Journal of Statistical Mechanics: Theory and Experiment
An IOP and SISSA journal

Multifractal detrended fluctuation analysis of sunspot time series

M Sadegh Movahed[1,2,3], G R Jafari[2,4], F Ghasemi[2], Sohrab Rahvar[1,2] and M Reza Rahimi Tabar[1,5]

[1] Department of Physics, Sharif University of Technology, PO Box 11365-9161, Tehran, Iran
[2] Institute for Studies in Theoretical Physics and Mathematics, PO Box 19395-5531, Tehran, Iran
[3] Iran Space Agency, PO Box 199799-4313, Tehran, Iran
[4] Department of Physics, Shahid Beheshti University, Evin, Tehran 19839, Iran
[5] CNRS UMR 6529, Observatoire de la Côte d'Azur, BP 4229, 06304 Nice Cedex 4, France
E-mail: m.s.movahed@mehr.sharif.edu, gjafari@gmail.com, f_ghasemi@mehr.sharif.edu, rahvar@sharif.edu and rahimitabar@sharif.edu

Received 6 December 2005
Accepted 20 January 2006
Published 9 February 2006

Online at stacks.iop.org/JSTAT/2006/P02003
doi:10.1088/1742-5468/2006/02/P02003

Abstract. We use multifractal detrended fluctuation analysis (MF-DFA), to study sunspot number fluctuations. The result of the MF-DFA shows that there are three crossover timescales in the fluctuation function. We discuss how the existence of the crossover timescales is related to a sinusoidal trend. Using Fourier detrended fluctuation analysis, the sinusoidal trend is eliminated. The Hurst exponent of the time series without the sinusoidal trend is 0.12 ± 0.01. Also we find that these fluctuations have multifractal nature. Comparing the MF-DFA results for the remaining data set to those for shuffled and surrogate series, we conclude that its multifractal nature is almost entirely due to long range correlations.

Keywords: new applications of statistical mechanics

J. Stat. Mech. (2006) P02003

 1742-5468/06/P02003+17$30.00

Contents

1. Introduction

The important feature of the sun's outer regions is the existence of a reasonably strong magnetic field. To the lowest order of approximation, the sun's magnetic field is dipolar in character and is axisymmetric. The strength of the field at a typical point on the solar surface is approximately a few gauss. There is, however significant variation in this value and there are localized regions (called *sunspots*) in which the field can be much higher [1]. Because of the symmetry of the twisted magnetic lines at the origin of sunspots, they are generally seen in pairs or in groups of pairs on both sides of the solar equator. As the sunspot cycle progresses, spots appear closer to the sun's equator giving rise to the so-called 'butterfly diagram' in the time latitude distribution [2]. The twisted magnetic fields above sunspots are sites where solar flares are observed. It has been found that chromospheric flares show a very close statistical relationship with sunspots [1]. The number of sunspots is continuously changing in time in a random fashion and constitutes a typically random time series. Figure 1 shows the monthly measured number of sunspots in terms of time. The data belong to a data set collected by the Sunspot Index Data Center (SIDC) from 1749 up to the present [3].

Recently, the statistical properties of sun activity have been investigated by some methods in chaos theory [4] and multifractal analysis [5, 6]. The periodic occurrences of hemispheric sunspots have been analysed with respect to the changes in time using wavelets. The north–south asymmetries concerning solar activity and rotational behaviour have been investigated by using wavelets and autocorrelation functions [7]. Cross-correlation functions between monthly mean sunspot areas and sunspot numbers have been determined in some papers [8]. The evidence for the existence of 'active longitudes' on the sun is given by using the autocorrelation function of daily sunspot numbers [8, 9]. See also [10, 11] as regards the relation between sunspot number fluctuation and number of flares, their evolution steps, i.e. duration, rise times, decay times, event asymmetries.

doi:10.1088/1742-5468/2006/02/P02003

J. Stat. Mech. (2006) P02003

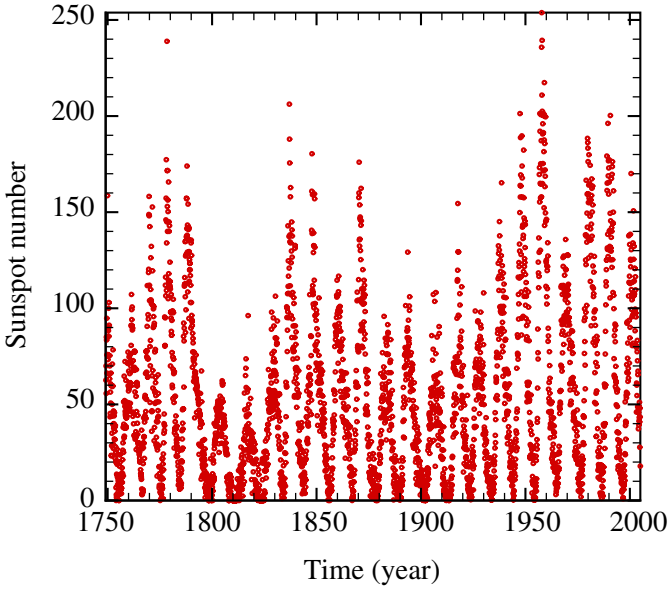

Figure 1. Observed spot numbers as a function of time.

In this paper we would like to characterize the complex behaviour of sunspot time series through the computation of the signal parameters—scaling exponents—which quantify the correlation exponents and multifractality of the signal. As shown in figure 1, the sunspot time series has a sinusoidal trend, with a frequency equal to the well-known cycle of sun activity, approximately 11 years. Because of the non-stationary nature of sunspot time series, and due to the finiteness of the available data sample, we should apply some methods which are insensitive to non-stationarities, like trends.

To eliminate the effect of the sinusoidal trend, we apply Fourier detrended fluctuation analysis (F-DFA) [12, 13]. After elimination of the trend we use multifractal detrended fluctuation analysis (MF-DFA) to analyse the data set. The MF-DFA methods are the modified version of detrended fluctuation analysis (DFA) used to detect multifractal properties of time series. The detrended fluctuation analysis (DFA) method introduced by Peng *et al* [14] has became a widely used technique for the determination of (mono)fractal scaling properties and the detection of long range correlations in noisy, non-stationary time series [14]–[18]. It has been applied successfully in diverse fields such as DNA sequences [14, 19], heart rate dynamics [20]–[22], neuron spiking [23], human gait [24], long time weather records [25], cloud structure [26], geology [27], ethnology [28], economical time series [29], solid state physics [30].

