



HAL
open science

Theoretical and numerical comparisons of the parameter estimator of the fractional Brownian motion

Jean-Marc Bardet

► **To cite this version:**

Jean-Marc Bardet. Theoretical and numerical comparisons of the parameter estimator of the fractional Brownian motion. Springer. Mathematical Structures and Applications, STEAM-H: Science, Technology, Engineering, Agriculture, Mathematics & Health, Springer, pp.153-173, In press, Mathematical Structures and Applications, 10.1007/978-3-319-97175-9 . hal-01674067

HAL Id: hal-01674067

<https://hal.science/hal-01674067>

Submitted on 1 Jan 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Theoretical and numerical comparisons of the parameter estimator of the fractional Brownian motion

JEAN-MARC BARDET ^a

^a *University Paris 1 Panthou-Sorbonne, France*

E-mail: bardet@univ-paris1.fr

This paper is dedicated to Norbert Hounkonnou who is not really a specialist of probability and statistics (...) but who often taught these topics and has mainly allowed them to significantly develop in Benin by his unwavering support to Masters and PhD of Statistics.

The fractional Brownian motion which has been defined by Kolmogorov [16] and numerous papers was devoted to its study since its study in Mandelbrot and Van Ness [19] present it as a paradigm of self-similar processes. The self-similarity parameter, also called the Hurst parameter, commands the dynamic of this process and the accuracy of its estimation is often crucial. We present here the main and used methods of estimation, with the limit theorems satisfied by the estimators. A numerical comparison is also provided allowing to distinguish between the estimators.

Introduction

The fractional Brownian motion (fBm for short) has been studied a lot since the seminal paper of Kolmogorov [16] and its resumption in Mandelbrot and Van Ness [19]. A simple way to present this extension of the classical Wiener Brownian motion is to define it from its both first moments. Hence, a fBm with parameter $(H, \sigma^2) \in (0, 1] \times (0, \infty)$ is a centered Gaussian process $X = \{X(t), t \in \mathbb{R}\}$ having stationary increments and such as

$$\text{cov}(X(s), X(t)) = \sigma^2 R_H(s, t) = \frac{1}{2} \sigma^2 (|s|^{2H} + |t|^{2H} - |t-s|^{2H}) \quad (s, t) \in \mathbb{R}^2.$$

As a consequence, $\text{Var}(X(t)) = \sigma^2 |t|^{2H}$ for any $t \in \mathbb{R}$, which induces that X is the only H -self-similar centered Gaussian process having stationary increments. More detail on this process can be found in the monograph of Samorodnitsky and Taqqu [25].

We consider the following statistical problem. We suppose that a trajectory $(X(1), \dots, X(N))$ of X is observed and we would like to estimate the parameters H and σ^2 which are unknown.

Remark 0.1 *It is also possible to consider an observed path $(X(1/N), X(2/N), \dots, X(1))$ of X . Using the self-similarity property of X , the distributions of $(X(1), \dots, X(N))$ and $N^H(X(1/N), X(2/N), \dots, X(1))$ are the same.*

This statistical model is parametric and it is natural to estimate $\theta = (H, \sigma^2)$ using a parametric method. Hence, in the one hand, in the forthcoming Section 1, two possible parametric estimators are presented. In the second hand, several other used and famous semi-parametric estimators are studied in Section 2. For each estimator, its asymptotic behavior is stated and the main references recalled. Finally, Section 3 is devoted to a numerical study from Monte-Carlo experiments, allowing to obtain some definitive conclusions with respect to the estimators.

1 Two classical parametric estimators

Since X is a Gaussian process for which the distribution is integrally defined when θ is known, a parametric method such as Maximum Likelihood Estimator (MLE) is a natural choice of estimation and it provides efficient estimators. As X is a process having stationary increments, it is appropriate to define Y the process of its increments, i.e. the fractional Gaussian noise,

$$Y = \{Y(t), t \in \mathbb{R}\} = \{X(t) - X(t-1), t \in \mathbb{R}\},$$

with covariogram $r_Y(k) = \text{cov}(Y(t), Y(t+k))$ satisfying

$$r_Y(k) = \frac{1}{2} \sigma^2 (|k+1|^{2H} + |k-1|^{2H} - 2|k|^{2H}) \quad k \in \mathbb{R}.$$

The spectral density f_Y of Y is defined for $\lambda \in [-\pi, \pi]$ by (see Sinaï [21] or Fox and Taqu [11]):

$$f_H(\lambda) := \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} r_Y(k) e^{ik\lambda} = 2\sigma^2 \Gamma(2H+1) \sin(\pi H) (1 - \cos \lambda) \sum_{k \in \mathbb{Z}} |\lambda + 2k\pi|^{-1-2H}. \quad (1.1)$$

Hence, since $(X(1), \dots, X(N))$ is observed, $(Y(1), \dots, Y(N))$ is observed (we assume $X(0) = 0$).

1.1 Maximum Likelihood Estimation

The likelihood $L_{(Y(1), \dots, Y(N))}(y_1, \dots, y_N)$ of $(Y(1), \dots, Y(N))$ is the Gaussian probability density of $(Y(1), \dots, Y(N))$, which can be written

$$\begin{aligned} L_{(Y(1), \dots, Y(N))}(y_1, \dots, y_N) &= \frac{(2\pi)^{N/2}}{\det(\Sigma_N(H, \sigma^2))} \times \\ &\exp\left(-\frac{1}{2}(y_1, \dots, y_N) \Sigma_N^{-1}(H, \sigma^2) (y_1, \dots, y_N)'\right), \end{aligned}$$

where the definite positive covariance matrix $\Sigma_N(H, \sigma^2)$ is such as

$$\Sigma_N(H, \sigma^2) = (r(|i-j|))_{1 \leq i, j \leq N}.$$

Then the MLE $\widehat{\theta}_N = (\widehat{H}_N, \widehat{\sigma}_N^2)$ of θ is defined as:

$$(\widehat{H}_N, \widehat{\sigma}_N^2) = \text{Arg} \max_{(H, \sigma^2) \in (0,1) \times (0,\infty)} L_{(Y(1), \dots, Y(N))}(Y(1), \dots, Y(N)).$$

As it is generally done, it can be more convenient to minimize the contrast defined by $-2 \log(L_{(Y(1), \dots, Y(N))}(Y(1), \dots, Y(N)))$.

