Statistica Sinica **33** (2023), 331-351
doi:https://doi.org/10.5705/ss.202020.0323

ON THE CONSISTENCY OF LEAST SQUARES ESTIMATOR IN MODELS SAMPLED AT RANDOM TIMES DRIVEN BY LONG MEMORY NOISE: THE JITTERED CASE

Héctor Araya[1], Natalia Bahamonde[2], Lisandro Fermín[3],
Tania Roa[1] and Soledad Torres[3]

[1]*Universidad Adolfo Ibáñez,* [2]*Pontificia Universidad Católica de Valparaíso and* [3]*Universidad de Valparaíso*

Abstract: In numerous applications, data are observed at random times. Our main purpose is to study a model observed at random times that incorporates a long-memory noise process with a fractional Brownian Hurst exponent H. We propose a least squares estimator in a linear regression model with long-memory noise and a random sampling time called "jittered sampling". Specifically, there is a fixed sampling rate $1/N$, contaminated by an additive noise (the jitter) and governed by a probability density function supported in $[0, 1/N]$. The strong consistency of the estimator is established, with a convergence rate depending on N and the Hurst exponent. A Monte Carlo analysis supports the relevance of the theory and produces additional insights, with several levels of long-range dependence (varying the Hurst index) and two different jitter densities.

Key words and phrases: Least squares estimator, long-memory noise, random times, regression model.

1. Introduction

In research areas such as finance, network traffic, meteorology, and astronomy, among others, it has been noticed that observations can be carried out by sampling with random disturbances. Examples of this sampling method are data behavior until it is necessary to increase the sampling frequency, measurements obtained at random times, and defining a stopping time when a particular event occurs. In particular, Nieto-Barajas and Sinha (2015) discuss a Bayesian interpolation of unequally spaced time series. The case of paleoclimate time series was considered by Max-Moerbeck et al. (2014) and Ólafsdóttir, Schulz and Mudelsee (2016), who estimate the significance of cross-correlations in unevenly sampled astronomical time series. Finally, in the area of computer science, we can mention the works of Chang (2014) and Zhao, Chen and Nakagawa (2014).

Corresponding author: Lisandro Fermín, CIMFAV, Universidad de Valparaíso, Valparaíso, 2362905, Chile. E-mail: lisandro.fermin@uv.cl.

The study of statistical models in the aforementioned situations is quite promising, with open problems such as statistical inference and the limit behavior of the estimators.

We propose taking a first step in this direction by studying a least squares (LS) estimator in a simple regression model with long-memory noise and observations measured at random times.

Previous works in this direction include Vilar (1995), who studied the nonparametric kernel estimator of the regression function, $m(x) = \mathbb{E}(Y|X = x)$, under mixing dependence conditions. In addition, the Ornstein–Uhlenbeck process driven by Brownian motion was studied by Vilar and Vilar (2000).

Masry (1983) studied the problem of estimating an unknown probability density function based on n independent observations sampled at random times.

Using a wavelet analysis, Bardet and Bertrand (2010) studied a nonparametric estimator of the spectral density of a Gaussian process with stationary increments, including the case of fractional Brownian motion (fBm), from the observation of one path at some particular class of random discrete times. They prove a central limit theorem and provide an application to biological data.

Philippe, Robet and Viano (2020) give the latest works on this topic, and study the preservation of memory in a statistical model. With respect to the problem of parameter estimation in time series that may be represented as a trend plus long-memory noise, see also the works of Baillie and Chung (2002), Brockwell (2007), and Lobato and Velasco (2002), among others.

We consider a jittered sampling (JS) scheme, which we define properly in Section 2. The term *jitter* is related to temporal variability during the sending of digital signals or to small variations in the accuracy of the clock signal; see Bellhouse (1981) and the references therein. The term has also recently appeared in works related to the analysis of computational images, such as Khan (2017), Krune et al. (2016), and Subr et al. (2014)

Our main purpose in studying a model with long-memory noise is to characterize the strong correlations between observations, or persistence, by a slow decay of the correlations. To explain this phenomenon in a model, it is common to represent it using the Hurst exponent H, which takes values in $[0, 1]$. In particular, the long-range dependence can be seen when $H \in (1/2, 1)$. Since the work of Mandelbrot and Van Ness (1968), the effect of long-range dependence has been the topic of numerous. One of the most popular Gaussian stochastic processes with long-memory is the fBm. Extensions to the fBm with the same covariance structure are Rosenblatt, Tudor (2008) and Hermite, Tudor (2013).

With these motivations in mind, we consider the following simple regression model:

$$Y_{\tau_i} = a\tau_i + \Delta B^H_{\tau_i}, \quad i = 0, \ldots, N-1, \tag{1.1}$$

where $a \in \mathbb{R}$ is the drift parameter of the model, $\Delta B^H_{\tau_i} = B^H_{\tau_{i+1}} - B^H_{\tau_i}$, and $\tau := \{\tau_i, 0 \leq i \leq N-1\}$ is the random time given by the JS.

The main interest in this work is to prove the strong consistency of the LS estimator in a random sampled linear regression model with long-memory noise and an independent set of random times given by JS. Recall that the process $Y := \{Y_{\tau_i}, 0 \leq i \leq N-1\}$ defined by equation (1.1) has long-range dependence and is nonstationary in the weak sense.

The remainder of the paper proceeds as follows. In Section 2, we define the random times within the random sampled regression model with long-memory noise and describe our notation. Section 3 is devoted to our main results. We use a LS procedure to obtain the parameter estimation and analyze the almost sure convergence using JS random time defined in the above section. In Section 4, a simulation study is presented to illustrate the performance of the estimator, considering different values of H and JS random time. Finally in Section A, we present the proof of Lemma 1, which was established in Section 3 and is necessary to prove the almost sure convergence of the estimator.

2. Preliminaries

In this section, we introduce the main tools from the stochastic calculus needed in the remainder of the paper. We present the random noise evaluated at the JS random time that we consider throughout this work.

The long-memory process, B^H, with Hurst parameter $H \in (1/2, 1)$, is a centered process with the following properties:

(HN1) The covariance structure is given by

$$R_H(t,s) := \mathbb{E}\left[B^H_t B^H_s\right] = \frac{\sigma^2}{2}\left[t^{2H} + s^{2H} - |t-s|^{2H}\right]. \tag{2.1}$$

(HN2) It is a self-similar process (with index H) with weakly stationary increments.

Remark 1. For example, B^H can represent the well known fBm. In the fBm framework, if $H = 1/2$, then B^H is a standard Brownian motion. Other types of long-memory processes with the same covariance structure as (2.1) are Hermite and Rosenblatt processes. In addition, the process B^H is not a semimartingale if $H \neq 1/2$. Hence, we cannot apply the classical Itô calculus to B^H.

The random time sequence $\tau = \{\tau_i;\ i = 0, \ldots, N-1\}$ is strictly increasing; here, N represents the sample size and also the sampling frequency or sampling rate, that is, the average number of samples obtained in $[0,1]$. In the following, we focus on the case where τ exhibits the following feature.

JS: We assume that we observe a certain process at irregular times τ, with period $\delta = 1/N > 0$, but contaminated by an additive noise ν, which represents possible measurement errors, satisfying the following hypothesis:

(HJ) $\nu = \{\nu_{i,N};\ 0 \le i \le N-1\}$ is a sequence of independent and identically distributed (i.i.d.) random variables with common density function $g_N(\cdot)$, depending on N, with support on $[0, 1/N]$.

