A generalized block bootstrap for seasonal time series

Anna E. Dudek
AGH University of Science and Technology
Dept. of Applied Mathematics
al. Mickiewicza 30, 30-059
Krakow, Poland
Efthathios Paparoditis
Dept. of Mathematics and Statistics
University of Cyprus
P.O.Box 20537
CY 1678 Nicosia, Cyprus

Jacek Leškow
Department of Econometrics
Institute of Mathematics
Cracow University of Technology,
Krakow, Poland

Efstathios Paparoditis
Dept. of Mathematics and Statistics
University of Cyprus
P.O.Box 20537
CY 1678 Nicosia, Cyprus

Dimitris N. Politis
Department of Mathematics
University of California, San Diego
La Jolla, CA 92093-0112, USA
email: dpolitis@ucsd.edu

September 5, 2013

Abstract

When time series data contain a periodic/seasonal component, the usual block bootstrap procedures are not directly applicable. We propose a modification of the block bootstrap—the Generalized Seasonal Block Bootstrap (GSBB)—and show its asymptotic consistency without undue restrictions on the relative size of the period and block size. Notably, it is exactly such restrictions that limit the applicability of other proposals of block bootstrap methods for time series with periodicities. The finite-sample performance of the GSBB is also illustrated by means of a small simulation experiment.

Keywords: Periodic time series, resampling, seasonality.

1 Introduction

Consider observations of the form $X_1, \ldots, X_n$ arising as a stretch of a time series $\{X_t, t \in \mathbb{Z}\}$. An interesting class of time series models involves the case where a seasonality effect is present. For example, consider the set-up where

$$X_t = \mu_t + \sigma_t Y_t \quad \text{where} \quad \mu_t = \mu_{t+d} \quad \text{and} \quad \sigma_t = \sigma_{t+d} \quad \text{for all} \quad t \in \mathbb{Z};$$

in the above, $d$ is an integer denoting the period of the deterministic (but unknown) functions $\mu_t$ and $\sigma_t$, and $\{Y_t, t \in \mathbb{Z}\}$ is a (strictly) stationary sequence with mean zero.

*Research partially supported by the Polish Ministry of Science and Higher Education and AGH local grant
†Research partially supported by NSF grant DMS-10-07513.
Typically, the period $d$ is known (or obvious from the data), and in many cases corresponds to a daily, weekly or annual periodicity; see e.g. Brockwell and Davis (1991). Except in the rather special case where $\mu_t$ and $\sigma_t$ are constant, model (1.1) is not stationary; hence, the usual block bootstrap procedures are not directly applicable—see e.g. Lahiri (2003) or Politis (2003) and the references therein.

To accommodate the presence of a seasonal mean, Politis (2001) proposed a ‘Seasonal Block Bootstrap’ (SBB) that is a version of Künsch’s (1989) Block Bootstrap (BB) with blocks whose size and starting points are restricted to be integer multiples of the period $d$. It is apparent that the SBB as stated above poses a restriction on the relative size of the period and block size, i.e., the block size $b$ must be at least of the order of the period $d$. Furthermore, the block size in the SBB can not be fine-tuned since only integer multiples of $d$ are allowable block sizes.

By contrast, Chan et al. (2004) proposed a BB procedure for periodic time series that works by dividing the observations within each period in blocks of length $b$ and resampling these blocks in a way that new periods of pseudo-observations are generated. Joining together the so-generated periods leads to a new series of pseudo-observations. However, the resampling procedure of Chan et al. (2004) is designed for periodic time series that have long periodicities, since it is assumed that the block length $b$ is much smaller compared to the period $d$. Furthermore, Chan et al. (2004) assume an $m$-dependence for the underlying process where the dependence span $m$ is again negligible compared to the length of the period. In a recent paper, Leškow and Synowiecki (2010) extended the applicability of the Chan et al. (2004) procedure to triangular arrays of periodic random variables allowing for a more general dependence structure, e.g., strong mixing, and showed that consistency of the Chan et al. (2004) BB procedure requires that the period length $d$ has to tend to infinity as the sample size $n$ increases. Consequently, for time series with a fixed periodicity, the aforementioned block bootstrap procedure is not consistent. These results heavily restrict the range of applicability of the Chan et al. (2004) block bootstrap, since periodic time series with a fixed periodicity are the rule rather than the exception in many interesting areas of applications.

The aim of this paper is to propose a new modification of the block bootstrap that is suitable for periodic time series with fixed length periodicities of arbitrary size as related to block size and sample size. Instead of dividing each period of the time series in small blocks of length $b$ and resampling these blocks in a way that reconstructs a new period of pseudo-observations, we propose to divide the series of observations in blocks of desired length (independent from periodicity) and resample these blocks in a way that retains periodicity. Since the block size can be chosen independent from the length of the periodicity, the usual asymptotic considerations for block size choice are possible avoiding the inconsistency problems of Chan et al. (2004) procedure, as well as the lack of fine-tuning in block size choice associated with the SBB of Politis (2001).

The proposed Generalized Seasonal Block Bootstrap (GSBB) algorithm is introduced in the next Section where its asymptotic validity in some important examples is shown. The results of a small simulation study illustrating the applicability of the new block bootstrap proposal are presented in Section 5.
The Generalized Seasonal Block Bootstrap Algorithm

Model (1.1) describes a seasonality in the mean and variance of the time series and is an example of a PC time series. Nevertheless, the seasonality effect could be more general, e.g. affecting the entire joint distribution structure of the underlying process. To capture an arbitrary seasonality effect, we consider a general (strictly) periodic time series model with period $d \in \mathbb{N}$, where

$$ (X_t, \ldots, X_{t+k}) \overset{\mathcal{L}}{=} (X_{t+d}, \ldots, X_{t+k+d}) \text{ for all } t \in \mathbb{Z} \text{ and } k \in \mathbb{N}. $$

Here $\mathcal{L}$ denotes equality of the probability laws.

If the period length $d = 1$, then conditions (2.1) reduce to weak stationarity and (2.2) becomes strict stationarity.

For time series that are PC and/or satisfy (2.2) we propose the Generalized Seasonal Block Bootstrap (GSBB) algorithm that constructs a new series of pseudo-observations $X_1^*, X_2^*, \ldots, X_n^*$ in the following fashion.

### Generalized Seasonal Block Bootstrap (GSBB) algorithm

**Step 1:** Choose a (positive) integer block size $b(< n)$, and let $l = \lfloor n/b \rfloor$ where $\lfloor \cdot \rfloor$ denotes integer part.

**Step 2:** For $t = 1, b + 1, 2b + 1, \ldots, lb + 1$, let

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

where $k_t$ is a discrete uniform random variable taking values in the set

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

with $R_{1,n} = \lfloor (t - 1)/d \rfloor$ and $R_{2,n} = \lfloor (n - b - t)/d \rfloor$. As usual, the random variables $k_1, k_2, \ldots$ are taken to be independent.

**Step 3:** Join the $l + 1$ blocks $(X_{k_t}, X_{k_t+1}, \ldots, X_{k_t+b-1})$ thus obtained together to form a new series of bootstrap pseudo-observations $X_1^*, X_2^*, \ldots, X_n^*$ from which only the first $n$ points $X_1^*, X_2^*, \ldots, X_n^*$ are retained so that the bootstrap series has same length as the original. Notice that if $n$ is an integer multiple of $b$, then the whole last block is superfluous.
Another way of visualizing Step 2 in the construction of the bootstrap series \(X_1^*, X_2^*, \ldots, X_n^*\) is the following: for \(t = 1, b+1, 2b+1, \ldots, lb+1\), replace the block \((X_t, X_{t+1}, \ldots, X_{t+b-1})\) in the original series by the block \((X_{k_t}^*, X_{k_t+1}^*, \ldots, X_{k_t+b-1}^*)\) where \(k_t\) is as defined above. The above GSBB procedure defines a probability measure (conditional on the data \(X_1, \ldots, X_n\)) that will be denoted by \(P^*\); expectation and variance with respect to \(P^*\) are denoted by \(E^*\) and by \(\text{Var}^*\) respectively.

In contrast to the block bootstrap in the case of stationary non-periodic time series where the beginning \(t\) of the block selected to replace \((X_t, X_{t+1}, \ldots, X_{t+b-1})\) is chosen randomly from the set \(L_n = \{1, 2, \ldots, n-b+1\}\), in the case of periodic time series the beginning of this block is restricted to be randomly chosen from a set containing only periodic shifts of \(t\), i.e., shifts of \(t\) obtained by adding integer multiples of the periodicity \(d\) to \(t\). The set \(S_{t;n}\) in the above block bootstrap algorithm contains for each time point \(t\) all possible (backward and forward) periodic shifts of \(t\) determining, therefore, the starting points of all eligible blocks. Note that 

\[
|S_{t;n}| = R_{2,n} + R_{1,n} + 1 = O((n-b)/d)
\]

and that if \(d = 1\) then \(S_{t;n} \equiv L_n\). Furthermore, if the periodicity \(d\) of the observed series is an integer multiple of the block length \(b\), i.e., if \(d = kb\) for some \(k > 1, k \in \mathbb{N}\), then the above block bootstrap procedure is identical to that of Chen et al. (2004). However, this is a very specific case and \(d = kb\) is not required in the above block bootstrap algorithm where \(b\) can in fact be chosen independently of the periodicity \(d\). This is important since for instance, for time series with small periodicities (e.g., quarterly recorded economic time series) the block size \(b\) can be chosen larger than the periodicity and can not, therefore, be restricted to be smaller than the period \(d\).

