

# ANALYSIS OF A HURST PARAMETER ESTIMATOR BASED ON THE MODIFIED ALLAN VARIANCE

ALESSANDRA BIANCHI, STEFANO BREGNI, IRENE CRIMALDI, AND MARCO FERRARI

ABSTRACT. In order to estimate the Hurst parameter of Internet traffic data, it has been recently proposed a log-regression estimator based on the so-called modified Allan variance (MAVAR). Simulations have shown that this estimator achieves higher accuracy and better confidence when compared with other methods of common use. Here we link it to the wavelets setting and provide an asymptotic analysis in the case the signal process is a fractional Brownian motion. In particular we show that the MAVAR log-regression estimator is consistent and asymptotically normal, providing the related confidence intervals for a suitable choice on the regression weights. Finally, we show some numerical examples.

## 1. INTRODUCTION

Internet traffic, as well as many different kinds of real data (Hydrology, Economics, Biology), has been demonstrated to exhibit self-similarity and long-range dependence (LRD) on various time scales [1, 2, 3]. In a self-similar random process, a dilated portion of a realization, by the scaling Hurst parameter  $H$ , has the same statistical characterization than the whole. On the other hand, the LRD is commonly equated to an asymptotic power-law decrease of the spectral density of a related stationary random process, and it is thus characterized by the exponent  $\alpha$  of such a power-law.

Though a self-similar process can not be stationary (and thus nor LRD), these two proprieties are often related in the following sense. Under the hypothesis that a self-similar process has stationary (or weakly stationary) increments, the scaling parameter  $H$  enters in the description of the spectral density of the increments, providing an asymptotic power-law with exponent  $\alpha = 2H - 1$ . The most paradigmatic example of this connection is given by the fractional Brownian motion and by its increment process, the fractional Gaussian noise [4].

Among the different techniques introduced in the literature in order to estimate the Hurst parameter  $H$ , here we focus on a method based on the log-regression of the Modified Allan Variance (MAVAR). The MAVAR is a well known time-domain quantity generalizing the classic Allan variance [5, 6]. It has been proposed for the first time as a traffic analysis tool in [7], and then its performance has been evaluated by simulation [7, 8] and also on real IP traffic [9]. These works have pointed out the high accuracy

---

*Date:* February 8, 2012.

*2000 Mathematics Subject Classification.* 62M10,62M15,62G05.

*Key words and phrases.* Hurst parameter, long-range dependence, self-similarity, modified Allan variance, parameter estimation, wavelets, fractional Brownian motion.

Work partially supported by the GDRE 224 GREFI-MEFI-CNRS-INdAM and by the GNAMPA 2011 Projects.

of the method in estimating the parameter  $H$ , and have shown that it achieves a highest confidence if compared with the well-established log-diagram based on Daubechies wavelets. In [10], the authors substantiate these results from the theoretical point of view, studying the limit properties of the MAVAR log-regression estimator. In particular, under the assumption that the signal process is a fractional Brownian motion, they deduce its precise asymptotic normality.

In the present work we provide a wavelet representation of the MAVAR estimator that can be seen as a generalization of the well known connection between Allan variance and Haar-wavelets family [11]. Applying the asymptotic results of [10] to the internet traffic data setting, we show that under the fractional Brownian motion hypothesis the MAVAR log-regression estimator is asymptotically unbiased and consistent. We provide explicit formulas for the related confidence intervals and numerical evaluations of such formulas. We show that the variances decrease with the rate predicted by their analytical asymptotes. Besides, we suggest different choices for the regression-weights.

## 2. SELF-SIMILARITY AND LONG-RANGE DEPENDENCE

According to [3], we consider a centered self-similar real-valued stochastic process  $X = \{X(t), t \in \mathbb{R}\}$ , with  $X(0) = 0$ , that can be interpreted as the signal process. By *self-similarity* of  $X$  we refer to the existence of a parameter  $H \in (0, 1)$ , called *Hurst index* or *Hurst parameter* of the process, such that, for all  $a > 0$ , it holds

$$\{X(t), t \in \mathbb{R}\} \stackrel{d}{=} \{a^{-H}X(at), t \in \mathbb{R}\}. \quad (1)$$

Assuming further that the process  $X$  has *weakly stationary* increments, we get the following expression for the autocovariance function

$$\text{Cov}(X(s), X(t)) = \frac{\sigma_H^2}{2} (|t|^{2H} - |t-s|^{2H} + |s|^{2H}), \quad (2)$$

with  $\sigma_H^2 := \mathbb{E}[X^2(1)]$ . Denoting by  $Y_\tau$  the  $\tau$ -increment process of  $X$ , defined as  $Y_\tau(t) = \frac{X(t+\tau) - X(t)}{\tau}$ , it also turns out that the autocovariance function of  $Y_\tau$ , given by  $R_{Y_\tau}(t) = \text{Cov}(Y_\tau(s), Y_\tau(s+t))$ , satisfies asymptotically the following power law [1]

$$R_{Y_\tau}(t) \sim \sigma_H^2 H(2H-1)|t|^{2H-2} \quad \text{as } |t| \rightarrow +\infty.$$