Thus, the sequence of random times τ_i, for $0 \le i \le N-1$, is given by

$$\tau_i = \frac{i}{N} + \nu_{i,N}, \quad i = 0, \ldots, N-1. \tag{2.2}$$

Remark 2. Some examples of distributions that satisfy **(HJ)** are the following:

1. Uniform distribution on $[0, 1/N]$,
2. Triangular distribution with parameters $(0, 1/2N, 1/N)$. The probability density function is given by

$$f_X(x) = \begin{cases} 0, & x < 0, \\ 4N^2 x, & 0 \le x < \frac{1}{2N}, \\ 2N, & x = \frac{1}{2N}, \\ 4N^2, \left(\frac{1}{N} - x\right) & \frac{1}{2N} < x \le \frac{1}{N}. \end{cases}$$

3. Raised cosine distribution with parameters $\mu = 1/2N$ and $s = 1/2N$. The probability density function is given by

$$f_X(x) = N\left[1 + \cos\left(\frac{x - 1/2N}{1/2N}\pi\right)\right], \quad 0 \le x \le \frac{1}{N}.$$

Remark 3. Note that the hypothesis **(HJ)** implies $\mathbb{E}(\nu_{i,N}) \le 1/N$ and $Var(\nu_{i,N}) \le 1/N^2$.

Finally, we also assume the following hypothesis:

(HJN) The random time sequence τ and the long-memory noise B^H are independent.

Remark 4. In the deterministic case, that is, $\tau_i = i/N$, for $0 \leq i \leq N$, for the Brownian motion case, $H = 1/2$, the rate of the L^2-convergence for the LS estimator is obtained from the property of independent increments. In fact, we have $\hat{a}_N - a = \sum_{i=0}^{N-1} \tau_i \Delta B_{\tau_i} / \sum_{i=0}^{N-1} \tau_i^2 = (6/((N+1)(2N+1))) \sum_{i=0}^{N-1} i \Delta B_{\tau_i}$ and

$$\mathbb{E}\left[(\hat{a}_N - a)^2\right] = \mathbb{E}\left[\left(\frac{6}{(N+1)(2N+1)} \sum_{i=0}^{N-1} i \Delta B_{\tau_i}\right)^2\right]$$
$$= \frac{6}{(N+1)(2N+1)} \leq \frac{6}{N^2}.$$

A direct application of the Borel–Cantelli lemma allows us to obtain the almost sure convergence of $\hat{a}_N$ to a. In the fBm case, the increments are no longer independent. However, for $H > 1/2$, a simple modification to the procedure given in the main result (Theorem 1) of this article (Section 3) allow us to obtain L^2 and the almost sure convergence of $\hat{a}_N$ to a.

In particular, Araya et al. (2023), study the consistency of the estimated drift parameter, when $\nu_{i,N}$ has a uniform distribution.

3. Main Results

In this section, we provide our main results. We study the LS estimation (LSE) for the random sampled linear regression model (1.1), with random times given by JS and long-memory noise. We prove that the LSE is an unbiased estimator for a, and that $\hat{a}_N$ converges almost surely to a (strongly consistent). Note that all the results presented here can be extended to a noise with the same covariance structure as the fBm, such as Rosenblatt and Hermite processes.

For the estimation of the drift parameter a in the model (1.1), the LS estimator is determined by

$$\hat{a}_N = \frac{\sum_{i=0}^{N-1} \tau_i Y_{\tau_i}}{\sum_{i=0}^{N-1} \tau_i^2}. \tag{3.1}$$

Recall that from (1.1) and (3.1), we have

$$\hat{a}_N - a = \frac{(1/N) \sum_{i=0}^{N-1} \tau_i \Delta B_{\tau_i}^H}{(1/N) \sum_{i=0}^{N-1} \tau_i^2} := \frac{A_N}{D_N}. \tag{3.2}$$

To study the asymptotic behavior of (3.2), we analyze the numerator and the denominator separately.

Remark 5. Note that $1 - \tau_{N-1} \leq 1/N \to 0$ as $N \to \infty$. Then, the almost sure convergence of $\tau_{N-1} \to 1$ is ensured as N goes to infinity.

Theorem 1. *Let τ be given by* (2.2)*. Assume that the regression model* (1.1) *satisfies hypotheses **(HN1)**, **(HN2)**, **(HJ)**, and **(HJN)**. Then, the LS estimator $\hat{a}_N$ given in* (3.1) *of the drift parameter a in the model* (1.1) *is strongly consistent, that is,*

$$\hat{a}_N \xrightarrow[N\to\infty]{a.s.} a.$$

Proof. To prove our main theorem, we need an auxiliary lemma related to the almost sure convergence of the denominator D_N given in (3.2). The proof of this lemma is provided in the Appendix.

Lemma 1. *Let D_N be defined in* (3.2)*. Let $\tau = \{\tau_i; 0 \leq i \leq N-1\}$ be the random sampling times defined by* (2.2)*. If τ satisfies hypothesis **(HJ)**, then*

$$D_N \xrightarrow[N\to\infty]{a.s.} \frac{1}{3}.$$

Remark 6. A direct computation gives the convergence of $D_N \to 1/3$ if we consider deterministic times $\tau_i = i/N$.

Hence, by Lemma 1, it remains to study the asymptotic behavior of A_N as $N \to \infty$.

It is quite easy to see by the definition of A_N and conditioning on τ that $\mathbb{E}\left[A_N\right] = 0$.

Let us compute $\mathbb{E}\left[A_N^2\right]$:

$$\begin{aligned}
\mathbb{E}\left[A_N^2\right] &= \mathbb{E}\left[\frac{1}{N^2}\sum_{i=0}^{N-1}\tau_i^2\left(B^H_{\tau_{i+1}} - B^H_{\tau_i}\right)^2\right] \\
&\quad + \mathbb{E}\left[\frac{1}{N^2}\sum_{0\leq i,j\leq N-1;|i-j|=1}\tau_i\tau_j\left(B^H_{\tau_{i+1}} - B^H_{\tau_i}\right)\left(B^H_{\tau_{j+1}} - B^H_{\tau_j}\right)\right] \\
&\quad + \mathbb{E}\left[\frac{1}{N^2}\sum_{0\leq i,j\leq N-1;|i-j|\geq 2}\tau_i\tau_j\left(B^H_{\tau_{i+1}} - B^H_{\tau_i}\right)\left(B^H_{\tau_{j+1}} - B^H_{\tau_j}\right)\right] \\
&:= \mathbb{E}\left[A_N^{(1)}\right] + \mathbb{E}\left[A_N^{(2)}\right] + \mathbb{E}\left[A_N^{(3)}\right],
\end{aligned} \tag{3.3}$$

where we split the sum into three terms associated with the distance of the indices. First, we study the first term in (3.3):

$$\mathbb{E}\Big[A_N^{(1)}\Big] = \frac{1}{N^2}\sum_{i=0}^{N-1}\mathbb{E}\Bigg[\mathbb{E}\Bigg[\left|\frac{i}{N}+\nu_{i,N}\right|^2\left|B^H_{(i+1)/N+\nu_{i+1,N}} - B^H_{i/N+\nu_{i,N}}\right|^2\Bigg|\,\nu_{i,N}=s_i, \nu_{i+1,N}=s_{i+1}\Bigg]\Bigg]$$

$$= \frac{1}{N^2}\sum_{i=0}^{N-1}\int_0^{1/N}\!\!\int_0^{1/N}\left|\frac{i}{N}+s_i\right|^2\mathbb{E}\left[\left|B^H_{(i+1)/N+s_{i+1}} - B^H_{i/N+s_i}\right|^2\right] g_N(s_i)g_N(s_{i+1})ds_ids_{i+1}$$

$$= \frac{1}{N^2}\sum_{i=0}^{N-1}\int_0^{1/N}\!\!\int_0^{1/N}\left|\frac{i}{N}+s_i\right|^2\left|\frac{i+1}{N}+s_{i+1}-\frac{i}{N}-s_i\right|^{2H} g_N(s_i)g_N(s_{i+1})ds_ids_{i+1}$$

$$= \frac{1}{N^2}\sum_{i=0}^{N-1}\int_0^{1/N}\!\!\int_0^{1/N}\left|\frac{i}{N}+s_i\right|^2\left|\frac{1}{N}+s_{i+1}-s_i\right|^{2H} g_N(s_i)g_N(s_{i+1})ds_ids_{i+1}.$$

