Comparison of different methods for computing scaling parameter in the presence of trends.

María E. Torres(1,✉) y Patrice Abry (2)

(1)Laboratorio de Señales y Dinámicas No Lineales - Facultad de Ingeniería, Universidad Nacional de Entre Ríos,
CC 47, Suc. 3 (CP 3100), Paraná, Entre Ríos, Argentina- (E-mail: metorres@ceride.gov.ar)
(2) CNRS, UMR 5672, Laboratoire de Physique, Ecole Normale Supérieure de Lyon,
69364 Lyon Cedex 7, France. (E-mail: Patrice.Abry@ens-lyon.fr)

Abstract

In a large number of analysis of biological signals, such as ElectroEncephaloGrams or Heart Rate Variability, phenomena such as long range correlation or long term memory or self similarity or more generally scaling phenomena have been observed and are often used as a basis for diagnosis and the assertion of major physiological facts. It is therefore an important issue to be able to perform a relevant analysis (detection, estimation) of such phenomena on empirical data. For a reliable use of long-range correlation analysis, it is essential to be able to distinguish between trends (whatever their origin - external, apparatus - or internal but related to other biological mechanism such as breathing...) and long-range fluctuations intrinsic to the data. In this paper we compare estimates of the Hurst parameter based on i) Detrending Fluctuation Analysis, ii) wavelet analysis, iii) discrete variations. We discuss their capabilities for a reliable estimation of the scaling exponent, depending on the signal length and for different trends, using simulated and real data.

I. Introduction

In a large number of analysis of biological signals, such as Electroencephalograms or Heart Rate Variability, phenomena such as long range correlation or long term memory or self similarity or more generally scaling phenomena have been observed and are often used as a basis for diagnosis and the assertion of major physiological facts. It is therefore an important issue to be able to perform a relevant analysis (detection, estimation) of such phenomena on empirical data. However, the empirical analysis of long memory and more generally of scaling involves important difficulties, mainly because of potential confusion with non stationarities (whatever their origin - external, apparatus - or internal but related to other biological mechanism such as breathing...). Smooth deterministic trends (such as polynomial or sinusoidal trends) possibly superimposed to the data to be analyzed are likely either to be taken for long memory or to significantly alter the quality of the scaling exponent estimation when scaling actually exists. For a reliable use of long-range correlation analysis, it is therefore essential to be able to distinguish between trends and long-range fluctuations intrinsic to the data.

To address those issues we use fractional Brownian motion, the only Gaussian self similar process, as the reference model for scaling phenomena and long range memory and we consider three different estimators for the self-similarity parameter based on i) Detrended Functional Analysis, ii) Wavelet Analysis, iii) Discrete Variations. The aim is to compare the estimation performance of these three estimators when applied to short duration data to which deterministic trends are superimposed.

The organization of the paper is as follows: in section II we recall some theoretical concepts. In section III we briefly summarize the three estimation methods here discussed. Results and discussions are presented in sec. IV. A large number of realizations of fractional Brownian motion plus trends are numerically synthesized, to which the three estimators are applied. Comparisons of estimation performance are obtained by averaging over the realizations and studied. An application to real physiological data is presented. In section V we present our conclusions.

II. Fractional Brownian motion

Self-similar processes have been widely used as the first natural model that can be thought of model scaling. Among them, fractional Brownian motion (hereafter fBm), the only Gaussian self-similar process with stationary increments, is of great theoretical and practical interest. It is essentially characterized by a single self similarity parameter H, it accounts for long range dependence, it can be easily numerically synthesized and numerous theoretical results related to its properties are available in the literature. It is defined as follows [4,9,11,12,16]:

✉ Author to whom the correspondence has to be submitted: M. E. Torres.

metorres@ceride.gov.ar

This work is supported by U.N.E.R., under Project PID 07/060. The author also thanks the CNRS support for making possible her stage at the ENSL during which this work has been performed.

Definition. Self-similarity. Let Y_t be a stochastic process with continuous time parameter t. Y_t is called self-similar with self-similarity parameter H, if and only if the original process Y_t. And the rescaled process with time scale ct, $c^{-H}\,Y_{ct}$ (for any positive stretching factor), have all their finite dimensional distributions equal.