### 3 Estimation under a simple model with seasonal trend

As usual, the validity of a bootstrap procedure must be established in a case-by-case basis. For this reason, we now re-visit the simple set-up of model (1.1) with \(\sigma_t\) assumed constant, i.e.,

\[
(3.1) \quad X_t = \mu_t + Y_t \quad \text{where} \quad \mu_t = \mu_{t+d} \quad \text{for all} \quad t \in \mathbb{Z}.
\]

In model (3.1), attention focuses on estimation of the seasonal trend component \(\mu_i, i = 1, \ldots, d\), and the overall mean \(\bar{\mu} = d^{-1} \sum_{i=1}^{d} \mu_i\). Point estimation is straightforward by means of averages of the ‘sampled’ series; in other words, define

\[
(3.2) \quad \hat{\mu}_i = w_i^{-1} \sum_{j=0}^{w_i-1} X_{i+jd} \quad \text{and} \quad \hat{\mu} = d^{-1} \sum_{i=1}^{d} \hat{\mu}_i,
\]

where \(w_i\) is the number of occurrence in the considered time series observations having marginal distribution \(\mathcal{L}(X_i)\).

The GSBB is useful in getting interval estimates for \(\mu_i\) and \(\mu\) by means of successfully approximating the distribution of \(\hat{\mu}_i\) and \(\hat{\mu}\); to do this, from each bootstrap pseudo-series \(X_1^*, X_2^*, \ldots, X_n^*\) we construct
bootstrap versions of $\hat{\mu}_i$ and $\hat{\mu}$ by:

$$
(3.3) \quad \hat{\mu}^*_i = w_i^{-1} \sum_{j=0}^{w_i-1} X_{i+jd}^* \quad \text{and} \quad \hat{\mu}^* = d^{-1} \sum_{i=1}^d \hat{\mu}^*_i.
$$

Denote the vectors $\mu = (\mu_1, \ldots, \mu_d)$, and $\hat{\mu} = (\hat{\mu}_1, \ldots, \hat{\mu}_d)$. Let $\mathcal{L}(\sqrt{n}(\hat{\mu} - \mu))$ denote the probability law of $\sqrt{n}(\hat{\mu} - \mu)$, and $\mathcal{L}^*(\sqrt{n}(\hat{\mu}^* - E^*\hat{\mu}^*))$ its bootstrap counterpart conditionally on the observed time series $X_1, X_2, \ldots, X_n$; similarly, define $\mathcal{L}(\sqrt{n}(\hat{\mu} - \mu))$, and $\mathcal{L}^*(\sqrt{n}(\hat{\mu}^* - E^*\hat{\mu}^*))$. Finally, let $d_2(\cdot, \cdot)$ denote Mallows’ metric, and $\alpha_Y(k)$ be the Rosenblatt strong mixing coefficients for stationary series $\{Y_t\}$. If $n$ is large, accurate point estimation is possible, and a central limit theorem and its bootstrap analog hold true under the assumptions of our Theorem 3.1.

**Theorem 3.1.** Assume that for some $\delta > 0$, $E|Y_t|^{4+\delta} < \infty$ and $\sum_{k=1}^{\infty} k\alpha_Y^{2/(4+\delta)}(k) < \infty$. If $b \to \infty$ as $n \to \infty$ such that $b = o(n)$, then

$$
(3.4) \quad d_2(\mathcal{L}(\sqrt{n}(\hat{\mu} - \mu)), \mathcal{L}^*(\sqrt{n}(\hat{\mu}^* - E^*\hat{\mu}^*))) \xrightarrow{p} 0,
$$

$$
(3.5) \quad d_2(\mathcal{L}(\sqrt{n}(\hat{\mu} - \mu)), \mathcal{L}^*(\sqrt{n}(\hat{\mu}^* - E^*\hat{\mu}^*))) \xrightarrow{p} 0,
$$

where $w = \lfloor n/d \rfloor$.

As is well known, convergence in the Mallows’ metric $d_2$ implies weak convergence as well as convergence of the first two moments; cf. Bickel and Freedman (1981). If convergence of bootstrap moments is not required, the moment and mixing assumptions of the above Theorem could be relaxed to $E|Y_t|^{2+\delta} < \infty$ and $\sum_{k=1}^{\infty} k\alpha_Y^{2/(2+\delta)}(k) < \infty$ as in Radulovic (1996) and Synowiecki (2007).

It is well-known that the degree of overlap among blocks to be bootstrapped plays a major role in efficiency: maximum overlap leads to maximum efficiency. Let the GSBB block length be $b = sd + r$, where $s \in \mathbb{Z}$ and $r = 0, \ldots, d - 1$. The GSBB overlap is equal to 0 when $b \leq d$. For $s > 0$ and $r > 0$ the proportion of overlap between successive blocks used by the GSBB is $(b - \min\{r, d - r\})/b$ which tends to one as $b \to \infty$. Finally, when $s > 0$ and $r = 0$ the overlap proportion is $(b - d)/b$ which again tends to one as $b \to \infty$. Therefore, the GSBB has (almost) full overlap in its successive blocks which hints to its efficient use of the available information—see e.g. Lahiri (2003).

**Remark 3.1.** Choosing the block length optimally (in some sense) is always an important and difficult problem in applications. Different methods for block size choice in the context of stationary data (or regression with stationary errors) are described in Nordman and Lahiri (2012) and the references therein. Some of these methods can conceivably be applicable in our context after the seasonal components have been estimated and removed, i.e., apply the different methods on (our estimate of) the stationary series $Y_t$ in model (1.1) or (3.1). In particular, such an application of the method of Politis and White (2004) is explored in our simulation section.
4 Asymptotic results for PC processes

A PC process is an example of a nonstationary process. The $\alpha$-mixing function for such processes is defined as follows (Doukhan (1994))

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

The seasonal means are defined as $\mu_i = E(X_i)$ for $i = 1, \ldots, d$ and the overall mean $\bar{\mu}$ is equal to $d^{-1}\sum_{i=1}^d \mu_i$. The estimators of the seasonal means and the overall mean are given by (3.2) and their bootstrap counterparts by (3.3).

Under summability condition for the autocovariance function (see Lenart et al. (2008), Lemma 1) the asymptotic variance of $1/\sqrt{n} \sum_{i=1}^n X_i$ is of the form

$$\sigma^2 = \frac{1}{d} \sum_{s=1}^d \sum_{k=-\infty}^{\infty} \text{Cov}(X_s, X_{s+k}).$$

Under regularity conditions the central limit theorem can be shown

$$\sup_{x \in \mathbb{R}^d} |P(\sqrt{w} (\hat{\mu} - \mu) \leq x) - \Phi_d(0, \Sigma)| \to 0,$$

$$\sup_{t \in \mathbb{R}} |P(\sqrt{n} (\hat{\mu} - \mu) \leq t) - \Phi(0, \sigma^2)| \to 0,$$

where $\Phi$ is a cumulative normal distribution function, $\Sigma = (\sigma_{ij})$ and $i, j = 1, \ldots, d$. Moreover, $\sigma_i = \sigma_i^2 = \text{Var}(X_i)$ and $\sigma_{ij} = \sum_{k=-\infty}^{\infty} \text{Cov}(X_i, X_{j+k})$. For example, sufficient conditions for (4.1) and (4.2) are $\sup_t E|X_t|^{2+\delta} < \infty$ and $\sum_{k=1}^{\infty} k \alpha_X^{\delta/(2+\delta)}(k) < \infty$ (see Theorem 2.1 in Guyon (1995)).

The Theorem below is dedicated to a general PC process that includes model (1.1) as a special case.

**Theorem 4.1.** Let $\{X_t, t \in \mathbb{Z}\}$ be a PC time series. Assume that for some $\delta > 0$, $\sup_t E|X_t|^{4+\delta} < \infty$ and $\sum_{k=1}^{\infty} k \alpha_X^{\delta/(4+\delta)}(k) < \infty$. Then, GSBB for the overall and seasonal means is consistent in the sense of (3.4) and (3.5).

Note that the real-world limit distributions of eq. (4.1) and (4.2) are Gaussian, and therefore continuous; hence, by Polya’s theorem, the weak convergences proved in Theorems 3.1 and 4.1 are uniform, implying that asymptotically valid confidence intervals for each $\mu_i$ and for $\bar{\mu}$ can be constructed using the quantiles of the GSBB distribution in place of the true ones. However, the consistency of the joint GSBB of distribution of $\hat{\mu}_i, i = 1, \ldots, d$ implies that simultaneous confidence intervals for $\{\mu_i, i = 1, \ldots, d\}$ can also be constructed, thus leading to an asymptotically valid confidence band for the trend function $\mu_t, t \in \mathbb{Z}$.

To construct the aforementioned confidence band, the continuous mapping theorem is invoked together with Theorem 4.1 to yield the following corollary.

