

A Comparative Evaluation of Long-memory Estimators through Wavelet-based Simulation¹

Una Valutazione Comparativa di Stimatori di Memoria Lunga attraverso Simulazione basata su Wavelet

Livia De Giovanni

Università del Molise. Via F. De Sanctis, 86100 Campobasso.
e-mail: degiovanni@unimol.it

Maurizio Naldi

Università di Roma "Tor Vergata". Via del Politecnico 1, 00133 Roma.
e-mail: naldi@disp.uniroma2.it

Riassunto: Il problema dell'inferenza statistica del parametro di memoria lunga in processi stazionari del secondo ordine è stato risolto nel caso parametrico mediante lo stimatore di Whittle. Per il caso semiparametrico le tecniche esistenti sono basate su una analisi asintotica e quindi non comparabili statisticamente. In questo lavoro viene pertanto presentata una analisi comparativa numerica di due stimatori semiparametrici, basati rispettivamente su decomposizione wavelet e sullo stimatore di Whittle, mediante simulazione basata su wavelet.

Keywords: Wavelets, Long Memory, Statistical Inference, Telecommunications.

1. Introduction

Recent measurements of data traffic in telecommunications networks show that the packet arrival process exhibits long range dependence (LRD or long memory) increments. Statistical inference is applied to evaluate the network performance (adversely affected by LRD), in particular to estimate the LRD parameter. If the spectral density function of the process is known this problem is solved by the Whittle estimator (parametric approach). In a semiparametric framework, when the spectral density function is known only close to the origin, existing methods have been validated only through an asymptotic analysis. The aim of this paper is to evaluate two semiparametric estimators (see Section 3) for finite sample size. Wavelets are used to simulate the long memory process (see Section 2). The evaluation results are reported in Section 4.

2. Wavelet-based analysis and synthesis

The wavelet-based *analysis* of a function (extended in (Cambanis 1995) to stochastic processes) provides the following generalized Fourier series representation, where the $\psi_{jk}(x)$ are the wavelet functions (i.e. the orthogonal basis functions) and the symbol $\langle a, b \rangle$ denotes the inner product of the functions a and b (Daubechies, 1992):

¹ The present paper is financially supported by MIUR, also in the context of the ALINWEB Project

$$f(x) = \sum_j \sum_k d_{jk} \psi_{jk}(x), \quad d_{jk} = \frac{\langle f, \psi_{jk} \rangle}{\langle \psi_{jk}, \psi_{jk} \rangle} \quad (1).$$

The whole set of wavelets can be derived from a single function, named the mother wavelet $\psi(x)$, through operations of scaling and dilation/contraction. The coefficients $\{d_{jk}\}$ represent the Discrete Wavelet transform (DWT) of $f(x)$. Among the several possibilities for the choice of the wavelets the Haar wavelets are often used because their support is finite and the coefficient computation procedure is simple. While the trigonometric functions employed in the Fourier series are extremely localized in frequency but completely unlocalized in time, wavelets are well localized both in time and frequency. Wavelets are also used for the *synthesis* of a stochastic process, i.e. to simulate a process with assigned mean, variance and LRD parameter.

3. Semiparametric estimators of long-memory

An important class of LRD processes is derived by self similar processes. A continuous parameter stochastic process $\{Z(t); t \geq 0\}$ is H self similar ($0 < H < 1$) if for any positive c the process $\{c^{-H} Z(ct); t \geq 0\}$ is equal in distribution to the original process $\{Z(t); t \geq 0\}$. A discrete parameter (second order) stationary process $\{X_j; j \geq 1\}$ with variance σ^2 is LRD if the correlation coefficients $r(k) = \sigma^{-2} \{E(X_j X_{j+k}) - E(X_j)^2\}$ have the form $r(k) \approx c_r k^{-(1-\alpha)}$ for some $0 < \alpha < 1$ as $k \rightarrow \infty$. If the original process $\{Z(t); t \geq 0\}$ is H self similar with stationary increments, the correlation coefficients $r(k)$ of the increments process $\{X_j = Z(j) - Z(j-1); j \geq 1\}$ have the form $r(k) \approx c_r k^{-2(1-H)}$ as $k \rightarrow \infty$ (Beran, 1994). The case $1/2 < H < 1$ corresponds to LRD ($H = 0.5(1+\alpha)$). An equivalent definition of LRD involves the spectrum $f(\lambda) \approx c_f \lambda^{-\alpha}$ ($\lambda \rightarrow 0$, $c_f > 0$).