Definition. Stationary increments. Let Y_t be a stochastic process with continuous time parameter t. Y_t is said to have stationary increments if and only if the finite dimensional distributions of the process $X_t = Y_{t+1}-Y_t$ do not depend on *t*.

Definition Fractional Brownian motion: Fractional Brownian motion B_H is the only (up to a multiplicative constant) Gaussian self-similar process with stationary increments.

The increment process of fBm $G_H(k) = B_H(k+1)-B_H(k)$ is known as the fractional Gaussian noise (hereafter, fGn). It can easily be shown [4,16] that its correlation function reads:

$$r_{G_H}(k) = s^2\!\big/\!2\,\left[(k+1)^{2H} - 2\,k^{2H} + (k-1)^{2H}\right],$$

and asymptotically decreases as H (2H-1) $k^{(2-2H)}$ when $k\to\infty$. For 1/2<H<1, this means that the correlation functions decays to zero so slowly that $\sum_{-\infty}^{\infty} r(k) = \infty$, and fGn is said to have long memory or long-range dependence. For H=1/2, the samples $G_H(k)$ are uncorrelated. If 0<H<1/2 the correlations are summable, in this case the process is said to have short-range dependence but still presents a power law decrease for its correlation function.

III. Self similarity parameters estimation methods

Various methods were proposed in the literature to estimate the self-similarity parameter of fBm. For a review, see e.g. [17,18]. We here concentrate on three of them known from the literature to present the most interesting characteristics in terms of performance as well as robustness with respects to additional trends. They are based on detrended functional analysis (hereafter DFA) [10,13], wavelet analysis (hereafter WA) [3] , and discrete variations (hereafter DV) [5] and their definitions are briefly recalled below.

III.1 Detrending Fluctuation Analysis

Given a time series (Y_i), the profile of the record of length n is determined: $X(i) = \sum_{k=1}^{i} Y_k - \overline{Y}$, where $\overline{Y}$ indicates the mean value of (Y_i). The profile *X(i)* is cut into $n_s \cong [n/s]$ non-overlapping segments of equal length *s* for different values of *s*. Next, the local trend for each segment is calculated by an unweighted linear fit of the corresponding data. The detrended time series of segment duration s is defined as: $X_s(i) = X(i) - p_N(i)$, where $p_N(i)$ is the fitting polynomial in the η-th segment. For each of the n_s segments the variance of the DFA fluctuation function is computed as $F(s) = 1\!/\!n_s \sum_{h=1}^{n_s} F_s^2(h)$, where $F_s^2(h) = 1\!/\!s \sum_{i=1}^{s} X_s^2[(h-1)s+i]$ is the fluctuation function. It can be shown that $F(s) \cong s^{\alpha}$ for large s values, where the fluctuation exponent α is related to the correlation exponent [10]. When Y is fBm with parameter H, it has been shown that $\hat{H}_{DFA} = 1 - \alpha/2$ [4]. The estimator for H consists in an unweighted linear fit in the coordinate log_2(s) *versus* log_2 F(s):

$$\hat{H}_{DFA} = 1 - 1\!/\!2\,\log(F(s)). \qquad [1]$$

Practical parameters: As we are interest in comparing with wavelet based methods, we cut the profile in logarithmically equally spaced segments between decades $s=2^{smin}$ and $s=2^{smax}$. The values of smin and smax have been selected such that $2^{smin} \le 2+N$, where N is the polynomial order, and smax so as to have at least 10 windows of maximum length. We have selected smin = 3 and smax=$\log_2$(n)-5 ($\log_2$(n)-4 for short signals).

