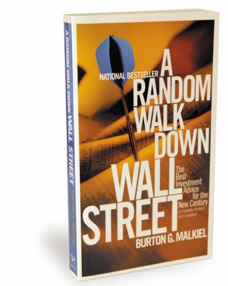
Detrending Moving Average Algorithm: from finance to genomics and disordered materials

Anna Carbone
Physics Department
Politecnico di Torino
www.polito.it/noiselab

September 27, 2009

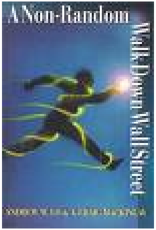
Publications

1. *Second-order moving average and scaling of long-range correlated series*
   E. Alessio, A. Carbone, G. Castelli, and V. Frappietro, Eur. Phys. Jour. B **27**, 197 (2002).
2. *Scaling of long-range correlated noisy signals: application to financial markets*
   A. Carbone and G. Castelli, Proc. of SPIE, **407**, 5114 (2003).
3. *Analysis of the clusters formed by the moving average of a long-range correlated stochastic series*
   A. Carbone, G. Castelli and H. E. Stanley, Phys. Rev. E **69**, 026105 (2004).
4. *Time-Dependent Hurst Exponent in Financial Time Series*
   A. Carbone, G. Castelli and H. E. Stanley, Physica A **344**, 267 (2004).
5. *Directed self-organized critical patterns emerging from fractional Brownian paths*
   A. Carbone and H. E. Stanley, Physica A **340**, 544 (2004).
6. *Spatio-temporal complexity in the clusters generated by fractional Brownian paths*
   A. Carbone and H.E. Stanley, Proc. of SPIE 5471, 1 (2004).
7. *Quantifying signals with power-law correlations*
   L. Xu, P. Ch. Ivanov, C. Zhi, K. Hu, A. Carbone, H. E. Stanley, Phys. Rev. E **71**, 051101 (2005).
8. *Scaling properties and entropy of long range correlated series*
   A. Carbone and H.E. Stanley, Physica A **384**, 21 (2007).
9. *Tails and Ties*
   A. Carbone, G. Kaniadakis, and A.M. Scarfone, Eur. Phys. J. B **57**, 121-125 (2007).
10. *Where do we stand on econophysics ?*
    A. Carbone, G. Kaniadakis and A.M. Scarfone, Physica A **382**, xi-xiv (2007).
11. *Detrending Moving Average (DMA) Algorithm: a closed form approximation of the scaling law*
    S. Arianos and A. Carbone, Physica A **382**, 9 (2007).
12. *Algorithm to estimate the Hurst exponent of high-dimensional fractals*
    A. Carbone, Phys. Rev. E **76**, 056703 (2007).
13. *Cross-correlation of long-range correlated series*
    S. Arianos and A. Carbone, J. Stat. Mech: Theory and Experiment **P03037**, (2009).



A Random Walk Down Wall Street

First edition published in the 70's by Burton Malkiel on the wave of the academic success of the influential survey on the Efficient Market Hypothesis (EMH) by Eugene Fama (1970).



A Non-Random Walk Down Wall Street

First edition published in 1999 by Andrew Lo and Craig MacKinley

A Moving Average Walk ... Down Wall Street

Surveys on the use of Technical Analysis by Finance Professionals [1] [2] [3]

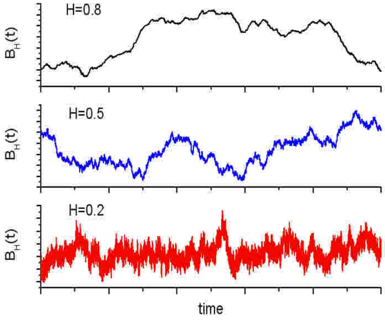
Horizons	min %	max%
intraday	50%	100%
1 week	50%	80%
1 month	40%	70%
3 months	30%	50%
6 months	20%	30%
1 year	2%	25%

