ELSEVIER

Physica A 300 (2001) 310–324

PHYSICA A

www.elsevier.com/locate/physa

Scaling features of texts, images and time series

Alexey N. Pavlov[a,*], Werner Ebeling[b], Lutz Molgedey[b], Amir R. Ziganshin[a], Vadim S. Anishchenko[a]

[a] *Nonlinear Dynamics Laboratory, Department of Physics, Saratov State University, Astrakhanskaya St. 83, 410026, Saratov, Russia*

[b] *Institute of Physics, Humboldt-University Berlin, Invalidenstr. 110, D-10099, Berlin, Germany*

Received 9 January 2001

Abstract

In the given paper, we consider the scaling features of long letter sequences like human writings, discretized images and discretized financial data. Using several approaches we show that the symbolic strings and time series being analyzed have a complex multiscale structure and demonstrate different scalings for large and small fluctuations. We discuss complex phenomena in the scaling behavior of partition functions in the case of high frequency DAX-future data.

PACS: 05.40.+j; 05.45.Tp; 02.50.Ey

Keywords: Scaling; Multifractality; Long-range correlations; Wavelet transform

1. Introduction

During the last few years there has been intense discussion about the scaling features of complex processes and natural sequences of different origin termed multifractal. Multiscale phenomena [1–8] occur in many fields of modern science. They take place in biosequences [9–11] and physiological signals [12,13], in fully developed turbulence [14–16] and random walks on random fractals [17,18], in cloud structure [19,20] and Brownian motion [21], in diffusion-limited aggregates [22] and chaotic attractors [23,24], etc.

Many time series in nature are nonstationary and inhomogeneous. Such signals have different scaling properties for different subsets of the data and require often a large number of characteristics to quantify the peculiarities of their complex structure [12].

* Corresponding author. Fax: +7-8452-514549.
E-mail address: pavlov@chaos.ssu.runnet.ru (A.N. Pavlov).

PII: S0378-4371(01)00341-7

At present, several approaches have been developed giving an opportunity to analyze multifractal objects (see, e.g. [25]). One of the commonly used approaches is the wavelet transform modulus maxima (WTMM) method [15,21,25–27] allowing to investigate, in practice, multiscale phenomena for any of the experimental signals including nonstationary data. The use of the wavelet analysis referred often as the "mathematical microscope" is now an attractive tool in time series processing. WTMM-method supposes the characterization of a time series by means of the function $D(h)$, where h is the Hölder exponent of some subset of the data which has the fractal dimension D [12]. The values of h quantify the local scaling and do not coincide for inhomogeneous processes. The function $D(h)$ called "singularity spectrum" together with the scaling exponents $\tau(q)$ of partition functions [21] constructed from the wavelet coefficients provide an informative characterization of the time series being investigated (and the same approach can also be applied to symbolic sequences).

In WTMM-method, the local singular behavior of the data is analyzed in terms of the wavelet transform coefficients. Another possible way to quantify the scaling properties of time series consists in the use of different variants of fluctuation analysis [13,28–40] and, in particular, in the use of the modified root mean square analysis of a random walk, termed detrended fluctuation analysis (DFA) [30–40]. This method is also appropriate to the processing of nonstationary data being an effective tool for detecting the presence of long-range correlations.

Although the scaling features can be studied using different algorithms, each of the methods provides us with an additional information about the complex structure of time series. Let us mention, for this purpose, the method of dynamic entropies [41–48]. The given approach allows to quantify how the value of conditional entropy h_n scales with the block length, i.e., how the average predictability of the state followed n former states scales with increasing n.

In the given paper, we discuss the scaling features of objects of different origin (human writings, images, financial data, etc.). The main purpose of our work is the simultaneous application and comparison of the above mentioned approaches for particular cases, which gives the possibility to describe different sides of the complex structure of the objects being investigated. The paper is organized as follows: Section 2 briefly describes the methods being used for analysis. The results of processing literary texts and images are the subject of Section 3. In Section 4, we focus on the features of the scaling behavior of financial data.

2. Methods

2.1. WTMM-method

A wavelet-based multifractal formalism originally proposed in Ref. [26] can be considered as a generalization of the well-known box-counting methods to fractal signals using wavelets instead of boxes. Since the analyzing wavelets represent well localized functions, they are appropriate to the processing of inhomogeneous time series. We shall limit ourselves to describe very briefly some details of the WTMM-algorithm. A complete description can be found in Ref. [21].

At the first step of this method, the wavelet transform of a function $f(x)$ is performed:

$$T_\psi[f](x_0,a)=\frac{1}{a}\int_{-\infty}^{\infty} f(x)\psi_m^*\left(\frac{x-x_0}{a}\right)\mathrm{d}x\,, \tag{1}$$

where $f(x_0)$ is a distribution at a point x_0 (for symbolic sequences one often uses random walk displacement as $f(x)$), ψ is the wavelet mother function chosen e.g. as follows

$$\psi^{(m)}=(-1)^m\frac{\partial^m}{\partial x^m}\left[\exp\left(-\frac{x^2}{2}\right)\right]\,, \tag{2}$$

and a is the scale parameter. As it was proved in Ref. [21], a local singular behavior of $f(x)$ around $x=x_0$ can be characterized by the Hölder exponent $h(x_0)$

$$T_\psi[f](x_0,a)\sim a^{h(x_0)},\quad a\to 0^+ \tag{3}$$

which quantifies the "strength" of a singularity of $f(x)$.

