Ann. Inst. Statist. Math.
Vol. 55, No. 4, 765–796 (2003)

RESAMPLING TIME SERIES USING MISSING VALUES TECHNIQUES

ANDRÉS M. ALONSO[1,2], DANIEL PEÑA[2] AND JUAN ROMO[2]

[1]*Department of Mathematics, Universidad Autónoma de Madrid, 28049-Madrid, Spain*
[2]*Department of Statistics and Econometrics, Universidad Carlos III de Madrid, 28903-Getafe, Spain*

(Received October 17, 2001; revised January 31, 2003)

Abstract. Several techniques for resampling dependent data have already been proposed. In this paper we use missing values techniques to modify the moving blocks jackknife and bootstrap. More specifically, we consider the blocks of deleted observations in the blockwise jackknife as missing data which are recovered by missing values estimates incorporating the observation dependence structure. Thus, we estimate the variance of a statistic as a weighted sample variance of the statistic evaluated in a "complete" series. Consistency of the variance and the distribution estimators of the sample mean are established. Also, we apply the missing values approach to the blockwise bootstrap by including some missing observations among two consecutive blocks and we demonstrate the consistency of the variance and the distribution estimators of the sample mean. Finally, we present the results of an extensive Monte Carlo study to evaluate the performance of these methods for finite sample sizes, showing that our proposal provides variance estimates for several time series statistics with smaller mean squared error than previous procedures.

Key words and phrases: Jackknife, bootstrap, missing values, time series.

1. Introduction

The classical jackknife and bootstrap, as proposed by Quenouille (1949), Tukey (1958) and Efron (1979), are inconsistent in the case of dependent observations. During recent years these methods have been modified in order to account for the dependence structure of the data. The main existing procedures could be broadly classified as model based and model free. Model based procedures fit a model to the data and resample the residuals which mimic the i.i.d. errors of the model (see, e.g., Freedman (1984), Efron and Tibshirani (1986), Bose (1990) and Kreiss and Franke (1992)). Model free procedures consider blocks of consecutive observations and resample from these blocks as in the independent case (see, e.g. Carlstein (1986), Künsch (1989) and Liu and Singh (1992)). Sherman (1998) compares these approaches in terms of efficiency and robustness and concludes that for moderate sample sizes the model based variance estimators provide a small gain under the correct model and, under mild misspecification, have bias similar to model free estimators while being more variable.

In this paper we are interested in the moving blocks jackknife (MBJ) and the moving blocks bootstrap (MBB) introduced in Künsch (1989) and also in Liu and Singh (1992). These methods allow us to estimate the variance of statistics defined by functionals of finite dimensional marginal distributions, which include robust estimators of location and scale, least-squares estimators of the parameters of an AR model and certain versions of the sample correlations.

As is usual in jackknife methods, the variance estimator is obtained by a weighted sample variance of the statistic evaluated in a sample where some observations (blocks of consecutive observations, in this case) are deleted or downweighted. Künsch (1989) showed that the MBJ which smoothes transitions between observations left out and observations with full weight, reduces the bias. Other bias reducing resampling methods are: linear combinations of block bootstrap estimators with different block sizes, proposed by Politis and Romano (1995), and the matched-block bootstrap of Carlstein *et al.* (1998) that suggests using some block joining rule favoring blocks that are more likely to be close.

When the time series has strong dependence structure, computing autocovariance by deleting blocks of observations is likely to produce bias. An alternative procedure is to assume that the block of observations is missing. For independent data, deleting observations is equivalent to assuming that these observations are missing, but for autocorrelated data, as shown in Peña (1990), the two procedures are very different. Deleting a block of data effectively means replacing the observations in the block with their marginal expectation. Treating the block as missing is equivalent to substituting the observations in the block by their conditional expectations given the rest of the data. This is the procedure we propose in this paper. In our case, the observations left out in the MBJ are considered as missing data and they are replaced with a missing value estimate which takes into account the data dependence structure. Thus, the variance estimator is a weighted sample variance of the statistic evaluated in a "complete" series. This procedure could be interpreted as smooth transition between the two parts with full weight in the blockwise jackknife.

Also, we extend this idea to the blockwise bootstrap, defining a block of missing values between the blocks that form the bootstrap resample. Thus, the procedure resembles a block joining engine. In some sense, the matched-block bootstrap has a common point with the procedure that we propose in this paper, in particular with their autoregressive matching.

In Section 2 we define the MBJ with missing values techniques (M^2BJ) and the moving block bootstrap with missing values techniques (M^2BB). In Section 3 we present the missing values estimation procedures. In Section 4 the results of consistency of both methods as variance and distribution estimators for the sample mean are presented. Finally, the results of a simulation study comparing the MBJ and the M^2BJ, and the MBB and the M^2BB are presented in Section 5. All proofs are given in an Appendix.

2. Resampling algorithms

2.1 *Moving missing block jackknife*

Let $X_1, \ldots, X_N$ be observations from a stationary process $\{X_t\}_{t\in\mathbb{Z}}$ with joint distribution ρ. Let us suppose that the statistic T_N, whose variance or distribution we want to estimate, is defined by $T_N = T_N(\rho^N)$, where ρ^N is the empirical measure of $X_1, \ldots, X_N$. As noted by Künsch (1989), it is impossible to estimate ρ^N without assuming some structure for the stationary processes. Thus, we suppose that T_N can be written as a functional of empirical m-dimensional distributions, i.e. $T_N = T(\rho_N^m)$, where $\rho_N^m = n^{-1}\sum_{t=1}^n \delta_{Y_t}$ is an empirical m-dimensional marginal measure, $n = N - m + 1$, $Y_t = (X_t, \ldots, X_{t+m-1})$ are blocks of m consecutive observations and δ_y denotes the point mass at $y \in \mathbb{R}^m$.

The MBJ deletes or downweights blocks of m-tuples in the calculation of ρ_N^m:

$$\rho_N^{m,(j)} = (n - \|w_n\|_1)^{-1} \sum_{t=1}^{n} (1 - w_n(t-j))\delta_{Y_t}, \tag{2.1}$$

where $\|w_n\|_1 = \sum_{i=1}^{l} w_n(i)$ and $j = 0, 1, \ldots, n-l$. The weights satisfy $0 \le w_n(i) \le 1$ for $i \in \mathbb{Z}$, and $w_n(i) > 0$ iff $1 \le i \le l$, and l is the length of the downweighted block. Note that $w_n(i) = 1$ for $1 \le i \le l$ corresponds to the deletion of blocks, in such a case the optimal order of $l = l(n)$ is $O(n^{1/3})$. Bühlmann and Künsch (1994) propose a method for selecting the block length in blockwise bootstrap which can be modified for blockwise jackknife. Künsch (1989) suggests using $w_n(i) = h((i-1)/2)$ where function $h : (0,1) \to (0,1)$ is symmetric about $x = 1/2$.

The MBJ variance estimator of T_N is defined as

$$\widehat{\sigma}^2_{Jack} = (n - \|w_n\|_1)^2 n^{-1} (n - l + 1)^{-1} \|w_n\|_2^{-2} \sum_{j=0}^{n-l} (T_N^{(j)} - T_N^{(\cdot)})^2, \tag{2.2}$$

where $T_N^{(j)} = T_N(\rho_N^{m,(j)})$ is the j-th jackknife pseudo-value, $T_N^{(\cdot)} = (n-l+1)^{-1} \sum_{j=0}^{n-l} T_N^{(j)}$ and $\|w_n\|_2^2 = \sum_{i=1}^{l} w_n(i)^2$.

In our approach we will use the following expression to calculate ρ_N^m:

$$\tilde{\rho}_N^{m,(j)} = n^{-1} \left(\sum_{t=1}^{n} (1 - w_n(t-j))\delta_{Y_t} + \sum_{t=1}^{n} w_n(t-j)\delta_{\widehat{Y}_{t,j}} \right), \tag{2.3}$$

where $\widehat{Y}_{t,j}$ is an estimate of Y_t supposing that Y_t is a missing value in the j-th sample, and then calculate $\widetilde{T}_N^{(j)} = T_N(\tilde{\rho}_N^{m,(j)})$, for $j = 0, 1, \ldots, n-l$. $\widehat{Y}_{t,j}$ is a missing value estimate which takes into account the data dependence structure. In Section 3, we present in detail a method for obtaining $\widehat{Y}_{t,j}$ for stationary and invertible linear processes. Note that in (2.3), instead of eliminating the blocks indexed by $j+1, \ldots, j+l$, we consider those $l + m - 1$ consecutive observations as missing in the time series sequence. The M^{2}BJ and variance estimator is defined by

$$\widetilde{\sigma}^2_{Jack} = n(n - l + 1)^{-1} \|w_n\|_2^{-2} \sum_{j=0}^{n-l} (\widetilde{T}_N^{(j)} - \widetilde{T}_N^{(\cdot)})^2. \tag{2.4}$$

Also, we are interested in the distribution of T_N. We define the following jackknife-histograms, as in the subsampling method of Politis and Romano (1994):

$$H_N(x) = (n - l + 1)^{-1} \sum_{j=0}^{n-l} 1\{\tau_l l^{-1} (n-l)(T_N^{(j)} - T_N) \le x\}, \tag{2.5}$$

for the MBJ, and

$$\widetilde{H}_N(x) = (n - l + 1)^{-1} \sum_{j=0}^{n-l} 1\{\tau_l l^{-1} (n-l)(\widetilde{T}_N^{(j)} - T_N) \le x\}, \tag{2.6}$$

for the M^{2}BJ, where τ_l is an appropriate normalizing constant (typically $\tau_l = \sqrt{l}$), and $1\{E\}$ denotes the indicator of the event E.

2.2 *Moving missing block bootstrap*

In the case of bootstrap, we will use the circular block bootstrap (CBB) of Politis and Romano (1992) and Shao and Yu (1993) which can be described as follows. First, the sample is "extended" with $l-1$ observations:

$$X_{i,n} = \begin{cases} X_i & \text{if } i \in \{1,\dots,n\} \\ X_{i-n} & \text{if } i \in \{n+1,\dots,n+l-1\}. \end{cases} \tag{2.7}$$

Second, blocks of l consecutive observations $Z_{i,n} = (X_{i,n},\dots,X_{i+l-1,n})$ are defined. Then $\{Z_{i,n}\}_{i=1}^{n}$ is used for obtaining resamples $(Z_1^*,\dots,Z_s^*)$ such that $\Pr^*\{Z_j^* = Z_{i,n}\} = 1/n$, and this implies that $\Pr^*\{X_j^* = X_i\} = 1/n$. The number s of blocks in the bootstrap resample is selected such that $n \approx sl$. Then, the bootstrap estimator is $T_N^* = T_N(\rho_N^*)$, where $\rho_N^* = n^{-1}\sum_{t=1}^{n}\delta_{Z_t^*}$. The bootstrap variance and distribution of T_N^*,

$$\mathrm{Var}^*(T_N^*) = \mathrm{E}^*[(T_N^* - \mathrm{E}^*[T_N^*])^2] \tag{2.8}$$

and

$$\Pr^*\{(sl)^{1/2}(T_N^* - \mathrm{E}^*[T_N^*]) \le x\} \tag{2.9}$$

are used as variance and distribution estimators of T_N.

Other blockwise bootstraps have been proposed; for instance, the moving blocks bootstrap (MBB) of Künsch (1989) and Liu and Singh (1992), the non-overlapping block bootstrap (NBB) based on Carlstein (1986), and the stationary bootstrap (SB) of Politis and Romano (1994*b*).

The method that we propose can be described as follows: given a CBB resample $(Y_1^*,\dots,Y_s^*)$, i.e., s blocks of l consecutive observations, the idea of moving missing blocks bootstrap (M^{2}BB) is to introduce a block $\widehat{Y}_j^*$ of k "observations" between two consecutive blocks. For simplicity, we will use a fixed block size k for the blocks included and we will always introduce a final block in order to have ks missing observations. Thus, the M^{2}BB resample is $(Y_1^*,\widehat{Y}_1^*,Y_2^*,\dots,\widehat{Y}_{s-1}^*,Y_s^*,\widehat{Y}_s^*)$. Notice that the M^{2}BB resample has $s(l+k)$ observations, meanwhile the CBB resample has sl observations.

Another way of interpreting the M^{2}BB resample is to put $l+k$ as the block size in the CBB, and then to consider the last k observations in each block as missing values. Notice that it is possible to implement M^{2}BB using other blockwise bootstraps as the above mentioned procedures.

The M^{2}BB estimator is $\widetilde{T}_N^* = T_N(\widetilde{\rho}_N^*)$, where $\widetilde{\rho}_N^* = n^{-1}\sum_{t=1}^{n}\delta_{\widehat{Y}_t^*}$, and $\widehat{Y}_t^* = Y_t^*$ if $t \in \{1,\dots,l,l+k+1,\dots,2l+k,\dots,(s-1)(l+k)+1,\dots,(s-1)(l+k)+l\}$ and $\widehat{Y}_t^*$ is properly an estimate, otherwise. Then the bootstrap variance and distribution of $\widetilde{T}_N^*$,

$$\mathrm{Var}^*(\widetilde{T}_N^*) = \mathrm{E}^*[(\widetilde{T}_N^* - \mathrm{E}^*[\widetilde{T}_N^*])^2] \tag{2.10}$$

and

$$\Pr^*\{(s(l+k))^{1/2}(\widetilde{T}_N^* - \mathrm{E}^*[\widetilde{T}_N^*]) \le x\}, \tag{2.11}$$

are used as variance and distribution estimators of T_N.

3. Missing values techniques

There are a number of alternatives which can be used to obtain $\widehat{Y}_t$ for stationary and invertible linear processes, see e.g. Harvey and Pierse (1984), Peña and Maravall

(1991), and Beveridge (1992), and for some nonlinear processes as in Abraham and Thavaneswaran (1991). In this paper we will use the generalized least squares method presented in Peña and Maravall (1991).

If $\{X_t\}_{t\in\mathbb{Z}}$ is a stationary process that admits an AR(∞) representation: $\Phi(B)(X_t - \mu) = e_t$, where $\Phi(B) = \sum_{j=0}^{\infty} \phi_j B^j$, B is the backshift operator and $\mathrm{E}[X_t] = \mu$, let $z_t = X_t - \mu$, and assume that the finite series z_t has m missing values at times $T_1, T_2, \ldots, T_m$ with $T_i < T_j$. We fill the holes in the series with arbitrary numbers v_{T_i} and construct an "observed" series Z_t by:

$$Z_t = \begin{cases} z_t + \omega_t, & \text{if} \quad t \in \{T_1, T_2, \ldots, T_m\} \\ z_t, & \text{otherwise} \end{cases} \tag{3.1}$$

where $v_t = z_t + \omega_t$ and ω_t is an unknown parameter. In matrix notation, we have

$$Z = z + H\omega, \tag{3.2}$$

where Z and z are the series expressed as a $N \times 1$ vector, H is $N \times m$ matrix such that $H_{T_i,i} = 1$ and $H_{i,j} = 0$ otherwise, and ω is a $m \times 1$ vector of unknown parameters. Let Σ be the $N \times N$ autocovariance matrix of the series z_t, then the generalized least squares estimator of ω is

$$\widehat{\omega} = (H'\Sigma^{-1}H)^{-1}H'\Sigma^{-1}Z, \tag{3.3}$$

and the missing values estimates are obtained by

$$\widehat{Z} = Z - H\widehat{\omega} = Z - H(H'\Sigma^{-1}H)^{-1}H'\Sigma^{-1}Z. \tag{3.4}$$

Note that $\widehat{Z}$ obtained in (3.4) does not depend on the "arbitrary" value of ω. On the other hand, expression (3.4) assumes that μ and Σ are known. In order to make the above estimation method feasible we propose to replace them with the sample mean $\bar{X}$ and with an autoregressive estimator $\widehat{\Sigma}$, respectively (see, Lemmas 4.3 and 4.4).