**Corollary 4.1.** Under the conditions of Theorem 4.1 we have

$$\sup_t |P(\sqrt{w} \max_i |\hat{\mu}_i - \mu_i| \leq t) - P^*(\sqrt{w} \max_i |\hat{\mu}_i^* - E^* \hat{\mu}_i^*| \leq t)| \xrightarrow{P} 0.$$
4.1 Smooth function of means

Below we present results concerning consistency of GSBB for smooth functions of the overall mean and seasonal means. For the almost periodically correlated (APC) processes, Synowiecki (2008) showed consistency of MBB for the overall mean with block size $b$ of order $o(n/\log n)$. APC processes have almost periodic mean and covariance functions and are generalization of periodically correlated (PC) processes. Definition and properties of almost periodic functions can be found in Besicovitch (1932).

**Theorem 4.2.** Let $\{X_t, t \in \mathbb{Z}\}$ be a PC time series that fulfills assumptions of Theorem 4.1. Suppose that function $H: \mathbb{R} \to \mathbb{R}$ is:

(i) differentiable in a neighborhood

$$N_H = \{x \in \mathbb{R} : |x - \mu| < 2\eta\} \quad \text{for some } \eta > 0$$

(ii) $H'(\bar{\mu}) \neq 0$

(iii) the first-order derivative $H'$ satisfies a Lipschitz condition of order $\kappa > 0$ on $N_H$.

If $b \to \infty$ as $n \to \infty$ such that $b = o(n)$, but $b = o(n/\log n)$ and $b^{-1} = o(\log^{-1} n)$, then GSBB is consistent i.e.

$$d_2 \left( \mathcal{L} \left( \sqrt{n} \left( H \left( \bar{\mu} \right) - H \left( \mu \right) \right) \right), \mathcal{L}^* \left( \sqrt{w} \left( H \left( \bar{\mu}^* \right) - H \left( \mu^* \right) \right) \right) \right) \overset{p}{\to} 0.$$

Note that the condition $b^{-1} = o(\log^{-1} n)$ is equivalent to $l = o(n/\log n)$, where $l$ is the number of blocks of length $b$.

**Theorem 4.3.** Let $\{X_t, t \in \mathbb{Z}\}$ be a PC time series that fulfills assumptions of Theorem 4.1. Suppose that function $H: \mathbb{R}^d \to \mathbb{R}^s$ is:

(i) differentiable in a neighborhood

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

(ii) $\nabla H(\mu) \neq 0$

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

If $b \to \infty$ as $n \to \infty$ such that $b = o(n)$, but $b = o(n/\log n)$ and $b^{-1} = o(\log^{-1} n)$, then GSBB is consistent i.e.

$$d_2 \left( \mathcal{L} \left( \sqrt{w} \left( H \left( \bar{\mu} \right) - H \left( \mu \right) \right) \right), \mathcal{L}^* \left( \sqrt{w} \left( H \left( \bar{\mu}^* \right) - H \left( \mu^* \right) \right) \right) \right) \overset{p}{\to} 0,$$

where $\hat{\mu} = (\hat{\mu}_1, \ldots, \hat{\mu}_d)$ and $w = \lfloor n/d \rfloor$. 


4.2 Least squares estimators

Suppose that eq. (1.1) holds and that
\[ \mu_t = \sum_{j=1}^{p} \beta_j m_j(t) \quad \text{for all} \quad t \in \mathbb{Z}, \]
where \( m_j(\cdot), j = 1, \ldots, p, \) are known functions of period \( d, \) and \( \beta_j, j = 1, \ldots, p, \) are unknown coefficients. To ensure identifiability we assume that \( p \leq d \) here, and that \( \text{rank}(M) = p \) where \( M \) is the \( d \times p \) matrix whose \( j \)-th row is given by \( (m_j(1), \ldots, m_j(p)) \). For instance, \( m_1(\cdot) \) may be the constant function while \( m_j(\cdot) \) for \( j > 1 \) some other functions of interest, e.g., trigonometric functions.

Let \( \hat{\beta}_j \) denote the least squares (LS) estimator of \( \beta_j \), i.e., the minimizer of \( \sum_{t=1}^{N} (X_t - \sum_{j=1}^{p} c_j m_j(t))^2 \) with respect to \( c_1, c_2, \ldots, c_p \). Let \( \beta = (\beta_1, \ldots, \beta_p) \) and \( \hat{\beta} = (\hat{\beta}_1, \ldots, \hat{\beta}_p) \) and denote by \( \hat{\beta}^* \) be the LS estimator of \( \beta \) computed from the GSBB bootstrap series \( X_1^*, X_2^*, \ldots, X_l^* \). It is easy to see that \( \hat{\beta} \) is a continuous function of \( \hat{\mu} \). Thus, by the continuous mapping theorem, the following result is an immediate corollary of Theorem 4.1.

**Corollary 4.2.** Under the conditions of Theorem 4.1 we have, as \( n \to \infty \), that
\[
(4.6) \quad d_2 \left( \mathcal{L}(\sqrt{n}(\hat{\beta} - \beta)), \mathcal{L}^*(\sqrt{n}(\hat{\beta}^* - E^*\hat{\beta}^*)) \right) \overset{P}{\to} 0.
\]

**Remark 4.1.** If \( \sigma_t \) could be assumed to be constant, i.e., if (3.1) were true, then the results of Nordman and Lahiri (2012) on different block bootstrap methods would be directly applicable. In the general case, however, the GSBB method gives an immediate and elegant solution as the above Corollary shows. Nordman and Lahiri (2012) also study the Tapered Block Bootstrap of Paparoditis and Politis (2001) and show that it requires a particular modification in order to maintain its validity and superior performance in a regression setting with stationary errors. The implementation of tapering in connection with the GSBB is possible in principle but the details are not trivial; future work may address a Tapered GSBB.

4.3 Triangular arrays with growing period

Finally, we present some additional results for the case when the period \( d \) is no longer constant, but is growing together with the sample size. Corresponding theorems for PBB can be found in Leškow and Synowiecki (2010).

Let \( \{X_{n,t} : t = 1, \ldots, m_n\} \) be an array of real valued random variables. We assume that the mean function \( \mu_{n,t} = E(X_{n,t}) \) in each row is periodic in variable \( t \) with period \( d_n \). The mean of elements in \( n \)-th row we denote by \( \overline{X}_n = 1/m_n \sum_{t=1}^{m_n} X_{n,t} \) and its bootstrap counterpart obtained with GSBB method by \( \overline{X}^* \). Moreover, \( \overline{\mu}_n = 1/d_n \sum_{t=1}^{d_n} \mu_{n,t}. \)

**Theorem 4.4.** Let \( \{X_{n,t} : t = 1, \ldots, m_n\} \) be an array of real valued random variables, where the mean function \( \mu_{n,t} = E(X_{n,t}) \) is periodic in variable \( t \) with period \( d_n \). Assume that

(i) \( m_n = w_n d_n, d_n \to \infty \) and \( w_n \to \infty \) as \( n \to \infty; \)
there exists $\delta > 0$ such that

$$\sup_{t,n} E |X_{t,n}|^{4+\delta} < \infty \quad \text{and} \quad \sup_{n} \sum_{k=1}^{m_n} \tau \alpha_n^{\delta/(4+\delta)}(\tau) = K < \infty;$$

(iii) the variance converges uniformly i.e. for any sequence $\{k_n\}$ such that $k_n < m_n$ and $k_n \to \infty$, we have

$$\sup_{s=1,\ldots,m_n-k_n+1} \text{Var} \left( \frac{1}{\sqrt{k_n}} \sum_{t=s}^{s+k_n-1} X_{n,t} \right) \to \sigma^2,$$

where $\sigma^2 > 0$.

If $b_n \to \infty$ as $n \to \infty$ such that $b_n = o(m_n)$, then GSBB is consistent i.e.

$$d_2 \left( \mathcal{L} \left( \sqrt{m_n} (\overline{X}_n - \overline{\mu}_n) \right), \mathcal{L}^* \left( \sqrt{m_n} (\overline{X}_n^* - \overline{E}^* \overline{X}_n^*) \right) \right) \to 0.$$

**Corollary 4.3.** In Theorem 4.4 the assumption (iii) can be replaced by the assumption

(iii’) the triangular array $\{X_{n,t} : t = 1, \ldots, m_n\}$ is row-wise periodically correlated with the same period $d_n$

$$\text{Var} \left( \frac{1}{\sqrt{m_n}} \sum_{t=1}^{m_n} X_{n,t} \right) \to \sigma^2,$$

where $\sigma^2 > 0$ and $\sup_{n,\tau} \left\{ \tau^{1+\zeta} \alpha_n^{1/2}(\tau) \right\} < \infty$ for some $\zeta > 0$.

5 Numerical results

A small simulation study was performed to calculate the actual coverage probability (ACP) and the mean length of the confidence intervals obtained by GSBB method. Three different time series with periodic structure were considered:

(M1) $X_t = U_t ((t \mod d) + 1) + \sin(2\pi t/d)$,

where $U_t$ are independent random variables uniformly distributed on the interval $[-0.5; 0.5]$.

(M2) $X_t = 0.5X_{t-1} + 0.25X_{t-2} + (t \mod d) + \varepsilon_t$,

where $\varepsilon_t$ are independent standard normal distribution random variables.