The second step of WTMM-algorithm assumes the building up of a partition function $Z(q,a)$ being the sum of the qth powers of the local maxima of $|T_\psi[f](b,a)|$ at the scale a. Following [21],

$$Z(q,a)\sim a^{\tau(q)}\,, \tag{4}$$

where each value of $\tau(q)$ defined for some q is the scaling exponent characterizing the power-law behavior of $Z(q,a)$. The variations of powers allow to obtain a linear function $\tau(q)$ and constant value $h=\mathrm{d}\tau/\mathrm{d}q$ in the case of monofractal objects and a nonlinear dependence $\tau(q)$ with a large number of Hölder exponents for multifractals. Since the analyzed time series has different scaling properties for different subsets in the last case, one uses the singularity spectrum $D(h)$ to characterize statistical properties of these subsets. This function denotes that some subset of points x, for which $h(x)=h$, has the Hausdorff dimension D [21]. The relationship between the main quantities of the WTMM-method is defined by the Legendre transform:

$$D(h)=qh-\tau(q)\,. \tag{5}$$

2.2. Fluctuation analysis of sequences

The scaling features of time series are studied often by means of different variants of fluctuation analysis [13,28–37]. Peng et al. [28] proposed a useful approach allowing to characterize the scale invariant long-range correlations in symbolic sequences, for example, in DNA. According to this approach, a nucleotide chain is transformed first into a binary sequence $u(i)$, where $u(i) = +1$ if pyrimidine occurs at position i and $u(i) = -1$ if purine occurs at i. After this transformation, the running sum called *fractal landscape* or *DNA walk* [28], is calculated: $y(l) = \sum_{i=1}^{l} u(i)$. An important statistical quantity characterizing this walk is the root mean square fluctuation $F(l)$ about the average of the displacement defined as proposed in Ref. [28]. From the viewpoint of detection of the presence of long-range correlation, the behavior of $F(l)$ for $l \gg 1$ is the focus of interest.

An analysis of fluctuations is not restricted by the consideration of binary sequences only. Let us discuss the generalization of the technique [28] on the case of compositions (for example, texts consisting of many different symbols). This generalization has been considered in Ref. [46] and was based on the method [28] using the invariant representation proposed by Voss [49]. The given procedure consists in the following: Instead of the original string consisting of λ different symbols, we generate λ strings on the binary alphabet (0,1) ($\lambda = 32$ for texts). In the first string, we place a "1" on all positions where there is an "a" in the original text and a "0" on all other positions. The same procedure is also carried out for the remaining symbols. Then we generate random walk processes corresponding to these strings moving one step upwards for any "1" and remaining on the same level for any "0". The resulting move over a distance l is called $y(k,l)$, where k denotes the symbol. Then by defining a λ-dimensional vector space considering $y(k,l)$ as the component k of the state vector at the (discrete) "time" l we can map the text to a trajectory. Any position in the text corresponds to a random vector $\mathbf{y}(l)$ in the state space. The mean-square displacement for symbol k is determined as

$$F^2(k,l) = \langle y^2(k,l) \rangle - (\langle y(k,l) \rangle)^2 , \tag{6}$$

where the brackets denote averaging over all the initial positions. It is expected that $F(k,l)$ follows a power law

$$F(k,l) \sim l^{\alpha(k)} , \tag{7}$$

where $\alpha(k)$ is the scaling exponent for symbol k. Here, $\alpha(k) = 0.5$ for a normal random walk and $\alpha(k) \neq 0.5$ at the presence of correlations. (Alternatively, we can consider a power-law dependence on l for $F^2(k,l)$ [46].) Besides the individual exponents for the letters, it is possible to estimate even an averaged exponent α for the state space.

Instead of speaking about random walks we can reformulate the given problem in another language, in terms of the statistical-mechanical theory of fluctuations of particle numbers and compositions. Let us consider a text with the total length L. Then the total number of letters is $N = L$ and the density is equal to "1". However, the density of different symbols may fluctuate along the string. In Ref. [46], for example, the fluctuating local density of blanks in "Moby Dick" has been considered and the existence of rather long wave fluctuations has been pointed out. The research [46] considered the averaged over windows of length 4000 local frequency of the blanks (and other letters) in the text of "Moby Dick" in dependence on the position along the text. The original symbolic string shows a large-scale structure extending over many windows. This reflects the fact that in some parts of the texts we have many short words, e.g. in conversations (yielding the peaks of the space frequency), and in others we have more long words, e.g. in descriptions and in philosophical considerations (yielding the minima of the space frequency). The shuffled text shows a much weaker nonuniformity of symbols; the lower the shuffling level, the larger is the uniformity. More uniformity means less fluctuations and more similarity to a Bernoulli sequence. For the case of DNA-sequences no analogies of pages, chapters etc. are known. Nevertheless, the reaction on shuffling is similar to those of the texts.