[1] Charts, Noise and Fundamentals in the London Foreign Exchange Market. H.L. Allen and M.P. Taylor, The Economic Journal, **100**, 49-59 (1990)

[2] Technical trading rule profitability and foreign exchange intervention, B. LeBaron, Journal of International Economics **49**, 125 (1999)

[3] The Obstinate Passion of foreign Exchange Professional: Technical Analysis. Lukas Menkhoff and Mark P. Taylor, Journal of Economic Literature **45**, 4 (2007)

Fractional Brownian or Levy Walk

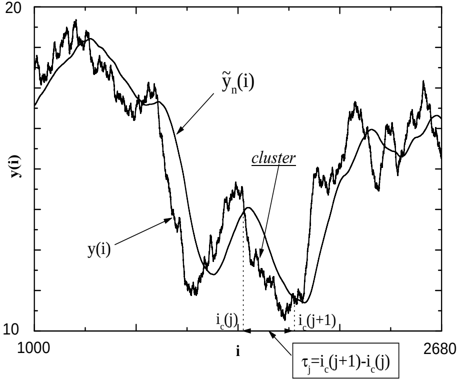


$H > 0.5$ positive correlation (persistence)

$H = 0.5$ Random walk

$H < 0.5$ negative correlation (anti-persistence)

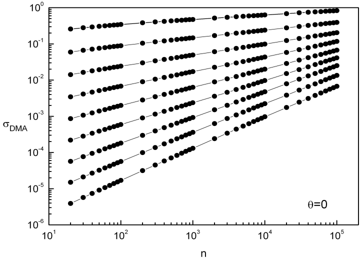
Detrending Moving Average Algorithm (DMA)



$$\sigma^2_{DMA} = \frac{1}{N}\sum_{i=n}^{N}\left[y(i) - \widetilde{y}_n(i)\right]^2$$

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

$$\sigma^2_{DMA} \sim n^{2H}$$



$$\sigma^2_{DMA} \sim n^{2H}$$

The log-log plot is a straight-line whose slope can be used to calculate the Hurst exponent H of the time series.

Self-similarity of a time series x(t)

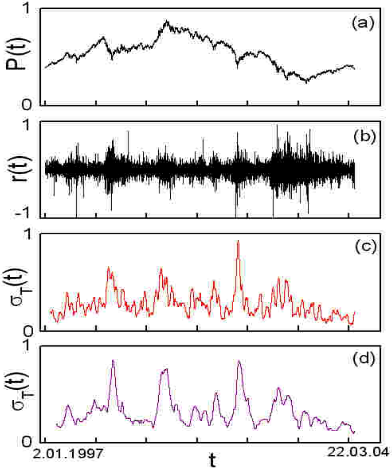
σ^2_{DMA} is **generalized variance** for nonstationary signals. It can be derived from the auto-correlation function $C_{xx}(t,\tau)$, which measures the self-similarity of a signal:

$$\sigma_{xx}(t,\tau) \equiv \left\langle [x(t) - \widetilde{x}(t)][x^*(t+\tau) - \widetilde{x}^*(t+\tau)] \right\rangle$$

By taking $\tau = 0$ and $x^* = x(t)$, $\widetilde{x}_n(i) = \frac{1}{n}\sum_{k=0}^{n} x(i-k)$ the autocovariance $C_{xx}(t,\tau)$ reduces to the function σ^2_{DMA} [4].

[4] S. Arianos and A. Carbone *Physica A* **382**, 9 (2007)

DAX (Deutscher AktienindeX)



Prices

(a) $P(t)$

Log Returns

(b) $r(t) = lnP(t + t') - lnP(t)$

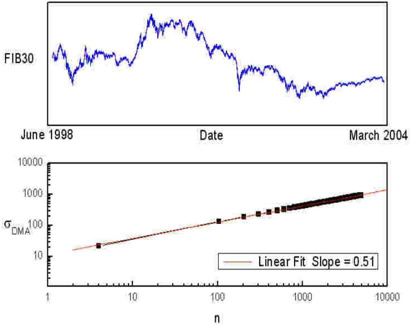
Volatility

$\sigma_T(t)^2 = \frac{1}{T-1}\sum_{t=1}^{T}[r(t) - \overline{r(t)_T}]^2$

(c) with $T = 300min$ (one half trading day)

