

Estimation on unevenly spaced time series

Liudas Giraitis^{1*} and Fulvia Marotta²

¹Queen Mary University of London, ²University of Oxford

May 18, 2023

Abstract

In many different fields realizations of stationary time series might be recorded at irregular points in time, resulting in observed unevenly spaced samples. These missing observations can happen for several reasons, depending on the mechanisms that record the data or external conditions that force the missing observations. In this paper, we first focus on the question if we can estimate the mean of a stationary time series when data are not equally spaced. We show that any unevenly spaced sample can be used to estimate the mean of an underlying stationary linear time series. Specifically, we do not impose any restrictions on sampling structure and times, as long as they are independent of the underlying time series. We provide an expression for the sample mean estimator and we establish its asymptotic properties and the central limit theorem. Subsequently we studentize estimation which allows to build confidence intervals for the mean. Finite sample properties of the estimator for the mean are investigated in a Monte Carlo study which confirms good performance of such estimation procedure.

Keywords: time series, missing data, sample mean.

JEL: C12, C22

MOS subject classification: 37M10

*Correspondence to: Liudas Giraitis, Queen Mary University of London. Email: L.Giraitis@qmul.ac.uk

1 Introduction

Time series data are usually recorded at regularly spaced time intervals. However, for a variety of reasons in many different fields data can be observed only at irregularly spaced times and some observations are missing. The first strand of the literature on modelling such series recasts them as a sample from a partially observed stationary time series, in which the value of the observed process is set to zero when the observation is missing, see Jones (1962), Parzen (1963) and Dunsmuir and Robinson (1981a, 1981b, 1981b). It assumes that the missing data indicator follows some specific deterministic or stochastic pattern, and estimation and inference is conducted under assumption that the observed process is *asymptotically stationary*.

Jones (1962) and Parzen (1963) have studied spectral estimation of the partially observed process. In particular, Parzen (1963), in his work on spectral analysis for time series with missing observations, was focusing on estimating the autocovariances of the underlying process when missing data pattern is cyclical and deterministic. Such sampling structure has a potential application when dealing with daily financial data, in which the weekend data are absent. For other examples, see e.g. Datta and Wu (2012).

Dunsmuir and Robinson (1981a, 1981b, 1981b) have established asymptotic properties of estimators of autocovariances and outlined the estimation method for fitting a parametric model to the underlying stationary time series in the presence of a random missing data pattern, when the missing data indicator is an asymptotically (second-order) stationary sequence. Their estimation methodology involves the use of the spectral density and covariances function for increasing number of lags, which results in overly strong assumptions on the pattern of missing data.

The second strand of the literature uses the framework of subordinated processes introduced by Clark (1973). In this framework, processes are observed in the operational time scale and subsequently transformed, using directing processes, into (unobserved) processes in the standard calendar time. An example of application of this model in intra-daily financial markets can be found in Ghysels (1997).

To our knowledge, only Brillinger (1973) has considered estimation of the mean of a stationary time series using unevenly spaced data. He derived non standard central limit theorems for the sample mean estimator based on several general assumptions. The first assumption is on the summability of the cumulants of the underlying time series. The second one imposes a specific generating mechanism which ensures that the sampling times satisfy an assumption of asymptotic stationarity and are independent of an underlying stationary time series.

In this paper we focus solely on the estimation of the mean μ of an underlying stationary time series $\{x_t\}$ from unevenly spaced data. The main novelty of this paper threefold. Firstly, we allow for a very general setting of sampling times covering most of applications. We show that as sample size increases, the sample mean of such data is asymptotically normality distributed as long as the sampling times are independent of an underlying stationary process $\{x_t\}$.

Secondly, we assume that the underlying series $\{x_t\}$ is a linear weakly dependent process with $Ex_t^2 < \infty$, which allows us to derive the asymptotic normality result for the sample mean. Recall that Brillinger (1973) imposes assumptions on cumulants and

requires existence of all moments of $\{x_t\}$.

Thirdly, we introduce a simple to use studentization procedure, that allows to compute confidence intervals for the mean. It requires no additional assumption on x_t and sampling times besides the existence of $Ex_t^4 < \infty$. Monte Carlo study confirms good performance of the studentized sample mean estimation procedure for various missing data patterns.

The paper is structured as follows. Section 2 outlines the setting for estimation of the mean and presents asymptotic results for building confidence intervals. Section 3 reports Monte Carlo simulation results that verify our theoretical results and validate their use in finite samples. Section 4 contains the proofs.

Throughout this paper, we denote by \rightarrow_p and \rightarrow_D the convergence in probability and distribution, respectively, while C denotes generic constants.

2 Estimation of the mean

In this section we show that, under minimal assumptions, estimation of the mean of a stationary linear time series can be conducted using unevenly spaced samples as long as the missing data pattern does not depend on an unobserved stationary time series. In Section 2.1, we provide estimates for standard errors for our estimate which allow to build confidence intervals for the mean.

Let $\{x_t\}$ be a stationary linear process with mean $Ex_t = \mu$ defined as

$$x_t = \mu + \sum_{k=0}^{\infty} a_k \varepsilon_{t-k}, \quad t \in \mathbb{Z}, \quad (1)$$

where $\{\varepsilon_t\}$ is a sequence of independent identically distributed (i.i.d.) random variables with zero mean, variance $E\varepsilon_1^2 = \sigma_\varepsilon^2$ and $a_k, k \geq 0$ are real numbers such that $\sum_{k=0}^{\infty} a_k^2 < \infty$. Suppose that over time period $[1, \dots, N]$ we collect a sample z_1, \dots, z_n with $n \leq N$ from $\{x_t\}$ of unevenly spaced data which can be written as

$$z_j = x_{k_j}, \quad j = 1, \dots, n, \quad \text{where} \quad k_1 < k_2 < \dots < k_n. \quad (2)$$

We will represent it as a partially observed process

$$y_j = h_j x_j, \quad j = 1, \dots, N,$$

where the weights h_j take value 0 or 1. In (2) they are defined as

$$h_j = \begin{cases} 1 & \text{for } j = k_1, k_2, \dots, k_n, \quad \text{where } k_1 < k_2 < \dots < k_n \leq N, \\ 0 & \text{otherwise.} \end{cases}$$

Notice that

$$n = \sum_{j=1}^N h_j.$$

We are interested in estimation and inference on $\{x_t\}$ using unevenly spaced data without imposing additional assumptions on sampling times $k_1 < k_2 < \dots < k_n$ or missing data indicator function h_t except that they are not dependent on the underlying process $\{x_t\}$. We do not make assumptions on the type of missing data pattern, i.e. whether it is systematical and random nor properties of $\{h_t\}$. Besides assumption that $\{h_t\}$ is independent of $\{x_t\}$, we only assume, that with the expansion of time period $[1, \dots, N]$, the sample size $n \rightarrow_p \infty$ increases. Contrary to the literature, e.g. Brillinger (1973) or Datta and Wu (2012), we do not impose any other assumptions on location times of missing data, or $\{h_t\}$. Such general setting does not allow using spectral domain approach, that is popular in the literature on missing data. We show, that, after some simple adjustments, our time domain estimation procedure allows standard basic estimation and inference on the mean of $\{x_t\}$ using unevenly spaced data.

We start with the estimation of the mean μ . Denote by

$$\bar{z}_n = \frac{1}{n} \sum_{t=1}^n z_t$$

the sample mean estimate based on a sample z_1, \dots, z_n . We are interested in establishing asymptotic properties of the sample mean estimator \bar{z}_n of μ based on unevenly spaced data, in particular, consistency,

$$\bar{z}_n \rightarrow_p \mu,$$

and asymptotic normality,

$$\frac{\bar{z}_n - \mu}{(\text{var}(\bar{z}_n | h))^{1/2}} \rightarrow_D \mathcal{N}(0, 1). \quad (3)$$

Furthermore, if (3) holds, we wish to estimate the standard error $(\text{var}(\bar{z}_n | h))^{1/2}$ from the data z_1, \dots, z_n , which is needed to conduct inference about μ .

We need the following additional assumptions on $\{x_t\}$. We assume that a_k 's in (1) are summable:

$$\sum_{k=0}^{\infty} |a_k| < \infty. \quad (4)$$

The linear process (1) has the autocovariance function:

$$\gamma_X(k) = \text{var}(\varepsilon_1) \sum_{j=0}^{\infty} a_j a_{j+k}, \quad k = 0, 1, 2, \dots$$

Under (4), the autocovariance function is summable:

$$\sum_{k \in \mathbb{Z}} |\gamma_X(k)| = G_0 < \infty. \quad (5)$$

The linear process $\{x_t\}$, that satisfies (5), has bounded spectral density:

$$\begin{aligned} f_X(u) &= \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} e^{iku} \gamma_X(k), \quad u \in \Pi = [-\pi, \pi], \\ f_X(u) &\leq \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} |\gamma_X(k)| \leq \frac{G_0}{2\pi}. \end{aligned} \quad (6)$$

In addition, we assume that the spectral density f_X is bounded from below: i.e. there exist a constant $c > 0$ such that

$$0 < c \leq f_X(u) \quad \text{for all } u \in [-\pi, \pi]. \quad (7)$$

These conditions are sufficient to establish the following results on the estimation of the mean. Since the missing data pattern indicator h_t is allowed to be a random process, we use notation $E[\dots|h]$, $\text{var}(\dots|h)$ to denote the conditional mean and variance of \bar{z}_n given $\{h_t\}$.