From the properties of the long-memory noise, from hypotheses **(HN1)**, **(HN2)**, and **(HJN)**, and because the domain of the random variables $\nu_{i,N}$ is $[0,1/N]$, we obtain

$$\mathbb{E}\left[A_N^{(1)}\right] \le \frac{1}{N^2}\sum_{i=0}^{N-1}\int_0^{1/N}\int_0^{1/N}\left|\frac{i}{N}+\frac{1}{N}\right|^2\left|\frac{1}{N}+\frac{1}{N}\right|^{2H} g_N(s_i)g_N(s_{i+1})ds_ids_{i+1}$$

$$= \frac{1}{N^2}\sum_{i=0}^{N-1}\left|\frac{i}{N}+\frac{1}{N}\right|^2\left|\frac{1}{N}+\frac{1}{N}\right|^{2H}\int_0^{1/N}\int_0^{1/N} g_N(s_i)g_N(s_{i+1})ds_ids_{i+1}$$

$$\le \frac{2^{2H}}{N^{4+2H}}\frac{N(N+1)(2N+1)}{6} \le \frac{C_1(H)}{N^{1+2H}}, \tag{3.4}$$

with $C_1(H) = 2^{2H}/3$.

Second, we consider the case of $|i-j|=1$ in (3.3). For simplicity, we take $j<i$, that is, $j=i-1$; the other case can be treated the same way. Therefore,

$$\mathbb{E}\left[A_N^{(2)}\right] = \frac{2}{N^2}\sum_{i=0}^{N-2}\mathbb{E}\left[\tau_{i+1}\tau_i\left(B^H_{\tau_{i+2}} - B^H_{\tau_{i+1}}\right)\left(B^H_{\tau_{i+1}} - B^H_{\tau_i}\right)\right] \tag{3.5}$$

$$= \frac{2}{N^2}\sum_{i=0}^{N-2}\int_0^{1/N}\int_0^{1/N}\int_0^{1/N}\left(\frac{i+1}{N}+s_{i+1}\right)\left(\frac{i}{N}+s_i\right)\times \mathbb{E}\left[\left(B^H_{(i+2)/N+s_{i+2}} - B^H_{(i+1)/N+s_{i+1}}\right)\left(B^H_{(i+1)/N+s_{i+1}} - B^H_{i/N+s_i}\right)\right] g_N(s_i)g_N(s_{i+1})g_N(s_{i+2})ds_ids_{i+1}ds_{i+2},$$

where in the last term, we apply the conditional expectation with respect to $\nu_{i,N} = s_i$, $\nu_{i+1,N} = s_{i+1}$ and $\nu_{i+2,N} = s_{i+2}$. Because

$$\begin{aligned}&\mathbb{E}\left[\left(B^H_{(i+2)/N+s_{i+2}} - B^H_{(i+1)/N+s_{i+1}}\right)\left(B^H_{(i+1)/N+s_{i+1}} - B^H_{i/N+s_i}\right)\right] \\ &= \frac{1}{2}\left[\left|s_{i+2} - s_i + \frac{2}{N}\right|^{2H} - \left|s_{i+2} - s_{i+1} + \frac{1}{N}\right|^{2H} - \left|s_{i+1} - s_i + \frac{1}{N}\right|^{2H}\right],\end{aligned}$$

we have

$$\mathbb{E}\left[\left(B^H_{(i+2)/N+s_{i+2}} - B^H_{(i+1)/N+s_{i+1}}\right)\left(B^H_{(i+1)/N+s_{i+1}} - B^H_{i/N+s_i}\right)\right] \leq \frac{3^{2H}}{2N^{2H}}, \tag{3.6}$$

Plugging inequality (3.6) into equation (3.5) yields

$$\begin{aligned}\mathbb{E}\left[A_N^{(2)}\right] &\leq \frac{2}{N^2}\sum_{i=0}^{N-2}\int_0^{1/N}\int_0^{1/N}\int_0^{1/N}\left[\frac{i+1}{N} + s_{i+1}\right]\left[\frac{i}{N} + s_i\right]\frac{3^{2H}}{2N^{2H}} \\ &\quad \times g_N(s_i)g_N(s_{i+1})g_N(s_{i+2})ds_i ds_{i+1} ds_{i+2} \\ &\leq \frac{3^{2H}}{N^{2+2H}}\sum_{i=0}^{N-2}\int_0^{1/N}\int_0^{1/N}\int_0^{1/N}\left[\frac{i+1}{N} + \frac{1}{N}\right]\left[\frac{i}{N} + \frac{1}{N}\right] \\ &\quad \times g_N(s_i)g_N(s_{i+1})g_N(s_{i+2})ds_i ds_{i+1} ds_{i+2} \\ &= \frac{3^{2H}}{N^{4+2H}}\sum_{i=0}^{N-2}(i+2)(i+1) \leq \frac{3^{2H}}{3N^{4+2H}}(N-1)N(N+1) \\ &\leq \frac{C_2(H)}{N^{1+2H}},\end{aligned} \tag{3.7}$$

where $C_2(H) = 3^{2H-1}$. Finally, we consider the case $|i-j| \geq 2$ in (3.3). Conditioning on $\nu_{i,N} = s_i, \nu_{i+1;N} = s_{i+1}, \nu_{j,N} = s_j$, and $\nu_{j+1,N} = s_{j+1}$, we get

$$\begin{aligned}\mathbb{E}\left[A_N^{(3)}\right] &= \frac{1}{N^2}\mathbb{E}\left[\sum_{0\leq i,j\leq N-1;|i-j|\geq 2}\tau_i\tau_j\left(B^H_{\tau_{i+1}} - B^H_{\tau_i}\right)\left(B^H_{\tau_{j+1}} - B^H_{\tau_j}\right)\right] \\ &= \frac{1}{N^2}\sum_{0\leq i,j\leq N-1;|i-j|\geq 2}\int_0^{1/N}\int_0^{1/N}\int_0^{1/N}\int_0^{1/N}\left(\frac{i}{N} + s_i\right)\left(\frac{j}{N} + s_j\right) \\ &\quad \times\mathbb{E}\left[\left(B^H_{(i+1)/N+s_{i+1}} - B^H_{i/N+s_i}\right)\left(B^H_{(j+1)/N+s_{j+1}} - B^H_{j/N+s_j}\right)\right] \\ &\quad \times g_N(s_i)g_N(s_{i+1})g_N(s_j)g_N(s_{j+1})ds_i ds_{i+1} ds_j ds_{j+1} \\ &= \frac{1}{N^2}\sum_{0\leq i,j\leq N-1;|i-j|\geq 2}\int_0^{1/N}\int_0^{1/N}\int_0^{1/N}\int_0^{1/N}\left(\frac{i}{N} + s_i\right)\left(\frac{j}{N} + s_j\right)\end{aligned}$$

$$\times\ \mathbf{I_{i,j}}\ \ g_N(s_i)g_N(s_{i+1})g_N(s_j)g_N(s_{j+1})ds_ids_{i+1}ds_jds_{j+1}, \tag{3.8}$$

where

$$\begin{aligned}\mathbf{I_{i,j}} &:= \mathbb{E}\left[\left(B^H_{(i+1)/N+s_{i+1}} - B^H_{i/N+s_i)}\right)\left(B^H_{(j+1)/N+s_{j+1}} - B^H_{j/N+s_j}\right)\right] \\ &= \frac{1}{2}\left[\left|\frac{i-j+1}{N} + s_{i+1} - s_j\right|^{2H} + \left|\frac{i-j-1}{N} + s_i - s_{j+1}\right|^{2H}\right. \\ &\quad \left. - \left|\frac{i-j}{N} + s_{i+1} - s_{j+1}\right|^{2H} - \left|\frac{i-j}{N} + s_i - s_j\right|^{2H}\right].\end{aligned}$$

