Scaling properties of long-range correlated noisy signals: application to financial markets

Scaling properties of long-range correlated noisy signals: application to financial markets

Anna Carbone and Giuliano Castelli

Physics Department and National Institute of Matter Physics (INFM),
Politecnico di Torino, Corso Duca degli Abruzzi 24, I-10129 Torino, Italy

ABSTRACT

Long-range correlation properties of financial stochastic time series $y(i)$ have been investigated with the main aim to demonstrate the ability of a recently proposed method to extract the scaling parameters of a stochastic series. According to this technique, the Hurst coefficient H is calculated by means of the following function:

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

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. The method is called *Detrending Moving Average Analysis* (DMA) on account of the several analogies with the well-known *Detrended Fluctuation Analysis* (DFA). The DMA technique has been widely tested on stochastic series with assigned H generated by suitable algorithms. It has been demonstrated that the ability of the proposed technique relies on very general grounds: the function $C_n(i) = y(i) - \widetilde{y}_n(i)$ generates indeed a sequence of clusters with power-law distribution of amplitudes and lifetimes. In particular the exponent of the distribution of cluster lifetime varies as the fractal dimension $2 - H$ of the series, as expected on the basis of the box-counting method. In the present paper we will report on the scaling coefficients of real data series (the BOBL and DAX German future) calculated by the DMA technique.

Keywords: Time series analysis, Systems obeying scaling laws, Complex systems

1. INTRODUCTION

Long-memory stochastic processes are ubiquitous in fields as different as condensed matter, biophysics, social science, climate change, finance[1–23]. Their statistical properties and, in particular, scaling exponents other than continuing to draw the attention of the physicist community, have recently demonstrated to be a powerful tool for practical purposes. The scaling analysis of medical series (heart-rate dynamics, lung inflation) supplies in-depth information on the disease. In finance, the series of the volatility are characterized by a degree of persistence higher than the price returns. For these reasons, it is crucial to develop even more accurate and fast algorithms able to extract the fractal dimension D, the Hurst exponent H or the scaling exponent α of a random sequence. The methods of extraction of the scaling exponents from a random series $y(i)$ usually exploit suitable statistical functions of $y(i)$. *Detrended Fluctuation Analysis (DFA)* and *Rescaled Range Analysis (R/S)* are the most popular scaling techniques to estimate the power-law correlation exponents from the random signals in the time domain.

In a recent work[22], a method for the analysis of the persistence, with particular features of accuracy and speed, has been proposed. Up to now such technique has been applied only to artificially generated series with the main aim to investigate the general properties[23] and the computational performances of the algorithm.[22]

In the present work, the DMA analysis will be applied to financial random series. As above mentioned, in addition to the remarkable interest of the physics community toward the fundamental implications of the topics, the demonstration of a scaling technique, gaining in execution speed, is particularly interesting in view of online trading application and in any other field, where the performance velocity is a discriminating issue.

E-mail: anna.carbone@polito.it

Noise in Complex Systems and Stochastic Dynamics, Lutz Schimansky-Geier, Derek Abbott, Alexander Neiman, Christian Van den Broeck, Editors, Proceedings of SPIE Vol. 5114 (2003)

2. DFA AND DMA SCALING TECHNIQUES

For the sake of clarity, we will briefly review the Detrended Fluctuation Analysis (DFA) and Detrending Moving Average Analysis (DMA) algorithms.

According to the DFA technique, after dividing the series in equal size boxes, an interpolating linear or cubic function $y_{pol}(i)$, representing the local trend of the random series, is calculated in each box. The function:

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

is then calculated over all the boxes of equal size n. Repeating the calculation over different size boxes, a relationship as:

$$DFA \propto n^H \tag{2}$$

is obtained. For long-memory correlated processes, it is $0 < H < 0.5$ for negative persistence, and $0.5 < H < 1$ for positive persistence. $H = 0.5$ characterizes fully uncorrelated signals.