Define the sum

$$W_N = \sum_{j=1}^N h_j x_j.$$

Notice that

$$W_N = z_1 + \dots + z_n \quad \text{and} \quad \bar{z}_n = \frac{W_N}{n}.$$

Since

$$\text{var}(\bar{z}_n | h) = \frac{\text{var}(W_N | h)}{n^2},$$

the results for the sum W_N in Theorem 1 imply the results for the sample mean \bar{z}_n in Corollary 2.

In this paper we assume that sample size $n \rightarrow_p \infty$ increases as $N \rightarrow \infty$.

Theorem 1 *Let $\{x_j\}$ be as in (1), assumptions (5) and (7) hold and $\{h_t\}$ be independent of $\{x_t\}$. Then, there exist constants $c > 0$, $C > 0$ that do not depend on n and $\{h_t\}$ such that*

$$cn \leq \text{var}(W_N | h) \leq Cn. \quad (8)$$

Moreover, as $N \rightarrow \infty$,

$$\frac{W_N - n\mu}{\sqrt{\text{var}(W_N | h)}} \rightarrow_D N(0, 1), \quad (9)$$

where

$$\text{var}(W_N | h) = \sum_{t,s=1}^N h_t h_s \gamma_X(t-s).$$

Corollary 2 *Suppose that assumptions of Theorem 1 are satisfied. Then as $n \rightarrow \infty$, there exist constants $c > 0$, $C > 0$ such that*

$$cn^{-1} \leq \text{var}(\bar{z}_n | h) \leq Cn^{-1}. \quad (10)$$

Moreover, as $N \rightarrow \infty$,

$$\frac{\bar{z}_n - \mu}{\sqrt{\text{var}(\bar{z}_n | h)}} = \frac{n}{\sqrt{\text{var}(W_N | h)}} (\bar{z}_n - \mu) \rightarrow_D N(0, 1). \quad (11)$$

The main novelty of our paper is showing that asymptotic normality (11) and its studentized version see Theorem 6 below, are valid for any sampling times h_t as long as they are independent of $\{x_t\}$.

Remark 3 Implementation of (11) for building confidence intervals for μ requires estimation of the standard error $\sqrt{\text{var}(\bar{z}_n | h)}$ established in Theorem 6. Notice that under our assumptions on h_t the asymptotics

$$\text{var}(\bar{z}_n | h) = bn(1 + o_p(1)) \quad \text{for some } b > 0 \quad (12)$$

may not exist.

Remark 4 The literature on missing data often imposes overly restrictive assumptions on h_t , in particular, that $\{h_t\}$ is an asymptotically stationary process, i.e. the limits, $n/N \rightarrow \alpha$ and $n^{-1} \sum_{t=k+1}^N h_t h_{t-k} \rightarrow \kappa(k)$, both exist. Brillinger (1973) used this type of setting to establish the asymptotic normality for the sample mean \bar{z}_n of unevenly spaced observations sampled from a continuous time stationary process. Datta and Du (2012) showed that the sample mean estimator \bar{z}_n is \sqrt{n} -consistent and the variance satisfies (12) with $b = \sum_{k=-\infty}^{\infty} \kappa(k) \gamma_X(k)$, and derived a HAC type estimator for b . Our Theorem 6 shows that the finite sample variance $\text{var}(\bar{z}_n | h)$ can be estimated from data for any $\{h_t\}$ that is independent of $\{x_t\}$ which is a valuable result from the application point of view.

2.1 Estimation of the standard errors

Building confidence intervals for the mean μ using normal approximation (11) requires estimation of the standard error $\sqrt{\text{var}(\bar{z}_n | h)} = n^{-1} \sqrt{\text{var}(W_N | h)}$.

Next we proceed to estimation of $s_{X,N}^2 = \text{var}(W_N | h)$. We have

$$\begin{aligned} \text{var}(W_N | h) &= \sum_{t,k=1}^N h_t h_k \gamma_X(t-k) \\ &= \sum_{t=k=1}^N h_t^2 \gamma_X(0) + 2 \sum_{1 \leq k < t \leq N} h_t h_k \gamma(t-k) \\ &= \gamma_X(0) \sum_{t=1}^N h_t^2 + 2 \sum_{v=1}^{N-1} \left[\sum_{t=v+1}^N h_t h_{t-v} \right] \gamma_X(v). \end{aligned}$$

We will write

$$s_{X,N}^2 := \text{var}(W_N | h) = \tau_{N,0} + 2 \sum_{v=1}^{N-1} \tau_{N,v}, \quad (13)$$

where

$$\tau_{N,v} = \left(\sum_{t=v+1}^N h_t h_{t-v} \right) \gamma_X(v).$$

We will estimate $\tau_{N,\nu}$ using the estimator

$$\tilde{\tau}_{N,\nu} = \sum_{t=\nu+1}^N (y_t - \bar{z}_n) (y_{t-\nu} - \bar{z}_n) h_t h_{t-\nu},$$

where $\bar{z}_n = n^{-1} \sum_{t=1}^n z_t$. The next proposition derives the following property of $\tilde{\tau}_{N,\nu}$.

Proposition 5 *Suppose that assumptions of Theorem 1 are satisfied and $E\varepsilon_1^4 < \infty$. Let $\nu \in \{0, \dots, N-1\}$ be a fixed integer. Then,*

$$E[|\tilde{\tau}_{N,\nu} - \tau_{N,\nu}| |h] \leq Cn^{1/2}, \quad (14)$$

where C does not depend on h_t 's, n , N and ν .

To evaluate $s_{X,N}^2$, we will use the estimator

$$\tilde{s}_{X,N,q}^2 = \tilde{\tau}_{N,0} + 2 \sum_{\nu=1}^q \tilde{\tau}_{N,\nu}, \quad (15)$$

where $q = q_n \rightarrow \infty$, $q = o(n)$ is the bandwidth parameter.

Theorem 6 *Let assumptions of Theorem 1 be satisfied and $E\varepsilon_1^4 < \infty$. Assume that $q \rightarrow \infty$ and $q = o(n)$. Then*

$$\tilde{s}_{X,N,q}^2 = s_{X,N}^2(1 + o_p(1)). \quad (16)$$

Moreover, as $N \rightarrow \infty$,

$$\frac{n}{\sqrt{\tilde{s}_{X,N,q}^2}} (\bar{z}_n - \mu) \rightarrow_D N(0, 1). \quad (17)$$

Notice that Theorem 6 focuses on estimation of the finite sample variance $\text{var}(\bar{z}_n | h)$ given in (13) and does not require existence of the limit $n^{-1} s_{X,N,q}^2$. Monte Carlo study shows that under missing data the variance $\text{var}(\bar{z}_n | h)$ can be estimated using smaller values of the bandwidth q than in stationary case in [2], e.g. the bandwidth $q = \sqrt{n}$ performs well in simulations. Observe, that missing data may complicate data based selection of q .

Next we introduce a special class of h_t 's for which the limit $n^{-1} s_{X,N}^2 \rightarrow_p b$ exists with $b = v_X^2$, where v_X^2 is the long-run variance of $\{x_t\}$. We believe this result is of independent interest although we do not use it in the estimation of the mean. We will say that the sample z_1, \dots, z_n contains M data clusters, if h_t , $t = 1, \dots, N$ indicates presence of M data blocks with no missing data. For example, $M = 3$ for $\{h_t\} = \{1, 0, 1, 1, 0, 0, 0, 0, 1, 1\}$, and $M = 2$ for $\{h_t\} = \{1, 1, 0, 0, 0, 0, 1, 1, 1\}$.

Theorem 7 *Let assumptions of Theorem 1 be satisfied and $E\varepsilon_1^4 < \infty$. Suppose that the number of data clusters M in the sample z_1, \dots, z_n has property $M = o_p(\sqrt{n})$. Then, as $N \rightarrow \infty$,*

$$n^{-1} s_{X,N,q}^2 \rightarrow_p v_X^2 = \sum_{k \in \mathbb{Z}} \gamma_X(k). \quad (18)$$

Moreover, as $N \rightarrow \infty$,

$$\frac{\sqrt{n}}{v_X}(\bar{z}_n - \mu) \rightarrow_D N(0, 1). \quad (19)$$

Theorem 7 shows that if the number of data clusters is low, the asymptotic behavior of the sample mean is the same as in the case of evenly spaced data.

3 Monte Carlo simulations

In this section we conduct simulations to examine finite sample properties of the sample mean estimator \bar{z}_n based on unevenly spaced sample z_1, \dots, z_n from a stationary time series $\{x_j\}$. We obtain z_1, \dots, z_n by generating a sample y_1, \dots, y_N of size $N = 1000$ of partially observed process

$$y_j = h_j x_j,$$

where $\{h_j\}$ is a missing data indicator and $\{x_j\}$ is a stationary zero mean time series. In our setting $\{h_j\}$ and $\{x_j\}$ are mutually independent, $\bar{z}_n = n^{-1} \sum_{j=1}^N y_j$ and $n = \sum_{j=1}^N h_j$.

We focus on the following models for $\{x_j\}$:

AR(1)model : $x_t = \phi x_{t-1} + \varepsilon_t$ with parameters $\phi = -0.6, -0.3, 0, 0.5, 0.9$,

AR(2)model : $x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \varepsilon_t$ with parameters
 $(\phi_1, \phi_2) = (-0.6, 0.1), (-0.3, 0.6), (0.1, 0.4), (0.1, 0.8),$

MA(1)model : $x_t = \varepsilon_t + \theta \varepsilon_{t-1}$ with parameters $\theta = -0.5, -0.2, 0.4, 0.7$,

where $\varepsilon_j \sim \text{i.i.d. } N(0, 1)$.

