Fractals, Vol. 3, No. 2 (1995) 315–328

SMOOTHING DIMENSIONS FOR TIME SERIES CHARACTERIZATION

FLORIN MUNTEANU, CRISTIAN IOANA,
CRISTIAN ŞUŢEANU and EDMOND CREŢU
Romanian Academy, Institute of Geodynamics,
19-21 Jean-Louis Calderon Str., Bucharest-37,
R 70201 Romania

Abstract

The paper refers to a recently introduced fractal signal analysis method, which relies on scaling properties of signal parameters regarding the cutoff frequency used to smoothen the signal's graph. The relations between the smoothing dimensions and other exponents (fractal dimension df, power spectrum exponent ß, Hurst exponent H) are determined theoretically and tested by numerical experiments.

In this context, a new analysis method was introduced recently[10]. It highlights scaling properties of the parameters attached to the time series, computed after successive steps of signal graph 'smoothing'. Since this evaluation procedure proves to be easy to use and leads to reliable results for many classes of signals, we intend to present here some theoretical links between the proposed smoothing dimensions and well known fractal exponents and, on the other hand, to put the method to practical tests, in order to evaluate some of its capabilities and limits.

1.2 Smoothing Dimensions

Given the original collected signal $s(t)$, we start in each case from the centred signal:

$$x = s - \frac{1}{T}\int_0^T s(t)dt. \tag{1}$$

The method makes use of a lowpass (LP) filter with selectable cutoff frequency f_c. After having passed the signal $x(t)$ through the LP filter, we compute parameters corresponding to the filtered output signal $y_{fc}(t)$. Thus, we can compute its Euclidian norm in L_2-space:

$$N_t(y_{fc}) = \sqrt{\int_0^T y_{fc}^2(t)\,dt}. \tag{2}$$

After successive filtering iterations with different cutoff frequencies f_c applied to $x(t)$, we look for a power law dependence:

$$N_t(y_{fc}) \sim (f_c)^{dn}. \tag{3}$$

In practice, one has to handle discrete time series $\{x_k\}$, $k = 1, \ldots, M$, and Eq. (2) becomes:

$$N_t(y_{fc}) = \sqrt{\sum_{k=1}^{M} y_k^2}. \tag{4}$$

Alternatively, we can compute the Fourier transform:

$$Y_{fc}(\omega) = \mathcal{F}(y_{fc}) = \int_{-\infty}^{\infty} y_{fc}(t)\, e^{-j\omega t}dt \tag{5}$$

and the parameter

$$I(y_{fc}) = \int_0^{\infty} |Y_{fc}(\omega)|\, d\omega, \tag{6}$$

looking again for a power law relation:

$$I(y_{fc}) \sim (f_c)^{di}. \tag{7}$$

In this case, for discrete time series, Eq.(7) becomes:

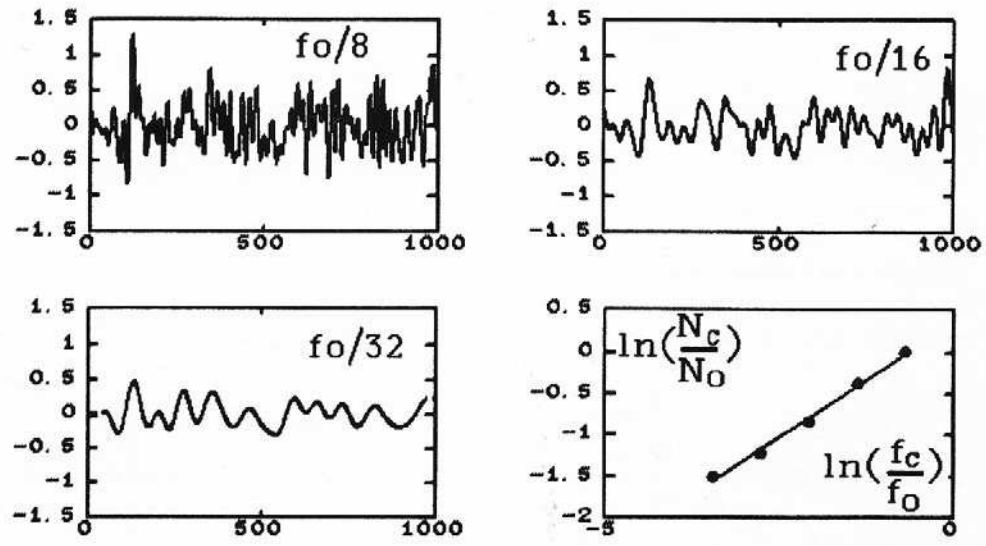
$$I(Y_{fc}) = \sum_{k=1}^{M/2} |Y_k|. \tag{8}$$

Still another possibility is to compute the curve length of the graph corresponding to the smoothened signal and to check whether it scales with the cutoff frequency f_c.