In order to quantify these findings let us define the number of letters of the kind k inside a substring of length l by $N(k,l)$. In the limit $l \to \infty$ we get the average density $n(k) = \lim_{l\to\infty} N(k,l)/l$. Since we have λ different symbols, we obtain in this way a λ-dimensional composition space. Let us now consider the fluctuations of $N(k,l)$ as a function of l. We expect that $N(k,l)$ fluctuates around the mean value $\langle N(k)\rangle = ln(k)$. Further, we assume that the mean square fluctuations scale with certain power of the mean (particle) numbers

$$\langle [N(k,l) - \langle N(k)\rangle]^2\rangle = \mathrm{const}\langle N(k)\rangle^{2\alpha_k} . \tag{8}$$

The exponent α_k is called the characteristic root mean square fluctuation exponent. In an analog we consider the sum of the mean square fluctuations defining an exponent α by

$$\sqrt{\sum \langle [N(k,l) - \langle N(k)\rangle]^2\rangle} = \mathrm{const}\, l^{\alpha} . \tag{9}$$

One can easily observe that the above definitions give the same numbers for the α-coefficient as the mapping to a random walk in a λ-dimensional space. In this way we have demonstrated the close connection to the concept of the fluctuation theory of particle numbers.

2.3. DFA-algorithm

The approach [28] and its generalizations cannot be applied reliably to inhomogeneous sequences (or time series) since the presence of nonstationarity affects the results of an analysis of statistical fluctuations. To take into account such a problem, a new method, termed detrended fluctuation analysis, has been proposed [30–37]. Like wavelets, DFA-technique gives the possibility to study nonstationary data. At present, it has been successfully applied to characterize the scaling properties in the processes of different origin, e.g. in physiological signals [13,30,39], DNA sequences [31–35], meteorological time series [35,36], etc. [37,38,40].

The main idea of the method consists in the following. The time series $a(i)$ being investigated is first integrated to obtain the dependence $y(k) = \sum_{i=1}^{k} [a(i) - a_{\mathrm{ave}}]$, where a_{ave} is the mean value and is next divided into boxes of length n. The local trend $y_n(k)$

estimated via least-squares fitting is removed from each box (using as a rule a linear dependence although polynomials may be also appropriate). After such procedures the root-mean-square fluctuations of the analyzed time series have the form [30]:

$$F(n) = \sqrt{\frac{1}{N}\sum_{k=1}^{N}[y(k) - y_n(k)]^2} \tag{10}$$

where N is the total number of points.

At the presence of scaling in time series, the dependence $\lg(F)$ vs. $\lg(n)$ has a linear segment with the slope α, called scaling exponent for fluctuations $F(n)$. Following [50–52], there exists a relationship between α and the scaling exponents for power spectrum and correlation function. In the absence of local trend both approaches [28,30] lead to the same value of α.

Unlike WTMM-method, the discussed technique allows to obtain usually only one exponent which is important for the analysis of monofractal objects. As it was already mentioned, multifractal signals require a large number of quantities to characterize their complex structure since the scaling properties are different for large (positive q in Eq. (4)) and for small (negative q) fluctuations.

Although DFA-algorithm is simpler and measures a single dependence $\lg(F)$ vs. $\lg(n)$, its local slopes may not coincide for time series with multiscale structure. This allows us to quantify the features of the power-law behavior for short and for long scales [31]. That is why we can consider DFA-method as a potential tool in the study of multiscale phenomena as well. Note, that both the above discussed methods (WTMM and DFA) lead to the values $h \approx \alpha \approx 0.5$ for uncorrelated processes (other values of these quantities mark different types of power-law correlations) [12,30].

3. Applications to human writings and images

Starting with Shannon's paper [53] human writings are a subject of many studies. One of the attractive questions being discussed in the last few years is the origin of long-range correlations in texts. In particular, the paper [54] considers the problem whether the scaling behavior is based on correlations below the sentence level or not. In Ref. [55] it was claimed that the origin of the long-range correlations in a text is in relation with the ideas expressed by an author. The work [46] has showed that the long correlations reflected in scaling features for power spectrum or other measures are connected with large-scale structures beyond the sentence level.

The presence of long-range order in human writings can be quantified using different approaches. Aiming to discuss the features of the scaling behavior for texts, we shall mention some results of recent studies for the book "Moby Dick" by Melville ($N \approx 1,170,000$ letters). As it was shown in Ref. [46], at least in a reasonable approximation the scaling of the entropy against the word length is given for large n by root laws of the type

$$h_n = \frac{c_1}{\sqrt{n}} + c_2 \,. \tag{11}$$

A power-law decay of h_n allows to state the effect of a rather long memory tail.

Calculation of the power spectrum $S(f)$ for "Moby Dick" performed in Ref. [47] has shown that the given dependence reminds in double-logarithmic plot ($\log(S)$ vs. $\log(f)$), a piecewise linear behavior (i.e., has different scaling properties for different ranges of frequencies) which is typical for multifractal structures.

Let us now consider this question on the basis of other approaches. Since the application of different numerical recipes to symbolic strings requires their initial transformation into series of numbers, first, we need to discuss how such a transformation can be realized. One often uses random walk models for this purpose. The given representation of symbolic strings is rather typical for biosequences. In Ref. [46], random walks have been successfully applied to study the effects of long correlations in human writings (e.g. in "Moby Dick"). The estimation of scaling exponent performed in this paper has led to the value $\alpha \approx 0.62$ for root mean square displacement (or ≈ 1.24 for mean square fluctuations).

