Empirical Analysis of Time Series

Ausloos, M
Universite de Liege

2002

Empirical Analysis of Time Series

M. Ausloos
SUPRAS (Services Universitaires Pour la Recherche et les Applications en Supraconductivité) & GRASP (Group for Research in Applied Statistical Physics), University of Liège, Belgium

Time series occur in many fields of biology, physics, chemistry, engineering. Much work has been recently performed in statistical physics using specific mathematical techniques on various time series pertaining to so-called nonlinear phenomena. Several methods, beyond the Fourier transform, are presented here. To distinguish between noise and deterministic content is the major challenge. Various phenomena are used for illustration. Some emphasis on findings and still questions will be drawn from problems in finance due to the existence (or not) of long-, medium-, short-range (power-law or not) correlations in such economic systems. The Fourier transform, the Hurst rescaled range, the instantaneous detrended fluctuations, the moving averages, and the Zipf-plots analysis methods will be recalled. They raise questions about fractional Brownian motion properties, or in sorting out correlation ranges and predictability. Among spectacular results, the possibility of crash predictions will be indicated when there is an underlying discrete scale invariance. Other time series for meteorology and electronics phenomena are also presented in order to discuss stratus cloud breaking and dielectric breakdown through avalanches for illustration purpose and to indicate that there are other widely open fields of possible investigations.

Key words:

I. Introduction

It is hereby intended to review, as I did in the oral contribution, in a rather practical way, and within a useful framework, numerical investigations on time series, with some emphasis on financial applications and materials science phenomena. Emphasis will be on outlining findings reported in easily available previous publications. No attempt is made at providing an extensive review of the literature. Moreover not all practical time series analysis techniques will be covered, the more so since excellent talks were given by H. Kantz, and several other techniques are likely reviewed in his contribution to the school proceedings. More specifically the following section II is divided into several subsections entitled (i) Fourier transform, (ii) Hurst exponent, (iii) discrete scale invariance, (iv) detrended fluctuation analysis, (v) moving average technique, (vi) Zipf method. Section III is pertinent to financial time series and contains two subsections, pointing to (1) the unusual log-periodic time series near some financial crashes, (2) the powerful instantaneous detrended fluctuation analysis method. In section IV, practical subjects in meteorology and electronics are mentioned without introducing any new technique, but in view of illustrations and in order to indicate that there are other widely open fields of possible investigations, with the recalled techniques. There are many others of interest, like i-variability diagram, the multifractal and the wavelet techniques, - they are not discussed nor even presented here. A few reference books can be recalled [1-5], but many others are also found to be excellent ones.

As a preliminary statement let it be recalled that error bar sizes are not often mentioned in modern physics papers, but let authors be encouraged to make appropriate statistical tests in order to convince readers of the confidence with which their numerical results are obtained or displayed. In contrast to economy, where statistical analysis has overshadowed a true or critical examination of hypotheses, thus of models, statistics is often lacking in physics reports because experimental evidence or apparent similarities are often accepted as valid arguments in proving a thesis. This should be improved, or revived, in future physics work, in order to convince in a more acceptable way. We agree, in physics, that error bars are to be taken as merely indicative of departure from truth. However when physicists intend to introduce physics models and theoretical ideas into other fields they should abide by the others rules and admit statistical evidence as an argument. That is one way to do better science, and numerics often impresses. Again several books on statistical tests can be usefully mentioned, but there are many other good ones than those given here [6-7].

In many cases, that I describe or have treated, even though not all analyzed data techniques and results have been checked with respect to standard physical signals (like fractional Brownian motion, white noise, etc.) most results when published were thought to be sufficiently in agreement with what should be thereby expected, and that is enough! Nevertheless I emphasize that it is necessary to warn against mistreatment of data. One can sometimes obtain quite varied mean values, large error bars or even unreliable parameters. One has to wonder why, and whether it makes sense. It is however extremely difficult to give general rules about how to be satisfied, without relying on abstract statistical inferences, with extracted data/parameter values from nonlinear fit techniques. However, it seems that at this time there is no strict need to request extreme precision from the techniques, because the intuition has still to play a great role in marginal fields of science.

Practical or physics model aspects are not intended here. In particular, safety factors on (financial investment) strategies have to be used when going beyond an estimated risk or a physical conclusion about underlying phenomena. Models can be more easily suggested in order to explain findings related to meteorology or electronics features. Indeed, Ising-like [8], percolation-like [9], sand pile-like [10] models are an invaluable source of analogies and can be used as standard for explaining results and suggesting predictions.

II. Introduction to Time Series Analysis

In order to obtain universal laws in dynamic stochastic systems one has to distinguish true noise from chaotic behavior, and sort out causality [11] and coherent sequences from random ones [12-14] in experimentally obtained signals. The stochastic aspects are not only found in the statistical distribution of underlying frequencies characterizing the Fourier transform of the signal, but also in the amplitude fluctuation distribution and high moments or correlation functions. Remember that amplitude, frequency and phases are the oldest physical characterizations of a time dependent signal. All developments and models should take care of the distributions and relationships between these quantities [14,15].

II.A) Fourier Transform

For non periodic signals, the Fourier transform (FT) [16] has been introduced in order to sort out the distribution of frequencies of interest, i.e. the density of modes, either containing bands, isolated frequencies or both. This distribution can indicate the type of persistence of a phenomenon. The density of modes might be also examined to find whether the frequencies are distributed in a geometrical progression, rather than following an ordinary/usual arithmetic progression, i.e. whether the phenomena might be log-periodical, like in earthquakes or stock market crashes. The Fourier technique is often used nowadays through the Fast Fourier Transform (FFT) method. Yet one should warn the reader about frequent mishandling of the technique.
As a practical example, consider the evolution of IBM share price from Jan. 1, 1990 to Dec. 31, 2000 (Fig.1) A FFT of this signal is shown in Fig.2. Often, in the scaling hypothesis framework [17] one considers that a (FT) spectrum should look like a power law with an exponent β. One says that the power spectrum is self-affine. Notice the poor fit to a power law of the IBM spectrum. A smoothing the spectrum via integration over appropriate choice of the bin sizes can reduce the error bars, but this can lead to drastic changes of the power law exponent.
Notice that one should not always consider that measurements are taken at equal time intervals, and be aware that this introduces spurious frequencies. For financial time series, there are e.g. holidays and weekends. Quotations of share price or stock market indices are given at various ticking times, not always regularly, depending on the orders which are given.
The fractal dimension [18-22] D is often used to characterize the apparent roughness of a signal. Several methods are used for measuring D, like the box counting method, which is not quite efficient; many others are found in the literature as seen in Refs. [19-22]. For topologically one dimensional systems, the fractal dimension D is related to the