The missing patterns used in the exercise are described by series of six indicator functions $h_{\ell,j}$. We include a benchmark of no missing data, 3 deterministic missing patterns and 2 stochastic missing patterns defined as follows.

- 1 No missing data: $h_{0,j} = 1$ for all $j = 1, \dots, N$.
- 2 Cyclical missing pattern: for every cycle of m observations we observe $\alpha < m$ observations, and $m - \alpha$ observations are missing, i.e. $h_1 = \dots = h_\alpha = 1$ and $h_{\alpha+1} = \dots = h_m = 0$.

In our simulations, we include a deterministic cyclical pattern $h_{1,j}$ with $m = 7$ and $\alpha = 5$, and a cyclical pattern $h_{2,j}$ with $m = 10$ and $\alpha = 5$.

- 3 Finite number of missing data blocks: $h_{3,j} = 0$ for $j \in [100, \dots, 150], [200, \dots, 250], [300, \dots, 450], [600 - 750], [800 - 825]$ and $[925 - 950]$.
- 4 Stochastic missing pattern: h_j is a stochastic process defined as follows:

$$h_{4,j} = I(|\eta_j| \leq 0.5), \eta_j \sim \text{i.i.d. } N(0, 1),$$

$$h_{5,j} = I(\eta_j \leq 0.1 + 0.5(j/N)), j = 1, \dots, N, \eta_j \sim \text{i.i.d. } N(0, 1),$$

where $\{\eta_j\}$ is independent of $\{\varepsilon_j\}$. The first example of stochastic missing pattern $h_{4,j}$ with the cutoff of $|\eta_j|$ set at 0.5 implies that more than half of the sample is missing. The second example $h_{5,j}$ instead illustrates the case when the sample has a lower fraction of missing data as time increases, implying more missing data in the early part of the sample.

Table 1 summarizes the six types of missing data patterns described above. Figure 1 shows examples of plots of $y_j = h_j x_j$, $j = 1, \dots, 500$ with the systematic missing pattern $h_{1,j}$, large missing blocks $h_{3,j}$ and the stochastic missing pattern $h_{4,j}$. The panel in gray depicts values of the indicator $h_j = 1$ (white when $h_j = 0$).

Table 2 reports the bias and the standard deviation $\text{sd}(\bar{z}_n)$ of the sample mean estimator \bar{z}_n based on 1000 replications for six different missing data patterns h_0, \dots, h_5 , when $\{x_t\}$ is a stationary AR(1) time series. The actual number of observations n is reported (for stochastic h_j the average actual number of observations over the 1000 replications). Bias is very small for all the missing patterns and for different values of the autoregressive parameter ϕ . Standard deviation increases in the case of $\phi = 0.9$ when persistence is high and ϕ close to the boundary of stationarity.

Table 3 contains the Monte Carlo results which permit to assess the precision of the normal approximation (17) for

$$t_n := \frac{n(\bar{z}_n - \mu)}{\sqrt{\hat{s}_{X,N,q}^2}}.$$

To evaluate the precision of the estimation of the finite sample variance $s_{X,N}^2$ by $\hat{s}_{X,N,q}^2$, the table reports their the average ratio and the relative MSE of $\hat{s}_{X,N,q}^2$:

$$E\left(\frac{\hat{s}_{X,N,q}^2}{s_{X,N}^2}\right), \quad \text{RMSE} = E\left[\left(\frac{\hat{s}_{X,N,q}^2}{s_{X,N}^2} - 1\right)^2\right].$$

Two values of the bandwidth, $q = \sqrt{n}$ and $q = n^{2/3}$, are considered.

For $q = \sqrt{n}$, the empirical and nominal coverage probabilities for the mean μ are relatively close, which confirms validity of the normal approximation (17). The estimator $\hat{s}_{X,N,q}^2$ of $s_{X,N,q}^2$ appears to perform also well since the RMSE is low and the average ratio stays very close to one. When q increases to $q = n^{2/3}$ the coverage probabilities start deteriorating.

Table 4 report the results on coverage probabilities for the mean and estimation of the variance $s_{X,N}^2$ for $\{x_j\}$ following AR(2) and MA(1) models. In these simulations $\hat{s}_{X,N,q}^2$ is estimated using $q = \sqrt{n}$. Similarly as for AR(1) model, coverage probabilities are close to the nominal, indicating that the normal approximation is precise enough. In addition, the RMSE is low for both for AR(2) and MA(1) models and the average ratio is close to 1, confirming similar results as for AR(1) model in Table 3. The corresponding simulation results for $q = n^{2/3}$ are slightly worse, results are available upon request.

Finally, Table 5 allows to verify whether the asymptotic property (18), $n^{-1}\hat{s}_{X,N,q}^2 \rightarrow_p v_X^2$, of Theorem 7 is valid for missing data patterns $h_{0,t}, h_{1,t}, \dots, h_{5,t}$. In the table we report

the Monte Carlo average ratio

$$r_N = E\left(\frac{\hat{s}_{X,N,q}^2}{nv_X^2}\right).$$

Simulations show that for h_0 (no missing data) and h_3 (when large blocks of data are missing) the ratio r_N is very close to 1 which is in agreement with the theory. Other missing patterns (both deterministic and stochastic) do not satisfy assumptions of the theorem and, as expected, we observe some deviations of r_N from 1, i.e the property (18) does not hold.

| h | missing pattern |
|-------|--|
| h_0 | no missing data |
| h_1 | cyclical: 2 obs missing each 7 |
| h_2 | cyclical: 5 obs missing each 10 |
| h_3 | big blocks of missing data |
| h_4 | $h_j = I(\eta_j \leq 0.5), j = 1, \dots, N, \eta_j \sim \text{i.i.d. } N(0, 1)$ |
| h_5 | $h_j = I(\eta_j \leq 0.1 + 0.5(j/N)), j = 1, \dots, N, \eta_j \sim \text{i.i.d. } N(0, 1)$ |

Table 1: Missing data patterns h_0, \dots, h_5 .

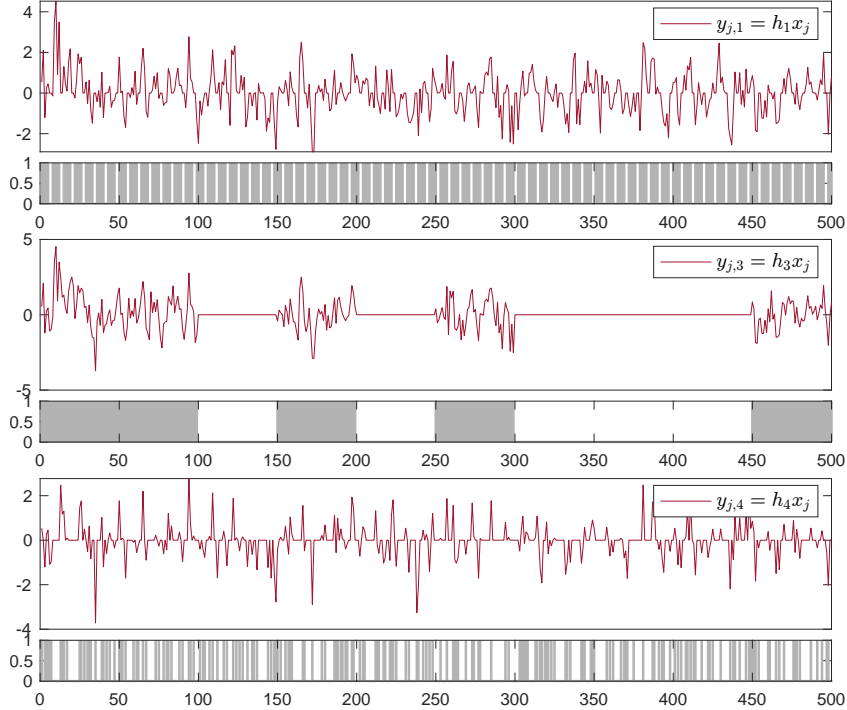


Figure 1: Plots of $y_j = h_j x_j$ with missing data patterns $h_{1,j}$, $h_{3,j}$ and $h_{4,j}$, $x_j \sim AR(1)$, $\phi = 0.5$, $N = 500$.

| n | h | ϕ | $E(\bar{z}_n) - \mu$ | $sd(\bar{z}_N)$ | n | h | ϕ | $E(\bar{z}_n) - \mu$ | $sd(\bar{z}_n)$ |
|------|-------|--------|----------------------|-----------------|-----|-------|--------|----------------------|-----------------|
| 1000 | h_0 | -0.6 | 0.00 | 0.02 | 550 | h_3 | -0.6 | 0.00 | 0.03 |
| | | -0.3 | 0.00 | 0.03 | | | -0.3 | 0.00 | 0.03 |
| | | 0 | 0.00 | 0.03 | | | 0 | 0.00 | 0.04 |
| | | 0.5 | 0.00 | 0.07 | | | 0.5 | 0.00 | 0.08 |
| | | 0.9 | -0.01 | 0.32 | | | 0.9 | -0.01 | 0.39 |
| 644 | h_1 | -0.6 | 0.00 | 0.02 | 372 | h_4 | -0.6 | 0.00 | 0.05 |
| | | -0.3 | 0.00 | 0.03 | | | -0.3 | 0.00 | 0.05 |
| | | 0 | 0.00 | 0.04 | | | 0 | 0.00 | 0.04 |
| | | 0.5 | 0.00 | 0.07 | | | 0.5 | 0.00 | 0.07 |
| | | 0.9 | 0.00 | 0.32 | | | 0.9 | 0.01 | 0.31 |
| 500 | h_2 | -0.6 | 0.00 | 0.03 | 635 | h_5 | -0.6 | 0.00 | 0.04 |
| | | -0.3 | 0.00 | 0.04 | | | -0.3 | 0.00 | 0.04 |
| | | 0 | 0.00 | 0.04 | | | 0 | 0.00 | 0.04 |
| | | 0.5 | 0.00 | 0.08 | | | 0.5 | 0.00 | 0.07 |
| | | 0.9 | 0.00 | 0.32 | | | 0.9 | 0.00 | 0.32 |