When we apply this method to the j-th jackknife resample, the observations $X_{j+1}, \ldots, X_{j+l}$ are considered as $m = l$ consecutive missing values and the matrix $H = H_j$ takes the form

$$H_j = \begin{bmatrix} 0_{j\times l} \\ I_{l\times l} \\ 0_{N-(l+j)\times l} \end{bmatrix}_{N\times l}. \tag{3.5}$$

In the case of the bootstrap, we have $m = k\lceil n/(l+k)\rceil$ missing observations, where l is the length of the block in the bootstrap resample and k is the number of missing observations between two consecutive blocks. The matrix H is fixed and has the following expression

$$H = \begin{bmatrix} 0_{l\times k} & 0_{l\times k} & \cdots & 0_{l\times k} & 0_{l\times k} \\ I_{k\times k} & 0_{k\times k} & \cdots & 0_{k\times k} & 0_{k\times k} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0_{l\times k} & 0_{l\times k} & \cdots & 0_{l\times k} & 0_{l\times k} \\ 0_{k\times k} & 0_{k\times k} & \cdots & 0_{k\times k} & I_{k\times k} \end{bmatrix}. \tag{3.6}$$

We will use $v_t = z_t$ (then $\omega_t = 0$) in expression (3.4), in such a case

$$z - \widehat{Z} = H(H'\Sigma^{-1}H)^{-1}H'\Sigma^{-1}z, \tag{3.7}$$

and since $z = X - \mu$ and defining $\widehat{Z} = \widehat{X} - \mu$, we have

$$X - \widehat{X} = H(H'\Sigma^{-1}H)^{-1}H'\Sigma^{-1}(X - \mu), \tag{3.8}$$

which is a more tractable expression. For the bootstrap, the X in (3.8) is replaced with the X^*, forming the bootstrap resample.

Instead of expression (3.4), we could use the following nonparametric interpolator proposed by Bosq (1996):

$$\widehat{Z}_\tau = \frac{\sum_{t\in\mathcal{O}_\tau} Z_t \mathbb{K}((Z_{\tau,s} - Z_{t,s}))}{\sum_{t\in\mathcal{O}_\tau} \mathbb{K}((Z_{\tau,s} - Z_{t,s}))}, \tag{3.9}$$

where $\mathbb{K}$ is a strictly positive s-dimensional kernel, $Z_{\tau,s}$ denotes the s observed values near to Z_τ and $\mathcal{O}_\tau = \{t : Z_t \text{ and } Z_{t,s} \text{ are observed}\}$. Expression (3.9) may be interpreted as an approximation of $\mathrm{E}[Z_\tau \mid Z_{\tau,s}]$. The implementation of $\mathrm{M^2BJ}$ and $\mathrm{M^2BB}$ using (3.9) as an interpolator will be subject of future research.

4. Consistency results

We now study consistency for the sample mean of the proposed missing values approaches for jackknife and bootstrap. This case corresponds to $m = 1$, $T(F^1) = \int x dF^1(x) = \mathrm{E}[X_t] = \mu$. We will show that both procedures provide consistent estimators of the variance and the distribution of the sample mean. Theorems 4.1 and 4.3 present the fundamental results for the jackknife and Theorems 4.4 and 4.5 for the bootstrap. Also in Theorem 4.2 we establish the consistency of the MBJ of Künsch (1989) as a distribution estimator of linear statistics. Notice that Theorems 4.1, 4.2 and 4.3 can be extended to statistics with linear influence function.

Starting with the MBJ with missing values replacement we have that, according to (2.3), the statistic evaluated in the j-th completed resample is

$$\begin{aligned}\widetilde{T}_N^{(j)} &= n^{-1}\left(\sum_{t=1}^{n}(1 - w_n(t-j))X_t + \sum_{t=1}^{n} w_n(t-j)\widehat{X}_{t,j}\right) \\ &= T_N - n^{-1}\sum_{t=1}^{n} w_n(t-j)(X_t - \widehat{X}_{t,j}),\end{aligned} \tag{4.1}$$

where $T_N = n^{-1}\sum_{t=1}^{n} X_t$. First, we will consider the expression $\sum_{j=0}^{n-l}(\widetilde{T}_N^{(j)} - T_N)^2$. The use of T_N as a central measure seems more natural than $\widetilde{T}_N^{(\cdot)}$ because $T_N = T(F_N^1)$ (see Liu and Singh (1992)). We have that

$$n(\widetilde{T}_N^{(j)} - T_N) = -\sum_{t=1}^{n} w_n(t-j)(X_t - \widehat{X}_{t,j}) = -w_{n,j}'(X - \widehat{X}_j), \tag{4.2}$$

where $w_{n,j} = (w_n(1-j), \ldots, w_n(n-j))'$ and $\widehat{X}_j = (\widehat{X}_{1,j}, \ldots, \widehat{X}_{n,j})'$, with

$$\widehat{X}_{t,j} = \begin{cases} X_t & \text{if } w_n(t-j) = 0 \\ \widehat{X}_t & \text{if } w_n(t-j) > 0. \end{cases} \tag{4.3}$$

In order to prove the consistency of jackknife variance estimator we will use the following proposition established in Berk (1974):

PROPOSITION 4.1. *Suppose that* $\{X_t\}_{t\in\mathbb{Z}}$ *is a linear process such that* $\sum_{i=0}^{\infty}\phi_i x_{t-i} = e_t$, *where* $\{e_t\}_{t\in\mathbb{Z}}$ *are independent and identically distributed r.v.'s with* $\mathrm{E}[e_t] = 0$ *and* $\mathrm{E}[e_t^2] = \sigma^2$, *and* $\phi_0 = 1$. *Assume also that* $\Phi(z) = \sum_{i=0}^{\infty}\phi_i z^i$ *is bounded away from zero for* $|z| \leq 1$. *Then, there are constants* F_1 *and* F_2, $0 < F_1 < F_2$, *such that*

$$2\pi F_1 \leq \|\Sigma\|_{spec} \leq 2\pi F_2 \quad \text{and} \quad (2\pi F_2)^{-1} \leq \|\Sigma^{-1}\|_{spec} \leq (2\pi F_1)^{-1}, \tag{4.4}$$

where Σ *is the autocovariance matrix of* $\{X_t\}_{t\in\mathbb{Z}}$ *and* $\|\Sigma\|_{spec} = \max\{\sqrt{\lambda} : \lambda$ *is eigenvalue of* $\Sigma'\Sigma\}$ *denotes the spectral norm.*

Condition (4.4) allows us to establish the asymptotical unbiasedness of $\widetilde{\sigma}^2_{Jack}$. We replace in (2.4) $\widetilde{T}_N^{(\cdot)}$ with T_N and under standard assumptions we prove in Corollary 4.1 that the effect of this substitution is negligible.

LEMMA 4.1. *If the conditions of Proposition* 4.1 *hold, and assuming that* $w_n(i) = 1$ *if and only if* $1 \leq i \leq l$, $l = l(n) \to \infty$, *and* $\sum_{m=1}^{\infty} m|\gamma_m| < \infty$, *then* $\mathrm{E}[n\widetilde{\sigma}^2_{Jack}] \to \sigma^2_\infty = \sum_{m=-\infty}^{+\infty}\gamma_m$.

Now, we must prove that $\mathrm{Var}(n\widetilde{\sigma}^2_{Jack}) \to 0$. We have that

$$\begin{aligned} &\mathrm{Var}(n\widetilde{\sigma}^2_{Jack}) \\ &\quad = (n-l+1)^{-2}\|w_n\|_2^{-4}\sum_{j=0}^{n-l}\sum_{i=0}^{n-l}\mathrm{cov}(n^2(\widetilde{T}_N^{(j)} - T_N)^2, n^2(\widetilde{T}_N^{(i)} - T_N)^2). \end{aligned} \tag{4.5}$$

Note that $n^2(\widetilde{T}_N^{(j)} - T_N)^2 = \widetilde{w}'_{n,j}(X-\mu)(X-\mu)'\widetilde{w}_{n,j}$, where $\widetilde{w}_{n,j} = \Sigma^{-1}H_j (H_j'\Sigma^{-1}H_j)^{-1}H_j'w_{n,j}$; thus the only difference with Theorem 3.3 of Künsch (1989) is replacing $w_{n,j}$ with $\widetilde{w}_{n,j}$. A crucial aspect in his proof is the number of non zero elements ($l = o(n)$) in the vector $w_{n,j}$. The following lemma establishes that $\widetilde{w}_{n,j} = \bar{w}_{n,j} + o(l^{-1/2})$, where $\bar{w}_{n,j}$ has at most $l + 4\lceil l^{1/2}\rceil$ non zero elements.

LEMMA 4.2. *Under the conditions of Lemma* 4.1, *and assuming that* $\sum_{m=1}^{\infty} m^2|\gamma_m| < \infty$, *we have that* $\widetilde{w}_{n,j} = \bar{w}_{n,j} + o(l^{-1/2})$ *uniformly in* $j = 0, 1, \ldots, n-l$.

The next result follows by combining the previous Lemmas 4.1 and 4.2 and Theorem 3.3 of Künsch (1989).

THEOREM 4.1. *Under the conditions of Lemma* 4.2, *and assuming that* $\mathrm{E}[|X_t|^{6+\delta}] < \infty$ *with* $\delta > 0$, $\sum_{m=1}^{\infty} m^2\alpha_m^{\delta/(6+\delta)} < \infty$ *where* α_m *are the strong mixing coefficients, and* $l = o(n)$, *it follows that* $n\widetilde{\sigma}^2_{Jack} \xrightarrow{P} \sigma^2_\infty$.

COROLLARY 4.1. *Under the conditions of Theorem* 4.1, *and assuming that* $l = o(n^{1/2})$, *we have that*

$$n\widetilde{\sigma}^2_{Jack} = n^2(n-l+1)^{-1}\|w_n\|_2^{-2}\sum_{j=0}^{n-l}(\widetilde{T}_N^{(j)} - T_N)^2 + o_P(l^{-1}).$$

The previous results assume that the matrix Σ is known; the next lemma shows that the consistency result obtained in Theorem 4.1 holds if we replace Σ with an autoregressive estimator $\widehat{\Sigma}$, i.e., the $n\times n$ autocovariance matrix of an AR(p) process, with $p = p(n)$. We will use the matrix column-sum norm $\|A\|_{col} = \max\{\sum_{i=1}^n |a_{ij}| : j = 1,\ldots,n\}$, and the vector maximum norm $\|X\|_\infty = \max\{x_i : i = 1,\ldots,n\}$.

LEMMA 4.3. *Under the conditions of Theorem* 4.1, *and assuming that* $\|\Sigma^{-1}\|_{col} < M < \infty$, $p = o((n/\log n)^{1/6})$, *and* $l = o((n/\log n)^{2/9})$, *it follows that*

$$n\widetilde{\sigma}^2_{Jack} - n\widehat{\widetilde{\sigma}}^2_{Jack} = o_P(1), \tag{4.6}$$

where

$$\widehat{\widetilde{\sigma}}^2_{Jack} = n(n-l+1)^{-1}\|w_n\|_2^{-2}\sum_{j=0}^{n-l}(\widehat{\widetilde{T}}_N^{(j)} - T_N)^2, \tag{4.7}$$

and $\widehat{\widetilde{T}}_N^{(j)} = w_{n,j}H_j(H_j'\widehat{\Sigma}^{-1}H_j)^{-1}H_j'\widehat{\Sigma}^{-1}(X-\bar{X})$.

The condition $\|\Sigma^{-1}\|_{col} < M < \infty$ is satisfied by stationary and invertible ARMA processes. This is a direct consequence of the representation of Σ^{-1} in Galbraith and Galbraith (1974). Note that the proof is still valid if $\|\Sigma^{-1}\|_{col} = O(l^{1/4-\alpha})$ for some α such that $0 < \alpha < 1/4$.

Now, we prove that the moving block jackknife (MBJ) of Künsch (1989) could be used as an estimator of the distribution of a linear statistic. We will use the analogy between the subsampling method of Politis and Romano (1994*a*) and the blockwise jackknife. First, we introduce some notation: $T_{N,t} = T_b(X_t,\ldots,X_{t+b-1})$ is the estimator of $T(\rho)$ based on the block or subsample $(X_t,\ldots,X_{t-b+1})$. Let $J_b(\rho)$ be the sampling distribution of

$$\tau_b(T_{N,1} - T(\rho)), \tag{4.8}$$

where τ_b is the normalizing constant. Also define the corresponding cumulative distribution function

$$J_b(x,\rho) = \Pr{}_\rho\{\tau_b(T_{N,1} - T(\rho)) \le x\}, \tag{4.9}$$

and denote $J_N(\rho)$ the sampling distribution of $\tau_n(T_N - T(\rho))$. The approximation of $J_N(\rho)$ proposed by subsampling is

$$L_N(x) = (N-b+1)^{-1}\sum_{t=1}^{N-b+1} 1\{\tau_b(T_{N,t} - T_N) \le x\}. \tag{4.10}$$

The only essential assumption in Politis and Romano's approach is that there exists a limiting law $J(\rho)$ such that $J_N(\rho)$ converges weakly to a limit law $J(\rho)$, as $n \to \infty$.

For simplicity, we only prove the consistency of MBJ for linear statistics, as the sample mean

$$T_N = n^{-1} \sum_{t=1}^{n} f(Y_t), \tag{4.11}$$

where $Y_t = (X_t, \ldots, X_{t+m-1})$, $n = N - m + 1$ and f is a continuous function on $\mathbb{R}^m$.

In the MBJ we have l deleted blocks $(Y_{j+1}, \ldots, Y_{j+l})$ which corresponds to $b = l + m - 1$ consecutive observations. Using (2.1), we have

$$\begin{aligned} T_N^{(j)} &= (n-l)^{-1} \sum_{t=1}^{n} (1 - w_n(t-j)) f(Y_t) \\ &= (n-l)^{-1} n T_N - (n-l)^{-1} \sum_{t=j+1}^{j+l} f(Y_t) \\ &= (n-l)^{-1} n T_N - (n-l)^{-1} T_{N,j+1}. \end{aligned} \tag{4.12}$$

Assuming without loss of generality that $m = 1$,

$$T_N^{(j)} - T_N = -l(n-l)^{-1}(T_{N,j+1} - T_N), \tag{4.13}$$

$$-\tau_l l^{-1}(n-l)(T_N^{(j)} - T_N) = \tau_l (T_{N,j+1} - T_N), \tag{4.14}$$

and

$$L_N(x) = (n-l+1)^{-1} \sum_{j=0}^{n-l} 1\{\tau_l l^{-1}(n-l)(T_N - T_N^{(j)}) \le x\}. \tag{4.15}$$

The MBJ analogous to $L_N(x)$ is

$$H_N(x) = (n-l+1)^{-1} \sum_{j=0}^{n-l} 1\{\tau_l l^{-1}(n-l)(T_N^{(j)} - T_N) \le x\}. \tag{4.16}$$

We obtain consistency under the following assumption:

ASSUMPTION 4.1. There exists a symmetric limiting law $J(\rho)$ such that $J_n(\rho)$ converges weakly to a limit law $J(\rho)$, as $n \to \infty$.

The following theorem shows that moving blocks jackknife-histograms are consistent estimators of the distribution.

THEOREM 4.2. *Suppose that Assumption* 4.1 *holds and that* $\tau_l/\tau_n \to 0$, $l/n \to 0$ *and* $l \to \infty$ *as* $n \to \infty$. *Also assume that the* α*-mixing sequence satisfies that* $\alpha_X(k) \to 0$ *as* $k \to \infty$.