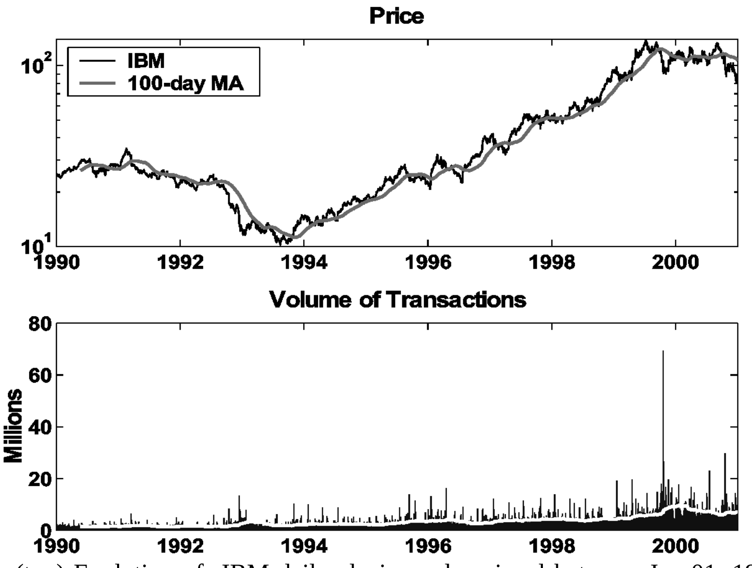

Fig. 1: (top) Evolution of IBM daily closing value signal between Jan 01, 1990 and Dec 31, 2000, with a moving averages, $M_\tau^{(t)}$ = 100 days; (bottom) volume of transactions with moving average over 100 days.

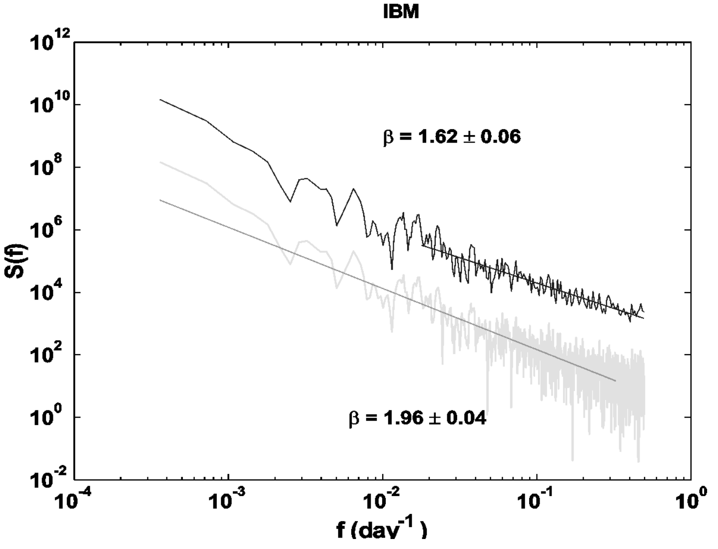

Fig. 2: Fast Fourier Transform of IBM signal (Fig. 1) on a log-log plot. The non smoothed spectrum and a smoothed one (through a running sum over bins with equal sizes in log scale) are shown to give different best fits to a power law.

exponent β by β = 5-2D. A Brownian motion is characterized by D = 3/2, and a white noise by D= 2.0 [20,21]. The FT of financial and other not so classical science signals often lead to large error bars, as in Fig.2; observe that a non smoothen spectrum seems to give a best fit with a slope (1.96) but a log bin size smoothening gives 1.62, in different ranges. Thus a FT is not the best way to measure D, nor to obtain a scaling range.
The coherence and roughness aspect of a signal are masked in this one-shot analysis. Only the persistence behavior is touched upon [15]. Nevertheless the FT technique should be used as the very first step of any signal analysis in order to observe what periods or frequencies exist in a signal, and also what are the missing ones, - in other words whether there are cycles. If so, folding of signals in the appropriate interval may help in further analysis. This is very relevant when a detrended analysis has to be attempted next.

II.B) Hurst Exponent

Hurst developed the rescale range analysis, a statistical method to analyze long records of natural phenomena [23,24]. The range R is the difference between the minimum and maximum of accumulated values (or of the cumulative sum of X(t,τ)) of the phenomenon at each discrete time t value over a time span τ; the standard deviation S is estimated as usual from the observed values $X_i(t)$. Practically, one should be listing the differences between the observed value at a discrete time t over an interval with a size □□on which the mean (not the linear trend) has been taken. The upper y_M and lower y_m values in that interval define the range $R = y_M - y_m$. This means that for a (discrete) self-affine signal y(t), the neighborhood of a particular point on the signal can be rescaled by a factor λ using the roughness (or Hurst) exponent Hu, thereby defining a new signal λ^{-Hu} y(λt). For the exponent value Hu, the frequency dependence of the signal so obtained should be undistinguishable from the original one, i.e. y(t).
Hurst found that the ratio R/S is described for a large number of natural phenomena by an empirical relation such that R/S behaves like a power law of τ with an exponent, ... Hu. The Hurst exponent Hu for statistically independent processes with finite variances, i.e. in the absence of long-run statistical dependencies, should give Hu = 0.5.
The relation between the Hurst exponent and the fractal dimension is simply D=2-Hu for univariate series. That means that for statistically independent fractional Brownian movement, with D=1.5, the Hurst exponent should be Hu=0.5, as said before. A Hurst exponent of 0.5<Hu<1 corresponds to a profile-like curve showing persistent behavior. Persistence means that if the curve has been increasing for a given time step, it is expected to continue for the next time step. A Hurst exponent of 0<Hu<0.5 indicates antipersistent behavior. After a decrease, an increase is expected. An antipersistent behavior has a rather high fractal dimension (1.5<D<2), corresponding to an apparently very ugly (in my opinion) profile, to a “curve” which fills up the plane.