Table 2: Bias and standard deviation of the sample mean estimator \bar{z}_n . Model: $y_j = h_j x_j$ with missing data patterns $h_{0,j}, \dots, h_{5,j}$, $x_j \sim AR(1)$, $N = 1000$, 1000 replications.

| $q = \sqrt{n}$ | | | | | | $q = n^{2/3}$ | | | | | |
|----------------|------|-------|--------|-------------------|-------------------|---------------|---|-------------------|-------------------|------|---|
| N | n | h | ϕ | CP _{95%} | CP _{99%} | RMSE | $E\left(\frac{\hat{s}_{X,N}^2}{s_{X,N}^2}\right)$ | CP _{95%} | CP _{99%} | RMSE | $E\left(\frac{\hat{s}_{X,N}^2}{s_{X,N}^2}\right)$ |
| 1000 | 1000 | h_0 | -0.6 | 0.91 | 0.96 | 0.17 | 0.93 | 0.84 | 0.91 | 0.36 | 0.81 |
| | | | -0.3 | 0.92 | 0.97 | 0.13 | 0.94 | 0.83 | 0.90 | 0.37 | 0.79 |
| | | | 0 | 0.93 | 0.98 | 0.13 | 0.92 | 0.85 | 0.91 | 0.34 | 0.80 |
| | | | 0.5 | 0.91 | 0.97 | 0.13 | 0.93 | 0.85 | 0.92 | 0.35 | 0.82 |
| | | | 0.9 | 0.93 | 0.98 | 0.11 | 0.89 | 0.87 | 0.93 | 0.31 | 0.79 |
| | 644 | h_1 | -0.6 | 0.91 | 0.97 | 0.13 | 0.94 | 0.88 | 0.94 | 0.29 | 0.84 |
| | | | -0.3 | 0.93 | 0.98 | 0.11 | 0.96 | 0.85 | 0.93 | 0.27 | 0.82 |
| | | | 0 | 0.93 | 0.97 | 0.12 | 0.93 | 0.87 | 0.93 | 0.29 | 0.85 |
| | | | 0.5 | 0.93 | 0.98 | 0.10 | 0.95 | 0.88 | 0.93 | 0.27 | 0.85 |
| | | | 0.9 | 0.93 | 0.98 | 0.09 | 0.86 | 0.87 | 0.94 | 0.27 | 0.83 |
| | 500 | h_2 | -0.6 | 0.92 | 0.97 | 0.11 | 0.94 | 0.88 | 0.94 | 0.27 | 0.88 |
| | | | -0.3 | 0.93 | 0.97 | 0.10 | 0.95 | 0.86 | 0.94 | 0.24 | 0.88 |
| | | | 0 | 0.92 | 0.97 | 0.09 | 0.93 | 0.88 | 0.95 | 0.23 | 0.88 |
| | | | 0.5 | 0.93 | 0.98 | 0.09 | 0.95 | 0.90 | 0.95 | 0.23 | 0.87 |
| | | | 0.9 | 0.92 | 0.99 | 0.09 | 0.86 | 0.90 | 0.96 | 0.20 | 0.88 |
| | 550 | h_3 | -0.6 | 0.91 | 0.96 | 0.17 | 0.94 | 0.85 | 0.92 | 0.31 | 0.84 |
| | | | -0.3 | 0.92 | 0.97 | 0.14 | 0.94 | 0.87 | 0.93 | 0.29 | 0.84 |
| | | | 0 | 0.92 | 0.96 | 0.14 | 0.91 | 0.86 | 0.93 | 0.27 | 0.85 |
| | | | 0.5 | 0.93 | 0.98 | 0.14 | 0.91 | 0.87 | 0.94 | 0.29 | 0.82 |
| | | | 0.9 | 0.90 | 0.97 | 0.12 | 0.84 | 0.89 | 0.94 | 0.27 | 0.83 |
| | 413 | h_4 | -0.6 | 0.93 | 0.98 | 0.09 | 0.96 | 0.91 | 0.97 | 0.18 | 0.89 |
| | | | -0.3 | 0.93 | 0.97 | 0.07 | 0.96 | 0.90 | 0.96 | 0.19 | 0.88 |
| | | | 0 | 0.91 | 0.97 | 0.07 | 0.95 | 0.91 | 0.96 | 0.20 | 0.88 |
| | | | 0.5 | 0.94 | 0.99 | 0.08 | 0.96 | 0.90 | 0.97 | 0.20 | 0.89 |
| | | | 0.9 | 0.91 | 0.97 | 0.08 | 0.84 | 0.90 | 0.97 | 0.17 | 0.91 |
| | 635 | h_5 | -0.6 | 0.93 | 0.98 | 0.10 | 0.94 | 0.87 | 0.93 | 0.29 | 0.86 |
| | | | -0.3 | 0.94 | 0.99 | 0.10 | 0.96 | 0.88 | 0.93 | 0.28 | 0.89 |
| | | | 0 | 0.92 | 0.97 | 0.10 | 0.96 | 0.87 | 0.93 | 0.28 | 0.86 |
| | | | 0.5 | 0.92 | 0.97 | 0.10 | 0.96 | 0.88 | 0.94 | 0.26 | 0.87 |
| | | | 0.9 | 0.93 | 0.98 | 0.08 | 0.87 | 0.87 | 0.94 | 0.24 | 0.82 |

Table 3: Empirical 95% and 99% coverage probabilities for the mean $\mu = Ex_j$ under unevenly spaced data based on statistic t_n . RMSE and average $E[\hat{s}_{X,N,q}^2/s_{X,N}^2]$ for the estimator $\hat{s}_{X,N}^2$. Model: $y_j = h_j x_j$ with missing data patterns $h_{0,j}, \dots, h_{5,j}$, $x_j \sim AR(1)$, $N = 1000$, 1000 replications.

| AR(2) | | | | | | | | | MA(1) | | | | |
|-------|------|-------|----------|----------|-------------------|-------------------|------|---|----------|-------------------|-------------------|------|---|
| N | n | h | ϕ_1 | ϕ_2 | CP _{95%} | CP _{99%} | RMSE | $E\left(\frac{\hat{s}_{X,N,q}^2}{s_{X,N}^2}\right)$ | θ | CP _{95%} | CP _{99%} | RMSE | $E\left(\frac{\hat{s}_{X,N,q}^2}{s_{X,N}^2}\right)$ |
| 1000 | 1000 | h_0 | -0.6 | 0.1 | 0.90 | 0.96 | 0.18 | 0.92 | -0.5 | 0.91 | 0.97 | 0.16 | 0.94 |
| | | | -0.3 | 0.6 | 0.87 | 0.94 | 0.21 | 0.74 | -0.2 | 0.93 | 0.98 | 0.12 | 0.93 |
| | | | 0.1 | 0.4 | 0.92 | 0.98 | 0.11 | 0.97 | 0.4 | 0.93 | 0.97 | 0.12 | 0.92 |
| | | | -0.1 | 0.8 | 0.93 | 0.98 | 0.11 | 0.89 | 0.7 | 0.93 | 0.98 | 0.11 | 0.95 |
| | 644 | h_1 | -0.6 | 0.1 | 0.91 | 0.97 | 0.12 | 0.93 | -0.5 | 0.91 | 0.97 | 0.12 | 0.94 |
| | | | -0.3 | 0.6 | 0.94 | 0.99 | 0.13 | 1.13 | -0.2 | 0.94 | 0.98 | 0.10 | 0.96 |
| | | | 0.1 | 0.4 | 0.94 | 0.98 | 0.10 | 0.95 | 0.4 | 0.92 | 0.98 | 0.10 | 0.93 |
| | | | -0.1 | 0.8 | 0.94 | 0.99 | 0.08 | 0.98 | 0.7 | 0.92 | 0.98 | 0.10 | 0.95 |
| | 500 | h_2 | -0.6 | 0.1 | 0.94 | 0.98 | 0.10 | 0.95 | -0.5 | 0.93 | 0.98 | 0.10 | 0.96 |
| | | | -0.3 | 0.6 | 0.92 | 0.98 | 0.08 | 0.91 | -0.2 | 0.93 | 0.98 | 0.09 | 0.95 |
| | | | 0.1 | 0.4 | 0.94 | 0.98 | 0.09 | 0.95 | 0.4 | 0.93 | 0.98 | 0.09 | 0.95 |
| | | | -0.1 | 0.8 | 0.93 | 0.98 | 0.07 | 0.91 | 0.7 | 0.93 | 0.98 | 0.08 | 0.94 |
| | 550 | h_3 | -0.6 | 0.1 | 0.91 | 0.96 | 0.19 | 0.94 | -0.5 | 0.92 | 0.97 | 0.19 | 0.94 |
| | | | -0.3 | 0.6 | 0.85 | 0.93 | 0.28 | 0.70 | -0.2 | 0.93 | 0.97 | 0.15 | 0.95 |
| | | | 0.1 | 0.4 | 0.92 | 0.97 | 0.13 | 0.94 | 0.4 | 0.91 | 0.97 | 0.14 | 0.94 |
| | | | -0.1 | 0.8 | 0.91 | 0.96 | 0.13 | 0.86 | 0.7 | 0.93 | 0.97 | 0.14 | 0.92 |
| | 407 | h_4 | -0.6 | 0.1 | 0.93 | 0.98 | 0.09 | 0.97 | -0.5 | 0.89 | 0.96 | 0.20 | 0.89 |
| | | | -0.3 | 0.6 | 0.90 | 0.96 | 0.12 | 0.78 | -0.2 | 0.90 | 0.96 | 0.18 | 0.90 |
| | | | 0.1 | 0.4 | 0.93 | 0.98 | 0.08 | 0.96 | 0.4 | 0.90 | 0.96 | 0.21 | 0.89 |
| | | | -0.1 | 0.8 | 0.93 | 0.98 | 0.08 | 0.96 | 0.7 | 0.89 | 0.96 | 0.20 | 0.90 |
| | 635 | h_5 | -0.6 | 0.1 | 0.94 | 0.98 | 0.10 | 0.95 | -0.5 | 0.94 | 0.99 | 0.10 | 0.95 |
| | | | -0.3 | 0.6 | 0.90 | 0.96 | 0.14 | 0.80 | -0.2 | 0.94 | 0.99 | 0.10 | 0.96 |
| | | | 0.1 | 0.4 | 0.92 | 0.98 | 0.10 | 0.95 | 0.4 | 0.91 | 0.98 | 0.10 | 0.94 |
| | | | -0.1 | 0.8 | 0.91 | 0.97 | 0.10 | 0.90 | 0.7 | 0.91 | 0.98 | 0.10 | 0.96 |