(M3) $X_t = \varepsilon_t + 0.4\varepsilon_{t-1} + 0.6\varepsilon_{t-2} + (t \mod d) + \cos(2\pi t/d)\varepsilon_t$,

where $\varepsilon_t$ are independent standard normal distribution random variables and $\varepsilon_t$ are independent random variables uniformly distributed on $[-1, 1]$.

For models (M1), (M2) and (M3) time series of length $n = 250$ and $n = 1000$ are generated while the number of bootstrap samples has been set equal to $B = 500$. For model (M1) three period lengths
$d \in \{10, 50, 100\}$ while for models (M2) and (M3) four period lengths $d \in \{5, 10, 50, 100\}$ have been considered. For $d = 5$ the block lengths $b$ are chosen from the set $\{5, 10, 25, 50, 125\}$ while for $d = 10$ from the set $\{2, 5, 25, 50, 125\}$. Finally for $d = 50$ and $d = 100$ the block lengths considered are those from the set $\{5, 10, 25, 50, 125\}$. In all situation a 95% pointwise equal-tailed percentile bootstrap confidence intervals for $\bar{p}$ as well as a 95% simultaneous equal-tailed percentile bootstrap confidence intervals for $\mu_i$ ($i = 1, \ldots, d$) were calculated. The ACP and the mean length of all confidence intervals obtained were calculated using 500 iterations.

In our study we focused only on GSBB method, but for PC time series the consistency for seasonal means has been proved also for vectorized MBB (see results in Synowiecki (2008)). However, for reasonable sample sizes, the effective sample size of the vectorized time series to be bootstrapped can be too low for asymptotics to manifest themselves. For example, in our simulation when $n = 250$ and period lengths $d = 10, 50, 100$, the sample size for the the vector valued series would be 25, 5, and 2 respectively which is too low to make any sort of block bootstrap reasonable. In the case where $n = 1000$, these values are 100, 20, and 10; only in the case of sample size 100 could we conceive of doing a block bootstrap. So all in all, a comparison with the vectorized MBB would not be possible in our simulation except for one instance ($n = 1000$ and $d = 10$) where the sample size is so much larger than the period.

The results are presented in Figures 1-6. For model (M1) with $n = 1000$, and $d = 10$ among all considered block sizes $b$, it seems that $b = 5$ leads to the best results both, for the global mean and for the seasonal means. One may notice that the ACP are too low in these cases. On the other hand, for $d = 50$ and $d = 100$ the optimal $b$ is 10 for the global mean and 125 for the seasonal means. Only in the latter case ACP is a bit too high.

In the case $n = 250$ we also get ACPs too low. The optimal values of $b$ are 2, 10, 5 (for the global mean) and 5, 50, 25 (for the seasonal means) for $d = 10, 50, 100$, respectively.

In the case of model (M2) and $n = 1000$ the ACPs are too low for the global mean and the seasonal means with $d = 5$ and $d = 10$. On the other hand, for $d = 50$ and $d = 100$ the ACPs are too high in the seasonal means case. The optimal values of $b$ are 50, 125, 50, 50 (for the global mean) and 50, 50, 10, 125 (for the seasonal means) for $d = 10, 50, 100$, respectively.

For model (M2) and $n = 250$ again only for $d = 50$ and $d = 100$ the ACPs are too high in the seasonal means case. In all other cases the confidence intervals are too narrow. The optimal value of $b$ is 25 (for the global mean independently on the considered value of $d$) and 50, 5, 25, 50 (for the seasonal means) for $d = 5, 10, 50, 100$, respectively.

In the case of model (M3) ACPs are a bit lower than the nominal coverage probability, but they are definitely higher than for model (M2). Except $d = 5$ the optimal block length choices are the same independently on the sample length. For $d = 5$ the optimal $b$ is 2 ($n = 250$) and 5 ($n = 1000$). For $d = 10, 50, 1000$ these values are 5, 5, 25, respectively. For seasonal means ACPs with optimal $b$ are very close to 95%. The only quite natural exception is $n = 250$ with $d = 100$. The optimal values of $b$ are 5, 25, 25, 50 ($n = 250$) and 10, 5, 25, 50 ($n = 1000$).

The general undercoverage of the GSBB confidence intervals is quite noticeable; the worst case is with model (M2) and $n = 250$ where the ACP for a global mean with optimal $b$ is only about 0.85. It is
interesting to investigate whether this undercoverage is due to the GSBB construction per se or to a deficiency of the generic block bootstrap procedure, i.e., the MBB. To do this, we ran an additional simulation using MBB on the AR and MA parts of (M2) and (M3) models i.e., having removed the seasonal parts of these models. Using the notation introduced, while defining models (M1)-(M3) we have

\[
\text{AR}(2): X_t = 0.5X_{t-1} + 0.25X_{t-2} + \varepsilon_t,
\]

\[
\text{MA}(2): X_t = \varepsilon_t + 0.4\varepsilon_{t-1} + 0.6\varepsilon_{t-2} + \varepsilon_t.
\]

We focused only on \(n = 250\) and \(b \in \{5, 10, 25, 50\}\), and give the results in Table 1. The empirical ACPs of the MBB for AR are maximally about 0.86, while for MA about 0.93. These numbers are in accordance with the ACPs achieved by the GSBB in the seasonal case. Therefore we can conclude that the observed undercoverages are due to the MBB procedure, and not to the GSBB construction; not surprisingly, the MBB/GSBB finds most difficulty in approximating the AR(2) model. The tapered block bootstrap of Paparoditis and Politis (2001) might do better in that respect but the theory behind it is—for the moment—intractable in the seasonal case.

<table>
<thead>
<tr>
<th>model</th>
<th>b</th>
<th>5</th>
<th>10</th>
<th>25</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR(2)</td>
<td></td>
<td>0.738</td>
<td>0.824</td>
<td>0.840</td>
<td>0.855</td>
</tr>
<tr>
<td>MA(2)</td>
<td></td>
<td>0.935</td>
<td>0.928</td>
<td>0.914</td>
<td>0.882</td>
</tr>
</tbody>
</table>

Table 1: ACPs of pointwise equal-tailed percentile bootstrap confidence intervals for \(\mu\) for block length \(b \in \{5, 10, 25, 50\}\). Nominal coverage probability is 95%.

Moreover, the optimal \(b\) is sometimes bigger and sometimes smaller than a period length indicating that the method of Chan et al. (2004) would not be applicable in some cases as it requires that \(b\) is of smaller order than \(d\). In addition, the optimal \(b\) is not necessarily an integer multiple of \(d\) suggesting that the SBB of Politis (2001)—although applicable—would have to be applied using a suboptimal choice of \(b\). Additionally, in all considered cases (especially when \(n = 1000\)) the ACP curves for the simultaneous confidence bands are very flat and become flatter near the point of optimal \(b\) as \(d\) increases. This means that the simultaneous confidence problem is less sensitive to the choice of \(b\).

Finally, we revisit the issue of optimal, data-based block size choice as discussed in Remark 3.1. For concreteness, we focused on the method of Politis and White (2004) taking into account the correction discussed in Patton et al. (2009). For each model, we removed the estimated periodic mean and divided by estimated standard deviation (if it is not constant) in order to get an approximation of the stationary series \(Y_t\); the Politis and White (2004) method was then applied to our estimated \(Y_t\) series. The results of 500 Monte Carlo trials are presented in Table 2.

In the case of model (M1) the mean and median of optimal \(b\) choice is almost always 1 independently
on the considered sample size. It is not surprising as (M1) is based on an approximately i.i.d. sample. For such \( b \) the actual coverage probabilities are quite close to nominal ones. For \( n = 250, d = 50 \) and \( n = 1000, d = 100 \) the median is equal to 2. The main exception is for \( d = 100 \) and \( n = 250 \), when this value is 27. In this case within the sample we have only 2.5 periods; hence, the estimators of the periodic means and the periodic standard deviations are expected to be poor—being averages of 2 or 3 points only—and this adversely affects the block length choice regardless of model. The same phenomenon is true for models (M2) and (M3) when \( d = 100 \) and \( n = 250 \). With the exception of this problematic case, the optimal block length choices obtained using the Politis and White (2004) method seem to work well in terms of the previously presented ACP curves; this is especially true for the longer samples. Only for (M3) and \( d = 100 \) with \( n = 1000 \) the choice \( b = 11 \) seems to be a local minimum on the APC curve giving low coverage probability. Even with \( n = 250 \) values of optimal \( b \) that give ACP close to maximal reachable are in a range of one standard deviation from the mean value. Although the Politis and White (2004) method focuses on finding the optimal \( b \) for variance estimation, it seems to work reasonably well also in terms of the quantile estimation involved in our confidence interval construction.