For $i-j=2$ (equivalently for $i-j=-2$), we directly get

$$\begin{aligned}\mathbf{I_{i,j}} &:= \frac{1}{2}\left[\left|\frac{3}{N} + s_{i+1} - s_j\right|^{2H} + \left|\frac{1}{N} + s_i - s_{j+1}\right|^{2H}\right. \\ &\quad \left. - \left|\frac{2}{N} + s_{i+1} - s_{j+1}\right|^{2H} - \left|\frac{1}{N} + s_i - s_j\right|^{2H}\right] \\ &\leq \frac{1}{2}\left[\left|\frac{3}{N} + s_{i+1} - s_j\right|^{2H} + \left|\frac{1}{N} + s_i - s_{j+1}\right|^{2H}\right] \\ &\leq \frac{1}{2}\left[\left|\frac{4}{N}\right|^{2H} + \left|\frac{2}{N}\right|^{2H}\right] = \frac{2^{2H-1}(2^{2H}+1)}{N^{2H}}.\end{aligned} \tag{3.9}$$

For $|i-j|>2$, applying the Taylor expansion to the function x^{2H} yields

$$\begin{aligned}&\left|\frac{i-j+1}{N} + s_{i+1} - s_j\right|^{2H} - \left|\frac{i-j}{N} + s_i - s_j\right|^{2H} \\ &= 2H\left|\frac{i-j}{N} + s_i - s_j\right|^{2H-1}\left(s_{i+1} - s_i + \frac{1}{N}\right) + R^1_N\end{aligned}$$

and

$$\begin{aligned}&\left|\frac{i-j}{N} + s_{i+1} - s_{j+1}\right|^{2H} - \left|\frac{i-j-1}{N} + s_i - s_{j+1}\right|^{2H} \\ &= 2H\left|\frac{i-j-1}{N} + s_i - s_{j+1}\right|^{2H-1}\left(s_{i+1} - s_i + \frac{1}{N}\right) + R^2_N.\end{aligned}$$

Therefore,

$$\mathbf{I_{i,j}} = 2H\left(s_{i+1} - s_i + \frac{1}{N}\right)\left[\left|\frac{i-j}{N} + s_i - s_j\right|^{2H-1}\right.$$

$$- \left|\frac{i-j-1}{N} + s_i - s_{j+1}\right|^{2H-1}\Bigg] + R_N^1 - R_N^2.$$

Again, applying the Taylor theorem to the function x^{2H-1}, we obtain

$$\mathbf{I_{i,j}} = 2H(2H-1)\left(s_{i+1} - s_i + \frac{1}{N}\right)\left(s_{j+1} - s_j + \frac{1}{N}\right) \times \left|\frac{i-j-1}{N} + s_i - s_{j+1}\right|^{2H-2} + R_N^3 + R_N^1 - R_N^2,$$

which from $2H - 2 < 0$, and for $|i-j| > 2$, implies that

$$\mathbf{I_{i,j}} \leq \frac{C_3(H)}{N^2}\left|\frac{i-j-1}{N} + s_i - s_{j+1}\right|^{2H-2} \leq \frac{C_3(H)}{N^{2H}}, \tag{3.10}$$

where $C_3(H) = 8H(2H-1)$. Note that the remainder terms R_N^1, R_N^2, and R_N^3 are of order $N^{-2H}$and uniformly independent on i and j. Plugging (3.9) and (3.10) into the expression (3.8) and considering $C_4(H) = C_3(H) + 2^{2H-1}(2^{2H}+1)$, we obtain

$$\begin{aligned}\mathbb{E}\left[A_N^{(3)}\right] &\leq \frac{C_4(H)}{N^{2H+2}} \sum_{0\leq i,j\leq N-1;|i-j|\geq 2} \int_0^{1/N}\int_0^{1/N}\int_0^{1/N}\int_0^{1/N}\left(\frac{i}{N} + s_i\right)\left(\frac{j}{N} + s_j\right) \\ &\quad \times g_N(s_i)g_N(s_{i+1})g_N(s_j)g_N(s_{j+1})ds_i ds_{i+1} ds_j ds_{j+1} \\ &\leq \frac{C_4(H)}{N^{2H+2}} \sum_{0\leq i,j\leq N-1;|i-j|\geq 2} \left(\frac{i+1}{N}\right)\left(\frac{j+1}{N}\right). \end{aligned} \tag{3.11}$$

Moreover, the above sum in (3.11) can be computed as follows:

$$\frac{1}{N^2} \sum_{0\leq i,j\leq N-1;|i-j|\geq 2} \left(\frac{i+1}{N}\right)\left(\frac{j+1}{N}\right) = \frac{1}{N^4}\left(\frac{N^4}{4} - \frac{N^3}{6} - \frac{N^2}{4} + \frac{N}{6}\right) \leq \frac{1}{4}.$$

Therefore,

$$\mathbb{E}\left[A_N^{(3)}\right] \leq \frac{C_5(H)}{N^{2H}}, \tag{3.12}$$

where $C_5(H) = C_4(H)/4$. Substituting (3.4), (3.7), and (3.12) into the equation in (3.3), we obtain

$$\mathbb{E}\left[(A_N)^2\right] \leq \frac{C_1(H) + C_2(H)}{N^{1+2H}} + \frac{C_5(H)}{N^{2H}}. \tag{3.13}$$

Because $H > 1/2$, the L^2 rate of A_N is faster than $1/N$. A direct application of the Borell–Cantelli lemma yields $A_N \xrightarrow[N\to\infty]{a.s.} 0$.

Remark 7. Under a random sampling scheme, the L^2-convergence is of order $1/N^{2H}$. The Borel–Cantelli lemma allows us to get the almost sure convergence of order $1/N^{2H-1}$.

When $\tau_i = i/N$, from remark 6 and with a small modification, we can ensure the same L^2 and almost sure convergence given in Theorem 1.

For $H < 1/2$ (anti-persistent case), the same arguments as before assure us that A_N converges to zero in L^2, and then in probability. Furthermore, D_N converges almost surely to 1/3, and therefore in probability. Using Slutsky's Theorem, we obtain the convergence in probability of $\hat{a}_N$ to a. However, our method does not obtain the almost sure convergence in this case ($H < 1/2$).

4. Simulation Study

In this section, we present a Monte Carlo simulation study to evaluate the performance of the LS estimator for a finite sample ($N = 10, 15, 20, 25, 50, 75$, and 100) in a linear regression model (1.1). We also perform a study for a large sample size ($N = 300$). In both cases, 1,000 replicates of the $\hat{a}_N$ are studied.

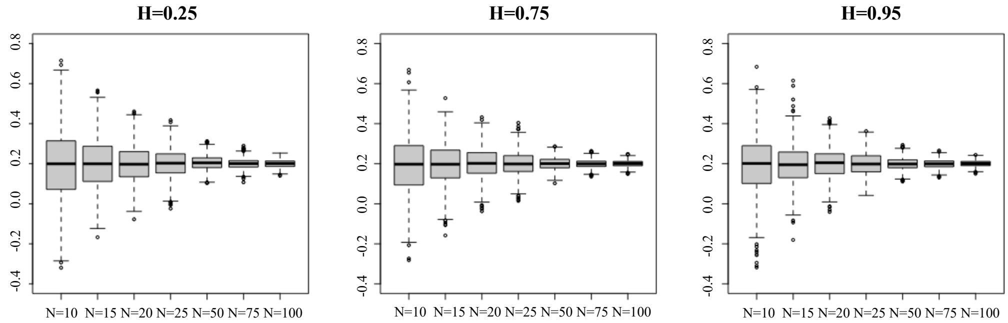
The deterministic case: We consider the model defined by equation (1.1) observed at equally spaced times, that is, $\tau_i = i/N$, for $i = 0, \ldots, N-1$.

The uniform and triangular cases: We review a parametric estimation when the model is observed at random sampling times and considering the uniform and triangular distributions, both with support on $[0, 1/N]$.