II.C) Discrete Scale Invariance

A system with discrete scale invariance (DSI) is one for which a partial breaking of the scaling symmetry, is reflected in the existence of a hierarchy of characteristic scales I0, I0l, I0l2, ..., where l is a preferred scaling ratio and I0 a microscopic cut-off. Signatures of discrete scale invariance have recently been found in a variety of systems ranging from rupture, earthquakes, and financial market crashes to growth phenomena and “animals” in percolation science [25]. DSI is a quite general feature, albeit subtle phenomenon. Indeed, the practical problem in uncovering an underlying discrete scale invariance is that standard ensemble averaging procedures destroy it as if it was pure noise. This is due to the fact, that while l only depends on the underlying physics, I0 on the contrary is realization-dependent.
The evolution of a y(t) signal with a DSI can be in fact taken as a power law with a complex exponent m + i ω.

$$y = A + B\left(\frac{t_c - t}{t_c}\right)^{-m}\left[1 + C\cos(\omega \ln\left(\frac{t_c - t}{t_c}\right) + \varphi\right] \quad \text{for } t<t_c \tag{1}$$

here t_c is some sort of crash-time or rupture point, while A, B, m, C, ω and ϕ are free parameters. A useful approximation is when m = 0, such that

$$y = A + B\ \ln\left(\frac{t_c - t}{t_c}\right)\left[1 + C\cos(\omega \ln\left(\frac{t_c - t}{t_c}\right) + \varphi\right] \quad \text{for } t<t_c \tag{2}$$

Thus y(t) can diverge at t= t_c, if the exponent m is positive (or remains at a finite value otherwise) but with an extra contribution which is periodic like. The period of these oscillations become shorter as one approaches the rupture point at t= t_c. This law is similar to that of critical points at so-called second order phase transitions [17], and generalizes the scaleless situation [25]. This relationship was already proposed in order to fit experimental measurements of sound wave rate emissions prior to the rupture of heterogeneous composite stressed up to failure [26]. The same type of complex power law behavior has been observed as a precursor of the Kobe earthquake in Japan [27]. Such log-periodic corrections have been recently reported in biased diffusion on random lattices [28]. As a consequence, the phenomenon could be viewed as taking place on a discrete fractal system.
A schematic representation of such a system is a Cayley tree as studied by Amaral and coworkers can be found in [29]. Another representative of a Cayley tree leading to a discrete fractal representation, i.e. a Sierpinski Gasket, can be found in [30], where the Oct. 97 financial crash has been discussed (see also Sect. III).

II.D) Detrended Fluctuation Analysis

The Detrended Fluctuation Analysis (DFA) technique [31] consists in dividing a time series or random one-variable sequence y(t) of length N into N/m equal size nonoverlapping boxes. The variable t is discrete, evolves by a single unit at each time step between t=1 and t=N. No data point is supposed to be missing. In other words, when applied to financial data, breaks due to holidays and weekends are disregarded. Nevertheless, the τ units are said to be days in the following, a week has often ... 5 days, and a year about ... 250 days. Thus let each box contain τ points and N/τ be an integer. The local trend in each τ-size box is assumed to be linear, i.e. it is taken as z(t)= a t + b. In each τ-size box one next calculates the root mean square deviation between y(t) and z(t). The detrended fluctuation function $F(\tau)$ is then calculated following

$$F^2(\tau) = \frac{1}{\tau}\sum_{t=k\tau+1}^{(k+1)\tau} \left|y(t)-z(t)\right|^2 \qquad k = 0,1,\ \ldots\left(\frac{N}{\tau}-1\right) \tag{3}$$

Averaging $F^2(\tau)$ over all N/τ box sizes centered on time τ gives the fluctuations $<F^2(\tau)>$ as a function of τ. The calculation is repeated for all possible different values of τ. A power law behavior is expected as

$$\left(< F^2(\tau)\right)^{1/2} \approx \tau^{\alpha}$$

An exponent α = 1/2 in a certain range of τ values implies the existence of long-range correlations in that time interval as in the fractional Brownian motion [21]. Such correlations are said to be “persistent” or “antipersistent” when they correspond to α >1/2 and α <1/2 respectively. E.g. the case of the IBM signal (Fig. 1), a DFA leads to a fine straight line fit log τ = 5 and log τ = 100, i.e. for an interval of time between 4 days and 100 days (Fig. 3). This interval is called the scaling range. Outside the scaling range the error bars are larger due to so called finite size effects, and/or the lack of numerous data points. Notice that the IBM signal is rather Brownian motion like.

II.E) Moving Average Analysis

Consider a time series y(t) given at N discrete times t, and ask the question whether there is a trend in the signal. This is answered by looking at the so-called moving average [32]. The series (or signal) moving average M_τ (t) over a time interval τ is defined as

$$M_\tau(t) = \frac{1}{\tau}\sum_{i=t}^{t+\tau-1} y(i-\tau) \qquad t = \tau+1,\ldots, N \tag{4}$$

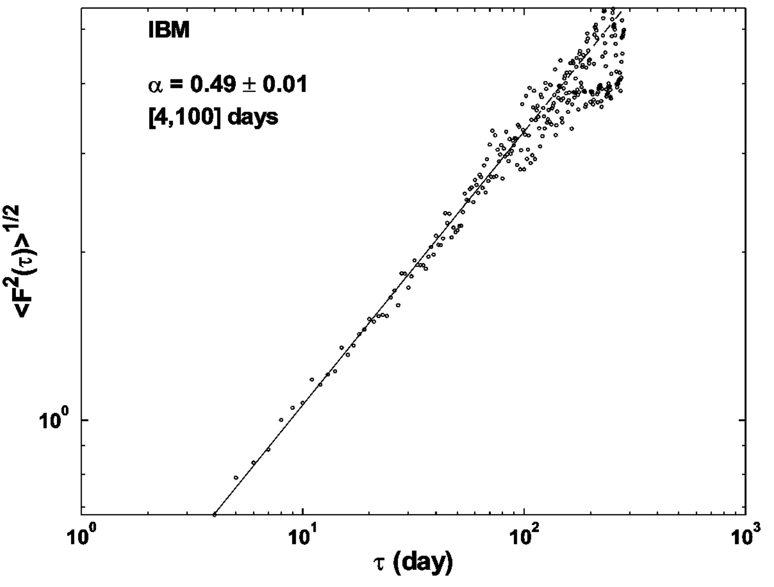

Fig. 3: Linearly Detrended Fluctuation Analysis function of IBM signal (Fig.1) on a log-log plot. The Brownian motion behavior corresponds to a slope 1/2 on this type of plot. The best linear fit seems to give a slope = 0.49±0.01 between 4 and 100 days.

i.e. the average of y(t) over the last τ data points. For simplicity we suppose that the ticking times are equally spaced. One can easily show that if the signal y(t) increases (decreases) with time, M_τ (t) <y(t) (M_τ (t) >y(t)). Thus, a moving average captures the trend of the signal over a given time interval τ. The IBM daily closing value price signal between Jan. 01, 1990 and Dec. 31, 2000 is shown in Fig. 1 (top figure) together with the τ=100 day moving average taken from Yahoo [33]. The bottom band in Fig.1 shows the daily volume given in millions. Sometimes the volume of transactions can also be analyzed in terms of a moving average. Due to the large number involved, one often takes the log of the volume before making a moving average. Be aware that this is a non linear transformation of the primary data. One can also sometimes use a log transformation for the financial data themselves, before making moving averages, in particular when looking for financial returns).