III.2 Wavelet analysis

Let $y_{j,k}(t) = 2^{-j/2}\, y(2^{-j}t - k)$ denotes the templates of the mother wavelet y . Let $\{d_Y(j,k) = \langle Y, y_{j,k}\rangle\}$ be the set of the so-called discrete wavelet coefficients of the process *Y*. Let $S_n(j) = \frac{1}{n_j}\sum_k |d_Y(j,k)|^2$, where n_j is the available number of wavelet coefficients at octave *j*. Essentially $n_j = 2^{-j}\,n$ where *n* is the length of the data. The estimator for *H* consists in an unweighted linear fit in the coordinates $log_2(2^j) = j$ versus $log_2\,S_n(j)$:

$$\hat{H}_W = 1/2 \sum_{j=j_1}^{j_2} w_j \log_2 S_n(j) - 1/2, \qquad [2]$$

where $w_j = (S_0 j - S_1)/(S_0 S_2 - S_1^2)$, with $S_m = \sum_{j_1}^{j_2} j^m$, m = 0,1,2.

When Y is fBm with parameter H, it has been shown that $\hat{H}_W$ is an efficient, weakly biased, robust estimator for H [2,1,5].

Practical parameters: in this paper, we used Daubechies wavelets with *N* vanishing moments and have selected $j_1 = 3$ and $j_2 = \log_2(n)-5$.

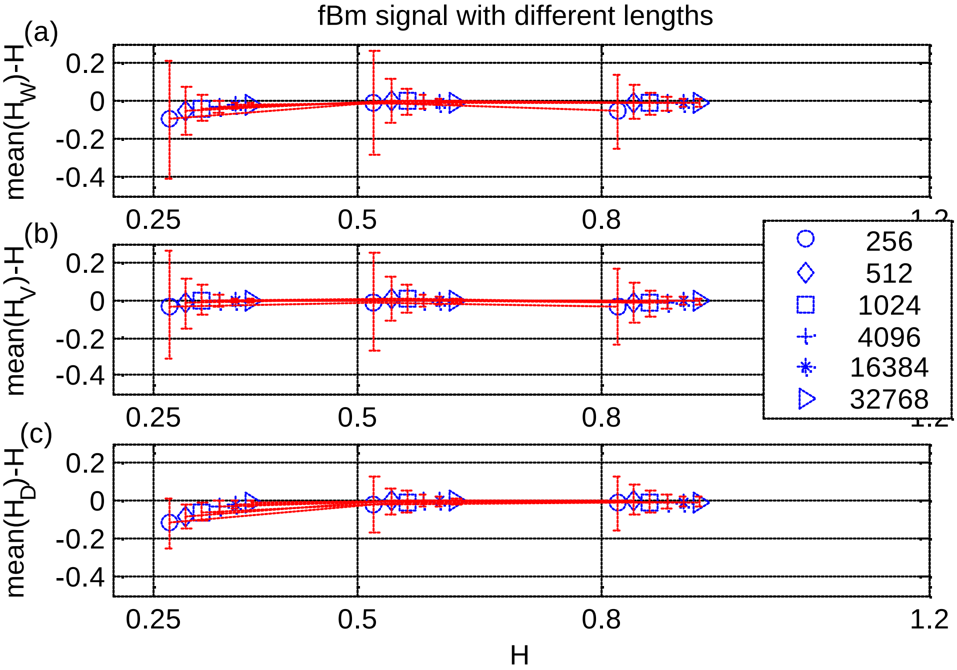

Figure 1 : Hurst exponent computed for 200 realizations of fBm, simulated for each different length. (a) Bias and std corresponding to $\hat{H}_D$. (b) Bias and std corresponding to $\hat{H}_W$. (c) Bias and std corresponding to $\hat{H}_V$.

III.3 Discrete Variations

Let $f(t) = \sum_1 a_1 d(t-1)$, where δ(t) stands for the Dirac mass. It is requested that the a_l satisfy:

$$\sum_{j=0}^{1} j^r a_j = 0, \quad \text{for } r = 0,\dots,N\text{-}1; \text{ and } \quad \sum_{j=0}^{1} j^N a_j \neq 0.$$

Note that the integer *N* plays a role totally equivalent to that of the number of vanishing moments in the wavelet method case. Let $f_{j,k}(t) = 2_{j/2} f(2^{-j} t\text{-}k)$ denotes the templates of f.