1. *If* x *is a continuity point of* $J(\cdot, \rho)$, *then* $H_N(x) \to J(x, \rho)$ *in probability.*
2. *If* $J(\cdot, \rho)$ *is continuous, then* $\sup_x |H_N(x) - J(x, \rho)| \to 0$ *in probability.*

Also, we could use the M^{2}BJ method as distribution estimator. We establish consistency for the sample mean. The MBJ and the M^{2}BJ statistics satisfy

$$T_N^{(j)} - T_N = -(n-l)^{-1} \sum_{t=1}^{n} w_n(t-j)(X_t - T_N), \tag{4.17}$$

and

$$\widetilde{T}_N^{(j)} - T_N = -n^{-1} \sum_{t=1}^{n} w_n(t-j)(X_t - \widehat{X}_{t,j}). \tag{4.18}$$

Therefore,

$$\begin{aligned} l^{-1/2} n(\widetilde{T}_N^{(j)} - T_N) &= l^{-1/2}(n-l)(T_N^{(j)} - T_N) \\ &\quad + l^{-1/2} \sum_{t=1}^{n} w_n(t-j)(\widehat{X}_{t,j} - T_N). \end{aligned} \tag{4.19}$$

The following proposition proves that the second term in the right hand side of (4.19) is $o_P(1)$.

PROPOSITION 4.2. *Suppose that* $w_n(i) = 1$ *iff* $1 \le i \le l$, $\sum_{k=1}^{\infty} k|\gamma_k| < \infty$, *and* $\|\Sigma^{-1}\|_{col} < M < \infty$. *Also assume that* $l/n \to 0$ *and* $l \to \infty$ *as* $n \to \infty$. *Then* $l^{-1/2} \sum_{t=1}^{n} w_n(t-j)(\widehat{X}_{t,j} - T_N) = o_P(1)$ *uniformly in* j.

Consistency follows now from Theorem 2.1 in Politis and Romano (1994*a*), Proposition 4.2, and the asymptotic equivalence lemma (cf. Lemma 4.7 of White (1984)), i.e. having two random sequences satisfying $A_n - B_n = o_P(1)$ and $B_n \xrightarrow{d} B$ then we conclude that $A_n \xrightarrow{d} B$.

THEOREM 4.3. *Under the conditions in Proposition* 4.2, *for all* x

$$\widetilde{H}_N(x) = (n-l+1)^{-1} \sum_{j=0}^{n-l} 1\{\tau_l l^{-1}(n-l)(\widetilde{T}_N^{(j)} - T_N) \le x\} \to J(x,\rho) \tag{4.20}$$

in probability.

Remark 4.1. In the proof of Lemma 4.3 we obtained that $n(\widetilde{T}_N^{(j)} - T_N) - n(\widehat{\widetilde{T}}_N^{(j)} - T_N)$ is $o_P(l^{3/2} \max(l,p)^{1/2} (n/\log n)^{-1/3})$, thus (4.20) holds if we replace Σ with $\widehat{\Sigma}$, and $l = o((n/\log n)^{2/9})$.

Now, we prove that the M^{2}BB give consistent estimators of the variance and the distribution of the sample mean. We have the following CBB and M^{2}BB statistics:

$$\bar{X}_{n,s}^{*} = (sl)^{-1} \sum_{i=1}^{s} \sum_{j=1}^{l} X_{(i-1)l+j}^{*}, \tag{4.21}$$

and

$$\tilde{\bar{X}}^*_{n,s} = (s(l+k))^{-1}\sum_{i=1}^{s}\left(\sum_{j=1}^{l} X^*_{(i-1)(l+k)+j} + \sum_{j=l+1}^{l+k}\widehat{X}^*_{(i-1)(l+k)+j}\right), \tag{4.22}$$

where $\widehat{X}^*_t$ is an estimate of the "missing observation" X^*_t, that takes into account the dependence structure on the original process $\{X_t\}$.

We could write the M²BB analogous to $(sl)^{1/2}(\bar{X}^*_{n,s} - \bar{X}_n)$ as follows:

$$\begin{aligned}(s(l+k))^{1/2}(\tilde{\bar{X}}^*_{n,s} - \bar{X}_n) &= (s(l+k))^{-1/2}\left(\sum_{i=1}^{s}\sum_{j=1}^{l}(X^*_{(i-1)(l+k)+j} - \bar{X}_n)\right.\\ &\qquad\left. + \sum_{i=1}^{s}\sum_{j=l+1}^{l+k}(\widehat{X}^*_{(i-1)(l+k)+j} - \bar{X}_n)\right)\\ &= (l/(l+k))^{1/2}(sl)^{1/2}(\bar{X}^*_{n,s} - \bar{X}_n)\\ &\quad + (s(l+k))^{-1/2}\sum_{i=1}^{s}\sum_{j=l+1}^{l+k}(\widehat{X}^*_{(i-1)(l+k)+j} - \bar{X}_n).\end{aligned} \tag{4.23}$$

Notice that if $k/l \to 0$ as $n \to \infty$ we have that $l/(l+k) \to 1$ and thus the first term in (4.23) satisfies the conditions in Theorem 1 in Politis and Romano (1992). The following proposition establishes that the second term in (4.23) is $o_P(1)$.

PROPOSITION 4.3. *Suppose that* $\sum_{m=1}^{\infty} m|\gamma_m| < \infty$, $\sum_{m=1}^{\infty} m^2\alpha_m^{\delta/(6+\delta)} < \infty$, *where* α_m *are the strong mixing coefficients, and* $\|\Sigma^{-1}\|_{col} < M < +\infty$. *Also assume that* $l/n \to 0$, $l \to \infty$, *and* $k/l \to 0$ *as* $n \to \infty$. *Then* $(s(l+k))^{-1/2}\sum_{i=1}^{s}\sum_{j=l+1}^{l+k}\cdot(\widehat{X}^*_{(i-1)(l+k)+j} - \bar{X}_n) = o_P(1)$.

The previous results assume that the matrix Σ is known; the next lemma is the analogous to Lemma 4.3. Notice that in M²BJ the number of missing values is l and in M²BB is sk.

LEMMA 4.4. *Under the conditions of Proposition* 4.3, *and assuming that* $\|\Sigma^{-1}\|_{col} < M < \infty$, $p = o((n/\log n)^{1/6})$, *and* $sk = o((n/\log n)^{2/9})$, *it follows that*

$$(s(l+k))^{1/2}(\tilde{\bar{X}}^*_{n,s} - \bar{X}_n) - (s(l+k))^{1/2}(\hat{\tilde{\bar{X}}}^*_{n,s} - \bar{X}_n) = o_P(1), \tag{4.24}$$

where $\hat{\tilde{\bar{X}}}^*_{n,s} = (s(l+k))^{-1}\sum_{i=1}^{s}(\sum_{j=1}^{l} X^*_{(i-1)(l+k)+j} + \sum_{j=l+1}^{l+k}\widehat{\widehat{X}}^*_{(i-1)(l+k)+j})$, *and* $\widehat{\widehat{X}}$ *is obtained replacing* Σ *with its autoregressive estimator* $\widehat{\Sigma}^{-1}$ *in* (3.4).

Now, using the statement (1) of Theorem 1 in Politis and Romano (1992), Proposition 4.3 and the Cauchy-Schwarz inequality we have that

THEOREM 4.4. *Under the conditions in Proposition* 4.3, *for all* x

$$\mathrm{Var}^*((s(l+k))^{1/2}(\tilde{\bar{X}}^*_{n,s} - \bar{X}_n)) \xrightarrow{P} \sigma^2_\infty. \tag{4.25}$$

And from statement (2) of Theorem 1 in Politis and Romano (1992), Proposition 4.3 and the asymptotic equivalence lemma, we obtain the consistency result.

THEOREM 4.5. *Under the conditions in Proposition* 4.3, *for all* x

$$\Pr{}^*\{(s(l+k))^{1/2}(\tilde{\bar{X}}^*_{n,s} - \bar{X}_n) \leq x \mid X_1, \ldots, X_n\} - \Pr\{n^{1/2}(\bar{X}_n - \mu) \leq x\} \to 0, \tag{4.26}$$

for almost all sample sequences $X_1, \ldots, X_N$.

In the next section we present an extensive simulation comparing the proposed moving missing blocks procedures with the MBJ and MBB methods. The theoretical comparison and confirmation of the superiority of the proposed methods, even in the case of linear processes, involves some unmanageable expressions that made those comparisons difficult. On the other hand, notice that using the nonparametric approach mentioned at end of Section 3 it is possible to implement the moving missing blocks methods without the linearity assumption.

5. Simulations

In this section, we compare the performance of the moving blocks jackknife (MBJ) and the moving blocks jackknife with missing values replacement (M^2BJ), and the

Table 1. MBJ and M^2BJ to estimate σ_N^2 in the case of the sample mean. (+) denotes cases where the MBJ outperforms the M^2BJ. (*) denotes cases where the M^2BJ outperforms the MBJ.

Model	N	$\log \sigma_N^2$	E_1	SD_1	L_1	MSE_1	E_2	SD_2	L_2	MSE_2
M1	480	3.21	2.95	0.25	35	0.130 (0.011)	2.90	0.24	45	0.152 (0.011)
M1	120	3.18	2.57	0.36	15	0.502 (0.030)	2.51	0.34	20	0.568 (0.032)
M2	480	1.68	1.61	0.09	2	0.012 (0.001)	1.65	0.09	7	0.009 (0.001)*
M2	120	1.70	1.57	0.18	2	0.050 (0.004)	1.55	0.18	7	0.056 (0.006)
M3	480	1.82	1.67	0.14	15	0.045 (0.003)	1.64	0.16	20	0.056 (0.004)
M3	120	1.83	1.52	0.17	4	0.119 (0.006)	1.38	0.23	10	0.251 (0.020)+
M4	480	−1.16	−1.09	0.16	25	0.030 (0.003)	−1.13	0.12	10	0.017 (0.002)*
M4	120	−1.13	−1.09	0.22	15	0.050 (0.005)	−1.07	0.16	4	0.029 (0.003)*
M5	480	−3.14	−3.04	0.22	85	0.057 (0.005)	−3.09	0.22	35	0.048 (0.005)
M5	120	−2.93	−2.57	0.18	30	0.166 (0.008)	−2.89	0.31	15	0.100 (0.012)*
M6	480	−3.05	−2.96	0.21	60	0.054 (0.004)	−3.03	0.10	5	0.011 (0.001)*
M6	120	−2.92	−2.85	0.21	30	0.051 (0.006)	−2.88	0.19	2	0.038 (0.004)*
M7	480	−1.81	−1.72	0.19	40	0.043 (0.004)	−1.71	0.15	15	0.032 (0.003)*
M7	120	−1.76	−1.64	0.24	20	0.073 (0.007)	−1.67	0.25	8	0.073 (0.007)
M8	480	1.18	1.09	0.11	8	0.020 (0.001)	1.05	0.11	10	0.030 (0.002)+
M8	120	1.17	0.99	0.16	4	0.058 (0.005)	0.90	0.22	7	0.118 (0.011)+

E_1, SD_1 and MSE_1 denotes the average, standard deviation and the mean square error of the statistic $\log N\hat{\sigma}_N^2$ in the 1000 replications using MBJ method. E_2, SD_2 and MSE_2 are the corresponding statistics using M^2BJ method.

Table 2. MBJ and M^2BJ to estimate σ_N^2 in the case of the median. (*) denotes cases where the M^2BJ outperforms the MBJ.

Model	N	$\log\sigma_N^2$	E_1	SD_1	L_1	MSE_1	E_2	SD_2	L_2	MSE_2
M1	480	3.31	3.04	0.36	35	0.202 (0.022)	3.37	0.31	15	0.101 (0.010)*
M1	120	3.29	2.67	0.52	20	0.654 (0.054)	2.81	0.48	15	0.458 (0.054)*
M2	480	2.04	2.02	0.43	35	0.183 (0.023)	1.93	0.42	20	0.187 (0.024)
M2	120	2.05	1.97	0.56	15	0.315 (0.035)	1.98	0.52	10	0.273 (0.033)
M3	480	2.05	1.91	0.42	20	0.194 (0.024)	1.90	0.39	15	0.177 (0.022)
M3	120	2.03	1.71	0.50	10	0.356 (0.038)	1.77	0.52	6	0.337 (0.036)
M4	480	0.33	0.24	0.56	40	0.329 (0.049)	0.28	0.49	30	0.243 (0.045)*
M4	120	0.34	0.29	0.65	15	0.426 (0.070)	0.18	0.57	15	0.352 (0.050)
M5	480	−0.10	−0.20	0.53	40	0.295 (0.037)	−0.15	0.42	20	0.180 (0.024)*
M5	120	−0.10	−0.22	0.72	15	0.528 (0.056)	−0.05	0.56	10	0.317 (0.035)*
M6	480	−0.28	−0.45	0.57	50	0.349 (0.036)	−0.36	0.53	20	0.288 (0.029)
M6	120	−0.28	−0.35	0.70	15	0.501 (0.060)	−0.12	0.54	10	0.321 (0.034)*
M7	480	0.18	0.06	0.52	55	0.285 (0.042)	0.17	0.45	25	0.201 (0.025)*
M7	120	0.19	−0.01	0.68	20	0.497 (0.071)	0.09	0.58	15	0.342 (0.062)*
M8	480	1.45	1.33	0.44	30	0.208 (0.026)	1.38	0.44	15	0.200 (0.023)
M8	120	1.43	1.25	0.57	10	0.360 (0.039)	1.24	0.47	8	0.263 (0.030)*

E_1, SD_1 and MSE_1 denotes the average, standard deviation and the mean square error of the statistic $\log N\hat{\sigma}_N^2$ in the 1000 replications using MBJ method. E_2, SD_2 and MSE_2 are the corresponding statistics using M^2BJ method.

Table 3. MBJ and M^2BJ to estimate σ_N^2 in the case of the variance.

Model	N	$\log\sigma_N^2$	E_1	SD_1	L_1	MSE_1	E_2	SD_2	L_2	MSE_2
M1	480	4.22	3.91	0.33	20	0.209 (0.016)	3.94	0.34	30	0.197 (0.017)
M1	120	4.11	3.31	0.55	9	0.941 (0.050)	3.43	0.59	15	0.810 (0.052)
M2	480	3.98	3.73	0.26	15	0.134 (0.010)	3.71	0.28	25	0.151 (0.011)
M2	120	3.98	3.38	0.46	9	0.576 (0.035)	3.35	0.47	15	0.616 (0.040)
M3	480	3.04	2.86	0.19	10	0.069 (0.005)	2.86	0.20	15	0.072 (0.006)
M3	120	3.02	2.58	0.36	7	0.327 (0.021)	2.59	0.37	10	0.327 (0.026)
M4	480	4.25	3.95	0.36	20	0.217 (0.016)	3.94	0.36	30	0.220 (0.018)
M4	120	4.25	3.51	0.58	10	0.883 (0.049)	3.50	0.56	15	0.891 (0.050)
M5	480	2.07	1.94	0.16	6	0.043 (0.003)	1.93	0.16	7	0.047 (0.003)
M5	120	2.08	1.82	0.29	4	0.149 (0.011)	1.79	0.31	5	0.178 (0.013)
M6	480	2.39	2.24	0.20	7	0.059 (0.005)	2.22	0.21	15	0.073 (0.006)
M6	120	2.40	2.10	0.34	5	0.205 (0.018)	2.05	0.37	8	0.255 (0.020)
M7	480	3.47	3.28	0.28	20	0.112 (0.009)	3.26	0.27	20	0.116 (0.009)
M7	120	3.48	3.04	0.47	9	0.423 (0.030)	2.97	0.43	7	0.447 (0.030)
M8	480	2.08	1.95	0.17	7	0.046 (0.004)	1.95	0.18	9	0.049 (0.004)
M8	120	2.06	1.83	0.33	4	0.162 (0.011)	1.82	0.33	5	0.168 (0.012)

E_1, SD_1 and MSE_1 denotes the average, standard deviation and the mean square error of the statistic $\log N\hat{\sigma}_N^2$ in the 1000 replications using MBJ method. E_2, SD_2 and MSE_2 are the corresponding statistics using M^2BJ method.

Table 4. MBJ and M^2BJ to estimate σ_N^2 in the case of the autocovariance of order 1. (*) denotes cases where the M^2BJ outperforms the MBJ.

