Eur. Phys. J. B **27**, 197–200 (2002)
DOI: 10.1140/epjb/e20020150

THE EUROPEAN PHYSICAL JOURNAL B
EDP Sciences

Second-order moving average and scaling of stochastic time series

E. Alessio[1], A. Carbone[2,a], G. Castelli[2], and V. Frappietro[1]

[1] Metronome-Ricerca sui Mercati Finanziari, via Bogino 23, 10123 Torino, Italy
[2] Istituto Nazionale per la Fisica della Materia (INFM) and Dipartimento di Fisica, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy

Received 31 December 2001

Abstract. Long-range correlation properties of stochastic time series $y(i)$ have been investigated by introducing the function $\sigma_{MA}^2 = \frac{1}{N_{max}-n}\sum_{i=n}^{N_{max}}[y(i)-\widetilde{y}_n(i)]^2$, where $\widetilde{y}_n(i)$ is the moving average of $y(i)$, defined as $1/n\sum_{k=0}^{n-1} y(i-k)$, n the moving average window and N_{max} is the dimension of the stochastic series. It is shown that, using an appropriate computational procedure, the function σ_{MA} varies as n^H where H is the Hurst exponent of the series. A comparison of the power-law exponents obtained using respectively the function σ_{MA} and the Detrended Fluctuation Analysis has been also carried out. Interesting features denoting the existence of a relationship between the scaling properties of the noisy process and the moving average filtering technique have been evidenced.

PACS. 05.40.-a Fluctuation phenomena, random processes, noise, and Brownian motion – 05.45.Tp Time series analysis

1 Introduction

Long-memory stochastic processes, owing to their intriguing statistical properties, continue to attract the interest of the physicist community. These processes, which are encountered in fields as different as condensed matter, biophysics, social science, climate change, finance, are usually characterized either by the fractal dimension D, or by the Hurst H or by the scaling exponent α of their power law statistics.

Since the scaling exponents are related to the universality class of the system, their knowledge allows to understand the fundamental processes ruling the system dynamics. Conversely, these parameters appear to be very helpful also for practical purposes. For example, it has been observed that the scaling exponent can distinguish between healthy and sick heart beat rate. Hurst exponents $H \gg 1/2$ have also been found in financial time series, where the knowledge of H can help to identify markets with higher degree of persistence [1–8].

A number of frequency, time and, even, *integrated* domain approaches have been proposed to estimate these exponents from random data sequence [8–13]. The procedures to measure the scaling exponents of a stochastic sequence $y(i)$ consist in calculating appropriate statistical functions from the whole signal. These functions show a power-law dependence on the scale size. Restricting our discussion to time domain, Detrended Fluctuation Analysis (DFA) and Rescaled Range Analysis (R/S) are the most popular scaling methods to estimate power-law correlation exponents from random signals. The R/S technique consists in the following steps. The stochastic time series $y(i)$ with ($i = 1, 2, ...N_{max}$) is divided into boxes of equal size n. The functions:

$$X_i = \sum_{j=kn+1}^{i}[y(j) - \langle y\rangle] \tag{1}$$

and

$$S = \sqrt{\frac{1}{n}\sum_{j=1}^{n}[y(j)-\langle y\rangle]^2} \tag{2}$$

are calculated in the kth box. In equations (1, 2), $\langle y\rangle$ represents the average value of the time series $y(i)$ over each box and is given by $1/n\sum_{i=kn+1}^{(k+1)n} y(i)$. The Rescaled Range function is defined by:

$$R/S = 1/S\left[\max_{kn+1\ll i\ll(k+1)n} X_i - \min_{kn+1\ll i\ll(k+1)n} X_i\right]. \tag{3}$$

The function R/S is then averaged over all the boxes of equal size n. By iterating the calculation of $\langle R/S\rangle$ for different box amplitudes n, a relationship between $\langle R/S\rangle$ and n is obtained, that in the presence of scaling is of power-law type. According to the DFA technique, after dividing the series in equal size boxes as done for the R/S