Let $\{V_Y(j,k) = \langle Y, f_{j,k} \rangle\}$ be the set of the discrete variation coefficients of the process *Y*. Let $T_n(j) = 1/n_j \sum_k |V_Y(j,k)|^2$, where n_j is the available number of coefficients at octave j. Essentially $n_j=2^{-j}\,n$ where n is the length of the data. The estimator for H consists in an unweighted linear fit in the coordinates $log_2(2^j) = j$ *versus* $log_2 T_n(j)$:

$$\hat{H}_V = 1/2 \sum_{j=j_1}^{j_2} w_j \log_2 S_n(j) - 1/2 \ , \qquad [3]$$

where the w_j are defined as above.

When Y is fBm with parameter H, it has been shown that $\hat{H}_W$ is an efficient, unbiased, robust estimator for H [8, 5].

Practical parameters: in this paper, we used sequences a_l such that { s $a_l = (-1)^l\, l\,!\, /\, (N\,!\, (N-l)\,!)$, l = 0,…,N }, where s is a constant such that $\sum_l a^2 l = 1$ and have selected j_1 and j_2 as above .

IV. Results and discussion

IV.1 Methodology

In order to study the estimation performance of the three estimators described above, we synthesized *200* realizations of fBm, $B_H(k)$, for different values of *H* and different length *n*, using the circulant matrix method [19]. We superimposed additive trends T(k) to $B_H(k)$. The time series to be analyzed Y is obtained as: $Y_k = B_H(k) + \sigma_T\, T(k)$ where $\sigma_T = (\max(T)-\min(T))/\mathrm{std}(B_H)$. The mean values and standard deviation for the three estimators $\hat{H}_W$, $\hat{H}_V$ and $\hat{H}_D$ are obtained from averaging over the realizations.

The practical parameters values of the parameters chosen here are: *nbreal* = 200 (for fBm) and 100 (for fBm+trend), $2^8 \leq n \leq 2^{15}$, H=[0.1, 0.25, 0.4, 0.5, 0.65, 0.7, 0.8, 0.9, 0.95], N=1,2,3,4; σ_T has been computed in each realization.

In Section IV.2, $T(t) \approx 0$ and in Section IV.3 $T(t) = A \sin(2\pi f t)$ with A=1, $f=2^{-11}$.