<table>
<thead>
<tr>
<th>model</th>
<th>( d )</th>
<th>( n = 250 )</th>
<th>( n = 1000 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean</td>
<td>median</td>
<td>sd</td>
</tr>
<tr>
<td>(M1)</td>
<td>10</td>
<td>1.64</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>1.89</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>26.80</td>
<td>27</td>
</tr>
<tr>
<td>(M2)</td>
<td>5</td>
<td>15.10</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>15.20</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>16.40</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>27.00</td>
<td>27</td>
</tr>
<tr>
<td>(M3)</td>
<td>5</td>
<td>6.76</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>6.80</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>7.12</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>27.00</td>
<td>27</td>
</tr>
</tbody>
</table>

Table 2: Summary statistics for optimal block length choices obtained using Politis and White (2004) method for series described by models (M1)-(M3) after subtraction the periodic means and division by the periodic standard deviations. The number of Monte Carlo trials was 500.

6 Appendix

**Proof of Theorem 3.1:** Without loss of generality we assume that \( n = dw \) and \( n = lb \), \( w, l \in \mathbb{N} \). Moreover, we assume for the global mean that \( \mu = 0 \). To simplify the notation we show the assertion of the theorem for a circular version of block bootstrap method. The idea of wrapping block in the circle before blocking them was introduced by Politis and Romano (1992).

First we prove (3.5). The proof is based on arguments similar to those used by Leškow and Synowiecki (2007) and (2010).
Figure 1: Model (M1): ACPs of pointwise equal-tailed percentile bootstrap confidence intervals for $\bar{\pi}$ (left-hand side) and ACPs of simultaneous equal-tailed percentile bootstrap confidence intervals for $\mu_i$ ($i = 1, \ldots, d$) (right-hand side) vs. block length $b$. From top results for $d = 10, 50, 100$, respectively. Sample size $n = 250$ (black) and $n = 1000$ (grey). Nominal coverage probability is 95%.
Figure 2: Model (M2): ACPs of pointwise equal-tailed percentile bootstrap confidence intervals for $\pi$ (left-hand side) and ACPs of simultaneous equal-tailed percentile bootstrap confidence intervals for and $\mu_i$ ($i = 1, \ldots, d$) (right-hand side) vs. block length $b$. From top results for $d = 5, 10, 50, 100$, respectively. Sample size $n = 250$ (black) and $n = 1000$ (grey). Nominal coverage probability is 95%.
Figure 3: Model (M3): ACPs of pointwise equal-tailed percentile bootstrap confidence intervals for $\pi$ (left-hand side) and ACPs of simultaneous equal-tailed percentile bootstrap confidence intervals for and $\mu_i$ ($i = 1, \ldots, d$) (right-hand side) vs. block length $b$. From top results for $d = 5, 10, 50, 100$, respectively. Sample size $n = 250$ (black) and $n = 1000$ (grey). Nominal coverage probability is 95%.
Figure 4: Model (M1): mean lengths of pointwise equal-tailed percentile bootstrap confidence intervals for $\bar{\mu}$ (left-hand side) and mean lengths of simultaneous equal-tailed percentile bootstrap confidence intervals for $\mu_i$ ($i = 1, \ldots, d$) (right-hand side) vs. block length $b$. From top results for $d = 10, 50, 100$, respectively. Sample size $n = 250$ (black) and $n = 1000$ (grey).
Figure 5: Model (M2): mean lengths of pointwise equal-tailed percentile bootstrap confidence intervals for $\pi$ (left-hand side) and mean lengths of simultaneous equal-tailed percentile bootstrap confidence intervals for $\mu_i$ ($i = 1, \ldots, d$) (right-hand side) vs. block length $b$. From top results for $d = 5, 10, 50, 100$, respectively. Sample size $n = 250$ (black) and $n = 1000$ (grey).
Figure 6: Model (M3): mean lengths of pointwise equal-tailed percentile bootstrap confidence intervals for $\mu$ (left-hand side) and mean lengths of simultaneous equal-tailed percentile bootstrap confidence intervals for $\mu_i$ ($i = 1, \ldots, d$) (right-hand side) vs. block length $b$. From top results for $d = 5, 10, 50, 100$, respectively. Sample size $n = 250$ (black) and $n = 1000$ (grey).
Denote by $Z_{t,b}$ the sum

$$Z_{t,b} = X_t + \ldots + X_{t+b-1}.$$ 

Notice that

$$\hat{\mu} = \frac{1}{n} \sum_{s=0}^{l-1} Z_{1+sb,b}.$$ 

For $j = 1, b+1, \ldots, (l-1)b+1$ the random variables $Z^*_{j,b}$ are conditionally independent with common distribution

$$P^* (Z^*_{j,b} = Z_{j+kd,b}) = \frac{1}{w} \quad \text{for } k = 0, \ldots, w-1.$$ 

The bootstrap version of the estimator $\hat{\mu}$ is of the form

$$\hat{\mu}^* = \frac{1}{n} \sum_{s=0}^{l-1} Z^*_{1+sb,b}.$$ 

First note that because GSBB replicates ideally the periodic structure in the bootstrap sample, we can assume that we sample from the blocks with mean value equal to zero.

By Corollary 2.4.8 in Araujo and Giné (1980) to get (3.5), we need to show that for any $\delta > 0$

\[ \sum_{k=0}^{l-1} P^* \left( \frac{1}{\sqrt{n}} |Z^*_{1+kb,b}| > \delta \right) \xrightarrow{p} 0, \]

\[ \sum_{k=0}^{l-1} E^* \left( \frac{1}{\sqrt{n}} Z^*_{1+kb,b} \mathbf{1} |Z^*_{1+kb,b}| > \sqrt{\delta} \right) \xrightarrow{p} 0, \]

\[ \sum_{k=0}^{l-1} \text{Var}^* \left( \frac{1}{\sqrt{n}} Z^*_{1+kb,b} \mathbf{1} |Z^*_{1+kb,b}| \leq \sqrt{\delta} \right) \xrightarrow{p} \sigma^2. \]

To prove (6.1) notice that

\[ \sum_{k=0}^{l-1} P^* \left( \frac{1}{\sqrt{n}} |Z^*_{1+kb,b}| > \delta \right) = \sum_{k=0}^{l-1} \frac{1}{w} \sum_{s=0}^{w-1} P \left( |Z_{1+kb+s,b}| > \sqrt{\delta} \right). \]

Moreover,

\[ E \left| \sum_{k=0}^{l-1} \frac{1}{w} \sum_{s=0}^{w-1} \mathbf{1} |Z_{1+kb+s,b}| > \sqrt{\delta} \right| \leq \sum_{k=0}^{l-1} \frac{1}{w} \sum_{s=0}^{w-1} P \left( |Z_{1+kb+s,b}| > \sqrt{\delta} \right) \leq \sum_{k=0}^{l-1} \frac{1}{w} \sum_{s=0}^{w-1} E \left| Z_{1+kb+s,b} \right|^3 \frac{1}{n^{3/2} \delta^3}. \]

Since $E \left| 1/\sqrt{n} Z_{1+kb+s,b} \right|^4$ is bounded (see Kim (1994)), we get

\[ \sum_{k=0}^{l-1} \frac{1}{w} \sum_{s=0}^{w-1} E \left| Z_{1+kb+s,b} \right|^3 \frac{1}{n^{3/2} \delta^3} = O \left( \frac{1}{\sqrt{l}} \right). \]
which establishes (6.1).

To see (6.2) notice that

\[
\sum_{k=0}^{l-1} \frac{1}{w} \sum_{t=0}^{w-1} \frac{1}{n} Z_{1+kk+td,b}^* 1_{Z_{1+kk+td,b}^* > \sqrt{n} \delta}
= \sum_{k=0}^{l-1} \frac{1}{w} \sum_{t=0}^{w-1} \frac{1}{n} Z_{1+kk+td,b} 1_{Z_{1+kk+td,b} > \sqrt{n} \delta},
\]

By Hölder’s inequality, the expected absolute value of the last expression is less or equal to

\[
\sum_{k=0}^{l-1} \frac{1}{w} \sum_{t=0}^{w-1} \frac{1}{\sqrt{l}} \left| E \left( \frac{1}{\sqrt{b}} Z_{1+kk+td,b} \right) \right|^2 P \left( \left| Z_{1+kk+td,b} \right| > \sqrt{n} \delta \right).
\]

Additionally,

\[
P \left( \left| Z_{1+kk+td,b} \right| > \sqrt{n} \delta \right) \leq \frac{E \left( \frac{1}{\sqrt{b}} Z_{1+kk+td,b} \right)^4}{l^2 \delta^4},
\]

which means that (6.4) is \(O(1/\sqrt{l})\).