[a] e-mail: acarbone@polito.it

technique, a polynomial function $y_{pol}(i)$ interpolating the sequence in each box is calculated. The interpolating curve $y_{pol}(i)$ represents the local trend in each box. The average value of the function:

$$DFA = \sqrt{\frac{1}{N_{max}} \sum_{i=1}^{N_{max}} [y(i) - y_{pol}(i)]^2}, \qquad (4)$$

is calculated over all the boxes of equal size. Repeating the calculation over boxes of different size, a relationship as $DFA \propto n^H$ is obtained for long-memory correlated processes. In particular, $0 < H < 0.5$ and $0.5 < H < 1$ correspond respectively to negative (antipersistence) and positive correlation (persistence), while $H = 0.5$ corresponds to fully uncorrelated signals.

In a recent paper, Vandewalle and Ausloos reported on interesting features of the moving average $\widetilde{y}_n(i)$ of a time series $y(i)$ [14]. Moving average is a well-known low-pass filter defined, for discrete signals, by:

$$\widetilde{y}_n(i) = \frac{1}{n} \sum_{k=0}^{n-1} y(i-k). \qquad (5)$$

They observed that the density ρ of crossing points between moving averages with time windows respectively equal to n_1 and n_2 can be indeed expressed as:

$$\rho = \frac{1}{n_2} [(\Delta n)(1 - \Delta n)]^{H-1} \qquad (6)$$

with $\Delta n = (n_2 - n_1)/n_2$ and $(n_2 \gg n_1)$. The authors used the equation (6) to extract the Hurst exponent of correlated time series obtaining results comparable with the DFA ones. On this account, they concluded that, even though not yet rigorously established from a theoretical point of view, moving averages show unexpected physical meaning deserving further investigation.

The results of paper [14] motivated our work. In the limit of $n \to 0$, the moving average tends to the series itself (*i.e.* $\widetilde{y}_n(i) \to y(i)$). The crossing points correspond to the zeroes of the *first-order* difference between $y(i)$ and $\widetilde{y}_n(i)$. In the following section we will report on a systematic analysis of the properties of the *second-order* difference of the process $y(i)$ with respect to $\widetilde{y}_n(i)$. The results will shed more light on the basic idea of Vandewalle and Ausloos derived from equation (6).

2 Moving average technique and stochastic series scaling

In Figure 1, a random sequence $y(i)$ with $H = 0.3$ is plotted (solid lines). Dotted and dashed lines correspond respectively to the moving averages $\widetilde{y}_n(i)$ calculated for $n = 10$ and $n = 30$. In Figure 2, a random sequence $y(i)$ with $H = 0.7$ is plotted (solid lines). Dotted and dashed lines correspond respectively to the moving averages $\widetilde{y}_n(i)$ calculated for $n = 500$ and $n = 5000$. The series have been generated using the Random Midpoint Displacement

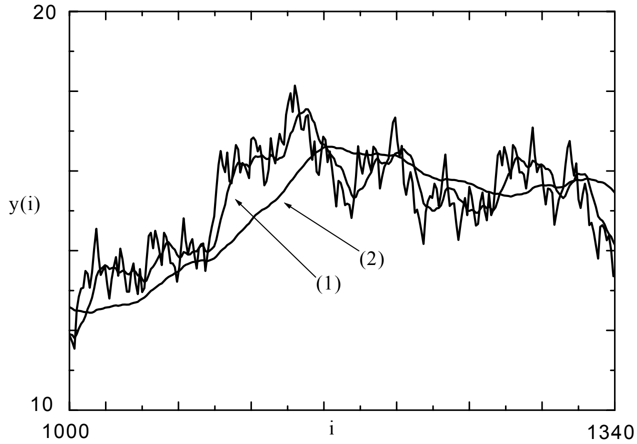

Fig. 1. Stochastic series $y_n(i)$ obtained by using the Random Midpoint Displacement algorithm with $H = 0.3$. The size of the series is $N_{max} = 2^{19}$. Curves (1) and (2) are the moving averages $\widetilde{y}_n(i)$, calculated using equation (5) respectively with $n = 10$ and $n = 30$.