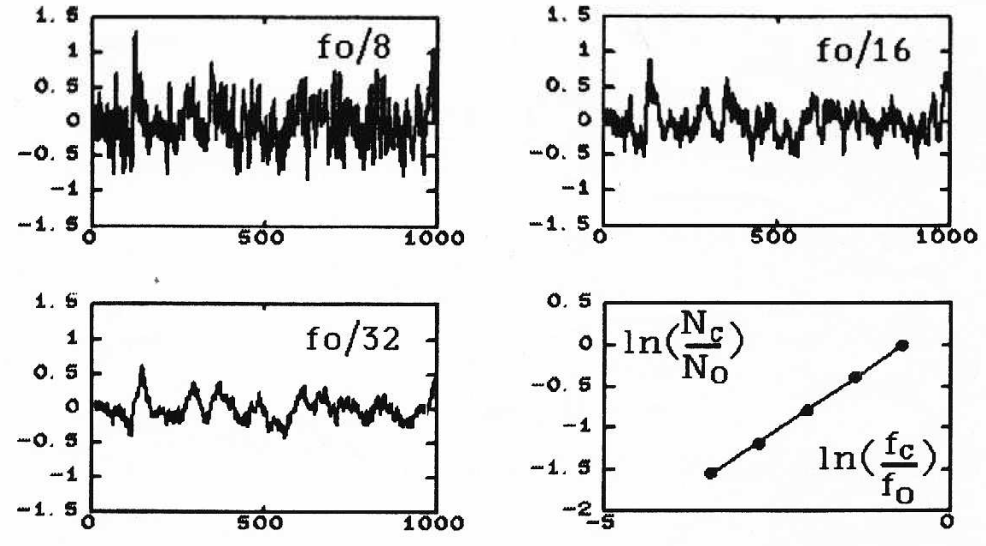
When relations like Eqs.(3) and (7) hold, we call the exponents 'smoothing dimensions'.

1.3 Smoothing Methods

The smoothing itself can be performed in different ways. One can use a *'frequency domain smoothing'* by lowering at each step the cutoff frequency f_c of the LP filter.



(a)



(b)

Fig.1 Graph of a signal after consecutive smoothing steps and determination of the smoothing dimension *dn*: (a) frequency domain smoothing; (b) time domain smoothing. Input signal generated by a cellular automaton.

One can also pass a 'smoothing window' (sliding mean) along the signal, the window width $\tau=1/f$ being increased at each step accordingly ('*time domain smoothing*'). In the latter case, the output signal is:

$$y_\tau(t) = \frac{1}{\tau}\int_{t-\tau}^{t} x(u)\,du \tag{9}$$

or, for discrete time series:

$$y_p(n) = \frac{1}{p}\sum_{n-p+1}^{n} x_i \tag{10}$$

If we choose to compute, for instance, the norm $N(f_p)$ - where $f_p = 1/p$ - we search for a dependence of the type:

$$N(f_p) \sim f_p^{dn} \sim p^{-dn}. \tag{11}$$

A sample from the same input signal, processed by the two smoothing methods, is shown in Fig.1. While the obtained exponent is roughly the same for both methods, the correlation of the regression line is always better for the time domain smoothing, pleading for this method.

2. THEORETICAL ASPECTS

An important problem is the link between the smoothing dimensions (*dn* and *di* are treated here) and exponents already often used in practice: the fractal dimension df and the exponent of the power spectrum ß.

Given an ideal LP filter:

$$H_{fc}(\omega) = u(\omega+\omega_c) - u(\omega-\omega_c) \tag{12}$$

where $\omega_c = 2\,\pi f_c$ and:

$$u(x) = \begin{cases} 1, & x>0 \\ 0, & x\le 0\,, \end{cases} \tag{13}$$

the Fourier transform of the filtered signal is:

$$Y_{fc}(\omega) = X(\omega)\,H_{fc}(\omega) \tag{14}$$

with:

$$X(\omega) = \mathcal{F}(x). \tag{15}$$

Considering the theorem of Parseval, Eq.(2) leads to:

$$N_t(y_{fc}) = \frac{1}{\sqrt{2\pi}} \sqrt{\int_0^{\infty} |Y_{fc}(\omega)|^2 \, d\omega} \sim \sqrt{\int_0^{\omega_c} |X(\omega)|^2 \, d\omega} = \sqrt{\int_0^{\omega_c} S(\omega) d\omega} \tag{16}$$