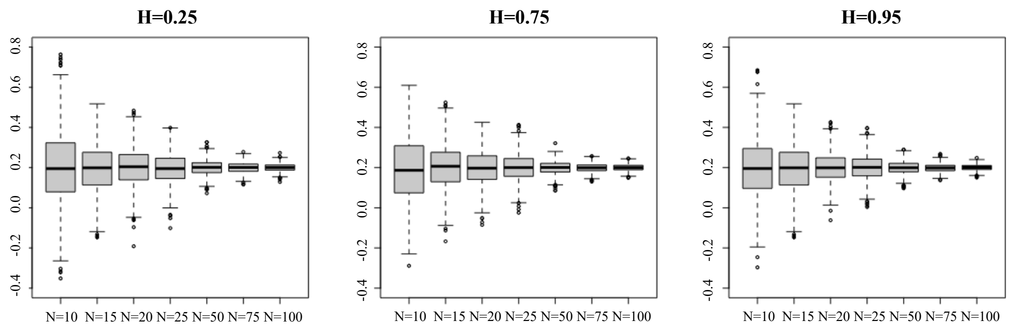
4.1. Small sample size

When considering the jittered sampling scheme, the squared error decreases as the sample size increases. Figures 1 and 2 show box plots of $M = 1{,}000$ trials of $\hat{a}_N$ for $N = 10, 15, 20, 25, 50, 75, 100$, for three different values of H, under three different scenarios (deterministic, uniform, and triangular cases), and for initial values of $a = 0.2$ and $a = 2$.

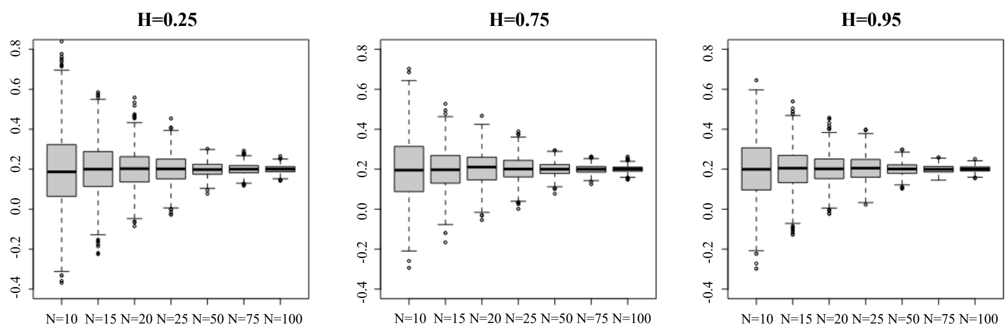
As N increases, the variability of $\hat{a}_N$ around the real value of the parameter, decreases; this situation holds in the same way for the different values of H and for the different cases considered. There is a remarkable reduction of the interquartile range for the three different scenarios considered as the value of N increases. From (3.13), we have that the upper bound of the convergence rate of $|\hat{a}_N - a|$ is C/N^{2H-1} for almost sure convergence, and $1/N^{2H}$ for convergence in probability. This upper bound is obtained for each N. Note that as long as H



(a) Deterministic



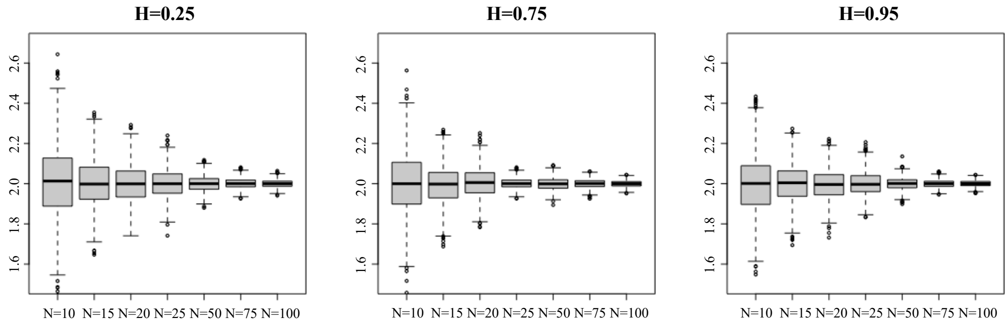
(b) Uniform



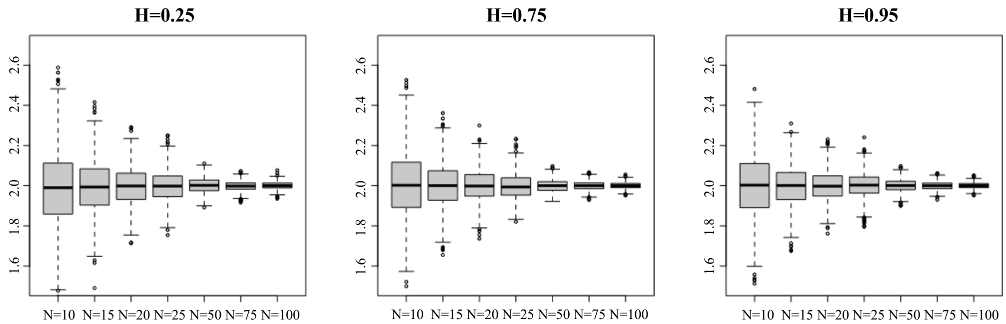
(c) Triangular

Figure 1. Box plots for $\hat{a}_N$ when $N = 10, 15, 20, 25, 50, 75, 100$, for different values of H, under the deterministic, uniform, and deterministic cases, for $a = 0.2$.

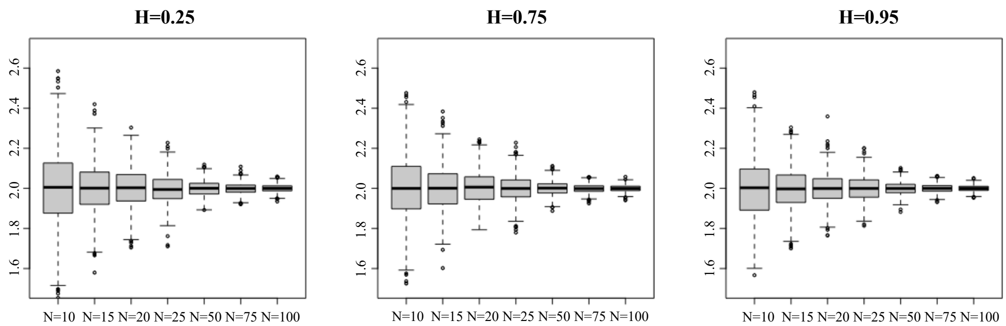
and N increase the estimation of the real parameter is more accurate. Even for small values of N, the estimation is close to the real parameter.



(a) Deterministic



(b) Uniform



(c) Triangular

Figure 2. Box plots for $\hat{a}_N$ when $N = 10, 15, 20, 25, 50, 75, 100$, for different values of H, under the deterministic, uniform, and deterministic cases, for $a = 2$.

4.2. Large sample size

For all the simulations shown, we consider $M = 1{,}000$ replicates of the model with the parameter $a = 0.2$ and $a = 2$, and with different values of N from $N = 3$ to $N = 300$. We also consider different values of the Hurst parameter: $H = 0.05$, $H = 0.25$, and $H = 0.45$ (anti-persistent cases), and $H = 0.55$, $H = 0.75$ and

Table 1. Uniform and triangular cases: Mean, SD, and kurtosis with $a = 0.2$.