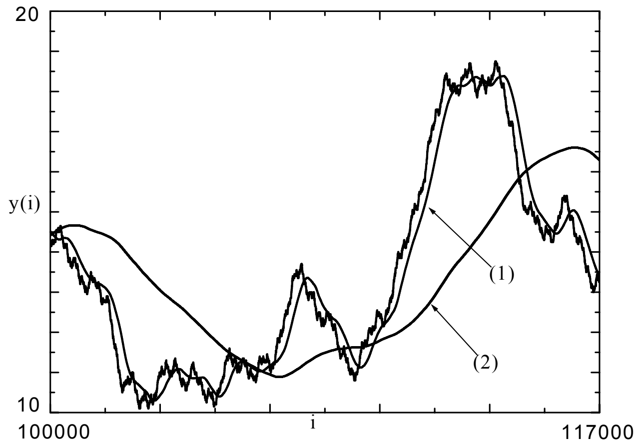

Fig. 2. Stochastic series $y_n(i)$ obtained by using the Random Midpoint Displacement algorithm with $H = 0.7$. The size of the series is $N_{max} = 2^{19}$. Curves (1) and (2) are the moving averages $\widetilde{y}_n(i)$, calculated using equation (5) respectively with $n = 500$ and $n = 5000$.

(RMD) algorithm. Both the series of Figures 1 and 2 have size $N_{max} = 2^{19}$.

As stated in the introduction, in this paper we will report on a study of the second-order difference of the noisy signal $y(i)$ with respect to the moving average $\widetilde{y}_n(i)$ defined by:

$$\sigma_{MA}^2 = \frac{1}{N_{max} - n} \sum_{i=n}^{N_{max}} [y(i) - \widetilde{y}_n(i)]^2 \qquad (7)$$

i.e. the variance of $y(i)$ with respect to the moving average $\widetilde{y}_n(i)$. By using equation (7), we have performed the following computational procedure:

1. a stochastic series $y(i)$ with an assigned Hurst exponent has been generated (as already specified, in this work we have used the RMD algorithm),

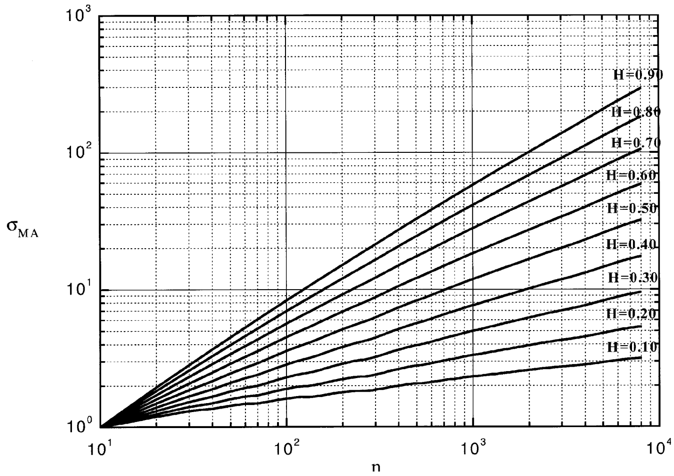

Fig. 3. Power-law dependence of the function σ_{MA}, defined by equation (8), on the moving average box n. Curves have been obtained using the procedure proposed in Section 2, with artificially generated series having $N_{max} = 2^{18}$ and Hurst coefficient H varying between 0.05 and 0.95.