In particular, if  $H \in (\frac{1}{2}, 1)$ , the process  $Y_\tau$  displays *long-range dependence*, in the sense that there exists  $\alpha = 2H - 1 \in (0, 1)$  such that the spectral density of the process  $Y_\tau$ ,  $f_{Y_\tau}(\lambda)$ , satisfies the condition

$$f_{Y_\tau}(\lambda) \sim c|\lambda|^{-\alpha} \quad \text{as } \lambda \rightarrow 0,$$

for some finite constant  $c \neq 0$ . The parameter  $\alpha$  is often referred as *memory parameter* of the process  $Y_\tau$  [12, 13, 14]. Thus, under the assumption that  $X$  is a self-similar process with weakly stationary increments, we embrace the two main empirical properties of a wide collection of real data.

A basic example of the connection between these two properties is provided by the *fractional Brownian motion*  $B_H = \{B_H(t), t \in \mathbb{R}\}$ , that is a centered Gaussian process

with autocovariance function given by (2) with

$$\sigma_H^2 = \frac{1}{\Gamma(2H + 1) \sin(\pi H)}. \quad (3)$$

It can be shown that  $B_H$  is a self-similar process with Hurst index  $H \in (0, 1)$ , which corresponds, for  $H = 1/2$ , to the standard Brownian motion. Moreover, its increment process

$$G_{\tau, H}(t) = \frac{B_H(t + \tau) - B_H(t)}{\tau},$$

called *fractional Gaussian noise*, turns out to be a weakly stationary Gaussian process [4, 15], displaying long memory for  $H > \frac{1}{2}$ .

### 3. THE MODIFIED ALLAN VARIANCE

In this section we introduce and recall the main properties of the Modified Allan variance (MAVAR) [5, 16], and of the log-regression estimator of the Hurst parameter based on it [7, 8, 9].

**3.1. Definition of MAVAR and related estimator.** Let  $\tau_0 > 0$  be the “sampling period” and define the sequence of times  $\{t_k\}_{k \geq 1}$  taking  $t_1 \in \mathbb{R}$  and setting  $t_i - t_{i-1} = \tau_0$ , i.e.  $t_i = t_1 + \tau_0(i - 1)$ . For any integer  $p \geq 1$ , we set  $\tau = \tau_0 p$  and define the modified Allan variance (MAVAR) as

$$\sigma_{p, \tau_0}^2 := \frac{1}{2\tau^2} \mathbb{E} \left[ \left( \frac{1}{p} \sum_{i=1}^p (X_{t_i+2\tau} - 2X_{t_i+\tau} + X_{t_i}) \right)^2 \right] \quad (4)$$

where  $\mathbb{E}[\cdot]$  is the mean value operator over the signal process  $X$  [5]. For  $p = 1$  we recover the well-known Allan variance.

Let us assume that a finite sample  $X_1, \dots, X_n$  of the process  $X$  is given, and that the observations are taken at times  $t_1, \dots, t_n$ , with constant sampling period  $\tau_0$ . In other words we set  $X_i = X_{t_i}$  for  $i = 1, \dots, n$ . For  $k \in \mathbb{Z}$ , let us define

$$d_{p, \tau_0, k} := \frac{1}{\sqrt{2} \tau p} \sum_{i=1}^p (X_{k+i+2p} - 2X_{k+i+p} + X_{k+i}), \quad (5)$$

and notice that, from the hypotheses on  $X$  of Sec. 2,  $\{d_{p, \tau_0, k}\}_k$  is weakly stationary. Moreover, by definitions (4) and (5),  $\sigma_{p, \tau_0}^2 = \mathbb{E} [d_{p, \tau_0, k}^2]$ .

A standard estimator for the modified Allan variance (MAVAR estimator) is given by

$$\widehat{\sigma}_{p, \tau_0}^2(n) := \frac{1}{n_p} \sum_{k=0}^{n_p-1} d_{p, \tau_0, k}^2, \quad (6)$$

for  $p \in \{1, \dots, \lfloor n/3 \rfloor\}$  and  $n_p := n - 3p + 1$ , where the space-average  $\mathbb{E}[\cdot]$  is replaced by the empirical average over the observations sample.

**3.2. MAVAR and wavelet estimators.** Consider the generalized process  $Y = \{Y(t), t \in \mathbb{R}\}$  defined through the set of identities

$$\int_{t_1}^{t_2} Y(t)dt = X_{t_2} - X_{t_1}, \quad \forall t_1, t_2 \in \mathbb{R}. \quad (7)$$

In short, we write  $Y = \dot{X}$ . With this definition we can rewrite the MAVAR and its related estimator as functions of the process  $Y$ . In particular we can write

$$d_{p,\tau_0,k} = \frac{1}{\sqrt{2}p^2\tau_0} \sum_{i=1}^p \left( \int_{t_{i+k+p}}^{t_{i+k+2p}} Y(t)dt - \int_{t_{i+k}}^{t_{i+k+p}} Y(t)dt \right). \quad (8)$$