If for the power spectrum $S(\omega)$ the following relation holds:

$$S(\omega) \sim \omega^{\beta} , \tag{17}$$

we have:

$$N_t(y_{fc}) \sim f_c^{\frac{\beta+1}{2}} . \tag{18}$$

From Eqs. (3) and (18) we obtain:

$$dn = \frac{\beta+1}{2} . \tag{19}$$

The relation between the fractal dimension df and the β exponent of the power spectrum[5]:

$$df = \frac{5+\beta}{2} \tag{20}$$

gives, together with Eq. (19):

$$df = 2 + dn . \tag{21}$$

A similar approach is possible for the smoothing dimension di. Eqs. (6) and (14) give:

$$I(y_{fc}) = \int_0^{\omega_c} |X(\omega)| \, d\omega \tag{22}$$

and again, if Eq. (17) holds, then:

$$I(y_{fc}) \sim f_c^{\frac{\beta}{2}+1} . \tag{23}$$

From Eqs. (7) and (23) we have:

$$di = \frac{\beta}{2} + 1 \tag{24}$$

leading to the following relation between di and the fractal dimension df:

$$df = di + 1.5 . \tag{25}$$

3. NUMERICAL EXPERIMENTS AND DISCUSSION

3.1 Applicability ranges

In practice, signal processing implies discretization of the continuous input and numerical treatment of the obtained time series. In this case, the analytical relations must be replaced by discrete ones.

For instance, Eq. (17) becomes:

$$S_k \sim k^{\beta},\ k=1,..,M/2 \tag{26}$$

and the analytical integrals are replaced by cumulative sums, e.g. instead of Eq. (22) we have:

$$I(y_{fc}) = \sum_{k=1}^{k_c} |X_k| . \tag{27}$$

In this context, we must check under what circumstances Eqs. (26) and (27) still lead to a relation like Eq. (18). In order to establish this, we performed numerical experiments.

First, we built the vector:

$$\{ v_k(\beta) \mid v_k = k^{\beta} \},\ k = 1, .., 1024;\ \beta \in [-3,3] \tag{28}$$

and, for a given β, we computed the vector σ corresponding to the cumulative sum:

$$\{ \sigma_k(\beta) \mid \sigma_k = \sum_{i=1}^{k} v_i \},\ k = 1, .., 1024 . \tag{29}$$

We refer to this method to sum up the elements of v as to 'left cumulative sum' (LCS). In terms of the signal spectrum, it is equivalent to a successive LP filtering.

The question arising now is whether, and in what conditions, $\sigma(\beta)$ can be approximated with a power law having the exponent $\beta_C = \beta + 1$. Examples for different values of β are shown in Fig. 2a, where the saturation effect (appearing when β descends under -1) is visible. Fig. 3 also illustrates this problem: one can see that the expected power law behaviour can be found for $\beta > -1$.

Since one is especially interested in the cases for which β has lower values, i.e. $\beta \in [-3,-1]$, some authors studied another way of processing the vector $v(\beta)$, which works in the desired domain but still implies problems[11]. Analytically, if:

$$f(\xi) = \xi^{\beta},\ \beta < -1 , \tag{30}$$

then:

$$\int_{\xi}^{\infty} f(\theta)\, d\theta \sim \xi^{\beta+1} . \tag{31}$$

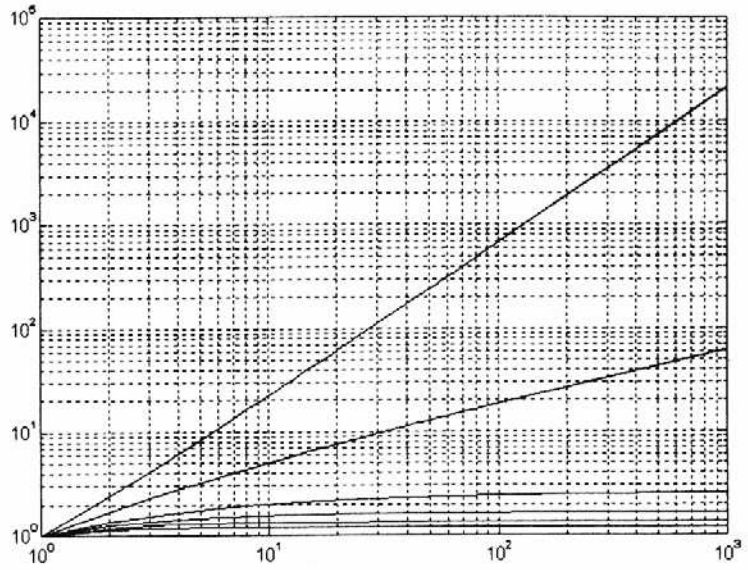
Numerically, we compute $\eta(\beta)$:

$$\{ \eta_k(\beta) \mid \eta_k = \sum_{i=k}^{M} v_i \},\ k = 1, .., M;\ (M = 1024) . \tag{32}$$

