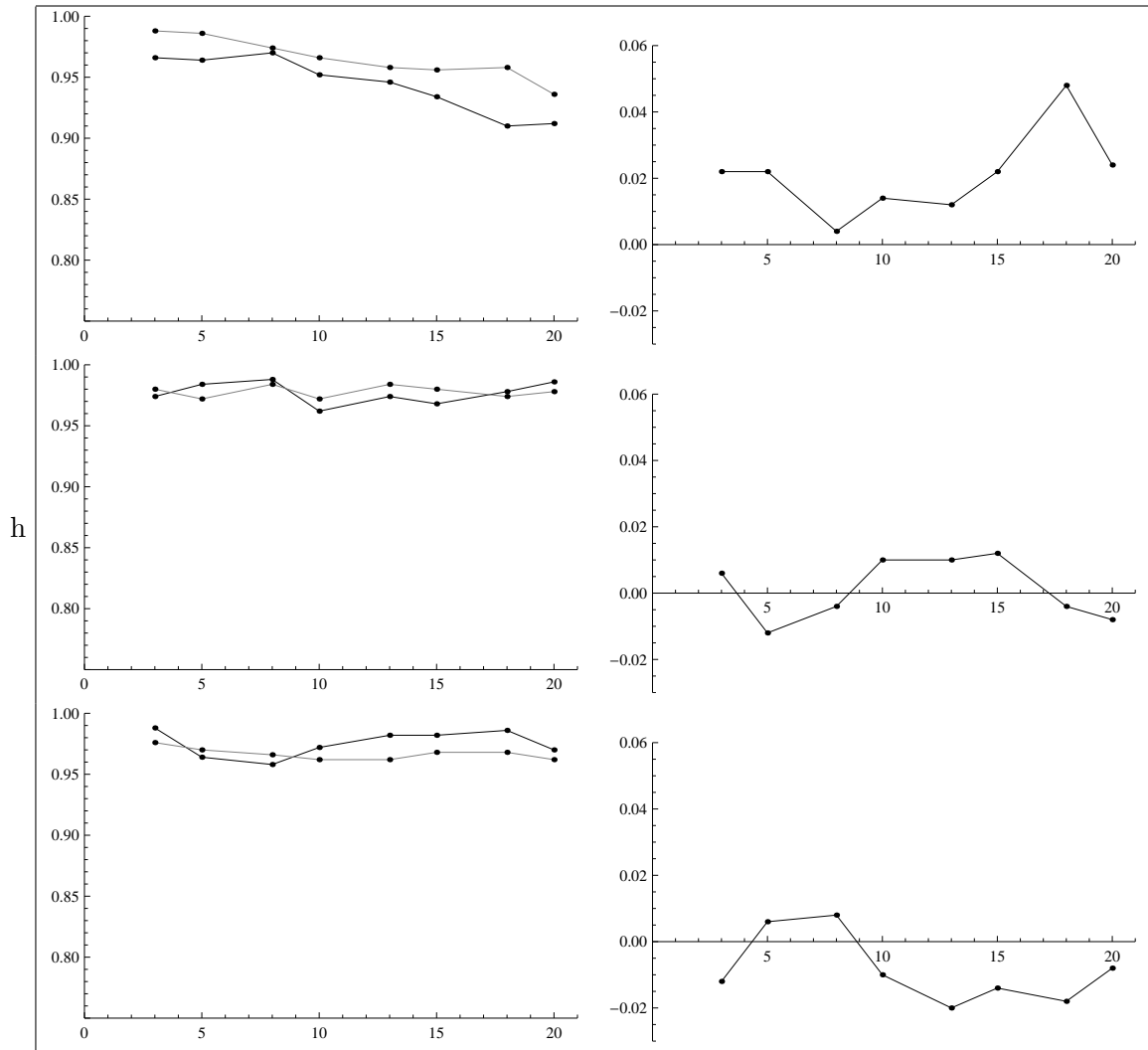


Figure 15: Model V4: ACPs of simultaneous equal-tailed percentile bootstrap confidence intervals for σ_i^2 ($i = 1, \dots, d$) vs. block length b (left column). cGSBB method (black) and ECBB (gray). Right column: differences between ACPs obtained for ECBB and cGSBB. From top results for $n = 60, 120, 240$, respectively. Nominal coverage probability is 95%.

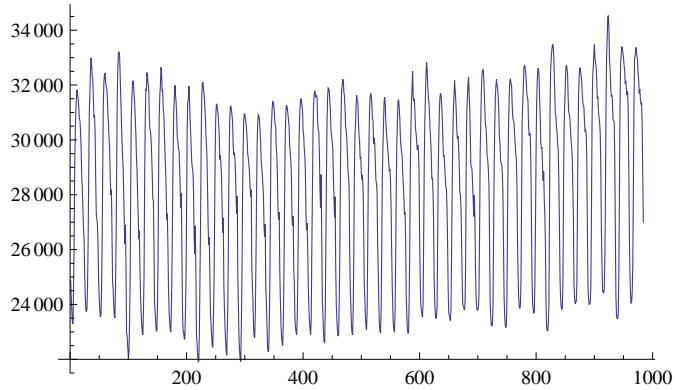


Figure 16: Volumes of energy traded hourly on the Nord Pool Spot Exchange from July 6th 2010 to August 31st 2010.