Now we claim that, for  $p$  fixed, this function can be put in correspondence with a family of discrete wavelet transforms of the process  $Y$ , indexed by  $\tau_0$  and  $k$ . To see that, let us fix  $j \in \mathbb{N}$  and set  $\tau_0 = 2^j$  and  $t_1 = 2^j$ , so that  $t_i = 2^j i$ , for all  $i \in \mathbb{N}$ . With this choice on the sequence of times, it is not difficult to construct a mother wavelet  $\psi(s)$  such that

$$\begin{aligned} d_{k,j} &:= d(2^j, p, k) = \langle Y; \psi_{k,j} \rangle \\ \text{with } \psi_{k,j}(s) &:= 2^{-j} \psi(2^{-j}s - k). \end{aligned} \quad (9)$$

An easy check shows that the function  $\psi(s) := \sum_{i=1}^p \psi^i(s)$ , where

$$\psi^i(s) := \frac{1}{\sqrt{2}p^2} (I_{[i+p, i+2p]}(s) - I_{[i, i+p]}(s)), \quad (10)$$

is a proper mother wavelet satisfying Eq. (9). Notice also that the components  $\psi^i$ ,  $i = 1, \dots, p$ , of  $\psi$  are suitably translated and re-normalized Haar functions. In the case  $p = 1$ , corresponding to the classical Allan variance, the mother wavelet is exactly given by the Haar function, as was already pointed out in [11].

Although the MAVAR can be related to the above Haar-type wavelets family, we will show that the MAVAR and wavelets log-regression estimators do not match as the regression runs on different parameters. For the wavelet-based estimators  $p$  is fixed and the regression parameter is  $j$  (related to  $\tau_0$ ), while for the MAVAR estimator (see Eq. (14)) the regression is on  $p$  with  $\tau_0$  fixed. Because of this difference, it not possible to apply the results available in the wavelets framework [12, 13, 14]. Hence we developed new arguments, which in turns allow us to avoid some technical troubles due to the poor regularity of the Haar-type functions.

**3.3. The MAVAR log-regression estimator.** As proven in [10], applying the covariance formula (2) we get, for  $H \in (1/2, 1)$ ,

$$|\sigma_{p,\tau_0}^2 - \sigma_H^2 \tau^{2H-2} K(H)| \leq \sigma_H^2 \tau^{2H-2} O_H(p^{-1}), \quad (11)$$

where

$$K(H) := \frac{2^{2H+4} + 2^{2H+3} - 3^{2H+2} - 15}{2(2H+1)(2H+2)}. \quad (12)$$

This asymptotic relation suggests the following estimation method for the parameter  $H$ .

Let  $n$  be the sample size, i.e. the number of the observations, choose  $\bar{p}, \bar{\ell} \in \mathbb{N}$  and an increasing sequence  $\{a_\ell\}_{\ell \in \mathbb{N}}$  such that  $1 \leq \bar{p}a_{\bar{\ell}} \leq p_{\max}(n) = \lfloor n/3 \rfloor$ . Let  $\underline{w} = (w_0, \dots, w_{\bar{\ell}})$  be a vector of weights satisfying the conditions

$$\sum_{\ell=0}^{\bar{\ell}} w_\ell = 0 \quad \text{and} \quad \sum_{\ell=0}^{\bar{\ell}} w_\ell \log(a_\ell) = 1. \quad (13)$$

The MAVAR log-regression estimator associated to the weights  $\underline{w}$  is defined as

$$\hat{\mu}_n(\tau_0, \bar{p}, \underline{w}) := \sum_{\ell=0}^{\bar{\ell}} w_\ell \log(\hat{\sigma}_{a_\ell \bar{p}, \tau_0}^2(n)). \quad (14)$$

Roughly speaking, the idea behind this definition is to use the approximation  $\hat{\sigma}_{a_\ell \bar{p}, \tau_0}^2(n) \simeq \sigma_{a_\ell \bar{p}, \tau_0}^2(n)$  in order to get, by (11) and (13),

$$\begin{aligned} \hat{\mu}_n(\tau_0, \bar{p}, \underline{w}) &\simeq \sum_{\ell=0}^{\bar{\ell}} w_\ell \log(\sigma_{a_\ell \bar{p}}^2) \\ &\simeq \sum_{\ell=0}^{\bar{\ell}} w_\ell \log(\sigma_H^2(\tau_0 a_\ell \bar{p})^\mu K(H)) = \mu, \end{aligned}$$

where  $\mu := 2H - 2$ . Thus, given the data  $X_1, \dots, X_n$  the following procedure is used to estimate  $H$ :

- compute the modified Allan variance by (6) for integer values  $a_\ell \bar{p}$ , with  $1 \leq a_\ell \bar{p} \leq p_{\max}(n) = \lfloor n/3 \rfloor$ ;
- compute the weighted MAVAR log-regression estimator by (14) in order to get an estimate  $\hat{\mu}$  of  $\mu$ ;
- estimate  $H$  by  $\hat{H} = (\hat{\mu} + 2)/2$ .

#### 4. THE ASYMPTOTIC NORMALITY OF THE ESTIMATOR

In [10], under the assumption that  $X$  is a *fractional Brownian motion with Hurst index*  $H \in (1/2, 1)$ , two convergence results are proven in order to justify the above approximations and to get the rate of convergence of  $\hat{\mu}_n(\tau_0, \bar{p}, \underline{w})$  toward  $\mu = 2H - 2$ .