Another way of transforming a symbolic string into a series of numbers (accepted in neurodynamics or cardiology) consists in the analysis of so called "point processes" (i.e., of the processes where the information carriers are time intervals between some "events" [56]). We can consider, for example, the appearance of letter "a" in text (any other letter or combination of symbols) as such an "event". A series of integer numbers being the intervals between successively appearing "events" can be studied using standard methods of time series analysis. We have investigated "point processes" corresponding to different letters in "Moby Dick" and other texts (e.g. tragedies by Shakespeare, bible, etc.) in order to analyze the scaling features in literary English. Since the obtained results are qualitatively the same, we shall limit ourselves to describe the typical scaling behavior using Melville's book.

As it was mentioned in the previous section, the DFA-algorithm is applied as a rule to extract a single scaling exponent α from the experimental data. But, in the presence of a multiscale structure, the local slopes of $\log(F)$ vs. $\log(n)$ do not coincide for large and small scales (Fig. 1). We can see also from this figure that a shuffling of original text on the sentence level destroys long correlations and keeps only the short-range ones leading to the multiscale structure for shuffled data as well. The shuffling has been performed as follows: First, we generated a series of normally distributed random numbers $r(j), j = 1, \ldots, M$, where M is the number of sentences. Next, the series $r(j)$ has been rewritten in the order of increasing values, so we have obtained the random sequence of indexes. The sentences in the original text have been shuffled according to this sequence and we have used such a shuffling procedure further for words, letters, etc.

The pictures similar to Fig. 1 have been obtained after averaging the results of processing different "event" series (using DFA-method and taking a large number of

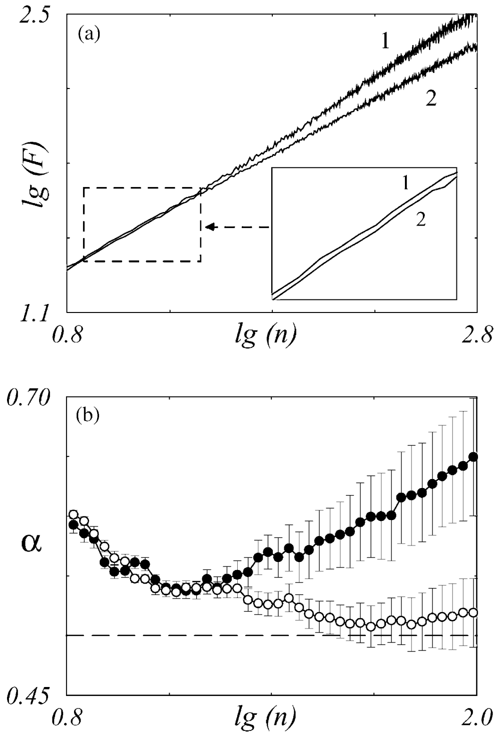

Fig. 1. DFA-analysis of the book "Moby Dick" by Melville: (a) The dependence $\lg(F)$ vs. $\lg(n)$ for the original text (1) and after shuffling on the sentence level (2). "Event" series has been recorded as intervals between the letters combination "th"; (b) Local scaling exponents for the original (black circles) and for the shuffled text (white circles).

letter combinations). The presence of multiscaling in literary English can be tested also by WTMM-algorithm. Fig. 2 shows the singularity spectra obtained for point processes (recorded as intervals between the letters combination "th") in the original text and after shuffling on the word level. As $\psi^{(m)}$ we have chosen WAVE-wavelet ($m = 1$) and MHAT-wavelet ($m = 2$). We can see from the given pictures that the original text can be really considered in the frame of multifractal concept since the dependence $D(h)$ does not consist of a single point. The shuffling on the word level destroys long correlations (the values of Hölder exponents are close to 0.5 for positive q), however, short-range correlations below the word level are retained. As a consequence, $h > 0.5$ for negative q.

So, we can see that human writings are an example of information carriers with multiscale structures. Processing different texts (or different "event" series) change the values of Hölder exponents and scaling exponents for fluctuations $F(n)$, but the results will be qualitatively similar. Note that in the analogy with human writings we can study different images using the same approaches. For example, Fig. 3 shows the singularity spectrum obtained for the photo of Saratov conservatory (black circles). White circles

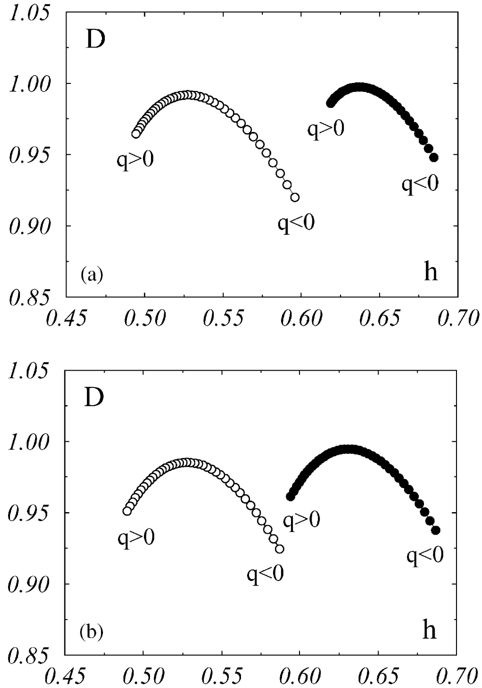

Fig. 2. Multifractal analysis of the book "Moby Dick": (a) The dependence $D(h)$ for "event" series recorded as intervals between the letters combination "th" in the original text (black circles) and in the text shuffled on the word level (white circles). As $\psi^{(m)}$ we have chosen WAVE-wavelet ($m = 1$); (b) The same as in (a) for MHAT-wavelet ($m = 2$).