Finally, from the Conditional Slutsky's Theorem we have that

$$\sqrt{w} \left| \frac{1}{v_1^*} \mathbf{E}^* \left(\sum_{i=1}^l \tilde{U}_i^* \right) - \hat{\mu} \right| \xrightarrow{P^*} 0 \quad \text{as } n \rightarrow \infty$$

and simultaneously we get (12). Since to obtain the consistency of $\hat{\mu}^{*ECBB}$ it is enough to follow the proof presented in Dudek et al. (2014a), we decided to omit technical details. \square

Proof of Theorems 2.2 and 2.3:

Since Theorem 2.2 is a special case of Theorem 2.3, we present only the proof of the latter one. The asymptotic normality of $\hat{B}(s, \tau)$ for strictly periodic time series under ϕ -mixing assumption was shown in Hurd and Miamee (2007) (see Proposition 9.13). Below we present the asymptotic result for α -mixing PC time series.

Theorem 5.2 *Assume that $\{X_t, t \in \mathcal{Z}\}$ is a PC, WP(3), WP(4) and α -mixing time series such that*

(i) $\sup_{t \in \mathcal{Z}} \mathbf{E}|X_t|^{4+2\delta} < \infty$ for some $\delta > 0$;

(ii) $\sum_{\tau=1}^{\infty} \alpha_X^{\delta/(2+\delta)}(\tau) < \infty$;

Then

$$\sqrt{w} \left(\hat{B}(s, \tau) - B(s, \tau) \right) \xrightarrow{d} N_r(0, \Sigma(s, \tau))$$

for any s, τ for which $\det \Sigma(s, \tau) > 0$, where $\Sigma(s, \tau) = [\sigma_{ij}]_{i,j=1,\dots,r}$, $\sigma_{ii} = \sigma^2(s_i, \tau_i)$ and

$$\sigma_{ij} = \sum_{k=-\infty}^{\infty} \text{Cov}(Z(s_i, \tau_i), Z(s_j + kd, \tau_j)),$$

where $Z(s, \tau) = (X_s - \mu_s)(X_{s+\tau} - \mu_{s+\tau})$.

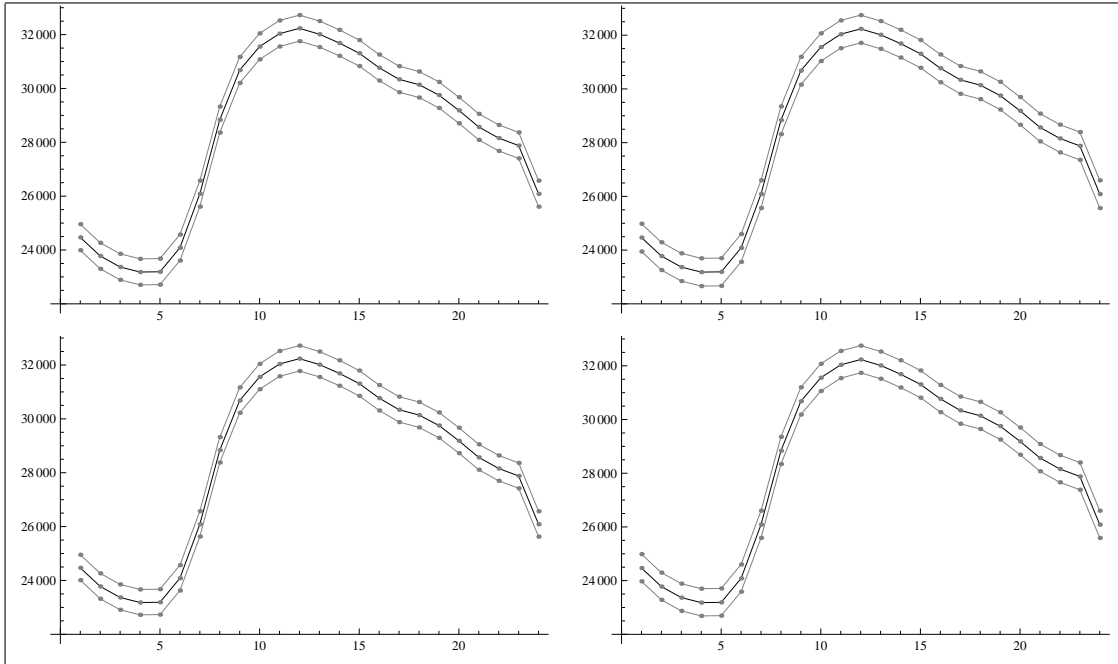


Figure 17: Estimated values of μ_i ($i = 1, \dots, 24$) (black) together with 95% simultaneous equal-tailed percentile bootstrap confidence intervals for (grey). Results for cGSBB and ECBB in first and second row, respectively. Block length $b = 9$ (left column) and $b = 31$ (right column).

Proof:

At first we consider one dimensional case, i.e. $r = 1$.

