Statistica Sinica **18**(2008), 559-579

CONFIDENCE INTERVALS FOR THE LONG MEMORY PARAMETER BASED ON WAVELETS AND RESAMPLING

Pier Luigi Conti, Livia De Giovanni, Stilian A. Stoev and Murad S. Taqqu

Università di Roma "La Sapienza", Università degli Studi LUMSA, University of Michigan and Boston University

Abstract: We study the problem of constructing confidence intervals for the long-memory parameter of stationary Gaussian processes with long-range dependence. The focus is on confidence intervals for the wavelet estimator introduced by Abry and Veitch(1998). We propose an approximation to the distribution of the estimator based on subsampling and use it to construct confidence intervals for the long-memory parameter. The performance of these confidence intervals, in terms of both coverage probability and length, is studied by using a Monte Carlo simulation. The proposed confidence intervals have more accurate coverage probability than the method of Veitch and Abry (1999), and are easy to compute in practice.

Key words and phrases: Hurst parameter, long-range dependence, resampling, wavelets.

1. Introduction and Motivation

An estimator of the long-memory parameter of stationary processes with long-range dependence (LRD processes, for short) based on a wavelet decomposition has been proposed and studied in Abry and Veitch (1998), Veitch and Abry (1999) and Veitch, Abry and Taqqu (2003). To obtain the estimator, one performs a discrete wavelet transform and a linear regression. The slope of the linear regression is related to the long-memory parameter and the intercept is related to the variance of the process. The resulting estimator will be called the (linear) wavelet estimator.

Under the idealized assumption of independence of the wavelet coefficients, the estimator of the long-memory parameter is unbiased and of minimum (or close to minimum) variance, and the estimator of the intercept is asymptotically unbiased and efficient (Veitch and Abry (1999)). Abry and Veitch suggest using confidence intervals for the long-memory parameter of stationary LRD processes based on these results.

In reality the wavelet coefficients are not independent. Bardet, Lang, Moulines and Soulier (2000) have performed statistical analysis of these linear wavelet estimators by taking into account the correlations among wavelet coefficients, at fixed scales as well as between different scales. Consistency and asymptotic normality of these estimators have been obtained under appropriate regularity conditions on the spectral density of the process. Unfortunately, constructing confidence intervals based on these results is difficult in practice. This is so because the resulting asymptotic variances depend on the unknown value of the long-memory parameter and, even if their values were known, evaluation is rather cumbersome. Thus, while the results of Abry and Veitch are easy to apply, they are only based on idealized assumptions. While the results of Bardet et al. (2000) are more precise, they cannot be well-implemented in practice.

We investigate here an approach, based on subsampling, that provides a practical way of computing confidence intervals with asymptotically correct coverage probabilities. The idea behind subsampling is as follows. Let θ be an unknown parameter, estimated by $\widehat{\theta}_n$, and let $Q_n(x) = P(\widehat{\theta}_n \leq x)$ denote the distribution function of $\widehat{\theta}_n$. One would like to construct confidence intervals for θ using $Q_n(x)$, but $Q_n(x)$ is unknown. One uses instead the empirical distribution function $\widehat{Q}_n(x)$ obtained by splitting the sample size of size n into N overlapping subsamples of length l_n, and using the estimates $\widehat{\theta}_{l_n,i}$, $i = 1, \ldots, N$, based on these subsamples.

Our goal is to study the resampling distribution of the linear wavelet estimator (namely the estimator of Abry and Veitch), based on the subsampling (sampling window) method (Hall, Jing and Lahiri (1998) and Lahiri (2003)). We show that such a resampling distribution is consistent. Consequently, confidence intervals based on this empirical distribution have asymptotically correct coverage probability. We then compare the performance of the confidence intervals based on the results of Abry and Veitch with the ones based on our resampling distribution technique. We do so by using Monte Carlo simulations. The Abry and Veitch type confidence intervals tend to have too small coverage probabilities. On the other hand, the ones based on resampling have essentially accurate coverage probabilities, that are typically slightly larger than the nominal values.

The paper is organized as follows. In Section 2, we briefly review the notion of long-range dependence (LRD) for stationary processes. The linear wavelet-based estimator of Abry and Veitch is described in Section 3. In Section 4, we study an approximation of the probability distribution of the linear wavelet estimator in terms of subsampling. Section 5 contains the proofs of our results. Finally, in Section 6, we present and discuss Monte Carlo simulation results.

2. Self-similarity and Long-memory

An important class of long-memory stationary processes is derived from self-similar processes. A continuous parameter stochastic process $(Z(t);\ t \geq 0)$ is self-similar, with self-similarity parameter (Hurst parameter) $0 < H < 1$, if for any positive real number c, the process $(c^{-H}Z(ct);\ t \geq 0)$ has the same finite-dimensional distributional distributions as the original process $(Z(t);\ t \geq 0)$; see, e.g., Beran (1994), Samorodnitsky and Taqqu (1994), Taqqu (2003), Embrechts and Maejima (2003). In other words, self-similar processes are invariant in distribution under judicious scaling of time and space.

A discrete parameter (second order) stationary process $(X_j;\ j \geq 1)$, with variance $\sigma^2 = E[X_j^2] - [EX_j]^2$, is a long-range dependent (LRD, or long-memory) process if its correlation coefficients $\rho(k) = \sigma^{-2}\{E[X_j\, X_{j+k}] - [EX_j]^2\}$ take the form $\rho(k) \sim c_r\, k^{-(1-\alpha)}$, for some $0 < \alpha < 1$, as k goes to infinity, c_r being an absolute constant (Beran (1994) and Taqqu (2003)).

Suppose that the original process $(Z(t);\ t \geq 0)$ has stationary increments and is self-similar with Hurst parameter H. If $1/2 < H < 1$, then its increments $(X_j = Z(j) - Z(j-1);\ j \geq 1)$ form a LRD process, with correlation coefficients $\rho(k) \sim c_r\, k^{-2(1-H)}$ as k tends to infinity. On the other hand, if $H = 1/2$, the $\rho(k)$'s are zero for every $k \geq 1$, and thus the process is short-range dependent. When $1/2 < H < 1$, the Hurst parameter H and the long-memory parameter α are linked by the relation:

$$H = \frac{1+\alpha}{2}. \tag{1}$$

An equivalent definition of long-range dependence involves the spectral density $f(\nu)$ of the process $X = (X_j;\ j \geq 1)$. Namely, X is LRD if $f(\nu)$ is unbounded at the origin and $f(\nu) \sim c_f\, |\nu|^{-\alpha}$ as ν tends to zero, where c_f a positive constant and where $0 < \alpha < 1$ (see, e.g., Beran (1994) and Samorodnitsky and Taqqu (1994)).

3. Wavelet-based Estimator of the Long-memory Parameter

The discrete wavelet coefficients of a function $x(t) \in L^2(\mathbb{R})$ are

$$d_{jk} = \int_{-\infty}^{+\infty} x(t)\, \psi_{jk}(t)\, dt, \quad j,\, k \in \mathbb{Z}, \tag{2}$$

where the $\psi_{jk}(t) \in L^2(\mathbb{R})$ are basis elements called wavelet functions. The set of wavelets ψ_{jk} is derived from a single function $\psi_0(t)$, the mother wavelet, via scaling and dilation:

$$\psi_{jk}(t) = 2^{-j/2}\, \psi_0(2^{-j}\, t - k), \quad j, k \in \mathbb{Z}.$$

The coefficients $d_{jk},\ j, k \in \mathbb{Z}$ are referred to as the Discrete Wavelet Transform (DWT) of $x(t)$.