Table 4: Empirical 95% and 99% coverage probabilities for the mean $\mu = Ex_j$ under unevenly spaced data. RMSE and average $E[\hat{s}_{X,N,q}^2/s_{X,N}^2]$ for the estimator $\hat{s}_{X,N}^2$. Model: $y_j = h_j x_j$ with missing data patterns $h_{0,j}, \dots, h_{5,j}$, $x_j \sim AR(2)$ and $x_j \sim MA(1)$, $N = 1000$, 1000 replications, $q = \sqrt{n}$.

| $E\left(\tilde{s}_{X,N,q}^2/nv_X^2\right)$ | | | | | $E\left(\tilde{s}_{X,N,q}^2/nv_X^2\right)$ | | | | |
|--|-------|--------|----------------|---------------|--|-------|--------|----------------|---------------|
| n | h | ϕ | $q = \sqrt{n}$ | $q = n^{2/3}$ | n | h | ϕ | $q = \sqrt{n}$ | $q = n^{2/3}$ |
| 1000 | h_0 | -0.6 | 1.04 | 1.03 | 550 | h_3 | -0.6 | 1.00 | 1.06 |
| | | -0.3 | 1.00 | 1.00 | | | -0.3 | 1.01 | 0.96 |
| | | 0 | 1.01 | 1.02 | | | 0 | 0.99 | 0.99 |
| | | 0.5 | 1.00 | 1.02 | | | 0.5 | 0.98 | 0.99 |
| | | 0.9 | 0.95 | 0.99 | | | 0.9 | 0.86 | 0.86 |
| 644 | h_1 | -0.6 | 1.25 | 1.25 | | h_4 | -0.6 | 2.61 | 2.51 |
| | | -0.3 | 1.12 | 1.14 | | | -0.3 | 1.51 | 1.38 |
| | | 0 | 1.00 | 0.99 | | | 0 | 0.95 | 0.88 |
| | | 0.5 | 0.80 | 0.79 | | | 0.5 | 0.56 | 0.54 |
| | | 0.9 | 0.67 | 0.70 | | | 0.9 | 0.34 | 0.34 |
| 500 | h_2 | -0.6 | 1.45 | 1.45 | 635 | h_5 | -0.6 | 2.19 | 2.09 |
| | | -0.3 | 1.14 | 1.14 | | | -0.3 | 1.36 | 1.26 |
| | | 0 | 1.00 | 1.00 | | | 0 | 1.01 | 1.00 |
| | | 0.5 | 0.75 | 0.75 | | | 0.5 | 0.74 | 0.76 |
| | | 0.9 | 0.47 | 0.50 | | | 0.9 | 0.62 | 0.62 |

Table 5: Average ratio $E[\tilde{s}_{X,N,q}^2/nv_X^2]$ for the estimator $\tilde{s}_{X,N}^2$. Model: $y_j = h_j x_j$ with missing data patterns $h_{0,j}, \dots, h_{5,j}$, $x_j \sim AR(1)$, $N = 1000$, 1000 replications, $q = \sqrt{n}$.

4 Proofs

This section contains proofs of Theorems 1, 6, 7 and Proposition 5.

Proof of Theorem 1. We start with the proof of (8). The conditional variance of W_N can be written as

$$\begin{aligned} s_{X,N}^2 &:= \text{var}(W_N | h) = \text{var}\left(\sum_{j=1}^N h_j x_j | h\right) = \sum_{j,k=1}^N h_j h_k \text{cov}(x_j, x_k) \\ &= \sum_{j,k=1}^N h_j h_k \gamma_X(j-k). \end{aligned}$$

The spectral density f_X of $\{x_j\}$ has property $\gamma_X(k) = \int_{-\pi}^{\pi} e^{iuk} f_X(u) du$, $k \in \mathbb{Z}$. Therefore,

$$\text{var}(W_N | h) = \int_{-\pi}^{\pi} \left| \sum_{j=1}^N h_j e^{iju} \right|^2 f_X(u) du.$$

By assumption (7), $f_X(u) \geq c_{\min} > 0$. Therefore,

$$\begin{aligned} \text{var}(W_N | h) &\geq c_{\min} \int_{-\pi}^{\pi} \left| \sum_{j=1}^N h_j e^{iju} \right|^2 du \\ &= c_{\min} \int_{-\pi}^{\pi} \sum_{j,k=1}^N h_j h_k e^{iu(j-k)} du = c_{\min} 2\pi \sum_{j=k=1}^N h_j^2 = c_{\min} 2\pi n, \end{aligned}$$

because $\int_{-\pi}^{\pi} e^{iju} du = 0$ if $j \neq 0$ and $\sum_{j=1}^N h_j^2 = n$. Hence,

$$\text{var}(W_N | h) \geq c_{\min} 2\pi n.$$

Next we bound $\text{var}(W_N | h)$ from above. By (6), $f_X(u) \leq G_0/2\pi$. Hence, by (20),

$$\text{var}(W_N | h) \leq G_0(2\pi)^{-1} \int_{-\pi}^{\pi} \left| \sum_{j=1}^N h_j e^{iju} \right|^2 du = G_0 \sum_{j=1}^N h_j^2 = G_0 n.$$

This completes the proof of (8).

Proof of (9). For simplicity of notation, set

$$a_j = 0, \quad j = 0, -1, -2, \dots \quad (20)$$

Write

$$W_N - n\mu = \sum_{t=1}^N h_t (x_t - \mu) = \sum_{t=1}^N h_t \sum_{j=0}^{\infty} a_j \varepsilon_{t-j} = \sum_{j=-\infty}^N \left\{ \sum_{t=1}^N h_t a_{t-j} \right\} \varepsilon_j = \sum_{j=-\infty}^N c_{Nj} \varepsilon_j,$$

where $c_{Nj} = \sum_{t=1}^N h_t a_{t-j}$. Then we can write

$$s_{X,N}^{-1}(W_N - n\mu) = \sum_{j=-\infty}^N s_{X,n}^{-1} c_{Nj} \varepsilon_j = S_{1,N} + S_{2,N},$$

where

$$S_{1,N} = \sum_{j=-\infty}^{-\log N-1} s_{X,N}^{-1} c_{Nj} \varepsilon_j, \quad S_{2,N} = \sum_{j=-\log N}^N s_{X,N}^{-1} c_{Nj} \varepsilon_j.$$

We will show that, as $N \rightarrow \infty$,

$$S_{1,N} = o_p(1), \quad (21)$$

$$S_{2,N} \rightarrow_D \mathcal{N}(0, 1), \quad (22)$$

which implies

$$s_{X,n}^{-1}(W_N - n\mu) \rightarrow_D \mathcal{N}(0, 1),$$

verifying (9).

We start with the proof of (21). Observe that the weights $s_{X,N}^{-1} c_{Nj}$ depend on h_t 's and do not depend on $\{\varepsilon_j\}$. Then,

$$E[S_{1,N}^2 | h] = \sum_{j,k=-\infty}^{-\log N-1} s_{X,N}^{-2} c_{Nj} c_{Nk} E[\varepsilon_j \varepsilon_k] = \sigma_\varepsilon^2 s_{X,N}^{-2} \sum_{j=-\infty}^{-\log N-1} c_{Nj}^2. \quad (23)$$

We have,

$$\sum_{j=-\infty}^{-\log N-1} c_{Nj}^2 \leq \sum_{j=-\infty}^{-\log N-1} \left(\sum_{t=1}^N h_t a_{t-j} \right)^2 \leq \sum_{j=-\infty}^{-\log N-1} \sum_{t,s=1}^N h_t h_s |a_{t-j} a_{s-j}|.$$

Notice that in this sum, $s - j \geq s + \log N \geq \log N$ and $t - j \geq \log N$, and $h_s \leq 1$. Therefore,

$$\sum_{j=-\infty}^{-\log N-1} c_{Nj}^2 \leq \sum_{t=1}^N h_t \left(\sum_{j \geq \log N} |a_j| \right) \left(\sum_{s \geq \log N} |a_s| \right) = n \left(\sum_{j \geq \log N} |a_j| \right)^2.$$

By (8), $s_{X,N}^{-2} \leq Cn^{-1}$. Hence, as $N \rightarrow \infty$,

$$\begin{aligned} E[S_{1,N}^2 | h] &= s_{X,N}^{-2} \sigma_\varepsilon^2 \sum_{j=-\infty}^{-\log N-1} c_{Nj}^2 \leq \sigma_\varepsilon^2 C \left(\sum_{j \geq \log N} |a_j| \right)^2, \\ E[S_{1,N}^2] &= E[E[S_{1,N}^2 | h]] \leq C \left(\sum_{j \geq \log N} |a_j| \right)^2 = o(1), \end{aligned} \quad (24)$$

by (4), which proves (21).