We start with the decomposition of the estimator $\widehat{B}(s, \tau)$. Note that

$$\begin{aligned}
 \widehat{B}(s, \tau) &= \frac{1}{w} \sum_{k=0}^{w-1} (X_{s+kd} - \widehat{\mu}_s) (X_{s+\tau+kd} - \widehat{\mu}_{s+\tau}) \\
 &= \frac{1}{w} \sum_{k=0}^{w-1} (X_{s+kd} - \mu_s) (X_{s+\tau+kd} - \mu_{s+\tau}) - \frac{1}{w} \sum_{k=0}^{w-1} (X_{s+kd} - \mu_s) (\widehat{\mu}_{s+\tau} - \mu_{s+\tau}) \\
 &\quad - \frac{1}{w} \sum_{k=0}^{w-1} (\widehat{\mu}_s - \mu_s) (X_{s+\tau+kd} - \mu_{s+\tau}) + \frac{1}{w} \sum_{k=0}^{w-1} (\widehat{\mu}_s - \mu_s) (\widehat{\mu}_{s+\tau} - \mu_{s+\tau}) \\
 &= \widetilde{B}(s, \tau) - I - II + III.
 \end{aligned}$$

Moreover,

$$\begin{aligned}
 \mathbb{E} |I| &\leq \sqrt{\mathbb{E} (\widehat{\mu}_{s+\tau} - \mu_{s+\tau})^2 \mathbb{E} \left(\frac{1}{w} \sum_{k=0}^{w-1} (X_{s+kd} - \mu_s) \right)^2} \\
 &= \sqrt{\text{Var} (\widehat{\mu}_{s+\tau}) \mathbb{E} \left(\frac{1}{w} \sum_{k=0}^{w-1} (X_{s+kd} - \mu_s) \right)^2}.
 \end{aligned}$$

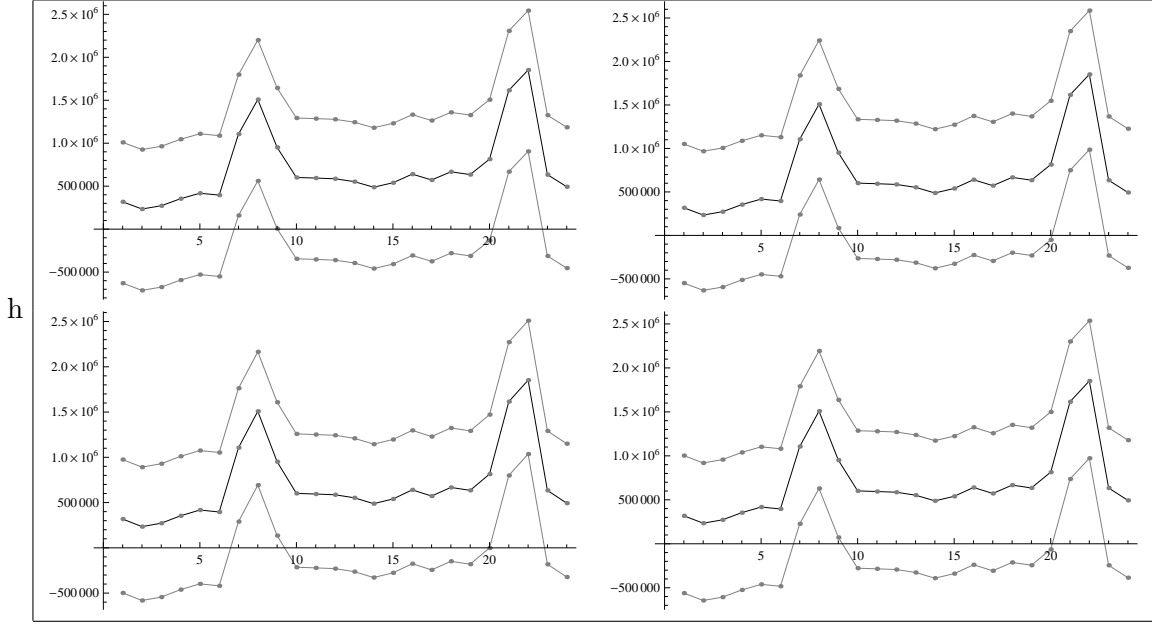


Figure 18: Estimated values of σ_i^2 ($i = 1, \dots, 24$) (black) together with 95% simultaneous equal-tailed percentile bootstrap confidence intervals for (grey). Results for cGSBB and ECBB in first and second row, respectively. Block length $b = 9$ (left column) and $b = 31$ (right column).

Under assumptions (i) and (ii), $E\left(\frac{1}{\sqrt{w}} \sum_{k=0}^{w-1} (X_{s+kd} - \mu_s)\right)^2$ is bounded (see Kim (1994)) by a constant independent on n . Additionally, using Lemma A.0.1 from Politis et al. (1999) we have that for any $s = 1, \dots, d$

$$\begin{aligned} \text{Var}(\hat{\mu}_s) &= \text{Var}\left(\frac{1}{w_s} \sum_{i=0}^{w_s-1} X_{s+id}\right) = \frac{1}{w_s} \frac{1}{w_s} \sum_{i=0}^{w_s-1} \sum_{j=0}^{w_s-1} \text{Cov}(X_{s+id}, X_{s+jd}) \\ &\leq \frac{1}{w_s} \frac{1}{w_s} \sum_{i=0}^{w_s-1} \sum_{k=-w_s+2}^{w_s-2} |\text{Cov}(X_{s+id}, X_{s+(i+|k|)d})| \\ &\leq \frac{1}{w_s} \frac{1}{w_s} \sum_{i=0}^{w_s-1} \sum_{k=-w_s+2}^{w_s-2} \alpha_X^{\frac{\delta}{2+\delta}}(|k|) = O\left(\frac{1}{w_s}\right) = O\left(\frac{1}{w}\right). \end{aligned}$$