In particular, it is shown that if  $\bar{p} = \bar{p}(n)$  is a sequence of integers such that  $\bar{p}(n) \rightarrow +\infty$ ,  $n\bar{p}(n)^{-1} \rightarrow +\infty$  and  $n\bar{p}(n)^{-3} \rightarrow 0$  as  $n \rightarrow +\infty$ , then (for a fixed  $\bar{\ell}$ )

$$\rho_n(\underline{w}, H)^{-1} (\hat{\mu}_n(\bar{p}, \underline{w}) - \mu) \xrightarrow[n \rightarrow +\infty]{d} \mathcal{N}(0, 1) \quad (15)$$

with

$$\rho_n(\underline{w}, H) \sim c(\underline{w}, H) \sqrt{\frac{\bar{p}}{n}} \xrightarrow[n \rightarrow +\infty]{} 0,$$

where  $c(\underline{w}, H)$  is a suitable constant depending on  $\underline{w}$  and  $H$ . Two important consequences of (15) are the following:

- 1) The MAVAR log-regression estimator is *consistent*, i.e. the bias  $(\hat{\mu}_n - \mu)$  converges in probability to zero.

- 2) Given an estimate of  $H$ , say  $\hat{H}$ , and the corresponding estimate of the normalizing coefficients  $\rho_n(\underline{w}, \hat{H})$ , we get the *asymptotic confidence interval* for the parameter  $H$ :

$$\hat{H} - q_{1-\beta/2} \frac{\rho_n(\underline{w}, \hat{H})}{2} \leq H \leq \hat{H} + q_{1-\beta/2} \frac{\rho_n(\underline{w}, \hat{H})}{2} \quad (16)$$

where  $q_{1-\beta/2}$  is the  $(1 - \beta/2)$ -quantile of the standard normal distribution. The length of the confidence interval is  $q_{1-\beta/2} \rho_n(\underline{w}, \hat{H})$ .

The coefficient  $\rho_n^2(\underline{w}, H)$  can be approximated by the following quantity (see [10])

$$\frac{1}{K(H)^2} \sum_{\ell=0}^{\bar{\ell}} \sum_{\ell'=0}^{\bar{\ell}} \left( \frac{a_{\ell \vee \ell'}}{a_{\ell \wedge \ell'}} \right)^{2H+2} \frac{w_\ell w_{\ell'}}{n_\ell n_{\ell'}} \sum_{k=0}^{n_{\ell \vee \ell'} - 1} \sum_{k'=0}^{n_{\ell \wedge \ell'} - 1} \Phi_H \left( \frac{k-k'}{\bar{p} a_{\ell \vee \ell'}}, \frac{|a_\ell - a_{\ell'}|}{a_{\ell \vee \ell'}} \right), \quad (17)$$

with  $K(H)$  given in (12),  $\ell \vee \ell' = \max\{\ell, \ell'\}$ ,  $\ell \wedge \ell' = \min\{\ell, \ell'\}$ ,  $n_\ell := n - 3a_\ell \bar{p} + 1$  and

$$\Phi_H(x, y) := \int_{\mathbb{R}} \left[ \gamma_H(x, 0, r) \gamma_H(0, y, r) \int_{-\infty}^r \gamma_H(x, 0, v) \gamma_H(0, y, v) dv \right] dr.$$

with

$$\begin{aligned} \gamma_H(x, y, v) := & [(H + 1/2)\Gamma(H + 1/2)\sigma_H]^{-1} \left\{ [(x + 3(1 - y) - v)^+]^{H+1/2} + \right. \\ & \left. - 3 [(x + 2(1 - y) - v)^+]^{H+1/2} + 3 [(x + (1 - y) - v)^+]^{H+1/2} - [(x - v)^+]^{H+1/2} \right\} \end{aligned}$$

## 5. WEIGHTS

The explicit expression of the weights clearly depends on the sequence  $\{a_\ell\}$ . In our investigation we considered the linear progression,  $a_\ell = 1 + \ell$ , and the geometrical progression,  $a_\ell = r^\ell$  with  $r > 1$ . The latter has provided the better numerical results which are then presented in the next section. Here we focus on the geometrical progression sequence and we give explicit formulas for three particular weights proposed in the literature for the log-regression procedure.