There can be as many moving averages as τ intervals. Thus the need for specifying such values at first. The shorter the τ interval the more sensitive the moving average. However, a too short moving average may give false messages about the long time trend of the signal.

Other features of the trends are the intersections between two moving averages $M_{\tau 1}$ and $M_{\tau 2}$ which are usually due to drastic changes in the trend of y(t). Consider two moving averages of y(t) signal with $\tau_1 < \tau_2$. If y(t) increases for a long period of time before decreasing rapidly, $M_{\tau 1}$ will cross $M_{\tau 2}$ from above. This event is called a "death cross" in

empirical finance [34] In contrast, when $M_{\tau 1}$ crosses $M_{\tau 2}$ from below, the crossing point coincide with an upsurge of the signal y(t). This event is called a "gold cross" [34].
It has been of interest to us to study the density ρ of crossing points between two moving averages as a function of the size difference of the τ's defining the moving averages [35,36]. It was unexpectedly found that such a density obeys a simple law,

$$\rho = \frac{1}{\tau_2}\left[\Delta\tau(1-\Delta\tau)\right]^{Hu-1}$$

As observed on Fig. 4 for $\Delta\tau$ and τ_2 = 100 days. Notice that the above law generalizes to fractional Brownian motion that found in Feller [6] from usual Brownian motion. The above law has not yet been derived theoretically for fractional Brownian motion. The numerical results for the IBM signal are shown in Fig. 4. Based on this finding, a new and efficient approach has been suggested in [35] in order to estimate an exponent that characterizes the roughness of a signal ... the Hu exponent Following that line of investigation the case of the IBM roughness exponent can be deduced from Fig. 3, for the time interval considered here above.

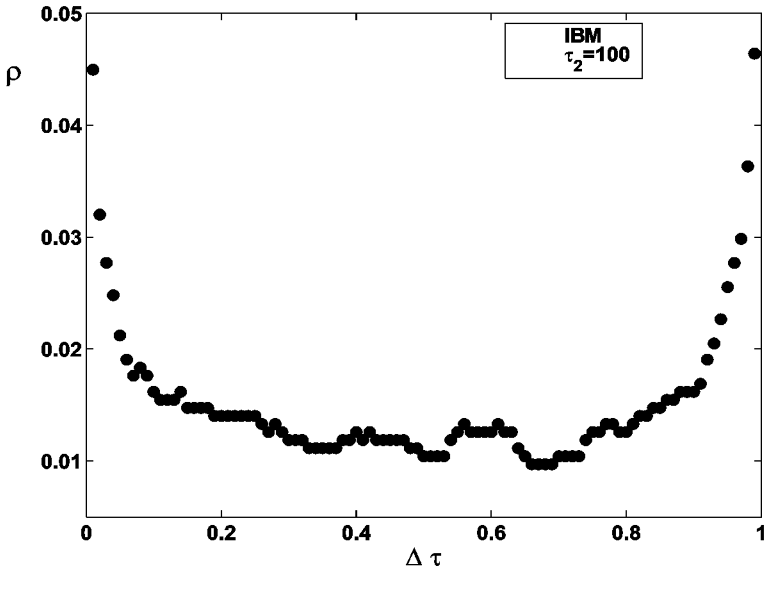

Fig. 4: The density ρ of crossing points between two moving averages as a function of the relative difference $\Delta\tau$ with fixed τ_2=100 days.

II.F) Zipf Analysis

In order to study in a pedestrian way various complex signals or “texts” [37], and economy (size of sales and firms) data [38] one relies also over the so-called Zipf-plot analysis [37]. This technique was originally introduced in the context of natural languages and is performed by calculating the frequency of occurrence f of each word in a given text. By sorting out the words according to their frequency, a rank r can be assigned to each word, with r=1 for the most frequent one. For natural languages, one observes a power law

$$f \sim r^{-\zeta}$$

with an exponent ζ close to one for any language. The occurrence of this power law has been suggested [37] to be due to the hierarchical structure of the text as well as the presence of long range correlations (sentences, and logical structures therein). Therefore it is of interest, just like a FT to make a Zipf plot from the start of a stochastic signal analysis. The case of the IBM signal is shown in Fig. 5. One recognizes easily that there is some relationship between a Zipf plot and an integrated distribution function of a signal fluctuations under very specific conditions.

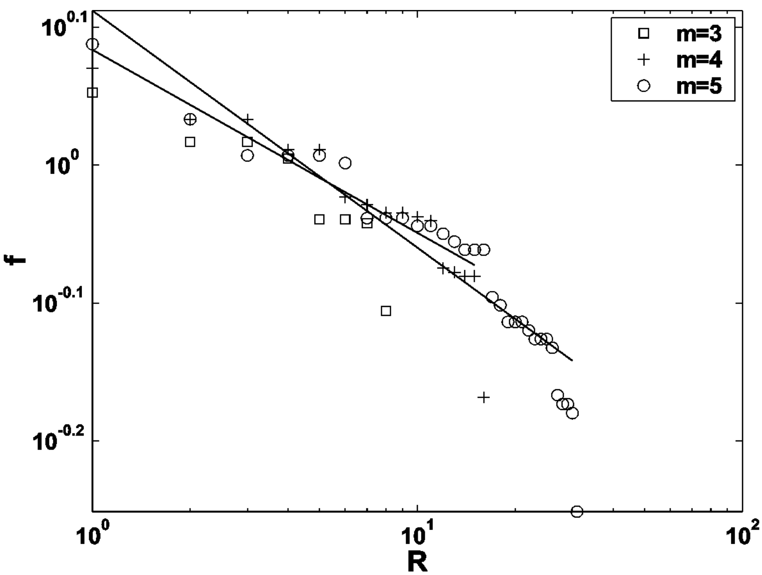

Fig. 5: Zipf plot of IBM signal (Fig.1) for different length of the “words” with $\zeta_4 = 0.13 \pm 0.01$ for m=4 (crosses) and $\zeta_5 = 0.17 \pm 0.01$ for m=5 (circles).