Model	N	$\log \sigma_N^2$	E_1	SD_1	L_1	MSE_1	E_2	SD_2	L_2	MSE_2
M1	480	4.17	3.85	0.34	20	0.216 (0.018)	3.85	0.32	20	0.205 (0.016)
M1	120	4.04	3.24	0.58	9	0.960 (0.056)	3.31	0.53	10	0.808 (0.050)
M2	480	3.69	3.45	0.26	15	0.124 (0.010)	3.43	0.28	25	0.147 (0.011)
M2	120	3.67	3.08	0.46	8	0.553 (0.034)	3.06	0.47	15	0.579 (0.040)
M3	480	2.66	2.51	0.21	9	0.067 (0.006)	2.49	0.19	10	0.066 (0.005)
M3	120	2.62	2.21	0.38	5	0.317 (0.022)	2.22	0.39	8	0.317 (0.025)
M4	480	4.20	3.91	0.38	25	0.232 (0.018)	3.90	0.37	30	0.229 (0.018)
M4	120	4.19	3.43	0.58	10	0.914 (0.052)	3.38	0.52	10	0.927 (0.049)
M5	480	1.52	1.40	0.19	6	0.054 (0.004)	1.40	0.14	2	0.035 (0.003)*
M5	120	1.51	1.25	0.29	2	0.150 (0.012)	1.29	0.28	2	0.124 (0.011)
M6	480	2.05	1.91	0.21	9	0.064 (0.006)	1.88	0.21	10	0.074 (0.006)
M6	120	2.04	1.69	0.36	3	0.250 (0.020)	1.70	0.41	8	0.290 (0.025)
M7	480	3.34	3.13	0.29	20	0.125 (0.011)	3.12	0.28	20	0.127 (0.010)
M7	120	3.33	2.85	0.52	9	0.501 (0.035)	2.83	0.49	7	0.484 (0.035)
M8	480	1.52	1.39	0.18	4	0.050 (0.004)	1.46	0.17	3	0.032 (0.003)*
M8	120	1.48	1.21	0.32	2	0.173 (0.014)	1.26	0.29	2	0.135 (0.012)*

E_1, SD_1 and MSE_1 denotes the average, standard deviation and the mean square error of the statistic $\log N\hat{\sigma}_N^2$ in the 1000 replications using MBJ method. E_2, SD_2 and MSE_2 are the corresponding statistics using M^2BJ method.

moving blocks bootstrap (MBB) and the moving blocks bootstrap with missing values replacement (M^2BB). We consider the following autoregressive models $X_t = \sum_{i=1}^{p} \phi_i X_{t-i} + e_t$:

- (M1) AR(1) $\phi_1 = 0.8$, e_t i.i.d. $\mathcal{N}(0,1)$.
- (M2) AR(2) $\phi_1 = 1.372$, $\phi_2 = -0.677$, e_t i.i.d. $\mathcal{N}(0, 0.4982)$.
- (M3) AR(5) $\phi_1 = 0.9$, $\phi_2 = -0.4$, $\phi_3 = 0.3$, $\phi_4 = -0.5$, $\phi_5 = 0.3$, e_t i.i.d. $\mathcal{N}(0,1)$.
- (M4) AR(1) $\phi_1 = -0.8$, e_t i.i.d. $\mathcal{N}(0,1)$.

Models M1–M3 are the same as in Bühlmann (1996) and Bühlmann and Künsch (1994). In all of them the largest root is around 0.8. M4 is included because it presents a considerable amount of repulsion, and Carlstein *et al.* (1998) show that this feature causes problems for the matching block bootstrap. The models M2–M4 exhibit a "damped-periodic" autocorrelation function, where the correlations can be negative. In M1 all the autocorrelations are positive. We also consider the following "dual" moving average models $X_t = e_t + \sum_{i=1}^{q} \theta_i e_{t-i}$:

- (M5) MA(1) $\theta_1 = -0.8$, e_t i.i.d. $\mathcal{N}(0,1)$.
- (M6) MA(2) $\theta_1 = -1.372$, $\theta_2 = 0.677$, e_t i.i.d. $\mathcal{N}(0, 0.4909)$.
- (M7) MA(5) $\theta_1 = -0.9$, $\theta_2 = 0.4$, $\theta_3 = -0.3$, $\theta_4 = 0.5$, $\theta_5 = -0.3$, e_t i.i.d. $\mathcal{N}(0,1)$.
- (M8) MA(1) $\theta_1 = 0.8$, e_t i.i.d. $\mathcal{N}(0,1)$.

For M^2BJ and M^2BB, we use an autoregressive estimator for the autocovariance matrix Σ, choosing the order p of the approximating autoregressive process by minimizing the BIC (cf. Schwarz (1978)) in a range $0 \le p \le 10 \log_{10} n$. As in Bühlmann and Künsch (1994) we choose the sample size $N = 480$ and $N = 120$. Our results are based on

Table 5. MBJ and M^2BJ to estimate σ_N^2 in the case of the autocovariance of order 5. (*) denotes cases where the M^2BJ outperforms the MBJ.

Model	N	$\log \sigma_N^2$	E_1	SD_1	L_1	MSE_1	E_2	SD_2	L_2	MSE_2
M1	480	3.79	3.39	0.40	25	0.319 (0.022)	3.52	0.33	15	0.176 (0.014)*
M1	120	3.55	2.72	0.59	7	1.041 (0.055)	2.97	0.49	9	0.578 (0.042)*
M2	480	3.50	3.22	0.29	20	0.157 (0.012)	3.30	0.24	15	0.097 (0.008)*
M2	120	3.43	2.78	0.48	8	0.651 (0.039)	2.91	0.44	10	0.469 (0.031)*
M3	480	2.39	2.20	0.20	10	0.078 (0.006)	2.42	0.17	5	0.029 (0.003)*
M3	120	2.30	1.92	0.38	6	0.278 (0.022)	2.22	0.37	6	0.144 (0.016)*
M4	480	3.84	3.46	0.41	20	0.311 (0.019)	3.56	0.34	15	0.198 (0.013)*
M4	120	3.79	2.92	0.58	7	1.109 (0.053)	3.18	0.55	10	0.684 (0.042)*
M5	480	1.37	1.25	0.14	4	0.031 (0.003)	1.46	0.12	2	0.023 (0.002)*
M5	120	1.33	1.14	0.29	3	0.121 (0.009)	1.33	0.27	2	0.072 (0.007)*
M6	480	1.69	1.55	0.18	6	0.050 (0.004)	1.77	0.15	4	0.029 (0.003)*
M6	120	1.65	1.40	0.35	5	0.185 (0.016)	1.61	0.32	4	0.102 (0.010)*
M7	480	2.78	2.54	0.27	10	0.133 (0.009)	2.74	0.22	7	0.048 (0.005)*
M7	120	2.73	2.28	0.44	6	0.399 (0.029)	2.53	0.39	6	0.193 (0.021)*
M8	480	1.36	1.26	0.15	5	0.033 (0.003)	1.45	0.13	2	0.025 (0.002)*
M8	120	1.28	1.12	0.31	3	0.122 (0.011)	1.27	0.27	2	0.073 (0.008)*

E_1, SD_1 and MSE_1 denotes the average, standard deviation and the mean square error of the statistic $\log N\hat{\sigma}_N^2$ in the 1000 replications using MBJ method. E_2, SD_2 and MSE_2 are the corresponding statistics using M^2BJ method.

1000 simulations, and block size range from $l = 1$ to $l = 95$ for $N = 480$, and from $l = 1$ to $l = 30$ for $N = 120$. The statistics T_N included in the simulation study are the sample mean, median, variance, and autocovariance of order 1 and 5. Notice that in the case of h-th autocovariance, a block size l corresponds to l blocks of size h in MBJ, and $l + h - 1$ missing observations in M^2BJ. We report the estimates for the variance of these statistics and, as recommended in Carlstein *et al.* (1998), we measure the accuracy using the mean square error (MSE) of the logarithm of the variance.

The simulations have been carried out as follows. First, for each model Mi ($i = 1, \ldots, 8$), $N_T = 1000$ replications have been generated. In each replication the value of the statistic T_N is computed and the "true" value of the variance of this statistic is calculated by

$$\sigma_N^2 = N \frac{\sum_1^{N_T} (T_N^{(i)} - \overline{T}_N)^2}{N_T}$$

where $\overline{T}_N = \sum_1^{N_T} T_N^{(i)}/N_T$. The log of this value, $\log \sigma_N^2$, is reported in all the tables for each model and sample size N.

Second, in the jackknife simulations (Tables 1 to 5) an estimate of the variance is computed by the following steps: (1) For each model Mi ($i = 1, \ldots, 8$) generate a sample of size N; (2) Select the length l, build the $N - l + 1$ jackknifed series and compute in each series the value of the statistic T_N; (3) Compute the estimated jackknife variance by (2.2) and (2.4), multiply by the sample size N and call them $\hat{\sigma}_N^2$ and $\tilde{\sigma}_N^2$ respectively; and (4) Repeat the steps (1) to (3) 1000 times for each possible value l. The statistics given are E_1, SD_1, the average and standard deviation of the statistic $\log N\hat{\sigma}_N^2$ in the 1000

Table 6. MBB and M^{2}BB to estimate σ_N^2 in the case of the sample mean. (∗) denotes cases where the M^{2}BB outperforms the MBB.

Model	N	$\log\sigma_N^2$	E_1	SD_1	L_1	MSE_1	E_2	SD_2	L_2	MSE_2
M1	480	3.21	2.93	0.24	30	0.135 (0.010)	3.00	0.19	5 15	0.079 (0.006)∗
M1	120	3.18	2.56	0.36	15	0.521 (0.031)	2.76	0.31	5 9	0.275 (0.018)∗
M2	480	1.68	1.60	0.11	2	0.018 (0.001)	1.70	0.10	1 1	0.011 (0.001)∗
M2	120	1.70	1.83	0.20	3	0.055 (0.005)	1.65	0.19	1 1	0.040 (0.004)∗
M3	480	1.82	1.66	0.16	15	0.052 (0.004)	1.87	0.11	3 5	0.014 (0.001)∗
M3	120	1.83	1.51	0.18	4	0.130 (0.007)	1.86	0.17	2 3	0.028 (0.003)∗
M4	480	−1.16	−1.08	0.16	25	0.033 (0.003)	−1.09	0.16	1 15	0.029 (0.003)
M4	120	−1.13	−1.10	0.24	15	0.057 (0.005)	−1.03	0.20	1 8	0.049 (0.004)
M5	480	−3.14	−2.84	0.17	60	0.117 (0.006)	−3.12	0.22	1 55	0.050 (0.005)∗
M5	120	−2.93	−2.57	0.19	30	0.167 (0.008)	−2.90	0.24	2 30	0.056 (0.007)∗
M6	480	−3.05	−2.96	0.22	60	0.056 (0.005)	−2.96	0.11	1 10	0.020 (0.002)∗
M6	120	−2.92	−2.85	0.21	30	0.049 (0.006)	−2.94	0.22	1 15	0.047 (0.004)
M7	480	−1.81	−1.74	0.20	40	0.045 (0.004)	−1.76	0.18	1 25	0.035 (0.004)∗
M7	120	−1.76	−1.79	0.27	25	0.072 (0.007)	−1.72	0.26	1 15	0.068 (0.007)
M8	480	1.18	1.09	0.13	9	0.025 (0.002)	1.17	0.08	1 1	0.006 (0.001)∗
M8	120	1.17	0.98	0.17	4	0.064 (0.005)	1.15	0.12	1 1	0.015 (0.002)∗

E_1, SD_1 and MSE_1 denotes the average, standard deviation and the mean square error of the statistic $\log N\widehat{\sigma}_N^2$ in the 1000 replications using MBB method. E_2, SD_2 and MSE_2 are the corresponding statistics using M^{2}BB method.

replications, and $MSE_1 = (\log\sigma_N^2 - E(\log\widehat{\sigma}_N^2))^2 + SD(\log\widehat{\sigma}_N^2)^2$, the mean squared error. The value of l given in L_1 is the block size producing the minimum MSE. The values E_2, SD_2, MSE_2, L_2 have the same interpretation and are computed for the proposed method based on $\widetilde{\sigma}_N^2$. The results with the relative mean square error $RMSE = MSE(\widehat{\sigma}_N^2)/\sigma_N^4$ are similar and therefore are omitted from the tables.

Third, in the bootstrap simulations (Tables 6 to 10) the estimate of the variance of the statistic is computed as follows: (1) For each model Mi ($i = 1,\ldots,8$) generate a sample of size N; (2) Choose the block length l (l and k in M^{2}BB), build $B = 250$ bootstrap samples by randomly selecting blocks with replacement among the blocks of observations and compute in each bootstrap sample the value of the statistic T_N; (3) Obtain the estimated bootstrap variance by (2.8) and (2.10), multiply by the sample size N and call them $\widehat{\sigma}_N^2$ and $\widetilde{\sigma}_N^2$, respectively; and (4) Repeat the steps (1) to (3) 1000 times. The values reported in the tables have the same interpretation as the jackknife ones. The only difference is that for the method M^{2}BB in the column corresponding to L_2, we also report the value of k, the optimal length of the missing value block (k takes values in $\{1,\ldots,5\}$). Note that the MBB is equivalent to M^{2}BB with $k = 0$.

Due to the large number of simulations, we find a significant difference between the two methods in almost all cases. However, we are interested in large differences, e.g. $MSE(\widehat{\sigma}_N^2)/MSE(\widetilde{\sigma}_N^2) > 1.25$, i.e., gain of at least a 25 percent. Additionally, we could use a smaller number of simulations, as in Bühlmann and Künsch (1994) and Bühlmann (1997); in such a case, the results are similar to those of the previous approach.

Our main conclusions for jackknife methods are as follows: (a) In the cases where

Table 7. MBB and M^2BB to estimate σ_N^2 in the case of the median. (*) denotes cases where the M^2BB outperforms the MBB.

Model	N	$\log \sigma_N^2$	E_1	SD_1	L_1	MSE_1	E_2	SD_2	L_2	MSE_2
M1	480	3.31	3.06	0.29	30	0.147 (0.013)	3.10	0.19	5 9	0.084 (0.006)*
M1	120	3.29	2.76	0.44	15	0.471 (0.035)	2.90	0.36	5 10	0.277 (0.021)*
M2	480	2.04	2.08	0.24	15	0.058 (0.007)	1.98	0.14	1 1	0.023 (0.002)*
M2	120	2.05	2.07	0.33	9	0.111 (0.013)	1.94	0.23	1 1	0.066 (0.006)*
M3	480	2.05	1.89	0.23	15	0.075 (0.008)	2.06	0.14	3 5	0.021 (0.003)*
M3	120	2.03	1.80	0.29	4	0.137 (0.013)	2.06	0.22	2 3	0.051 (0.006)*
M4	480	0.33	0.44	0.29	8	0.097 (0.009)	0.26	0.18	3 5	0.037 (0.004)*
M4	120	0.34	0.45	0.38	4	0.156 (0.016)	0.26	0.27	3 5	0.077 (0.009)*
M5	480	−0.10	−0.04	0.30	15	0.093 (0.009)	−0.22	0.16	3 5	0.038 (0.003)*
M5	120	−0.10	0.05	0.37	7	0.162 (0.016)	−0.12	0.23	3 5	0.054 (0.005)*
M6	480	−0.28	−0.23	0.30	15	0.092 (0.009)	−0.25	0.21	2 5	0.046 (0.004)*
M6	120	−0.28	−0.18	0.40	9	0.168 (0.019)	−0.37	0.27	3 5	0.084 (0.007)*
M7	480	0.18	0.20	0.29	20	0.085 (0.011)	0.14	0.18	3 5	0.033 (0.003)*
M7	120	0.19	0.27	0.40	9	0.167 (0.019)	0.15	0.24	3 5	0.061 (0.007)*
M8	480	1.45	1.36	0.21	6	0.053 (0.005)	1.39	0.12	1 1	0.017 (0.001)*
M8	120	1.43	1.27	0.28	3	0.105 (0.010)	1.39	0.16	1 1	0.028 (0.003)*

E_1, SD_1 and MSE_1 denotes the average, standard deviation and the mean square error of the statistic $\log N\hat{\sigma}_N^2$ in the 1000 replications using MBB method. E_2, SD_2 and MSE_2 are the corresponding statistics using M^2BB method.