Thus, $E|I| = O(1/w)$. Analogically one may get that $E|II| = O(1/w)$. Finally,

$$E|III| \leq \sqrt{E(\hat{\mu}_{s+\tau} - \mu_{s+\tau})^2} E(\hat{\mu}_{s+\tau} - \mu_{s+\tau})^2 \leq \sqrt{\text{Var}(\hat{\mu}_s) \text{Var}(\hat{\mu}_{s+\tau})} = O\left(\frac{1}{w}\right).$$

Therefore, $\sqrt{w} \left(E\left(\hat{B}(s, \tau)\right) - B(s, \tau)\right) = O_P(1/\sqrt{w})$ and

$$\sqrt{w} \left(\hat{B}(s, \tau) - B(s, \tau)\right) = \sqrt{w} \left(\tilde{B}(s, \tau) - B(s, \tau)\right) + O_P\left(\frac{1}{\sqrt{w}}\right).$$

Since X_t is WP(4), the series $Z_k(s, \tau) = (X_{s+kd} - \mu_s)(X_{s+\tau+kd} - \mu_{s+\tau})$ is PC. Moreover, $Z_k(s, \tau)$ is α -mixing with the mixing coefficients $\alpha_Z(t) = \alpha_X(td - \tau)$. Finally, $Z_k(s, \tau)$ has the absolute moments of

order $2 + \delta$ uniformly bounded. Under assumption (i) we have

$$\begin{aligned} \mathbb{E} |Z_k(s, \tau)|^{2+\delta} &= \mathbb{E} |(X_{s+kd} - \mu_s)(X_{s+\tau+kd} - \mu_{s+\tau})|^{2+\delta} \\ &\leq \sqrt{\mathbb{E} |X_{s+kd} - \mu_s|^{4+2\delta} \mathbb{E} |X_{s+\tau+kd} - \mu_{s+\tau}|^{4+2\delta}} \end{aligned}$$

and hence $\sup_k \mathbb{E} |Z_k(s, \tau)|^{2+\delta} < \infty$.

As a straightforward consequence of Lemma 1 from Lenart et al. (2008) we get that

$$\text{Var} \left(\frac{1}{\sqrt{w}} \sum_{k=1}^w Z_k(s, \tau) \right) \rightarrow \sigma^2(s, \tau) \quad \text{as} \quad n \rightarrow \infty.$$

Finally, the conclusion of the theorem is obtained using Theorem 3.3.1 from Guyon (1995).

To get the thesis for the multidimensional case one needs to use Cramér-Wold device and the fact that $Z_k(s_i, \tau_i)$ and $Z_k(s_j, \tau_j)$ for $i \neq j$ are jointly cyclostationary (for definition see Definition 1 in Lenart et al. (2008)). Using Lemma 2 from Lenart et al. (2008) one may easily obtain the thesis, thus we omit the details. \square

Notice that under the assumptions of Theorem 2.3 the asymptotic normality of $\widehat{B}(s, \tau)$ holds. Without loss of generality and for the sake of simplicity from now on we assume that the considered time series X_t is zero mean.

First we prove Theorem 2.3 in the one dimensional case for the ECBB. Without loss of generality we take $s = 1$ and $\tau \geq 0$. For the sake of simplicity, we omit the superscript *ECBB*. We show that

$$\begin{aligned} \sup_{x \in \mathcal{R}} \left| P \left(\sqrt{w} \left(\widehat{B}(1, \tau) - B(1, \tau) \right) \leq x \right) \right. \\ \left. - P^* \left(\sqrt{w} \left(\widehat{B}^*(1, \tau) - \mathbb{E}^* \widehat{B}^*(1, \tau) \right) \leq x \right) \right| \xrightarrow{p} 0. \end{aligned} \quad (14)$$

To get the desired consistency we introduce the modified versions of $\widehat{B}^*(1, \tau)$. To understand the idea of their construction let us recall formula (9)

$$\widehat{B}^*(1, \tau) = \frac{1}{v_1^*} \sum_{t \notin C_{1,b,\tau}^*} X_t^* X_{t+\tau}^* + \frac{1}{v_1^*} \sum_{t \in C_{1,b,\tau}^*} X_t^* X_{t+\tau}^*.$$

New estimators $\widehat{B}_1^*(1, \tau)$ and $\widetilde{B}^*(1, \tau)$ are created by replacing v_1^* by w and then removing the second summand on the right-hand.