A simple extension of the Zipf analysis is to consider words strictly made of m characters without considering the white spaces. In the latter case, a power law in f^{β} is observed in correlated sequences. A conjecture

$$\zeta = |2Hu-1|$$

has been proposed [39,40] that Hu is the so-called Hurst exponent measuring the long-range correlations lying in a signal. Such an analysis has been e.g. applied to correlated systems like DNA sequences [41].
A random walk can be easily translated into a sequence made of two characters 'u' and 'd' representing the up and down fluctuations respectively. For a true random walk, the exponent ζ is expected to be zero, i.e. the Zipf plot is horizontal since all patterns of length m should appear with the same probability. However, the situation is different for biased random sequences, as in financial markets [43,44].
In view of this, more elaborated signal transformations can be considered [43,44] like replacing respectively small up, large up, small down and large down fluctuations by e.g. 'u', 'a', 'd', 'b' characters. The text (or signal) can be analyzed according to the number of characters k in the alphabet and the length m of letters in a " word ". The ζ exponent is thus generalized and so depends on k and m. We have also found that it can apparently depend on normalization rules for the frequencies. A good question not fully answered is the role of the signal trend(s) in the resulting value of $\zeta(k,m)$.

III. Financial Aspects Applications

From the above two peculiarities should draw some reader attention. On one hand the distribution of frequencies in a signal might not be a simple law. In fact it was pointed out that the signal can be log-periodic, if there is some underlying DSI. This has been usefully taken into account to analyze big financial crashes (Sect. III.A). On the other hand, the non-stationarity of signals, i.e. the time dependence of the characteristic parameters (D, β, Hu, ...) as a function of time should be envisaged here also such peculiarity can be turned into some advantage in e.g. devising an investment strategy, see Sect. III.B.

III.A) Financial Crashes

Much has already been written about big financial crashes, and our purpose is surely not to review the necessarily controversial statements made about such drastic events, in the physics literature. We only point out that there is some similarity between 1929, 1987, 1997 and other crashes. All of these contain a log-periodic ingredients in the signal, whatever the index which has been examined, DJIA, S&P500, Nikkei, DAX, ... To make things short, we have used the simplified law written here above in which we assume that the financial crash occurs after some euphoria time described by a divergent

log behavior of whatever index. Thus six parameters are involved into describing the index evolution, starting from Eq. (2).
It should be stressed that the numerical parameter values in Eq.(1) are not robust against small data perturbations. A nonlinear seven parameter fit is hardly stable indeed from a numerical point of view. E.g., removing the contribution of the oscillations in Eq.(1), i.e. setting C=0 thereby reducing the number of free parameters to that in Eq.(2) leads to an exponent m=0.7 quite large. In fact, various values of m were reported to be ranging from 0.53 to 0.06 for various indices and crashes; the sign of m is not even well asserted. The optimum test we used consisted in separating the most diverging term from the others and next searching for the correction to the classical scaling law. The technique in order to find tc and the crash amplitude consisted thus into two distinct problems: (i) a fit to the divergence, and (ii) a fit to the oscillating signal. A search was made for the best common solution in particular involving t_c.... and that occurred [45,46]! Our analysis suggested as early as August 1997 that a crash or a rupture point was highly probable between the end of October and mid-November 1997.
Moreover for both investigated time period cases, i.e. 80-87 and 90-97, it was found that the value of λ seems to be almost a constant [47] corresponding to ω ca. 6. Thus a set log-periodic oscillations seems realistically probable with a simple parameter. An analysis along similar lines of thought, emphasizing a no-divergence at t_c was discussed for the Nikkei [48,49] and NASDAQ April 2000 crash [50].
It has been emphasized elsewhere that the parameter λ is a hint toward the most relevant connectivity of the market, i.e. a measure of the number of main speculators/actors [51], - by analogy with avalanches of sandpiles built on fractal lattices. Other authors find a ω value ca. 8-12 [52]. Other considerations on imprints of log-periodic self-similarity in the stock market can be found in the references quoted here above and in [53-56] as well.

III.B) DFA-based Strategy

The DFA method can be used to probe instantaneous variations of the nature of correlations. In order to do so one has to first find whether a scaling law holds and in what range, then take a box of the size equal to the upper scaling range and repeat the analysis moving the box along the data. larger error bars do occur, due to the a priori restricted finite size of the data, but an alpha evolution can then be done. The variations in ($d\alpha/dt$) could be related to economic events [57] following political events and some panic storm spreading over financial markets.
For example, in the case of the IBM data hereby illustrated, we have constructed a so-called observation box (a time window) of 2 year width placed at the beginning of the data, and calculated α for the data contained in that box. Then, we have moved this box by 1 point (1 day) toward the right along the financial sequence and again calculated α. Iterating this procedure for the Fig.1 sequence, one can obtain an "instantaneous measurement" of the degree of long-range correlations. The results are shown in Figure 6, as the α-curve is smoothed through 40 days moving average. The α exponent value obvi-

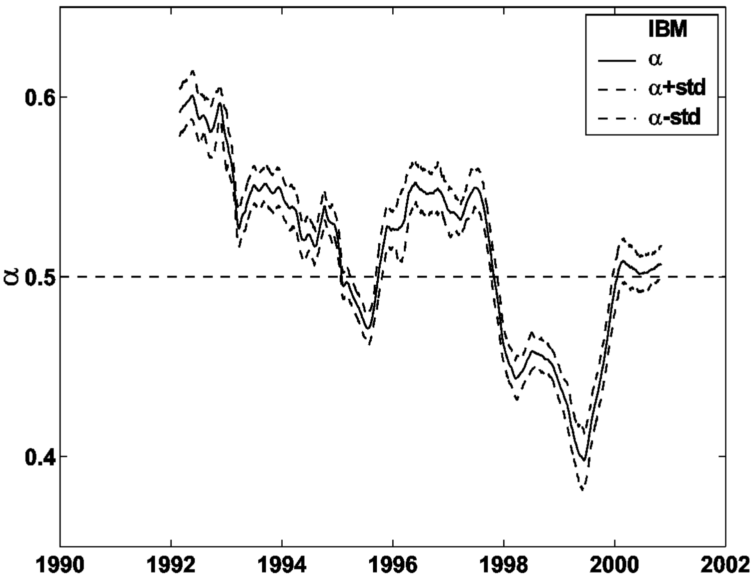

Fig. 6: Time dependence of the characteristic (Hurst) exponent following the DFA method for the IBM signal corresponding to data in Fig. 1.

ously varies with the date but is mostly near 1/2. However, both around spring 1995 and spring 1999, the exponent α has a sharp minimum below the α =1/2 value. The second event is the most dramatic one. This is similar to what is also observed along DNA sequences where the α-exponent drops below 1/2 in so-called non-coding regions [31].