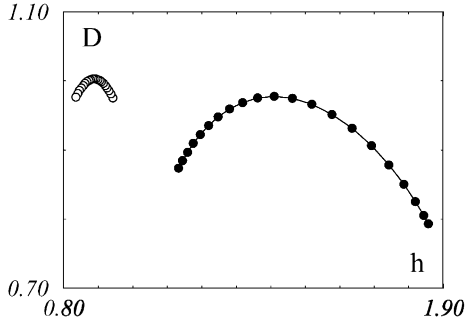

Fig. 3. Multifractal analysis of images: The singularity spectrum (black circles) for the photo of Saratov conservatory. This photo was transformed into series of numbers (from 0 to 255) being tints of the color of black-and-white image. Unlike 2D-method [19,20], we have restricted ourselves for the simplest case (1D-method). For comparison, the results for a fragment of Leonardo picture "Virgin of the Rocks" are presented (white circles).

in Fig. 3 mark the corresponding results for a fragment of the well-known picture "Virgin of the Rocks" by Leonardo showing that the presence of multiscale structure can be quantified for paintings as well.

4. Applications to financial data

The peculiarity of statistical methods is that they find successful applications at the analysis of objects of different origin. Using the same approaches we can study literary texts, biosequences, physical processes, mathematical functions, etc. In this section, we would like to consider a rather specific type of experimental data demonstrating complex scaling behavior, namely, financial time series.

Our experiments have been performed for daily stock index data: Dow Jones (DJ) 1900–2000 S_t (27,044 trading days) and DAX-future data recorded each 10 s during 60 trading days. To remove the trend from the time series (which grows exponentially for DJ-data) one often uses logarithmic price changes

$$x_t = \ln(S_t) - \ln(S_{t-1}) , \tag{12}$$

where S_{t-1} is the last measurement of the stock index, S_t is the new value. Next step for entropy analysis consists in transformation of real numbers x_t into a symbolic string A_t using alphabet of length λ. In Ref. [57], this transformation has been performed as follows: $\lambda = 3$ and $A_t = 0$; $x_t < -0.0025$ (strong decrease in the stock value), $A_t = 2$; $x_t > 0.0034$ (strong increase), $A_t = 1$ (intermediate).

A direct application of the entropy concept to DJ-data in order to analyze the scaling features is difficult in connection with a sufficiently short length of time series, which limits the abilities in considering large blocks due to the increasing of statistical errors with n [48].

At the same time, this concept can quantify the local predictability in time series. As it was shown in Ref. [57], although the averaged predictability is close to zero, there exist certain patterns of stock movements $A_1 \ldots A_n$ behind which the local predictability reaches 8%. This is a notable value for the stock market, which is usually purely random. Moreover, higher local predictabilities coincide with larger levels of significance [57]. Fig. 4a shows that in a time window averaged local uncertainty has certain periods of higher averaged predictabilities which were relatively small in the last few decades (after the beginning of computerized trading).

Other approaches being discussed above are less sensitive to the length of time series. Both, DFA- and WTMM-methods show that the DJ stock index data demonstrate rather uncorrelated behavior (Hölder exponents h and scaling exponents for fluctuations α are close to 0.5). We can therefore conclude that processing financial data for short time scales may be only potentially interesting. Note also that the properties of DJ-data are clearly different for the first few decades of 20th century and for the last few decades, which can be illustrated by WTMM-approach: We see the presence of correlations in