Thus,

$$\widehat{B}_1^*(1, \tau) = \frac{1}{w} \sum_{t \notin C_{1,b,\tau}^*} X_t^* X_{t+\tau}^* + \frac{1}{w} \sum_{t \in C_{1,b,\tau}^*} X_t^* X_{t+\tau}^*$$

and

$$\widetilde{B}^*(1, \tau) = \frac{1}{w} \sum_{t \notin C_{1,b,\tau}^*} X_t^* X_{t+\tau}^*.$$

Estimator $\widetilde{B}^*(1, \tau)$ is based only on those elements $X_t^* X_{t+\tau}^*$ for which X_t^* and $X_{t+\tau}^*$ belong to the same block. Those that we skip are based on this part of the bootstrap sample, for which the dependence

structure contained in the original data was destroyed as a result of joining selected blocks. Similar ideas of asymptotically equivalent estimators can be found in Dudek (2015) and Dudek et al. (2014b) for the estimator of the autocovariance function coefficients.

We show that

$$\left| \sqrt{w} \left(\widehat{B}_1^*(1, \tau) - \widehat{B}(1, \tau) \right) - \sqrt{w} \left(\widetilde{B}^*(1, \tau) - \mathbf{E}^* \widetilde{B}^*(1, \tau) \right) \right| \xrightarrow{P^*} 0 \quad (15)$$

and hence to get (14) it will be enough to show consistency of $\widetilde{B}^*(1, \tau)$.

Before we proceed we decompose $\widetilde{B}^*(1, \tau)$ into sums over disjoint blocks of the length b . We have

$$\begin{aligned} \widetilde{B}^*(1, \tau) &= \frac{1}{w} \sum_{t \notin C_{1,b,\tau}^*} X_t^* X_{t+\tau}^* = \frac{1}{w} \left(\sum_{\substack{t \notin C_{1,b,\tau}^* \\ 1 \leq t \leq b}} X_t^* X_{t+\tau}^* + \cdots + \sum_{\substack{t \notin C_{1,b,\tau}^* \\ (l-1)b+1 \leq t \leq lb}} X_t^* X_{t+\tau}^* \right) = \\ &= \frac{1}{w} \left(\widetilde{Z}_{1,b}^* + \widetilde{Z}_{b+1,b}^* + \cdots + \widetilde{Z}_{(l-1)b+1,b}^* \right). \end{aligned}$$

The first subscript in $\widetilde{Z}_{i,b}^*$ indicates the index of the first observation in the block. Thus, $\widetilde{Z}_{i,b}^*$ is the sum of the observations from the first season contained in the block $B_i^* = (X_i^*, \dots, X_{i+b-1}^*)$. Note that $\widetilde{Z}_{i,b}^*$ are conditionally independent and

$$P^* \left(\widetilde{Z}_{i,b}^* = \widetilde{Z}_{j,b}^* \right) = \frac{1}{n} \quad \text{for } j = 1, \dots, n,$$

where $\widetilde{Z}_{i,b}$ is the corresponding sum over the block B_i for the original data. Moreover, let

$$Z_{i,b} = \widetilde{Z}_{i,b} - \mathbf{E} \widetilde{Z}_{i,b}, \quad Z_{i,b}^* = \widetilde{Z}_{i,b}^* - \mathbf{E}^* \widetilde{Z}_{i,b}^*.$$

Note that

$$\sqrt{w} \left| \mathbf{E}^* \widetilde{B}^*(1, \tau) - \widehat{B}(1, \tau) \right| \longrightarrow 0 \quad \text{as } n \longrightarrow \infty.$$

and hence condition (15) is equivalent to

$$\frac{1}{w} \text{Var}^* \left(\sum_{t \in C_{1,b,\tau}^*} X_t^* X_{t+\tau}^* \right) \xrightarrow{P^*} 0 \quad \text{as } n \longrightarrow \infty.$$

In the proof of Theorem 2.1 we showed that $v_1^* \sim w$ as $n \rightarrow \infty$. Hence we get that $w/(v_1^*)^2 \sim 1/w$ as $n \rightarrow \infty$ and

$$\left| \frac{1}{w} \text{Var}^* \left(\sum_{t \in C_{1,b,\tau}^*} X_t^* X_{t+\tau}^* \right) - \frac{1}{w} \text{Var}^* \left(\sum_{t \in C_{1,b,\tau}^*} X_t^* X_{t+\tau}^* \right) \right| \xrightarrow{P^*} 0.$$

Moreover,

$$\begin{aligned} & \frac{1}{w} \text{Var}^* \left(\sum_{t \in C_{1,b,\tau}^*} X_t^* X_{t+\tau}^* \right) \\ &= \frac{2}{w} \sum_{k=0}^{l-2} \sum_{k' \geq k} \text{Cov}^* \left(\sum_{i \in TI_{1,k}} X_{(k+1)b-\tau+i}^* X_{(k+1)b+i}^*, \sum_{j \in TI_{1,k'}} X_{(k'+1)b-\tau+j}^* X_{(k'+1)b+j}^* \right), \end{aligned}$$

where $TI_{1,k} = \{i : 1 \leq i \leq \tau, (k+1)b - \tau + i \in TI_1\}$ and $TI_{1,k'} = \{j : 1 \leq j \leq \tau, (k'+1)b - \tau + j \in TI_1\}$. The only non-zero elements in the sum on the right hand-side are for $k' = k$ and $k' = k + 1$. In the first case we have