The asymptotic behavior of this estimator has been first obtained in Dahlhaus [9]. The main results are the following:

Theorem 1.1 *The estimator $(\widehat{H}_N, \widehat{\sigma}_N^2)$ is asymptotically efficient and satisfies*

$$\begin{pmatrix} \widehat{H}_N \\ \widehat{\sigma}_N^2 \end{pmatrix} \xrightarrow[N \rightarrow \infty]{a.s.} \begin{pmatrix} H \\ \sigma^2 \end{pmatrix} \quad (1.2)$$

and

$$\sqrt{N} \left(\begin{pmatrix} \widehat{H}_N \\ \widehat{\sigma}_N^2 \end{pmatrix} - \begin{pmatrix} H \\ \sigma^2 \end{pmatrix} \right) \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \mathcal{N}_2(0, \Gamma_0^{-1}(H, \sigma^2)), \quad (1.3)$$

where $\Gamma_0^{-1}(H, \sigma^2)$ is the limit of $\frac{1}{N} I_N^{-1}(H, \sigma^2)$ and $I_N(H, \sigma^2)$ is the Fisher information matrix of $(Y(1), \dots, Y(N))$. Moreover,

$$\Gamma_0(H, \sigma^2) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \left(\frac{\partial}{\partial \theta} \log f_{\theta}(\lambda) \right) \left(\frac{\partial}{\partial \theta} \log f_{\theta}(\lambda) \right)' d\lambda, \quad (1.4)$$

and $f_{\theta}(\lambda) = \sigma^2 g_H(\lambda)$ where

$$g_H(\lambda) = 2 \sin(\pi H) \Gamma(2H+1) (1 - \cos \lambda) \sum_{j \in \mathbb{Z}} |\lambda + 2\pi j|^{-2H-1}.$$

The asymptotic covariance $\Gamma_0(H, \sigma^2)$ can not be really simplified, we just can obtain:

$$\Gamma_0^{-1}(H, \sigma^2) = \frac{1}{\frac{1}{2}a_H - b_H^2} \begin{pmatrix} \frac{1}{2} & -\sigma^2 b_H \\ -\sigma^2 b_H & \sigma^4 a_H \end{pmatrix} \quad (1.5)$$

where $a_H = \frac{1}{4\pi} \int_{-\pi}^{\pi} \left(\frac{\partial}{\partial H} \log g_H(\lambda) \right)^2 d\lambda$ and $b_H = \frac{1}{4\pi} \int_{-\pi}^{\pi} \frac{\partial}{\partial H} \log g_H(\lambda) d\lambda$.

1.2 Whittle estimation

The MLE is an asymptotically efficient estimator but it has two main drawbacks: first, it is a parametric estimator which can only be used, stricto sensu, to fBm and its use is numerically limited since the computation of the likelihood requires to inverse the matrix $\Sigma_N(H, \sigma^2)$ and this is extremely time consuming when $N \geq 5000$ and impossible

when $N \geq 10000$ (with a 2014-software). In Whittle [28], a general approximation of the likelihood for Gaussian stationary processes Y has been first proposed. This consists on writing for Y depending on a parameter vector θ :

$$-\frac{1}{N} \log L_{(Y(1), \dots, Y(N))}(Y(1), \dots, Y(N)) - \frac{1}{2} \log(2\pi) \\ \stackrel{\mathcal{D}}{\underset{N \rightarrow \infty}{\approx}} \widehat{U}_N(\theta) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \left(\log(f_{\theta}(\lambda)) + \frac{\widehat{I}_N(\lambda)}{f_{\theta}(\lambda)} \right) d\lambda$$

where $\widehat{I}_N(\lambda) = \frac{1}{2\pi N} \left| \sum_{k=1}^N Y(k) e^{-ik\lambda} \right|^2$ is the periodogram of Y .

Then $\widetilde{\theta}_N = (\widetilde{H}_W, \widetilde{\sigma}_N^2) = \text{Argmin}_{\theta \in \Theta} \widehat{U}_N(\theta)$ is called the Whittle estimator of θ . In case of the fractional Gaussian noise, Dahlhaus [9] achieved the results of Fox and Taqqu [11] and proved the following limit theorem:

Theorem 1.2 *The estimator $(\widetilde{H}_W, \widetilde{\sigma}_N^2)$ is asymptotically efficient and satisfies*

$$\begin{pmatrix} \widetilde{H}_W \\ \widetilde{\sigma}_N^2 \end{pmatrix} \xrightarrow[N \rightarrow \infty]{a.s.} \begin{pmatrix} H \\ \sigma^2 \end{pmatrix} \quad (1.6)$$

and

$$\sqrt{N} \left(\begin{pmatrix} \widetilde{H}_W \\ \widetilde{\sigma}_N^2 \end{pmatrix} - \begin{pmatrix} H \\ \sigma^2 \end{pmatrix} \right) \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \mathcal{N}_2(0, \Gamma_0^{-1}(H, \sigma^2)), \quad (1.7)$$

with $\Gamma_0(H, \sigma^2)$ defined in (1.4).

Hence, the Whittle estimator $(\widetilde{H}_W, \widetilde{\sigma}_N^2)$ has the same asymptotic behaviour than the MLE while its numerical accuracy is clearly better: in case of fractional Gaussian noise, and therefore in case of the fBm, the Whittle estimator has to be preferred to the MLE.

2 Other classical semi-parametric estimators

As we said previously, we present now some classical semi-parametric methods frequently used for estimating the parameter H of a fBm, but also applied to other long-range dependent or self-similar processes.

2.1 R/S and modified R/S statistics

The first estimator which has been applied to a fBm, and more precisely to a fractional Gaussian noise has been the R/S estimator. This estimator defined by the hydrologist Hurst [14] was devoted to estimated the long-range dependent parameter (also called the Hurst parameter) of a long memory process. Lo [18] introduced the modified R/S statistic for a times series X which is defined as

$$\widehat{Q}_N(q) = \frac{1}{\widehat{s}_{N,q}} \left(\max_{1 \leq k \leq N} \sum_{i=1}^k (X(i) - \bar{X}_N) - \min_{1 \leq k \leq N} \sum_{i=1}^k (X(i) - \bar{X}_N) \right) \quad (2.1)$$

where $\bar{X}_N = \frac{1}{N}(X(1) + \dots + X(N))$ is the sample mean and $\hat{s}_{N,q}^2$ is an estimator of $\sigma^2 = \sum_{i \in \mathbb{Z}} \text{cov}(X(0), X(j))$ defined by

$$\hat{s}_{N,q}^2 = \frac{1}{N} \sum_{i=1}^N (X(i) - \bar{X}_N)^2 + 2 \sum_{i=1}^q \omega_i(q) \hat{\gamma}_N(i) \quad (2.2)$$

$$\text{with } \begin{cases} \omega_i(q) = 1 - i/(q+1) \\ \hat{\gamma}_N(i) = \frac{1}{N} \sum_{j=1}^{N-i} (X(j) - \bar{X}_N)(X(j+i) - \bar{X}_N) \end{cases} \quad (2.3)$$

The classical R/S statistic corresponds to $q = 0$. In such case, we have the following asymptotic behavior when X is a fGn:

Proposition 2.1 *If X is a fGn with parameter $H > 1/2$, then (see Li et al., 2011)*

$$\frac{1}{N^H} E[\hat{Q}_N(0)] \xrightarrow{N \rightarrow \infty} E[\max_{0 \leq t \leq 1} B_H(t) - \min_{0 \leq t \leq 1} B_H(t)], \quad (2.4)$$

with B_H the fractional bridge with parameter H , i.e., $B_H(t) = X_H(t) - tX_H(1)$ for $t \in [0, 1]$ where X_H is a standardized fBm of parameter 1.