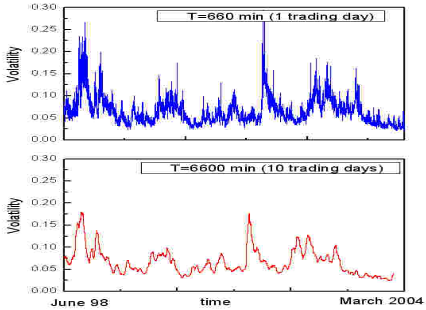
(d) with $T = 660min$ (one trading day)

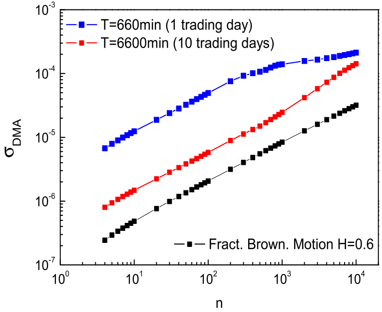
FIB30 futures: PRICES



The FIB30 is a future contract on the MIB30 index, which considers the thirty firms with higher capitalization and trading (the top 30 blue-chip index) of the MIBTEL (www.borsaitaliana.it) (Since 2004 it is named *S&P* MIB index).

FIB30 futures: VOLATILITIES





Cross-Similarity of two stochastic series x(t) and y(t)

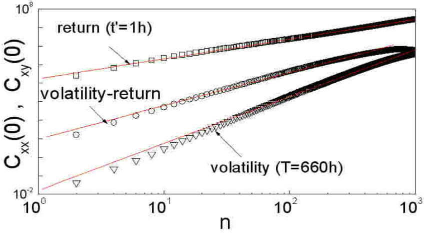
In many cases, the coupling between two systems (time series) might be of interest. A measure of **cross-correlation** can be implemented [5] by defining:

$$\sigma_{xy}(t,\tau) \equiv \left\langle [x(t) - \widetilde{x}(t)][y^*(t+\tau) - \widetilde{y}^*(t+\tau)] \right\rangle$$

Now it must be $\tau \neq 0$. Moreover $x^* = x(t)$ and $y^* = y(t)$
$\widetilde{x}_n(i) = \frac{1}{n}\sum_{k=0}^{n} x(i-k)$ and $\widetilde{y}_n(i) = \frac{1}{n}\sum_{k=0}^{n} y(i-k)$.

[5] S. Arianos and A.Carbone, J. Stat. Mech.: Theory and Experiment **P03037**, (2009).

Cross-Similarity of return and volatility

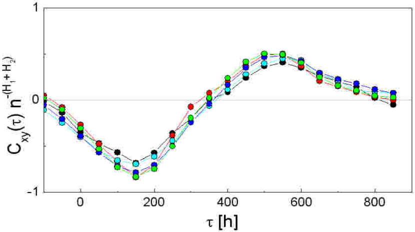


$$\sigma^2_{DMA} \sim n^{H_1+H_2}$$

The log-log plot is a straight-line. The slope is given by the sum of the Hurst exponents H_1 and H_2 of the two time series.[6]

[6] S. Arianos and A.Carbone, J. Stat. Mech: Theory and Experiment **P03037**, (2009).