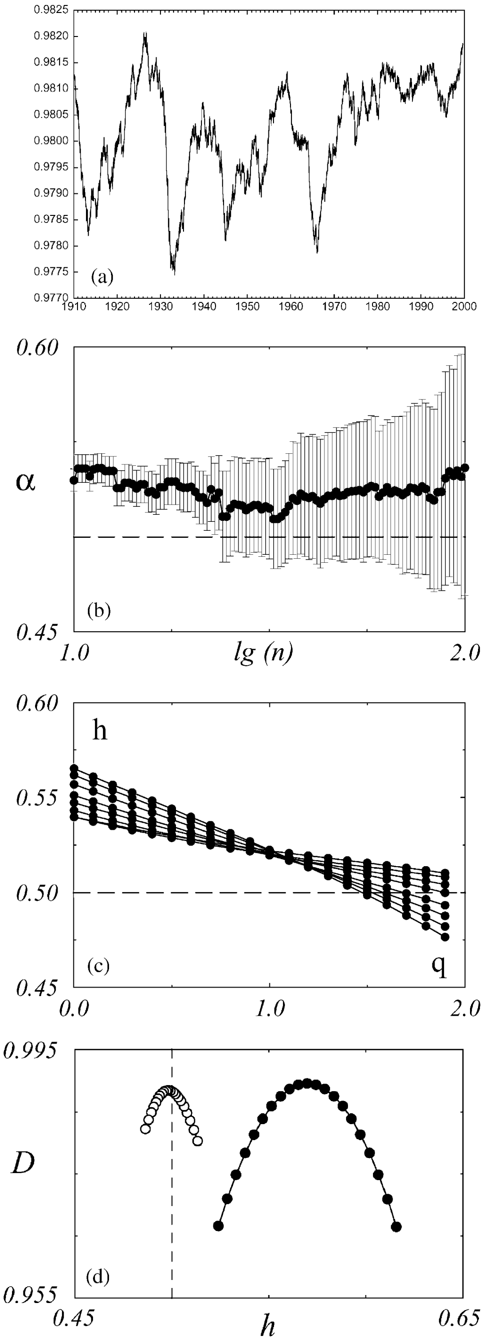

Fig. 4. Analysis of DJ-data. (a) Entropy approach: moving exponential average of the local uncertainty h_5 with an half-life period of 5 years for daily data 1900–2000; (b) The results of DFA-algorithm: local slopes of the dependencies $\lg(F)$ vs. $\lg(n)$; (c) Multifractal analysis: the values of Hölder exponents vs. q at the variation of the range of scales used for fitting; (d) Singularity spectrum for the first 8000 days of DJ-data (black circles) and for the last 8000 days (white circles).

time series for the first 8000 days (Fig. 4d, black circles) and the behavior which is close to an uncorrelated one during the last 8000 days (Fig. 4d, white circles).

Now we turn to DAX-future data. As it was already mentioned, the original tick-by-tick data were resampled to equidistant times of an interval of 10 s during 60 trading days. Thus, we have about 2500 measurements per day. This is small enough to consider scaling properties for dynamic entropy [58]. Further, we shall discuss the results obtained by means of other algorithms. Again, our analysis has shown that the behavior of financial data being discussed is close to an uncorrelated one ($h \approx \alpha \approx 0.5$). For some days, Hölder exponents and scaling exponents for fluctuations are clearly different from 0.5. However, this is connected perhaps with a nonstationarity of x_t during these days (something similar to the switching process) while the corresponding results of the given computations need further analysis.

Fig. 5 presents the results obtained for typical days when the analyzed time series can be considered as stationary. The scaling exponent α is close to 0.5 in a width range of scales (Fig. 5a).

WTMM-algorithm maybe a more appropriate tool to consider short scales. We can see from Fig. 5b that DAX-future data show nontrivial scaling behavior for the partition function constructed from wavelet coefficients. The dependence $\ln(Z(q,a))$ vs. $\ln(a)$ is not linear (as it is expected usually) and we have piecewise linear segments for negative values of q reflecting the features of the scaling behavior for small fluctuations. Fitting by linear functions in a wide range of scales can lead this case to the complex dependence of singularity spectrum $D(h)$ (Fig. 5c) being typical for multifractal structures. (The problems which arise in some cases as to how to interpret the results obtained using multifractal concept are discussed e.g. in Ref. [59].) Fittings performed separately for two linear segments of the dependence $\ln(Z)$ vs. $\ln(a)$ lead to the Hölder exponents $h \approx 0.5$ for large scales and $h \approx 0$ for $\ln(a) \leqslant 1$ (where the scale a is taken in the number of measurements, i.e., in the number of points with the time interval 10 s). That is why we can conclude that the chance of prediction for the analyzed data exists perhaps only for very short time intervals (up to 20–30 s). When considering other days we can obtain slightly different results for the predictability horizon showing however the complex scaling behavior for wavelet coefficients (and as a consequence, the complex behavior for partition functions).

5. Conclusions

In this work, several algorithms have been used to quantify the scaling features in the long sequences like human writings and financial data. The simultaneous application of different approaches allowed us to describe the details of the complex structure of the analyzed symbolic sequences and time series. It was shown that literary English can be studied in the frames of multifractal concept and the scaling properties of the given symbolic strings do not coincide for short and for long scales demonstrating the inhomogeneity of texts on different levels. This was testified by both, DFA- and

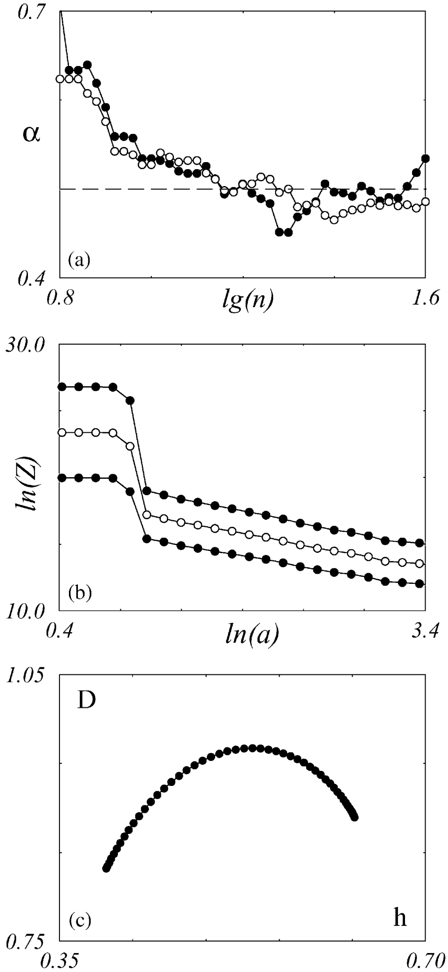

Fig. 5. Analysis of DAX-future data (typical results): (a) Local scaling exponents for logarithmic price changes x_t (black circles) and for shuffled time series (white circles); (b) The behavior of partition function for negative values of q; (c) The singularity spectrum obtained when fitting is performed over a full-range of scales.