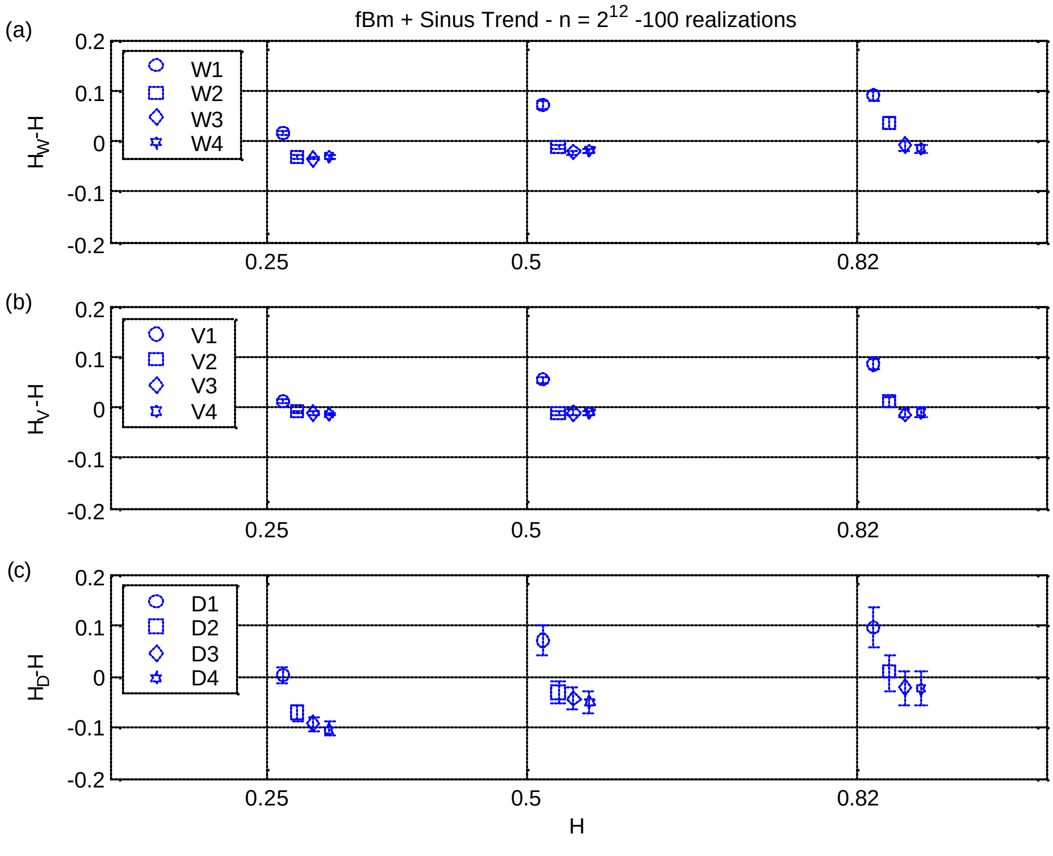

Figure 2: Hurst exponent computed for 100 realizations of fBm (length $n=2^{12}$) with a sinusoidal trend superimposed. Bias and std corresponding to : (a) $\hat{H}_D$. (b) $\hat{H}_W$ .and (c) $\hat{H}_V$.

IV.2 Numerical simulations I: Dependency on the signal length

In this section, no trend has been superimposed to the increments of fBm ($T(t) \approx 0$) and we compare the performance of the estimator with respect to the observation duration n. Biases and standard deviations for the three estimators are compared in Fig. 1 below.

In fig. 1 we present the results obtained for H= 0.25, 0.5 and 0.8. We can appreciate that the bias is related with the signal length. We notice that for large n values, the three estimators present quasi-negligible biases. However, when the number of samples n decreases, all three estimators show a bias that increases while H decreases. It is also to be noted that for the smallest values of H and n, the bias of $\hat{H}_D$ is significantly large compared to that of $\hat{H}_V$ and $\hat{H}_W$, which remain comparable with a slight advantage to the former.

IV.3 Numerical simulations II: Sinusoidal trend superimposed to fGn

In this section, we study the effects of the presence of a superimposed trend $T(t) = A \sin(2\pi f t)$, with A=1, and $f = 2^{-11}$ on the performance of the estimators.

In fig. 2 the results obtained for 100 realizations of length 2^{12} are displayed, comparing the performance of our estimators as functions of the order. Again we notice that $\hat{H}_V$ and $\hat{H}_W$ are negligibly biased for order ≥ 2, and that their standard deviations are smaller than the ones corresponding to $\hat{H}_D$. All the estimators show a bias that increases while H increases when the order (moment) is 1, but it decreases as H increases for orders (moments) larger than 1. The bias in this cases is negligible for the estimators based on wavelet and discrete variations, with an advantage in the last one, given that $\hat{H}_V$ presents a bias of almost the same magnitude for all the H values considered.

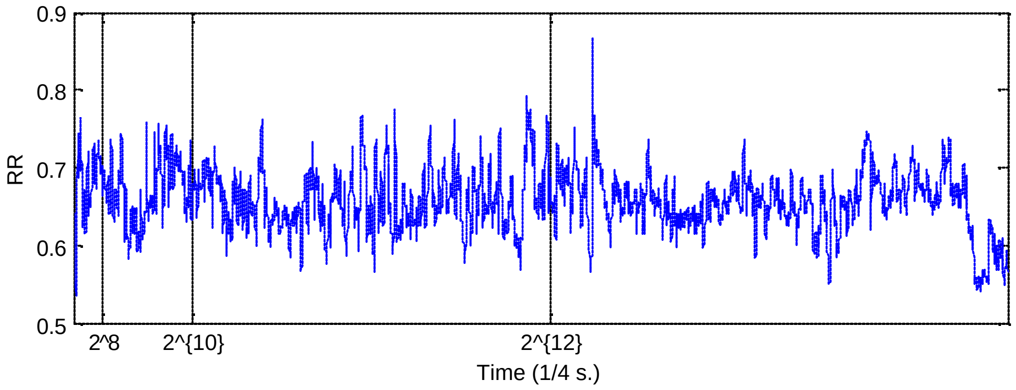

Figure 3: Segment of a RR signal.