Interestingly this analysis can be used for inventing an investment strategy. Indeed if the value of the exponent α is greater than 1/2, it can be expected that the measurement is in a so-called region of persistent fluctuations, and conversely if α is smaller than 1/2, the measurement day is in a so-called non-persistent fluctuation region. In other words, there is an expectation to have the IBM share price within an antipersistent or persistent sequence. In order to predict the (shortest range) next (up or down) fluctuation we should know whether the measurement day fluctuation has been up or down. A two by two matrix can be constructed in order to build an investment rule. For example we decide to take for granted as an investment rule that when α is greater than 1/2 we buy the IBM if the previous fluctuation was down. As a rule we have taken an arbitrary error bar (3.0%) on the value of α = 1/2 in order to reduce Brownian noise in the vicinity of that value. If α is in the "no-confidence interval" [0.47,0.53], we remain inactive, neither buy nor sell IBM.

In so doing one can display Fig. 7 the capital gain over the time interval which is considered. Notice that the strategy can only be implemented after the first day of the window size, i.e. here starting on Jan 01, 1992. We have assumed that there is no transaction fee. The gain data is normalized to a value of 1 at the starting date. The results are

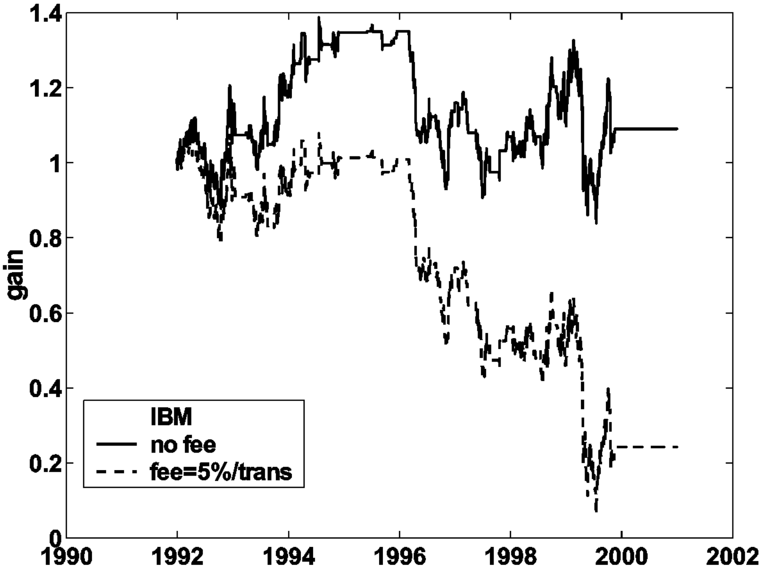

Fig. 7: Theoretical gain for a strategy based on the time dependence of the characteristic (Hurst) exponent time dependence (see text) following the DFA method for the IBM signal displayed in Fig.1

rather remarkable at the beginning of the investment period, though get worse thereafter. No need to say that ta real strategy should be more elaborate, for example combining it with conclusions drawn from the moving average analysis discussed here above. See [58-60] for other applications.

IV. Practical Subjects: Breakdown through Avalanches

There are many applications of such analysis techniques to data arising from biology, geology and geophysics, and understood within fractals and chaos frameworks. For the latter subjects see [61]. In the case of ion transport through membranes, I indicated in the lecture and shown recent results [62] to which only reference is made here. Rather I emphasize two cases, discussed in the lectures but which show the versatility of the DFA technique.

IV.A) Meteorology

Among cases in which catastrophes, like rupture, occur the liquid water content and brightness temperature of stratus clouds [63,64] have been studied [65,66]. The data stems from microwave radiometer measurements. We chose a long set of data points, i.e the period 14:00 UTC April 3 to 24:00 UTC April 5, 1998 for a Oklahoma cloud. A rather sharp transition to clear sky occurred on April 5, 1998, around the 50-th hour. Such a time evolution over a 58 h time period, consists of 10 381 data points [65]. The infrared thermometer brightness temperature (BT) quantifies the cloud at the base as if it is a black body planar object. It does not entirely describe the cloud inner dynamics, in contrast to the liquid water path (LWP) which clearly contains information about the bulk structure of the cloud. It could be argued that the LWP better captures the inner dynamics of the transition from "solid" stratus cloud to "broken" cloud. Nevertheless the cloud surface is typically better "seen" that the inner structure, whence the interest of examining BT data.

To search for existence of correlated sequences we have followed the technique described in Sect. III.B, i.e. constructed a so-called observation box (a window) of size w placed at the beginning of the data; then calculate the DFA exponent for the data in that finite size box and move the box by m minutes toward the right along the signal sequence. By iterating this procedure for the data sequence an instantaneous measurement is obtained for the degree of local long-range correlations [57]. The results from this DFA analysis [65] indicate two well defined regions of scaling with different values of α. The first one corresponds to the first two days when thick stratus clouds existed. The average value of the local scaling exponent over this period is $\alpha = 0.34 \pm 0.02$ with a sharp drop below $\alpha = 0.1$ for the clear sky day (see precise values in [65,66]). These values of \alpha are well defined for a scaling time (range) interval extending between 2 and 25 minutes for the various m and w combinations. For BT, similar values of α (ca. 0.3) are found as well. Note that these values can be interpreted as the H_1 parameter of the multifractal analysis of LWP by Ivanova and Ackerman [67].

Also the value of α being close to 0.3, indicates a very large antipersistence, thus a set of fluctuations tending to induce a greater stability of the system and greater antipersistence of the prevailing meteorology, - in contrast to a persistence of the system fluctuations which would drag the system out of equilibrium. This implies a specific dynamics to be usefully inserted as ingredients of models. The very well defined increase to a value ca. 0.5 is to be interpreted as the existence of a driving ingredient toward more stochasticity and less predictability.

In summary, long-range power law correlations have been shown to occur in stratus cloud system. The appearance of broken clouds and clear sky following a period of thick stratus can thus be interpreted as a non equilibrium transition or a sort of fracture process in more conventional physics. The existence of a crossover suggests [65] two types of correlated events as in classical fracture processes.