In applications, one typically uses Daubechies$-M$ mother wavelets, which have $M \geq 1$ zero moments:

$$\int_{-\infty}^{+\infty} t^s\, \psi_0(t)\, dt = 0, \quad s = 0,\, 1,\, \ldots,\, M-1,$$

see Daubechies (1992). The corresponding wavelet functions ψ_{jk}, $j, k \in \mathbb{Z}$, are an orthogonal basis of $L^2(\mathbb{R})$. The Daubechies mother wavelets are continuous except in the case $M = 1$, where they reduce to the Haar wavelet $\psi_0(t) = I_{(0 \leq t < 1/2)} - I_{(1/2 \leq t < 1)}$. They also have bounded support, and the greater the M, the wider the support. In contrast with the trigonometric functions used in Fourier series, which are perfectly localized in frequency but non-localized in time, the Daubechies wavelets are well-localized in time and approximately localized in frequency. Thus, the wavelet coefficients capture both time- and frequency-domain characteristics of the function $x(t)$.

The wavelet coefficients of a second order long-memory process have received considerable attention, because their special properties can be used to estimate the long-memory parameter α.

The wavelet transform in (2) is defined only for continuous-time processes. In the LRD setting, Veitch, Abry and Taqqu (2000) propose the following natural procedure. Given a discrete parameter stationary long-memory process $(X_j;\ j \geq 1)$, construct a continuous-time process

$$\widetilde{X}(t) = \sum_n X_n \operatorname{sinc}(t - n),$$

where $\operatorname{sinc}(t) = (\pi t)^{-1} \sin \pi t$. The spectral densities of the processes (X_j) and $(\widetilde{X}(t))$ coincide on the interval $(-1/2,\, 1/2]$, and moreover $X_n = \widetilde{X}(n)$ for every integer n. Thus, the long-range dependence behavior of the two processes is identical. In the sequel, we define the wavelet transform of the discrete-time process X as that of its continuous-time counterpart $\widetilde{X}$. This corresponds to using the initialization option in the `LDestimate` software of Abry and Veitch (see http://www.cubinlab.ee.mu.oz.au/~darryl/LDestimate_code.tar.gz).

The wavelet coefficients of a discrete parameter (second order) stationary long-memory process $(X_j;\ j \geq 1)$ have the following two properties (see Flandrin (1992), Masry (1993), Tewfik and Kim (1992), Abry, Flandrin, Taqqu and Veitch (2003)).

P1 If $M \geq (\alpha - 1)/2$, then the wavelet coefficients $(d_{jk};\ k = 0, \pm 1, \ldots)$ at a fixed scale j form a stationary process with $E[d_{jk}]= 0$ and

$$E[d_{jk}^2] \sim 2^{j\alpha}\, c_f\, C(\alpha,\, \psi_0) \tag{3}$$

as j tends to infinity. The quantity $C(\alpha,\,\psi_0)$ is defined as

$$C(\alpha,\,\psi_0) = \int_{-\infty}^{+\infty} \nu^{-\alpha}\,|\Psi_0(\nu)|^2\,d\nu,$$

where $\Psi_0(\nu)$ is the Fourier transform of the mother wavelet $\psi_0(t)$ (Abry et al. (2003)).

P2 If $M \geq \alpha/2$, then $E[d_{jk}\,d_{jk'}]$ goes to 0 as $|k-k'|$ tends to infinity and $\sum_{k=0}^{\infty}|Ed_{jk}d_{j0}| < \infty$, (Flandrin (1992), Masry (1993) and Tewfik and Kim (1992)). This indicates that $(d_{jk};\,k=0,\,\pm 1,\,\dots)$ has short-range dependence.

In view of **P1**, Veitch and Abry (1999) propose to estimate the pair $(\alpha,\,c_f)$ by a (weighted) linear regression of $\log_2 E[d_{jk}^2]$ on $\log_2 2^j = j$. Since

$$\log_2 E[d_{jk}^2] \sim j\alpha + \log_2 c_f C(\alpha,\,\psi_0), \tag{4}$$

the slope of this regression provides an estimate of α, and its intercept is related to $c_f\,C(\alpha,\,\psi_0)$

The quantity $\log_2 E[d_{jk}^2]$ is estimated by its moment estimator

$$\log_2\left(\frac{1}{\nu_j(n)}\sum_{k=1}^{\nu_j(n)} d_{jk}^2\right) - g_j,$$

where $\nu_j(n) = n/2^j$ is the number of available coefficients at scale j, and g_j is a correction term to reduce bias, since $\log_2(E[\,\cdot\,]) \neq E[\log_2(\cdot)]$ (see Abry et al. (2003, p.541)). The linear regression becomes

$$E[Y_j] = \theta + \alpha\,j, \tag{5}$$

where

$$Y_j = \log_2\left(\frac{1}{\nu_j(n)}\sum_{k=1}^{\nu_j(n)} d_{jk}^2\right) - g_j, \quad \theta = \log_2\left(c_f\,C(\alpha,\,\psi_0)\right). \tag{6}$$

The plot of Y_j against the scale index j is called the log-scale diagram.

Relation (5) is only approximate since (4) is an asymptotic relation. It can hold already, however, for small values of j, for example, when the time series (X_j) is obtained as the increment of a self-similar process with stationary increments, as indicated in Section 2.

In practice, if the sample size is n, (5) is used for all $j \geq j_1(n)$, where $j_1(n)$ is the lowest scale where $E\,Y_j$ is approximately linear in j. Precise asymptotic assumptions on $j_1(n)$ are given below.

In order to apply the weighted least squares method to estimate θ and α, and to study the statistical properties of the resulting estimators, Veitch and Abry (1999) make the following supplementary simplifying assumptions.

C1 The process $(X_j;\ j \geq 1)$ is a Gaussian stochastic process. As a consequence, the process $(d_{jk};\ j \geq 1;\ k = 0, \pm 1, \pm 2, \ldots)$, obtained through the linear transformation (2), is Gaussian.

C2 The random variables d_{jk}s, for every fixed scale index j, are independent and identically distributed.

C3 The processes $(d_{jk};\ k = 0, \pm 1, \pm 2, \ldots)$ and $(d_{j'k};\ k = 0, \pm 1, \pm 2, \ldots)$ are independent for every $j \neq j'$.

The variances, σ_j^2, of the Y_j have been computed explicitly by Veitch and Abry (1999) under assumptions $C1$-$C3$. The weighted least squares estimator of $(\theta,\ \alpha)$ takes the form:

$$\begin{bmatrix} \widehat{\theta}_n \\ \widehat{\alpha}_n \end{bmatrix} = \left(\sum_j \frac{1}{\sigma_j^2} \begin{bmatrix} 1 & j \\ j & j^2 \end{bmatrix} \right)^{-1} \left(\sum_j \frac{y_j}{\sigma_j^2} \begin{bmatrix} 1 \\ j \end{bmatrix} \right).$$

From (6), one obtains an estimator of c_f. The corresponding estimator of the Hurst parameter $H = (1+\alpha)/2$ is $\widehat{H}_n = (1 + \widehat{\alpha}_n)/2$.

Veitch and Abry (1999) have shown that, under $P1$-$P2$ and $C1$-$C3$, the estimators $\widehat{\alpha}_n$ and $\widehat{\theta}_n$ are unbiased, asymptotically efficient, and asymptotically normally distributed. Confidence intervals based on these results have been obtained.