Next we prove (22). Denote $\xi_{N,j} = s_{X,N}^{-1} c_{Nj} \varepsilon_j$. Notice that $\{\xi_{N,j}, j = 1, \dots, N\}$ is a triangular array of martingale differences (m.d.). Indeed, by assumption, $\{h_t\}$ and $\{\varepsilon_t\}$ are mutually independent. Denote by \mathcal{F}_{Nj} the σ -field generated by $\{\varepsilon_s, s \leq j; h_k, k = 1, \dots, N\}$. Then, $\xi_{N,j}$ is an m.d. sequence with respect to the σ -field \mathcal{F}_{Nj} :

$$E[\xi_{N,j} | \mathcal{F}_{N,j-1}] = s_{X,N}^{-1} c_{Nj} E[\varepsilon_j | \mathcal{F}_{N,j-1}] = 0. \quad (25)$$

Hence, $S_{2,N} = \sum_{j=-\log N}^N \xi_{N,j}$ is a sum of *m.d.* variables $\xi_{N,j}$. Therefore, by the central limit theorem for martingale differences (see Corollary 3.1 in [11]), to prove (22) it suffices to show that for any $\epsilon > 0$,

$$\sum_{j=-\log N}^N E[\xi_{N,j}^2 | \mathcal{F}_{N,j-1}] \rightarrow_p 1, \quad (26)$$

$$\sum_{j=-\log N}^N E[\xi_{N,j}^2 I(\xi_{N,j}^2 \geq \epsilon) | \mathcal{F}_{N,j-1}] \rightarrow_p 0. \quad (27)$$

We have

$$E[\xi_{N,j}^2 | \mathcal{F}_{N,j-1}] = s_{X,N}^{-2} c_{Nj}^2 E[\epsilon_j^2 | \mathcal{F}_{N,j-1}] = \sigma_\epsilon^2 s_{X,N}^{-2} c_{Nj}^2.$$

Therefore,

$$\sum_{j=-\log N}^N E[\xi_{N,j}^2 | \mathcal{F}_{N,j-1}] = s_{X,N}^{-2} \sum_{j=-\infty}^N \sigma_\epsilon^2 c_{Nj}^2 - E[S_{1,N}^2 | h], \quad (28)$$

by (23). Observe that

$$\begin{aligned} s_{X,N}^{-2} \sigma_\epsilon^2 \sum_{j=-\infty}^N c_{Nj}^2 &= s_{X,N}^{-2} \sigma_\epsilon^2 \sum_{j=-\infty}^N \sum_{t,s=1}^N h_t h_s a_{t-j} a_{s-j} \\ &= s_{X,N}^{-2} \sum_{t,s=1}^N h_t h_s \gamma_X(t-s) = 1, \end{aligned} \quad (29)$$

see (20). In (24) we showed that $E[S_{1,N}^2 | h] = o(1)$. This together with (28) proves (26).

It remains to verify (27). By (8), $s_{X,N}^2 \geq cn$, while by assumption (4),

$$|c_{Nj}| \leq \sum_{t=1}^N h_t |a_{t-j}| \leq \sum_{t=0}^{\infty} |a_t| = c_* < \infty.$$

Therefore, $\xi_{N,j}^2 = s_{X,N}^{-2} c_{Nj}^2 \epsilon_j^2 \leq (cn)^{-1} c_*^2 \epsilon_j^2$. Hence,

$$\begin{aligned} E[\xi_{N,j}^2 I(\xi_{N,j}^2 \geq \epsilon) | \mathcal{F}_{N,j-1}] &\leq E[\xi_{N,j}^2 I(\epsilon_j^2 \geq ncc_*^{-2}\epsilon) | \mathcal{F}_{N,j-1}] \\ &\leq s_{X,N}^{-2} c_{Nj}^2 \delta_{nj}, \quad \delta_{nj} = E[\epsilon_j^2 I(\epsilon_j^2 \geq ncc_*^{-2}\epsilon) | h]. \end{aligned}$$

Notice that $\delta_{nj} = \delta_{n1} = o_p(1)$ because $n \rightarrow_p \infty$ and n is independent of $\{\epsilon_t\}$. Together with (29), this implies

$$\sum_{j=-\log N}^N E[\xi_{N,j}^2 I(\xi_{N,j}^2 \geq \epsilon) | \mathcal{F}_{N,j-1}] \leq \delta_{n1} s_{X,N}^{-2} \sum_{j=-\log N}^N c_{Nj}^2 = \delta_{n1} \sigma_\epsilon^{-2} = o_p(1),$$

which completes the proof of (27) and of the theorem. \square

Proof of Proposition 5. Denote

$$\widehat{\tau}_{N,\nu} = \sum_{t=\nu+1}^N (y_t - \mu)(y_{t-\nu} - \mu)h_t h_{t-\nu}. \quad (30)$$

To prove (14), we show that

$$E[|\widehat{\tau}_{N,\nu} - \tau_{N,\nu}| |h] \leq Cn^{1/2}, \quad (31)$$

$$E[|\widetilde{\tau}_{N,\nu} - \widehat{\tau}_{N,\nu}| |h] \leq C, \quad (32)$$

where C does not depend on h_t 's, n , N and ν . Then,

$$E[|\widetilde{\tau}_{N,\nu} - \tau_{N,\nu}| |h] \leq E[|\widehat{\tau}_{N,\nu} - \tau_{N,\nu}| |h] + E[|\widetilde{\tau}_{N,\nu} - \widehat{\tau}_{N,\nu}| |h] \leq Cn^{1/2}.$$

First we prove (31). Without loss of generality, assume that $Ex_t = 0$. Notice that

$$E[\widehat{\tau}_{N,\nu} |h] = \sum_{t=\nu+1}^N h_t h_{t-\nu} E(x_t x_{t-\nu}) = \sum_{t=\nu+1}^N h_t h_{t-\nu} \gamma_X(\nu) = \tau_{N,\nu}.$$

Denote $w_{\nu,t} = x_t x_{t-\nu}$. Then, by (30), we can write

$$\widehat{\tau}_{N,\nu} - \tau_{N,\nu} = \sum_{t=1}^N h_t h_{t-\nu} (w_{\nu,t} - Ew_{\nu,t}).$$

Definition (1) of x_t implies that $\{w_{\nu,t}\}$ is a stationary time series. Under (5), we have $\sum_{k \in \mathbb{Z}} |\gamma_X(k)| < \infty$. Then, Theorem 4.5.2 in Giraitis, Surgailis, Koul (GSK, 2012) implies that $\gamma_w(k) = \text{Cov}(w_{\nu,t}, w_{\nu,t+k})$ has property

$$\sum_{k \in \mathbb{Z}} |\gamma_w(k)| = C_* < \infty.$$

Thus,

$$\begin{aligned} E[(\widehat{\tau}_{N,\nu} - \tau_{N,\nu})^2 |h] &= E[(\sum_{t=1}^N h_t h_{t-\nu} (w_{\nu,t} - Ew_{\nu,t}))^2 |h] \\ &= E[\sum_{t,s=1}^N (h_t h_{t-\nu})(h_s h_{s-\nu})(w_{\nu,t} - Ew_{\nu,t})(w_{\nu,s} - Ew_{\nu,s}) |h] \leq \sum_{t,s=1}^N h_t |\gamma_w(t-s)| \\ &\leq (\sum_{t=1}^N h_t) \{ \sum_{s=-\infty}^{\infty} |\gamma_w(s)| \} = C_* n. \end{aligned}$$

This implies (31):

$$E[|\widehat{\tau}_{N,\nu} - \tau_{N,\nu}| |h] \leq (E[(\widehat{\tau}_{N,\nu} - \tau_{N,\nu})^2 |h])^{1/2} \leq (C_* n)^{1/2}.$$

Next we prove (32). Again, without loss of generality set $\mu = Ex_t = 0$. Notice that

$$(x_t - \bar{z}_n)(x_{t-\nu} - \bar{z}_n) - x_t x_{t-\nu} = \bar{z}_n^2 - \bar{z}_n(x_t + x_{t-\nu}).$$

Therefore, we can write

$$\begin{aligned}
\tilde{\tau}_{N,\nu} - \hat{\tau}_{N,\nu} &= \sum_{t=\nu+1}^N h_t h_{t-\nu} ((x_t - \bar{z}_n)(x_{t-\nu} - \bar{z}_n) - x_t x_{t-\nu}) \\
&= \sum_{t=\nu+1}^N h_t h_{t-\nu} (\bar{z}_n^2 - \bar{z}_n(x_t + x_{t-\nu})) \\
&= \left(\sum_{t=\nu+1}^N h_t h_{t-\nu} \right) \bar{z}_n^2 - \bar{z}_n \left(\sum_{t=\nu+1}^N h_t h_{t-\nu} x_t \right) - \bar{z}_n \left(\sum_{t=\nu+1}^N h_t h_{t-\nu} x_{t-\nu} \right).
\end{aligned}$$