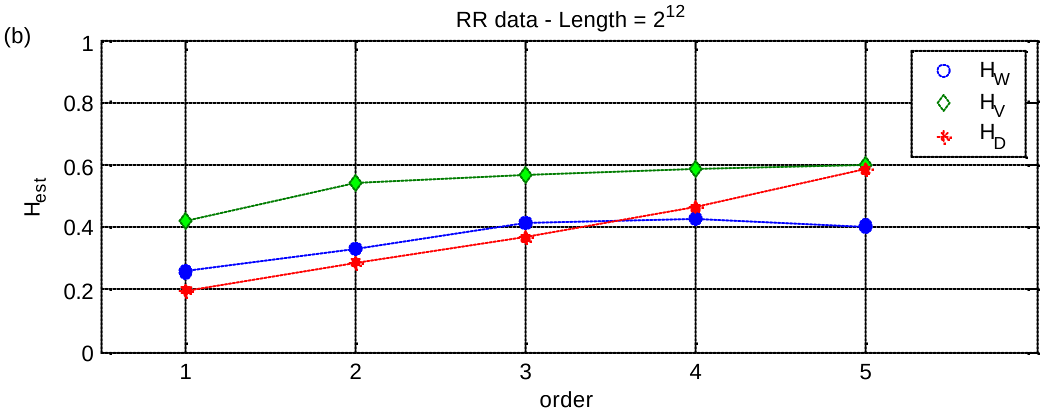

Figure 4: Estimated value of H for a segment of the RR signal shown in (a) of length 2^{12}.

Parameters values considered: s_{min}= 3; s_{max}=12-4 = 8; and orders 1, 2, 3, 4 and 5.

IV.4 Real Physiological Data: RR signals

Let us now consider a RR time series obtained from Physionet European ECG data base [15], corresponding to heart beat fluctuations.

In [6, 7,14] the authors assert that the time series of human heart rate show “noisy” fluctuations. According to classical physiologic paradigms based on homeostasis, such systems should be designed to damp out noise and settle down to a constant equilibrium-like state. However, analysis of heartbeat fluctuations under apparently steady-state conditions reveals, according to the same authors, the presence of long-range correlations. The discovery of such long-range organization poses a remarkable challenge to contemporary efforts to understand and eventually simulate physiological control systems. Plausible models must account for such long-range ``memory".

We observe in fig.3 a segment of a normal RR time series. For this data we have computed the Hurst exponent considering different signal lengths $n = 2^8, 2^{10}, 2^{12}, 2^{13}$ and 2^{14}. In each case we compare the three methods for different polynomial and wavelet orders N=1, 2, 3, 4, 5.

As can be observed in fig. 4, the estimated values of H for the three methods here discussed are highly sensitive to the order N used in the estimation, while applied to the RR segment of length 2^{12} displayed in fig. 4.a. In fact, looking at this figure we are not in conditions to assert if $\hat{H}$ is larger than 0.5 or not. The discrete variations method offers $\hat{H}_V > 0.5$ for moments N=2-5, while the wavelet based method gives $\hat{H}_W < 0.5$ for all the moment values here considered. As far as DFA is concerned, $\hat{H}_D$ steadily increases as the order N goes from 1 to 5.

Let us know observe in fig. 5 the behavior of the three methods for different lengths and orders (moments). We can appreciate that for N=1, both wavelet and DFA methods offer estimations of H <0.5, while increments method does the same only for signals of length $n>2^{10}$. As far as we increase the order (moment) we observe that, independently of the signal length, all the H estimates are also increased. In particular, $\hat{H}_D$ and $\hat{H}_V$ arrive to values slightly larger than 0.5 in the case N=5.