We have reported on a novel technique based on the following function[22]:

$$DMA = \sqrt{\frac{1}{N_{max} - n_{max}} \sum_{i=n_{max}}^{N_{max}} [y(i) - \widetilde{y}_n(i)]^2} \tag{3}$$

The Eq. (3) defines a generalized variance of the random series $y(i)$ with respect to the moving average $\widetilde{y}_n(i)$. The moving average $\widetilde{y}_n(i)$ is calculated for different values of the boxes n. n_{max} is the maximum value of n. Then the DMA function is calculated over all the boxes of equal size n.

It can be observed that the function DMA presents the same structure of the DFA. In the Eq. (1) and (3), the functions $y_{pol}(i)$ and $\widetilde{y}_n(i)$ act as low-pass filters of the random signal $y(i)$, and thus, $y_{pol}(i)$ and $\widetilde{y}_n(i)$ represent the trend of $y(i)$.

As already said, by means of the Eq.(3), the following computational procedure can be implemented. The moving averages $\widetilde{y}_n(i)$ with different values of n are calculated for the series $y(i)$. The function DMA, defined by the Eq.(3), is then calculated over the time interval $[n_{max}, N_{max}]$. For each moving average $\widetilde{y}_n(i)$, the values of DMA corresponding to each $\widetilde{y}_n(i)$ are plotted as a function of n on log-log axes. The most remarkable property of the plot so obtained is that the function DMA exhibits a power-law dependence with exponent H on n, i.e.:

$$DMA \propto n^H. \tag{4}$$

On account of this relationship, the function DMA allows to estimate the scaling exponent H as done by the DFA technique. Due to the several analogies with the DFA, this technique has been named *Detrending Moving Average Analysis* (DMA).

The described algorithm has been tested over several artificially generated random series with different H.[22] Very accurate results in a wide range of values of the series size N_{max} and of the scaling window size n have been obtained.

In a recent work,[23] it has been demonstrated that the function $C_n(i) = y(i) - \widetilde{y}_n(i)$ generates a sequence of clusters with power-law distribution of the amplitudes and lifetimes. In particular, the exponent of the distribution of the cluster lifetimes is equal to the fractal dimension $2 - H$ of the series, as expected on the basis of the box-counting technique.[2]

The results obtained using the DMA algorithm are strictly related to the property of the density of crossing points between $y(i)$ and $\widetilde{y}_n(i)$ reported by Vandewalle and Ausloos.[10] However, the relationship $DMA \propto n^H$, satisfied by the function DMA, better evidences the scaling properties of $y(i)$ and the relationship with the other techniques.

Finally, another characteristics of the DMA function is the high speed of execution, worthy of note in view of the application. The DMA algorithm is indeed at least 50 times faster than the DFA algorithm. This fact follows in a straightforward manner if the speed of execution of the moving average $\widetilde{y}_n(i)$ with respect to the linear or polynomial $y_{pol}(i)$ algorithm is taken in mind.

Results concerning the application of the DMA technique to real data sequence have been not yet published, therefore, in the present work, we will report on a study of the Hurst coefficients of financial series based on the DMA algorithm. As said in the Introduction, complex signals are found in several topics, as well as climate, geology, biology, however financial analysis is a field where the algorithm performances are a stringent requirement.

3. APPLICATIONS TO FINANCIAL SERIES

The DMA algorithm has been applied to the German Bobl and Dax future data (sampled every minute). The Bobl future is a derivative of a ten years maturity, 5% coupon German Government security. The Dax Future is the derivative of the main German stock index and thereby represents a measure of expectations of both stock market growth and in general of economic growth in German and generally in the European Union area.

Let $p(t)$ be the price at time t, the *log-return* $G_{\Delta t}(t)$ is defined as[5,6,14–16]:

$$G_{\Delta t}(t) = log\ p(t+\Delta t) - log\ p(t) \tag{5}$$

where Δt is the time sampling interval. The volatility is taken as the average of $|G(t)|$ over a time window $T = n\Delta t$:

$$v_T(t) = \frac{1}{n}\sum_{t'=t}^{t+n-1} |G_{\Delta t}(t')|^2 \tag{6}$$

where n is an integer.