Bardet et al. (2000) have studied the asymptotic properties of the estimator $\widehat{\alpha}_n$ of α without assuming $C2$ and $C3$, which are approximations only. They proved consistency and asymptotic normality under some regularity conditions on the spectral density $f(\lambda) = \lambda^{1-2H} f_*(\lambda)$, as long as (i) the scale index $j_1(n)$ goes sufficiently fast to infinity (see (7) below) with the sample size n, and (ii) $n^{-1}\, 2^{j_1(n)}$ goes to zero. To describe the regularity conditions on the spectral function f, denote by

$$\Delta_h^r g(x) = \sum_{k=0}^{r} \binom{r}{k} (-1)^{r-k} g(x + kh)$$

the rth order differences of a function g. The *generalized Lipschitz space* $Lip^*(\alpha,\ p)$ is the space of functions in $L^p(\mathbb{R})$ such that

$$\|\Delta_h^r g(\cdot)\|_p = \left(\int |\Delta_h^{\lceil \alpha \rceil} g(x)|^p \, dx \right)^{\frac{1}{p}} \leq M\, h^{\alpha},$$

for some positive M, where $r = \lceil \alpha \rceil$ denotes the smallest integer bigger than α. On the space $Lip^*(\alpha,\, p)$, take

$$|g|_{Lip^*(\alpha, L_p)} = \sup_{t>0} \sup_{0<h\le t} \frac{1}{t^\alpha} \|\Delta_h^r g(\cdot)\|_p .$$

The regularity conditions on the spectral density f are as follows.

$S1$. For some $1 < \delta \le 2$, $f_* \in Lip^*(1,\, 1) \cap Lip^*(\delta,\, 1) \cap L^\infty(\mathbb{R})$;

$S2$. $f_*(0) \neq 0$;

$S3$. there are positive real constants η, β, β', and a real constant $f_*^{(\beta)}(0) \neq 0$ such that $1 < \beta < \beta'$ and $\sup_{0<\lambda\le\eta} |f_*(\lambda) - f_*(0) - f_*^{(\beta)}(0)\lambda^\beta|/\lambda^{\beta'} < \infty$;

$S4$. there exist $\epsilon > 0$, $0 \le \gamma < 1$, and $C(f_*')$ such that f_* is differentiable on $(0, \epsilon)$ with $|f_*'(\lambda)| \le C(f_*')/\lambda^\gamma$.

The main result of Bardet et al. (2000) (stated in terms of $\widehat{H}_n$ and H instead of $\widehat{\alpha}_n$ and α) can be formulated as follows. If

$$\lim_{n\to\infty} 2^{-j_1(n)(1+2\beta)} n = 0 \quad \text{and} \quad \lim_{n\to\infty} \nu_{j_1(n)}(n) = \lim_{n\to\infty} 2^{-j_1(n)} n = \infty, \tag{7}$$

then

$$\sqrt{\nu_{j_1(n)}(n)}\, (\widehat{H}_n - H) \xrightarrow{d} N(0,\, \sigma^2(H)) \quad as\ n \to \infty, \tag{8}$$

where the symbol $N(\mu,\, \sigma^2)$ denotes a normal distribution with mean μ and variance σ^2 and $\nu_{j_1(n)}(n)$ denotes the number of wavelet coefficients available at scale $j_1 = j_1(n)$. To simplify notation we shall often write $\nu_{j_1}(n)$ instead of $\nu_{j_1(n)}(n)$. The asymptotic variance $\sigma^2(H)$ that appears in (8) depends on H. More importantly, one cannot use $\sigma^2(\widehat{H}_n)$ to estimate $\mathrm{Var}\,(\widehat{H}_n)$, because $\mathrm{Var}\,(\widehat{H}_n)$ involves unknown constants that depend on n. The asymptotic confidence intervals, moreover, are based on a Gaussian distribution which may not be valid for small values of $\nu_{j_1(n)}(n)$. For these reasons, it is virtually impossible to obtain an accurate confidence interval for H by using (8) with $\sigma^2(H)$ simply replaced by $\sigma^2(\widehat{H}_n)$. This is why here we propose to use resampling techniques, in particular, the sampling window method (Lahiri (2003)) to obtain confidence intervals for H.

4. Approximating the Distribution of $\widehat{H}_n$ via Subsampling

Denote by

$$T_n = \sqrt{\nu_{j_1(n)}(n)}\, (\widehat{H}_n - H) \tag{9}$$

the "centered version" of the estimators $\widehat{H}_n$, based on n observations from a long-memory Gaussian stationary process $(X_j;\ j \geq 1)$, and denote by

$$Q_n(x) = P(T_n \leq x) \tag{10}$$

its distribution function (d.f.). We assume that $T_n \xrightarrow{d} N(0, \sigma^2(H))$ as $n \to \infty$. This is the case, for example, under the assumptions of Bardet et al. (2000). The distribution $Q_n(x)$ of T_n, for finite n, is unknown. We are going to estimate it by a distribution $\widehat{Q}_n(x)$ defined in the sequel.

Let $B_i = (X_i, \ldots, X_{i+l-1})$, $i = 1, \ldots, N$, be a collection of $N = n - l + 1$ overlapping blocks of length l, for some given integer $l = l_n$ $(1 \leq l \leq n)$. Finally, let $\widehat{H}_{l,i}$ be the estimator of H based on the data in block B_i. Observe that the blocks B_i overlap and that they have, in fact, adjacent starting points.

A "sub-sample copy" of T_n, based on B_i, is given by

$$\widehat{T}_{l,i} = \sqrt{\nu_{j_1(l)}(l)}\, (\widehat{H}_{l,i} - \widehat{H}_n). \tag{11}$$

The subsampling estimator of the distribution function $Q_n(x) = P(T_n \leq x)$ of T_n, based on the sub-samples B_i's, is simply the empirical distribution function (EDF) of the $\widehat{T}_{l,i}$'s in (11), that is,

$$\widehat{Q}_n(x) = \frac{1}{N} \sum_{i=1}^{N} I_{(\widehat{T}_{l,i} \leq x)}, \quad x \in I\!R. \tag{12}$$

While $Q_n(x)$ is unknown, the subsampling distribution $\widehat{Q}_n(x)$ can be computed from the data.

The following result shows that $\widehat{Q}_n$ is a consistent estimator of Q_n. Hall et al. (1998) proved the consistency of the sample mean, and (8) implies that this consistency extends to the estimator $\widehat{H}_n$. The intuitive reason behind the consistency of $\widehat{Q}_n$ is that the $\widehat{H}_{l,i}$'s are functions of eventually non–overlapping blocks of X_j's. Thus, as in the case of the sample mean, the complete regularity of the Gaussian process $(X_j;\ j \geq 1)$ implies the consistency of the empirical distribution $\widehat{Q}_n$. Details are provided in the proof of Proposition 1. Let $\widetilde{\log f}$ be the harmonic conjugate of $\log f$, the logarithm of the spectral density function of the process $(X_j;\ j \geq 1)$. The key assumption is the same as in Hall et al. (1998), namely that $\widetilde{\log f}$, say, possesses a continuous branch on $[-\pi, \pi]$. This implies that the (Gaussian) process $(X_j;\ j \geq 1)$ is completely regular (Ibragimov and Rozanov (1978, pp.178-179)).