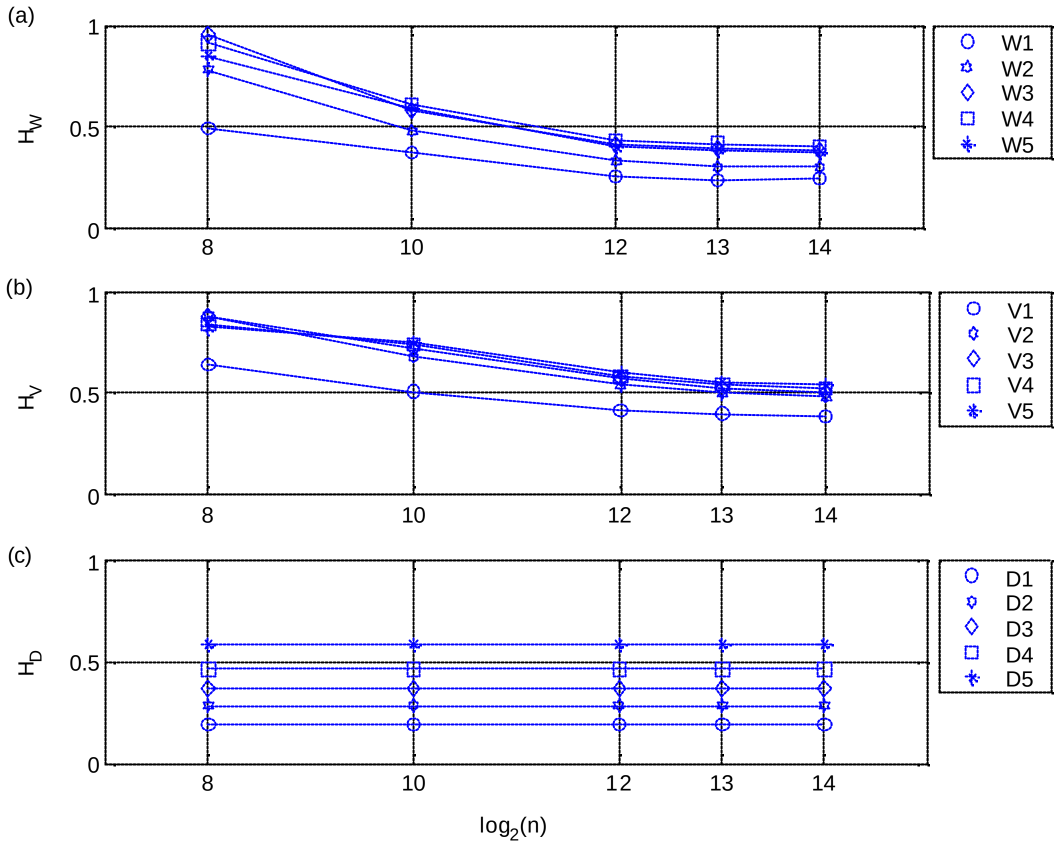

Figure 5: Estimated values of H for segments of the RR signal shown in (a) of lengths 2^8, 2^{10}, 2^{12}, 2^{13} y 2^{14}. Parameters values considered: $s_{min}= 3$; $s_{max}=$ length - 4; and orders 1, 2, 3, 4 and 5.

If we should only consider the estimations based on the discrete variations based methods we should say that this RR time series corresponds to a LRD process probably moving towards to an anti persistency state. However, DFA results for N<5 say that a LRD is not such a clear option since the very beginning of the time series. According to the results obtained for the synthesized signals in previous sections, the former one does not appear to be an appropriate interpretation of the situation, given the amount of data corresponding to this hypothesis (less than 2^{12}). Further and more careful analysis of these signals should be performed.

V. Conclusions

The goal of this paper was to compare different methods to estimate Hurst exponent, and to discuss their efficiency while applied to time series of different lengths. In particular, we focused in short length signals, as many times happens in the available real physiological data. We have discussed the behavior of three estimators in the case of 200 synthesized fBm realizations and 100 fBm with a superimposed sinusoidal trend. We have seen that the three methods here discussed not only are highly dependent on the signal length but also on the order or number of moments (polynomial, wavelet regularity or discrete variations). In the case of real data, we have also shown that for the same signal, depending on the methods and the parameters used, one can get results indicating either a self-similar short-range dependency or a long-range dependency. These analyses indicate that more care should be taken while trying to obtain physiological conclusions from results obtained using a single method or more than one, but with short length signals. The authors have observed similar behavior in other data provided from other colleagues suspecting LRD or self-similarity on them.