Deterministic	$H = 0.05$	$H = 0.25$	$H = 0.45$	$H = 0.55$	$H = 0.75$	$H = 0.95$
Mean	0.1998	0.1997	0.2000	0.2001	0.1997	0.2001
SD	0.0070	0.0061	0.0059	0.0056	0.0053	0.0049
Kurtosis	-0.0653	0.0631	-0.1805	-0.1815	0.0401	-0.0528
Uniform	$H = 0.05$	$H = 0.25$	$H = 0.45$	$H = 0.55$	$H = 0.75$	$H = 0.95$
Mean	0.1996	0.2000	0.1998	0.2003	0.2001	0.1999
SD	0.0069	0.0061	0.0057	0.0054	0.0053	0.0050
Kurtosis	0.0058	-0.0309	-0.2094	-0.2633	-0.0163	-0.2246
Triangular	$H = 0.05$	$H = 0.25$	$H = 0.45$	$H = 0.55$	$H = 0.75$	$H = 0.95$
Mean	2.0002	1.9996	2.0005	1.9997	2.0002	2.0001
SD	0.0070	0.0060	0.0058	0.0059	0.0052	0.0051
Kurtosis	-0.0400	-0.1758	0.0772	0.3426	-0.0475	-0.0785

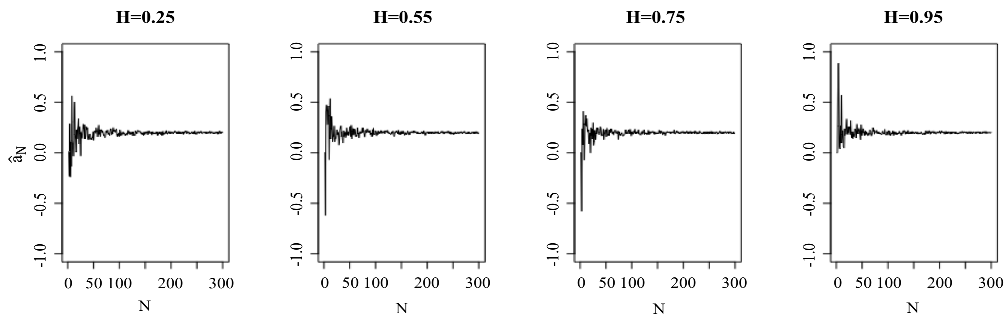
$H = 0.95$ (long-memory cases). For a practical reason, in Figure 3, four values of H are considered for the graphs, in Figure 4 three different values of H have been considered. For the tables, all of the aforementioned values of H are considered.

Figure 3 shows the value of $\hat{a}_N$ for each value of N from 3 to 300. For all the values of H, the true value of the parameter is reached, even when noise driven by an anti-persistent process ($H = 0.25$) is considered. For $a = 0.2$ and $a = 2$, the uniform scheme converges faster to the real parameter. This confirms the discussion on the convergence speed of the different values of H, and shows that the explicit scheme can perform better for some selections of the long-memory parameter. Tables 1 and 2 summarize the simulation results for the estimation according to equation (3.1). The performance statistics presented are the mean, standard deviation (SD), and kurtosis from $M = 1{,}000$ trials for $N = 300$ fixed of $\hat{a}_N$. The kurtosis refers to the difference between the kurtosis of a Gaussian distribution and that of the random variable a_N.

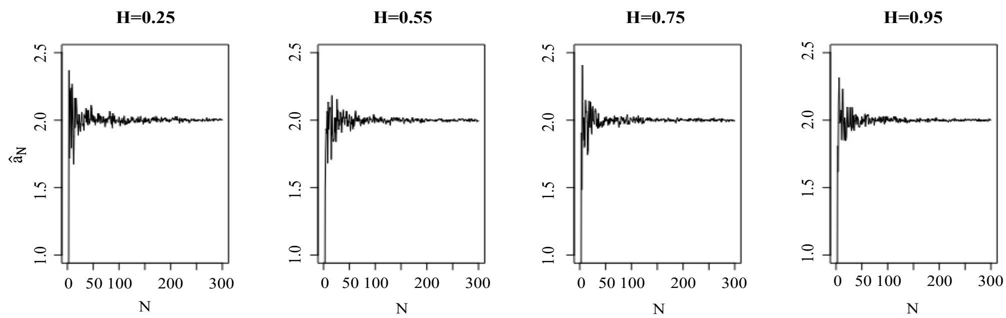
Overall, SD decreases as the value of H increases for all H, which is expected because, in the context of long-range dependence processes, it is quite common for the process to be less noisy. This is a reflection of the consistency of the estimator. On the other hand, when H increases towards one, the empirical estimation exhibits better behavior, which is reflected in a more accurate solution.

Figure 4 shows the frequency histograms (sampling distribution) of $M = 1{,}000$ values of the variable $N(\hat{a}_N - a)$ generated for different values of H. We take the values $a = 0.2$ and $a = 2$, and the random times follow uniform and triangular distributions, respectively.

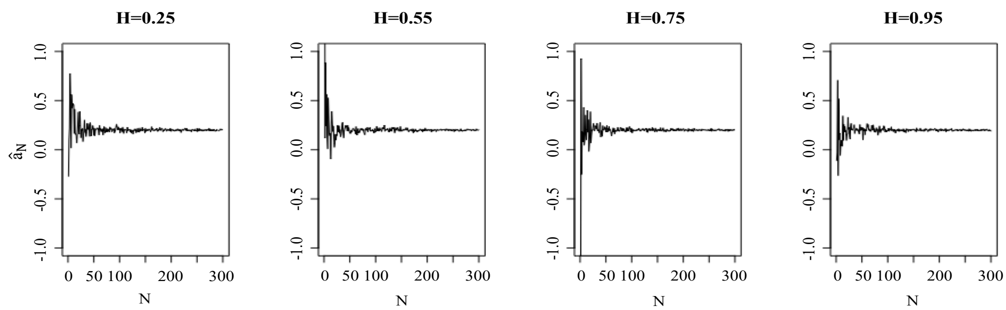
The empirical (solid line) and theoretical (dotted line) distributions overlap. The theoretical distribution considered is a normal distribution with parameters



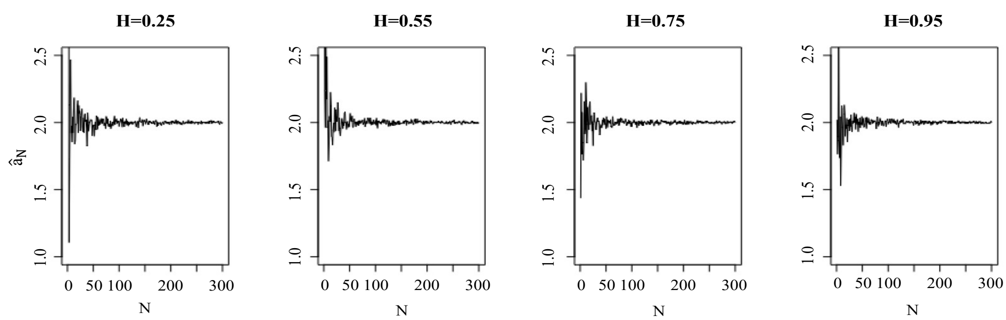
(a) Uniform, $a = 0.2$.



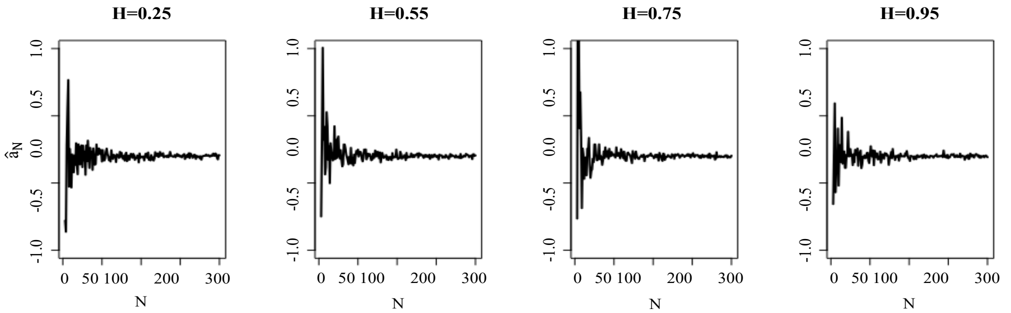
(b) Uniform, $a = 2$.