To prove (6.3), notice that

\[
\begin{align*}
\sum_{k=0}^{l-1} \text{Var}^* \left( \frac{1}{\sqrt{n}} Z_{1+kk+td,b}^* \right) & = \sum_{k=0}^{l-1} \frac{1}{w} \sum_{t=0}^{w-1} \frac{1}{n} Z_{1+kk+td,b}^* \left| Z_{1+kk+td,b}^* \leq \sqrt{n} \delta \right. \\
& - \sum_{k=0}^{l-1} \left( \frac{1}{w} \sum_{t=0}^{w-1} \frac{1}{\sqrt{n}} Z_{1+kk+td,b} 1_{Z_{1+kk+td,b} \leq \sqrt{n} \delta} \right)^2 \\
& = I + II,
\end{align*}
\]

with an obvious notation for \(I\) and \(II\). Notice that \(I\) can be rewritten as follows

\[
I = \sum_{k=0}^{l-1} \frac{1}{w} \sum_{t=0}^{w-1} \frac{1}{n} Z_{1+kk+td,b}^2 1_{Z_{1+kk+td,b} \leq \sqrt{n} \delta} - \sum_{k=0}^{l-1} \frac{1}{w} \sum_{t=0}^{w-1} \frac{1}{n} Z_{1+kk+td,b} 1_{Z_{1+kk+td,b} > \sqrt{n} \delta}.
\]

First we show that the second expression tends to 0 in probability. Its expected absolute value is less or equal to

\[
\sum_{k=0}^{l-1} \frac{1}{w} \sum_{t=0}^{w-1} \frac{1}{n} E \left| Z_{1+kk+td,b}^2 1_{Z_{1+kk+td,b} > \sqrt{n} \delta} \right| \leq \sum_{k=0}^{l-1} \frac{1}{w} \sum_{t=0}^{w-1} \frac{1}{n} E \frac{1}{\sqrt{b}} Z_{1+kk+td,b}^2 1_{Z_{1+kk+td,b} > \sqrt{n} \delta}^3 = O \left( \frac{1}{\sqrt{l}} \right).
\]

To prove that \(I\) tends to \(\sigma^2\) in probability, we use Lemma 5 from Leśkow and Synowiecki (2010). Note that the set of subindexes \(S = \{1+kb+td, t = 0, \ldots, w-1, k = 1, \ldots, l\}\) has at most \(n\) different elements and not less then \(w\). Denote the number of elements of \(S\) by \(n_0\). Moreover, we split \(S\) in two subsets
\[ S = S^1 \cup S^2, \] where \( S^1 = \{ s \in S : s \leq n - b + 1 \} \) and \( S^2 = \{ s \in S : s > n - b + 1 \} \). Also denote by \( n_{0,1} \) and \( n_{0,2} \) the number of elements of subsets \( S^1 \) and \( S^2 \), respectively.

We define the array \( \{ Q_{n,s}^1, s \in S^1 \} \), where

\[ Q_{n,s}^1 = \frac{1}{b} Z_{s,b}^2, \quad s \in S^1. \]

As a result of Lemma A.6 from Synowiecki (2007), we get

\[ \frac{1}{n_{0,1}} \sum_{t \in S^1} \mathbb{E} \left( Q_{n,s}^1 \right) = \frac{1}{n_{0,1}} \sum_{s \in S^1} \operatorname{Var} \left( \frac{1}{\sqrt{b}} Z_{s,b} \right) \rightarrow \sigma^2, \]

where \( \sigma^2 \) is the asymptotic variance of \( L \left( \sqrt{w} (\mu - \bar{\mu}) \right) \).

Moreover, \( \sup_s \mathbb{E} \left| Q_{n,s}^1 \right|^2 \) is bounded by the constant that is independent of \( n \).

Additionally, the considered array is \( \alpha \)-mixing with \( \alpha_{Q^1}(\tau) \leq \alpha_Y (\tau - b + 1) \), which means that \( 1/n_{0,1} \sum_{s \in S^1} Q_{n,s}^1 \) tends to \( \sigma^2 \) in probability.

Additionally,

\[ \frac{1}{n_0} \sum_{s \in S} \frac{1}{\sqrt{b}} Z_{s,b} = \frac{n_{0,1}}{n_0} \frac{1}{n_{0,1}} \sum_{s \in S^1} \frac{1}{\sqrt{b}} Z_{s,b} + \frac{n_{0,2}}{n_0} \frac{1}{n_{0,2}} \sum_{s \in S^2} \frac{1}{\sqrt{b}} Z_{s,b}. \]

Since \( n_{0,1}/n_0 \) tends to 1 as \( n \to \infty \), \( n_{0,2}/n_0 \) tends to 0 as \( n \to \infty \), applying Markov’s inequality we get that \( I \) tends to \( \sigma^2 \) in probability.

Now we only need to show that \( II \) tends to 0 in probability. We have

\[ P \left( |II| > \varepsilon \right) \leq \frac{1}{\ell \varepsilon} \sum_{k=0}^{l-1} \mathbb{E} \left( \frac{1}{w} \sum_{t=0}^{w-1} \frac{1}{\sqrt{b}} Z_{1+k\ell+b+t\ell} 1_{|Z_{1+k\ell+b+t\ell}| \leq \sqrt{n} \delta} \right)^2 \leq \frac{2}{\ell \varepsilon} \sum_{k=0}^{l-1} \mathbb{E} \left( \frac{1}{w} \sum_{t=0}^{w-1} \frac{1}{\sqrt{b}} Z_{1+k\ell+b+t\ell} \right)^2 + \frac{2}{\ell \varepsilon} \sum_{k=0}^{l-1} \mathbb{E} \left( \frac{1}{w} \sum_{t=0}^{w-1} \frac{1}{\sqrt{b}} Z_{1+k\ell+b+t\ell} 1_{|Z_{1+k\ell+b+t\ell}| > \sqrt{n} \delta} \right)^2. \]

Notice that

\[ \sqrt{\mathbb{E} \left( \frac{1}{w} \sum_{t=0}^{w-1} \frac{1}{\sqrt{b}} Z_{1+k\ell+b+t\ell} 1_{|Z_{1+k\ell+b+t\ell}| > \sqrt{n} \delta} \right)^2} \leq \frac{1}{w} \sum_{t=0}^{w-1} \sqrt{\mathbb{E} \left( \frac{1}{\sqrt{b}} Z_{1+k\ell+b+t\ell} 1_{|Z_{1+k\ell+b+t\ell}| > \sqrt{n} \delta} \right)^2} \leq \frac{1}{w} \sum_{t=0}^{w-1} \sqrt{\frac{b}{n \delta^2} \mathbb{E} \left( \frac{1}{\sqrt{b}} Z_{1+k\ell+b+t\ell} \right)^4} = O \left( \frac{1}{\sqrt{n}} \right). \]
Moreover,
\[
\frac{1}{l} \sum_{k=0}^{l-1} \mathbb{E} \left( \frac{1}{w} \sum_{t=0}^{w-1} \frac{1}{\sqrt{b}} Z_{1+kb+td,b} \right)^2 \leq \frac{1}{lw^2} \sum_{k=0}^{l-1} \sum_{t=0}^{w-1} \sum_{\tau=0}^{w-1} \left| \text{Cov} \left( \frac{1}{\sqrt{b}} Z_{1+kb+td,b}, \frac{1}{\sqrt{b}} Z_{1+kb+(t+\tau)d,b} \right) \right| \leq \frac{8}{lw^2} \sum_{k=0}^{l-1} \sum_{t=0}^{w-1} \sum_{\tau=0}^{w-1} \left( \sup_{k,t} \mathbb{E} \left| \frac{1}{\sqrt{b}} Z_{1+kb+td,b} \right|^4 \right)^{1/2} \alpha \frac{1/2}{w} (\tau d - b + 1).
\]

The last inequality is a consequence of an inequality for $\alpha$-mixing sequences with bounded fourth moments (see Doukhan (1994)).

Finally, for some $\zeta > 1/2$ we get
\[
\frac{1}{l} \sum_{k=0}^{l-1} \mathbb{E} \left( \frac{1}{w} \sum_{t=0}^{w-1} \frac{1}{\sqrt{b}} Z_{1+kb+td,b} \right)^2 \leq \frac{C_2}{w} \sum_{\tau=q+2}^{w-1} \alpha \frac{1/2}{w} (\tau d - b + 1) + O \left( \frac{q+1}{w} \right) \leq \frac{C_2}{w} \sum_{\tau=1}^{w-1} \frac{1}{\tau^c d^\delta} + O \left( \frac{q+1}{w} \right),
\]

where $q$ is a natural number such that $(q - 1)d \leq b \leq qd$. Using Toeplitz lemma we have
\[
\frac{C_2}{w} \sum_{\tau=1}^{w-1} \frac{1}{\tau^c d^\delta} \to 0.
\]

and also the demanded convergence of II to zero in probability and therefore (3.5). If we assume additionally that $\sum_{k=1}^{\infty} \alpha_y^{\delta/(6+\delta)}(k) < \infty$, we can easily show that the last expression is $O(1/w)$.

Now we prove (3.4). The first step is to show consistency of our bootstrap method in the one dimensional case, e.g.,
\[
d_0 \left( \mathcal{L} \left( \sqrt{w} \left( \hat{\mu}_i - \mu_i \right) \right), \mathcal{L}^* \left( \sqrt{w} \left( \hat{\mu}_i^* - \text{E}^* \hat{\mu}_i^* \right) \right) \right) \overset{P}{\to} 0,
\]

where $i = 1, \ldots, d$. Without the loss of generality we take $i = 1$.

In each block of length $b$, there are $v = \lfloor w/l \rfloor$ observations stemming from the marginal distribution $\mathcal{L}(X_1)$. The rest $r$ observations $(r = \lfloor w/l - \lfloor w/l \rfloor \rfloor l)$ are divided into $r$ blocks. So we have $l - r$ blocks with $v$ observations and $r$ blocks with $v + 1$ observations having marginal distribution $\mathcal{L}(X_1)$. Let $\tilde{U}_i$ be the sum of those observations in the $i$-th block e.g. block of the form $(X_{1+(i-1)b}, \ldots, X_{ib})$, where $i = 1, \ldots, l$. Let $U_i = \tilde{U}_i - \text{E}\tilde{U}_i$ for $i = 1, \ldots, l$. When the number of summands equal $v$, we denote the sum $\tilde{U}_i$ by $U_i^I$ and otherwise by $U_i^{II}$.