Proposition 1. *Assume that:*

(a) *the process $(X_j;\ j \geq 1)$ possesses a spectral function f such that $\widetilde{\log f}$ is continuous on $[-\pi, \pi]$;*

(b) *conditions $S1$-$S4$ hold;*
(c) ψ_0 *has $m \geq 2$ vanishing moments with*

$$(1+|t|)^r \{|\widehat{\psi}_0(t)| + |\widehat{\psi}_0'(t)| + |\widehat{\psi}_0''(t)|\} \leq C_\psi \quad \forall\, x \in \mathbb{R},$$

where $r \geq 2$, C_ψ is an appropriate positive constant, and $\widehat{\psi}_0(x) = \int \exp\{-2i\pi tx\}$ $\psi_0(t)dt$ is the Fourier transform of ψ_0;
(d) $l^{-1} + n^{-1} l$ *tends to zero, $j_1(n)$ goes to infinity, and $\nu_{j_1(n)}(n)/\nu_{j_1(l)}(l)$ tends to zero as n goes to infinity.*
Then

$$\sup_{x \in \mathbb{R}} \left| \widehat{Q}_n(x) - Q_n(x) \right| \xrightarrow{p} 0 \quad \text{as } n \to \infty. \tag{13}$$

Proof. The proof is given in Section 5.

It follows from Proposition 1 that the quantiles of $\widehat{Q}_n$ behave asymptotically as the quantiles of Q_n. More precisely, the following result holds.

Proposition 2. *Assume the same regularity conditions as in Proposition 1, and set $\widehat{Q}_n^{-1}(u) = \inf\{x : \ \widehat{Q}_n(x) \geq u\}$ and $Q_n^{-1}(u) = \inf\{x : \ Q_n(x) \geq u\}$. Then*

$$\widehat{Q}_n^{-1}(u) - Q_n^{-1}(u) \xrightarrow{p} 0 \quad \text{as} \quad n \to \infty,\ \forall\, u \in (0,\, 1). \tag{14}$$

Proof. The proof is given in Section 5.

Propositions 1 and 2 suggest that the distribution function Q_n may be approximated by the EDF $\widehat{Q}_n$ obtained through subsampling, and that the quantile Q_n^{-1} may be approximated by the sample quantile $\widehat{Q}_n^{-1}$. Hence, for every $0 < \gamma < 1$ the interval

$$\left(\widehat{H}_n + \frac{1}{\sqrt{\nu_{j_1(n)}(n)}}\, \widehat{Q}_n^{-1}(\frac{\gamma}{2}),\ \widehat{H}_n + \frac{1}{\sqrt{\nu_{j_1(n)}(n)}}\, \widehat{Q}_n^{-1}(1-\frac{\gamma}{2}) \right) \tag{15}$$

is a confidence interval for H with asymptotic coverage probability $1-\gamma$.

5. Proof of the Propositions

Proof of Proposition 1. The proof is based on ideas similar to those in Hall et al. (1998). Define first the quantities

$$\overline{T}_{l,i} = \sqrt{\nu_{j_1(l)}(l)}\, (\widehat{H}_{l,i} - H), \quad i = 1,\, \ldots,\, N;$$

$$\overline{Q}_n(x) = \frac{1}{N} \sum_{i=1}^{N} I_{(\overline{T}_{l,i} \leq x)}, \quad x \in I\!R.$$

Our first goal is to show that

$$\left|\overline{Q}_n(x) - Q_n(x)\right| \xrightarrow{p} 0 \text{ as } n \to \infty, \forall\, x \in I\!R. \tag{16}$$

Since $\overline{T}_{l,i}$ has the same distribution as T_l (see (9)), we have $E[\overline{Q}_n(x)] = Q_l(x)$, and hence

$$E[(\overline{Q}_n(x) - Q_n(x))^2] = \mathrm{Var}\,(\overline{Q}_n(x)) + (Q_l(x) - Q_n(x))^2. \tag{17}$$

Since $Q_n(x)$ is a non-random function which, by (8), converges to a limit, we have for $l = l(n) \to \infty$,

$$|Q_l(x) - Q_n(x)| \to 0 \text{ as } n \to \infty. \tag{18}$$

Moreover, we may express the variance in (17) as

$$\begin{aligned}
&\mathrm{Var}\,(\overline{Q}_n(x)) \\
&\quad = \frac{1}{N^2}\sum_{i=1}^{N}\mathrm{Var}\left(I_{(\overline{T}_{l,i}\le x)}\right) + \frac{2}{N^2}\sum_{i=1}^{N}\sum_{j>i}\mathrm{Cov}\left(I_{(\overline{T}_{l,i}\le x)}, I_{(\overline{T}_{l,j}\le x)}\right) \\
&\quad = \frac{1}{N}\mathrm{Var}\left(I_{(\overline{T}_{l,1}\le x)}\right) + \frac{2}{N^2}\sum_{k=1}^{N-1}(N-k)\,\mathrm{Cov}\left(I_{(\overline{T}_{l,1}\le x)}, I_{(\overline{T}_{l,k+1}\le x)}\right) \\
&\quad \le \frac{1}{N}\mathrm{Var}\left(I_{(\overline{T}_{l,1}\le x)}\right) + \mathrm{Var}\left(I_{(\overline{T}_{l,1}\le x)}\right)\frac{2}{N^2}\sum_{k=1}^{l}(N-k) \\
&\qquad + \frac{2}{N^2}\sum_{k=l+1}^{N-1}(N-k)\left|\mathrm{Cov}\left(I_{(\overline{T}_{l,1}\le x)}, I_{(\overline{T}_{l,k+1}\le x)}\right)\right| \\
&\quad \le \left(\frac{1}{N} + \frac{2\,l}{N}\right)\mathrm{Var}\, I_{(\overline{T}_{l,1}\le x)} + \frac{2}{N}\sum_{k=l+1}^{N-1}\left|E\left[I_{(\overline{T}_{l,1}\le x)}\, I_{(\overline{T}_{l,k+1}\le x)}\right] - Q_l(x)^2\right|.
\end{aligned} \tag{19}$$

Using results in Ibragimov and Rozanov (1978, pp.178-179), the Gaussianity of the process $(X_j;\ j \ge 1)$ and the continuity of $\widetilde{\log f}$ imply that $(X_j;\ j \ge 1)$ is completely regular. Hence, denoting by $\mathcal{F}(i\ j)$ the set of functions Y measurable w.r.t. the σ-field $\mathcal{A}(i\ j)$ generated by $X_i, \ldots, X_j$ (possibly with $j = \infty$), with $EY = 0,\ EY^2 = 1$, from Theorem V.5.7 and relationships (IV.1.9), (IV.1.16) in Ibragimov and Rozanov (1978), we have

$$\begin{aligned}
&\sup_{A_1\in\mathcal{A}(1\ l),\, A_2\in\mathcal{A}(l+h\ \infty)} |P(A_1\cap A_2) - P(A_1)\,P(A_2)| \\
&\quad \le \sup_{Y_1\in\mathcal{F}(1\ l),\, Y_2\in\mathcal{F}(l+h\ \infty)} |EY_1\,Y_2| \to 0
\end{aligned}$$

as $h \to \infty$. From this result, it follows that the second term in (19) tends to zero as N goes to infinity. Indeed,

$$S_h := \sup_{k \geq h} \left| E\left[I_{(\overline{T}_{l,1} \leq x)} I_{(\overline{T}_{l,k+1} \leq x)} \right] - Q_l(x)^2 \right| \to 0$$

as $h \to \infty$. Therefore, the second term in (19) is bounded above by $4h/N + 2S_h$, which vanishes as $N \to \infty$, for any $h = h(N)$ such that $h(N)/N \to 0$ and $h(N) \to \infty$ as $N \to \infty$.

Furthermore, all other terms in (19) go to zero as n increases, as a consequence of the assumption $l^{-1} + n^{-1}\, l \to 0$. Thus, $\mathrm{Var}(\overline{Q}_n(x)) \to 0$, as $n \to \infty$, and hence (18) and (17) imply (16).