Table 8. MBB and M^2BB to estimate σ_N^2 in the case of the variance.

Model	N	$\log \sigma_N^2$	E_1	SD_1	L_1	MSE_1	E_2	SD_2	L_2	MSE_2
M1	480	4.22	3.88	0.33	20	0.219 (0.017)	3.86	0.33	3 20	0.233 (0.018)
M1	120	4.11	3.28	0.53	8	0.970 (0.050)	3.24	0.51	3 7	1.026 (0.054)
M2	480	3.98	3.73	0.27	15	0.139 (0.011)	3.74	0.29	3 20	0.139 (0.011)
M2	120	3.98	3.34	0.47	10	0.622 (0.036)	3.38	0.47	3 10	0.581 (0.037)
M3	480	3.04	2.83	0.20	9	0.082 (0.006)	2.83	0.20	1 10	0.081 (0.006)
M3	120	3.02	2.55	0.37	8	0.353 (0.023)	2.55	0.36	1 7	0.350 (0.022)
M4	480	4.25	3.94	0.35	20	0.222 (0.017)	3.92	0.36	4 25	0.238 (0.019)
M4	120	4.25	3.50	0.56	9	0.875 (0.050)	3.46	0.56	4 7	0.948 (0.053)
M5	480	2.07	1.95	0.17	9	0.044 (0.004)	1.95	0.17	1 8	0.043 (0.004)
M5	120	2.08	1.81	0.28	3	0.152 (0.012)	1.80	0.26	1 2	0.146 (0.011)
M6	480	2.39	2.24	0.20	7	0.062 (0.005)	2.25	0.18	2 4	0.050 (0.005)
M6	120	2.40	2.09	0.36	6	0.225 (0.019)	2.10	0.33	2 4	0.199 (0.017)
M7	480	3.47	3.26	0.27	15	0.118 (0.009)	3.26	0.27	1 15	0.115 (0.009)
M7	120	3.48	3.01	0.46	8	0.438 (0.030)	2.98	0.46	1 7	0.464 (0.029)
M8	480	2.08	1.93	0.18	5	0.052 (0.004)	1.95	0.18	1 5	0.048 (0.004)
M8	120	2.06	1.80	0.31	3	0.168 (0.012)	1.79	0.30	1 4	0.163 (0.012)

E_1, SD_1 and MSE_1 denotes the average, standard deviation and the mean square error of the statistic $\log N\hat{\sigma}_N^2$ in the 1000 replications using MBB method. E_2, SD_2 and MSE_2 are the corresponding statistics using M^2BB method.

Table 9. MBB and M^2BB to estimate σ_N^2 in the case of the autocovariance of order 1. (*) denotes cases where M^2BB outperforms the MBB.

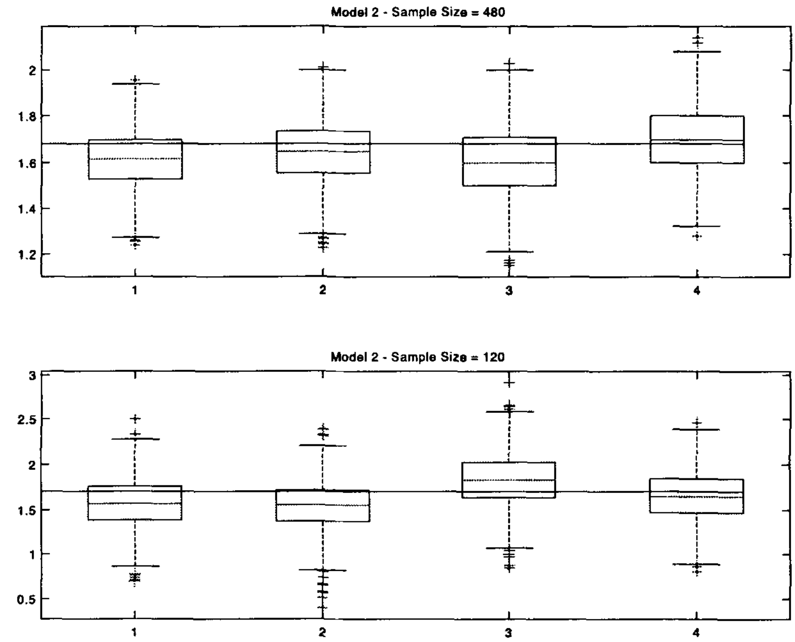
Model	N	$\log \sigma_N^2$	E_1	SD_1	L_1	MSE_1	E_2	SD_2	L_2	MSE_2
M1	480	4.17	3.79	0.35	25	0.268 (0.019)	3.80	0.32	3 20	0.245 (0.019)
M1	120	4.04	3.13	0.61	15	1.195 (0.065)	3.14	0.53	3 7	1.081 (0.059)
M2	480	3.69	3.41	0.27	20	0.153 (0.011)	3.43	0.24	3 9	0.125 (0.009)
M2	120	3.67	2.99	0.46	10	0.678 (0.036)	3.11	0.45	3 7	0.519 (0.034)*
M3	480	2.66	2.45	0.23	20	0.098 (0.007)	2.48	0.22	2 10	0.080 (0.006)
M3	120	2.62	2.11	0.39	8	0.414 (0.025)	2.18	0.37	2 5	0.336 (0.025)
M4	480	4.20	3.84	0.37	20	0.263 (0.019)	3.86	0.37	4 25	0.247 (0.019)
M4	120	4.19	3.35	0.62	15	1.081 (0.058)	3.37	0.57	4 7	1.002 (0.056)
M5	480	1.52	1.36	0.19	9	0.065 (0.005)	1.39	0.18	1 5	0.049 (0.004)*
M5	120	1.51	1.20	0.32	5	0.194 (0.013)	1.24	0.30	1 3	0.159 (0.012)
M6	480	2.05	1.88	0.24	15	0.086 (0.008)	1.90	0.19	2 4	0.057 (0.005)*
M6	120	2.04	1.65	0.40	8	0.315 (0.025)	1.73	0.35	2 3	0.217 (0.020)*
M7	480	3.34	3.09	0.30	20	0.151 (0.011)	3.10	0.28	2 15	0.133 (0.010)
M7	120	3.33	2.77	0.50	10	0.564 (0.037)	2.81	0.53	2 10	0.544 (0.037)
M8	480	1.52	1.34	0.20	8	0.073 (0.005)	1.41	0.20	1 8	0.051 (0.005)*
M8	120	1.48	1.09	0.29	3	0.231 (0.014)	1.16	0.28	1 1	0.179 (0.013)*

E_1, SD_1 and MSE_1 denotes the average, standard deviation and the mean square error of the statistic $\log N\hat{\sigma}_N^2$ in the 1000 replications using MBB method. E_2, SD_2 and MSE_2 are the corresponding statistics using M^2BB method.

Table 10. MBB and M^2BB to estimate σ_N^2 in the case of the autocovariance of order 5.

Model	N	$\log \sigma_N^2$	E_1	SD_1	L_1	MSE_1	E_2	SD_2	L_2	MSE_2
M1	480	3.79	3.30	0.39	35	0.389 (0.022)	3.31	0.37	3 25	0.366 (0.022)
M1	120	3.55	2.60	0.47	5	1.133 (0.045)	2.59	0.53	2 10	1.204 (0.054)
M2	480	3.50	3.16	0.29	30	0.201 (0.013)	3.18	0.29	4 25	0.189 (0.012)
M2	120	3.43	2.68	0.38	5	0.702 (0.032)	2.76	0.44	3 10	0.643 (0.035)
M3	480	2.39	2.15	0.18	15	0.092 (0.006)	2.19	0.20	2 15	0.081 (0.006)
M3	120	2.30	1.94	0.30	5	0.212 (0.015)	1.92	0.31	1 9	0.241 (0.016)
M4	480	3.84	3.40	0.39	30	0.351 (0.020)	3.40	0.39	2 25	0.337 (0.019)
M4	120	3.79	2.89	0.46	5	1.034 (0.045)	2.92	0.54	3 10	1.053 (0.050)
M5	480	1.37	1.29	0.14	5	0.025 (0.002)	1.29	0.13	1 6	0.022 (0.002)
M5	120	1.33	1.24	0.25	5	0.070 (0.006)	1.22	0.24	1 4	0.069 (0.006)
M6	480	1.69	1.56	0.15	5	0.038 (0.003)	1.60	0.17	2 9	0.034 (0.003)
M6	120	1.65	1.51	0.27	5	0.095 (0.011)	1.50	0.28	2 6	0.100 (0.012)
M7	480	2.78	2.52	0.24	15	0.124 (0.008)	2.54	0.26	2 15	0.124 (0.008)
M7	120	2.73	2.40	0.41	5	0.273 (0.018)	2.34	0.37	1 4	0.292 (0.019)
M8	480	1.36	1.26	0.13	5	0.027 (0.002)	1.28	0.13	1 9	0.024 (0.002)
M8	120	1.28	1.18	0.25	5	0.073 (0.007)	1.17	0.24	1 4	0.069 (0.007)

E_1, SD_1 and MSE_1 denotes the average, standard deviation and the mean square error of the statistic $\log N\hat{\sigma}_N^2$ in the 1000 replications using MBB method. E_2, SD_2 and MSE_2 are the corresponding statistics using M^2BB method.

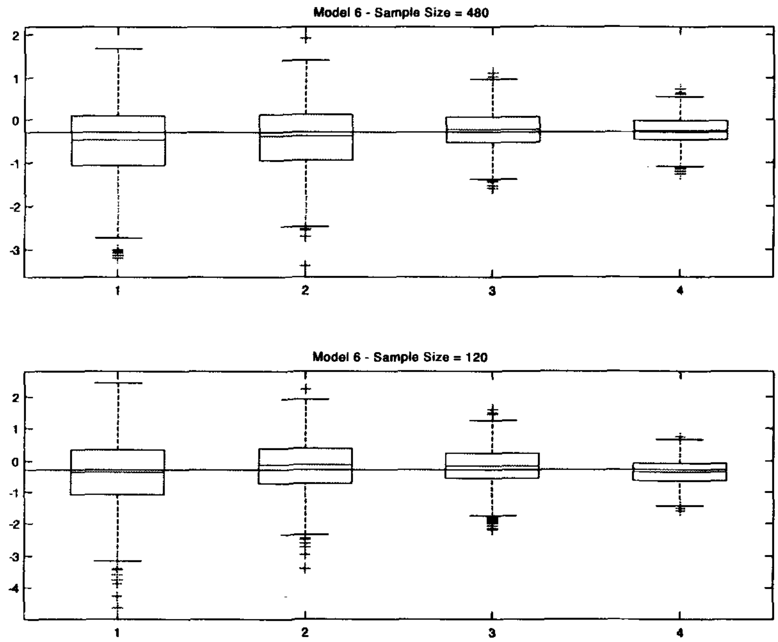


Column 1: MBJ, Column 2: M^2BJ, Column 3: MBB and Column 4: M^2BB.

Fig. 1. Moving blocks methods to estimate σ_N^2 in the case of the sample mean.

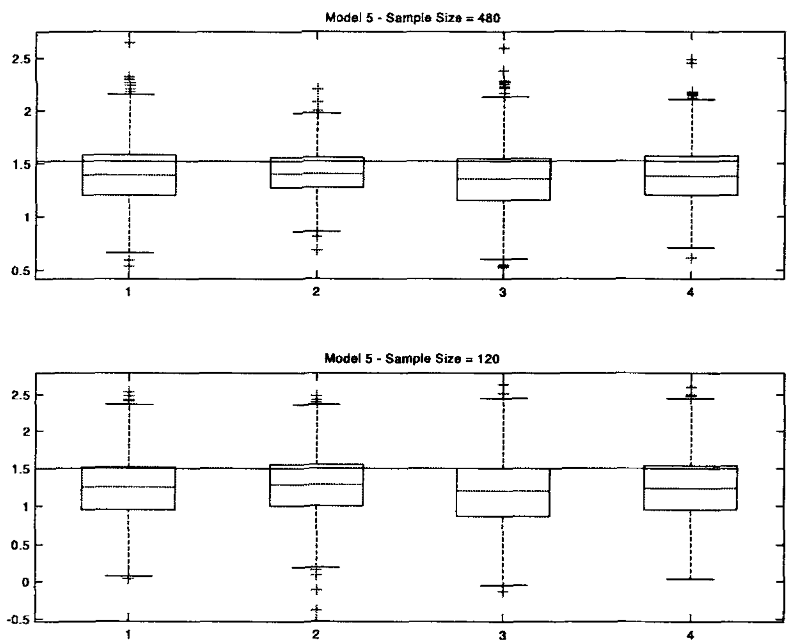
there is a substantial difference between the two methods, the missing values replacement generally gives the least MSE. In particular, only in the case of the sample mean and models M3 and M8, the MBJ have a better performance; (b) for the median, and models M1 and M5–M8, the M^2BJ outperforms the MBJ; (c) the methods are "equivalent" for the variance, but for the first autocovariance the proposed method gives a large improvement in three cases; and (d) for the autocovariance of order 5, which is the statistic that depends on the largest m-dimensional marginal distribution, in all cases and sample sizes the M^2BJ has a significantly smaller MSE than MBJ. We can conclude that the proposed method works better in general than previous procedures and that the advantages are especially great for autocovariance, especially for lags greater than 2. Other simulation studies (not shown here) have confirmed this advantage of the proposed method in estimating the variance of autocovariance for lags larger than 2. As for the optimal value of l, it is larger in MBJ than in M^2BJ. In Figs. 1–3, columns 1 and 2, we illustrate the performance of MBJ and M^2BJ methods in different scenarios. It is observed that M^2BJ tends to reduce the bias, the variability or both.

In the comparison of bootstrap methods, we observe that: (a) In the cases where there is a substantial difference between the two methods, the missing values replacement always gives least MSE; (b) for the mean, in almost all models, and for the median, in all models, the M^2BB outperforms the MBB; (c) the methods are "equivalent" in the variance and the autocovariance of order 5 (although the M^2BB outperforms the MBB when the sample size is large, 480) but for the first order autocovariance the M^2BB outperforms the MBB in all the cases and the differences are significantly larger for moving average models. In Figs. 1–3, columns 3 and 4, we illustrate the performance



Column 1: MBJ, Column 2: M^2BJ, Column 3: MBB and Column 4: M^2BB.

Fig. 2. Moving blocks methods to estimate σ_N^2 in the case of the median.



Column 1: MBJ, Column 2: M^2BJ, Column 3: MBB and Column 4: M^2BB.

Fig. 3. Moving blocks methods to estimate σ_N^2 in the case of the autocovariance of order 1.

of MBB and M^{2}BB methods in different scenarios. It is observed that M^{2}BB tends to reduce the bias, the variability or both. Similar conclusions are obtained using optimal order block length $l = n^{1/3}$ and $k = 1$ missing value in each block. The empirical selection of the number of missing values in each block could be solve by cross–validation or by the jackknife–after–bootstrap method. The jackknife–after–bootstrap appears as a computationally efficient method for estimating the bias and the variance of block bootstrap quantities as was illustrated in Lahiri (2002).

6. Conclusions

We have presented a modification of the idea of using blocks for jackknife and bootstrapping estimation in time series. In the jackknife method, instead of deleting observations, we propose to assume that these observations are missing values. For independent data both procedures are equivalent, but for correlated data they are not. It has been shown, that with this procedure better results can be obtained in the model free estimation of the variance of the autocovariance of a stationary process. The advantages are especially important for larger lags. The consistency of the estimation of the variance and distribution of the sample mean has been established.

In the block bootstrap case we propose to assume that there are missing observations between two consecutive blocks. In this way, the dependence structure among observations is better preserved and it has been shown that this procedure leads, in general, to better estimation than previous procedures, especially for large sample sizes. Consistency of the estimation of the variance and distribution of the sample mean has been proved.