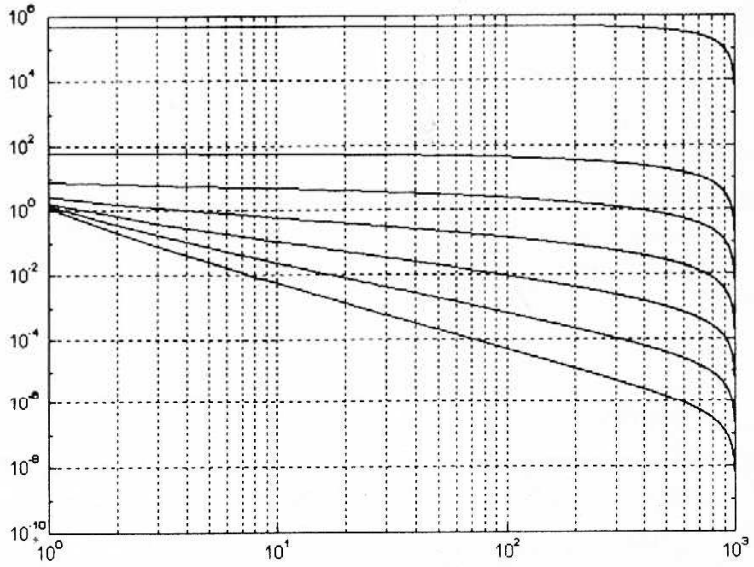
and check the relation:

$$\eta_k(\beta) \sim k^{\beta_C},\ \beta_C = \beta + 1 . \tag{33}$$

We refer to this method (which for a signal spectrum is equivalent to successive highpass filtering steps) as to 'right cumulative sum' (RCS). Examples regarding the application of this procedure on the vector $v(\beta)$ are given in Fig.2b.



(a)



(b)

Fig. 2 Numerical results for the determination of the exponent $ß_c$ by the two methods: (a) LCS. ß having the values -3, -2.5, -2, -1.5, -1, -0.5; (b) RCS. ß having the values -3, -2.5, -2, -1.5, -1, -0.5, 1.

An overview of results obtained with both methods, for $\beta \in [-3,3]$, is given in Fig. 3. One can notice that despite the advantages it offers for $\beta < -1$, the RCS method implies considerable errors, even for low values of β (the regression itself rises delicate problems, with questionable outcomes[11]). In contrast, the LCS method works well for $\beta > -1$, but is completely unusable at lower values.

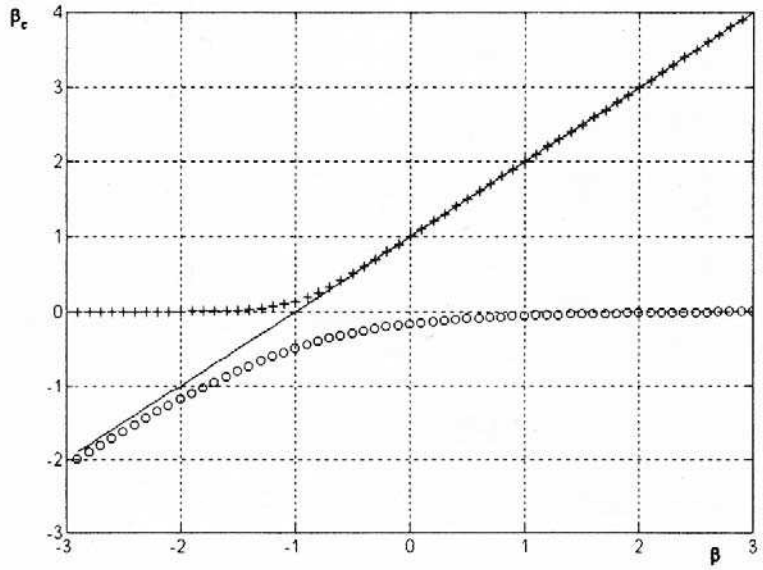


Fig.3 Results regarding the exponent β_C obtained by numerical integration according to the procedure LCS, i.e. from low frequencies up (+), and RCS - from high frequencies down (o).

These observations are important because the method for finding the smoothing dimension (by successive LP filtering steps) automatically implies the use of the LCS method. This means that there are limitations regarding the range in which the smoothing dimensions can be useful when computed according to the procedure exposed so far; the limits are given by:

$$dn: \quad \beta > -1 \tag{34}$$

and

$$di: \quad \beta > -2 . \tag{35}$$

In other words, from this point of view, *di* for instance can be accurately computed in the manner described above only for signals with a fractal dimension $df > 1.5$.

In order to shift the applicability range of the smoothing dimensions, we apply the smoothing process not to the time series x itself, but to its numerical derivative:

$$x' = \{x_k' \mid x_k' = x_{k+1} - x_k ; x_M' = x_1 - x_M\}, \; k = 1, .., M-1 . \tag{36}$$

This is supported by the following analytical results.