The volatility $v_T(t)$ of the financial return is not unambiguously defined, it is however out of the scope of the present work an in-depth discussion on this topics.[16]

In Figs. (1) and (2), the plot of the prices and of the volatility, according to the definition given by the Eq.(6), of the Bobl futures of the German market are shown. The size of the series is $N_{max} = 559939$. The volatility window T is equal to 10000 and the sampling interval Δt is equal to 1. In figure (3) and (4) , the log-log plot of the curves defined by the Eqs.(1-4) are shown. H is obtained by the slope of the straight lines. The H values, calculated by the DMA and DFA algorithms, for the series of prices are respectively equal to $H_{DMA} = 0.48$ and $H_{DFA} = 0.49$. The H values, calculated by the DMA and DFA algorithms, for the series of volatilities are equal to $H_{DMA} = 0.71$ and $H_{DFA} = 0.71$. The scaling box amplitudes ranged from $n = 100$ to $n = 10000$ with step 100 for both techniques.

In Figs. (5) and (6), the plot of the prices and of the volatility, according to the definition given by the Eq.(6), of the DAX futures of the German market are shown. The size of the series is $N_{max} = 741657$. The volatility window T is equal to 10000 and the sampling interval Δt is equal to 1. In Figs. (7) and (8) , the log-log plot of the curves defined by the Eqs.(1,4) is shown. H is obtained by the slope of the straight lines. The H values, calculated by the DMA and DFA algorithms, for the series of prices are respectively equal to $H_{DMA} = 0.47$ and $H_{DFA} = 0.49$. The H values, calculated by the DMA and DFA algorithms, for the series of volatilities are equal to $H_{DMA} = 0.81$ and $H_{DFA} = 0.86$. The scaling box amplitudes ranged from $n = 100$ to $n = 10000$ with step 100 for both techniques. The values of the Hurst exponent H for the German futures analyzed in this work by the DMA technique are in very good agreement with those obtained by the DFA and with the findings of other authors.[15] It can be noted by observing the Figs. Figs. (7) and (8) that the DFA curves are noisier than the DMA curves.

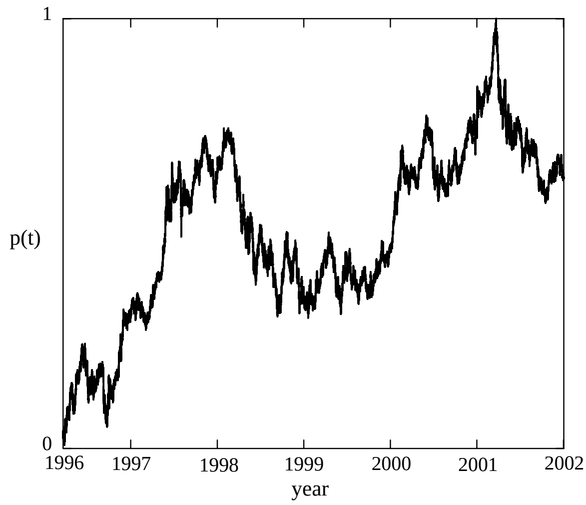

Figure 1. Stochastic series of the prices $p(t)$ of the Bobl German future.

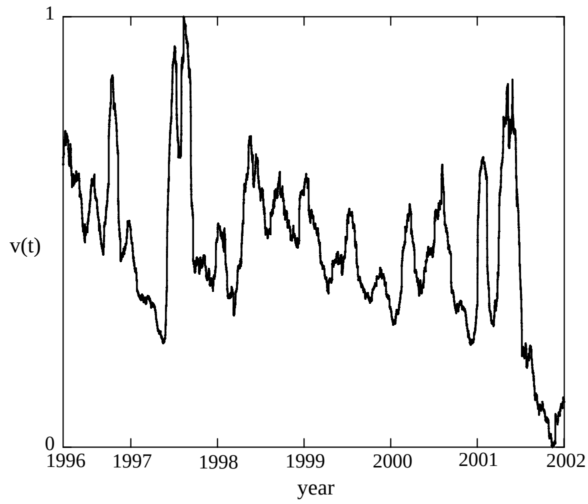

Figure 2. Stochastic series of the Bobl German future volatility $v(t)$ according to the definition (6). The volatility window T is 10000 and the sampling interval Δt is 1.