We now show that

$$\left| \widehat{Q}_n(x) - \overline{Q}_n(x) \right| \xrightarrow{p} 0, \quad \text{as } n \to \infty,\ \forall\, x \in \mathbb{R}. \tag{20}$$

Observe first that

$$\left| \widehat{Q}_n(x) - \overline{Q}_n(x) \right| \leq \frac{1}{N} \sum_{i=1}^{N} \left| I_{\left(\nu_{j_1(l)}(l)^{1/2}(\widehat{H}_{l,i} - \widehat{H}_n) \leq x\right)} - I_{\left(\nu_{j_1(l)}(l)^{1/2}(\widehat{H}_{l,i} - H) \leq x\right)} \right| . \tag{21}$$

and that

$$\begin{aligned} &\left| I_{\left(\nu_{j_1(l)}(l)^{1/2}(\widehat{H}_{l,i} - \widehat{H}_n) \leq x\right)} - I_{\left(\nu_{j_1(l)}(l)^{1/2}(\widehat{H}_{l,i} - H) \leq x\right)} \right| \\ &\quad = \left| I_{\left(\nu_{j_1(l)}(l)^{1/2}(\widehat{H}_{l,i} - H) \leq x + \nu_{j_1(l)}(l)^{1/2}(\widehat{H}_n - H)\right)} - I_{\left(\nu_{j_1(l)}(l)^{1/2}(\widehat{H}_{l,i} - H) \leq x\right)} \right| \\ &\quad \leq A_i + B_i, \end{aligned} \tag{22}$$

where

$$\begin{aligned} A_i &= I_{\left(x < \nu_{j_1(l)}(l)^{1/2}(\widehat{H}_{l,i} - H) \leq x + \nu_{j_1(l)}(l)^{1/2}(\widehat{H}_n - H)\right)}, \\ B_i &= I_{\left(x + \nu_{j_1(l)}(l)^{1/2}(\widehat{H}_n - H) < \nu_{j_1(l)}(l)^{1/2}(\widehat{H}_{l,i} - H) \leq x\right)}. \end{aligned}$$

Now, for every $\epsilon > 0$ and for all $i = 1, \ldots, N$, we have

$$\begin{aligned} EA_i &= P(A_i > 0) \\ &\leq P(\sqrt{\nu_{j_1(l)}(l)} |\widehat{H}_n - H| > \epsilon) + P(x < \sqrt{\nu_{j_1(l)}(l)} (\widehat{H}_{l,i} - H) \leq x + \epsilon) \\ &= P(\sqrt{\nu_{j_1(l)}(l)} |\widehat{H}_n - H| > \epsilon) + P(x < \sqrt{\nu_{j_1(l)}(l)} (\widehat{H}_{l,1} - H) \leq x + \epsilon), \end{aligned} \tag{23}$$

where the equality in (23) follows from $\widehat{H}_{l,i} \stackrel{d}{=} \widehat{H}_{l,1}$, $i = 1, \ldots, N$, since the process $(X_j;\ j \geq 0)$ is stationary.

From (8) and assumption (d), one gets

$$\sqrt{\nu_{j_1(l)}(l)}(\widehat{H}_n - H) = \sqrt{\frac{\nu_{j_1(l)}(l)}{\nu_{j_1(n)}(n)}}\sqrt{\nu_{j_1(n)}(n)}(\widehat{H}_n - H) \xrightarrow{p} 0 \text{ as } n \to \infty,$$

and hence the first term in (23) converges to zero as n increases, for every positive ϵ. As far as the second term in (23) is concerned, (8) implies that

$$\lim_{n\to\infty} P\left(x < \sqrt{\nu_{j_1(l)}(l)}(\widehat{H}_{l,1} - H) \le x + \epsilon\right) = P(x < N(0, \sigma^2(H)) \le x + \epsilon).$$

Thus, as $n \to \infty$ and letting ϵ go to zero, we obtain that $EA_i = P(A_i > 0)$ tends to zero as n increases, uniformly in $i = 1, \ldots, N$. Similarly, $E[B_i] = P(B_i > 0) \to 0$ as $n \to \infty$, uniformly in $i = 1, \ldots, N$. The conclusion (20) follows from (22) and (21).

In view of (16) and (20), the difference $|\widehat{Q}_n(x) - Q_n(x)|$ tends to zero in probability as n tends to infinity, for every fixed x. By (9) and (10), $Q_n(x)$ tends to the normal distribution function $\Phi_{0,\sigma^2(H)}(x)$ with mean zero and variance $\sigma^2(H)$ as n increases, for every real x. Hence $\widehat{Q}_n(x)$ tends in probability to the same limit, for every real x. Finally, using the uniform continuity of $\Phi_{0,\sigma^2(H)}$, and taking into account that

$$\sup_{x\in\mathbb{R}} |\widehat{Q}_n(x) - Q_n(x)| \le \sup_{x\in\mathbb{R}} \left\{ |Q_n(x) - \Phi_{0,\sigma^2(H)}(x)| + |\widehat{Q}_n(x) - \Phi_{0,\sigma^2(H)}(x)| \right\},$$

and that $\widehat{Q}_n(x)$ tends in probability to $\Phi_{0,\sigma^2(H)}(x)$, (13) is obtained.

Proof of Proposition 2. The proof uses standard techniques. In view of (8) and Proposition 1 it follows that Q_n and $\widehat{Q}_n$ both converge ($\widehat{Q}_n$ in probability) to a normal distribution function with mean zero and variance $\sigma^2(H)$. Using Lemma 1.5.6 in Serfling (1980), and taking into account that $\Phi_{0,\sigma^2(H)}$ is continuous and strictly increasing, it is immediate to see that

$$Q_n^{-1}(u) \to \Phi^{-1}_{0,\sigma^2(H)}(u) \ \ \forall\, 0 < u < 1, \ \ as\ n \to \infty.$$

In the same way, it is not difficult to show that

$$\widehat{Q}_n^{-1}(u) \xrightarrow{p} \Phi^{-1}_{0,\sigma^2(H)}(u) \ \ \forall\, 0 < u < 1, \ \ as\ n \to \infty,$$

from which conclusion (14) follows.

6. Simulation results

We performed a simulation study in order to evaluate the actual coverage probability of the confidence intervals (15). In particular, we focus on simulating traces from a Fractional Brownian motion. Although our remarks depend on this particular choice, some useful indications on the validity of the proposed confidence intervals are obtained. Of course, in order to reach wider and stronger conclusions, simulations from different LRD processes, and from other combinations of H, n, and l would be necessary.

To synthesize fractional Gaussian noise (i.e., the increments of a Fractional Brownian motion), approximate synthesis methods with reasonable computational loads are known (see Park and Willinger (2000) and Bardet, Lang, Oppenheim, Philippe and Taqqu (2003)), but their approximation errors cannot be easily controlled. For this reason we have used the circulant matrix method that possesses a good computational efficiency. In order to generate a trace of length n, this method essentially performs an embedding of the correlation matrix of the original process into a non-negative definite matrix R of size $m \geq 2(n-1)$, which is a circulant. The main advantage is that circulants are easy to diagonalize by using the discrete Fourier transform. Therefore, when m is an integer power of 2, the complexity of the circulant matrix embedding method coincides with the complexity of the Fast Fourier Transform, that is $O(m \log_2 m)$. A detailed description of the method is given in Davies and Harte (1987), Bardet et al. (2003)); properties are studied in Craigmile (2003).