Using this asymptotic behavior of the expectation, an estimator of H has been defined. First the trajectory $(X(1), \dots, X(N))$ is divided in K blocks of length N/K and $\hat{Q}_{n_i}(0)$ is averaged for several values of n_i such as $n_i \xrightarrow{N \rightarrow \infty} \infty$. Then, a log-log regression of $(\hat{Q}_{n_i}(0))$ onto (n_i) provides a slope \hat{H}_{RS} which is an estimator of H since $\log E[\hat{Q}_{n_i}(0)] \simeq H \log(n_i) + C$ (see for instance Taqqu et al. [23]). Note that even in the "simple" case of the fGn, there still do not really exist a convincing asymptotic study of such an estimator.

Lo [18] and numerical experiments in Taqqu et al. [23] have shown that this estimator is not really accurate and Lo [18] proposed an extension of this R/S statistic: this is the modified R/S statistic. We have the following asymptotic behavior (see Giraitis et al. [12] and Li et al. [17]):

Proposition 2.2 *If $q \xrightarrow{N \rightarrow \infty} \infty$ and $q/N \xrightarrow{N \rightarrow \infty} 0$, and X is a fGn, then:*

$$\bullet \text{ if } H \leq 1/2, \quad N^{-1/2} \hat{Q}_N(q) \xrightarrow[N \rightarrow \infty]{\mathcal{D}} U_{R/S} \quad (2.5)$$

$$\bullet \text{ if } H > 1/2, \quad q^{H-1/2} N^{-H} \hat{Q}_N(q) \xrightarrow[N \rightarrow \infty]{\mathcal{D}} Z_{R/S} \quad (2.6)$$

where $U_{R/S} = \max_{0 \leq t \leq 1} B_{1/2}(t) - \min_{0 \leq t \leq 1} B_{1/2}(t)$ and $Z_{R/S} = \max_{0 \leq t \leq 1} B_H(t) - \min_{0 \leq t \leq 1} B_H(t)$, with B_H the fractional bridge with parameter H .

Then, using several values of q , (q_1, \dots, q_m) , a log-log regression of $\hat{Q}_N(q_i)$ onto q_i provides an estimator \hat{H}_{RSM} of H (the slope of the regression line is $H - \frac{1}{2}$). But there do not exist more precise result about the convergence rate of such estimator of H in the literature. Moreover, in Teverovsky et al. [24], the difficulty of selecting a right range of values for q_i is highlighting.

As a conclusion, we can say that R/S or modified R/S statistics provide estimation of H but these estimators are not really accurate.

2.2 Second-order quadratic variations

Contrary to the R/S method, the second-order quadratic variations can be directly applied to self-similar processes, and hence in particular to fBm.

For presenting this method, introduced by Guyon and Leòn [13] and Istas and Lang [15], first, for $a \in \mathbb{N}^*$, define the second-order quadratic variations of $X = (X(i))_{i \in \mathbb{Z}}$ by

$$V_t^{(a)} := (X(t+2a) - 2X(t+a) + X(t))^2 \text{ for } t \in \mathbb{N}^*. \quad (2.7)$$

The key-point of this method is the following property:

Property 2.3 *If X is a second moment order H -self-similar process having stationary increments, with $EX^2(1) = \sigma^2$, then for all $a \in \mathbb{N}^*$ and $t \in \mathbb{N}^*$*

$$E(V_t^{(a)}) = \sigma^2 (4 - 2^{2H}) a^{2H}. \quad (2.8)$$

Therefore $\log E(V_t^{(a)}) = C + 2H \log a$ for any (a, t) . This provides an idea of estimating H : if $E(V_t^{(a_i)})$ can be estimated for several different scales a_i , then the slope of the log-log-regression of $(\widehat{E}(V_t^{(a_i)}))$ onto (a_i) is $2\widehat{H}_N$, which is an estimator of H .

The common choice for scales are $a_i = i$ and the estimator of $E(V_t^{(i)})$ is the empirical mean of $V_t^{(i)}$,

$$S_N(i) = \frac{1}{N-2i} \sum_{k=1}^{N-2i} V_k^{(i)}. \quad (2.9)$$

Then a central limit theorem can be established for $i \in \mathbb{N}^*$:

$$\sqrt{N}(S_N(i) - \sigma^2(4 - 2^{2H})i^{2H}) \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \mathcal{N}(0, \gamma(i))$$

with $\gamma(i) = \frac{1}{2} \sigma^4 i^{4+2H} \sum_{\ell=-\infty}^{\infty} (|\ell+2|^{2H} + |\ell-2|^{2H} - 2|\ell+1|^{2H} - 2|\ell-1|^{2H} + 6|\ell+2|^{2H})^2$.

Then we can define

$$\widehat{H}_N := \frac{1}{2} \frac{A}{AA^\top} (\log(S_N(i)))_{1 \leq i \leq p}^\top,$$

where $A := (\log i - \frac{1}{p} \sum_{j=1}^p \log j)_{1 \leq i \leq p} \in \mathbb{R}^p$ is a row vector, A^\top its transposed vector (vector-column).

As a consequence, it can be shown (see Bardet [2] or [8]) that, with $(a_0, a_1, a_2) = (1, -2, 1)$,

$$\sqrt{N}(\widehat{H}_N - H) \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \mathcal{N}(0, \Sigma(H)), \text{ where } \Sigma(H) := \frac{A\Gamma(H)A^\top}{4(A^\top A)^2} \text{ and } (2.10)$$

$$\Gamma(H) := \left(\frac{2}{i^{2H} j^{2H}} \sum_{k \in \mathbb{Z}} \left[\frac{\sum_{k_1, k_2=0}^q a_{k_1}^{(1)} a_{k_2}^{(1)} |ik_1 - jk_2 + k|^{2H}}{\sum_{k_1, k_2=0}^2 a_{k_1}^{(1)} a_{k_2}^{(1)} |k_1 - k_2|^{2H}} \right]^2 \right)_{1 \leq i, j \leq p}. \quad (2.11)$$

This method has a lot of advantages: low time-consuming, convergence rate close to MLE convergence rate...

However it is easy to slightly improve this estimator. First, as the asymptotic covariance $\Gamma(H)$ of $(S_N(i))$ is a function of H ; hence, using the estimator \hat{H}_N , this asymptotic covariance can be estimated. Hence, a pseudo-generalized estimator of H can be defined. More precisely, define

$$\hat{\Gamma}_N = \Gamma(\hat{H}_N). \quad (2.12)$$

Then the pseudo-generalized estimator \tilde{H}_N of H is defined by:

$$\tilde{H}_N = \frac{(B_N L)^\top \hat{\Gamma}_N^{-1} (\log(S_N(i)))_{1 \leq i \leq p}^\top}{2(B_N L)^\top \hat{\Gamma}_N^{-1} (B_N L)} \quad (2.13)$$

with $I_p = (1, 1, \dots, 1)^\top$, $L = (\log i)_{1 \leq i \leq p}$ and $B_N = I - \frac{I_p I_p^\top \hat{\Gamma}_N}{I_p^\top \hat{\Gamma}_N^{-1} I_p}$.