The paper is organized as follows. In section 2 we describe the MF-DFA and F-DFA methods in detail and show that the scaling exponents determined via the MF-DFA method are identical to those obtained by the standard multifractal formalism based on partition functions. We eliminate the sinusoidal trend via the F-DFA technique in section 3 and investigate the multifractal nature of the remaining fluctuation. In section 4, we examine the source of multifractality in sunspot data by comparing the MF-DFA results for the remaining data set to those obtained via the MF-DFA for shuffled and surrogate series. Section 5 closes with a discussion of the present results.

J. Stat. Mech. (2006) P02003

2. Multifractal detrended fluctuation analysis

The simplest type of multifractal analysis is based upon the standard partition function multifractal formalism, which has been developed for the multifractal characterization of normalized, stationary measurements [31]–[34]. Unfortunately, this standard formalism does not give correct results for non-stationary time series that are affected by trends or that cannot be normalized. Thus, in the early 1990s an improved multifractal formalism was developed, the wavelet transform modulus maxima (WTMM) method [35], which is based on the wavelet analysis and involves tracing the maxima lines in the continuous wavelet transform over all scales. Another method, multifractal detrended fluctuation analysis (MF-DFA), is based on the identification of the scaling of the qth-order moments depending on the signal length and is a generalization of the standard DFA using only the second moment $q = 2$.

The MF-DFA does not require the modulus maxima procedure, in contrast to the WTMM method, and hence does not require more effort in programming and computing than the conventional DFA. On the other hand, often experimental data are affected by non-stationarities like trends, which have to be well distinguished from the intrinsic fluctuations of the system in order to find the correct scaling behaviour of the fluctuations. In addition very often we do not know the reasons for underlying trends in collected data and even worse we do not know the scales of the underlying trends; also, usually the available record data set is small. For the reliable detection of correlations, it is essential to distinguish trends from the fluctuations intrinsic to the data. Hurst rescaled range analysis [36] and other non-detrending methods work well if the records are long and do not involve trends. But if trends are present in the data, they might give wrong results. Detrended fluctuation analysis (DFA) is a well-established method for determining the scaling behaviour of noisy data in the presence of trends without knowing their origin and shape [14, 21], [37]–[39].

2.1. Description of the MF-DFA

The modified multifractal DFA (MF-DFA) procedure consists of five steps. The first three steps are essentially identical to the conventional DFA procedure (see e.g. [14]–[18]). Suppose that x_k is a series of length N, and that this series is of compact support, i.e. $x_k = 0$ for an insignificant fraction of the values only.

- *Step 1.* Determine the 'profile'

$$Y(i) \equiv \sum_{k=1}^{i} [x_k - \langle x \rangle], \qquad i = 1, \ldots, N. \tag{1}$$

  Subtraction of the mean $\langle x \rangle$ is not compulsory, since it would be eliminated by the later detrending in the third step.

- *Step 2.* Divide the profile $Y(i)$ into $N_s \equiv \mathrm{int}\,(N/s)$ non-overlapping segments of equal lengths s. Since the length N of the series is often not a multiple of the timescale s considered, a short part at the end of the profile may remain. In order not to disregard this part of the series, the same procedure is repeated starting from the opposite end. Thereby, $2N_s$ segments are obtained altogether.

J. Stat. Mech. (2006) P02003

- *Step 3.* Calculate the local trend for each of the $2N_s$ segments by a least squares fit of the series. Then determine the variance

$$F^2(s,\nu) \equiv \frac{1}{s}\sum_{i=1}^{s}\{Y[(\nu-1)s+i]-y_\nu(i)\}^2\,, \tag{2}$$

for each segment ν, $\nu = 1,\ldots,N_s$, and

$$F^2(s,\nu) \equiv \frac{1}{s}\sum_{i=1}^{s}\{Y[N-(\nu-N_s)s+i]-y_\nu(i)\}^2\,, \tag{3}$$

for $\nu = N_s+1,\ldots,2N_s$. Here, $y_\nu(i)$ is the fitting polynomial in segment ν. Linear, quadratic, cubic or higher order polynomials can be used in the fitting procedure (conventionally called DFA1, DFA2, DFA3, ...) [14, 22]. Since the detrending of the time series is done by the subtraction of the polynomial fits from the profile, different order DFA differ in their capability of eliminating trends in the series. In (MF-)DFAm (mth-order (MF-)DFA) trends of order m in the profile (or, equivalently, of order $m-1$ in the original series) are eliminated. Thus a comparison of the results for different orders of DFA allows one to estimate the type of the polynomial trend in the time series [16, 17].

- *Step 4.* Average over all segments to obtain the qth-order fluctuation function, defined as

$$F_q(s) \equiv \left\{\frac{1}{2N_s}\sum_{\nu=1}^{2N_s}\left[F^2(s,\nu)\right]^{q/2}\right\}^{1/q}\,, \tag{4}$$

where, in general, the index variable q can take any real value except zero. For $q = 2$, the standard DFA procedure is retrieved. Generally we are interested in how the generalized q dependent fluctuation functions $F_q(s)$ depend on the timescale s for different values of q. Hence, we must repeat steps 2, 3 and 4 for several timescales s. It is apparent that $F_q(s)$ will increase with increasing s. Of course, $F_q(s)$ depends on the DFA order m. By construction, $F_q(s)$ is only defined for $s \geq m+2$.

- *Step 5.* Determine the scaling behaviour of the fluctuation functions by analysing log–log plots of $F_q(s)$ versus s for each value of q. If the series x_i are long range power law correlated, $F_q(s)$ increases, for large values of s, as a power law,

$$F_q(s) \sim s^{h(q)}. \tag{5}$$

In general, the exponent $h(q)$ may depend on q. For stationary time series such as fGn (fractional Gaussian noise), $Y(i)$ in equation (1) will be a fBm (fractional Brownian motion) signal, so $0 < h(q=2) < 1.0$. The exponent $h(2)$ is identical to the well-known Hurst exponent H [14, 15, 31]. Also for a non-stationary signal, such as fBm noise, $Y(i)$ in equation (1) will be a sum of fBm signals, so the corresponding scaling exponent of $F_q(s)$ is identified by $h(q=2) > 1.0$ [14, 40] (see the appendix for more details). In this case the relation between the exponents $h(2)$ and H will be $H = h(q=2)-1$. The exponent $h(q)$ is known as the generalized Hurst exponent. The autocorrelation function can be characterized by a power law $C(s) \equiv \langle n_k n_{k+s}\rangle \sim s^{-\gamma}$ with exponent $\gamma = 2-2H$. Its power spectra can be characterized by $S(\omega) \sim \omega^{-\beta}$ with frequency ω and $\beta = 2H-1$, In the non-stationary case, the correlation exponent and power spectrum scaling are $\gamma = -2H$ and $\beta = 2H+1$, respectively [14, 40].