The estimator of $\mu_1$ and its bootstrap version are of the form:
\[
\hat{\mu}_1 = \frac{1}{w} \sum_{i=1}^{l} \tilde{U}_i, \quad \hat{\mu}_1^* = \frac{1}{w} \sum_{i=1}^{l} \tilde{U}_i^*.
\]
To establish the assertion of the theorem we use Corollary 2.4.8 in Araujo and Giné (1980). For this we need to show that for any \( \delta > 0 \) the following conditions are fulfilled:

\[
\sum_{i=1}^{l} P^*(\frac{1}{\sqrt{w}} |U_i^*| > \delta) \xrightarrow{p} 0, \tag{6.6}
\]

\[
\sum_{i=1}^{l} E^*(\frac{1}{\sqrt{w}} U_i^* 1_{|U_i^*| > \sqrt{w}\delta}) \xrightarrow{p} 0, \tag{6.7}
\]

\[
\sum_{i=1}^{l} \text{Var}^*(\frac{1}{\sqrt{w}} U_i^* 1_{|U_i^*| \leq \sqrt{w}\delta}) \xrightarrow{p} \sigma_1^2. \tag{6.8}
\]

Since the \( \alpha \)-mixing coefficient of the random variables \( Y_1, Y_1 + d, \ldots, Y_1 + (w-1)d \) satisfies \( \alpha_1(\tau) = \alpha_Y(\tau d) \), we get that \( E \left( \frac{1}{\sqrt{vU}} I_{\left| U \right| > \sqrt{w}} \right) \) and \( E \left( \frac{1}{\sqrt{v+1U}} II_{\left| U \right| > \sqrt{w}} \right) \) are bounded (see Kim (1994)).

Since the main steps of the proof of (3.11) to (3.13) are using similar arguments as those needed for the proof of (3.5), we only show (6.6).

Notice that

\[
\sum_{i=1}^{l} P^*(\frac{1}{\sqrt{w}} |U_i^*| > \delta) = \sum_{j=1}^{w} P\left( |U_j^*| > \sqrt{w}\delta \right) + \sum_{j=1}^{w} P\left( |U_j^{II}| > \sqrt{w}\delta \right).
\]

The absolute expected value of the right hand side of the last equality above is less or equal to

\[
\begin{align*}
\frac{l-r}{w} \sum_{j=1}^{w} P\left( |U_j^*| > \sqrt{w}\delta \right) + \frac{r}{w} \sum_{j=1}^{w} P\left( |U_j^{II}| > \sqrt{w}\delta \right) & \leq \left( l - r \right) \sum_{j=1}^{w} \frac{v^{3/2}E\left[ \frac{1}{\sqrt{w}} U_j^* \right]^3}{\sqrt{w^3\delta^3}} + \frac{r}{w} \sum_{j=1}^{w} (v + 1)^{3/2} E\left[ \frac{1}{\sqrt{w+1}} U_j^{II} \right]^3 \leq O\left( \frac{1}{\sqrt{v}} \right).
\end{align*}
\]

To proceed with the proof of (3.4) we use Theorem A in Athreya (1987) and show first the pointwise convergence in probability of the characteristic function of the multidimensional bootstrap distribution. Notice that for any vector of constants \( c \in \mathcal{R}^d \) we have

\[
\phi_{\sqrt{w}(\hat{\mu}^*-E\hat{\mu}^*)}^*(c) = \phi_{\sqrt{w}c^T(\hat{\mu}^*-E\hat{\mu}^*)}^*(1).
\]

Let \( G_{t+sd} = c_t X_{t+sd} \), where \( s = 0, \ldots, w-1 \) and \( s = 1, \ldots, d \). The series \( \{G_t\} \) fulfills the assumptions of Theorem 3.1, which means that GSBB is consistent for the global mean \( \mu_G \). Moreover,

\[
\text{Var}\left( \sqrt{w} \hat{\mu}_G \right) \rightarrow c^T \Sigma c
\]

and

\[
\sqrt{w} \left( \hat{\mu}_G^* - E\hat{\mu}_G^* \right) = \sqrt{w} c^T \left( \hat{\mu}^* - E\hat{\mu}^* \right).
\]

Finally, by Theorem A in Athreya (1987) we get

\[
\phi_{\sqrt{w}c^T(\hat{\mu}^*-E\hat{\mu}^*)}^*(1) \xrightarrow{p} \phi_{N(0,c^T\Sigma c)}(1) = \phi_{N(0,\Sigma)}(c)
\]

23
and
\[
P^* \left( \sqrt{n} (\hat{\mu}^* - E^* \hat{\mu}^*) \leq x \right) \xrightarrow{p} F_{\mathcal{N}(0,\Sigma)}(x),
\]
for any \( x \in \mathbb{R}^d \), where \( F_{\mathcal{N}(0,\Sigma)}(x) \) is the cumulative distribution function of \( \mathcal{N}(0,\Sigma) \). Assertion (3.4) follows then from Polya’s theorem. \( \square \)

**Proof of Theorem 4.1:** Our assumptions imply the validity of (4.1) and (4.2). By the same reasoning as in the proof of Theorem 3.1, we obtain consistency of GSBB for the overall mean. For the seasonal means, we construct blocks \( \bar{U}_i \) from the observations that have the same expected value and variance, but not necessarily the distribution. Using same arguments together with Lemma A.6 from Synowiecki (2007) we get the thesis. Please note that we use Lemma A.6. for periodically correlated processes (PC) i.e. processes that have periodic mean and covariance functions. However, this Lemma is dedicated to the much wider class of the almost periodically correlated processes (APC) which have almost periodic mean and covariance functions. Definition and properties of almost periodic functions can be found in Besicovitch (1932). \( \square \)

**Proof of Theorems 4.2 and 4.3:** The proof is based on the proof of Theorem 4.1 from Lahiri (2003) and Theorem 3.5 from Synowiecki (2008). First we show (4.4).

Note that we can decompose the bootstrap statistic as
\[
\sqrt{n} \left( H \left( \hat{\mu}^* \right) - H \left( E^* \hat{\mu}^* \right) \right) = \sqrt{n} H' \left( \hat{\mu}^* \right) \left( \hat{\mu}^* - E^* \hat{\mu}^* \right) + R_n^*.
\]
Using a conditional Slutsky’s theorem (see Lahiri (2003) Lemma 4.1), it is enough to show that for any \( \varepsilon > 0 \)
\[
P^* \left( |R_n^*| > \varepsilon \right) \xrightarrow{p} 0.
\]
Denote \( t_n = \left| E^* \hat{\mu}^* - \bar{\mu} \right| \) and \( T_n^* = \sqrt{n} \left( \hat{\mu}^* - E^* \hat{\mu}^* \right) \).

Note that using Lagrange theorem, on the set \( \{ t_n \leq \eta \} \cap \{ \left| \hat{\mu}^* - E^* \hat{\mu}^* \right| \leq \eta \} \) we have
\[
|T_n^*| = \sqrt{n} \left| H' \left( \zeta \hat{\mu}^* + (1 - \zeta) E^* \hat{\mu}^* \right) \right| \leq C \left( \zeta \hat{\mu}^* + (1 - \zeta) E^* \hat{\mu}^* - \bar{\mu} \right) \left| T_n^* \right| = C \left( \hat{\mu}^* - E^* \hat{\mu}^* \right) + t_n \left| T_n^* \right| \leq C_1 \left( \left| \hat{\mu}^* - E^* \hat{\mu}^* \right|^\kappa + t_n^\kappa \right) |T_n^*|,
\]
where \( C, C_1 \) are some positive constants independent of \( n \) and \( \zeta \in [0,1] \).