One additional advantage of this approach is that we are always dealing with complete series and, therefore, the usual routines for computing statistics in a time series can be applied to the jackknife or bootstrap samples generated with the missing value approach. In particular, previous bootstrap procedures can be seen as particular cases in which the length of the missing values block is equal to zero.

Acknowledgements

We would like to thank the two referees and the coordinating editor for carefully reading and suggestions that greatly improved the paper. This research was partially supported by the Plan Nacional I+D+I project BEC 2000-0167 and the Cátedra de Calidad BBVA.

Appendix A

Proof of Lemma 4.1. Using (3.8) and (4.2), we obtain:

$$n(\widetilde{T}_N^{(j)} - T_N) = -w'_{n,j}H_j(H_j'\Sigma^{-1}H_j)^{-1}H_j'\Sigma^{-1}(X-\mu) \tag{A.1}$$

and

$$\mathrm{E}[n^2(\widetilde{T}_N^{(j)} - T_N)^2] = w'_{n,j}H_j(H_j'\Sigma^{-1}H_j)^{-1}H_j'w_{n,j}. \tag{A.2}$$

Let $\alpha_j = \{j+1, j+2, \ldots, j+l\}$. Using the formula for the inverse of a partitioned matrix,

$$(H_j'\Sigma^{-1}H_j)^{-1} = (\Sigma^{-1}(\alpha_j))^{-1} = \Sigma(\alpha_j) - \Sigma(\alpha_j, \alpha_j')\Sigma(\alpha_j')^{-1}\Sigma(\alpha_j', \alpha_j), \tag{A.3}$$

where $\Sigma(\alpha_j)$ is the principal submatrix of Σ with the elements indexed by α_j, and $\Sigma(\alpha_j, \alpha_j')$ is the result of taking the rows indicated by α_j and deleting the columns indicated by α_j. $\Sigma(\alpha_j', \alpha_j)$ and $\Sigma(\alpha_j')$ are defined analogously, cf. Horn and Johnson (1990). Note that $\Sigma^{-1}(\alpha_j)$ is a submatrix of Σ^{-1}, while $\Sigma(\alpha_j)^{-1}$ is the inverse of a submatrix of Σ.

Using (2.4), and (A.2)–(A.3), we get:

$$\begin{aligned} \mathrm{E}[n\widetilde{\sigma}^2_{Jack}] &= (n-l+1)^{-1}l^{-1}\sum_{j=0}^{n-l}\mathrm{E}[n^2(\widetilde{T}_N^{(j)} - T_N)^2] \\ &= l^{-1}w_n'\Sigma_{ll}w_n \\ &\quad - (n-l+1)^{-1}l^{-1}\sum_{j=0}^{n-l} w_n'\Sigma(\alpha_j,\alpha_j')\Sigma(\alpha_j')^{-1}\Sigma(\alpha_j',\alpha_j)w_n, \end{aligned} \tag{A.4}$$

where $w_n = (w_n(1), \ldots, w_n(l))' = H_j'w_{n,j} = 1_{l\times 1}$, and $\Sigma(\alpha_j) = \Sigma_{ll}$ is the $l \times l$ autocovariance matrix.

We now prove that $l^{-1}w_n'\Sigma_{ll}w_n \to \sigma^2_\infty$. We have that

$$\begin{aligned} l^{-1}w_n'\Sigma_{ll}w_n &= l^{-1}(l\gamma_0 + 2(l-1)\gamma_1 + \cdots + 2\gamma_{l-1}) \\ &= \sum_{m=-l+1}^{l-1}\gamma_m - 2l^{-1}\sum_{m=1}^{l-1} m\gamma_m, \end{aligned} \tag{A.5}$$

which has limit σ^2_∞ using that $l(n) \to \infty$ and $\sum_{m=1}^{+\infty} m|\gamma_m| < \infty$.

Now we prove that the second term in (A.4) goes to 0. First, note that

$$\begin{aligned} \|w_n'\Sigma(\alpha_j,\alpha_j')\|_2 &\le \|w_n'\Sigma(\alpha_j,\alpha_j')\|_1 \\ &= \sum_{m=1}^{j}\left|\sum_{i=1}^{l}\gamma_{m+i-1}\right| + \sum_{m=1}^{n-l-j}\left|\sum_{i=1}^{l}\gamma_{m+i-1}\right| \\ &\le 2\sum_{m=1}^{\max\{j,n-l-j\}}\left|\sum_{i=1}^{l}\gamma_{m+i-1}\right| \\ &\le 2\sum_{m=1}^{n-l}\sum_{i=1}^{l}|\gamma_{m+i-1}| \\ &\le 2\sum_{m=1}^{n} m|\gamma_m|. \end{aligned} \tag{A.6}$$

Second, we can write

$$\Sigma = \begin{bmatrix} A & B & C \\ B' & \Sigma(\alpha_j) & D \\ C' & D' & E \end{bmatrix},$$

then

$$\Sigma(\alpha_j') = \begin{bmatrix} A & C \\ C' & E \end{bmatrix}.$$

Define

$$\widetilde{\Sigma} = \begin{bmatrix} \Sigma(\alpha_j) & B' & D \\ B & A & C \\ D' & C' & E \end{bmatrix}.$$

Note that $\widetilde{\Sigma}$ is also symmetric and $x'\Sigma x = \widetilde{x}'\widetilde{\Sigma}\widetilde{x}$, where $x' = (x_1, x_2, \ldots, x_n)$ and $\widetilde{x}' = (x_{j+1}, \ldots, x_{j+l}, x_1, \ldots, x_j, x_{j+l+1}, \ldots, x_n)$. Then

$$\lambda_{\max}(\Sigma) = \max\left\{\frac{x'\Sigma x}{x'x} : x \neq 0\right\} = \max\left\{\frac{\widetilde{x}'\widetilde{\Sigma}\widetilde{x}}{\widetilde{x}'\widetilde{x}} : \widetilde{x} \neq 0\right\} = \lambda_{\max}(\widetilde{\Sigma})$$

and the same is true for $\lambda_{\min}(\Sigma)$. Thus, we have

$$\begin{aligned} &2\pi F_1 \le \lambda_{\min}(\widetilde{\Sigma}) \le \lambda_{\max}(\widetilde{\Sigma}) \le 2\pi F_2 \\ &(2\pi F_2)^{-1} \le \lambda_{\min}(\widetilde{\Sigma}^{-1}) \le \lambda_{\max}(\widetilde{\Sigma}^{-1}) \le (2\pi F_1)^{-1}. \end{aligned}$$

Since $\Sigma(\alpha_j')$ is a principal symmetric submatrix of $\widetilde{\Sigma}$, we have:

$$\begin{aligned} &\lambda_{\min}(\widetilde{\Sigma}) \le \lambda_{\min}(\Sigma(\alpha_j')) \quad \text{and} \quad \lambda_{\max}(\Sigma(\alpha_j')) \le \lambda_{\max}(\widetilde{\Sigma}) \\ &\lambda_{\max}(\Sigma(\alpha_j')^{-1}) \le \lambda_{\max}(\widetilde{\Sigma}^{-1}) \quad \text{and} \quad \lambda_{\min}(\widetilde{\Sigma}^{-1}) \le \lambda_{\min}(\Sigma(\alpha_j')^{-1}). \end{aligned}$$

Finally,

$$\begin{aligned} w_n'\Sigma(\alpha_j, \alpha_j')\Sigma(\alpha_j')^{-1}\Sigma(\alpha_j', \alpha_j)w_n &\le \|\Sigma(\alpha_j')^{-1}\|_{spec}\|w_n'\Sigma(\alpha_j, \alpha_j')\|_2^2 \\ &\le (2\pi F_1)^{-1}\left(2\sum_{m=1}^{n} m|\gamma_m|\right)^2, \end{aligned}$$

and thus the second term in (A.4) goes to 0 as l goes to infinity. □

PROOF OF LEMMA 4.2. Let

$$\Sigma^{-1} = \begin{bmatrix} A_1 & B_1 & C_1 \\ B_1' & \Sigma^{-1}(\alpha_j) & D_1 \\ C_1' & D_1' & E_1 \end{bmatrix};$$

then, using (3.5),

$$\begin{aligned} Z_j &= \Sigma^{-1}H_j(H_j'\Sigma^{-1}H_j)^{-1}H_j' \\ &= \begin{bmatrix} 0_{j\times j} & B_1(\Sigma^{-1}(\alpha_j))^{-1} & 0_{j\times N-l-j} \\ 0_{l\times j} & I_{l\times l} & 0_{l\times N-l-j} \\ 0_{N-l-j\times j} & D_1'(\Sigma^{-1}(\alpha_j))^{-1} & 0_{N-l-j\times N-l-j} \end{bmatrix}, \end{aligned} \tag{A.7}$$

and

$$\widetilde{w}_{n,j} = \begin{bmatrix} B_1(\Sigma^{-1}(\alpha_j))^{-1}1_{l\times 1} \\ 1_{l\times 1} \\ D_1'(\Sigma^{-1}(\alpha_j))^{-1}1_{l\times 1} \end{bmatrix}. \tag{A.8}$$

The elements in positions $j+1,\dots,j+l$ are all 1's, and the remaining elements depend on the product $\Sigma^{-1}(\alpha_j',\alpha_j)(\Sigma^{-1}(\alpha_j))^{-1}$, because $\left[\begin{smallmatrix}B_1\\D_1'\end{smallmatrix}\right] = \Sigma^{-1}(\alpha_j',\alpha_j)$. Using the expressions for the inverse of a partitioned matrix, we obtain

$$\Sigma^{-1}(\alpha_j',\alpha_j) = (\Sigma(\alpha_j',\alpha_j)\Sigma(\alpha_j)^{-1}\Sigma(\alpha_j,\alpha_j') - \Sigma(\alpha_j'))^{-1}\Sigma(\alpha_j',\alpha_j)\Sigma(\alpha_j)^{-1}$$
$$(\Sigma^{-1}(\alpha_j))^{-1} = \Sigma(\alpha_j) - \Sigma(\alpha_j,\alpha_j')\Sigma(\alpha_j')^{-1}\Sigma(\alpha_j',\alpha_j).$$

Let's denote $Q_j = (\Sigma(\alpha_j',\alpha_j)\Sigma(\alpha_j)^{-1}\Sigma(\alpha_j,\alpha_j') - \Sigma(\alpha_j'))^{-1}$; then

$$\text{(A.9)}\quad \begin{aligned}\Sigma^{-1}(\alpha_j',\alpha_j)(\Sigma^{-1}(\alpha_j))^{-1} &= Q_j\Sigma(\alpha_j',\alpha_j) - (I + Q_j\Sigma(\alpha_j'))\Sigma(\alpha_j')^{-1}\Sigma(\alpha_j',\alpha_j)\\ &= -\Sigma(\alpha_j')^{-1}\Sigma(\alpha_j',\alpha_j).\end{aligned}$$

Thus, we can concentrate our attention on $-\Sigma(\alpha_j')^{-1}\Sigma(\alpha_j',\alpha_j)$. We have that

$$\text{(A.10)}\qquad \Sigma(\alpha_j',\alpha_j) = \begin{bmatrix} \gamma_j & \gamma_{j+1} & \cdots & \gamma_{j+l-1}\\ \gamma_{j-1} & \gamma_j & \cdots & \gamma_{j+l-2}\\ \vdots & \vdots & & \vdots\\ \gamma_1 & \gamma_2 & \cdots & \gamma_l\\ \gamma_l & \gamma_{l-1} & \cdots & \gamma_1\\ \vdots & \vdots & & \vdots\\ \gamma_{n-2-j} & \gamma_{n-3-j} & \cdots & \gamma_{n-l-1-j}\\ \gamma_{n-1-j} & \gamma_{n-2-j} & \cdots & \gamma_{n-l-j} \end{bmatrix}_{n-l\times l}.$$

Let $\widetilde{\Sigma}(\alpha_j',\alpha_j)_{n-l\times l}$ be the matrix obtained by substituting 0 in every position of matrix $\Sigma(\alpha_j',\alpha_j)$ where the index m of γ_m satisfies $m > \lceil l^{1/2}\rceil$. The difference between $\Sigma(\alpha_j',\alpha_j)$ and $\widetilde{\Sigma}(\alpha_j',\alpha_j)$ satisfies

$$\text{(A.11)}\quad \begin{aligned}\|(\Sigma(\alpha_j',\alpha_j) - \widetilde{\Sigma}(\alpha_j',\alpha_j))w_n\|_2 &\le \|(\Sigma(\alpha_j',\alpha_j) - \widetilde{\Sigma}(\alpha_j',\alpha_j))w_n\|_1\\ &\le 2\sum_{m=\lceil l^{1/2}\rceil+1}^{n-1} m|\gamma_m| = o(l^{-1/2}),\end{aligned}$$

since $\sum_{m=1}^{\infty} m^2|\gamma_m| < \infty$ implies $\sum_{m=\lceil l^{1/2}\rceil+1}^{n-1} m|\gamma_m| = o(l^{-1/2})$.

On the other hand, $\Sigma(\alpha_j') = \begin{bmatrix}\Sigma_{j-1,j-1} & F\\ F' & \Sigma_{n-l-j-1,n-l-j-1}\end{bmatrix}$, where $\Sigma_{h,h}$ is the $h\times h$ autocovariance matrix. Define $\widetilde{\Sigma}(\alpha_j') = \begin{bmatrix}\Sigma_{j-1,j-1} & 0\\ 0' & \Sigma_{n-l-j-1,n-l-j-1}\end{bmatrix}$; as earlier, we have that

$$\text{(A.12)}\quad \begin{aligned}&\|\Sigma(\alpha_j')^{-1} - \widetilde{\Sigma}(\alpha_j')^{-1}\|_{spec}\\ &= \|\Sigma(\alpha_j')^{-1}(\Sigma(\alpha_j') - \widetilde{\Sigma}(\alpha_j'))\widetilde{\Sigma}(\alpha_j')^{-1}\|_{spec}\\ &\le \|\Sigma(\alpha_j')^{-1}\|_{spec}\|\Sigma(\alpha_j') - \widetilde{\Sigma}(\alpha_j')\|_{spec}\|\widetilde{\Sigma}(\alpha_j')^{-1}\|_{spec}\\ &\le (2\pi F_1)^{-2}(\|\Sigma(\alpha_j') - \widetilde{\Sigma}(\alpha_j')\|_{col}\|\Sigma(\alpha_j') - \widetilde{\Sigma}(\alpha_j')\|_{row})^{1/2}\end{aligned}$$

$$\leq (2\pi F_1)^{-2} \sum_{m=l+1}^{n-1} |\gamma_m|,$$

and $\sum_{m=1}^{\infty} m^2|\gamma_m| < \infty$ implies $\sum_{m=l+1}^{n-1} |\gamma_m| = o(l^{-2})$. Let $\Sigma^a_{h,h} = [\gamma^a]_{h\times h}$ be the $h \times h$ covariance matrix of an AR($\lceil l^{1/2}\rceil$) process such that $\Sigma^a_{\lceil l^{1/2}\rceil,\lceil l^{1/2}\rceil} = \Sigma_{\lceil l^{1/2}\rceil,\lceil l^{1/2}\rceil}$.