* The *first semiparametric estimator* is based on the wavelet representation (*wavelet based*). The wavelet coefficients of a discrete parameter (second order) stationary long memory process $\{X_j; j \geq 1\}$ satisfy the following two properties (Flandrin, 1992):

P1. Provided $N \geq 0.5(\alpha - 1)$ the wavelet coefficients d_{jk} with fixed scale index j form a stationary process satisfying $E(d_{jk}^2) = 2^{j\alpha} c_f C(\alpha, \psi_0)$ as $j \rightarrow \infty$ (this relationship actually holds true for j larger than some cutoff scale index j_1). $C(\alpha, \psi_0)$ stands for $\int v^{-\alpha} |\psi_0(v)|^2 dv$, $\psi_0(v)$ is the Fourier transform of the mother wavelet wavelet and N is a positive integer named the number of vanishing moments of the wavelet (an integer such that $\forall k = 0, 1, \dots, N-1, \int t^k \psi_0(t) dt = 0$).

P2. Provided $N \geq 0.5\alpha$, $E(d_{jk} d_{jk'}) \approx |k - k'|^{\alpha-1-2N}$ tends to 0 as $N, |k - k'|$ tend to infinity.

In P1, P2 α is the long memory parameter of the power law spectrum, j_2 is the largest j allowed by the given data ($\log_2(n)$, where n is the number of available data), and j_1 has to be identified from the data. From P1, Veitch and Abry (1999) propose to estimate (α, c_f) by a linear regression of $\log_2 E(d_{jk}^2)$ versus $\log_2 2^j = j$, where $\log_2 E(d_{jk}^2)$ is

estimated by $\log_2(1/n_j \sum_{k=1}^{n_j} d_{jk}^2) - g_j$, $n_j = n/2^j$ is the number of available coefficients at scale j and the g_j account for $\log_2 E(\cdot) \neq E[\log_2(\cdot)]$. Closed form solutions for the variances of $\log_2(1/n_j \sum_{k=1}^{n_j} d_{jk}^2)$ have been used in the weighted linear regression:

$$\log_2(1/n_j \sum_{k=1}^{n_j} d_{jk}^2) - g_j = j\alpha + \log_2 c_f C(\alpha, \psi_0) \quad (2)$$

Under the supplementary conditions that the process $\{X_j; j \geq 1\}$ is Gaussian, the random variables $\{d_{jk}\}$ with fixed j are independent identically distributed and that the processes $\{d_{jk}\}$ and $\{d_{j'k}\}, j \neq j'$, are independent, the estimator of $(\alpha, \log_2(c_f C(\alpha, \psi_0)))$ is unbiased, asymptotically Gaussian and efficient (Veitch and Abry, 1999).

* The *second semiparametric estimator* considered has been introduced by Robinson (Robinson (1995)) and is known as “*local Whittle*”. It is based on the following approximation of (-) the log-likelihood function for Gaussian processes:

$$Q(\vartheta) = \frac{1}{2\pi} \left[\int_{-\pi}^{\pi} \log f(\lambda, \vartheta) d\lambda + \int_{-\pi}^{\pi} \frac{I(\lambda)}{f(\lambda, \vartheta)} d\lambda \right] \quad (3)$$

where $I(\lambda)$ is the periodogram, $f(\lambda, \theta)$ the spectral density function at frequency λ , ϑ the vector of unknown parameters. The Whittle parametric estimator is the value of ϑ that minimizes (3). The “local” Whittle H estimator is the value that minimizes the discretized version of (3) on the Fourier frequencies $2\pi j/n; j=1, \dots, \lceil (n-1)/2 \rceil$, in which $f(\lambda, \theta)$ is $f(\lambda) \approx c_f \lambda^{-\alpha}$ and only the first l frequencies are included (l satisfies $(1/l + l/n) \rightarrow 0$ as $n \rightarrow \infty$). Robinson is able to show as $n \rightarrow \infty$ convergence in probability of $\hat{H}$ to H and moreover that $\sqrt{l}(\hat{H} - H_0) \xrightarrow{d} N(0, 1/4)$. No Gaussianity is needed.