- The simple linear regression weights are defined as

$$w_\ell^{SLR} := \frac{(\ell - m(\bar{\ell}))}{\log(r) \sum_{\ell=0}^{\bar{\ell}} (\ell - m(\bar{\ell}))^2}$$

with

$$m(\bar{\ell}) := (\bar{\ell} + 1)^{-1} \sum_{\ell=0}^{\bar{\ell}} \ell = \frac{\bar{\ell}}{2}$$

- Following Abry and Veitch [11], we can take

$$w_\ell^{AV} := \frac{(\ell - m(\bar{\ell})) r^{-\ell}}{\log(r) \sum_{\ell=0}^{\bar{\ell}} (\ell - m(\bar{\ell}))^2 r^{-\ell}}$$

where

$$m(\bar{\ell}) := \frac{\sum_{\ell=0}^{\bar{\ell}} \ell r^{-\ell}}{\sum_{\ell=0}^{\bar{\ell}} r^{-\ell}}.$$

- Following Fay, Moulines, Roueff and Taquq [17], we can compute a preliminary estimate, say  $\hat{H}^{(1)}$ , of  $H$  (for instance, applying the first or the second method) and then use it in order to compute the weights

$$\underline{w}^{FMRT} = w^{FMRT}(n, \hat{H}^{(1)}) := \frac{1}{\log(r)} D^{-1} B (B^T D^{-1} B)^{-1} \underline{b}$$

where

$$\underline{b} := \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad B := \begin{pmatrix} 1 & 1 & \dots & 1 \\ 0 & 1 & \dots & \bar{\ell} \end{pmatrix}^T,$$

and  $D = D(n, \hat{H}^{(1)})$  is the symmetric matrix with entries (for a generic  $H$ )

$$[D(n, H)]_{\ell', \ell} := \frac{r^{(\ell \vee \ell') 4H}}{r^{(\ell \wedge \ell') 4}} \frac{1}{n_{\ell} n_{\ell'}} \sum_{k=0}^{n_{\ell \vee \ell'} - 1} \sum_{k'=0}^{n_{\ell \wedge \ell'} - 1} \Phi_H \left( \frac{k-k'}{r^{\ell \vee \ell'} \bar{p}}, \frac{|r^{\ell} - r^{\ell'}|}{r^{\ell \vee \ell'}} \right).$$

## 6. NUMERICAL RESULTS

In this section we present some numerical results that provide the approximated variance of the MAVAR estimator,  $\rho_n^2(\underline{w}, H)$ , and the relative confidence intervals for the three different weights listed above. The numerical evaluations have been realized for different choices on the parameters, and the most interesting results are presented, with comments, in the next figures and tables.

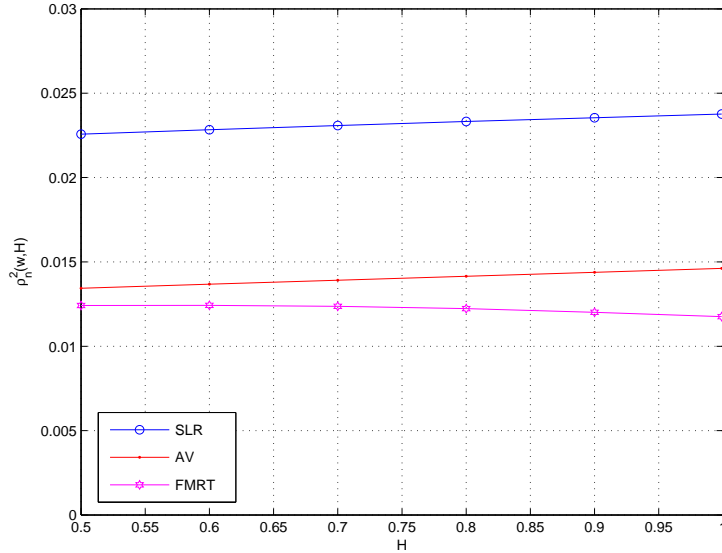


FIGURE 1. Trend of  $\rho_n^2(\underline{w}, H)$  as a function of the Hurst parameter  $H$ , with SLR-, AV- and FMRT-weights and geometrical progression  $p = \bar{p}r^\ell$ ,  $\bar{p} = \lfloor n^\delta \rfloor$  with  $\delta = 0.35$ ,  $r = 2$  and  $n = 4096$ .

The variance of the MAVAR estimator is almost unchanged with  $H$  as shown in Fig. 1, thus for our analysis we used a fixed  $H = 0.7$  to reduce the parameters space. We have first investigated the behavior of  $\rho_n^2(\underline{w}, H)$  as a function of  $n$  with  $p$  following a geometrical growth  $\bar{p}r^\ell$ , for  $0 \leq \ell \leq \bar{\ell}$ . In order to satisfy the hypotheses which are behind convergence (15), the value of  $\bar{p}$  has been chosen as  $\bar{p} = \bar{p}(n) = \lfloor n^\delta \rfloor$  with  $\delta = 0.35$ . (Formally, any value of  $\delta \in (1/3, 1)$  is admissible, but the best results are obtained for  $\delta$  close to  $1/3$ ). We used two distinct values for the parameter  $r$  of the geometrical progression,  $r = 1.1$  and  $r = 2$ , and for each of them we fixed a value  $\bar{\ell}$  with the only restriction that  $\bar{p}r^{\bar{\ell}} \leq \lfloor n/3 \rfloor$ .