The Fourier transform $X'(\omega)$ of the time derivative of $x(t)$:

$$X'(\omega) = j\,\omega\, X(\omega) \tag{37}$$

gives, together with Eq. (17):

$$S'(\omega) = |X'(\omega)|^2 \sim \omega^{\beta+2} = \omega^{\beta'} . \tag{38}$$

Here β' is the exponent of the power spectrum corresponding to the time derivative of $x(t)$:

$$\beta' = \beta + 2 . \tag{39}$$

If we denote by dn' and di' the smoothing dimensions attached to the time derivative of $x(t)$:

$$dn'(x) = dn\left(\frac{dx(t)}{dt}\right) \tag{40}$$

and

$$di'(x) = di\left(\frac{dx(t)}{dt}\right) , \tag{41}$$

then Eqs. (21) and (25) regarding the relations between the fractal dimension df and the smoothing dimensions become:

$$df(x) = 1 + dn'(x) \tag{42}$$

and

$$df(x) = 0.5 + di'(x) . \tag{43}$$

Condition (34) gives - together with the hypothesis (17) and Eqs. (16) and (39) - the new applicability range for the smoothing dimension dn, when this dimension is computed on the derivative of the signal:

$$dn' : \quad \beta > -3 . \tag{44}$$

Likewise, from Eqs. (17), (22), (35) and (39) we have for di:

$$di' : \quad \beta > -4 . \tag{45}$$

3.2 Application on Test Signals

In order to test the signal characterization capability of the smoothing dimensions and their links to fractal exponents, we performed numerical experiments on signals generated with imposed fractal parameters. Here we present results obtained on time series generated following the Weierstraß-Mandelbrot (W-M) algorithm[6] with random phase and for time series with an imposed exponent of the power spectrum[2].

We built W-M time series with the box dimension db comprised between 1 and 2, with a step of 0.05.

Let us first notice that Eq. (16) suggests two methods for computing the smoothing dimension dn in practice.

One way, possibly called 'time domain evaluation', consists in the direct implementation of Eqs. (4) and (3). This implies successive filtering steps (using either of the two smoothing methods presented in Sec. 1) and evaluation, after each step, of the norm attached to the obtained smoothened signal.

A second way - 'frequency domain evaluation' - is to proceed as follows. One computes left cumulative sums on the same power spectrum of the time series [Eq. (16)] and looks for the exponent of the power law according to Eq. (18).

As shown in Fig. 4, both methods applied to W-M time series lead to very close results and allow a rather accurate detection of the fractal dimension df computed according to Eq. (42).

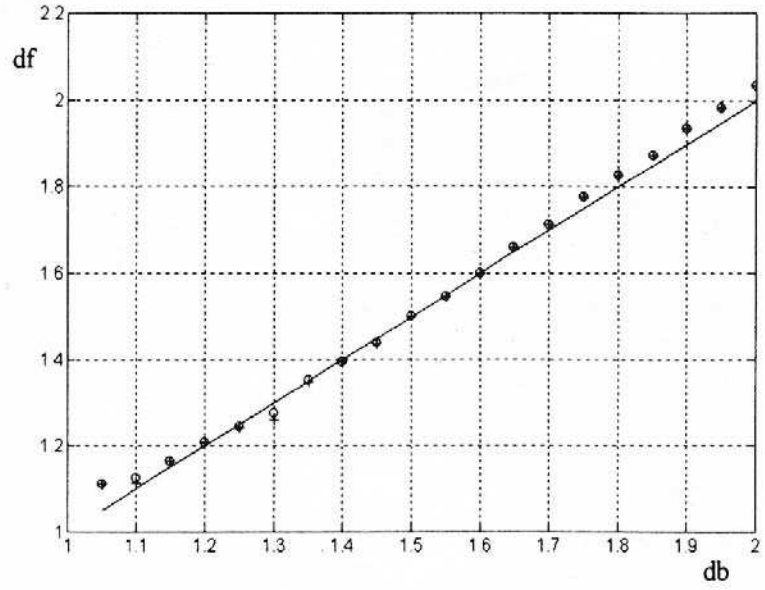


Fig. 4 Fractal dimension *df* determined from the smoothing dimension *dn'* [Eq.(42)] as a function of the imposed box dimension *db* (time series generated according to the W-M algorithm): (+) - time domain evaluation by 20 successive filtering operations with cutoff frequencies distributed according to a geometrical progression; (o) - frequency domain evaluation for the same 20 cutoff frequencies. The solid line represents the ideal (imposed) results.