We can assume that $\sum_{m=1}^{\infty} m^2|\gamma^a_m| < \infty$, see Bühlmann (1995). Define $\bar\Sigma(\alpha'_j) = \begin{bmatrix} \Sigma^a_{j-1,j-1} & 0 \\ 0' & \Sigma^a_{n-l-j-1,n-l-j-1} \end{bmatrix}$, then we have the following results:

$$\begin{aligned} (\text{A.13}) \qquad &\|\widetilde\Sigma(\alpha'_j)^{-1} - \bar\Sigma(\alpha'_j)^{-1}\|_{spec} \\ &= \|\widetilde\Sigma(\alpha'_j)^{-1}(\widetilde\Sigma(\alpha'_j) - \bar\Sigma(\alpha'_j))\bar\Sigma(\alpha'_j)^{-1}\|_{spec} \\ &\leq \|\widetilde\Sigma(\alpha'_j)^{-1}\|_{spec}\|\widetilde\Sigma(\alpha'_j) - \bar\Sigma(\alpha'_j)\|_{spec}\|\bar\Sigma(\alpha'_j)^{-1}\|_{spec} \\ &\leq 2(2\pi F_1)^{-2} \sum_{m=\lceil l^{1/2}\rceil+1}^{n-1} (|\gamma_m| + |\gamma^a_m|), \end{aligned}$$

and $\sum_{m=1}^{\infty} m^2|\gamma_m| < \infty$ and $\sum_{m=1}^{\infty} m^2|\gamma^a_m| < \infty$ imply that $\sum_{m=\lceil l^{1/2}\rceil+1}^{n-1} |\gamma_m| = o(l^{-1})$ and $\sum_{m=\lceil l^{1/2}\rceil+1}^{n-1} |\gamma^a_m| = o(l^{-1})$.

Note that $\bar\Sigma(\alpha'_j)^{-1}$ is a $\lceil l^{1/2}\rceil$-diagonal matrix; then $\bar\Sigma(\alpha'_j)^{-1}\widetilde\Sigma(\alpha'_j,\alpha_j)w_n$ has at most $4\lceil l^{1/2}\rceil$ non zero elements. Define $\bar w_{n,j}$ replacing in $\widetilde w_{n,j}$ the matrices $\Sigma(\alpha'_j)^{-1}$ and $\Sigma(\alpha'_j,\alpha_j)$ with $\bar\Sigma(\alpha'_j)^{-1}$ and $\widetilde\Sigma(\alpha'_j,\alpha_j)$, then $\bar w_{n,j}$ has at most $l + 4\lceil l^{1/2}\rceil$ non zero elements.

Finally,

$$\begin{aligned} \|\widetilde w_{n,j} - \bar w_{n,j}\|_2 &= \|\Sigma(\alpha'_j)^{-1}\Sigma(\alpha'_j,\alpha_j)w_n - \bar\Sigma(\alpha'_j)^{-1}\widetilde\Sigma(\alpha'_j,\alpha_j)w_n\|_2 \\ &\leq \|\Sigma(\alpha'_j)^{-1}\Sigma(\alpha'_j,\alpha_j)w_n - \bar\Sigma(\alpha'_j)^{-1}\Sigma(\alpha'_j,\alpha_j)w_n\|_2 \\ &\quad + \|\bar\Sigma(\alpha'_j)^{-1}\Sigma(\alpha'_j,\alpha_j)w_n - \bar\Sigma(\alpha'_j)^{-1}\widetilde\Sigma(\alpha'_j,\alpha_j)w_n\|_2 \\ &= \|(\Sigma(\alpha'_j)^{-1} - \bar\Sigma(\alpha'_j)^{-1})\Sigma(\alpha'_j,\alpha_j)w_n\|_2 \\ &\quad + \|\bar\Sigma(\alpha'_j)^{-1}(\Sigma(\alpha'_j,\alpha_j) - \widetilde\Sigma(\alpha'_j,\alpha_j))w_n\|_2 \\ &\leq \|\Sigma(\alpha'_j)^{-1} - \bar\Sigma(\alpha'_j)^{-1}\|_{spec}\|\Sigma(\alpha'_j,\alpha_j)w_n\|_2 \\ &\quad + \|\bar\Sigma(\alpha'_j)^{-1}\|_{spec}\|(\Sigma(\alpha'_j,\alpha_j) - \widetilde\Sigma(\alpha'_j,\alpha_j))w_n\|_2, \end{aligned}$$

and using (A.11)-(A.13) we have that $\|\widetilde w_{n,j} - \bar w_{n,j}\|_2 = o(l^{-1/2})$. □

Proof of Corollary 4.1. We have

$$(\text{A.14}) \qquad n\widetilde\sigma^2_{Jack} = n^2\|w_n\|_2^{-2}\left((n-l+1)^{-1}\sum_{j=0}^{n-l}(\widetilde T_N^{(j)} - T_N)^2 - (\widetilde T_N^{(\cdot)} - T_N)^2\right),$$

and it is enough to prove that $\widetilde S_N = n(\widetilde T_N^{(\cdot)} - T_N) = (n-l+1)^{-1}\sum_{j=0}^{n-l}\widetilde w_{n,j}(X-\mu)$ is $o_p(1)$. It's clear that $\mathrm{E}[\widetilde S_N] = 0$, and

$$(\text{A.15}) \qquad \mathrm{E}[\widetilde S_N^2] = (n-l+1)^{-2}\sum_{j=0}^{n-l}\sum_{i=0}^{n-l}\widetilde w'_{n,j}\Sigma\widetilde w_{n,i}.$$

As in Lemma 4.2, we can concentrate our attention on

$$(n-l+1)^{-2}\sum_{j=0}^{n-l}\sum_{i=0}^{n-l}\bar{w}_{n,i}'\Sigma\bar{w}_{n,j}. \tag{A.16}$$

Replace γ_m with a 0 in each position of the matrix Σ where $m > l$. Let $\widetilde{\Sigma}$ denote the resulting matrix. Since $\|\Sigma - \widetilde{\Sigma}\|_{spec} = o(l^{-1})$, then

$$(n-l+1)^{-2}\sum_{j=0}^{n-l}\sum_{i=0}^{n-l}\bar{w}_{n,j}'\Sigma\bar{w}_{n,i} = (n-l+1)^{-2}\sum_{j=0}^{n-l}\sum_{i=0}^{n-l}\bar{w}_{n,j}'\widetilde{\Sigma}\bar{w}_{n,i} + o(1). \tag{A.17}$$

On the other hand, note that $\bar{w}_{n,j}'\widetilde{\Sigma}\bar{w}_{n,i}$ is equal to a sum of between 1 and $l+4\lceil l^{1/2}\rceil$ non zero summands, where the size of the sum depends on the different values of i and j. Then

$$\sum_{i=0}^{n-l}|\bar{w}_{n,j}'\Sigma\bar{w}_{n,i}| \le 2C(1+2+\cdots+(l+4\lceil l^{1/2}\rceil)) = O(l^2), \tag{A.18}$$

and

$$(n-l+1)^{-2}\sum_{j=0}^{n-l}\sum_{i=0}^{n-l}\bar{w}_{n,j}'\widetilde{\Sigma}\bar{w}_{n,i} = O((n-l+1)^{-1}l^2), \tag{A.19}$$

where $C = \sum_{m=-l}^{l}|\gamma_m|\max\{1,(\pi F_1)^{-1}\sum_{m=1}^{\infty}m|\gamma_m|\}$. Finally, if $l = o(n^{1/2})$, we obtain that $\widetilde{S}_N \xrightarrow{P} 0$. □

Proof of Lemma 4.3. Under these assumptions, we have that (cf. Hannan and Kavalieris (1986) and Bühlmann (1995))

$$\max_{0\le m\le p}|\widehat{\gamma}_m - \gamma_m| = O((n/\log n)^{-1/2}) \quad \text{a.s.}, \tag{A.20}$$

and there exists a random variable n_1 such that:

$$\sup_{n\ge n_1}\sum_{m=0}^{\infty}m^2|\gamma_m| < +\infty \quad \text{a.s.} \tag{A.21}$$

Thus, we have that:

$$\begin{aligned}\|\Sigma - \widehat{\Sigma}\|_{col} &\le 2\sum_{m=0}^{\infty}|\widehat{\gamma}_m - \gamma_m| \\ &= 2\left(\sum_{m=0}^{p}|\widehat{\gamma}_m - \gamma_m| + \sum_{m=p+1}^{\infty}|\widehat{\gamma}_m - \gamma_m|\right) \\ &= O((n/\log n)^{-1/2})p + o(p^{-2}) \quad \text{a.s.} \\ &= o((n/\log n)^{-1/3}) \quad \text{a.s.}\end{aligned} \tag{A.22}$$

and

$$\|\Sigma^{-1} - \widehat{\Sigma}^{-1}\|_{col} \leq \|\Sigma^{-1}\|_{col}\|\Sigma - \widehat{\Sigma}\|_{col}\|\widehat{\Sigma}^{-1}\|_{col} = o((n/\log n)^{-1/3}) \quad \text{a.s.} \tag{A.23}$$

Define B_j, A_j, $\widehat{B}_j$ and $\widehat{A}_j$ by

$$B_j = A_j^2 = n^2(\widetilde{T}_N^{(j)} - T_N)^2 \tag{A.24}$$

and

$$\widehat{B}_j = \widehat{A}_j^2 = n^2(\widehat{\widetilde{T}}_N^{(j)} - T_N)^2. \tag{A.25}$$

Note that $|B_j - \widehat{B}_j| = |A_j - \widehat{A}_j||A_j + \widehat{A}_j|$. Next, we find a bound that does not depend on j.

Using vector $\bar{w}_{n,j}$ defined in Lemma 2, we have

$$|A_j| = |\widetilde{w}_{n,j}'(X - \bar{X})| \leq |\bar{w}_{n,j}'(X - \bar{X})| + |(\widetilde{w}_{n,j} - \bar{w}_{n,j})'(X - \bar{X})| = O_P(l^{1/2}) + O_P(l^{-1/2}). \tag{A.26}$$

For $|\widehat{A}_j|$, we can proceed in a similar way. Note that $\widehat{\Sigma}$ is an autoregressive estimator and then $\widehat{\Sigma}^{-1}$ is a $(2p+1)$-diagonal matrix. Therefore, $\widehat{\widetilde{w}}_{n,j}$ has at most $O(\max(l,p))$ non zero elements.

Lets denote $\alpha(p)$ the indexes of non zero elements in $\widehat{w}_{n,j}$ and define a vector $\widetilde{w}_{n,j}^{(p)}$ such that $\widetilde{w}_{n,j,\alpha(p)}^{(p)} = \widetilde{w}_{n,j}^{(p)}$ and $\widetilde{w}_{n,j,\alpha(p)'}^{(p)} = 0$. By definition, $\widetilde{w}_{n,j}^{(p)}$ has at most $O(\max(l,p))$ non zero elements and $\|\widetilde{w}_{n,j} - \widehat{\widetilde{w}}_{n,j}\|_1 = \|\widetilde{w}_{n,j}^{(p)} - \widehat{\widetilde{w}}_{n,j}\|_1 + \|\widetilde{w}_{n,j} - \widetilde{w}_{n,j}^{(p)}\|_1$.

Now,

$$|A_j - \widehat{A}_j| = |(\widetilde{w}_{n,j} - \widehat{\widetilde{w}}_{n,j})'(X - \bar{X})| \leq |(\widetilde{w}_{n,j}^{(p)} - \widehat{\widetilde{w}}_{n,j})'(X - \bar{X})| + |(\widetilde{w}_{n,j} - \widetilde{w}_{n,j}^{(p)})'(X - \bar{X})|, \tag{A.27}$$

$$\begin{aligned}\|\widetilde{w}_{n,j} - \widehat{\widetilde{w}}_{n,j}\|_1 &= \|\Sigma^{-1}H_j(H_j'\Sigma^{-1}H_j)^{-1}H_j'w_{n,j} - \widehat{\Sigma}^{-1}H_j(H_j'\widehat{\Sigma}^{-1}H_j)^{-1}H_j'w_{n,j}\|_1 \\ &\leq l(\|\Sigma^{-1} - \widehat{\Sigma}^{-1}\|_{col}\|(\Sigma^{-1}(\alpha_j))^{-1}\|_{col} \\ &\quad + \|\widehat{\Sigma}^{-1}\|_{col}\|(\Sigma^{-1}(\alpha_j))^{-1} - (\widehat{\Sigma}^{-1}(\alpha_j))^{-1}\|_{col})\end{aligned} \tag{A.28}$$

and

$$\begin{aligned}\|\Sigma^{-1}(\alpha_j)^{-1} - \widehat{\Sigma}^{-1}(\alpha_j)^{-1}\|_{col} &\leq \|\Sigma(\alpha_j) - \widehat{\Sigma}(\alpha_j)\|_{col} \\ &\quad + \|\Sigma(\alpha_j,\alpha_j')\Sigma(\alpha_j')^{-1}\Sigma(\alpha_j',\alpha_j) \\ &\quad - \widehat{\Sigma}(\alpha_j,\alpha_j')\widehat{\Sigma}(\alpha_j')^{-1}\widehat{\Sigma}(\alpha_j',\alpha_j)\|_{col} \\ &\leq O(l^{1/2})\|\Sigma - \widehat{\Sigma}\|_{col} \\ &= o_{a.s.}(l^{1/2}(n/\log n)^{-1/3}).\end{aligned} \tag{A.29}$$

Then,

$$\|\widetilde{w}_{n,j} - \widehat{\widetilde{w}}_{n,j}\|_1 = o_{a.s.}(l^{3/2}(n/\log n)^{-1/3}), \tag{A.30}$$

$$|A_j - \widehat{A}_j| = o_{a.s.}(l^{3/2}(n/\log n)^{-1/3})O_P(\max(l,p)^{1/2}), \tag{A.31}$$

$$|B_j - \widehat{B}_j| = (O_P(l^{1/2}) + O_P(\max(l,p)^{1/2})) \tag{A.32}$$

$$\times\, o_p(\max(l,p)^{1/2} l^{3/2} (n/\log n)^{-1/3})$$
$$= o_P(\max(l,p)^{1/2} l^2 (n/\log n)^{-1/3})$$
$$+\, o_P(\max(l,p) l^{3/2} (n/\log n)^{-1/3}),$$

and

$$|n\widetilde{\sigma}^2_{Jack} - n\widehat{\widetilde{\sigma}}^2_{Jack}| = o_P(\max(l,p)^{1/2} l (n/\log n)^{-1/3}) + o_P(\max(l,p) l^{1/2} (n/\log n)^{-1/3}). \tag{A.33}$$

To finish the proof, we only need to consider the two possible cases $l \le p$ and $l > p$. In the first case, we need that $l = O((n/\log n)^{1/4})$ which is trivially satisfied since $l = O(p)$, and in the second case we need that $l = O((n/\log n)^{2/9})$. Just rest to impose that $l = O((n/\log n)^{2/9})$. □

Proof of Theorem 4.2. We extensively use the relation between H_N and L_N and the symmetry of $J(\rho)$, i.e. $J(x,\rho) = 1 - J(-x,\rho)$ and the following result from Politis and Romano (1994*a*):

Theorem. (Politis and Romano (1994*a*)) *Assume that there exists a limiting law $J(\rho)$ such that $J_N(\rho)$ converges weakly to a limit law $J(\rho)$, as $n \to \infty$, and that $\tau_b/\tau_n \to 0$, $b/n \to 0$ and $b \to \infty$ as $n \to \infty$. Also assume that the α-mixing sequence satisfies that $\alpha_X(m) \to 0$ as $m \to \infty$.*

1. *If x is a continuity point of $J(\cdot,\rho)$, then $L_N(x) \to J(x,\rho)$ in probability.*
2. *If $J(\cdot,\rho)$ is continuous, then* $\sup_x |L_N(x) - J(x,\rho)| \to 0$ *in probability.*

By symmetry, if x is a continuity point of $J(\cdot,\rho)$, then $-x$ is also a continuity point. Then, using statement (1) of the theorem, we have

$$H_N(x) = 1 - L_N(-x) \to 1 - J(-x,\rho) = J(x,\rho) \text{ in probability.} \tag{A.34}$$

Using statement (2) of the theorem, we obtain convergence to 0 in probability, since

$$\begin{aligned} \sup_x |H_N(x) - J(x,\rho)| &= \sup_x |1 - L_N(-x) - (1 - J(-x,\rho))| \\ &= \sup_x |L_N(-x) - J(-x,\rho)|. \end{aligned} \tag{A.35}$$

Notice that if $\tau_l = \sqrt{l}$, using the fact that $l/n \to 0$, then the coefficient $\tau_l l^{-1}(n-l)$ is close to $\sqrt{n(n-l)l^{-1}}$ which is the standardizing constant of Wu (1990). □