IV.B) Electronics

Dielectric breakdown is recognized as a major reliability issue in VLSI (Very Large Scale Integrated) circuits technology [68-72]. However, the mechanisms responsible for breakdown of ultra-thin insulating layers are still not completely understood. Soft

breakdown is a phenomenon observed during constant current or voltage stress of a MOS (Metal Oxide Semiconductor) device with a nm gate oxide. Soft breakdown corresponds to a large increase of the stress induced leakage current and to the occurrence of fluctuations in the time dependence of the gate voltage VG or the gate current JG.
The gate voltage VG(t) after soft breakdown of MOS capacitors with a 2.4 nm SiO2 layer has been measured and reported in [73]. It was found that the VG(t) fluctuation distributions are non-Gaussian, and can be described by a Lévy stable distribution. The Lévy index μ is found to increase logarithmically with the current density J. For low current densities, the gate voltage VG(t) evolves erratically with frequent large drop and burst excursions. For large current density, the Gaussian value $\mu = 2$ should be expected. However, this regime was not observed since at high values of the current density a complete (thermal) breakdown of the gate oxide occurs. The Lévy index was seen to be non-universal and logarithmically increasing with J. When varying the sampling rate dt (the time resolution) of VG(t), the results were found to be robust : the amplitude of the fluctuations were only multiplied by a constant factor which depends on the sampling rate. Indeed, an erratic signal like VG(t) should be confined in an envelope for which the width grows as s~tHu, where Hu is the above Hurst exponent.
The long-range correlations in VG(t) were studied, and analyzed following the DFA technique. The variance F(t) of the signal VG(t) around a local trend z(t) was calculated as discussed in the previous sections. An exponent value Hu=0.25±0.04 was obtained and found to be apparently independent of J. The power spectrum analysis (see Sect. II.A) indicated that the voltage VG(t) is a "$1/f^{\beta}$-noise" with $\beta = 3/2$.
In previous reports [74,75], the relevance of the percolation framework [9] had been underlined for describing the properties of such ultra-thin gate oxides after the occurrence of soft breakdown, as if the process was of probabilistic/geometric nature. This approximation is removed when studying the dynamical properties like the time dependence of the gate voltage as in [73]. Whence a conjecture was presented: the fluctuations and long-range correlations in the gate voltage of ultra-thin gate oxides are most probably related to the trapping-detrapping of electrons in the electrical pathway formed between the electrodes of the capacitors at soft breakdown. When the current density is relatively low, this pathway is a nearly fractal object, like a usual percolating cluster. Therefore, the trapping/detrapping of electrical charges composing the backbone leads to drastic fluctuations and are expected to belong to an non-Gaussian distribution. When the current density is high, the percolating cluster is expected to better fill the available space such that the backbone is quasi-equivalent to the overall percolating cluster. Thus, the local detrapping/trapping of electrical charges does not affect drastically the backbone: a Gaussian behavior is expected.
The above physical interpretation in terms of trapping-detrapping phenomena can lead to imagine new dynamic percolation models for describing the properties of ultra-thin SiO2 layers after soft breakdown and similar ones in mechanics of rupture, e.g. the role of inclusions away from the crack tip. Within this perspective, self-organized critical (SOC) models of pinning/depinning [76] are relevant candidates: they produce $1/f^{\beta}$ noise indeed.

V. Conclusions

There are numerous methods of time series analysis of non linear dynamical systems. We have recalled a few of them quite appropriate in order to search for correlation ranges in fluctuations, and in order to characterize signals in a modern framework, like that of fractals. Illustration of concepts have used for financial series like that of the IBM share price. Moreover we have used the DFA method to search for scaling ranges if any and the type of behavior of correlated fluctuation sequences, sometimes leading to an investment strategy. We have also sorted out correlations and anticorrelations. Two unusual cases have also been recalled, i.e. one pertaining to the meteorology field, the other to material science. The results indicate a very complex behavior, but hint toward better modeling. Time series are full of hidden wonders.

Acknowledgements

Thanks to Dr. J. Holyst and W. Klonowski for inviting me to lecture on the above techniques, results and ideas. Thanks to the sponsors of EUROATTRACTOR. Thanks also to several coworkers who contributed to mentioned papers not only for many useful discussions but also for allowing me to pick up and lean on data from previous work, i.e. in particular thanks to M. Houssa, K. Ivanova, Z. Siwy, and N. Vandewalle. Text illustrations result from K. Ivanova hard and skillful work. Her critical reading should be also mentioned.

References

1. H. Kantz, Th. Schreiber: Nonlinear Time Series Analysis. Cambridge Univ. Press, Cambridge, 1997.
2. C. Diks: Nonlinear Time Series Analysis. World Scient., Singapore, 1999.
3. P.J. Brockwell, R.A. Davis: Introduction to time series and forecasting. Springer Text in Statistics. Springer, Berlin, 1998.
4. Ph. Franses: Time Series Models for Business and Economic Forecasting. Cambridge U. Press, Cambridge, 1998.
5. Ch. Gourieroux, A. Monfort: Time Series and Dynamic Models. Cambridge U. Press, Cambridge, 1997.
6. W. Feller: An Introduction to Probability Theory and its Applications, vol. I. Wiley, New York, 1950; ibid., An Introduction to Probability Theory and its Applications, vol. II. Wiley, New York, 1966.
7. A.C. Bajpai, I.M. Calus, J.A. Fairley: Statistical Methods for Engineers and Scientists. Wiley, Chichester, 1978.
8. B.M. McCoy, Tai Tsun Wun: The Two-Dimensional Ising Model. Harvard Univ Pr., Cambridge, 1973.