J. Stat. Mech. (2006) P02003

For monofractal time series, $h(q)$ is independent of q, since the scaling behaviour of the variances $F^2(s,\nu)$ is identical for all segments ν, and the averaging procedure in equation (4) will just give this identical scaling behaviour for all values of q. If we consider positive values of q, the segments ν with large variance $F^2(s,\nu)$ (i.e. large deviations from the corresponding fit) will dominate the average $F_q(s)$. Thus, for positive values of q, $h(q)$ describes the scaling behaviour of the segments with large fluctuations. For negative values of q, the segments ν with small variance $F^2(s,\nu)$ will dominate the average $F_q(s)$. Hence, for negative values of q, $h(q)$ describes the scaling behaviour of the segments with small fluctuations[6].

2.2. Relation to standard multifractal analysis

For a stationary, normalized series the multifractal scaling exponents $h(q)$ defined in equation (5) are directly related to the scaling exponents $\tau(q)$ defined by the standard partition function-based multifractal formalism as shown below. Suppose that the series x_k of length N is a stationary, normalized sequence. Then the detrending procedure in step 3 of the MF-DFA method is not required, since no trend has to be eliminated. Thus, the DFA can be replaced by the standard fluctuation analysis (FA), which is identical to the DFA except for a simplified definition of the variance for each segment ν, $\nu = 1, \ldots, N_s$. Step 3 now becomes (see equation (2))

$$F^2_{\mathrm{FA}}(s,\nu) \equiv [Y(\nu s) - Y((\nu-1)s)]^2. \tag{6}$$

Inserting this simplified definition into equation (4) and using equation (5), we obtain

$$\left\{ \frac{1}{2N_s} \sum_{\nu=1}^{2N_s} |Y(\nu s) - Y((\nu-1)s)|^q \right\}^{1/q} \sim s^{h(q)}. \tag{7}$$

For simplicity we can assume that the length N of the series is an integer multiple of the scale s, obtaining $N_s = N/s$, and therefore

$$\sum_{\nu=1}^{N/s} |Y(\nu s) - Y((\nu-1)s)|^q \sim s^{qh(q)-1}. \tag{8}$$

This corresponds to the multifractal formalism used e.g. in [32, 34]. In fact, a hierarchy of exponents H_q similar to our $h(q)$ has been introduced on the basis of equation (8) in [32]. In order to relate also to the standard textbook box counting formalism [31, 33], we employ the definition of the profile in equation (1). It is evident that the term $Y(\nu s) - Y((\nu-1)s)$ in equation (8) is identical to the sum of the numbers x_k within each segment ν of size s. This sum is known as the box probability $p_s(\nu)$ in the standard multifractal formalism

[6] For the maximum scale $s = N$ the fluctuation function $F_q(s)$ is independent of q, since the sum in equation (4) runs over only two identical segments ($N_s \equiv [N/s] = 1$). For smaller scales $s \ll N$ the averaging procedure runs over several segments, and the average value $F_q(s)$ will be dominated by the $F^2(s,\nu)$ from the segments with small (large) fluctuations if $q < 0$ ($q > 0$). Thus, for $s \ll N$, $F_q(s)$ with $q < 0$ will be smaller than $F_q(s)$ with $q > 0$, while they become equal for $s = N$. Hence, if we assume an homogeneous scaling behaviour of $F_q(s)$ following equation (5), the slope $h(q)$ in a log–log plot of $F_q(s)$ with $q < 0$ versus s must be larger than the corresponding slope for $F_q(s)$ with $q > 0$. Thus, $h(q)$ for $q < 0$ will usually be larger than $h(q)$ for $q > 0$.

for normalized series x_k,

$$p_s(\nu) \equiv \sum_{k=(\nu-1)s+1}^{\nu s} x_k = Y(\nu s) - Y((\nu-1)s). \tag{9}$$

The scaling exponent $\tau(q)$ is usually defined via the partition function $Z_q(s)$,

$$Z_q(s) \equiv \sum_{\nu=1}^{N/s} |p_s(\nu)|^q \sim s^{\tau(q)}, \tag{10}$$

where q is a real parameter as in the MF-DFA method, discussed above. Using equation (9) we see that equation (10) is identical to equation (8), and obtain analytically the relation between the two sets of multifractal scaling exponents,

$$\tau(q) = qh(q) - 1. \tag{11}$$

Thus, we observe that $h(q)$ defined in equation (5) for the MF-DFA is directly related to the classical multifractal scaling exponents $\tau(q)$. Note that $h(q)$ is different from the generalized multifractal dimensions

$$D(q) \equiv \frac{\tau(q)}{q-1} = \frac{qh(q)-1}{q-1}, \tag{12}$$

that are used instead of $\tau(q)$ in some papers. While $h(q)$ is independent of q for a monofractal time series, $D(q)$ depends on q in this case. Another way to characterize a multifractal series is via the singularity spectrum $f(\alpha)$, that is related to $\tau(q)$ via a Legendre transform [31, 33]:

$$\alpha = \tau'(q) \qquad \text{and} \qquad f(\alpha) = q\alpha - \tau(q). \tag{13}$$

Here, α is the singularity strength or Hölder exponent, while $f(\alpha)$ denotes the dimension of the subset of the series that is characterized by α. Using equation (11), we can directly relate α and $f(\alpha)$ to $h(q)$:

$$\alpha = h(q) + qh'(q) \qquad \text{and} \qquad f(\alpha) = q[\alpha - h(q)] + 1. \tag{14}$$

A Hölder exponent denotes monofractality, while in the multifractal case, the different parts of the structure are characterized by different values of α, leading to the existence of the spectrum $f(\alpha)$.

2.3. Fourier detrended fluctuation analysis

In some cases, there exist one or more crossover (time)scales $s_\times$ separating regimes with different scaling exponents [16, 17]. In this case investigation of the scaling behaviour is more complicated and different scaling exponents are required for different parts of the series [18]. Therefore one needs a multitude of scaling exponents (multifractality) for a full description of the scaling behaviour. A crossover can often arise from a variation in the correlation properties of the signal at different scales of time or space, or from trends in the data. To remove the crossover due to a trend such as a sinusoidal trend, Fourier detrended fluctuation analysis (F-DFA) is applied. F-DFA is a modified approach for the analysis of low frequency trends added to a noise in time series [12, 13, 41, 42].