$$\begin{aligned} & \text{Var}^* \left(\sum_{i \in TI_{1,k}} X_{(k+1)b-\tau+i}^* X_{(k+1)b+i}^* \right) = \\ &= \mathbb{E}^* \left(\sum_{i \in TI_{1,k}} X_{(k+1)b-\tau+i}^* X_{(k+1)b+i}^* \right)^2 - \left(\mathbb{E}^* \left(\sum_{i \in TI_{1,k}} X_{(k+1)b-\tau+i}^* X_{(k+1)b+i}^* \right) \right)^2 \\ &= \frac{1}{n} \sum_{j_1=1}^n \frac{1}{n} \sum_{j_2=1}^n \left(\sum_{i \in TI_{1,j_1}} X_{j_1+b-\tau+i-1} X_{j_2+i-1} \right)^2 \\ &\quad - \left(\frac{1}{n} \sum_{j_3=1}^n \frac{1}{n} \sum_{j_4=1}^n \left(\sum_{i \in TI_{1,j_3}} X_{j_3+b-\tau+i-1} X_{j_4+i-1} \right) \right)^2. \end{aligned}$$

$TI_{1,j_a} = \{i : 1 \leq i \leq \tau, (j_a + b - \tau + i - 1) \bmod d = 1\}$ for $a = 1, 3$. Since the set TI_{1,j_a} contains at most τ elements, using assumption (ii) the absolute expected value of the last expression can be bounded from above by $C_1 \tau^2$, where C_1 is some positive constant independent of n .

For $k' = k + 1$ we get

$$\begin{aligned}
& \text{Cov}^* \left(\sum_{i \in TI_{1,k}} X_{(k+1)b-\tau+i}^* X_{(k+1)b+i}^*, \sum_{j \in TI_{1,k+1}} X_{(k+2)b-\tau+j}^* X_{(k+2)b+j}^* \right) \\
&= \text{E}^* \left(\sum_{i_1 \in TI_{1,k}} X_{(k+1)b-\tau+i_1}^* X_{(k+1)b+i_1}^* \sum_{i_2 \in TI_{1,k+1}}^{\tau} X_{(k+2)b-\tau+i_2}^* X_{(k+2)b+i_2}^* \right) \\
&- \text{E}^* \left(\sum_{i_1 \in TI_{1,k}} X_{(k+1)b-\tau+i_1}^* X_{(k+1)b+i_1}^* \right) \text{E}^* \left(\sum_{i_2 \in TI_{1,k+1}}^{\tau} X_{(k+2)b-\tau+i_2}^* X_{(k+2)b+i_2}^* \right) \\
&= \frac{1}{n} \sum_{j_1=1}^n \frac{1}{n} \sum_{j_2=1}^n \frac{1}{n} \sum_{j_3=1}^n \left(\sum_{i_1 \in TI_{1,j_1}} X_{j_1+b-\tau+i_1-1} X_{j_2+i_1-1} \sum_{i_2 \in TI_{1,j_2}} X_{j_2+b-\tau+i_2-1} X_{j_3+i_2-1} \right) \\
&- \left(\frac{1}{n} \sum_{j_4=1}^n \frac{1}{n} \sum_{j_5=1}^n \left(\sum_{i_3 \in TI_{1,j_4}} X_{j_4+b-\tau+i_3-1} X_{j_5+i_3-1} \right) \right) \\
&\cdot \left(\frac{1}{n} \sum_{j_6=1}^n \frac{1}{n} \sum_{j_7=1}^n \left(\sum_{i_4 \in TI_{1,j_5}} X_{j_6+b-\tau+i_4-1} X_{j_7+i_4-1} \right) \right).
\end{aligned}$$

The absolute expected value of the right-hand side is less than or equal to $C_2 \tau^2$, where C_2 is some positive constant independent of n . Thus,

$$\text{E} \left| \frac{1}{w} \text{Var}^* \left(\sum_{t \in C_{1,b,\tau}^*} X_t^* X_{t+\tau}^* \right) \right| = O \left(\frac{l}{w} \right) = O \left(\frac{1}{b} \right)$$

and we get the desired convergence (15). Additionally, using the Conditional Slutsky's Theorem (see Lahiri 2003, p. 77) we get that it is enough to show the consistency of $\tilde{B}^*(1, \tau)$, i.e.

$$\begin{aligned}
& \sup_{x \in \mathcal{R}} \left| P \left(\sqrt{w} \left(\hat{B}(1, \tau) - B(1, \tau) \right) \leq x \right) \right. \\
& \quad \left. - P^* \left(\sqrt{w} \left(\tilde{B}^*(1, \tau) - \text{E}^* \tilde{B}^*(1, \tau) \right) \leq x \right) \right| \xrightarrow{P} 0.
\end{aligned} \tag{16}$$