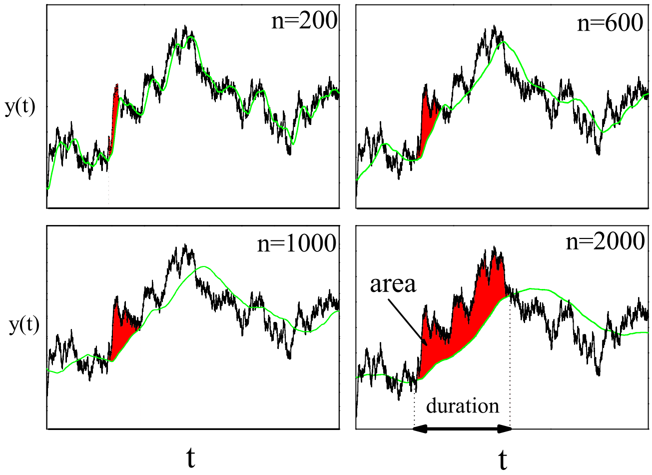
Leverage effect: DAX



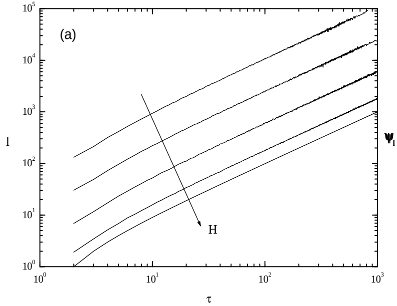
DMA Cross-correlation as a function of the lag τ for the return and volatility of the DAX series.[7]

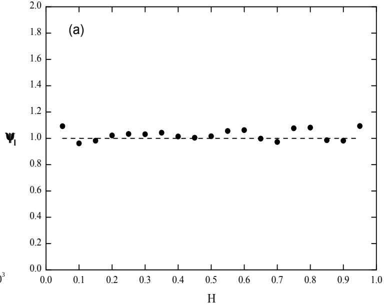
[7] S. Arianos and A.Carbone, J. Stat. Mech: Theory and Experiment **P03037**, (2009).

Scaling law of the moving average clusters [8 9]