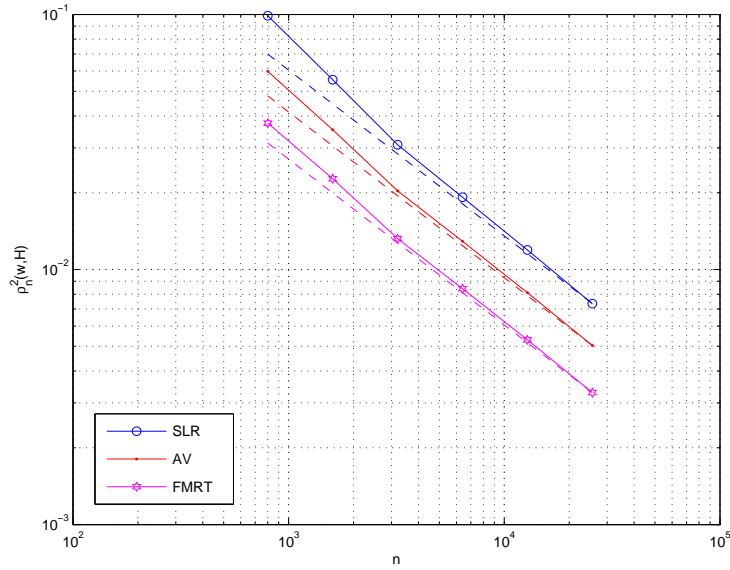


FIGURE 2. Trend of  $\rho_n^2(\underline{w}, H)$  as a function of  $n$ , with SLR-, AV- and FMRT-weights and geometrical progression  $p = \bar{p}r^\ell$ ,  $\bar{p} = \lfloor n^\delta \rfloor$  with  $\delta = 0.35$ ,  $r = 1.1$ ,  $\bar{\ell} = 30$  and  $H = 0.7$ .

In Fig. 2 we plot the results for  $r = 1.1$  and  $\bar{\ell} = 30$ : Each color and marker is associated to one of the three weights (in short, SLR-, AV- and FMRT-weights) of the previous section as listed in the legend; the lines with markers show the results obtained by numerical evaluation of (17), while the dashed lines represent the corresponding theoretical asymptotes  $n^{\delta-1} = n^{-0.65}$  (see Eq. (15)).

Using the same notation, in Fig. 3 we plot the results for  $r = 2$  and taking  $\bar{\ell} = 4$ . The two figures show that the approximation formula (17) for the variance  $\rho_n^2(\underline{w}, H)$ , that we have used for the numerical evaluation, provides results which are in quite good agreement with the theoretical behavior (dashed line), independently on the choice on the weights and on the other parameters. In particular, for  $n$  sufficiently large, we get very small values of  $\rho_n^2(\underline{w}, H)$  and thus small confidence intervals.

In tables 1 and 2, we list the value of the confidence intervals related to Fig. 2, namely for  $r = 1.1$  and, respectively,  $\bar{\ell} = 30$  and  $\bar{\ell} = 45$ . In Table 3 we list the value of the confidence intervals related to Fig. 3, namely for  $r = 2$  and  $\bar{\ell} = 4$ .



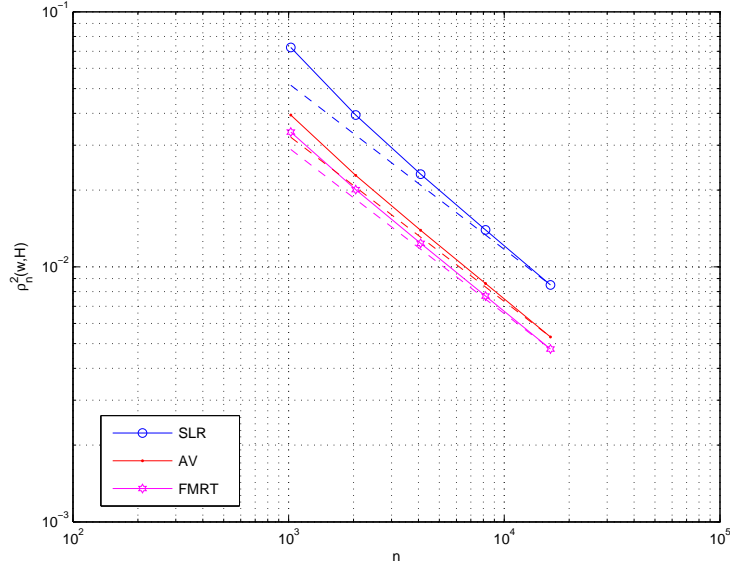


FIGURE 3. Trend of  $\rho_n^2(\underline{w}, H)$  as a function of  $n$ , with SLR-, AV- and FMRT-weights and geometrical progression  $p = \bar{p}r^\ell$ ,  $\bar{p} = \lfloor n^\delta \rfloor$  with  $\delta = 0.35$ ,  $r = 2$ ,  $\bar{\ell} = 4$  and  $H = 0.7$ .

$n$	3200	4096	6400	12800	25600
$\bar{p}$	16	18	21	27	34
$1.96\rho_n(w_{SLR}, H)$	0.3442	0.3197	0.2716	0.2141	0.1680
$1.96\rho_n(w_{AV}, H)$	0.2796	0.2605	0.2227	0.1765	0.1391
$1.96\rho_n(w_{FMRT}, H)$	0.2253	0.2101	0.1798	0.1427	0.1124

TABLE 1. Length of the confidence interval as a function of  $n$ , with SLR-, AV- and FMRT-weights and geometrical progression  $p = \bar{p}r^\ell$ ,  $\bar{p} = \lfloor n^\delta \rfloor$  with  $\delta = 0.35$ ,  $r = 1.1$ ,  $H = 0.7$  and  $\bar{\ell} = 30$ .