To get (16) we use Corollary 2.4.8 from Araujo and Giné (1980) and the fact that $v_1^*/w \rightarrow 1$ as $n \rightarrow \infty$. We need to show that for any $\nu > 0$

$$\sum_{k=0}^{l-1} P^* \left(\frac{1}{\sqrt{w}} |Z_{1+kb,b}^*| > \nu \right) \xrightarrow{P} 0, \tag{17}$$

$$\sum_{k=0}^{l-1} \text{E}^* \left(\frac{1}{\sqrt{w}} Z_{1+kb,b}^* \mathbf{1}_{\{|Z_{1+kb,b}^*| > \sqrt{w}\nu\}} \right) \xrightarrow{P} 0, \tag{18}$$

$$\sum_{k=0}^{l-1} \text{Var}^* \left(\frac{1}{\sqrt{w}} Z_{1+kb,b}^* \mathbf{1}_{\{|Z_{1+kb,b}^*| \leq \sqrt{w}\nu\}} \right) \xrightarrow{P} \sigma^2(1, \tau). \tag{19}$$

Using the same reasoning as in Dudek (2015) (see Theorem 2, proof of (14)-(16)) after some minor adjustments one may obtain (17)-(19). As the main steps of the proof are the same, we decided to omit technical details.

Finally, using same arguments as we presented in the proof of Theorem 2.1 one may obtain consistency of $\widehat{B}^*(1, \tau)$. To get the consistency in the multidimensional case, the Cramér-Wold device needs to be used. Same arguments as in the one dimensional case are used and hence again we skip the computation.

Following the main steps of just presented proof with minor adjustments corresponding to those presented in proof of Theorem 3.3 in Dudek et al. (2014b) one may obtain the consistency result for the GSBB. \square

Proof of Theorems 2.4 and 2.5:

To get thesis of both theorems it is enough to follow ideas presented in proofs of Theorem 4.3 from Dudek et al. (2014a) and Theorem 3.6 from Dudek (2015) and hence we omit the details. \square

References

- [Andrews(1984)] Andrews, D. (1984), 'Non-strong Mixing Autoregressive Processes', *J. Appl. Probab.*, 21, 930–934.
- [Antoni(2009)] Antoni, J. (2009), 'Cyclostationarity by examples', *Mech. Syst. Sig. Process.*, 23(4), 987–1036.
- [Araujo and Giné(1980)] Araujo, A. and Giné, E. (1980), *The central limit theorem for real and Banach valued random variables*, New York: Wiley.
- [Chan et al.(2004)Chan, Lahiri, and Meeker] Chan, V., Lahiri, S.N., and Meeker, W.Q (2004), 'Block bootstrap estimation of the distribution of cumulative outdoor degradation', *Technometrics*, 46, 215–224.
- [Broszkiewicz-Suwaj et al.(2004)Broszkiewicz-Suwaj, Makagon, Weron, and Wyżomańska] Broszkiewicz-Suwaj, E., Makagon, A., Weron, R., and Wyżomańska, A. (2004), 'On detecting and modeling periodic correlation in financial data', *Physica*, 336, 196–205.
- [Ciblat et al.(2002)Ciblat, Loubaton, Serpedin, and Giannakis] Ciblat, P. Loubaton, P., Serpedin, E., and Giannakis, G.B. (2002), 'Asymptotic analysis of blind cyclic correlation-based symbol-rate estimators', *IEEE Transactions on Information Theory*, 48(7), 1922–1934.
- [Dehay and Dudek(2015)] Dehay, D. and Dudek, A.E. (2015), 'Bootstrap method for Poisson sampling almost periodic process', *J. Time Ser. Anal.*, 36(3), 327–351.
- [Dehay and Dudek(2017)] Dehay, D. and Dudek, A.E. (2017), 'Bootstrap for the second-order analysis of Poisson-sampled almost periodic processes', *Electronic Journal of Statistics*, 11(1), 99–147.
- [Dehay and Hurd(1994)] Dehay, D. and Hurd, H.L. (1994), 'Representation and estimation for periodically and almost periodically correlated random processes', In: W. A. Gardner (ed.), *Cyclostationarity in Communications and Signal Processing*, chap. 6, IEEE Press, pp. 295–328.