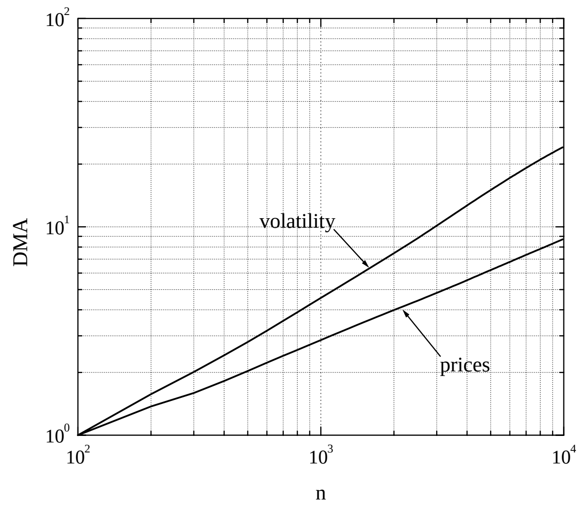

Figure 3. DMA functions for the stochastic series of figure (1) and (2) according to the equation (3).

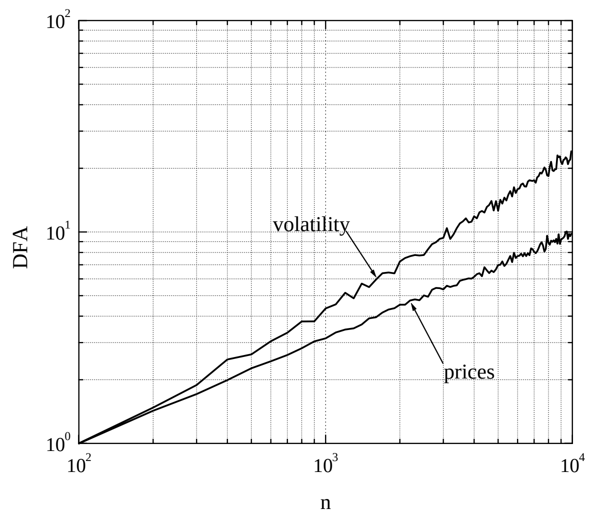

Figure 4. DFA functions for the stochastic series of figure (1) and (2) according to the equation (1).

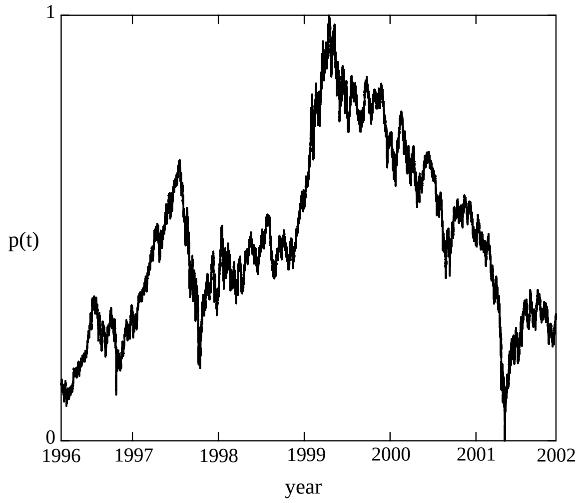

Figure 5. Stochastic series of the prices $p(t)$ of the DAX German future.