Two values of H were considered, $H = 0.6$ and $H = 0.8$, and two sample sizes: $n = 2^{11}$ and $n = 2^{15}$. For the above values of H and n, 1,000 independent traces of length n were generated.

The coverage probability of the (linear) wavelet estimator has been computed over the full sample of n observations for three different choices of the mother wavelet: Haar function, Daubechies–2, and Daubechies–4 wavelets, respectively (Daubechies (1992)). In order to test which scales should be used, the goodness-of-fit test proposed in Veitch et al. (2003) has been applied. Daubechies–2, and Daubechies–4 mother wavelets satisfy condition (c) of Proposition 1 (see Bardet et al. (2000)). The Haar wavelet possesses only one vanishing moment. It has been included in our simulation study because it is widely used in practice.

Table 1 displays the mean, the standard deviation, the skewness and kurtosis coefficients of the estimator $\widehat{H}_n$, computed on the basis of our simulated data. Furthermore, in order to evaluate how well the confidence intervals for H proposed in Veitch and Abry (1999) work in practice, the coverage probabilities have been computed on the basis of the simulated data, for two nominal confidence levels 0.95 and 0.99. The corresponding results are reported in Table 2, where the interval length is computed as $2\, z_{\alpha/2}\sigma(H)$, with $\Phi_{0,1}(z_{\alpha/2}) = 1 - \alpha/2$. The standard errors $\widehat{\sigma}$, denoted *Std* in Table 2 and shown in parentheses, are based on the binomial approximation. Namely, $\widehat{\sigma} = \sqrt{\widehat{p}(1-\widehat{p})/N}$, where $\widehat{p}$ is the estimated coverage probability, and where N is the number of independent replications considered.

Table 1. Mean, standard deviation, skewness, and kurtosis of $\widehat{H}_n$.

Mother wavelet	$H = 0.6$ *Mean*	*Std*	*Skew*	*Kurt*	$H = 0.8$ *Mean*	*Std*	*Skew*	*Kurt*
	$n = 2^{11} = 2,048$							
Haar	0.581	0.123	−0.130	−0.156	0.780	0.126	−0.153	−0.144
Daubechies 2	0.625	0.141	0.120	−0.089	0.827	0.145	0.107	−0.050
Daubechies 4	0.610	0.155	0.110	−0.130	0.814	0.164	0.099	−0.174
	$n = 2^{15} = 32768$							
Haar	0.595	0.075	−0.140	−0.143	0.794	0.079	−0.155	−0.154
Daubechies 2	0.602	0.070	0.150	−0.101	0.806	0.071	0.153	−0.098
Daubechies 4	0.602	0.073	0.190	−0.113	0.808	0.074	0.163	−0.164

Table 2. Coverage probabilities and lengths for the confidence intervals by Veitch and Abry (1999); in parentheses, standard errors are reported.

Mother wavelet	$H = 0.6$ *nom. lev.* 0.95	*nom. lev.* 0.99	$H = 0.8$ *nom. lev.* 0.95	*nom. lev.* 0.99
	$n = 2^{11}$			
Haar	0.941	0.982	0.919	0.978
	(0.0075)	(0.0042)	(0.0086)	(0.0046)
Daubechies 2	0.943	0.984	0.923	0.980
	(0.0073)	(0.0040)	(0.0084)	(0.0044)
Daubechies 4	0.942	0.983	0.931	0.973
	(0.0074)	(0.0041)	(0.0080)	(0.0051)
Interval length	0.049	0.065	0.049	0.065
	(0.0)	(0.0)	(0.0)	(0.0)
	$n = 2^{15}$			
Haar	0.942	0.981	0.917	0.972
	(0.0074)	(0.0043)	(0.0087)	(0.0052)
Daubechies 2	0.943	0.985	0.921	0.979
	(0.0073)	(0.0038)	(0.0085)	(0.0045)
Daubechies 4	0.941	0.982	0.933	0.976
	(0.0075)	(0.0042)	(0.0079)	(0.0048)
Interval length	0.022	0.029	0.022	0.029
	(0.0)	(0.0)	(0.0)	(0.0)

Some comments are in order.

- The confidence intervals of Veitch and Abry (1999) seem to be too narrow since their coverage probability tends to be smaller than the nominal one.

This may be explained as follows. First, the asymptotic variance of $\widehat{H}_n$ used by those authors is not correct, and this is the main source of inaccuracy. Second, the (simulated) distribution of $\widehat{H}_n$ over the 1,000 traces, for each value of H, exhibits skewness (skew) and negative kurtosis (kurt), so that the confidence intervals computed under the hypothesis of normality tend to be narrower than they should. Furthermore, the variance is taken to be the asymptotic variance under assumptions **C2** and **C3**, which are not true in reality.

- The coverage probability decreases as H increases. The reason is that the correlation in k of the d_{jk}'s, which is ignored in Veitch and Abry (1999), increases as H increases (see Flandrin (1992), Masry (1993) and Tewfik and Kim (1992)).
- As the sample size n increases, the (expected) length of the confidence intervals considered decreases. On the other hand, the coverage probabilities are essentially the same.

To test how the confidence intervals (15) based on subsampling perform, the simulated data described above have been used to evaluate their actual coverage probability.

Experience with the case of short-range dependent data suggests that the block size l should be considerably smaller than n. Following Hall et al. (1998), the choice of $l = cn^{1/2}$, with $c = 3$, 6 and 9 (with l possibly corrected to the closest power of 2) has been considered, leading to $l = 3\sqrt{2048} \approx 128 = 2^7$, $l = 6\sqrt{2048} \approx 256 = 2^8$, and $l = 9\sqrt{2048} \approx 512 = 2^9$ when $n = 2^{11}$, and to $l = 3\sqrt{32768} \approx 512 = 2^9$, $l = 6\sqrt{32768} \approx 1,024 = 2^{10}$, and $l = 9\sqrt{32768} \approx 2,048 = 2^{11}$ when $n = 2^{15}$. The number of blocks N is therefore 1921, 1793, and 1537 in the case $n = 2^{11}$, and 32257, 31745 and 30721 in the case $n = 2^{15}$.

The results we have obtained are summarized in Tables 3 and 4. Statistics of the distribution of the lengths of the confidence intervals based on the subsampling method have been computed for each block size and are presented in Tables 5 and 6. For a given sample of confidence intervals of lengths $L_j,\ j = 1, \ldots, N$, we compute the sample *mean* $\overline{L} = \sum_{j=1}^{N} L_j/N$ and the *range* $\max_{1\le j\le N} L_j - \min_{1\le j\le N} L_j$. In parentheses, *standard deviations* $s = \{\sum_{i=1}^{N}(L_i-\overline{L})^2/(N-1)\}^{1/2}$ of the lengths are also reported.

Here are some remarks.

- For each value of H and for each mother wavelet considered, the coverage probability of the confidence intervals based on the subsampling method behaves better (i.e., is closer to the fixed nominal level) than the coverage probability of the linear wavelet estimator based on the assumptions of Abry and Veitch and displayed in Table 2. This is because $\widehat{Q}_n(x)$ converges to the same limiting distribution function as $Q_n(x)$, which is the correct limit distribution of the error $T_n = \sqrt{\nu_{j_1(n)}(n)}(\widehat{H}_n - H)$. This limit distribution is normal with zero mean and variance $\sigma^2(H)$, which is different from the variance involved in the confidence intervals based on assumptions **C2** and **C3**. Therefore, the confidence intervals based on subsampling are asymptotically correct, whereas this is not the case for the ones proposed by Abry and Veitch.