9. D. Stauffer, A. Aharony: Introduction to Percolation Theory. Taylor & Francis, London, 1994, 2nd printing.
10. P. Bak: How Nature Works. Copernicus, New York, 1996.
11. C. Brooks, M.J. Hinich: J. Empir. Fin. 6: 385; 1999.
12. P.E. Rapp: Integrat. Physiol. Behav. Sci. 29: 311; 1994.
13. C.J. Cellucci, A.M. Albano, P.E. Rapp, R.A. Pittenger, R.C. Josiassen: Chaos 7: 414; 1997.
14. B.D. Malamud, D.L. Turcotte: J. Stat. Plann. Infer. 80: 173; 1999.
15. M. Ausloos: From Billiards to Monte Carlo. In: K.H.Hoffmann, M. Schreiber (Eds.) Computational Statistical Physics (pp. 153-168). Springer, Berlin, 2001.
16. T.W. Körner: Fourier Analysis. Cambridge Univ. Press, Cambridge, 1988.
17. H.E.Stanley: Phase transitions and critical phenomena. Oxford Univ. Press, Oxford, 1971.
18. B.B. Mandelbrot, D.E. Passoja, A.J. Paulay: Nature 308: 721; 1984.
19. K. J. Falconer: The Geometry of Fractal Sets. Cambridge Univ. Press, Cambridge, 1985.
20. M. Schroeder: Fractals, Chaos and Power Laws. W.H. Freeman and Co., New York, 1991.
21. B. J. West, B. Deering: The Lure of Modern Science: Fractal Thinking. World Scient., Singapore, 1995.
22. P. S. Addison: Fractals and Chaos. Inst. of Phys., Bristol, 1997.
23. H. E. Hurst: Trans. Amer. Soc. Civ. Engin. 116: 770; 1951.
24. H. E. Hurst, R.P. Black, Y.M. Simaika: Long Term Storage. Constable, London, 1965.
25. D. Sornette: Phys. Rep., 297: 239; 1998.
26. J.C. Anifrani, C. Le Floc'h, D. Sornette, B. Souillard: J. Phys. I (France) 5: 631; 1995.
27. A. Johansen, D. Sornette, H. Wakita, U. Tsunogai, W.I. Newman, H. Saleur: J. Phys. I (France) 6: 1391; 1996.
28. D. Stauffer, D. Sornette: Physica A, 252: 271; 1998.
29. L.A.N. Amaral, S.V. Buldyrev, S. Havlin, H. Leschhorn, P. Maass, M.A. Salinger, H.E. Stanley, M.H.R. Stanley: J. Phys. I France 7: 621; 1997.
30. N. Vandewalle, M. Ausloos, Ph. Boveroux, A. Minguet: Eur. J. Phys. B 9: 355-359; 1999.
31. C.-K. Peng, S.V. Buldyrev, S. Havlin, M. Simmons, H.E. Stanley, A.L. Goldberger: Phys. Rev. E 49: 1685; 1994.
32. A.G. Ellinger: The Art of Investment. Bowers & Bowers, London, 1971.
33. http://finance.yahoo.com
34. S.B. Achelis: In: http://www.equis.com/ free/taaz/
35. N. Vandewalle, M. Ausloos: Phys. Rev. E 58: 6832; 1998.
36. N. Vandewalle, M. Ausloos, Ph. Boveroux: Physica A 269: 170; 1999.
37. G.K. Zipf: Human Behavior and the Principle of Least Effort. Addison-Wesley, Cambridge, MA, 1949.
38. M.H.R. Stanley, S.V. Buldyrev, S. Havlin, R. Mantegna, M.A. Salinger, H.E. Stanley: Econom. Lett. 49: 453; 1995.
39. A. Czirok, R.N. Mantegna, S. Havlin, H.E. Stanley: Phys. Rev. E 52: 446; 1995.
40. G. Troll, P.B. Graben: Phys. Rev. E 57: 1347; 1998.
41. N. Vandewalle, M. Ausloos: In: M.M. Novak (Ed.) Fractals and Beyond (pp. 355). World Scientific, Singapore, 1998.

42. R.N. Mantegna, S.V. Buldyrev, A.L. Goldberger, S. Havlin, C.-K. Peng, M. Simmons, H.E. Stanley: Phys. Rev. Lett. 73: 3169; 1994.
43. N. Vandewalle, M. Ausloos: Physica A 268: 240; 1999.
44. M. Ausloos, K. Ivanova: Physica A 270: 526; 1999.
45. H. Dupuis: Trends-Tendances, Sept. 18, 97 ; n° 38, p. 26-27.
46. H. Dupuis: Trends-Tendances, Oct. 30, 97 ; n° 44, p. 11; editorial.
47. A. Johansen, D. Sornette: Int. J. Mod. Phys. C 10: 563; 1999.
48. D. Stauffer, R.B. Pandey: Int. J. Theor. Appl. Finance 3: 479; 2000.
49. A. Johansen, D. Sornette: Eur. J. Phys; B 17: 319; 2000.
50. M. Ausloos, K. Ivanova, N. Vandewalle: Crashes: symptoms, diagnoses and remedies. In: H. Takayasu (Ed.) Empirical sciences of financial fluctuations. Tokyo, Japan, Nov. 15-17, 2000 Proceedings. Springer Verlag, Berlin, 2002. http://xxx.lanl.gov/abs/cond-mat/0104127
51. J.M.R. Parrondo: http://seneca.fis.ucm.es/parr/BANCO/report/node7.html
52. A. Johansen, D. Sornette: http://netec.mcc.ac.uk/WoPEc/data/Papers/wpawuwpfi9907004.html
53. L. Laloux, M. Potters, R. Cont, J-P. Aguilar, J-Ph. Bouchaud: Eur. Letters 45 (1): 1; 1999.
54. D. Sornette, A. Johansen, J-Ph. Bouchaud: J. Physique (France) 6: 167; 1996.
55. S. Drozdz, F. Ruf, J. Speth, M. Wójcik: Eur. Phys. J. B 10: 589; 1999.
56. N. Vandewalle, M. Ausloos: Physica A 246: 454; 1997.
57. N. Vandewalle, M. Ausloos: Int. J. Comput. Anticipat. Syst. 1: 342; 1998.
58. M. Ausloos, K. Ivanova: Physica A 286: 353; 2000.
59. K. Ivanova, M. Ausloos: Eur. Phys. J. B 20: 537; 2001.
60. D.L. Turcotte: Fractals and Chaos in Geology and Geophysics. Cambridge University Press, Cambridge, 1997.
61. Z. Siwy, S. Mercik, K. Weron, M. Ausloos: Physica A 297: 79; 2001.
62. R.R. Rogers: Short Course in Cloud Physics. Pergamon Press, New York, 1976.
63. D. Andrews: An Introduction to Atmospheric Physics. Cambridge University Press, Cambridge, 2000.
64. K. Ivanova, M. Ausloos: Physica A 274: 349; 1999.
65. K. Ivanova, M. Ausloos, E.E. Clothiaux, T.P. Ackerman: Europhys. Lett. 52: 40; 2000.
66. K. Ivanova, T. Ackerman: Phys. Rev. E 59: 2778; 1999.
67. E. Harari: J. Appl. Phys. 49: 2478; 1978.
68. H. Takayasu: Phys. Rev. Lett. 54: 1099; 1985.
69. J.C. Lee, I.C. Chen, C. Hu: IEEE Trans. Electron Devices 35: 2268; 1988.
70. D.J. DiMaria, E. Cartier, D. Arnold: J. Appl. Phys. 73: 3367; 1993.
71. M. Depas, B. Vermeire, M.M. Heyns: J. Appl. Phys. 80: 382; 1996.
72. N. Vandewalle, M. Ausloos, M. Houssa, P.W. Mertens, M.M. Heyns: Appl. Phys. Lett. 74: 1579; 1999.
73. M. Houssa, T. Nigam, P.W. Mertens, M.M. Heyns: Solid State Phenom 249: 65-66; 1999.
74. R. Degraeve, G. Groeseneken, R. Bellens, M. Depas, H. Maes: IEDM Tech. Dig. IEEE, Piscataway, p. 863, 1995.
75. K. Sneppen: Phys. Rev. Lett. 69: 3539; 1992.