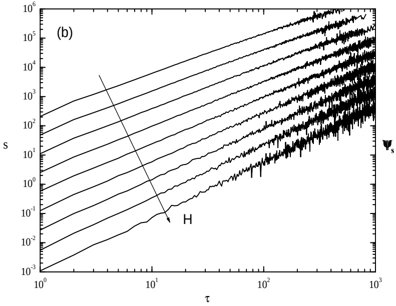
Scaling law of the cluster lengths ℓ

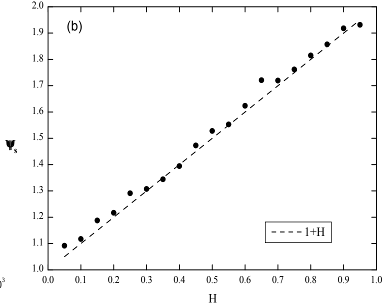




$$\ell \sim \tau^{\psi_\ell} \qquad \psi_\ell = 1$$

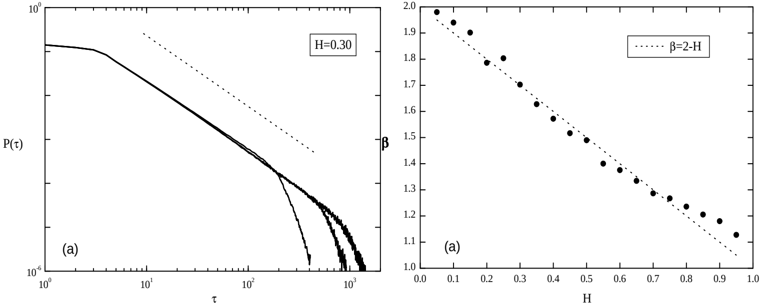
Scaling law of the cluster areas s





$$s \sim \tau^{\psi_s} \qquad \psi_s = 1 + H$$

Scaling law of the pdf of the cluster length and duration

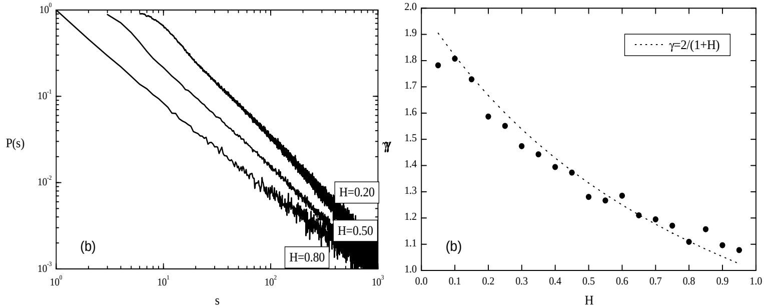


$$P(\tau, n) \sim \tau^{-\beta} \mathcal{F}(\tau, n)$$

with $\beta = 2 - H$ the fractal dimension of the time series and $\mathcal{F}(\tau, n)$ defined as:

$$\mathcal{F}(\tau, n) = e^{-\tau/\tau^*}.$$

Scaling law of the pdf of the cluster area

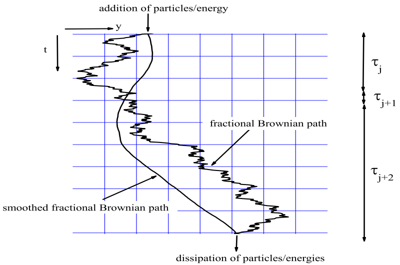


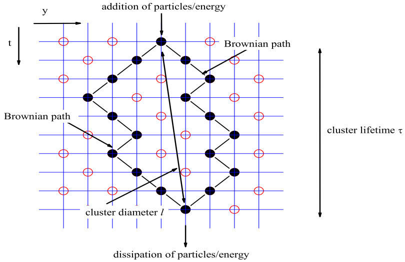
$$P(s,n) \sim s^{-\gamma} \mathcal{F}(s,n)$$

with $\gamma = 2/(1+H)$ and:

$$\mathcal{F}(s,n) = e^{-s/s^*}.$$

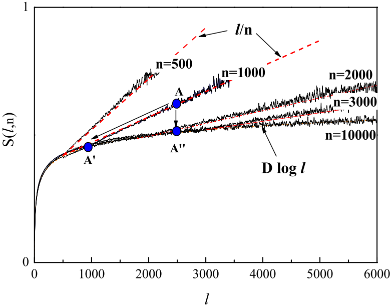
Self-organized criticality of the moving average clusters





	Moving Average Clusters	SOC Clusters
$\ell \sim \tau^{\psi_\ell}$	$\psi_\ell = 1$	$\psi_\ell = 1$
$s \sim \tau^{\psi_s}$	$\psi_s = 1 + H$	$\psi_s = 3/2$
$P(\ell, n) \sim \ell^{-\alpha}$	$\alpha = 2 - H$	$\alpha = 3/2$
$P(\tau, n) \sim \tau^{-\beta}$	$\beta = 2 - H$	$\beta = 3/2$
$P(s, n) \sim s^{-\gamma}$	$\gamma = 2/(1 + H)$	$\gamma = 4/3$

Entropy of long-range correlated time series[10]

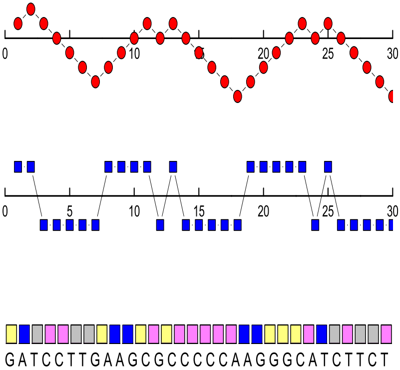


$$S(\ell, n) \equiv - \sum_{\mu(\ell,n)} P(\ell, n) \log P(\ell, n).$$

$$S(\ell, n) = S_0 + D \log \ell + \frac{\ell}{n}.$$

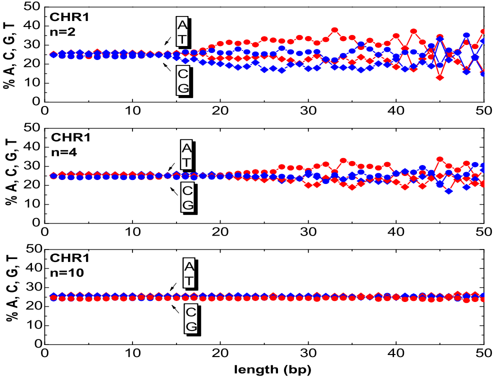
[10] A. Carbone and H.E. Stanley *Physica A* **385**, 21 (2007)