Then, from Bardet [2],

Proposition 2.4 *If X is a fBm of parameter H , then with $B = I - \frac{I_p I_p^\top \Gamma}{I_p^\top \Gamma^{-1} I_p}$,*

$$\sqrt{N}(\tilde{H}_N - H) \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \mathcal{N}(0; \Sigma'(H)), \quad \text{with } \Sigma'(H) = \frac{1}{4^\top (BL) G^{-1} (BL)}. \quad (2.14)$$

From Gauss-Markov Theorem, the asymptotic variance $\Sigma'(H)$ is smaller or equal to $\Sigma(H)$ and thus the estimator \tilde{H}_N is more accurate than \hat{H}_N .

Another improvement of this estimator consists on considering a number p of "scales" increasing with N : this is what we will use in simulations (theoretical results are not yet established, if they could be once).

2.3 Detrended Fluctuation Analysis (DFA)

The DFA method was introduced in Peng *al.* [20] in a biological frame. The aim of this method is to highlight the self-similarity of a time series with a trend. Let $(Y(1), \dots, Y(N))$ be a sample of a time series $(Y(n))_{n \in \mathbb{N}}$.

1. The first step of the DFA method is a division of $\{1, \dots, N\}$ in $[N/n]$ windows of length n (for $x \in \mathbb{R}$, $[x]$ is the integer part of x). In each window, the least squares regression line is computed, which represents the linear trend of the process in the window. Then, we denote by $\hat{Y}_n(k)$ for $k = 1, \dots, N$ the process formed by this piecewise linear interpolation. Then the DFA function is the standard deviation of the residuals obtained from the difference between $Y(k)$ and $\hat{Y}_n(k)$, therefore,

$$\hat{F}(n) = \sqrt{\frac{1}{n \cdot [N/n]} \sum_{k=1}^{n \cdot [N/n]} (Y(k) - \hat{Y}_n(k))^2}$$

2. The second step consists on a repetition of the first step with different values (n_1, \dots, n_m) of the window's length. Then the graph of the $\log \hat{F}(n_i)$ by $\log n_i$ is drawn. The slope of the least squares regression line of this graph provides an estimation of the self-similarity parameter of the process $(Y(k))_{k \in \mathbb{N}}$.

From the construction of the DFA method, it is interesting to define the restriction of the DFA function in a window. Thus, for $n \in \{1, \dots, N\}$, one defines the partial DFA function computed in the j -th window, *i.e.*

$$F_j^2(n) = \frac{1}{n} \sum_{i=n(j-1)+1}^{nj} (X(i) - \widehat{X}_n(i))^2, \quad \text{for } j \in \{1, \dots, [N/n]\}. \quad (2.15)$$

Then, it is obvious that

$$F^2(n) = \frac{1}{[N/n]} \sum_{j=1}^{[N/n]} F_j^2(n). \quad (2.16)$$

Let $\{X^H(t), t \geq 0\}$ be a FBM, built as a cumulated sum of stationary and centered fGn $\{Y^H(t), t \geq 0\}$. In Bardet and Kammoun [4], the following detailed asymptotic behavior of the DFA method is established. First some asymptotic properties of $F_1^2(n)$ can be established:

Proposition 2.5 *Let $\{X^H(t), t \geq 0\}$ be a fBm with parameters $0 < H < 1$ and $\sigma^2 > 0$. Then, for n and j large enough,*

$$\begin{aligned} 1. \quad E(F_1^2(n)) &= \sigma^2 f(H) n^{2H} \left(1 + O\left(\frac{1}{n}\right)\right), \quad \text{with } f(H) = \frac{(1-H)}{(2H+1)(H+1)(H+2)}, \\ 2. \quad \text{Var}(F_1^2(n)) &= \sigma^4 g(H) n^{4H} \left(1 + O\left(\frac{1}{n}\right)\right), \quad \text{with } g \text{ depending only on } H, \\ 3. \quad \text{cov}(F_1^2(n), F_j^2(n)) &= \sigma^4 h(H) n^{4H} j^{2H-3} \left(1 + O\left(\frac{1}{n}\right) + O\left(\frac{1}{j}\right)\right), \\ &\quad \text{with } h(H) = \frac{H^2(H-1)(2H-1)^2}{48(H+1)(2H+1)(2H+3)}. \end{aligned}$$

In order to obtain a central limit theorem for the logarithm of the DFA function, we consider a normalized DFA functions

$$\tilde{S}_j(n) = \frac{F_j^2(n)}{n^{2H} \sigma^2 f(H)} \quad \text{and} \quad \tilde{S}(n) = \frac{F^2(n)}{n^{2H} \sigma^2 f(H)} \quad (2.17)$$

for $n \in \{1, \dots, N\}$ and $j \in \{1, \dots, [N/n]\}$.

Under conditions on the asymptotic length n of the windows, one proves a central limit theorem satisfied by the logarithm of the empirical mean $\tilde{S}(n)$ of the random variables $(\tilde{S}_j(n))_{1 \leq j \leq [N/n]}$.

Proposition 2.6 *Under the previous assumptions and notations, let $n \in \{1, \dots, N\}$ be such that $N/n \rightarrow \infty$ and $N/n^3 \rightarrow 0$ when $N \rightarrow \infty$. Then*

$$\sqrt{\left[\frac{N}{n}\right]} \cdot \log(\tilde{S}(n)) \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \mathcal{N}(0, \gamma^2(H)),$$

where $\gamma^2(H) > 0$ depends only on H .

This result can be obtained for different lengths of windows satisfying the conditions $N/n \rightarrow \infty$ and $N/n^3 \rightarrow 0$. Let (n_1, \dots, n_m) be such different window lengths. Then, one can write for N and n_i large enough

$$\begin{aligned} \log(\tilde{S}(n_i)) &\simeq \frac{1}{\sqrt{[N/n_i]}} \cdot \varepsilon_i \implies \\ \log(F(n_i)) &\simeq H \cdot \log(n_i) + \frac{1}{2} \log(\sigma^2 f(H)) + \frac{1}{\sqrt{[N/n_i]}} \cdot \varepsilon_i, \end{aligned}$$

with $\varepsilon_i \sim \mathcal{N}(0, \gamma^2(H))$. As a consequence, a linear regression of $\log(F(n_i))$ on $\log(n_i)$ provides an estimation \hat{H}_{DFA} of H . More precisely,

Proposition 2.7 *Under the previous assumptions and notations, let $n \in \{1, \dots, N\}$, $m \in \mathbb{N}^* \setminus \{1\}$, $r_i \in \{1, \dots, [N/n]\}$ for each i with $r_1 < \dots < r_m$ and $n_i = r_i n$ be such that $N/n \rightarrow \infty$ and $N/n^3 \rightarrow 0$ when $N \rightarrow \infty$. Let \hat{H}_{DFA} be the estimator of H , defined as the slope of the linear regression of $\log(F(r_i \cdot n))$ on $\log(r_i \cdot n)$, i.e.*

$$\hat{H}_{DFA} = \frac{\sum_{i=1}^m (\log(F(r_i \cdot n)) - \overline{\log(F)}) (\log(r_i \cdot n) - \overline{\log(n)})}{\sum_{i=1}^m (\log(r_i \cdot n) - \overline{\log(n)})^2}.$$