J. Stat. Mech. (2006) P02003

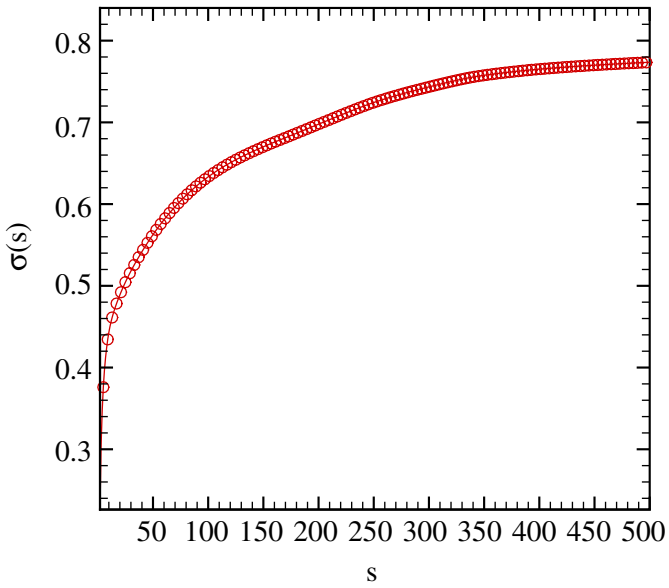

Figure 2. Behaviour of the standard deviation of the sunspot time series as a function of the timescale. It shows that this time series is not stationary and direct calculation of the correlation gives a strongly wrong result.

In order to investigate how we can remove trends having a low frequency periodic behaviour, we transform the data record to Fourier space, then we truncate the first few coefficients of the Fourier expansion and inverse Fourier transform the series. After removing the sinusoidal trends we can obtain the fluctuation exponent by using the direct calculation of the MF-DFA. If truncation numbers are sufficient, the crossover due to a sinusoidal trend in the log–log plot of $F_q(s)$ versus s disappears.

3. Analysis of sunspot time series

As mentioned in section 2, a spurious correlation may be detected if the time series is non-stationarity, so direct calculations of correlation behaviour, spectral density exponents, fractal dimensions etc do not give reliable results. It can be checked that the sunspot time series is non-stationary. One can verified the non-stationarity property experimentally by measuring the stability of the average and variance in a moving window, for example with scale s. Figure 2 shows that the standard deviation of the signal versus scale s is not saturated. Let us determine whether the data set has a sinusoidal trend or not. According to the MF-DFA1 method, generalized Hurst exponents $h(q)$ in equation (5) can be found by analysing log–log plots of $F_q(s)$ versus s for each q. Our investigation shows that there are three crossover timescales $s_\times$ in the log–log plots of $F_q(s)$ versus s for every q. These three crossovers divide $F_q(s)$ into four regions, as shown in figure 3 (for instance we took $q = 2$). The existence of these regions is due to the competition between the noise and sinusoidal trend. For $s < s_{1\times}$ and $s > s_{3\times}$, the noise has the dominating effect [17]. For $s_{1\times} < s < s_{2\times}$ and $s_{2\times} < s < s_{3\times}$, the sinusoidal trend dominates [17]. The value of $s_{2\times}$ is approximately equal to 130 months which is equal to the well-known cycle of sun activity. As mentioned before, for very small scales $s < s_{1\times}$ the effect of the sinusoidal trend is not pronounced, indicating that in this scale region the signal can be considered as noise

J. Stat. Mech. (2006) P02003

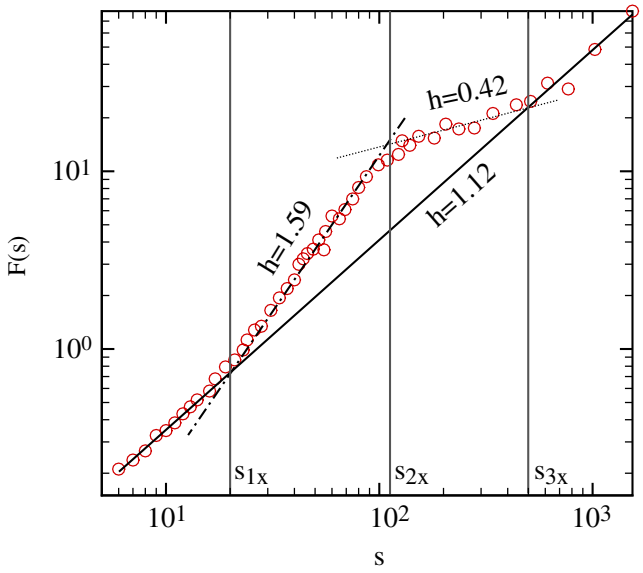

Figure 3. Crossover behaviour of the log–log plot of $F(s)$ versus s for the sunspot time series for $q = 2.0$. There are three crossover timescales in the plot of $F(s)$, at scales $s_{1\times}$, $s_{2\times}$ and $s_{3\times}$.

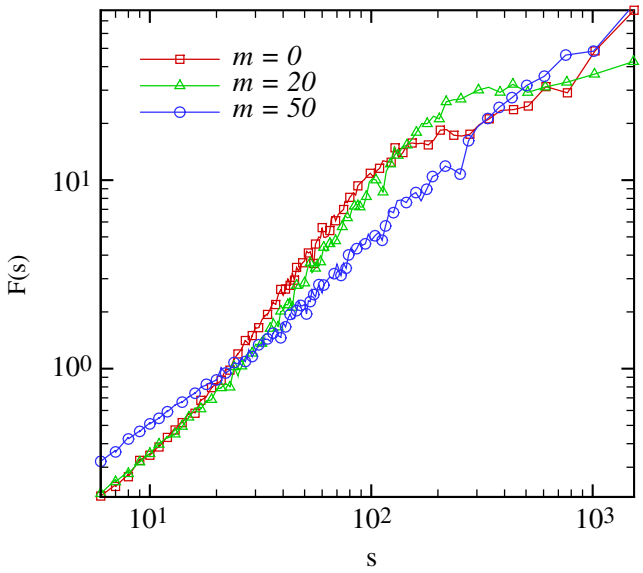

Figure 4. The MF-DFA1 functions $F_q(s)$ for the sunspot time series versus the timescale s in a log–log plot. The original time series $m = 0$, truncation of the first 20 terms $m = 20$ and 50 terms $m = 50$.

fluctuating around a constant which is filtered out by the MF-DFA1 procedure. In this region the generalized DFA1 exponent is $h(q = 2) = 1.12 \pm 0.01$, which confirms that the process is a non-stationary process with anticorrelation behaviour.