2. the moving averages $\widetilde{y}_n(i)$ with different values of n have been calculated for the series $y(i)$. The values of n range from 2 to 15000. In the following we will refer to the maximum value of n used as n_{max},
3. the function σ_{MA}, defined by the equation (7), is then calculated over the time interval $[n_{max}, N_{max}]$, *i.e.*:

$$\sigma_{MA} = \sqrt{\frac{1}{N_{max} - n_{max}} \sum_{i=n_{max}}^{N_{max}} [y(i) - \widetilde{y}_n(i)]^2} \quad (8)$$

for each moving average $\widetilde{y}_n(i)$,
4. the values of σ_{MA} corresponding to each $\widetilde{y}_n(i)$ are plotted as a function of n on log-log axes.

The described algorithm has been applied to a number of artificially generated random series with different sizes.

Results shown in Figure 3 refer to stochastic series with $N_{max} = 2^{18}$ and Hurst coefficients varying from 0.05 to 0.95 with step 0.05. For these curves, it is $n_{max} = 8000$ as it can be seen in Figure 3. Analogous behaviour has been observed with different values of n_{max}.

The most remarkable property of the curves plotted in Figure 4 is the power-law dependence on n, *i.e.*:

$$\sigma_{MA} \propto n^H. \quad (9)$$

The function σ_{MA} allows to estimate the scaling exponents analogously to the DFA and the to the R/S function.

In Figure 4, power-law exponents obtained using the σ_{MA} algorithm have been compared with the DFA ones. Data refer to artificial series with $N_{max} = 2^{17}$ and $n_{max} = 10000$. Dashed line represents the ideal behaviour, *i.e.* $H_{out} = H_{in}$. Both the σ_{MA} and the DFA curves deviate from the ideal behaviour and a crossover with the dashed line occurs. The onset of crossovers is a main concern of DFA and R/S scaling techniques, reducing the range of applicability of the algorithms and the accuracy of the estimation. It has been therefore extensively investigated by several authors [15–17]. By observing the results of Figure 4, the crossover between the σ_{MA} curve and the dashed line occurs at a value of $H \simeq 0.49$, while the crossover of the DFA curve occurs at $H \simeq 0.67$. This could be due to a better smoothing and detrending of the moving average filter respect to the polynomial fitting used by the DFA. We have ascribed the better performance of our algorithm to the following reasons. When using the DFA technique, linear or cubic polynomials fitting the stochastic series in the corresponding boxes, are used. Moving average filter adjusts the fitting curve dynamically (*i.e.* every time the discrete index i increases by a unity, the box window n adjust its position accordingly). Due to the continuous adjusting of the box position to the series, a higher accuracy should be expected.

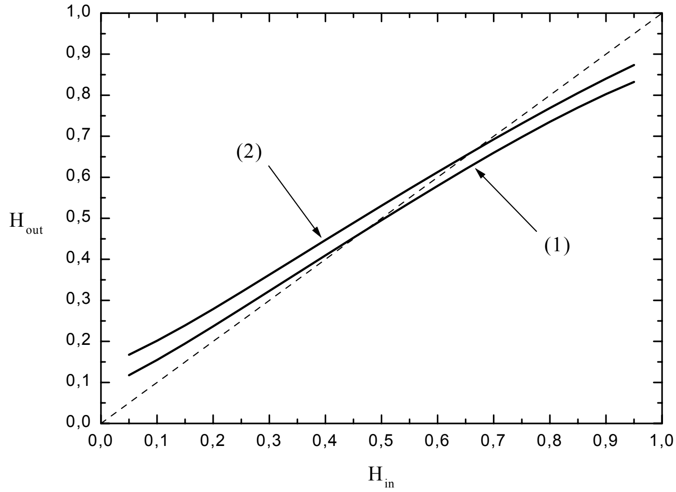

Fig. 4. Comparison between the estimated power-law exponents from artificially generated series using the technique σ_{MA} (curve (1)), proposed in Section 2, and the Detrended Fluctuation Analysis (DFA) (curve (2)). Dashed line represents the ideal behaviour ($H_{out} = H_{in}$). The occurrence of a crossover at $H = 0.49$ and at $H = 0.67$ respectively for the σ_{MA} and for the DFA curve can be observed.