Then \hat{H}_{DFA} is a consistant estimator of H such that, with $C(H, m, r_1, \dots, r_m) > 0$,

$$E[(\hat{H}_{DFA} - H)^2] \leq C(H, m, r_1, \dots, r_m) \frac{1}{[N/n]}. \quad (2.18)$$

Hence, this result shows that the convergence rate of \hat{H}_{DFA} is $\sqrt{N/n}$ that is a convergence rate $o(N^{1/3})$ from the condition $N/n^3 \rightarrow 0$. This is clearly less accurate than parametric estimators or even quadratic variations estimators. This estimator is devoted to trended long-range time series but even in such frame this estimator does not give satisfying results (see Bardet and Kammoun [4]).

2.4 Increment Ratio Statistic

The Increment Ratio (IR) statistic was first proposed in Surgailis *et al.* [22] in the frame of long-range dependent time series and extended to continuous time processes in Bardet and Surgailis [6]. For a time series $X = (X(k))_{k \in \mathbb{Z}}$ define the second order variation as in 2.7 $D_t^{(a)} = X(t+2a) - 2X(t+a) + X(t)$ for $t \in \mathbb{Z}$ and $a \in \mathbb{N}^*$. Assume that a trajectory $(X(0), X(1), \dots, X(N))$ is observed. Define for $a \in \mathbb{N}^*$,

$$R_N^{(a)} := \frac{1}{N-2a} \sum_{k=0}^{N-3a} \frac{|D_k^{(a)} + D_{k+1}^{(a)}|}{|D_k^{(a)}| + |D_{k+1}^{(a)}|}, \quad (2.19)$$

with the convention $\frac{0}{0} := 1$. Note the ratio on the right-hand side of (2.19) is either 1 or less than 1 depending on whether the consecutive increments $D_k^{(a)}$ and $D_{k+1}^{(a)}$ have

same signs or different signs; moreover, in the latter case, this ratio generally is small whenever the increments are similar in magnitude (“cancel each other”).

If X is a fBm with parameter $H \in (0, 1)$, then it is established in Bardet and Surgailis [6] that for any $H \in (0, 1)$,

$$R_N^{(1)} \xrightarrow[N \rightarrow \infty]{a.s.} \Lambda_2(H) \quad (2.20)$$

$$\sqrt{N}(R_N^{(1)} - \Lambda_2(H)) \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \mathcal{N}(0, \Sigma_2(H)). \quad (2.21)$$

The expressions of $\Lambda_2(H)$ and $\Sigma_2(H)$ are respectively given by:

$$\Lambda_2(H) := \lambda(\rho_2(H)), \quad (2.22)$$

$$\lambda(r) := \frac{1}{\pi} \arccos(-r) + \frac{1}{\pi} \sqrt{\frac{1+r}{1-r}} \log\left(\frac{2}{1+r}\right), \quad (2.23)$$

$$\rho_2(H) := \text{corr}(D_0^{(1)}, D_1^{(1)}) = \frac{-3^{2H(t)} + 2^{2H(t)+2} - 7}{8 - 2^{2H(t)+1}} \quad (2.24)$$

$$\text{and } \Sigma_2(H) := \sum_{j \in \mathbb{Z}} \text{cov} \left(\frac{|D_0^{(1)} + D_1^{(1)}|}{|D_0^{(1)}| + |D_1^{(1)}|}, \frac{|D_j^{(1)} + D_{j+1}^{(1)}|}{|D_j^{(1)}| + |D_{j+1}^{(1)}|} \right). \quad (2.25)$$

and their graphs in Figures 1 and 2.

Figure 1: The graph of $\Lambda_2(H)$.

The central limit (2.21) provides a way for estimating H : indeed, since $H \in (0, 1) \mapsto \Lambda_2(H)$ is an increasing \mathcal{C}^1 function, define

$$\widehat{H}_N = \Lambda_2^{-1}(R_N^{(1)}).$$

From the Delta-method, we obtain the following central limit theorem for \widehat{H}_N :

Figure 2: The graphs of $\sqrt{\Sigma_p(H)}$, $p = 1$ (with a pole at $3/4$) and $p = 2$ (with a pole at $7/4$) (from Stoncelis and Vaičiulis [26])

Proposition 2.8 For all $H \in (0, 1)$,

$$\sqrt{N}(\widehat{H}_N - H) \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \mathcal{N}(0, \gamma^2(H))$$

with $\gamma^2(H) = \Sigma_2(H) [(\Lambda_2^{-1})'(\Lambda_2(H))]^2$

In Bardet and Surgailis [7], in a quite similar frame, an improvement of \widehat{H}_N has been proposed. It consists first on obtaining a central limit theorem for the vector $(R_N^{(1)}, R_N^{(2)}, \dots, R_N^{(m)})$ with $m \in \mathbb{N}^*$ and not only for $R_N^{(a)}$ with $a = 1$. Hence, we obtain the following multidimensional central limit theorem:

$$\sqrt{N}((R_N^{(i)})_{1 \leq i \leq m} - (\Lambda_2^{(i)})_{1 \leq i \leq m}) \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \mathcal{N}(0, \Gamma_m(H))$$

where $\Gamma_m(H) = (\gamma_{ij}(H))_{1 \leq i, j \leq m}$ and

$$\Lambda_2^{(i)}(H) = \lambda(\rho_2^{(i)}(H))$$

$$\text{with } \rho_2^{(i)}(H) = \text{Cor}(D_0^{(i)}, D_1^{(i)}) = \frac{-|2i+1|^{2H} - |2i-1|^{2H} + 4|i+1|^{2H} + 4|i-1|^{2H} - 6}{8 - 2^{2H+1}}$$

$$\text{and } \gamma_{ij}(H) = \sum_{k \in \mathbb{Z}} \text{cov} \left(\frac{|D_0^{(i)} + D_1^{(i)}|}{|D_0^{(i)}| + |D_1^{(i)}|}, \frac{|D_k^{(j)} + D_{k+1}^{(j)}|}{|D_k^{(j)}| + |D_{k+1}^{(j)}|} \right).$$

Then we define:

$$\widehat{H}_N^{(i)} = [\Lambda_2^{(i)}]^{-1}(R_N^{(i)}),$$

and using again the Delta-Method we obtain another multidimensional central limit theorem

$$\sqrt{N}((\widehat{H}_N^{(i)})_{1 \leq i \leq m} - HI_m) \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \mathcal{N}(0, \Delta_m(H))$$

$$\text{and } \Delta_m(H) := \left(\left[\frac{\partial}{\partial x} (\Lambda_2^{(i)})^{-1} (\Lambda_2^{(i)}(H)) \right] \gamma_{ij}(H) \left[\frac{\partial}{\partial x} (\Lambda_2^{(j)})^{-1} (\Lambda_2^{(j)}(H)) \right] \right)_{1 \leq i, j \leq p}.$$