- [Doukhan(1994)] Doukhan, P. (1994), *Mixing. Properties and examples*. Lecture Notes Statist. Springer.
- [Dudek(2015)] Dudek, A.E. (2015), 'Circular block bootstrap for coefficients of autocovariance function of almost periodically correlated time series', *Metrika*, 78(3), 313–335.
- [Dudek(2016)] Dudek, A.E. (2015), 'First and second order analysis for periodic random arrays using block bootstrap methods', *Electronic Journal of Statistics*, 10(2), 2561–2583.
- [Dudek et al.(2015)Dudek, Hurd, and Wójtowicz] Dudek, A.E., Hurd, H., and Wójtowicz, W. (2015), 'PARMA models with applications in R', F. Chaari et al. (eds.) *Cyclostationarity: Theory and Methods - II*, Applied Condition Monitoring 3. Chapter 7, pp 131–153, DOI 10.1007/978-3-319-16330-7_7, Springer International Publishing Switzerland 2015.
- [Dudek and Leśkow(2011)] Dudek, A.E. and Leśkow, J., (2011), 'A bootstrap algorithm for data from a periodic multiplicative intensity model', *Comm Statist Theory Methods*, 40, 1468–1489.
- [Dudek et al.(2014)Dudek, Leśkow, Paparoditis, and Politis] Dudek, A.E., Leśkow, J., Paparoditis, E., and Politis, D., (2014a), 'A generalized block bootstrap for seasonal time series', *J. Time Ser. Anal.*, 35, 89–114.
- [Dudek et al.(2014)Dudek, Maiz, and Elbadaoui] Dudek, A.E. , Maiz, S., and Elbadaoui, M. (2014b), 'Generalized Seasonal Block Bootstrap in frequency analysis of cyclostationary signals', *Signal Process.*, 104C, 358–368.
- [Dudek and Potorski(2014)] Dudek, A.E. and Potorski, P. (2014), 'Simulation comparison of CBB and GSBB in overall mean estimation problem for PC time series', F. Chaari et al. (eds.) *Cyclostationarity: Theory and Methods*, Lecture Notes in Mechanical Engineering. Chapter 7, pp 95–106, DOI 10.1007/978-3-319-04187-2, Springer International Publishing Switzerland.
- [Gardner et al.(2006)Gardner, Napolitano, and Paura] Gardner, W.A., Napolitano, A., and Paura, L. (2006), 'Cyclostationarity: Half a century of research', *Signal Process.*, 86(4), 639–697.
- [Guyon(1995)] Guyon, X. (1995), *Random Fields on a Network*, New York: Springer-Verlag.
- [Hurd and Leśkow(1992)] Hurd, H.L. and Leśkow, J. (1992a), 'Estimation of the Fourier coefficient functions and their spectral densities for ϕ -mixing almost periodically correlated processes', *Statistics & Probability Letters*, 14, 299–306.
- [Hurd and Leśkow(1992)] Hurd, H.L. and Leśkow, J. (1992b), 'Strongly consistent and asymptotically normal estimation of the covariance for almost periodically correlated processes', *Statist. & Decision*, 10, 201–225.
- [Hurd and Miamee(2007)] Hurd, H.L. and Miamee, A.G. (2007), *Periodically Correlated Random Sequences: Spectral. Theory and Practice*, John Wiley.
- [Kim(1994)] Kim, T. (1994), 'Moment bounds for non-stationary dependent sequences', *J. Appl. Probab.*, 31, 731–742.
- [Künsch(1989)] Künsch, H. (1989), 'The jackknife and the bootstrap for general stationary observations', *Ann. Statist.*, 17, 1217–1241.
- [Lahiri(2003)] Lahiri, S.N. (2003), *Resampling Methods for Dependent Data*, Springer, New York.

- [Lenart(2011)] Lenart, Ł. (2011), 'Asymptotic distributions and subsampling in spectral analysis for almost periodically correlated time series', *Bernoulli*, 17(1), 290–319.
- [Lenart(2016)] Lenart, Ł. (2016), 'Generalized resampling scheme with application to spectral density matrix in Almost Periodically Correlated class of time series', *J. Time Ser. Anal.*, 37(3), 369–404.
- [Lenart et al.(2008)Lenart, Leśkow, and Synowiecki] Lenart, Ł., Leśkow, J., and Synowiecki, R. (2008), 'Subsampling in testing autocovariance for periodically correlated time series', *J. Time Ser. Anal.*, 29(6), 995–1018.
- [Lenart and Synowiecki(2010)] Leśkow, J. and Synowiecki, R. (2010), 'On bootstrapping periodic random arrays with increasing period', *Metrika*, 71, 253–279.
- [Liu and Singh(1992)] Liu, R. and Singh, K. (1992), 'Moving block jackknife and bootstrap capture weak dependence', *Exploring the Limits of Bootstrap*, Wiley Ser. Probab. Math. Statist. Probab. Math. Statist. Wiley, New York, pp 225–248.
- [Napolitano(2012)] Napolitano, A. (2012), *Generalizations of Cyclostationary Signal Processing: Spectral Analysis and Applications*, Wiley-IEEE Press.
- [Napolitano(2016)] Napolitano, A. (2016), 'Cyclostationarity: New trends and applications', *Signal Process.*, 120, 385–408.
- [Politis(2001)] Politis, D.N. (2001), 'Resampling time series with seasonal components', In *Frontiers in Data Mining and Bioinformatics: Proceedings of the 33rd Symposium on the Interface of Computing Science and Statistics*, Orange County, California, June 13-17, 2001, pp. 619–621.
- [Politis and Romano(1992)] Politis, D.N. and Romano, J.P. (1992), *A circular block-resampling procedure for stationary data*, *Exploring the Limits of Bootstrap*, Wiley Ser. Probab. Math. Statist. Probab. Math. Statist. Wiley, New York, pp 263–270.
- [Politis et al.(1999)Politis, Romano, and Wolf] Politis, D., Romano, J., and Wolf, M. (1999), *Subsampling*. New York: Springer-Verlag.
- [Politis and White(2004)] Politis, D.N. and White, H. (2004), 'Automatic block-length selection for the dependent bootstrap', *Econometric Reviews*, 23(1), 53–70.
- [Synowiecki(2007)] Synowiecki, R. (2007), 'Consistency and application of moving block bootstrap for nonstationary time series with periodic and almost periodic structure', *Bernoulli*, 13(4), 1151–1178.

Acknowledgements

Anna Dudek has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 655394.