WTMM-algorithms. The structural origin of the observed phenomena may be understood on the basis of shuffling experiments. For example, the shuffling of sentences destroys long correlations beyond the sentence level and we expect to obtain different scaling behaviors for large and small fluctuations. Less well understood is the origin of the observed inhomogeneities, and this question needs further research.

We have also chosen an object showing other types of scaling behavior, namely financial time series and are certain that the given data have non-trivial scaling properties for partition function allowing to obtain different results for singularity spectrum when considering different ranges of a. Although the shape of singularity spectrum (Fig. 5c) looks like a continuous function, we suppose that it may be connected with the shortcomings of standard multifractal analysis methods when the object to which they are applied is not multifractal. It is known [59], that the function $D(h)$ may have spurious points representing top envelope of true spectrum. From the piecewise linear structure of the partition function we expect to obtain different values of $\tau(q)$ for large and short scales. Fittings performed separately for each linear segment of $\ln(Z)$ vs. $\ln(a)$ verify that the scaling properties of DAX-future data are characterized more probably by two quantities ($h \approx 0.5$ and $h \approx 0$) instead of a continuous dependence $D(h)$.

The above mentioned gives the possibility to conclude that the objects being considered have a complex structure indeed and this structure is perhaps more complex than it is often supposed.

Acknowledgements

The authors thank A.B. Neiman for fruitful discussions. The research described in this publication was made possible in part for the authors from Saratov State University by Award No. REC-006 of the U.S. Civilian Research & Development Foundation for the Independent States of the Former Soviet Union (CRDF). Further, we thank the Deutsche Forschungsgemeinschaft (DFG) for the support within the Sonderforschungsbereich 555.

References

[1] B.B. Mandelbrot, Fractals and Multifractals: Noise, Turbulence and Galazies, Selecta, Vol. 1, Springer, New York, 1989.
[2] B.B. Mandelbrot, Multifractals anf 1/F Noise: Wild Self-Affinity in Physics, Springer, New York, 1999.
[3] A. Bunde, S. Havlin (Eds.), Fractals in Science, Springer, Berlin, 1994.
[4] T. Vicsek, Fractal Growth Phenomena, 2nd Edition, World Scientific, Singapore, 1993.
[5] T. Tel, Z. Naturforsh. 43a (1988) 1154.
[6] T.C. Halsey, M.H. Jensen, L.P. Kadanoff, I. Procaccia, B.I. Shraiman, Phys. Rev. A 33 (1986) 1141.
[7] U. Frish, G. Parisi, in: M. Ghil, R. Benzi, G. Parisi (Eds.), Turbulence and Predictability in Geophysical Fluid Dynamics and Climate Dynamics, North-Holland, Amsterdam, 1985, p. 71.
[8] M.H. Jensen, G. Paladin, A. Vulpiani, Phys. Rev. Lett. 67 (1991) 208.
[9] B.J. Strait, T.G. Dewey, Phys. Rev. E 52 (1995) 6588.
[10] C.L. Berthelsen, J.A. Glazier, S. Raghavachari, Phys. Rev. E 49 (1994) 1860.
[11] J.A. Glazier, S. Raghavachari, C.L. Berthelsen, M.H. Skolnick, Phys. Rev. E 51 (1995) 2665.
[12] P.Ch. Ivanov, L.A.N. Amaral, A.L. Goldberger, S. Havlin, M.G. Rosenblum, Z.R. Struzik, H.E. Stanley, Nature 399 (1999) 461.
[13] H.E. Stanley, L.A.N. Amaral, A.L. Goldberger, S. Havlin, P.Ch. Ivanov, C.-K. Peng, Physica A 270 (1999) 309.
[14] R. Benzi, L. Biferale, G. Paladin, A. Vulpiani, M. Vergassola, Phys. Rev. Lett. 67 (1991) 2299.
[15] J.F. Muzy, E. Bacry, A. Arneodo, Phys. Rev. Lett. 67 (1991) 3515.
[16] S.I. Vainshtein, K.R. Sreenivasan, R.T. Pierrehumbert, V. Kashyap, A. Juneja, Phys. Rev. E 50 (1994) 1823.
[17] E. Eisenberg, A. Bunde, S. Havlin, H.E. Roman, Phys. Rev. E 47 (1993) 2333.
[18] J. Dräger, A. Bunde, Phys. Rev. E 54 (1996) 4596.
[19] J. Arrault, A. Arneodo, A. Davis, A. Marshak, Phys. Rev. Lett. 79 (1997) 75.