We consider that reporting the Hurst exponent is meaningful if only one method and order is considered. Researchers should use several estimators in order to identify long-range dependency or self-similarity. A decomposition analysis of the signal should also be necessary, given that it provides useful information on the possible hidden trends. Finally we consider that more care is imposed when scientists try to derive physiological meanings from such kind of experiments, in particular when the results obtained are in contradiction with well established physiological laws.

Bibliography

[1] P. Abry, P. Flandrin, M. S. Taqqu, and D. Veitch. *Wavelets for the analysis, estimation and synthesis of scaling data*, Wiley Interscience, 2000.

[2] P. Abry, P.Gonçalvès and P.Flandrin. Wavelet-based spectral analysis of *1/f* processes. *ICASSP-93, 1993 IEEE International Conference on*, vol. III, pp 237-240, April 1993.

[3] P. Abry, P. Gonçalvès and P. Flandrin, Wavelets, spectrum estimation and *1/f* processes, in A. Antoniadis and G. Oppenheim, eds. *Wavelets and Statistics*, *Lectures Note in Statistics*, **103**, pp. 15-30. Springer-Verlag, New York, 1995.

[4] J. Beran. *Statistics for Long-Memory Processes*. New York: Chapman & Hall, 1994. (*CRC* Press, USA, 1998).

[5] Jean-François Coeurjolly. Estimating the parameters of a fractional Brownian motion by discrete variations of its sample paths. *Statistical Inference for Stochastic Processes*, **4**: 199-227, 2001.

[6] A. L. Goldberger *et al.* Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals, circulation. **101**(23): e215-e220.

[7] A. L. Goldberger {\it et al.}, Fractal dynamics in physiology: Alterations with disease and aging, 2002

[8] J. Istas and G. Lang. Quadratic variations and estimation of the local h\ older index of a Gaussian process. *Ann. Inst. Henri Poincaré*, **33**(4): 407-436, 1997.

[9] A. N. Kolmogorov. The local structure of turbulence in incompressible viscous fluid for very large Reynolds number. *Dokl. Akad. Nauk USSR*, **30**: 299-303, 1941.

[10] Jan W.Kantelhardt *et al.* Detecting long-range correlations with detrended fluctuation analysis. *Physica A*, **295**: 441-454, 2001.

[11] B.B.Mandelbrot. *Fractals: Form, chance and dimension*. W. H. Freeman, San Francisco, 1977.

[12] B.B.Mandelbrot. *The Fractal Geometry of Nature*. W. H. Freeman, San Francisco, 1983.

[13] C-K Peng *et al.* Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. CHAOS, **1**:82-87, 1995.

[14] C-K Peng *et al.* Fractal mechanisms in neural control: Human heartbeat and gait dynamics in health and disease. Chapter in *Self-Organized Biological Dynamics and Nonlinear Control*, Walleczek J, ed., Cambridge: Cambridge University Press, 2000.

[15] European *ST-T* database, 2003.
In http://www.physionet.org/physiobank/database/edb/.

[16] G. Samorodnitsky, M. S. Taqqu, *Stable Non-Gaussian Processes: Stochastic Models with Infinite Variance*. Chapman and Hall, New York, London, 1994.

[17] M. S. Taqqu, V. Teverovsky, and W. Willinger, Estimators for long-range dependence: an empirical study, *Fractals*, **3**(4):785-798, 1995. (Reprinted in C.J.G. Evertsz, H-O Peitgen and R.F. Voss, eds. *Fractal Geometry and Analysis*, World Scientific Publishing Co., Singapore, 1996.)

[18] V. Teverovsky, M. S. Taqqu, Testing for long-range dependence in the presence of shifting means or a slowly declining trend using a variance-type estimator, *Journal of Time Series Analysis*, **18**:279-304, 1997.

[19] A. T. A. Wood and G. Chan. Simulation of stationary gaussian processes in [0;1]d.,
J. of Computational and Graphical Statistics, **3**:409-432, 1994.