$n$	4096	8192	16384
$\bar{p}$	18	23	29
$1.96\rho_n(w_{SLR}, H)$	0.4601	0.3263	0.2425
$1.96\rho_n(w_{AV}, H)$	0.2562	0.1955	0.1509
$1.96\rho_n(w_{FMRT}, H)$	0.2005	0.1564	0.1220

TABLE 2. Length of the confidence interval as a function of  $n$ , with SLR-, AV- and FMRT-weights and geometrical progression  $p = \bar{p}r^\ell$ ,  $\bar{p} = \lfloor n^\delta \rfloor$  with  $\delta = 0.35$ ,  $r = 1.1$ ,  $H = 0.7$  and  $\bar{\ell} = 45$ .

Comparing the results displayed in tables 1-3, as  $r$  and  $\bar{\ell}$  vary, we can deduce that

$n$	4096	8192	16384
$\bar{p}$	18	23	29
$1.96\rho_n(w_{SLR}, H)$	0.2978	0.2315	0.1807
$1.96\rho_n(w_{AV}, H)$	0.2312	0.1818	0.1429
$1.96\rho_n(w_{FMRT}, H)$	0.2180	0.1718	0.1352

TABLE 3. Length of the confidence interval as a function of  $n$ , with SLR-, AV- and FMRT-weights and geometrical progression  $p = \bar{p}r^\ell$ ,  $\bar{p} = \lfloor n^\delta \rfloor$  with  $\delta = 0.35$ ,  $r = 2$ ,  $\bar{\ell} = 4$  and  $H = 0.7$ .

- the best (smaller) value of  $\rho_n^2(\underline{w}, H)$  is obtained at  $r = 2$  for the SLR- and AV-weights, and at  $r = 1.1$  for the FMRT-weight;
- the value of  $\rho_n^2(\underline{w}, H)$  is also sensitive of  $\bar{\ell}$  and in particular, for  $r = 1.1$ , the value  $\bar{\ell} = 45$  provides better results for the AV- and FMRT-weights, while  $\bar{\ell} = 30$  provides better results for the SLR-weight.

This last point suggests us to investigate the effect of an increase of  $\bar{\ell}$  over  $\rho_n^2(\underline{w}, H)$ . We thus evaluate the variance as a function of  $\bar{\ell}$ , taking fixed  $n = 4096$ ,  $H = 0.7$ , and  $r = 1.1$ . The trend is shown in Fig. 4 and Table 4 lists the related confidence intervals. As one can see, whereas with SLR-weight there exists an optimal choice of  $\bar{\ell}$  (approximately 20 in this setting), with the AV- and FMRT-weights the total variance keeps decreasing with increasing  $\bar{\ell}$ .

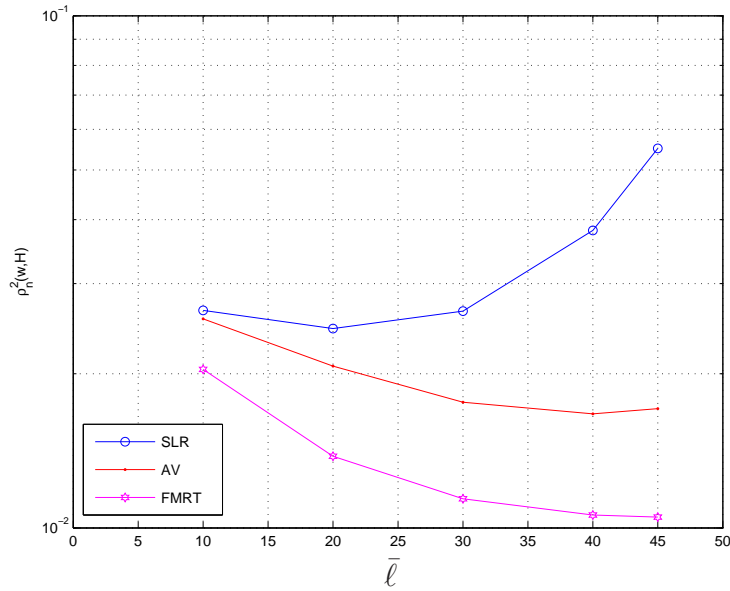


FIGURE 4. Trend of  $\rho_n^2(\underline{w}, H)$  as a function of  $\bar{\ell}$ , with SLR-, AV- and FMRT-weights and geometrical progression  $p = \bar{p}r^\ell$ . Here  $n = 4096$ ,  $\bar{p} = 18$ ,  $r = 1.1$  and  $H = 0.7$ .