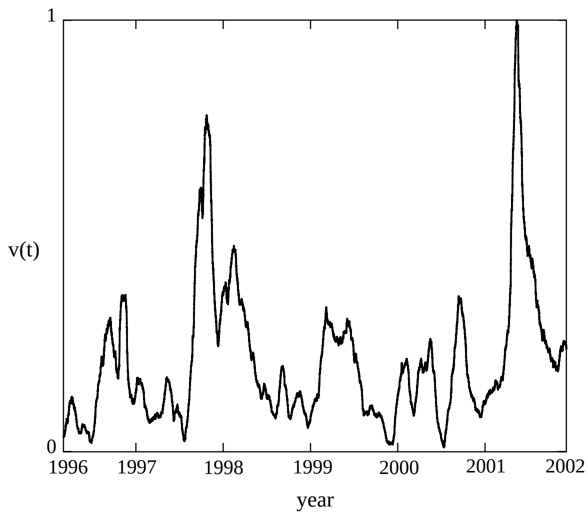

Figure 6. Stochastic series of the Dax German future volatility $v(t)$ according to the definition (6). The volatility window T is 10000 and the sampling interval Δt is 1.

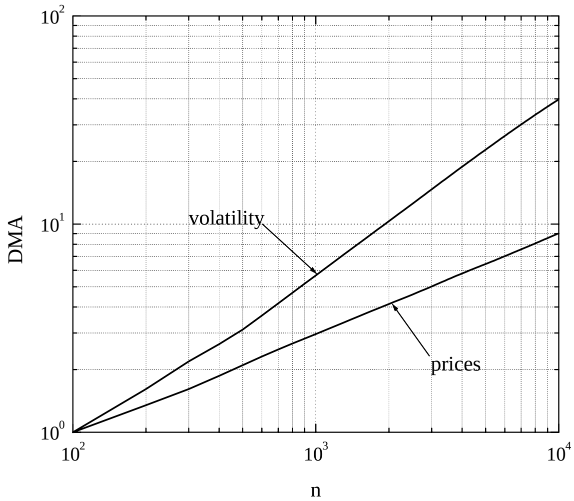

Figure 7. DMA functions for the stochastic series of figure (5) and (6) according to the equation (3).

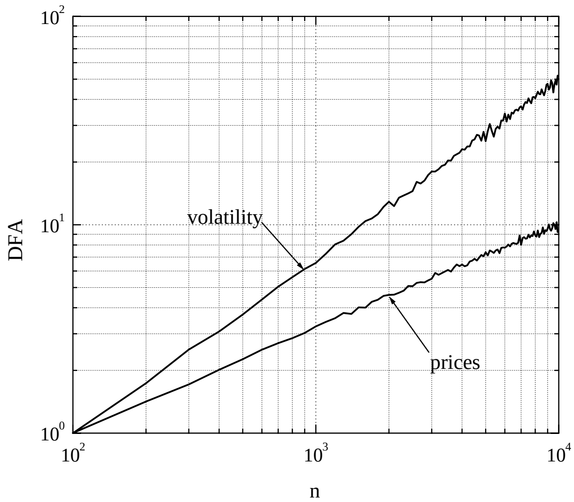

Figure 8. DFA functions for the stochastic series of figure (5) and (6) according to the equation (1).

4. CONCLUSION

We have reported on the scaling properties of long-range correlated stochastic series $y(i)$ as obtained by the computational procedure recently proposed by us.[22, 23]

This procedure makes use of the function DMA defined by the Eq.(3) that exhibits the remarkable properties to vary as a power-law, with exponent H, of the amplitude n of the moving average window.

The DMA algorithm has been applied to the German Bobl and Dax future data, sampled every minute. The Bobl future is a derivative of a ten years maturity, 5% coupon German Government security. The Dax Future is the derivative of the main German stock index and thereby represents a measure of expectations of both stock market growth and in general of economic growth in German and generally in the European Union area.

The DMA technique has revealed high accuracy and speed of execution.