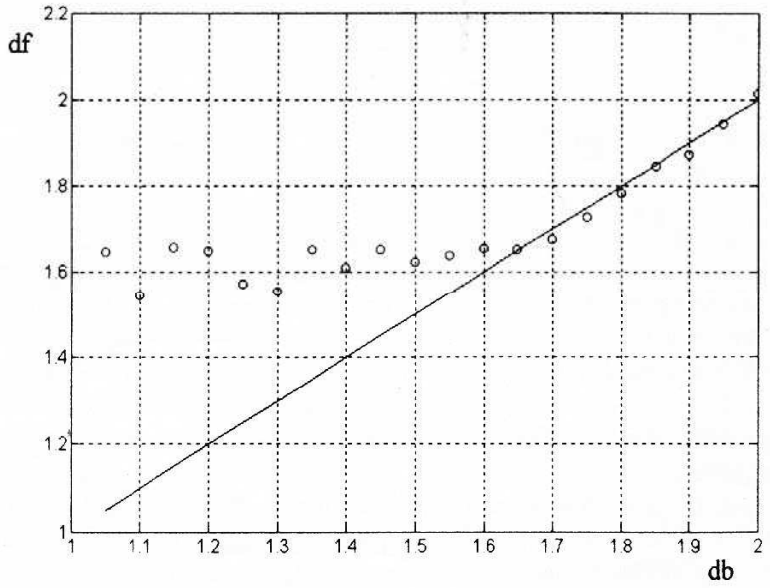


Fig. 5
Fractal dimension *df* deduced directly from the smoothing dimension *di*, without prior derivation of the signal; it is useful - according to Eq. (25) - for fractal dimensions above 1.5.

As stated in Sec. 1, the fractal dimension can also be determined on the original signal *x(t)*: for *di*, a fractal dimension larger than 1.5 is a condition for the accuracy of the determination. This assumption was also checked on W-M time series. The results are shown in Fig. 5.

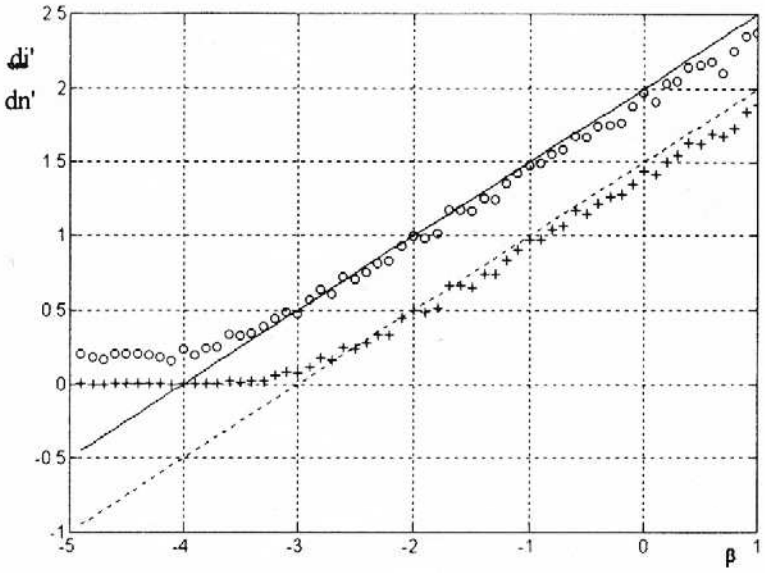


Fig. 6
Determination of the smoothing dimensions dn' (+) and di' (o) as a function of the β exponent of the power spectrum. Straight lines represent the theoretical results. Time series generated according to an algorithm that imposes a power law spectrum.

Another experiment consisted in the determination of the smoothing dimensions dn' and di' for time series generated according to an algorithm that imposes a power law spectrum with a given exponent[2]. One can notice that the validity ranges [Eqs. (44), (45)] are confirmed by these experiments. A slight underestimation is visible on this type of time series.

3.3 'Smoothing' Dimensions vs. 'Roughness' Exponent

As already mentioned, detecting long-range correlations in time series is important in many cases, this task being approached with various methods[12-15]. For instance, a practical problem in experimental research consists in evaluation of time series representing *almost* white noise, where one has to establish, as accurately as possible, if long-range correlations are present[9].

In this regard, the characterization of time series by means of the Hurst exponent H (developed for fractional Gaussian noises and their walks, fractional Brownian motions) is frequently used and well known for its capability to reveal persistence ($H > 0.5$) and antipersistence ($H < 0.5$) as compared to an uncorrelated random signal ($H = 0.5$)[6,16]. Therefore, we are interested in the link between the smoothing dimensions (e.g. dn) and H.

Let us notice from the start that, since H is linked to the fractal dimension df by the relation[6]:

$$df = 2 - H, \tag{46}$$

from Eq. (42) we have the (theoretical) link between H and dn':

$$dn' + H = 1 \tag{47}$$

and from Eq. (43):

$$di' + H = 1.5, \tag{48}$$

where, as discussed above, $dn'(x) = dn(x')$, $di'(x) = di(x')$ and x' is the difference vector of the initial time series x.