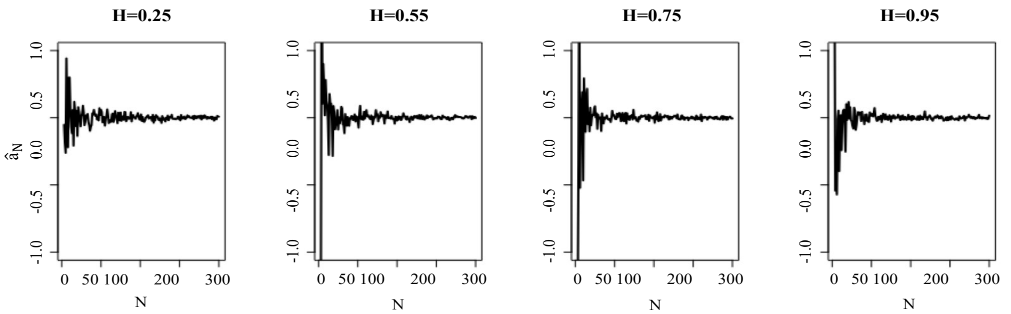
(c) Triangular, $a = 0.2$.



(d) Triangular, $a = 2$.



(e) Deterministic, $a = 0.2$.



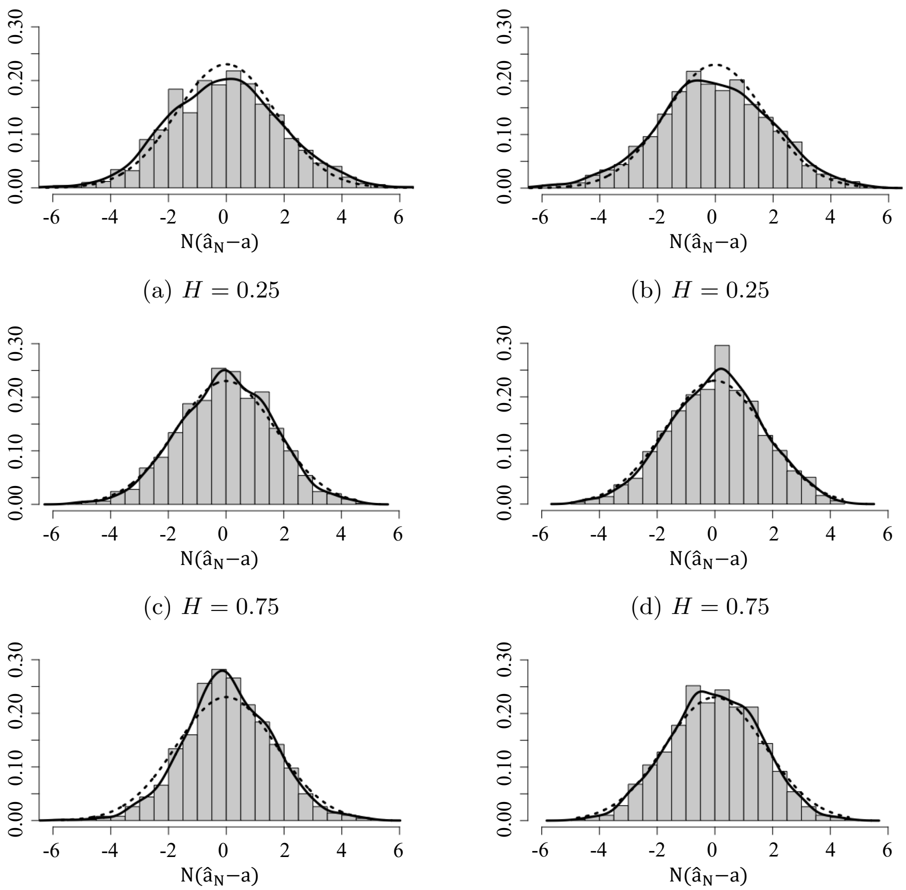
(f) Deterministic, $a = 2$.

Figure 3. Rate of convergence of $\hat{a}_N$ under deterministic, uniform, and triangular distributions and different values of H: (left) case $a = 0.2$, (right) case $a = 2$.

Table 2. Uniform and triangular cases: Mean, SD, and kurtosis with $a = 2$.

Deterministic	$H = 0.05$	$H = 0.25$	$H = 0.45$	$H = 0.55$	$H = 0.75$	$H = 0.95$
Mean	2.000	2.000	2.000	1.9998	2.0000	1.9998
SD	0.0069	0.0062	0.0060	0.0057	0.0054	0.0050
Kurtosis	-0.0804	-0.0948	-0.0326	0.1986	0.0797	-0.0822
Uniform	$H = 0.05$	$H = 0.25$	$H = 0.45$	$H = 0.55$	$H = 0.75$	$H = 0.95$
Mean	1.9999	2.0004	2.0003	2.0001	2.0002	2.0001
SD	0.0069	0.0061	0.0060	0.0056	0.0054	0.0049
Kurtosis	0.0401	-0.1856	-0.2662	0.0499	-0.0646	0.0181
Triangular	$H = 0.05$	$H = 0.25$	$H = 0.45$	$H = 0.55$	$H = 0.75$	$H = 0.95$
Mean	2.0002	1.9996	2.0005	1.9997	2.0002	2.0001
SD	0.0070	0.0060	0.0058	0.0059	0.0052	0.0051
Kurtosis	-0.0400	-0.1758	0.0772	0.3426	-0.0475	-0.0785

$\mu = 0$ and $\sigma^2 = 3$. The value of σ^2 comes from Roa, Torres and Tudor (2021), where the authors consider the case $H = 1/2$, the Brownian motion, and prove that in this case, $N(\hat{a}_N - a)$ converges in distribution to $\mathcal{N}(0, 3)$. Note that com-



(a) $H = 0.25$

(b) $H = 0.25$

(c) $H = 0.75$

(d) $H = 0.75$

(e) $H = 0.95$
Uniform distribution

(f) $H = 0.95$
Triangular distribution

Figure 4. (Left) Histograms for $N(\hat{a}_N - 0.2)$ under the uniform distribution and different values of H. (Right) Histograms for $N(\hat{a}_N - 2)$ under the triangular distribution and different values of H.

puting the asymptotic distributions in a general case is challenging, and thus left to future work. In conclusion, we have shown that the empirical estimation of the LS estimator in models driven by long-memory noise under the random scheme described by (2.2) guarantees stability and convergence, and is very accurate for any value $H \in (0, 1)$ and different values of a. Furthermore, when H increases toward one the estimator presents less variability. The same conclusion can be drawn when N approaches infinity. Therefore, the estimation procedure presented here is a good alternative for estimating parameters in a linear regression model with random times and long-range dependent noise.

Acknowledgements

This research was partially supported by Project MATHAMSUD 19-MATH-06 and 18-MATH-07. H. Araya was partially supported by Proyecto FONDECYT Post-Doctorado 3190465, N. Bahamonde was partially supported by FONDECYT Regular 1201898, T. Roa was partially supported by Beca CONICYT-PFCHA/Doctorado Nacional/2018-21180298, and S. Torres was partially supported by FONDECYT 1171335.