Notice that $\sum_{t=\nu+1}^N h_t h_{t-\nu} \leq \sum_{t=\nu+1}^N h_t = n$. Then, by Cauchy inequality,

$$\begin{aligned}
E[|\tilde{\tau}_{N,\nu} - \hat{\tau}_{N,\nu}| | h] &\leq E[n \bar{z}_n^2 | h] \\
&+ (E[\bar{z}_n^2 | h])^{1/2} \left\{ \left(E\left[\left(\sum_{t=\nu+1}^N h_t h_{t-\nu} x_t \right)^2 | h \right] \right)^{1/2} + \left(E\left[\left(\sum_{t=\nu+1}^N h_t h_{t-\nu} x_{t-\nu} \right)^2 | h \right] \right)^{1/2} \right\}.
\end{aligned} \tag{33}$$

By (10), $E[n \bar{z}_n^2 | h] \leq C < \infty$ and $E[\bar{z}_n^2 | h] \leq Cn^{-1}$, where C does not depend on n and $\{h_t\}$. The same argument as in the proof of (20) implies that

$$E\left[\left(\sum_{t=\nu+1}^N h_t h_{t-\nu} x_t \right)^2 | h \right] \leq C \sum_{t=\nu+1}^N h_t h_{t-\nu} \leq Cn, \tag{34}$$

$$E\left[\left(\sum_{t=\nu+1}^N h_t h_{t-\nu} x_{t-\nu} \right)^2 | h \right] \leq Cn, \tag{35}$$

where C do not depend on h_t 's, n , N and ν . Applying these bounds in (33), we obtain (32). This completes the proof of the proposition. \square

Proof of Theorem 6. Notice that (16) implies $\hat{s}_{X,N}^2 = \text{var}(W_N | h)(1 + o_p(1))$, which together with (11) proves (17).

Proof of (16). Define

$$\hat{s}_{X,N,q}^2 = \hat{\tau}_{N,0} + 2 \sum_{\nu=1}^q \hat{\tau}_{N,\nu},$$

where $\hat{\tau}_{N,\nu}$ are the same as in (30). We will show that

$$\hat{s}_{X,N,q}^2 = s_{X,N}^2 + o_p(n), \tag{36}$$

$$\hat{s}_{X,N,q}^2 = \hat{s}_{X,N,q}^2 + o_p(n). \tag{37}$$

Recall that by (8), $s_{X,N}^2 \geq Cn$, where C does not depend on n . Then,

$$\hat{s}_{X,N,q}^2 - s_{X,N}^2 = (\hat{s}_{X,N,q}^2 - s_{X,N}^2) + (\hat{s}_{X,N,q}^2 - \hat{s}_{X,N,q}^2) = o_p(n) = o_p(s_{X,N}^2). \tag{38}$$

Clearly, (38) proves (16).

Proof of (36). Without loss of generality assume that $Ex_t = 0$. Write

$$\hat{s}_{X,N,q}^2 - s_{X,N}^2 = (\hat{s}_{X,N,q}^2 - E[\hat{s}_{X,N,q}^2|h]) + (E[\hat{s}_{X,N,q}^2|h] - s_{X,N}^2).$$

To prove (36), it suffices to show that

$$|n^{-1}(E[\hat{s}_{X,N,q}^2|h] - s_{X,N}^2)| = o_p(1), \quad (39)$$

$$E[(n^{-1}(\hat{s}_{X,N,q}^2 - E[\hat{s}_{X,N,q}^2|h]))^2] = o_p(1). \quad (40)$$

First we prove (39). Recall that $E[\hat{\tau}_{N,\nu}|h] = \tau_{N,\nu}$. Thus,

$$E[\hat{s}_{X,N,q}^2|h] = E[\hat{\tau}_{N,0}|h] + 2 \sum_{\nu=1}^q E[\hat{\tau}_{N,\nu}|h] = \tau_{N,0} + 2 \sum_{\nu=1}^q \tau_{N,\nu}.$$

This together with (13), implies

$$E[\hat{s}_{X,N,q}^2|h] - s_{X,N}^2 = -2 \sum_{k=q+1}^{N-1} \tau_{N,\nu}.$$

Hence, we obtain

$$\begin{aligned} |n^{-1}(E[\hat{s}_{X,N,q}^2|h] - s_{X,N}^2)| &= 2 \left| \sum_{\nu=q+1}^{N-1} n^{-1} \tau_{N,\nu} \right| = 2 \left| \sum_{\nu=q+1}^{N-1} n^{-1} \sum_{t=\nu+1}^N h_t h_{t-\nu} \gamma_X(\nu) \right| \\ &\leq 2 \sum_{\nu=q+1}^{\infty} |\gamma_X(\nu)| (n^{-1} \sum_{t=\nu+1}^N h_t h_{t-\nu}) \leq 2 \sum_{\nu=q+1}^{\infty} |\gamma_X(\nu)| \rightarrow_p 0, \end{aligned}$$

since $\sum_{\nu=q}^{\infty} |\gamma_X(\nu)| \rightarrow 0$ as $q \rightarrow \infty$, and $n^{-1} \sum_{t=\nu+1}^N h_t h_{t-\nu} \leq 1$. This proves (39).

Next we show (40). Define for simplicity,

$$\begin{aligned} a_u &= 0, \quad u = 0, -1, -2, \dots, \\ \delta_{t,\nu} &= n^{-1} h_t h_{t-\nu} \text{ for } t = \nu + 1, \dots, N; = 0 \text{ for } t = 1, \dots, \nu. \end{aligned}$$

Using definitions (1) of x_t and (30) of $\hat{\tau}_{N,\nu}$, we obtain

$$\begin{aligned} n^{-1} \hat{\tau}_{N,\nu} &= \sum_{t=\nu+1}^N n^{-1} h_t h_{t-\nu} (x_t x_{t-\nu}) \\ &= \sum_{t=1}^N \delta_{\nu,t} \left(\sum_{u=0}^{\infty} a_u \varepsilon_{t-u} \sum_{v=0}^{\infty} a_v \varepsilon_{t-\nu-v} \right) = \sum_{t=1}^N \delta_{\nu,t} \sum_{u,v=0}^{\infty} a_u a_v (\varepsilon_{t-u} \varepsilon_{t-\nu-v}). \end{aligned}$$

Then setting $t - u = u'$ and $t - \nu - v = v'$, we write

$$n^{-1} \hat{\tau}_{N,\nu} = \sum_{u',v' \in \mathbb{Z}} v_{\nu}(u', v') \varepsilon_{u'} \varepsilon_{v'}, \quad v_{\nu}(u', v') := \sum_{t=1}^N \delta_{\nu,t} a_{t-u'} a_{t-\nu-v'}.$$

Therefore,

$$\begin{aligned} n^{-1} \sum_{v=1}^q \widehat{\tau}_{N,v} &= \sum_{u,v \in \mathbb{Z}} \left[\sum_{v=1}^q v_v(u, v) \right] \varepsilon_u \varepsilon_v, \\ n^{-1} \widehat{\tau}_{N,0} &= \sum_{u,v \in \mathbb{Z}} v_0(u, v) \varepsilon_u \varepsilon_v. \end{aligned}$$

Hence, we can write

$$q_N := n^{-1} (\widehat{s}_{X,N,q}^2 - E[\widehat{s}_{X,N,q}^2 | h]) = \sum_{u,v \in \mathbb{Z}} (v_0(u, v) + 2 \sum_{v=1}^q v_v(u, v)) (\varepsilon_u \varepsilon_v - E[\varepsilon_u \varepsilon_v]).$$

By Lemma 4.5.1 of Giraitis, Surgailis, Koul (2012),

$$E \left[\left(\sum_{u,v \in \mathbb{Z}} b_{u,v} (\varepsilon_u \varepsilon_v - E[\varepsilon_u \varepsilon_v]) \right)^2 \right] \leq C \sum_{u,v \in \mathbb{Z}} b_{u,v}^2 \quad (41)$$

for any non-random weights $b_{u,v}$, such that

$$\sum_{u,v \in \mathbb{Z}} b_{u,v}^2 < \infty,$$

where C does not depend on $b_{u,v}$ and N . Notice that the weights in the sum q_N involve h_t 's which by assumption are independent of $\{\varepsilon_t\}$. Then, using (41), we can bound

$$E[q_N^2 | h] \leq C \sum_{u,v \in \mathbb{Z}} (v_0(u, v) + 2 \sum_{v=1}^q v_v(u, v))^2,$$

where C does not depend on the weights in the sum. So,

$$\begin{aligned} E[q_N^2] = E[E[q_N^2 | h]] &\leq CE \left[\sum_{u,v \in \mathbb{Z}} (v_0(u, v) + 2 \sum_{v=1}^q v_v(u, v))^2 \right] \\ &\leq 2CE \left[\sum_{u,v \in \mathbb{Z}} v_0^2(u, v) \right] + 8CE \left[\sum_{u,v \in \mathbb{Z}} \left(\sum_{v=1}^q v_v(u, v) \right)^2 \right]. \quad (42) \end{aligned}$$

We will show that

$$E \left[\sum_{u,v \in \mathbb{Z}} v_0^2(u, v) \right] = o(1), \quad (43)$$

$$E \left[\sum_{u,v \in \mathbb{Z}} \left(\sum_{v=1}^q v_v(u, v) \right)^2 \right] = o(1). \quad (44)$$

The bounds (42), (43) and (44) imply (40).