$\bar{\ell}$	10	20	30	40	45
$1.96\rho_n(w_{SLR}, H)$	0.3197	0.3068	0.3191	0.3826	0.4605
$1.96\rho_n(w_{AV}, H)$	0.3136	0.2820	0.2600	0.2533	0.2568
$1.96\rho_n(w_{FMRT}, H)$	0.2799	0.2302	0.2093	0.2018	0.2011

TABLE 4. Length of the confidence interval as a function of  $\bar{\ell}$ , with SLR- AV- and FMRT-weights and geometrical progression  $p = \bar{p}r^{\ell}$ ,  $\bar{p} = 18$ ,  $n = 4096$ ,  $r = 1.1$ , and  $H = 0.7$ .

#### REFERENCES

- [1] ABRY, P., FLANDRIN, P., TAQQU, M.S. and VEITCH, D. (2000). Wavelets for the analysis, estimation and synthesis of scaling data. *Park, K., Willinger, W. (Eds.), Self-Similar Network Traffic and Performance Evaluation*, 39–88, Wiley (Interscience Division), New York.
- [2] PAXSON, V. and FLOYD, S. (1995). Wide-area traffic: the failure of Poisson modeling. *IEEE/ACM Trans. Networking* **3**, no. 6, 226–244.
- [3] PARK, K. and WILLINGER, W. (2000). Self-similar network traffic: an overview. *Park, K., Willinger, W. (Eds.), Self-Similar Network Traffic and Performance Evaluation*, 1–38, Wiley (Interscience Division), New York.
- [4] MANDELBROT, B. B. and VAN NESS, J. W. (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Rev.* **10**, 422–437.
- [5] BERNIER L. G. (1987). Theoretical Analysis of the Modified Allan Variance. *Proc. 41st Annual Frequency Control Symposium*.
- [6] BREGNI, S. (2002). Characterization and Modelling of Clocks, in *Synchronization of Digital Telecommunications Networks*, John Wiley & Sons.
- [7] BREGNI, S. and PRIMERANO, L. (2004). The modified Allan variance as timedomain analysis tool for estimating the Hurst parameter of long-range dependent traffic. *Proc. IEEE GLOBECOM*.
- [8] BREGNI, S. and JMODA, L. (2008). Accurate Estimation of the Hurst Parameter of Long-Range Dependent Traffic Using Modified Allan and Hadamard Variances. *IEEE Transactions on Communications* **56**, No. 11, 1900–1906.
- [9] BREGNI, S. and ERANGOLI, W. (2005). Fractional noise in experimental measurements of IP traffic in a metropolitan area network. *Proc. IEEE GLOBECOM*.
- [10] BIANCHI, A., CAMPANINO M. and CRIMALDI, I. (2011). Asymptotic normality of a Hurst parameter estimator based on the modified Allan variance. Submitted. Electronic preprint is available at <http://amsacta.cib.unibo.it/3140/>
- [11] ABRY, P. and VEITCH, D. (1998). Wavelet analysis of long-range-dependent traffic. *IEEE Trans. Inform. Theory* **44**, no. 1, 2–15.
- [12] MOULINES, E., ROUEFF, F. and TAQQU, M.S. (2007). Central limit theorem for the log-regression wavelet estimation of the memory parameter in the Gaussian semi-parametric context. *Fractals* **15**, no. 4, 301–313.
- [13] MOULINES, E., ROUEFF, F. and TAQQU, M.S. (2007). On the spectral density of the wavelet coefficients of long-memory time series with application to the log-regression estimation of the memory parameter. *J. Time Ser. Anal.* **28**, no. 2, 155–187. MR2345656
- [14] MOULINES, E., ROUEFF, F. and TAQQU, M.S. (2008). A wavelet Whittle estimator of the memory parameter of a nonstationary Gaussian time series. *Ann. Statist.* **36**, no. 4, 1925–1956.
- [15] YAGLOM, A. M. (1987). Correlation theory of stationary and related random functions. Vol. I. Basic results. *Springer Series in Statistics. Springer-Verlag, New York*.
- [16] ALLAN, D. W. and BARNES, J. A. (1981). A Modified Allan Variance with Increased Oscillator Characterization Ability. *Proc. 35th Annual Frequency Control Symposium*.

- [17] FAÿ, G., MOULINES, E., ROUEFF, F. and TAQQU, M.S. (2009). Estimators of long-memory: Fourier versus wavelets. *J. Econometrics* **151**, no. 2, 159–177.

A. BIANCHI, DIP. DI MATEMATICA - UNIVERSITÀ DI BOLOGNA, PIAZZA DI PORTA S. DONATO 5, 40126 BOLOGNA, ITALY

*E-mail address:* `alessandra.bianchi7@unibo.it`

S. BREGNI, DEPT. OF ELECTRONICS AND - POLITECNICO DI MILANO, PIAZZA LEONARDO DA VINCI 32, 20133 MILANO, ITALY

*E-mail address:* `bregni@elet.polimi.it`

I. CRIMALDI, IMT INSTITUTE FOR ADVANCED STUDIES, PIAZZA SAN PONZIANO 6, 55100 LUCCA, ITALY

*E-mail address:* `irene.crimaldi@imtlucca.it`

M. FERRARI, DEPT. OF ELECTRONICS AND - POLITECNICO DI MILANO, PIAZZA LEONARDO DA VINCI 32, 20133 MILANO, ITALY

*E-mail address:* `marco.ferrari@elet.polimi.it`