Appendix

A. Proof of Lemma 1

Proof. By definition of D_N given in equation (3.2), we have

$$\begin{aligned} D_N &= \frac{1}{N}\sum_{i=0}^{N-1}\left(\frac{i}{N}+\nu_{i,N}\right)^2 \\ &= \frac{1}{N}\sum_{i=0}^{N-1}\frac{i^2}{N^2}+\frac{2}{N}\sum_{i=0}^{N-1}\frac{i\nu_{i,N}}{N}+\frac{1}{N}\sum_{i=0}^{N-1}\nu_{i,N}^2 \\ &= I_N^{(1)}+I_N^{(2)}+I_N^{(3)}. \end{aligned}$$

First, $I_N^{(1)} = (1/N^3)\sum_{i=0}^{N-1} i^2 = (2N^3-3N^2+N)/6N^3$. Then,

$$\lim_{N\to\infty} I_N^{(1)} = \frac{1}{3}. \tag{A.1}$$

Now for $I_N^{(2)}$, let $\gamma > 0$ we use the fact that $0 \le \nu_{i,N} \le 1/N$ for all $i = 0,\dots,N-1$. Now, we will apply Chebyshev inequality as follows:

$$\begin{aligned} \mathbb{P}\left(\left|\frac{2}{N^2}\sum_{i=0}^{N-1} i\nu_{i,N}\right| > \frac{1}{N^\gamma}\right) &\le \frac{1}{N^{-2\gamma}}\mathbb{E}\left[\left(\frac{2}{N^2}\sum_{i=0}^{N-1} i\nu_{i,N}\right)^2\right] \\ &= \frac{4}{N^{4-2\gamma}}\mathbb{E}\left[\left(\sum_{i=0}^{N-1} i\nu_{i,N}\right)^2\right] \\ &\le \frac{4}{N^{4-2\gamma}}\left(\sum_{i=0}^{N-1}\frac{i}{N}\right)^2 \\ &= \frac{4}{N^{4-2\gamma}}\left(\frac{N-1}{2}\right)^2 \end{aligned}$$

$$\leq \frac{1}{N^{2-2\gamma}}.$$

Given that $\nu_{i,N}$ is a sequence of events in a probability space we are in position to use Borel Cantelli Lemma; we need to find a strictly positive γ, so that

$$\sum_{N\geq 1} \mathbb{P}\left(\left| \frac{2}{N^2} \sum_{i=0}^{N-1} i\nu_{i,N} \right| > \frac{1}{N^\gamma} \right) \leq \sum_{N\geq 1} \frac{1}{N^{2-2\gamma}} < \infty,$$

to ensure the convergence of the previous sum, it is necessary to find a value for γ such that $2 - 2\gamma > 1$, for $0 < \gamma < 1/2$, then

$$I_N^{(2)} = \frac{2}{N} \sum_{i=0}^{N-1} \frac{i\nu_{i,N}}{N} \xrightarrow[N\to\infty]{a.s.} 0. \tag{A.2}$$

For the third term $I_N^{(3)}$, we take into account that $\nu_{i,N} \in [0, 1/N]$ for all $i = 0, \ldots, N-1$. Consequently

$$\frac{1}{N} \sum_{i=0}^{N-1} \nu_{i,N}^2 \leq \frac{1}{N} \sum_{i=0}^{N-1} \frac{1}{N^2} = \frac{1}{N^2} \xrightarrow[N\to\infty]{a.s.} 0. \tag{A.3}$$

Finally, by (A.1), (A.2) and (A.3) the result is achieved.

References

Araya, H., Bahamonde, N., Fermin, L., Roa, T. and Torres, S. (2023). On the consistency of the least squares estimator in models sampled at random times driven by long memory noise: The renewal case. *Statistica Sinica* **33**.

Baillie, R. T. and Chung, S. K. (2002). Modeling and forecasting from trend-stationary long memory models with applications to climatology. *International Journal of Forecasting* **18**, 215–226.

Bardet, J. M. and Bertrand, P. R. (2010) . A non-parametric estimator of the spectral density of a continuous-time Gaussian process observed at random times. *Scandinavian Journal of Statistics* **37**, 458–476.

Bellhouse, D. R. (1981). Area estimation by point-counting techniques. *Biometrics* **37**, 303–312.

Brockwell, A. E. (2007). Likelihood-based analysis of a class of generalized long-memory time series models. *Journal of Time Series Analysis* **28**, 386–407.

Chang, C. (2014). Optimum preventive maintenance policies for systems subject to random working times, replacement, and minimal repair. *Computers & Industrial Engineering* **67**, 185–194.

Khan, M. (2017). Performance testing of computed radiography system and imaging plates. In *ASNT Annual Conference 2017*, 83–92. American Society for Nondestructive Testing, Columbus.

Krune, E., Krueger, B., Zimmermann, L., Voigt, K. and Petermann, K. (2016). Comparison of the jitter performance of different photonic sampling techniques. *Journal of Lightwave Technology* **34**, 1360–1367.

Lobato, I. and Velasco, C. (2000). Long memory in stock-market trading volume. *Journal of Business & Economic Statistics* **18**, 410–427.

Mandelbrot, B. and Van Ness, J. (1968). Fractional brownian motions, fractional noises and applications. *SIAM Review* **10**, 422–437.

Masry, E. (1983). Probability density estimation from sampled data. *IEEE Transactions on Information Theory* **29**, 696–709.

Max-Moerbeck, W., Richards, J. L., Hovatta, T., Pavlidou, V., Pearson, T. J. and Readhead, A. C. S. (2014). A method for the estimation of the significance of cross-correlations in unevenly sampled red-noise time series. *Monthly Notices of the Royal Astronomical Society* **445**, 437–459.

Nieto-Barajas, L. and Sinha, T. (2015). Bayesian interpolation of unequally spaced time series. *Stochastic Environmental Research and Risk Assessment* **29**, 577–587.

Ólafsdóttir, K. B., Schulz, M. and Mudelsee, M. (2016). REDFIT-X: Cross-spectral analysis of unevenly spaced paleoclimate time series. *Computers & Geosciences* **91**, 11–18.

Philippe, A., Robet, C. and Viano, M. (2020). Random discretization of stationary continuous time processes. *Metrika* **84**, 375–400.

Roa, T., Torres, S. and Tudor, C. (2021). Limit distribution of the least square estimator with observations sampled at random times driven by standard Brownian motion. *Communications in Statistics-Theory and Methods*. DOI: 10.1080/03610926.2021.1980044.

Subr, K., Nowrouzezahrai, D., Jarosz, W., Kautz, J. and Mitchell, K. (2014). Error analysis of estimators that use combinations of stochastic sampling strategies for direct illumination. *Computer Graphics Forum* **33**, 93–102.

Tudor C. A. (2008). Analysis of the Rosenblatt process. *ESAIM: Probability and Statistics* **12**, 230–257.

Tudor C. A. (2013). *Analysis of Variations for Self-Similar Processes: A Stochastic Calculus Approach.* Springer, Cham.

Vilar, J. A. (1995). Kernel estimation of the regression function with random sampling times. *Test* **4**, 137–178.

Vilar, J. A. and Vilar, J. M. (2000). Finite sample performance of density estimators from unequally spaced data. *Statistics and Probability Letters* **50**, 63–73.

Zhao, X., Chen, M. and Nakagawa, T. (2014). Optimal time and random inspection policies for computer systems. *Applied Mathematics and Information Sciences* **8**, 413–417.

Héctor Araya

Facultad de Ingeniería y Ciencias, Universidad Adolfo Ibáñez, Viña del Mar, 252001, Chile.

E-mail: hector.araya@uai.cl

Natalia Bahamonde

Instituto de Estadística, Pontificia Universidad Católica de Valparaíso, Valparaíso, 2340031, Chile.

E-mail: natalia.bahamonde@pucv.cl

Lisandro Fermín

CIMFAV, Instituto de Ingeniería Matemática, Facultad de Ingeniería, Universidad de Valparaíso, Valparaíso, 2362905, Chile.

E-mail: lisandro.fermin@uv.cl

Tania Roa

Facultad de Ingeniería y Ciencias, Universidad Adolfo Ibáñez, Viña del Mar, 252001, Chile.

E-mail: tania.roa@uai.cl

Soledad Torres

CIMFAV, Instituto de Ingeniería Matemática, Facultad de Ingeniería, Universidad de Valparaíso, Valparaíso, 2362905, Chile.

E-mail: soledad.torres@uv.cl

(Received March 2020; accepted June 2021)