To cancel the sinusoidal trend in MF-DFA1, we apply the F-DFA method to sunspot data. We truncate some of the first coefficients of the Fourier expansion of the sunspot series. According to figure 4, for eliminating the crossover scales, we need to remove

J. Stat. Mech. (2006) P02003

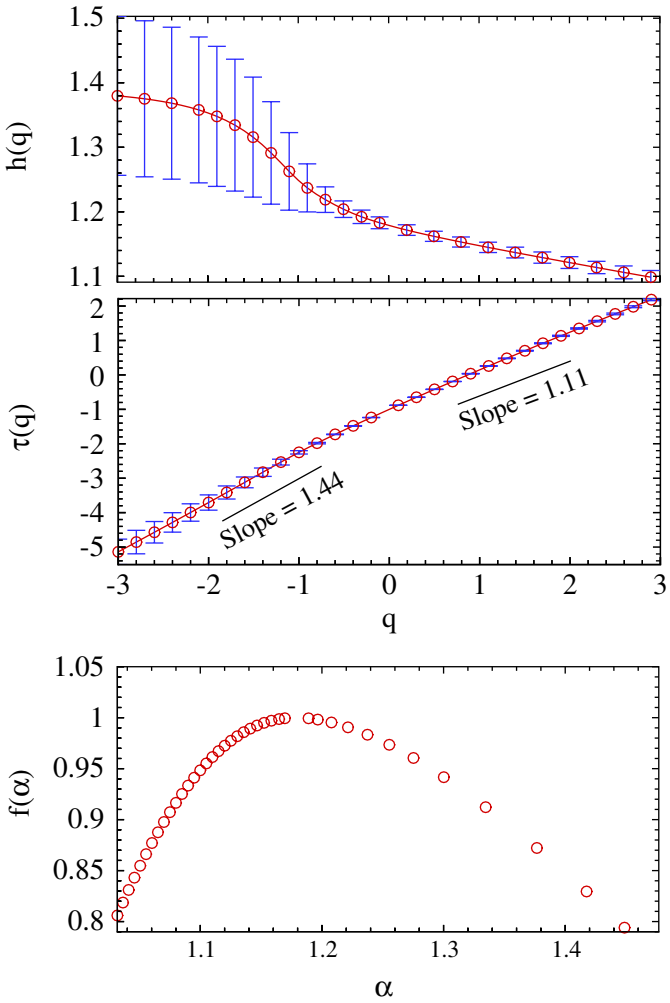

Figure 5. The q dependence of the generalized Hurst exponent $h(q)$, the corresponding $\tau(q)$ and the singularity spectrum $f(\alpha)$ are shown in the upper to lower panels respectively for sunspot time series without a sinusoidal trend.

approximately 50 terms of the Fourier expansion. Then, by inverse Fourier transformation, the noise without a sinusoidal trend is extracted.

The MF-DFA1 results of the remaining new signal are shown in figure 5. The sunspot time series is a multifractal process as indicated by the strong q dependence of the generalized Hurst exponents and $\tau(q)$ [43]. The q dependence of the classical multifractal scaling exponent $\tau(q)$ has different behaviours for $q < 0$ and $q > 0$. For positive and negative values of q, the slopes of $\tau(q)$ are 1.11 ± 0.01 and 1.44 ± 0.01, respectively. According to the relation between the Hurst exponent and $h(2)$, i.e. $h(q = 2) - 1 = H$, we find that the Hurst exponent is 0.12 ± 0.01. This result is equal to the value of the Hurst exponent on a small scale of MF-DFA1 for noise with a sinusoidal trend. The fractal dimension is obtained as $D_f = 2 - H = 1.88$ [16]. The values of quantities derived from the MF-DFA1 method are given in tables 1 and 2.

Usually, in the MF-DFA method, deviation from a straight line in the log–log plot of equation (5) occurs for small scales s. This deviation limits the capability of DFA

J. Stat. Mech. (2006) P02003

Table 1. The values of $h(q = 2)$, multifractal scaling and generalized multifractal exponents for $q = 2.0$ for original, surrogate and shuffled monthly sunspot time series obtained by MF-DFA1.

Data	h	τ	D
Sunspot	1.12 ± 0.01	1.24 ± 0.02	1.24 ± 0.02
Surrogate	1.13 ± 0.01	1.26 ± 0.02	1.26 ± 0.02
Shuffled	0.51 ± 0.01	0.02 ± 0.02	0.02 ± 0.02

Table 2. The values of the Hurst (H), power spectrum scaling (β) and autocorrelation scaling (γ) exponents for original, surrogate and shuffled monthly sunspot time series obtained by MF-DFA1.

Data	H	β	γ
Sunspot	0.12 ± 0.02	1.24 ± 0.02	-0.24 ± 0.02
Surrogate	0.13 ± 0.02	1.26 ± 0.02	-0.26 ± 0.02
Shuffled	0.51 ± 0.01	0.02 ± 0.02	0.98 ± 0.02

to determine the correct correlation behaviour for very short scales and in the regime of small s. The modified MF-DFA is defined as follows [16]:

$$F_q^{\mathrm{mod}}(s) = \frac{F_q(s)}{K_q(s)}, \\ = F_q(s) \frac{\langle [F_q^{\mathrm{shuf}}(s')]^2 \rangle^{1/2} \, s^{1/2}}{\langle [F_q^{\mathrm{shuf}}(s)]^2 \rangle^{1/2} \, s'^{1/2}} \qquad (\text{for } s' \gg 1), \tag{15}$$

where $\langle [F_q^{\mathrm{shuf}}(s)]^2 \rangle^{1/2}$ denotes the usual MF-DFA fluctuation function (defined in equation (4)) averaged over several configurations of shuffled data taken from the original time series, and $s' \approx N/40$. The value of the Hurst exponent obtained by modified MF-DFA1 methods for sunspot time series is 0.11 ± 0.01. The relative deviation of the Hurst exponent which is obtained by modified MF-DFA1 in comparison to MF-DFA1 for the original data is approximately 8.33%.

4. Comparison of the multifractality for original, shuffled and surrogate sunspot time series

As discussed in the section 3 the remaining data set after the elimination of the sinusoidal trend has a multifractal nature. In this section we are interested in determining the source of multifractality. In general, two different types of multifractality in time series can be distinguished:

(i) Multifractality due to a fatness of the probability density function (PDF) of the time series. In this case the multifractality cannot be removed by shuffling the series.

(ii) Multifractality due to different correlations in small and large scale fluctuations. In this case the data may have a PDF with finite moments, e.g. a Gaussian distribution.

J. Stat. Mech. (2006) P02003

Thus the corresponding shuffled time series will exhibit monofractal scaling, since all long range correlations are destroyed by the shuffling procedure. If both types of multifractality are present, the shuffled series will show weaker multifractality than the original series. The easiest way to clarify the type of multifractality is by analysing the corresponding shuffled and surrogate time series. The shuffling of time series destroys the long range correlation, therefore if the multifractality belongs only to the long range correlation, we should find $h_{\rm shuf}(q) = 0.5$. The multifractality nature due to the fatness of the PDF signals is not affected by the shuffling procedure. On the other hand, to determine the multifractality due to the broadness of the PDF, the phase of the discrete Fourier transform (DFT) coefficients of sunspot time series are replaced with a set of pseudo-independent distributed uniform $(-\pi, \pi)$ quantities in the surrogate method. The correlations in the surrogate series do not change, but the probability function changes to the Gaussian distribution. If multifractality in the time series is due to a broad PDF, $h(q)$ obtained by the surrogate method will be independent of q. If both types of multifractality are present in sunspot time series, the shuffled and surrogate series will show weaker multifractality than the original one.