It remains to verify (43) and (44). Notice that

$$\gamma_X(t - k) = \text{Cov}(x_t, x_k) = E\varepsilon_1^2 \sum_{u \in \mathbb{Z}} a_{t-u} a_{k-u}.$$

Then,

$$\begin{aligned}
\sum_{u,v \in \mathbb{Z}} v_0^2(u,v) &= \sum_{u,v \in \mathbb{Z}} \left(\sum_{t=1}^N \delta_{0,t} a_{t-u} a_{t-v} \right)^2 \\
&= \sum_{t,s=1}^N \delta_{0,t} \delta_{0,s} \left(\sum_{u \in \mathbb{Z}} a_{t-u} a_{s-u} \right) \left(\sum_{v \in \mathbb{Z}} a_{t-v} a_{s-v} \right) \\
&= \sigma_\varepsilon^{-4} \sum_{t,s=1}^N \delta_{0,t} \delta_{0,s} \gamma_X^2(t-s).
\end{aligned} \tag{45}$$

Recall, that $\delta_{0,s} \leq n^{-1}$, $\sum_{t=1}^N \delta_{0,t} \leq 1$ and

$$\sum_{s \in \mathbb{Z}} \gamma_X^2(s) \leq \gamma_X(0) \sum_{s \in \mathbb{Z}} |\gamma_X(s)| < \infty.$$

Hence, from (45), we obtain

$$\sum_{u,v \in \mathbb{Z}} v_0^2(u,v) \leq C n^{-1} \sum_{t=1}^N \delta_{0,t} \left\{ \sum_{s=-\infty}^{\infty} |\gamma_X(s)| \right\} \leq C n^{-1}.$$

Hence

$$E \left[\sum_{u,v \in \mathbb{Z}} v_0^2(u,v) \right] \leq C E[n^{-1}] \rightarrow 0,$$

since by assumption $n \rightarrow \infty$. This proves (43).

Next we prove (44). We have

$$\begin{aligned}
\sum_{u,v \in \mathbb{Z}} \left(\sum_{v=1}^q v_v(u,v) \right)^2 &= \sum_{u,v \in \mathbb{Z}} \left(\sum_{v=1}^q \left[\sum_{t=1}^N \delta_{v,t} a_{t-u} a_{t-v-v} \right] \right)^2 \\
&= \sum_{u,v \in \mathbb{Z}} \sum_{v,\delta=1}^q \sum_{t,s=1}^N \delta_{v,t} \delta_{\delta,s} a_{t-u} a_{s-u} a_{t-v-v} a_{s-\delta-v} \\
&= \sum_{v,\delta=1}^q \sum_{t,s=1}^N \delta_{v,t} \delta_{\delta,s} \left(\sum_{u \in \mathbb{Z}} a_{t-u} a_{s-u} \right) \left(\sum_{v \in \mathbb{Z}} a_{t-v-v} a_{s-\delta-v} \right) \\
&= \sigma_\varepsilon^{-4} \sum_{v,\delta=1}^q \sum_{t,s=1}^N \delta_{v,t} \delta_{\delta,s} \gamma_X(t-s) \gamma_X(t-s-v+\delta).
\end{aligned}$$

Using the bound $\delta_{\delta,s} \leq n^{-1}$, we obtain

$$\begin{aligned}
\sum_{u,v \in \mathbb{Z}} \left(\sum_{v=1}^q v_v(u,v) \right)^2 &\leq n^{-1} \sum_{v,\delta=1}^q \sum_{t,s=1}^N \delta_{v,t} |\gamma_X(t-s) \gamma_X(t-s-v+\delta)| \\
&\leq n^{-1} \sum_{v=1}^q \sum_{t=1}^N \delta_{v,t} \sum_{s \in \mathbb{Z}} |\gamma_X(s)| \sum_{\delta \in \mathbb{Z}} |\gamma_X(\delta)| \\
&\leq C n^{-1} \sum_{v=1}^q \sum_{t=1}^N \delta_{v,t} \leq C n^{-1} q.
\end{aligned}$$

Hence,

$$E\left[\sum_{u,v \in \mathbb{Z}} \left(\sum_{v=1}^q v_v(u,v)\right)^2\right] \leq C E[n^{-1}q] \rightarrow 0,$$

because by assumption $q < n$, $q/n = o_p(1)$. This completes the proof of (44), (40) and (36)

Proof of (37). We will show that

$$E\left[n^{-1}(\tilde{s}_{X,N,q}^2 - \hat{s}_{X,N,q}^2)\right] = o(1), \quad (46)$$

which implies (37). By definition,

$$\tilde{s}_{X,N,q}^2 - \hat{s}_{X,N,q}^2 = \tilde{\tau}_{N,0} - \hat{\tau}_{N,0} + 2 \sum_{v=1}^q (\tilde{\tau}_{N,v} - \hat{\tau}_{N,v}). \quad (47)$$

We showed in (32) that

$$E[|\tilde{\tau}_{N,v} - \hat{\tau}_{N,v}| | h] \leq C,$$

where C does not depend on h_t 's, N and v .

Before proceeding to the proof of (46), recall, that n and q depend on h_1, \dots, h_N and are random variables. Together with (47), this implies

$$\begin{aligned} E[|\tilde{s}_{X,N,q}^2 - \hat{s}_{X,N,q}^2| | h] &\leq E[|\tilde{\tau}_{N,0} - \hat{\tau}_{N,0}| | h] + 2 \sum_{v=1}^q E[|\tilde{\tau}_{N,v} - \hat{\tau}_{N,v}| | h] \\ &\leq C + 2Cq. \end{aligned}$$

Hence,

$$\begin{aligned} E[n^{-1}|\tilde{s}_{X,N,q}^2 - \hat{s}_{X,N,q}^2|] &= E[n^{-1}E[|\tilde{s}_{X,N,q}^2 - \hat{s}_{X,N,q}^2| | h]] \\ &\leq CE[n^{-1}(1+q)] = o(1) \end{aligned}$$

since $q \leq n$, $q = o_p(n)$. This proves (46) and completes the proof of the theorem. \square

Proof of Theorem 7. Denote $\tilde{h}_t = n^{-1/2}h_t$. First we will show that,

$$\tilde{h}_1 + \sum_{j=2}^n |\tilde{h}_j - \tilde{h}_{j-1}| = o_p\left(\sum_{j=1}^n \tilde{h}_j^2\right). \quad (48)$$

Notice that $\tilde{h}_j - \tilde{h}_{j-1} \neq 0$ when indexes j and $j-1$ belong to different clusters of data. Hence, the left hand side of (48) can be bounded by $1/\sqrt{n} + 2M/\sqrt{n} = o_p(1)$, because $M = o_p(\sqrt{n})$ by assumption of the theorem. Since $\sum_{j=1}^n \tilde{h}_j^2 = 1$, this verifies (48).

By (20),

$$n^{-1}s_{X,N}^2 = n^{-1} \sum_{j,k=1}^N h_j h_k \gamma_X(j-k) = \sum_{j,k=1}^N \tilde{h}_j \tilde{h}_k \gamma_X(j-k). \quad (49)$$

Property (5) implies, that the spectral density $f_X(u)$ is a continuous bounded function. Therefore, (18) follows using (49) and (48) by the same argument as in the proof of (2.20) of Proposition 2.2 in Abadir and *et al.* (2014).

Convergence (19) follows from (11) using (18). \square

Acknowledgments. We thank the Editor, Associate Editor and the referees for constructive comments and valuable suggestions.

References

- [1] Abadir, K.M., Distaso, W., Giraitis, L., Koul, H.L. (2014). Asymptotic normality for weighted sums of linear processes. *Econometric Theory*, **30**, 252-284.
- [2] Andrews, D.W.K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* **59**, 817–858.
- [3] Brillinger, D.R. (1973). Estimation of the mean of a stationary time series by sampling. *Journal of Applied Probability*, **10**, 419-431.
- [4] Clark, P.K. (1973). A subordinated stochastic process model with finite variance for speculative prices. *Econometrica*, **41**, 135-155.
- [5] Datta, D. D. and Du, W. (2012). Nonparametric HAC estimation for time series data with missing observations. *Working Paper*.
- [6] Dunsmuir, W. and Robinson, P.M. (1981). Estimation of time series models in the presence of missing data. *Journal of the American Statistical Association*, **76**, 560-568.
- [7] Dunsmuir, W. and Robinson, P.M. (1981a). Asymptotic theory for time series containing missing and amplitude modulated observations. *Sankhiā: The Indian Journal of Statistics*, **43**, 260-281.
- [8] Dunsmuir, W. and Robinson, P.M. (1981b). Parametric estimators for stationary time series with missing observations. *Advances of Applied Probability*, **13**, 129-146.
- [9] Ghysels, E., Gouriéroux, C., Jasiak, J. (1997). Kernel autocorrelogram for time-deformed processes. *Journal of Statistical Planning and Inference*, **68**, 167-191.
- [10] Giraitis, L., Koul, H.L., Surgailis, D. (2012). Large sample inference for long memory processes. *Imperial College Press, London*.
- [11] Hall, P. and Heyde, C. C. (1980). *Martingale Limit Theory and its Application*. Academic Press.
- [12] Jones, R.H. (1962). Spectral analysis with regularly missed observations. *Annals of Mathematical Statistics*, **33**, 455-461.
- [13] Parzen, E. (1963). On spectral analysis with missing observations and amplitude modulation. *The Indian Journal of Statistics*, **25**, 383-392.