Finally a pseudo-generalized least squares estimator of H can be constructed (like for \widehat{H}_{QV}). Indeed $\Delta_m(H)$ can be estimated by $\widehat{\Delta}_m = \Delta_m(\widehat{H}_N^{(1)})$. Then we define

$$\widehat{H}_{IR} = (I_m^\top (\widehat{\Delta}_m I_m)^{-1} I_m^\top (\widehat{\Delta}_m)^{-1} (\widehat{H}_N^{(i)})_{1 \leq i \leq m}) \quad (2.26)$$

and we obtain this proposition:

Proposition 2.9 For all $H \in (0, 1)$ and $m \in \mathbb{N}^*$,

$$\sqrt{N}(\widehat{H}_{IR} - H) \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \mathcal{N}(0, s^2),$$

with $s^2 = (I_m^\top (\Delta_m(H))^{-1} I_m)^{-1}$.

Then the convergence rate of \widehat{H}_{IR} is \sqrt{N} , confidence intervals can also be easily computed.

2.5 Wavelet based estimator

This approach was introduced for fBm by Flandrin [10], and popularized by many authors to other self-similar or long-range dependent processes (see for instance Veitch and Abry [27], Abry *et al.* [1] or Bardet *et al.* [5]). Here we are going to follow Bardet [3], which is especially devoted to fBm as an extension of Flandrin [10].

First we define a (mother) wavelet function ψ such as $\psi : \mathbb{R} \rightarrow \mathbb{R}$ is a piecewise continuous and piecewise left (or right)-differentiable in $[0, 1]$, such that $|\psi'_l(t)|$ is Riemann integrable in $[0, 1]$ with ψ'_l the left-derivative of ψ , with support included in $[0, 1]$ and Q first vanishing moments, i.e.

$$\int t^p \psi(t) dt = 0 \text{ for } p = 0, 1, \dots, Q-1 \quad (2.27)$$

$$\text{and } \int t^Q \psi(t) dt \neq 0. \quad (2.28)$$

For ease of writing, we have chosen a ψ supported in $[0, 1]$. But all the following results are still true, *mutatis mutandis*, if we work with any compactly supported wavelets. For instance, ψ can be any of the Daubechies wavelets.

Now we define the wavelet coefficients $d(a, i)$ of X where $a \in \mathbb{N}^*$ is called “scale” and $i \in \{1, 2, \dots, [N/a] - 1\}$ is called “shift” by:

$$d(a, i) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} \psi\left(\frac{t}{a} - i\right) X(t) dt = \frac{1}{\sqrt{a}} \int_0^a \psi\left(\frac{t}{a}\right) X(t + ai) dt. \quad (2.29)$$

, For each (a, i) , $d(a, i)$ is a zero-mean Gaussian variable and its variance is a self-similar deterministic function in a , independent of the shift i , since for any $a \in \{1, 2, \dots, [N/2]\}$ and $i \in \{1, 2, \dots, [N/a] - 1\}$,

$$Ed^2(a, i) = a^{2H+1} C_\psi(H) \text{ where } C_\psi(H) = -\frac{\sigma^2}{2} \int \int \psi(t) \psi(t') |t - t'|^{2H} dt dt'.$$

We assume now $C_\psi(H) > 0$ for all $H \in]0, 1[$. Now we consider an empirical variance $I_N(a)$ of $d(a, i)$ by

$$I_N(a) = \frac{1}{[N/a] - 1} \sum_{k=1}^{[N/a]-1} d^2(a, k). \quad (2.30)$$

Using properties of ψ and particularly condition $\int t^p \psi(t) dt = 0$ for $p = 0, 1, \dots, Q-1$, we can show that $\lim_{|i-j| \rightarrow \infty} |\text{cov}(\tilde{d}(a, i), \tilde{d}(a, j))| = 0$ and limit theorems for $I_N(a)$. More precisely,

Proposition 2.10 *Under the previous assumptions, for $1 \leq a_1 < \dots < a_m \in \mathbb{N}^*$, then*

$$\left[\sqrt{\frac{N}{a_i}} (\log I_N(a_i) - (2H+1) \log a_i - \log C_\psi(H)) \right]_{1 \leq i \leq m} \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \mathcal{N}_m(0; F), \quad (2.31)$$

with $F = (f_{ij})_{1 \leq i, j \leq m}$ the matrix with $D_{ij} = \text{GCD}(a_i, a_j)$,

$$f_{ij} = \frac{\sigma^4 D_{ij}}{2C_\psi^2(H) a_i^{2H+1/2} a_j^{2H+1/2}} \sum_{k=-\infty}^{\infty} \left(\int \int \psi(t) \psi(t') |kD_{ij} + a_i t - a_j t'|^{2H} dt dt' \right)^2.$$

When a trajectory $(X(1), \dots, X(N))$ is observed, d_X can not be computed and an approximation has to be considered. Indeed, the wavelet coefficients $d(a, i)$, computed from a continuous process, can not be directly obtained and only approximated coefficients can be computed from a time series. It requires to choose large enough scales to fit well. Here, we will work with approximated coefficients $e(a, i)$ defined by:

$$e(a, i) = \frac{1}{\sqrt{a_i}} \sum_{k=-\infty}^{\infty} \psi\left(\frac{k}{a} - i\right) X(k) = \frac{1}{\sqrt{a}} \sum_{k=0}^a \psi\left(\frac{k}{a}\right) X(k + ai). \quad (2.32)$$

Denote also

$$J_N(a) = \frac{1}{[N/a] - 1} \sum_{k=1}^{[N/a]-1} e^2(a, k). \quad (2.33)$$

The limit theorem of proposition 2.10 can be rewritten with $\tilde{e}(a, i)$ instead of $\tilde{d}(a, i)$. The main difference is the use of scales $a_1(N), \dots, a_m(N)$ satisfying $\lim_{N \rightarrow \infty} a_i(N) = \infty$. More precisely, the limit theorem is the following:

Proposition 2.11 *Let $n_1 < \dots < n_m$ be integer numbers and let $a_i(N) = n_i b(N)$ for $i = 1, \dots, m$ with $b(N)$ a sequence of integer numbers satisfying: $\lceil \frac{N}{b(N)} \rceil \geq 2$, $\lim_{N \rightarrow \infty} b(N) = \infty$ and $\lim_{N \rightarrow \infty} \frac{N}{b^3(N)} = 0$. Then, under previous assumptions,*

$$\sqrt{\frac{N}{b(N)}} (\log J_N(a_i(N)) - (2H+1) \log a_i(N) - \log C_\psi(H))_{1 \leq i \leq m} \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \mathcal{N}_m(0; G),$$

with $G=(g_{ij})_{1 \leq i, j \leq m}$ the matrix with $D_{ij}=\text{GCD}(n_i, n_j)$,

$$g_{ij} = \frac{\sigma^4 D_{ij}}{2C_\psi^2(H)n_i^{2H+1/2}n_j^{2H+1/2}} \sum_{k=-\infty}^{\infty} \left(\int \int \psi(t)\psi(t') |kD_{ij} + n_it - n_jt'|^{2H} dt dt' \right)^2.$$

From conditions on $b(N)$, the best convergence rate of this limit theorem is less than $N^{1/3}$ instead of \sqrt{N} without the discretization problem. It is an important difference for the following estimation of H .