REFERENCES

1. R.N. Mantegna and 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 and K. Wiesenfeld, "Self-organized criticality: An explanation of the 1/f noise", Phys. Rev. Lett. **59**, Issue 4, (1987)
4. Y. Liu, P. Gopikrishnan, P. Cizeau, M. Mayer, C.-K. Peng and H. E. Stanley, "Statistical properties of the volatility of price fluctuations ", Phys. Rev. E **60**, Issue 2, pp. 1390-1400 (1999)
5. P. Gopikrishnan., V. Plerou, L. A. N. Amaral, M. Mayer and H. E. Stanley, "Scaling of the distribution of fluctuations of financial market indices", Phys. Rev. E **60**, Issue 5, pp. 5305-5316 (1999)
6. C.-K. Peng, S.V. Buldyrev, A.L. Goldberger, S. Havlin, M. Simons, and H.E. Stanley, "Finite-size effects on long-range correlations: Implications for analyzing DNA sequences", Phys. Rev. E, **47**, Issue 5, pp. 3730-3733 (1993)
7. S.V. Buldyrev, A.L. Goldberger, S. Havlin, R.N. Mantegna, M.E. Matsa, C.-K. Peng, M. Simons, and H.E. Stanley, "Long-range correlation properties of coding and noncoding DNA sequences: GenBank analysis", Phys. Rev. E, **51**, Issue 5, pp. 5084-509 (1993)
8. C.-K. Peng, S.V. Buldyrev, S. Havlin, M. Simons, H.E. Stanley, and A.L. Goldberger, "Mosaic organization of DNA nucleotides", Phys. Rev. E, **49**, Issue 2, pp. 1685-1689 (1994)
9. 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, "Anomalous fluctuations in the dynamics of complex systems: from DNA and physiology to econophysics", Physica A, **224**, pp. 302-321 (1996)
10. N. Vandewalle and M. Ausloos, "Crossing of two mobile averages: A method for measuring the roughness exponent", Phys. Rev. E, **58**, Issue 5 pp. 6832-6834, (1998)
11. Kun Hu, Plamen Ch. Ivanov, Zhi Chen, Pedro Carpena, and H. Eugene Stanley, "Effect of trends on detrended fluctuation analysis", Phys. Rev. E **64**, 011114, (2001)
12. G. Rangarajan, M. Ding, "Integrated approach to the assessment of long range correlation in time series data", Phys. Rev. E **61**, Issue 5, 004991 (2000)
13. C. Heneghan and G. McDarby, "Coherence and stochastic resonance in a two-state system", Phys. Rev. E, **62**, Issue 5, pp. 6103-6110 (2000)
14. O. Biham, Z.-F. Huang, O. Malcai, S. Solomon, "Long-time fluctuations in a dynamical model of stock market indices", Phys. Rev. E, **64**, 026101, (2001).
15. P. Grau-Carles, "Long-range power-law correlations in stock returns", Physica A, **299**, 521, (2001).
16. Michel M. Dacorogna, Ramazan Genay, Ulrich Mller, Richard B. Olsen, Olivier V. Pictet, "An Introduction to High-Frequency Finance", Academic Press, (2001)
17. F. Lillo, J.D. Farmer, R.N. Mantegna, "Master curve for price-impact function", Nature **421**, 129 (2003);
18. J. P. Sethna, K. A. Dahmen and C. R. Myers, "Crackling noise", Nature **410**, pp. 242-250 (2001)

19. C.-K. Peng, J.M. Hausdorff, J.E. Mietus, S. Havlin, H.E. Stanley, and A.L. Goldberger, "Fractals in physiological control: from heartbeat to gait", in: Levy Flights and Related Topics in Physics (Proc. 1994 International Conference on Levy Flights), edited by M.F. Shlesinger, G.M. Zaslavsky and U. Frish (Springer, Berlin, 1995), pp. 315-330
20. M.S. Taqqu, V. Teverovsky, and W. Willinger, "Estimators for long-range dependence: an empirical study", Fractals, **3**, pp. 785 (1995)
21. K. Ivanova, M. Ausloos, "Low order variability diagrams for short range correlation evidence in financial data: BGL-USD exchange rate, Dow-Jones Industrial Average, Gold ounce price", Physica A, **265**, pp. 279-291 (1999)
22. E. Alessio, A. Carbone, G. Castelli and V.Frappietro, "Second-Order Moving Average and Scaling of Stochastic Time Series", Eur. Phys. Jour. B, **27**, 197, (2002).
23. A. Carbone and G. Castelli, "Scaling Properties of Long-Range Correlated Noisy Signals", http://arxiv.org/PS_cache/cond-mat/pdf/0303/0303465.pdf, submitted to Phys. Rev. E