Proof of Proposition 4.2. We have that

$$\begin{aligned} & l^{-1/2} \sum_{t=1}^{n} w_n(t-j)(\widehat{X}_{t,j} - T_N) \\ &= l^{-1/2} \sum_{t=1}^{n} w_n(t-j)(\widehat{X}_{t,j} - \mu) - l^{1/2}(T_N - \mu) \\ &= l^{-1/2} \sum_{t=1}^{n} w_n(t-j)(\widehat{X}_{t,j} - \mu) + O_P(l^{1/2} n^{-1/2}) \\ &= l^{-1/2} w'_{n,j}(\widehat{X}_j - \mu) + O_P(l^{1/2} n^{-1/2}), \end{aligned} \tag{A.36}$$

where $w'_{n,j} = (w_n(1-j), \ldots, w_n(n-j))$ and $\widehat{X}'_j = (\widehat{X}_{1,j}, \ldots, \widehat{X}_{n,j})$. Now,

$$\text{(A.37)} \quad \begin{aligned} l^{-1/2} w'_{n,j}(\widehat{X}_j - \mu) &= l^{-1/2} w'_{n,j}(I - H_j(H'_j \Sigma^{-1} H_j)^{-1} H'_j \Sigma^{-1})(X - \mu) \\ &= l^{-1/2}(w_{n,j} - \widetilde{w}_{n,j})'(X - \mu). \end{aligned}$$

From the proof of Lemma 4.2 we have that

$$\text{(A.38)} \quad \widetilde{w}_{n,j} = \Sigma^{-1} H_j (H'_j \Sigma^{-1} H_j)^{-1} H'_j w_{n,j} = \begin{bmatrix} B(\Sigma^{-1}(\alpha_j))^{-1} 1_{l\times 1} \\ 1_{l\times 1} \\ D'_1(\Sigma^{-1}(\alpha_j))^{-1} 1_{l\times 1} \end{bmatrix}$$

and

$$\text{(A.39)} \quad \begin{aligned} \|\Sigma^{-1}(\alpha'_j, \alpha_j)(\Sigma^{-1}(\alpha_j))^{-1}\|_1 &= \| - \Sigma(\alpha'_j)^{-1} \Sigma(\alpha'_j, \alpha_j)\|_1 \\ &\le 4M \sum_{m=1}^{\infty} m|\gamma_m|. \end{aligned}$$

Notice that bound (A.39) does not depend on j. Then,

$$\text{(A.40)} \quad \begin{aligned} \|w_{n,j} - \widetilde{w}_{n,j}\|_1 &= \left\| \begin{bmatrix} B(\Sigma^{-1}(\alpha_j))^{-1} 1_{l\times 1} \\ D'_1(\Sigma^{-1}(\alpha_j))^{-1} 1_{l\times 1} \end{bmatrix} \right\|_1 \\ &\le \|\Sigma^{-1}(\alpha'_j, \alpha_j)(\Sigma^{-1}(\alpha_j))^{-1}\|_1, \end{aligned}$$

and for some $0 < \varepsilon < 1/2$,

$$\text{(A.41)} \quad l^{-1/2}(w_{n,j} - \widetilde{w}_{n,j})'(X - \mu) = o_P(l^{-1/2+\varepsilon}).$$

Therefore,

$$\text{(A.42)} \quad l^{-1/2} \sum_{t=1}^{n} w_n(t-j)(\widehat{X}_{t,j} - T_N) = O_P(l^{1/2} n^{-1/2}) + o_P(l^{-1/2+\varepsilon}).$$

Then, only rest to use that $l = o(n)$. □

PROOF OF PROPOSITION 4.3. Assuming that $n = s(l+k)$, we have that

$$\text{(A.43)} \quad \begin{aligned} &(s(l+k))^{-1/2} \sum_{i=1}^{s} \sum_{j=l+1}^{l+k} (\widehat{X}^*_{(i-1)(l+k)+j} - \bar{X}_n) \\ &= (s(l+k))^{-1/2} W'(I_{s(l+k)} - H(H'\Sigma^{-1}H)^{-1} H'\Sigma^{-1})(X^* - \bar{X}), \end{aligned}$$

where $I_{s(l+k)}$ is the $s(l+k) \times s(l+k)$ identity matrix, $X^* = (X^*_1, \ldots, X^*_{s(l+k)})'$, $\bar{X} = \bar{X}_n 1_{n\times 1}$, and W is a $s(l+k) \times 1$ vector defined as

$$W' = (\underbrace{0, \ldots, 0}_{l\ times}, \underbrace{1, \ldots, 1}_{k\ times}, \ldots\ldots, \underbrace{0, \ldots, 0}_{l\ times}, \underbrace{1, \ldots, 1}_{k\ times}),$$

i.e., W indicates the missing observations positions.

Analogously to M^{2}BJ, the matrix $H(H'\Sigma^{-1}H)^{-1}H'\Sigma^{-1}$ have submatrices equal to the $k \times k$ identity matrix in the missing observations positions, and the remaining non-zero elements are elements of $-\Sigma(\alpha')^{-1}\Sigma(\alpha', \alpha)$, where $\alpha = (l+1, \ldots, l+k, 2l+k+1, \ldots, 2(l+k), \ldots, \ldots, sl+(s-1)k+1, \ldots, s(l+k))$. Therefore,

$$\begin{aligned} &W'(I - H(H'\Sigma^{-1}H)^{-1}H'\Sigma^{-1}) \\ &= (a_1, \ldots, a_l, \underbrace{0, \ldots, 0}_{k\ times}, \ldots, a_{(s-1)(l+k)+1}, \ldots\ldots, a_{(s-1)(l+k)+l}, \underbrace{0, \ldots, 0}_{k\ times}), \end{aligned} \tag{A.44}$$

where the a's are 0 or are the sum of one column of $-\Sigma(\alpha')^{-1}\Sigma(\alpha', \alpha)$, and they satisfy $\sum_{t=1}^{s(l+k)} |a_t| \leq 4M \sum_{m=1}^{\infty} m|\gamma_m|$. Then,

$$\mathrm{E}^*\left[(s(l+k))^{-1/2} \sum_{t=1}^{s(l+k)} a_t(X_t^* - \bar{X}_n)\right] = 0, \tag{A.45}$$

and

$$\begin{aligned} &\mathrm{E}^*\left[\left((s(l+k))^{-1/2} \sum_{t=1}^{s(l+k)} a_t(X_t^* - \bar{X}_n)\right)^2\right] \\ &= (s(l+k))^{-1} \sum_{t=1}^{s(l+k)} \sum_{r=1}^{s(l+k)} a_t a_r \mathrm{E}^*[(X_t^* - \bar{X}_n)(X_r^* - \bar{X}_n)] \\ &\leq (s(l+k))^{-1} \sum_{t=1}^{s(l+k)} \sum_{r=1}^{s(l+k)} |a_t a_r| \mathrm{E}^*[(X_1^* - \bar{X}_n)^2] \\ &= (s(l+k))^{-1} O(1) O_{a.s.}(1) = o_{a.s.}((s(l+k))^{-1+\varepsilon}), \end{aligned} \tag{A.46}$$

for some $0 < \varepsilon < 1$.

Finally, (A.45) and (A.46) imply that $(s(l+k))^{-1/2} \sum_{t=1}^{s(l+k)} a_t(X_t^* - \bar{X}_n) = o_P(1)$ for almost all sample sequences $X_1, \ldots, X_N$. □

PROOF OF LEMMA 4.4. We have,

$$\begin{aligned} &(s(l+k))^{1/2}(\tilde{\bar{X}}^*_{n,s} - \bar{X}_n) \\ &= (s(l+k))^{-1/2} \left(\sum_{i=1}^{s} \sum_{j=1}^{l} (X^*_{(i-1)(l+k)+j} - \bar{X}_n) \right. \\ &\qquad \left. + \sum_{i=1}^{s} \sum_{j=l+1}^{l+k} (\widehat{X}^*_{(i-1)(l+k)+j} - \bar{X}_n) \right), \end{aligned} \tag{A.47}$$

and

$$\begin{aligned} &(s(l+k))^{1/2}(\widehat{\bar{X}}^*_{n,s} - \bar{X}_n) \\ &= (s(l+k))^{-1/2} \left(\sum_{i=1}^{s} \sum_{j=1}^{l} (X^*_{(i-1)(l+k)+j} - \bar{X}_n) \right. \end{aligned} \tag{A.48}$$

$$+\sum_{i=1}^{s}\sum_{j=l+1}^{l+k}(\widehat{\bar{X}}^*_{(i-1)(l+k)+j}-\bar{X}_n)\Bigg).$$

Then,

$$(s(l+k))^{1/2}(\tilde{\bar{X}}^*_{n,s}-\bar{X}_n)-(s(l+k))^{1/2}(\widehat{\bar{X}}^*_{n,s}-\bar{X}_n)$$
$$=(s(l+k))^{-1/2}W'(H(H'\widehat{\Sigma}^{-1}H)^{-1}H'\widehat{\Sigma}^{-1}-H(H'\Sigma^{-1}H)^{-1}H'\Sigma^{-1})(X^*-\bar{X}), \quad (A.49)$$

$$\mathrm{E}^*[(s(l+k))^{1/2}(\tilde{\bar{X}}^*_{n,s}-\bar{X}_n)-(s(l+k))^{1/2}(\widehat{\bar{X}}^*_{n,s}-\bar{X}_n)]$$
$$=\mathrm{E}^*\left[(s(l+k))^{-1/2}\sum_{t=1}^{s(l+k)}(\widehat{a}_t-a_t)(X^*_t-\bar{X}_n)\right]$$
$$=(s(l+k))^{-1/2}\sum_{t=1}^{s(l+k)}(\widehat{a}_t-a_t)\mathrm{E}^*[X^*_t-\bar{X}_n]=0, \quad (A.50)$$

and

$$\mathrm{E}^*[((s(l+k))^{1/2}(\tilde{\bar{X}}^*_{n,s}-\bar{X}_n)-(s(l+k))^{1/2}(\widehat{\bar{X}}^*_{n,s}-\bar{X}_n))^2]$$
$$=(s(l+k))^{-1/2}\sum_{t=1}^{s(l+k)}\sum_{r=1}^{s(l+k)}(\widehat{a}_t-a_t)(\widehat{a}_r-a_r)\times\mathrm{E}^*[(X^*_t-\bar{X}_n)(X^*_r-\bar{X}_n)]$$
$$\le(s(l+k))^{-1}\sum_{t=1}^{s(l+k)}\sum_{r=1}^{s(l+k)}|(\widehat{a}_t-a_t)(\widehat{a}_r-a_r)|\mathrm{E}^*[(X^*_1-\bar{X}_n)^2]$$
$$=(s(l+k))^{-1}o_{a.s.}((sk)^3(s(l+k)\log(s(l+k)))^{-2/3})O_{a.s.}(1)$$
$$=o_{a.s.}(1). \quad (A.51)$$

Finally, (A.50) and (A.51) imply (4.24) for almost all sample sequences $X_1,\ldots,X_N$. □

References

Abraham, B. and Thavaneswaran, A. (1991). A nonlinear time series model and estimation of missing observations, *Ann. Inst. Statist. Math.*, **43**, 493–504.

Berk, K. N. (1974). Consistent autoregressive spectral estimates, *Ann. Statist.*, **2**, 489–502.

Beveridge, S. (1992). Least squares estimation of missing values in time series, *Comm. Statist. Theory Methods*, **21**, 3479–3496.

Bose, A. (1990). Bootstrap in moving average models, *Ann. Inst. Statist. Math.*, **42**, 753–768.

Bosq, D. (1996). *Nonparametric Statistics for Stochastic Processes: Estimation and Prediction*, Springer, New York.

Bühlmann, P. (1995). Moving-average representation of autoregressive approximations, *Stochastic Process. Appl.*, **60**, 331–342.

Bühlmann, P. (1996). Locally adaptive lag-window spectral estimation, *J. Time Ser. Anal.*, **17**, 247–270.

Bühlmann, P. (1997). Sieve bootstrap for time series, *Bernoulli*, **8**, 123–148.

Bühlmann, P. and Künsch, H. R. (1994). Block length selection in the bootstrap for time series, R.R. No. 72, Eidgenössische Technische Hochschule.

Carlstein, E. (1986). The use of subseries values for estimating the variance of a general statistics from a stationary sequence, *Ann. Statist.*, **14**, 1171–1194.

Carlstein, E., Do, K., Hall, P., Hesterberg, T. and Künsch, H. R. (1998). Matched-block bootstrap for dependent data, *Bernoulli*, **4**, 305–328.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife, *Ann. Statist.*, **7**, 1–26.

Efron, B. and Tibshirani, R. J. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy, *Statist. Sci.*, **1**, 54–77.

Freedman, D. A. (1984). On bootstrapping two-stage least-square estimates in stationary linear models, *Ann. Statist.*, **12**, 827–842.

Galbraith, R. F. and Galbraith, J. I. (1974). On the inverse of some patterned matrices arising in the theory of stationary time series, *J. Appl. Probab.*, **11**, 63–71.

Hannan, E. J. and Kavalieris, L. (1986). Regression, autoregression models, *J. Time Ser. Anal.*, **7**, 27–49.

Harvey, A. C. and Pierse, R. G. (1984). Estimating missing observations in economic time series, *J. Amer. Statist. Assoc.*, **79**, 125–131.

Horn, R. A. and Johnson, C. R. (1990). *Matrix Analysis*, Cambridge University Press, Cambridge.

Kreiss, J. P. and Franke, J. (1992). Bootstrapping stationary autoregressive moving average models, *J. Time Ser. Anal.*, **13**, 297–317.

Künsch, H. R. (1989). The jackknife and the bootstrap for general stationary observations, *Ann. Statist.*, **17**, 1217–1241.

Lahiri, S. N. (2002). On the jackknife-after-bootstrap method for dependent data and its consistency properties, *Econometric Theory*, **18**, 79–98.

Liu, R. Y. and Singh, K. (1992). Moving blocks jackknife and bootstrap capture weak dependence, *Exploring the Limits of Bootstrap* (eds. R. Lepage and L. Billard), 225–248, Wiley, New York.

Peña, D. (1990). Influential observations in time series, *J. Bus. Econom. Statist.*, **8**, 235–241.

Peña, D. and Maravall, A. (1991). Interpolation, outliers, and inverse autocorrelations, *Comm. Statist. Theory Methods*, **20**, 3175–3186.

Politis, D. N. and Romano, J. F. (1992). A circular block-resampling procedure for stationary data, *Exploring the Limits of Bootstrap* (eds. R. Lepage and L. Billard), 263–270, Wiley, New York.

Politis, D. N. and Romano, J. F. (1994*a*). Large sample confidence regions based on subsamples under minimal assumptions, *Ann. Statist.*, **22**, 2031–2050.

Politis, D. N. and Romano, J. F. (1994*b*). The stationary bootstrap, *J. Amer. Statist. Assoc.*, **89**, 1303–1313.

Politis, D. N. and Romano, J. F. (1995). Bias-corrected nonparametric spectral estimation, *J. Time Ser. Anal.*, **16**, 67–103.

Quenouille, M. (1949). Approximation test of correlation in time series, *J. Roy. Statist. Soc. Ser. B*, **11**, 18–84.

Schwarz, G. (1978). Estimating the dimension of a model, *Ann. Statist.*, **6**, 461–464.

Shao, Q.-M. and Yu, H. (1993). Bootstrapping the sample means for stationary mixing sequences, *Stochastic Process. Appl.*, **48**, 175–190.

Sherman, M. (1998). Efficiency and robustness in subsampling for dependent data, *J. Statist. Plann. Inference*, **75**, 133–146.

Tukey, J. (1958). Bias and confidence in not quite large samples, *Ann. Math. Statist.*, **29**, p. 614.

White, H. (1984). *Asymptotic Theory for Econometricians*, Academic Press, New York.

Wu, C. F. J. (1990). On the asymptotic properties of the jackknife histogram, *Ann. Statist.*, **18**, 1438–1452.