Genome Heterogeneity

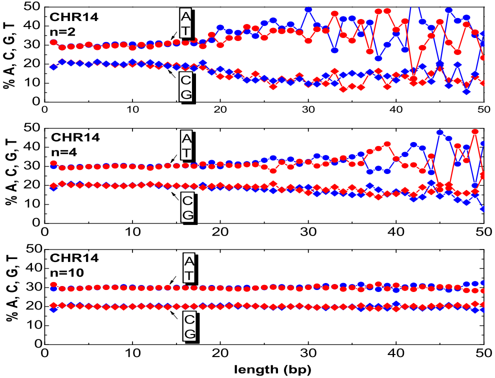


1. The sequence of the nucleotide bases ATGC is mapped to a numeric sequence.
2. If the base is a purine (A,G) is mapped to $+1$, otherwise if the base is a pyrimidine (C,T) is mapped to -1.
3. The sequence of $+1$ and -1 steps is summed and a random walk $y(x)$ is obtained.

Genome Heterogeneity

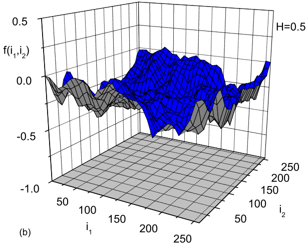


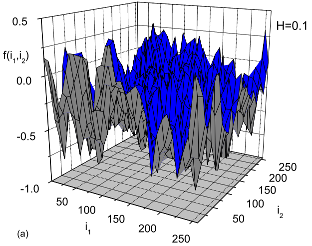
Genome Heterogeneity



DMA for high-dimensional fractals (e.g. selfsimilar networks, rough surfaces) [11]

Fractal surfaces

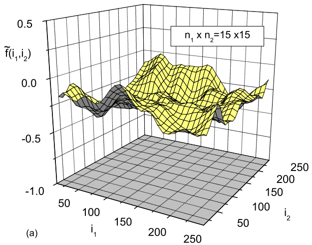


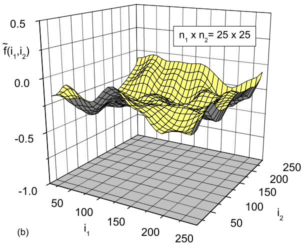


[11] A. Carbone *Phys. Rev. E* **76**, 056703 (2007)

DMA for high-dimensional fractals (e.g. selfsimilar networks, rough surfaces) [12]

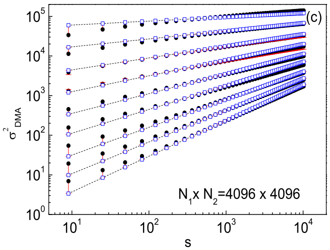
Moving Average Surface for the fractal surface with $H = 0.5$.





[12] A. Carbone *Phys. Rev. E* **76**, 056703 (2007)

DMA for high-dimensional fractals (e.g. selfsimilar networks, rough surfaces) [13]



The plot of σ^2_{DMA} vs. $\mathcal{S} = \sqrt{(n_1^2 + n_2^2)}$ is a straight line.
It corresponds to the scaling relation:

$$\sigma^2_{DMA} \sim \left[\sqrt{(n_1^2 + n_2^2)}\right]^{2H}$$

i.e.:

$$\sigma^2_{DMA} \sim \mathcal{S}^{2H}$$

[13] A. Carbone *Phys. Rev. E* **76**, 056703 (2007)

THANK YOU!