To check the nature of the multifractality, we compare the fluctuation function $F_q(s)$, for the original series (after cancellation of the sinusoidal trend) with the result for the corresponding shuffled, $F_q^{\rm shuf}(s)$, and surrogate series $F_q^{\rm sur}(s)$. Differences between these two fluctuation functions and the original one directly indicate the presence of long range correlations or broadness of the probability density function in the original series. These differences can be observed in a plot of the ratios $F_q(s)/F_q^{\rm shuf}(s)$ and $F_q(s)/F_q^{\rm sur}(s)$ versus s [43]. Since the anomalous scaling due to a broad probability density affects $F_q(s)$ and $F_q^{\rm shuf}(s)$ in the same way, only multifractality due to correlations will be observed in $F_q(s)/F_q^{\rm shuf}(s)$. The scaling behaviours of these ratios are

$$F_q(s)/F_q^{\rm shuf}(s) \sim s^{h(q)-h_{\rm shuf}(q)} = s^{h_{\rm cor}(q)}, \tag{16}$$

$$F_q(s)/F_q^{\rm sur}(s) \sim s^{h(q)-h_{\rm sur}(q)} = s^{h_{\rm PDF}(q)}. \tag{17}$$

If only fatness of the PDF is responsible for the multifractality, one should find $h(q) = h_{\rm shuf}(q)$ and $h_{\rm cor}(q) = 0$. On the other hand, deviations from $h_{\rm cor}(q) = 0$ indicate the presence of correlations, and q dependence of $h_{\rm cor}(q)$ indicates that multifractality is due to the long range correlation. If only correlation multifractality is present, one finds $h_{\rm shuf}(q) = 0.5$. If both distribution and correlation multifractality are present, both $h_{\rm shuf}(q)$ and $h_{\rm sur}(q)$ will depend on q. The q dependences of the exponent $h(q)$ for original, surrogate and shuffled time series are shown in figure 6. The q dependence of $h_{\rm cor}$ and $h_{\rm PDF}$ shows that the multifractality nature of sunspot time series is due to both broadness of the PDF and long range correlation. The absolute value of $h_{\rm cor}(q)$ is greater than $h_{\rm PDF}(q)$, so the multifractality due to the fatness is weaker than the multifractality due to the correlation. The deviations of $h_{\rm sur}(q)$ and $h_{\rm shuf}(q)$ from $h(q)$ can be determined by using the χ^2 test as follows:

$$\chi^2_\diamond = \sum_{i=1}^{N} \frac{[h(q_i) - h_\diamond(q_i)]^2}{\sigma(q_i)^2 + \sigma_\diamond(q_i)^2}, \tag{18}$$

where the symbols '$\diamond$' can be replaced by 'sur' and 'shuf', to determine the confidence level of $h_{\rm sur}$ and $h_{\rm shuf}$ for generalized Hurst exponents of the original series, respectively.

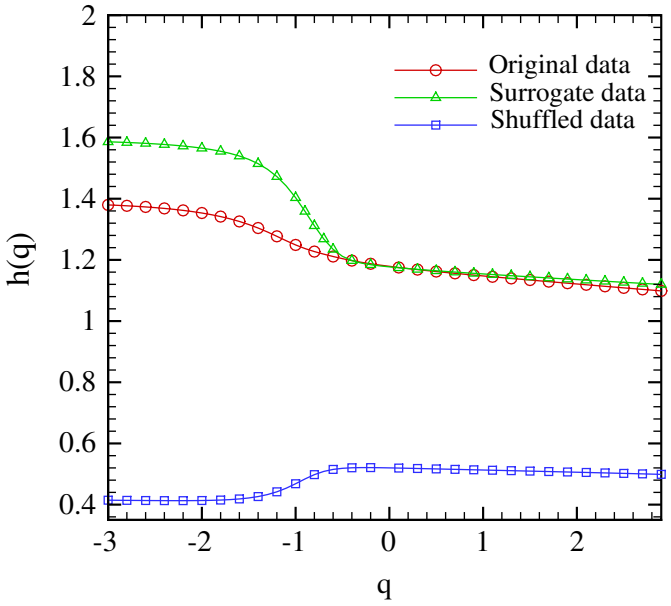

Figure 6. Generalized Hurst exponent, $h(q)$ as a function of q for the original, surrogate and shuffled data.

Table 3. The value of the Hurst exponent obtained using MF-DFA1 and modified MF-DFA1 for the original, shuffled and surrogate monthly sunspot time series.

Method	Sunspot	Surrogate	Shuffled
MF-DFA1	0.12 ± 0.01	0.13 ± 0.01	0.51 ± 0.01
Modified	0.11 ± 0.01	0.12 ± 0.01	0.50 ± 0.01

The values of the reduced chi square $\chi^2_{\nu\diamond} = \chi^2_{\diamond}/\mathcal{N}$ ($\mathcal{N}$ is the number of degrees of freedom) for shuffled and surrogate time series are 1653.47, 1.10, respectively. On the other hand, the widths of the singularity spectrum, $f(\alpha)$, i.e. $\Delta\alpha = \alpha(q_{\min}) - \alpha(q_{\max})$ for original, surrogate and shuffled time series, are approximately 0.44, 0.75 and 0.22 respectively. These values also show that the multifractality due to correlation is dominant [44].

The values of the generalized Hurst exponent $h(q = 2.0)$, multifractal scaling $\tau(q = 2)$ and generalized multifractal exponents $(D(q = 2))$ for the original, shuffled and surrogate sunspot time series obtained with the MF-DFA1 method are reported in table 1, The related scaling exponents are indicated in table 2. The values of the Hurst exponent obtained by MF-DFA1 and modified MF-DFA1 methods for original, surrogate and shuffled sunspot time series are given in table 3.

5. Conclusion

The MF-DFA method allows us to determine the multifractal characterization of the non-stationary and stationary time series. The concept of MF-DFA of sunspot time series can be used to gain deeper insight into the processes occurring in non-stationary

J. Stat. Mech. (2006) P02003

dynamical systems, such as sunspot formation. We have shown that the MF-DFA1 result for the monthly sunspot time series has three crossover timescales ($s_\times$). These crossover timescales are due to the sinusoidal trend. To minimize the effect of this trend, we have applied F-DFA to sunspot time series. Applying the MF-DFA1 method to truncated data, we demonstrated that the monthly sunspot time series is a non-stationary time series with anticorrelation behaviour. The q dependence of $h(q)$ and $\tau(q)$ indicated that the monthly sunspot time series has multifractal behaviour. By comparing the generalized Hurst exponent of the original time series with the shuffled and surrogate ones, we have found that multifractality due to the correlation makes a greater contribution than the broadness of the probability density function.