Indeed, Proposition 2.11 provides a method to estimate H from a linear regression. In fact, the central limit theorem of this proposition can be written :

$$\sqrt{\frac{N}{b(N)}} (Y_N - (2H+1)L - KI_m) \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \mathcal{N}_m(0; G),$$

with

- $K = \log C_\psi(H)$, $I_m = (1, \dots, 1)'$, and $L = (\log n_i)_{1 \leq i \leq m}$,
- $Y_N = (\log J_N(a_i(N)) - \log b(N))_{1 \leq i \leq m}$ and $M = (L, I_m)$.

Under assumptions, there exists $\theta = {}^t(2H+1, K)$, such that $Y_N = M\theta + \beta_N$, where β_N is a remainder which is asymptotically Gaussian. By the linearity of this model, one obtains an estimation $\hat{\theta}_1(N)$ of θ by the regression of Y_N on M and ordinary least squares (O.L.S.).

But we can also identify the asymptotic covariance matrix of β_N . Indeed, the matrix G is a function $G_\psi(n_1, \dots, n_m, H)$ and $\hat{G}(N) = G_\psi(n_1, \dots, n_m, \hat{H}_1(N))$ converges in probability to G . So, it is possible to determine an estimation $\hat{\theta}_2(N)$ of θ by generalized least squares (G.L.S.) of H by minimizing

$$\|Y_N - M\theta\|_{\hat{G}(N)^{-1}}^2 = (Y_N - M\theta)' \hat{G}(N)^{-1} (Y_N - M\theta).$$

Thus, from the classical linear regression theory and the Gauss-Markov's Theorem:

Proposition 2.12 *Under previous assumptions,*

1. *The O.L.S. estimator of H is \hat{H}_{OLS} such as $\hat{H}_{OLS} = \left(\frac{1}{2}, 0\right) (M' M)^{-1} M' Y_N - \frac{1}{2}$*

$$\text{and } \sqrt{\frac{N}{b(N)}} (\hat{H}_{OLS} - H) \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \mathcal{N}(0, \sigma_1^2), \text{ with } \sigma_1^2 = \frac{1}{4} (M' M)^{-1} M G M' (M' M)^{-1}.$$

2. *The G.L.S. estimator of H is \hat{H}_{Wave} such as $\hat{H}_{Wave} = \left(\frac{1}{2}, 0\right) (M' \hat{G}(N)^{-1} M)^{-1} M' \hat{G}(N)^{-1} Y_N - \frac{1}{2}$*

$$\text{and } \sqrt{\frac{N}{b(N)}} (\hat{H}_{Wave} - H) \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \mathcal{N}(0, \sigma_2^2), \text{ with } \sigma_2^2 = \frac{1}{4} (M' \hat{G}(N)^{-1} M)^{-1} \leq \sigma_1^2.$$

Hence, as for other semi-parametric estimator of H (see the DFA estimator of H), the convergence rates of \widehat{H}_{Wave} is $N^{1/3-\varepsilon}$, which is less accurate than the convergence rates of Whittle or generalized quadratic variations estimators.

3 Numerical applications and results of simulations

3.1 Concrete procedures of estimation of H

In the previous section, we theoretically defined several estimators of H from a trajectory $(X(1), \dots, X(N))$ of a fBm. Hereafter, we specify the concrete procedure of computation of these estimators:

- The Whittle estimator \widehat{H}_W does not require to select auxiliary parameters or bandwidths. However we can notice that the integrals are replaced by Riemann sums computed for $\lambda = \pi k/n, k = 1, \dots, N$.
- The classical R/S estimator \widehat{H}_{RS} has been computed by averaging on uniformly distributed windows (see Taqqu *et al.* [23]).
- The modified R/S estimator \widehat{H}_{RSM} has been computed using several uniformly distributed values of q around the optimal bandwidth $q = [N^{1/3}]$ as it is given by Lo [18]. More precisely we selected $q = \{[N^{0.3}], \dots, [N^{0.5}]\}$.
- The second order quadratic variations estimator \widehat{H}_{QV} requires the choice of the number of scales. After convincing auxiliary simulations, we selected $p = [3 \log(N)]$.
- The DFA estimator \widehat{H}_{DFA} requires the choice of windows. From the theoretical and numerical study in Bardet and Kammoun [4], we have chosen $n = \{[N^{0.3}], \dots, [N^{0.5}]\}$.
- The Increment Ratio estimator \widehat{H}_{IR} is computed with $M = 5$.
- The wavelet based estimator \widehat{H}_{Wave} is computed with $b(N) = [N^{0.3}]$ and $m = [2 * \log(N)]$.

3.2 Results of simulations

We generated 1000 independent replications of trajectories $(X(1), \dots, X(N))$ for $N = 500$ and $N = 5000$, with X a fBm of parameters $H = 0.1, 0.2, \dots, 0.9$. We applied the estimators of H to these trajectories and compared the Mean Square Error (MSE) for each of them.

4 Conclusion

We studied here several parametric and semi-parametric estimators of the Hurst parameter. In the part, we only consider the fBm. In such frame, we obtained:

H		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$N = 500$	\sqrt{MSE} for \hat{H}_W	0.0280	0.0246	0.0255	0.0277	0.0294	0.0301	0.0310	0.0319	0.0305
	\sqrt{MSE} for \hat{H}_{RS}	0.2148	0.1918	0.1682	0.1420	0.1141	0.0847	0.0566	0.0431	0.0724
	\sqrt{MSE} for \hat{H}_{RSM}	0.0642	0.0660	0.0675	0.0654	0.0651	0.0691	0.0831	0.1061	0.1437
	\sqrt{MSE} for \hat{H}_{QV}	0.0206	0.0263	0.0292	0.0326	0.0339	0.0346	0.0356	0.0361	0.0354
	\sqrt{MSE} for \hat{H}_{DFA}	0.0291	0.0424	0.0574	0.0723	0.0840	0.0994	0.1109	0.1195	0.1293
	\sqrt{MSE} for \hat{H}_{IR}	0.0411	0.0483	0.0549	0.0561	0.0612	0.0624	0.0644	0.0603	0.0511
	\sqrt{MSE} for \hat{H}_{Wave}	0.0970	0.1043	0.0791	0.0666	0.0608	0.0598	0.0606	0.0618	0.0637
	$N = 5000$	\sqrt{MSE} for \hat{H}_W	0.0246	0.0094	0.0088	0.0085	0.0087	0.0090	0.0092	0.0094
\sqrt{MSE} for \hat{H}_{RS}		0.1640	0.1427	0.1191	0.0972	0.0751	0.0515	0.0261	0.0185	0.0568
\sqrt{MSE} for \hat{H}_{RSM}		0.0296	0.0300	0.0303	0.0293	0.0275	0.0309	0.0352	0.0497	0.0757
\sqrt{MSE} for \hat{H}_{QV}		0.0059	0.0077	0.0321	0.0491	0.0013	0.0104	0.0157	0.0107	0.0111
\sqrt{MSE} for \hat{H}_{DFA}		0.0152	0.0271	0.0362	0.0457	0.0510	0.0573	0.0632	0.0682	0.0710
\sqrt{MSE} for \hat{H}_{IR}		0.0119	0.0149	0.0159	0.0177	0.0183	0.0194	0.0192	0.0198	0.0196
\sqrt{MSE} for \hat{H}_{Wave}		0.0666	0.0359	0.0293	0.0281	0.0286	0.0312	0.0328	0.0428	0.0472