4. A comparative evaluation through wavelet-based simulation

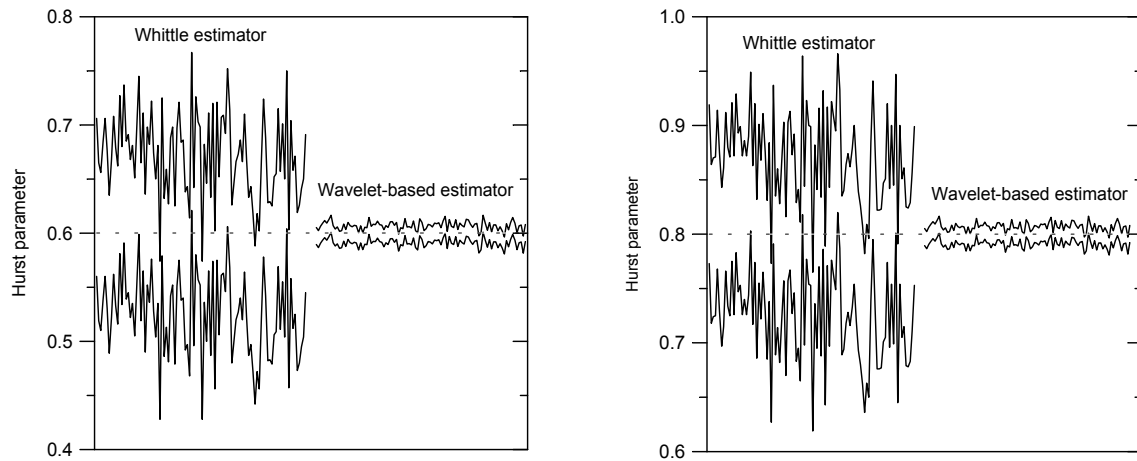
The coverage probability and size of the confidence intervals of the two estimators have been computed on a set of 100 sequences $\{x_i\}$ generated by wavelet-based simulation, for $H=0.6, 0.8, j=15$ (the sample size is $2^{15}=32768$). Each $\{x_i\}$ represents the number of fixed length packets arrived in a 0.1 second time interval. The mean value and the variance of the series to be generated have been obtained by a real trace of arrivals of packets on a 34 Mbit/s link with estimated offered traffic $\hat{\rho}$ (the product of the mean sample packet arrival rate and the mean sample packet service time) equal to 0.78. A process with a power-law behaviour over almost the full range of frequencies has been generated. To obtain each sequence a DWT $\{d_{jk}\}$, for which we assume a truncated Gaussian distribution to get non negative $\{x_i\}$ values, has been generated by the polar method and inverted. For the wavelet-based estimator a goodness of fit test allows to use all scales. For the “local” Whittle estimator we have chosen $l=180$. The results are

reported in Table 1. The confidence interval bounds are shown in Figure 1 versus the trace number. Though the coverage probability of the wavelet-based estimator is only marginally better than the local Whittle one, its confidence intervals are much narrower. The same results hold true increasing l up to the maximum allowed by the convergence condition $(1/l+1/n) \rightarrow 0$ as $n \rightarrow \infty$. The cut-off of scales penalizes the wavelet based estimator less than the convergence condition penalizes the “local Whittle” estimator.

Table 1: Coverage probabilities and size of 100 simulated confidence intervals

	Wavelet-based estimator		Local Whittle estimator	
	H=0.6	H=0.8	H=0.6	H=0.8
95% Coverage probability	0.94	0.98	0.94	0.98
95% Confidence interval size	0.014		0.146	
99% Coverage probability	0.94	0.97	0.90	0.98
99% Confidence interval size	0.020		0.192	

Figure 1: 95% Confidence interval bounds for 100 simulated long memory sequences



References

- Beran J. (1994) *Statistics for Long-Memory Processes*, Chapman and Hall.
- Cambanis S. (1995) On the continuous wavelet transform of second order random processes, *IEEE Trans. Info. Theory*, 41, 3, 628-642.
- Daubechies I. (1992) *Ten Lectures on Wavelets*, SIAM, Philadelphia.
- Flandrin P. (1992) Wavelet analysis and synthesis of fractional Brownian motion, *IEEE Trans. Info. Theory*, 38, 2, 910-917.
- Robinson P.M. (1995) Gaussian Semiparametric estimation of long range dependence, *The Annals of Statistics*, 23, 5, 1630-1661.
- Veitch D., Abry P. (1999) A Wavelet Based Joint Estimator of the Parameters of Long Range Dependence Traffic. *IEEE Trans. Info. Theory*, 45, 3, 878-897.