Acknowledgments

We would like to thank Sepehr Arbabi Bidgoli and Mojtaba Mohammadi Najafabadi for reading the manuscript and useful comments. This paper is dedicated to Dr Somaihe Abdolahi.

Appendix

In this appendix we derive the relation between the exponent $h(2)$ (DFA1 exponent) and the Hurst exponent of a fBm signal. We show that for such a non-stationary signal the average sample variance (equation (4)) for $q = 2$ is proportional to $s^{h(q)}$, where $h(q = 2) = H + 1$. It is shown that the averaged sample variance $F^2(s)$ behaves as

$$\begin{aligned} F^2(s) &\equiv \frac{1}{2N_s}\sum_{\nu=1}^{2N_s}\left[F^2(s,\nu)\right], \\ &= \left\langle\left[F^2(s,\nu)\right]\right\rangle_\nu, \\ &\equiv \mathcal{C}_H s^{2(H+1)}, \end{aligned} \tag{A.1}$$

where $F^2(s,\nu)$ is defined as in equation (2) and $\mathcal{C}_H$ is a function of Hurst exponent H.

To prove the statement we note that the data set $x(k)$ is a fractional Brownian motion (fBm); the partial sums $Y(i)$ (equation (1)) will be a summed fBm signal. In the DFA1, the fitting function will have the expression ($y_\nu = a_\nu + b_\nu i$). The slope b_ν and intercept a_ν of a least squares line on $Y(i)$ (from 0 to s) for every window (ν) are given by

$$\begin{aligned} b_\nu &= \frac{\sum_{i=1}^{s} Y(i)i - (1/s)\sum_{i=1}^{s} Y(i)\sum_{i=1}^{s} i}{\sum_{i=1}^{s} i^2 - (1/s)\left[\sum_{i=1}^{s} i\right]^2}, \\ &\simeq \frac{\sum_{i=1}^{s} Y(i)i - (s/2)\sum_{i=1}^{s} Y(i)}{s^3/12} \\ a_\nu &= \frac{1}{s}\sum_{i=1}^{s} Y(i) - \frac{1}{s}\sum_{i=1}^{s} i \simeq \frac{1}{s}\sum_{i=1}^{s} Y(i) - \frac{s}{2}, \end{aligned} \tag{A.2}$$

respectively.

Using equations (4) and (A.2), equation (A.1) can be written as follows:

$$\begin{aligned}\left\langle\left[F^2(s,\nu)\right]\right\rangle &= \left\langle\frac{1}{s}\sum_{i=1}^{s}(Y(i)-a-bi)^2\right\rangle \\ &\simeq \left\langle\frac{1}{s}\sum_{i=1}^{s}Y(i)^2\right\rangle + \left\langle a^2\right\rangle + \frac{s^2}{3}\left\langle b^2\right\rangle \\ &\quad - 2\left\langle\frac{a}{s}\sum_{i=1}^{s}y(i)\right\rangle - 2\left\langle\frac{b}{s}\sum_{i=1}^{s}iY(i)\right\rangle + s\left\langle ab\right\rangle, \\ &= \left\langle\frac{1}{s}\sum_{i=1}^{s}Y(i)^2\right\rangle - \frac{4}{s^2}\left\langle\left[\sum_{i=1}^{s}Y(i)\right]^2\right\rangle \\ &\quad - \frac{12}{s^4}\left\langle\left[\sum_{i=1}^{s}iY(i)\right]^2\right\rangle + \frac{12}{s^3}\left\langle\sum_{i=1}^{s}iY(i)\sum_{i=1}^{s}Y(i)\right\rangle, \\ &= \frac{A}{s} - \frac{4}{s^2}B - \frac{12}{s^4}D + \frac{12}{s^3}C \end{aligned} \tag{A.3}$$

where we have discarded the subscript ν for simplicity. Now let us calculate the functions A, B, C and D in equation (A.3). The increment of summed fBm and fBm signals, i.e.

$$\begin{aligned} x(i) &= Y(i) - Y(i-1) \\ u(i) &= x(i) - x(i-1), \end{aligned} \tag{A.4}$$

are a fBm $x(i)$ and fGn $u(i)$ noise, respectively. The correlations of $Y(i)$ and $x(i)$ are as follows [15]:

$$\begin{aligned}\langle x(i)x(j)\rangle &= \frac{\sigma^2}{2}\left[i^{2H}+j^{2H}-|i-j|^{2H}\right], \\ \langle Y(i)Y(j)\rangle &= \frac{\sigma^2}{(H+1)^2}(ij)^{H+1}, \end{aligned} \tag{A.5}$$

where $\sigma^2 = \langle u(i)^2\rangle$. Also the variance of a summed fBm signal is $\langle Y(i)^2\rangle = (\sigma^2/(H+1)^2)i^{2(H+1)}$ [14]. Finally, using equations (A.2) and (A.5), it can be easily shown that equation (A.3) can be written as follows:

$$\left\langle\left[F^2(s,\nu)\right]\right\rangle_\nu = \mathcal{C}_H s^{2(H+1)}, \tag{A.6}$$

where $\mathcal{C}_H$ is

$$\begin{aligned}\mathcal{C}_H = &\frac{\sigma^2}{(2H+3)(H+1)^2} - \frac{4\sigma^2}{[(H+1)(H+2)]^2} \\ &- \frac{12\sigma^2}{[(H+1)(H+3)]^2} + \frac{12\sigma^2}{(H+1)^2(H+2)(H+3)}. \end{aligned} \tag{A.7}$$

Therefore the standard DFA1 exponent for a non-stationary signal is related to its Hurst exponent as $h(q=2) = H+1$.