Table 1: Values of the (empirical) MSE for the estimators of H when X is a fBm of parameter H and $N = 500, N = 5000$

1. **Theoretically** Only 3 estimators have a \sqrt{N} convergence rate: \hat{H}_W, \hat{H}_{QV} and \hat{H}_{IR} . The first one is specific for fBm, both the other ones can be also applied to other processes. The worst estimators are certainly R/S and modified R/S estimators. Wavelet based and DFA estimators can finally be written as generalized quadratic variations but the works with semi-parametric convergence rate $o(N^{1/2})$, and second-order quadratic variation estimator is clearly more accurate.
2. **Numerically** The ranking between the estimators is clear and follows the theoretical study: the Whittle estimator \hat{H}_W provides the best results, followed by the second-order quadratic variation estimator \hat{H}_{QV} and the IR estimator \hat{H}_{IR} which provides accurate estimations, followed by \hat{H}_{Wave} which is still efficient. The DFA and R/S estimators are not really interesting.

References

- [1] P. Abry, P. Flandrin, M.S. Taqqu and D. Veitch, *Self-similarity and long-range dependence through the wavelet lens*, in Long-range Dependence: Theory and Applications, P. Doukhan, G. Oppenheim and M.S. Taqqu editors, Birkhäuser (2003).
- [2] J.-M. Bardet, Testing for the presence of self-similarity of Gaussian processes having stationary increments, *J. Time Ser. Anal.* **21**, 497–516, (2000).

- [3] J.-M. Bardet, Statistical study of the wavelet analysis of fractional Brownian motion, *IEEE Trans. Inform. Theory* **48**, 991–999 (2002).
- [4] J.-M. Bardet and I. Kammoun, Asymptotic Properties of the Detrended Fluctuation Analysis of Long Range Dependent Processes, *IEEE Trans. Inform. Theory* **54**, 1–13 (2008).
- [5] J.-M. Bardet, G. Lang, E. Moulines and P. Soulier, Wavelet estimator of long range dependent processes, *Statistical Inference for Stochastic Processes*, **3**, 85–99 (2000).
- [6] J.-M. Bardet and D. Surgailis, Measuring the roughness of random paths by increment ratios, *Bernoulli*, **17**, 749–780 (2010)
- [7] J.-M. Bardet and D. Surgailis, A new nonparametric estimator of the local Hurst function of multifractional processes, *Stochastic Processes and Applications*, **123**, 1004–1045 (2012)
- [8] J.-F. Coeurjolly, Identification of multifractional Brownian motion. *Bernoulli* **11**, 987–1008 (2005).
- [9] R. Dahlhaus, Efficient parameter estimation for self-similar processes. *Ann. Statist.* **17**, 1749–1766 (1989).
- [10] P. Flandrin, Wavelet Analysis and Synthesis of Fractional Brownian Motions. *IEEE Transactions on Information Theory* **38** (2): 910–917 (1992).
- [11] R. Fox and M.S. Taqqu, Large-sample properties of parameter estimates for strongly dependent Gaussian time series, *Ann. Statist.* **14**, 517–532 (1986).
- [12] L. Giraitis, P. Kokoszka, R. Leipus and G. Teyssire, Rescaled variance and related tests for long memory in volatility and levels. *J. Econometrics*, **112**, 265–294 (2003).
- [13] X. Guyon and J. Leon, Convergence en loi des h-variations d’un processus gaussien stationnaire. *Ann. Inst. Poincaré*, **25** 265–282 (1989).
- [14] H. Hurst, Long Term Storage Capacity of Reservoirs, *Transactions of the American Society of Civil Engineers*, **116**, 770–799 (1951).
- [15] J. Istas and G. Lang, Quadratic variations and estimation of the local Hölder index of a Gaussian process. *Ann. Inst. Poincaré* **33**, 407–436 (1997).
- [16] A.N. Kolmogorov, Wiener’sche Spiralen und einige andere interessante Kurven im Hilbertschen Raum, *C.R. (Doklady) Acad. Sci. URSS (N.S.)*, **26**, 115–118.
- [17] W. Li, C. Yua, A. Carriquiry and W. Kliemann, The asymptotic behavior of the R/S statistic for fractional Brownian motion, *Statistics and Probability letters*, **81**, 83–91 (2011).

- [18] A.W. Lo, Long-term memory in stock market prices, *Econometrica* **59**, 1279–1313 (1991).
- [19] B. Mandelbrot and J. Van Ness, Fractional Brownian motion, fractional noises and applications. *SIAM review* **10**, 422–437 (1968).
- [20] Peng C.K., Havlin S., Stanley H.E., Goldberger A.L., Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series, *Chaos* **5**, 82 (1995).
- [21] Y.G. Sinai, Self-Similar Probability Distributions. *Theory Probab. Appl.*, **21**, 64–80 (1976).
- [22] D. Surgailis, G. Teyssière and M. Vaičiulis, The increment ratio statistic, *J. Multiv. Anal.* **99**, 510–541 (2008).
- [23] M.S. Taqqu, V. Teverovsky and W. Willinger, Estimators for long-range dependence: an empirical study, *Fractals*, **3**, 785–798 (1995).
- [24] V. Teverovsky, M.S. Taqqu and W. Willinger, A critical look at Lo’s modified R/S statistic, *J. Statist. Plann. Inference* **80**, 211–227 (1999).
- [25] G. Samorodnitsky and M.S. Taqqu, *Stable non-Gaussian random processes*. Stochastic modeling. Chapman & Hall (1994).
- [26] M. Stoncelis and M. Vaičiulis, Numerical Approximation of Some Infinite Gaussian Series and Integrals, *Nonlinear Analysis: Modelling and Control*, **13**, 397–415 (2008).
- [27] D. Veitch and P. Abry, A wavelet-based joint estimator of the parameters of long-range dependence, *IEEE Trans. on Information Theory* **45**, 878–897 (1999).
- [28] P. Whittle, *Gaussian estimation in stationary time series*. Bulletin of the International Statistical Institute **39**, 105-129 (1962).