A numerical experiment consisted in the determination of H and dn' for time series corresponding to Brownian motions (1000 signals of 4096 samples), which theoretically give H=0.5. The mean values and the standard deviations for H and dn', respectively, were: $\langle H \rangle$=0.529, $\langle dn' \rangle$=0.533, std(H)=0.118, std(dn')=0.038. The overall dynamics (the difference between the maximum and the minimum value) was 0.700 for H and 0.247 for dn'. A sample from the series of determinations on these signals is shown in Fig. 7.

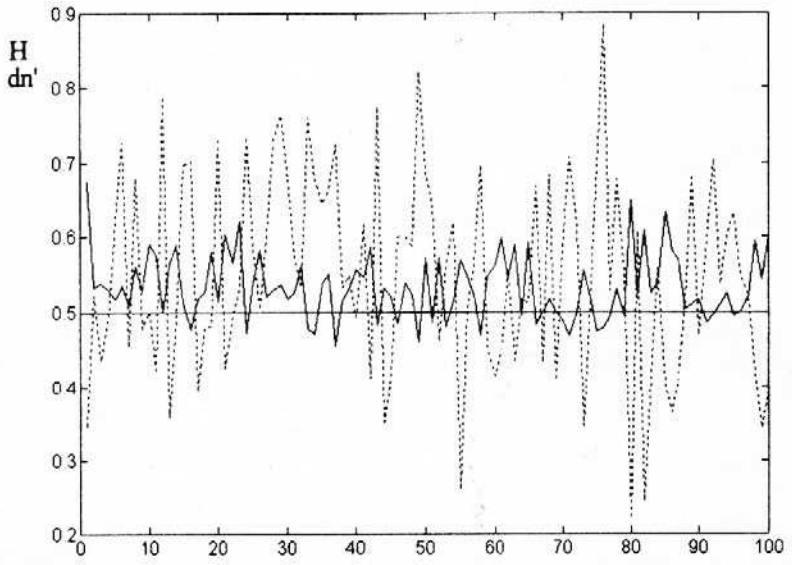


Fig. 7 Values for H (connected with dotted line) and dn' (connected with solid line) determined for time series representing Brownian motions (excerpt from a series of 1000 signals).

The mean values obtained experimentally for H and dn' also support Eq. (47): as one can see in Fig. 7, the ***smoothing*** dimension *(dn')* and the ***roughness***[17] exponent *(H)* are anticorrelated. In fact, a correlation test on the series of 1000 determinations for H and dn' emphasized a correlation coefficient equal to -0.7. The statistical parameters confirm the significantly lower scattering of the smoothing dimension dn' as compared to H, visible in Fig. 7. This feature - particularly important for signal classification - was also found on other types of signals.

A similar experiment was carried out for the smoothing dimension di'. The values for a series of 1000 Brownian motions of 4096 samples were: $\langle di' \rangle$=1.013, std(di')=0.014, overall dynamics:0.088. One can notice that the statistical parameters for di' are almost one order of magnitude better than those for H. They are even better than those for dn'. This could already have been expected, since the determination of dn' implies a supplementary power rising [it operates with the power spectrum - Eq. (16) , while di' uses the frequency spectrum - Eq. (22)], thus enhancing the effects of noise.

Another test consisted in the determination of both dn' and H for time series generated with an imposed exponent of the power spectrum[2], like those presented above. A synthetic image of the results is given in Fig. 8, which confirms the theoretical relation between dn' and H [Eq. (47)]. The higher scattering of the values of H as compared to dn' is striking. On the other hand, one can clearly see the delimitations of the validity ranges, which confirm the theoretical assumptions.

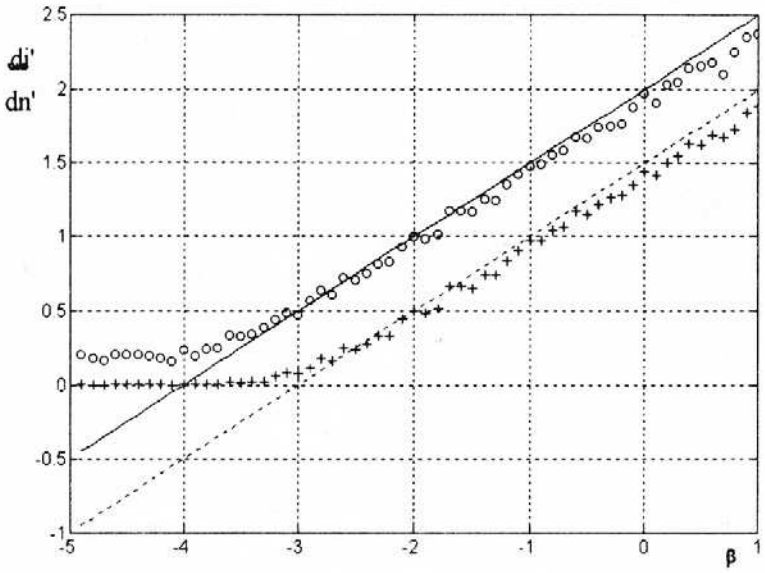


Fig. 6
Determination of the smoothing dimensions *dn'* (+) and *di'* (o) as a function of the β exponent of the power spectrum. Straight lines represent the theoretical results. Time series generated according to an algorithm that imposes a power law spectrum.