References

[1] Bray R J and Loughhead R E, 1979 *Sunspots* (New York: Dover)
[2] Petrovaye E, 2000 *ESA Publ. Sol. Phys.* **463** 3–14
[3] http://www.oma.be/KSB-ORB/SIDC/index.html
[4] Veronig A, Messerotti M and Hanslmeier A, 2000 *Astron. Astrophys.* **357** 337
[5] Abramenko V I, 2005 *Sol. Phys.* **228** 29–42
[6] Zhukov V I, 2003 *Preprint* astro-ph/0304456
[7] Temmer M, Veronig A, Rybák J and Hanslmeier A, *Solar variability: from core to outer frontiers*, 2002 *10th Eur. Solar Physics Mtg (9–14 September 2002, Prague, Czech Republic)* vol 2, ESA SP-506, ed A Wilson (Noordwijk: ESA Publications Division) pp 859–62
[8] Temmer M, Veronig A and Hanslmeier A, 2002 *Astron. Astrophys.* **390** 707
[9] Bogart R S, 1982 *Sol. Phys.* **76** 155
[10] Temmer M *et al*, 2002 *Preprint* astro-ph/0207239
[11] Temmer M, Veronig A and Hanslmeier A, 2003 *Sol. Phys.* **512** 111–29
[12] Nagarajan R and Kavasseri R G, 2004 *Preprint* cond-mat/0411543
[13] Chianca C V, Ticona A and Penna T J P, 2005 *Physica* A **357** 447–54
[14] Peng C K, Buldyrev S V, Havlin S, Simons M, Stanley H E and Goldberger A L, 1994 *Phys. Rev.* E **49** 1685
Ossadnik S M, Buldyrev S B, Goldberger A L, Havlin S, Mantegna R N, Peng C K, Simons M and Stanley H E, 1994 *Biophys. J.* **67** 64
[15] Taqqu M S, Teverovsky V and Willinger W, 1995 *Fractals* **3** 785
[16] Kantelhardt J W, Koscielny-Bunde E, Rego H H A, Havlin S and Bunde A, 2001 *Physica* A **295** 441
[17] Hu K, Ivanov P Ch, Chen Z, Carpena P and Stanley H E, 2001 *Phys. Rev.* E **64** 011114
[18] Chen Z, Ivanov P Ch, Hu K and Stanley H E, 2002 *Phys. Rev.* E **65** [physics/0111103]
[19] Buldyrev S V, Goldberger A L, Havlin S, Mantegna R N, Matsa M E, Peng C K, Simons M and Stanley H E, 1995 *Phys. Rev.* E **51** 5084
Buldyrev S V, Dokholyan N V, Goldberger A L, Havlin S, Peng C K, Stanley H E and Viswanathan G M, 1998 *Physica* A **249** 430
[20] Ivanov P Ch, Bunde A, Amaral L A N, Havlin S, Fritsch-Yelle J, Baevsky R M, Stanley H E and Goldberger A L, 1999 *Europhys. Lett.* **48** 594
Ashkenazy Y, Lewkowicz M, Levitan J, Havlin S, Saermark K, Moelgaard H, Thomsen P E B, Moller M, Hintze U and Huikuri H V, 2001 *Europhys. Lett.* **53** 709
Ashkenazy Y, Ivanov P Ch, Havlin S, Peng C K, Goldberger A L and Stanley H E, 2001 *Phys. Rev. Lett.* **86** 1900
[21] Peng C K, Havlin S, Stanley H E and Goldberger A L, 1995 *Chaos* **5** 82
[22] Bunde A, Havlin S, Kantelhardt J W, Penzel T, Peter J H and Voigt K, 2000 *Phys. Rev. Lett.* **85** 3736
[23] Blesic S, Milosevic S, Stratimirovic D and Ljubisavljevic M, 1999 *Physica* A **268** 275
Bahar S, Kantelhardt J W, Neiman A, Rego H H A, Russell D F, Wilkens L, Bunde A and Moss F, 2001 *Europhys. Lett.* **56** 454
[24] Hausdorff J M, Mitchell S L, Firtion R, Peng C K, Cudkowicz M E, Wei J Y and Goldberger A L, 1997 *J. Appl. Phys.* **82** 262
[25] Koscielny-Bunde E, Bunde A, Havlin S, Roman H E, Goldreich Y and Schellnhuber H J, 1998 *Phys. Rev. Lett.* **81** 729
Ivanova K and Ausloos M, 1999 *Physica* A **274** 349
Talkner P and Weber R O, 2000 *Phys. Rev.* E **62** 150
[26] Ivanova K, Ausloos M, Clothiaux E E and Ackerman T P, 2000 *Europhys. Lett.* **52** 40
[27] Malamud B D and Turcotte D L, 1999 *J. Stat. Plan. Inference.* **80** 173
[28] Alados C L and Huffman M A, 2000 *Ethnology* **106** 105
[29] Mantegna R N and Stanley H E, 2000 *An Introduction to Econophysics* (Cambridge: Cambridge University Press)
Liu Y, Gopikrishnan P, Cizeau P, Meyer M, Peng C K and Stanley H E, 1999 *Phys. Rev.* E **60** 1390
Vandewalle N, Ausloos M and Boveroux P, 1999 *Physica* A **269** 170
[30] Kantelhardt J W, Berkovits R, Havlin S and Bunde A, 1999 *Physica* A **266** 461
Vandewalle N, Ausloos M, Houssa M, Mertens P W and Heyns M M, 1999 *Appl. Phys. Lett.* **74** 1579
[31] Feder J, 1988 *Fractals* (New York: Plenum)
[32] Barabási A L and Vicsek T, 1991 *Phys. Rev.* A **44** 2730
[33] Peitgen H O, Jürgens H and Saupe D, 1992 *Chaos and Fractals* (New York: Springer) Appendix B

J. Stat. Mech. (2006) P02003

[34] Bacry E, Delour J and Muzy J F, 2001 *Phys. Rev.* E **64** 026103
[35] Muzy J F, Bacry E and Arneodo A, 1991 *Phys. Rev. Lett.* **67** 3515
[36] Hurst H E, Black R P and Simaika Y M, 1965 *Long-Term Storage: An Experimental Study* (London: Constable)
[37] Fano U, 1947 *Phys. Rev.* **72** 26
[38] Barmes J A and Allan D W, 1996 *Proc. IEEE* **54** 176
[39] Buldyrev S V, Goldberger A L, Havlin S, Mantegna R N, Matsa M E, Peng C K, Simons M and Stanley H E, 1995 *Phys. Rev.* E **51** 5084
[40] Eke A, Herman P, Kocsis L and Kozak L R, 2002 *Physiol. Meas.* **23** R1–38
[41] Koscielny-Bunde E, Roman H E, Bunde A, Havlin S and Schellnhuber H J, 1998 *Phil. Mag.* B **77** 1331
[42] Koscielny-Bunde E, Bunde A, Havlin S, Roman H E, Goldreich Y and Schellnhuber H J, 1998 *Phys. Rev. Lett.* **81** 729
[43] Gantelhardt J W, Zschiegner S A, Kosciliny-Bunde E, Bunde A, Pavlin S and Stanley H E, 2002 *Physica* A **316** 78–114
[44] Oświęcimka P *et al*, 2005 *Preprint* cond-mat/0504608

J. Stat. Mech. (2006) P02003