Table 3. Coverage probabilities for confidence intervals (15); $n = 2^{11}$; in parentheses, standard errors are reported.

Block size: $l = 128 = 2^7$				
Mother	$H = 0.6$		$H = 0.8$	
wavelet	*nom. lev.* 0.95	*nom. lev.* 0.99	*nom. lev.* 0.95	*nom. lev.* 0.99
Haar	0.950	0.993	0.954	0.998
	(0.0069)	(0.0025)	(0.0066)	(0.0014)
Daubechies 2	0.952	0.992	0.957	0.996
	(0.0067)	(0.0028)	(0.0064)	(0.0020)
Daubechies 4	0.956	0.991	0.960	0.996
	(0.0065)	(0.0030)	(0.0062)	(0.0020)
Block size: $l = 256 = 2^8$				
Mother	$H = 0.6$		$H = 0.8$	
wavelet	*nom. lev.* 0.95	*nom. lev.* 0.99	*nom. lev.* 0.95	*nom. lev.* 0.99
Haar	0.954	0.994	0.956	0.999
	(0.0066)	(0.0024)	(0.0065)	(0.0031)
Daubechies 2	0.955	0.992	0.955	0.997
	(0.0065)	(0.0028)	(0.0065)	(0.0017)
Daubechies 4	0.956	0.993	0.961	0.996
	(0.0065)	(0.0026)	(0.0061)	(0.0020)
Block size: $l = 512 = 2^9$				
Mother	$H = 0.6$		$H = 0.8$	
wavelet	*nom. lev.* 0.95	*nom. lev.* 0.99	*nom. lev.* 0.95	*nom. lev.* 0.99
Haar	0.955	0.994	0.957	0.993
	(0.0065)	(0.0024)	(0.0064)	(0.0026)
Daubechies 2	0.957	0.993	0.956	0.992
	(0.0064)	(0.0026)	(0.0065)	(0.0028)
Daubechies 4	0.957	0.994	0.959	0.997
	(0.0064)	(0.0024)	(0.0064)	(0.0017)

Table 4. Coverage probabilities for confidence intervals (15); $n = 2^{15}$; in parentheses, standard errors are reported.

Block size: $l = 512 = 2^9$				
Mother wavelet	$H = 0.6$		$H = 0.8$	
	nom. lev. 0.95	*nom. lev.* 0.99	*nom. lev.* 0.95	*nom. lev.* 0.99
Haar	0.953	0.993	0.955	0.996
	(0.0066)	(0.0025)	(0.0065)	(0.0020)
Daubechies 2	0.952	0.991	0.952	0.995
	(0.0067)	(0.0030)	(0.0067)	(0.0022)
Daubechies 4	0.954	0.993	0.956	0.996
	(0.0066)	(0.0025)	(0.0065)	(0.0020)

Block size: $l = 1,024 = 2^{10}$				
Mother wavelet	$H = 0.6$		$H = 0.8$	
	nom. lev. 0.95	*nom. lev.* 0.99	*nom. lev.* 0.95	*nom. lev.* 0.99
Haar	0.954	0.994	0.956	0.998
	(0.0066)	(0.0024)	(0.0065)	(0.0014)
Daubechies 2	0.952	0.991	0.953	0.996
	(0.0067)	(0.0030)	(0.0066)	(0.0020)
Daubechies 4	0.954	0.993	0.960	0.998
	(0.0066)	(0.0026)	(0.0062)	(0.0014)

Block size: $l = 2,048 = 2^{11}$				
Mother wavelet	$H = 0.6$		$H = 0.8$	
	nom. lev. 0.95	*nom. lev.* 0.99	*nom. lev.* 0.95	*nom. lev.* 0.99
Haar	0.955	0.995	0.957	0.998
	(0.0065)	(0.0022)	(0.0064)	(0.0014)
Daubechies 2	0.953	0.992	0.953	0.997
	(0.0066)	(0.0028)	(0.0066)	(0.0017)
Daubechies 4	0.956	0.993	0.960	0.999
	(0.0065)	(0.0026)	(0.0062)	(0.0010)

- The length of the confidence intervals based on the subsampling method (see Tables 5 and 6) is generally larger than that for the linear wavelet estimator based on the normal approximation (see Table 2). This may be due, in part, to the fact that the skewness and kurtosis of the actual distribution of $\widehat{H}_n$ is now taken into account.
- The actual coverage probability is always close (and typically slightly larger than) the nominal one (0.95 or 0.99).
- As l increases, the number of scales available in the computation of the wavelet estimator increases, too (from seven to nine), and consequently the variance of the estimator (and the length of the confidence interval, as well) decreases.

Table 5. Mean and range of the length of confidence intervals (15); $n = 2^{11}$. In parentheses, standard deviations of lengths are reported.

Block size: $l = 128 = 2^7$								
Mother wavelet	$H = 0.6$				$H = 0.8$			
	Mean	*Range*	*Mean*	*Range*	*Mean*	*Range*	*Mean*	*Range*
	nom. lev. 0.95		*nom. lev.* 0.99		*nom. lev.* 0.95		*nom. lev.* 0.99	
Haar	0.257	0.165	0.346	0.205	0.271	0.175	0.354	0.222
	(0.029)		(0.031)		(0.029)		(0.033)	
Daubechies 2	0.255	0.192	0.320	0.249	0.272	0.197	0.352	0.220
	(0.030)		(0.032)		(0.031)		(0.032)	
Daubechies 4	0.254	0.201	0.337	0.239	0.258	0.199	0.346	0.245
	(0.031)		(0.034)		(0.031)		(0.034)	

Block size: $l = 256 = 2^8$								
Mother wavelet	$H = 0.6$				$H = 0.8$			
	Mean	*Range*	*Mean*	*Range*	*Mean*	*Range*	*Mean*	*Range*
	nom. lev. 0.95		*nom. lev.* 0.99		*nom. lev.* 0.95		*nom. lev.* 0.99	
Haar	0.145	0.147	0.197	0.230	0.148	0.168	0.197	0.240
	(0.026)		(0.033)		(0.028)		(0.034)	
Daubechies 2	0.143	0.178	0.192	0.206	0.146	0.179	0.195	0.209
	(0.029)		(0.031)		(0.029)		(0.031)	
Daubechies 4	0.140	0.194	0.188	0.216	0.141	0.187	0.111	0.233
	(0.030)		(0.032)		(0.030)		(0.031)	

Block size: $l = 512 = 2^9$								
Mother wavelet	$H = 0.6$				$H = 0.8$			
	Mean	*Range*	*Mean*	*Range*	*Mean*	*Range*	*Mean*	*Range*
	nom. lev. 0.95		*nom. lev.* 0.99		*nom. lev.* 0.95		*nom. lev.* 0.99	
Haar	0.124	0.143	0.132	0.142	0.132	0.128	0.141	0.142
	(0.026)		(0.026)		(0.00061)		(0.00069)	
Daubechies 2	0.104	0.133	0.120	0.128	0.104	0.122	0.115	0.136
	(0.00064)		(0.00067)		(0.025)		(0.026)	
Daubechies 4	0.105	0.135	0.121	0.130	0.106	0.119	0.122	0.125
	(0.025)		(0.025)		(0.024)		(0.024)	

As an overall comment, the confidence intervals by Abry and Veitch are advantageous from a computational point of view. However, since they are based on an underestimated variance, they do not meet the nominal confidence level. On the other hand, although slightly more computationally expensive, the confidence intervals based on subsampling seem to be accurate. Their coverage probability essentially coincides with the nominal value. As far as the value of l is concerned we notice that, within the range of l values considered in our study, as l grows

Table 6. Mean and range of the length of confidence intervals (15); $n = 2^{15}$. In parentheses, standard deviations of lengths are reported.