Moreover,
\[
\{ |R_n^*| > \varepsilon \} = \begin{cases} |R_n^*| > \varepsilon, |\hat{\mu}^* - E^* \hat{\mu}^*| \leq \eta, t_n \leq \eta \} \cup \{ |R_n^*| > \varepsilon, |\hat{\mu}^* - E^* \hat{\mu}^*| > \eta, t_n \leq \eta \} \cup \{ |R_n^*| > \varepsilon, t_n > \eta \} \cup \{ |R_n^*| > \varepsilon, |\hat{\mu}^* - E^* \hat{\mu}^*| \leq \eta, t_n \leq \eta \} \cup \{ |\hat{\mu}^* - E^* \hat{\mu}^*| > \eta \} \cup \{ t_n > \eta \}.
\]
Using the decomposition above, we get

\[
P^* (|R_n^*| > \varepsilon) \leq P^* \left( |R_n^*| > \varepsilon, \left| \hat{\mu} - E^* \hat{\mu}^* \right| \leq \eta \right) 1_{t_n \leq \eta} + P^* \left( \left| \hat{\mu} - E^* \hat{\mu}^* \right| > \eta \right) + 1_{t_n > \eta} \leq
\]

\[
\leq P^* \left( C_1 \left( \left| \hat{\mu} - E^* \hat{\mu}^* \right| + t_n^\kappa \right) |T_n^*| > \varepsilon \right) + P^* \left( \left| \hat{\mu} - E^* \hat{\mu}^* \right| > \eta \right) + 1_{t_n > \eta} \leq
\]

\[
\leq P^* \left( C_1 n^{-\frac{\varepsilon}{2}} |T_n^*|^{1+\kappa} > \frac{\varepsilon}{2} \right) + P^* \left( C_1 t_n^\kappa |T_n^*| > \frac{\varepsilon}{2} \right) + P^* \left( n^{\frac{1}{2}} |T_n^*| > \eta \right) + 1_{t_n > \eta}.
\]

Since

\[
\left\{ C_1 t_n^\kappa |T_n^*| > \frac{\varepsilon}{2} \right\} \subset \left\{ t_n^\kappa > \frac{\eta}{\log n} \right\} \cup \left\{ |T_n^*| > \frac{\varepsilon}{2C_1 \eta} \log n \right\},
\]

we have

\[
P^* (|R_n^*| > \varepsilon) \leq 3P^* (|T_n^*| > C (\varepsilon, \kappa, \eta) \log n) + 2 \cdot 1_{t_n > \frac{\eta}{\log n}}.
\]

From consistency of GSBB the first summand of the right-hand side is \(o_P(1)\). For the second summand we can write

\[
\mathbb{E}\left( 1_{t_n > \frac{\eta}{\log n}} \right) \leq P\left( t_n > \frac{\eta}{\log n} \right) \leq \frac{\log^2 n}{\eta^2} \mathbb{E}\left( t_n^2 \right).
\]

To calculate \( \mathbb{E}(t_n^2) \) we introduce the following notation:

\[
n = lb + r_1, \quad n = wd + r_2, \quad b = vd + r,
\]

where as before \( l \) is the number of blocks of length \( b \) and \( w \) is the number of periods in the sample. Additionally, \( v \) is the number of periods in the block and \( r_1 \in [0, b) \) and \( r_2, r \in [0, d) \).

The expected value of our estimator is of the form

\[
\mathbb{E}^* \hat{\mu}^* = \frac{1}{n} \sum_{s=0}^{l} \mathbb{E}^* (Z_{s+sb}^*).
\]

For simplicity we omit the second subscript denoting the block length, because the last block is of length \( r_1 \). Moreover, we do not consider the circular version of GSBB, so we have \( n - b + 1 = (w - v)d + r_2 - r + 1 \) different blocks. Without loss of generality, we can assume that \( r_2 - r + 1 \in [0, d) \). Otherwise, \( n - b + 1 \) should be rewritten as \( (w - v - 1)d + d + r_2 - r + 1 \). Note that now we do not have the same probability of selection for each block. For blocks beginning with observations \( 1, \ldots, r_2 - r + 1 \) its value is \( 1/(w - v + 1) \) and for others \( 1/(w - v) \).

We also need to define the following sets to simplify further notation:

\[
A = \{s : s = 1 + tb, t = 0, \ldots, l\},
\]

\[
A^I = \{s : s \in A, s \mod d = g \leq r_2 - r + 1\},
\]

\[
A^{II} = A \setminus A^I,
\]

\[
G = \{g_0, \ldots, g_l : g_i = s_i \mod d, s_i \in A\}
\]
and whenever the result of modulo operation is equal to zero we take $d$ instead.

As the result we get

$$E(\hat{\nu} - \nu)^2 \leq 6E \left( \frac{1}{n} \sum_{s \in A} \frac{1}{w-v} \sum_{k=0}^{w-v-1} Z_{(s \mod d) + kd} - 1 \right)^2 +$$

$$+ 6E \left( \frac{1}{n} \frac{1}{w-v} \sum_{k=0}^{w-v-1} Z_{g_{k+1} + kd} \right)^2 +$$

$$+ 3E \left( \frac{1}{n} \sum_{s \in A} \sum_{k=0}^{w-v-1} \frac{1}{(w-v)(w-v+1)} Z_{(s \mod d) + kd} \right)^2 +$$

$$+ 3E \left( \frac{1}{n} \sum_{s \in A} \frac{1}{w-v+1} Z_{(s \mod d) + (w-v)d} \right)^2 .$$

We denote the terms on the right-hand side by $I, II, III$ and $IV$, respectively.

For any $s, w, v$ we have that $E(\hat{Z}_{(s \mod d) + (w-v)d})^2 = O(b^2)$. As a results $II - IV$ are of order $O(1/n^2)$.

The case of $I$ is a bit more complicated. It is equivalent to the situation when we consider time series in which each block of length $b$ can be chosen with probability $1/(w-v)$. The length of this series is $d + (w-v-1)d + b - 1 = wd + r - 1$. Remind that the set $G$ contains indexes that define the beginnings of considered blocks in each period. The maximal number of different elements in set $G$ is $d$. In $I$ we
take the sum of all blocks which beginnings corresponding to elements of $G$. Let’s consider for example the first element from this set $g_1$. We sum $w-v$ blocks of length $b$. As a result the sum all the elements of considered time series (of length $wd + r - 1$) can be written as

$$
(v + 1) \sum_{i=1}^{wd+r-1} X_i - \left( (v + 1) \sum_{i=1}^{g_1} X_i + (v - 1) \sum_{i=g_1+1}^{g_1+d-1} X_i + \cdots + \sum_{i=g_1+(v-1)d}^{g_1+vd-1} X_i + \sum_{i=g_1+(v-2)d}^{g_1+(vd-1)d-1} X_i \right) - 
$$

$$
= (v + 1) \sum_{i=1}^{wd+r-1} X_i - \left( (v + 1) \sum_{i=1}^{g_1} X_i + (v - 1) \sum_{i=g_1+1}^{g_1+d-1} X_i + \cdots + \sum_{i=g_1+(v-1)d}^{g_1+(vd-1)d-1} X_i \right) - 
$$

$$
= (v + 1) \sum_{i=1}^{wd+r-1} X_i - E_{g_1}
$$

assuming that $r < g_1$. We omit other cases because it will not affect the final order of consider expected value and would complicate a lot the notation. The set $A$ has $l + 1$ elements, but we consider the last element separately as the last block is shorter. Finally to get the order of $I$ we consider separately the difference between the first summand above and $\mu$ and the remaining part of expression above.

$$
E \left( \frac{(v + 1)l(wd + r - 1)}{(w - v)n} \bar{X} - \mu \right)^2 = E \left( \left( \frac{(v + 1)l(wd + r - 1)}{(w - v)n} - 1 \right) \bar{X} + P_n \right)^2 \leq 
$$

$$
\leq 2E \left( \left( \frac{(v + 1)l(wd + r - 1)}{(w - v)n} - 1 \right) \bar{X} \right)^2 + 2E (P_n)^2
$$

where $\bar{X} = 1/(w + r - 1) \sum_{i=1}^{wd+r-1} X_i$ and $P_n = \bar{X} - \mu$.

Moreover, $E(P_n)^2 = O(n^{-2})$ and $E(\bar{X})^2 = O(1)$, which means that

$$
E \left( \frac{(v + 1)l(wd + r - 1)}{(w - v)n} \bar{X} - \mu \right)^2 = O \left( \frac{b^2}{n^2} \right) + O \left( \frac{1}{n^2} \right).
$$

On the other hand, since $E(X_i)^2$ are uniformly bounded we get

$$
E \left( \sum_{i=0}^{l-1} E_{g_i} \right)^2 \leq C_1 \left\{ (v + 1)^2 + (v + v - 1 + \cdots + 1)^2 d^2 + (w - 2v - 1)^2 (d - r)^2 + 
$$

$$
+ (v + 1 + v + \cdots + 1)^2 d^2 \right\} \leq C_2 \left\{ v^2 + \frac{(v + 1)^2 v^2}{4} + (w - 2v - 1)^2 + \frac{(v + 1)^2 (v + 2)^2}{4} \right\},
$$

where $C_1$ and $C_2$ are positive constants independent of $n$.

Thus,

$$
E \left( \frac{1}{n(w - v)} \sum_{i=0}^{l} E_{g_i} \right)^2 = O \left( \frac{1}{n^2} \right) + O \left( \frac{b^2}{n^2} \right) + O \left( \frac{l^2}{n^2} \right).
$$

27
Finally, since \( E \left( \sqrt{\left| \hat{\mu} - \mu \right|^2} \right) = O(1/n) \) we get

\[
E \left( \left( E \hat{\mu} - \mu \right)^2 \right) = O \left( \frac{1}{n} \right) + O \left( \frac{1}{n^2} \right) + O \left( \frac{b^2}{n^2} \right) + O \left( \frac{l^2}{n^2} \right),
\]

which means that

\[
1_{t_n > \frac{a}{\log n}} = o_P(1).
\]

This ends proof of Theorem 4.2. Proof of Theorem 4.3 follows exactly the same steps after changing all absolute values to norm symbol.

**Proof of Theorem 4.4 and Corollary 4.3:** The proofs follow similar techniques as those presented in proof of Theorem 3.1 and techniques used by Leśkow and Synowiecki (2010).
References