Another experiment consisted in the determination of the smoothing dimensions *dn'* and *di'* for time series generated according to an algorithm that imposes a power law spectrum with a given exponent[2]. One can notice that the validity ranges [Eqs. (44), (45)] are confirmed by these experiments. A slight underestimation is visible on this type of time series.

3.3 'Smoothing' Dimensions vs. 'Roughness' Exponent

As already mentioned, detecting long-range correlations in time series is important in many cases, this task being approached with various methods[12-15]. For instance, a practical problem in experimental research consists in evaluation of time series representing *almost* white noise, where one has to establish, as accurately as possible, if long-range correlations are present[9].

In this regard, the characterization of time series by means of the Hurst exponent H (developed for fractional Gaussian noises and their walks, fractional Brownian motions) is frequently used and well known for its capability to reveal persistence ($H > 0.5$) and antipersistence ($H < 0.5$) as compared to an uncorrelated random signal ($H = 0.5$)[6,16]. Therefore, we are interested in the link between the smoothing dimensions (e.g. *dn*) and H.

Let us notice from the start that, since H is linked to the fractal dimension *df* by the relation[6]:

$$df = 2 - H, \tag{46}$$

from Eq. (42) we have the (theoretical) link between H and *dn'*:

$$dn' + H = 1 \tag{47}$$

and from Eq. (43):

$$di' + H = 1.5, \tag{48}$$

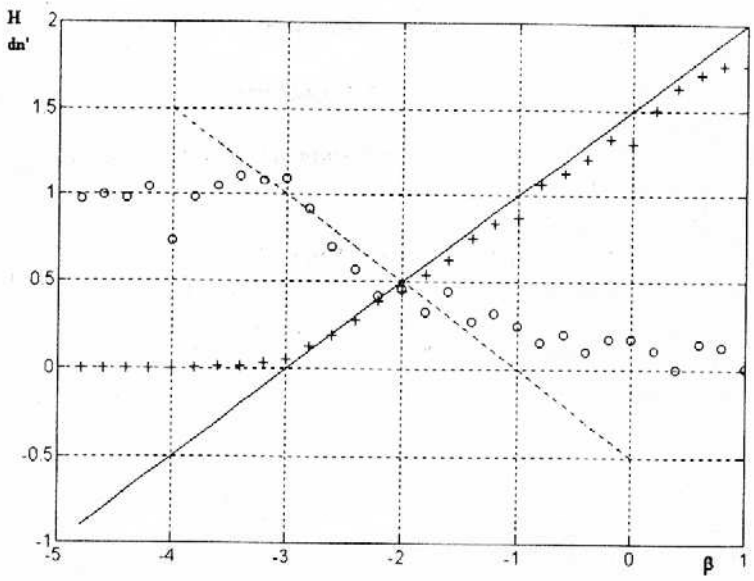


Fig. 8
Smoothing dimension *dn'* (+) and Hurst exponent *H* (o) as a function of the β-exponent of the power spectrum. Experiment on time series with imposed power spectrum[2].

4. CONCLUSIONS

The smoothing dimensions are directly linked to the fractal dimension *df* of the time series (*df* = 1 + *dn'*; *df* = 0.5 + *di'*). While the exponents *dn'* and *di'* are useful for fractal signals with a power spectrum exponent $\beta < -1$, the dimensions *dn* and *di* prove their effectiveness in a range closer to white noise.

They are easy to compute and lead to accurate results confirmed on different types of test signals. The possibility to use two different filtering methods and two evaluation methods increases the reliability of the results.

They also have a good capability to detect long-range correlations and are linked to the Hurst exponent *H* (*H* = 1 - *dn'*; *H* = 1.5 - *di'*). Moreover, numerical experiments performed on various types of signals show that the smoothing dimensions have a significantly lower scattering as compared to the Hurst exponent, an important feature in signal classification problems.

5. ACKNOWLEDGMENTS

The authors like to thank Dr. M. M. Novak for encouraging discussions.
They also acknowledge KENO Ltd., Romania, for the support of this work.

6. REFERENCES

1. B. B. Mandelbrot, *The Fractal Geometry of Nature* (Freeman, New York, 1982).
2. H. Takayasu, *Fractals in the Physical Sciences* (John Wiley and Sons, New York, 1992).