[20] A. Arneodo, N. Decoster, S.G. Roux, Phys. Rev. Lett. 83 (1999) 1255.
[21] J.F. Muzy, E. Bacry, A. Arneodo, Int. J. Bifurcation Chaos 4 (1994) 245.
[22] H.G.E. Hentschel, Phys. Rev. E 50 (1994) 243.
[23] V. Afraimovich, G.M. Zaslavsky, Phys. Rev. E 55 (1997) 5418.
[24] K.O. Wiklund, J.N. Elgin, Phys. Rev. E 54 (1996) 1111.
[25] J.F. Muzy, E. Bacry, A. Arneodo, Phys. Rev. E 47 (1993) 875.
[26] A. Arneodo, G. Grasseau, M. Holschneider, Phys. Rev. Lett. 61 (1988) 2281.
[27] A. Arneodo, Y.D. Aubenton-Carafa, B. Audit, E. Bacry, J.F. Muzy, C. Thermes, Physica A 249 (1998) 439.
[28] C.-K. Peng, S.V. Buldyrev, A.L. Goldberger, S. Havlin, F. Sciortino, M. Simons, H.E. Stanley, Nature 356 (1992) 168.
[29] E. Koscielny-Bunde, A. Bunde, S. Havlin, H.E. Roman, Y. Goldreich, H.-J. Schellnhuber, Phys. Rev. Lett. 81 (1998) 729.
[30] C.-K. Peng, S. Havlin, H.E. Stanley, A.L. Goldberger, CHAOS 5 (1995) 82.
[31] S.V. Buldyrev, A.L. Goldberger, S. Havlin, C.-K. Peng, M. Simons, H.E. Stanley, Phys. Rev. E 47 (1993) 4514.
[32] C.-K. Peng, S.V. Buldyrev, S. Havlin, M. Simons, H.E. Stanley, A.L. Goldberger, Phys. Rev. E 49 (1994) 1685.
[33] 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 (1995) 5084.
[34] H.E. Stanley, S.V. Buldyrev, A.L. Goldberger, S. Havlin, C.-K. Peng, M. Simons, Physica A 273 (1999) 1.
[35] S. Havlin, S.V. Buldyrev, A. Bunde, A.L. Goldberger, P.Ch. Ivanov, C.-K. Peng, H.E. Stanley, Physica A 273 (1999) 46.
[36] P. Talkner, R.O. Weber, Phys. Rev. E 62 (2000) 150.
[37] R.N. Mantegna, Z. Palágyi, H.E. Stanley, Physica A 274 (1999) 216.
[38] K. Ivanova, M. Ausloos, Physica A 274 (1999) 349.
[39] P.-A. Absil, R. Sepulchre, A. Bilde, P. Gérard, Physica A 272 (1999) 235.
[40] M. Ausloos, N. Vandewalle, Ph. Boveroux, A. Minguet, K. Ivanova, Physica A 274 (1999) 229.
[41] W. Ebeling, G. Nicolis, Europhys. Lett. 14 (1991) 191.
[42] W. Ebeling, G. Nicolis, Chaos, Solitons and Fractals 2 (1992) 635.
[43] W. Ebeling, T. Pöschel, Europhys. Lett. 26 (1994) 241.
[44] J. Freund, W. Ebeling, K. Rateitschak, Phys. Rev. E 54 (1996) 5561.
[45] K. Rateitschak, W. Ebeling, J. Freund, Europhys. Lett. 35 (1996) 401.
[46] W. Ebeling, A. Neiman, Physica A 215 (1995) 233.
[47] W. Ebeling, A. Neiman, T. Pöschel, in: M. Suzuki, N. Kawashima (Eds.), Coherent Approaches to Fluctuations, Proceedings of the Hayashibara Forum 95, World Scientific, Singapore, 1995, p. 59.
[48] W. Ebeling, Physica D 109 (1997) 42.
[49] R.F. Voss, Phys. Rev. Lett. 68 (1992) 3805.
[50] C.-K. Peng, J. Mietus, J.M. Hausdorff, S. Havlin, H.E. Stanley, A.L. Goldberger, Phys. Rev. Lett. 70 (1993) 1343.
[51] C.-K. Peng, S.V. Buldyrev, A.L. Goldberger, S. Havlin, M. Simons, H.E. Stanley, Phys. Rev. E 47 (1993) 3730.
[52] J.M. Hausdorff, C.-K. Peng, Phys. Rev. E 54 (1996) 2154.
[53] C. Shannon, Bell Syst. Tech. 30 (1951) 50.
[54] W. Li, K. Kaneko, Europhys. Lett. 17 (1992) 655.
[55] M. Amit, Y. Shmerler, E. Eisenberg, M. Abraham, N. Shnerb, Int. J. Bifurcation Chaos 2 (1994) 7.
[56] T. Sauer, Phys. Rev. Lett. 72 (1994) 3911.
[57] L. Molgedey, W. Ebeling, Eur. Phys. J. B 15 (2000) 733.
[58] L. Molgedey, W. Ebeling, Physica A 287 (2000) 420.
[59] D. Veneziano, G.E. Moglen, R.L. Bras, Phys. Rev. E 52 (1995) 1387.