Block size: $l = 512 = 2^9$								
Mother wavelet	$H = 0.6$				$H = 0.8$			
	Mean	*Range*	*Mean*	*Range*	*Mean*	*Range*	*Mean*	*Range*
	nom. lev. 0.95		*nom. lev.* 0.99		*nom. lev.* 0.95		*nom. lev.* 0.99	
Haar	0.127 (0.011)	0.026	0.168 (0.014)	0.039	0.156 (0.011)	0.025	0.172 (0.015)	0.048
Daubechies 2	0.151 (0.011)	0.027	0.201 (0.014)	0.041	0.153 (0.011)	0.026	0.205 (0.013)	0.037
Daubechies 4	0.207 (0.010)	0.020	0.213 (0.013)	0.032	0.209 (0.009)	0.018	0.224 (0.013)	0.032
Block size: $l = 1,024 = 2^{10}$								
Mother wavelet	$H = 0.6$				$H = 0.8$			
	Mean	*Range*	*Mean*	*Range*	*Mean*	*Range*	*Mean*	*Range*
	nom. lev. 0.95		*nom. lev.* 0.99		*nom. lev.* 0.95		*nom. lev.* 0.99	
Haar	0.095 (0.008)	0.013	0.122 (0.012)	0.032	0.098 (0.009)	0.017	0.126 (0.014)	0.042
Daubechies 2	0.104 (0.010)	0.021	0.135 (0.012)	0.031	0.106 (0.010)	0.022	0.138 (0.013)	0.036
Daubechies 4	0.130 (0.009)	0.018	0.170 (0.009)	0.016	0.131 (0.009)	0.019	0.172 (0.011)	0.025
Block size: $l = 2,048 = 2^{11}$								
Mother wavelet	$H = 0.6$				$H = 0.8$			
	Mean	*Range*	*Mean*	*Range*	*Mean*	*Range*	*Mean*	*Range*
	nom. lev. 0.95		*nom. lev.* 0.99		*nom. lev.* 0.95		*nom. lev.* 0.99	
Haar	0.063 (0.010)	0.018	0.082 (0.011)	0.022	0.066 (0.009)	0.021	0.082 (0.010)	0.027
Daubechies 2	0.070 (0.010)	0.023	0.089 (0.012)	0.029	0.072 (0.012)	0.024	0.090 (0.012)	0.032
Daubechies 4	0.083 (0.009)	0.019	0.108 (0.012)	0.030	0.084 (0.009)	0.019	0.109 (0.012)	0.031

the coverage probabilities are virtually stable, while the corresponding confidence intervals lengths decrease. Hence, it seems reasonable to adopt the heuristic *ad-hoc* rule $l = 9\sqrt{n}$. Finally, the choice of the mother wavelet does not seem to affect significantly the coverage probability. However, the Daubechies-2 wavelets turns out to work best in our simulation setting.

Acknowledgements

This research was partially supported by the NSF Grant DMS–0505747 at Boston University, and by a MIUR (Ministero dell'Istruzione, dell'Università e della Ricerca) grant. Thanks are due to two anonymous referees for their suggestions that considerably improved a previous version of the present paper.

References

Abry, P. and Veitch, D. (1998). Wavelet analysis of long range dependent traffic. *IEEE Trans. Inform. Theory* **44**, 2-15.

Abry, P., Flandrin, P., Taqqu, M. and Veitch, D. (2003). Self-similarity and long-range dependence through the wavelet lens. In *Theory and Applications of Long-Range Dependence* (Edited by P. Doukhan, G. Oppenheim and M. S. Taqqu). Birkhäuser, Basel.

Bardet, J. M., Lang, G., Moulines, E. and Soulier, P. (2000). Wavelet estimator of long-range dependent processes. *Stat. Inference Stoch. Process.* **3**, 85-99.

Bardet, J. M., Lang, G., Oppenheim, G., Philippe, A. and Taqqu, M. S. (2003). Generators of long range dependence: a survey. In *Theory and Applications of Long-Range Dependence* (Edited by P. Doukhan, G. Oppenheim and M.S. Taqqu). Birkhäuser, Basel.

Beran, J. (1994). *Statistics for Long-Memory Processes*. Chapman and Hall, London.

Craigmile, P. F. (2003). Simulating a class of stationary Gaussian processes using the Davies-Harte algorithm, with application to long memory processes. *J. Time Ser. Anal.* **24**, 505-511.

Daubechies, I. (1992). *Ten Lectures on Wavelets*. SIAM, Philadelphia.

Davies, R. B. and Harte, D. S. (1987). Tests for Hurst effect. *Biometrika* **74**, 95-101.

Embrechts, P. and Maejima, M. (2003). *Selfsimilar Processes*. Academic Press.

Flandrin, P. (1992). Wavelet analysis and synthesis of fractional brownian motion. *IEEE Trans. Inform. Theory* **138**, 910-917.

Hall, P., Jing, B. and Lahiri, S. N. (1998). On the sampling window method for long range-dependent data. *Statist. Sinica* **8**, 1189-1204.

Ibragimov, I. A. and Rozanov, Y. A. (1978). *Gaussian Random Processes*. Springer Verlag, New York.

Lahiri, S. N. (2003). *Reasampling Methods for Dependent Data*. Springer Verlag, New York.

Masry, E. (1993). The wavelet transform of stochastic processes with stationary increments and its application to fractional brownian motion. *IEEE Trans. Inform. Theory* **39**, 260-264.

Park, K. and Willinger, W. (2000). *Self-similar Network Traffic and Performance Evaluation*. Wiley, New York.

Samorodnitsky, G. and Taqqu, M. (1994). *Stable Non-Gaussian Random Processes*. Chapman and Hall, London.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley, New York.

Taqqu, M. S. (2003). Fractional Brownian motion and long-range dependence. In *Theory and Applications of Long-Range Dependence* (Edited by P. Doukhan, G. Oppenheim and M. S. Taqqu). Birkhäuser, Basel.

Tewfik, A. H. and Kim, M. (1992). Correlation structure of the discrete wavelet coefficients of fractional brownian motion. *IEEE Trans. Inform. Theory* **38**, 904-909.

Veitch, D. and Abry, P. (1999). A wavelet based joint estimator of the parameters of long range dependence traffic. *IEEE Trans. Inform. Theory*, special issue on "Multiscale Statistical Signal Analysis and its Applications" **45**, 878-897.

Veitch, D., Abry, P. and Taqqu, M. (2000). Meaningful MRA inisialisation for discrete time series. *Signal Processing* **80**, 1971-1983.

Veitch, D., Abry, P. and Taqqu, M. (2003). On the automatic selection of the onset of scaling. *Fractals* **11**, 377-390.

Dipartimento di Statistica, Probabilità e Statistiche Applicate, Sapienza Università di Roma P.le Aldo Moro, 5, 00185 Roma, Italy.

E-mail: pierluigi@uniroma1.it

Facoltà di Scienze della Formazione, Università degli Studi LUMSA, P.za delle Vaschette, 101, 00193 Roma, Italy.

E-mail: l.degiovanni@lumsa.it

Department of Statistics, University of Michigan, 445B West Hall, 1085 South University, Ann Arbor, MI 48109-1107, U.S.A.

E-mail: sstoev@unimich.edu

Department of Mathematics and Statistics, Boston University, 111 Cummington Street, Boston, MA 02215, U.S.A.

E-mail: murad@math.bu.edu

(Received January 2006; accepted November 2006)