3 Discussion and conclusion

We have reported on the scaling properties of long-range correlated stochastic series $y(i)$ as obtained by the computational procedure described in the previous section. This procedure makes use of the function σ_{MA} defined by the equation (7). We have found the remarkable result that the function σ_{MA} varies as a power-law of the amplitude n of the moving average window. The results obtained using the σ_{MA} algorithm are strictly related to the property of the density of crossing points between $y(i)$ and $\widetilde{y}_n(i)$ reported by [14]. However, the relationship $\sigma_{MA} \propto n^H$, satisfied by the function σ_{MA}, better evidences the link between $\widetilde{y}_n(i)$ and the scaling properties of $y(i)$.

To the best of our knowledge, this behaviour has never before been evidenced in the literature. In order to attain a full understanding of these results, a detailed analysis of the statistical properties of σ_{MA} and other related quantities is required.

The results of the present work confirm that a deeper theoretical insight of moving average filtering could reveal other interesting properties for both fundamental and application purposes.

References

1. R.N. Mantegna, H.E. Stanley, *An Introduction to Econophysics: Correlations and Complexity in Finance* (Cambridge University Press, 2000).
2. J. Feder, *Fractals* (Plenum, New York, 1988).
3. P. Bak, C. Tang, K. Wiesenfeld, Phys. Rev. Lett. **59**, 381 (1987).
4. J.P. Sethna, K.A. Dahmen, C.R. Myers, Nature **410**, 242 (2001).
5. Y. Liu, P. Gopikrishnan, P. Cizeau, M. Mayer, C.-K. Peng, H.E. Stanley, Phys. Rev. E **60**, 1390 (1999).
6. P. Gopikrishnan, V. Plerou, L.A.N. Amaral, M. Mayer, H.E. Stanley, Phys. Rev. E **60**, 5305 (1999).
7. C.-K. Peng, J.M. Hausdorff, J.E. Mietus, S. Havlin, H.E. Stanley, A.L. Goldberger in: *Levy Flights and Related Topics in Physics (Proc. 1994 International Conference on Levy Flights)*, edited by M.F. Shlesinger, G.M. Zaslavsky, U. Frish (Springer, Berlin, 1995), pp. 315–330.
8. C.-K. Peng, S.V. Buldyrev, A.L. Goldberger, S. Havlin, M. Simons, H.E. Stanley, Phys. Rev. E **47**, 3730 (1993).
9. S.V. Buldyrev, A.L. Goldberger, S. Havlin, R.N. Mantegna, M.E. Matsa, C.-K. Peng, M. Simons, H.E. Stanley, Phys. Rev. E **51**, 5084 (1993).
10. C.-K. Peng, S.V. Buldyrev, S. Havlin, M. Simons, H.E. Stanley, A.L. Goldberger, Phys. Rev. E **49**, 1685 (1994).
11. M.S. Taqqu, V. Teverovsky, W. Willinger, Fractals **3**, 785 (1995).
12. K. Ivanova, M. Ausloos, Physica A **265**, 279 (1999).
13. H.E. Stanley, V. Afanasyev, L.A.N. Amaral, S.V. Buldyrev, A.L. Goldberger, S. Havlin, H. Leschhorn, P. Maass, R.N. Mantegna, C.-K. Peng, P.A. Prince, M.A. Salinger, M.H.R. Stanley, G.M. Viswanathan, Physica A **224**, 302 (1996).
14. N. Vandewalle, M. Ausloos, Phys. Rev. E **58**, 6832 (1998).
15. Kun Hu, Plamen Ch. Ivanov, Zhi Chen, P. Carpena, H. Eugene Stanley, Phys. Rev. E **64**, 011114 (2001).
16. G. Rangarajan, M. Ding, Phys. Rev. E **61**, 004991 (2000).
17. C. Heneghan, G. McDarby, Phys. Rev. E **62**, 6103 (2000).