Wavelet-based fractal analysis of airborne pollen

M. E. Degaudenzi and C. M. Arizmendi
Departamento de Física, Facultad de Ingeniería, Universidad Nacional de Mar del Plata, Avenida J.B. Justo 4302, 7600 Mar del Plata, Argentina

(Received 28 August 1998; revised manuscript received 4 December 1998)

The most abundant biological particles in the atmosphere are pollen grains and spores. Self-protection of a pollen allergy is possible through information about future pollen contents in the air. In spite of the importance of airborne pollen concentration forecasting, it has not been possible to predict the pollen concentrations with great accuracy, and about 25% of daily pollen forecasts result in failures. Previous analyses of the dynamic characteristics of atmospheric pollen time series indicate that the system can be described by a low dimensional chaotic map. We apply a wavelet transform to study the multifractal characteristics of an airborne pollen time series. The information and the correlation dimensions correspond to a chaotic system showing a loss of information with time evolution. [S1063-651X(99)00606-6]

PACS number(s): 05.45.−a, 87.10.+e

I. INTRODUCTION

A pollen allergy is a common disease causing hay fever in 5–10 % of the population. Although not a life threatening disease, the symptoms can be very troublesome; furthermore, the costs to the social sector due to pollen related diseases are high. Self-protection of hay fever patients is possible through information about future pollen contents in the air [1].

Models to forecast pollen concentration in the air are principally based on pollen and atmospheric weather interactions. Several statistical techniques [2–4] have been used to predict future atmospheric pollen concentrations from weather conditions of the day and recent previous days. In spite of these attempts, it has not been possible to predict the pollen concentrations with great accuracy, and about 25% of the daily pollen forecasts have resulted in failures [4]. A reason for these failures could be that the methods used in airborne pollen forecasting are based on standard linear statistical techniques which are not suitable when the phenomenon to be forecasted is esentially nonlinear.

A previous analysis of the dynamic characteristics of a time series of atmospheric pollen was developed in Ref. [5], through the study of the correlation dimension [6,7]. The dimension found was of a low and noninteger value [5], which indicates that the system may be described by a nonlinear function of just a few variables relating the nearest pollen concentrations of the time series. The fact that the correlation dimension found was fractal predicts that this function, also called a map in nonlinear dynamics, can display chaotic behavior under certain circumstances. The existence of a low dimensional map suggests possibilities for short-term prediction [8] through the use of some nonlinear model. Artificial neural networks have been widely used to predict future values of chaotic time series identifying the nonlinear model by extracting knowledge from the past [9]. Very good pollen concentration forecasts were obtained using neural networks [10] and, in a previous work, the hypothesis that random fluctuations appearing in the pollen time series are produced by Gaussian noise was rejected [11].

To continue with the characterization of airborne pollen concentrations, the next step would be to characterize them as a multifractal. A very efficient method to obtain the $f(\alpha)$ singularity spectrum of a pollen time series relies on the use of a mathematical tool introduced in the early 1980's in signal analysis; the *wavelet transform*. The wavelet transform has been proved very efficient to detect singularities and to prove that fractals are indeed singular functions. Muzy, Bacry, and Arneodo [12,13] developed the *wavelet transform modulus maxima* (WTMM) method as a technique to study fractal objects. In this method the wavelet is used as an oscillating variant of the ''square'' function of a box. The WTMM method was succesfully applied to study fractal properties of diverse systems such as DNA nucleotide sequences [14,15], Modane turbulent velocity signal [12,16], and a cool flame experiment [17]. We apply the WTMM method to obtain the generalized fractal dimensions D_q associated with the pollen time series.

II. EXPERIMENTAL SETUP

The material used in this work was from our chaos study of pollen series [5]. Data of airborne pollen concentration were obtained with an automatic and volumetric Burkard pollen and spore trap, situated at the roof of the Facultad de Ciencias Exactas y Naturales of our University, 12 m above ground level. The area surrounding the sample is typical of Mar del Plata. The great distance from the sampling site to the emission sources makes the particular emission spectra unimportant.

Ten liters of air per minute were sucked through a 14 $\times 2$-mm^2 orifice, always orientated against the wind flow. The sucking rate is checked weekly. Behind the slit, a drum rotates at a speed of 2 mm per hour. The particles are collected on a cellophane tape (Melinex), 19 mm wide, just below the orifice. The sticky collecting surface comprises of nine parts vaseline, and one part paraffin in toluene. The exposed tape is removed from the drum, cut into pieces of 48 mm, corresponding to 24-h intervals, then embedded into a solution of polivinylalcohol (Gelvatol), water, and glycerol, and covered with a cover glass. Slides were studied as 12

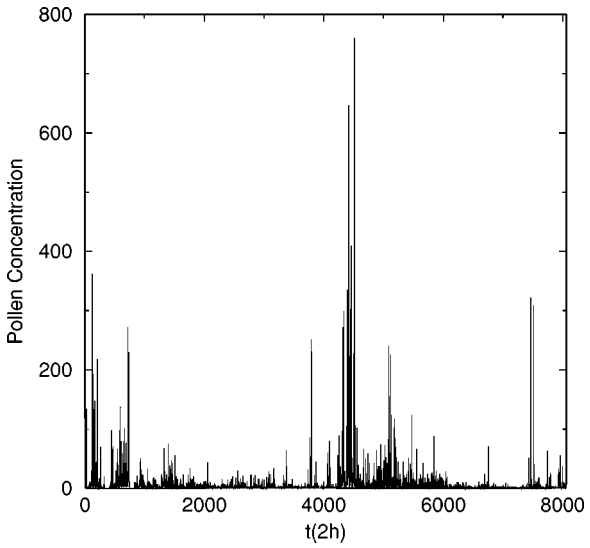

FIG. 1. Two years of an airborne pollen concentration time series. The time step is 2 h. The units of pollen concentration are pollen grains.

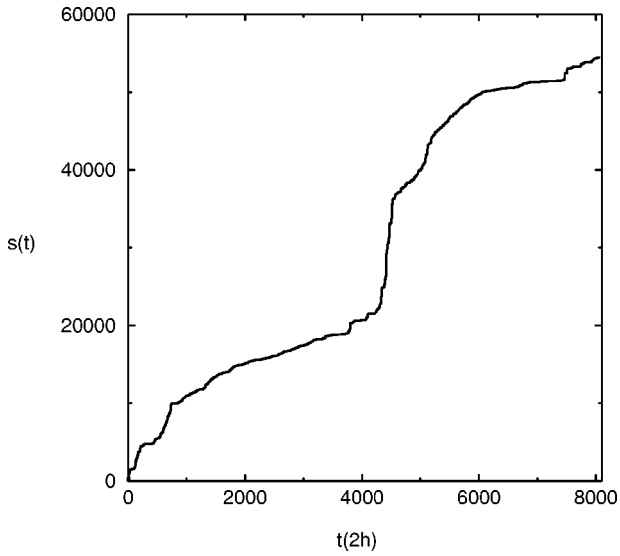

FIG. 2. Distribution function associated with a pollen concentration time series. The WTMM method was applied to this function. The units of the distribution function are pollen grains times hours.

transects per day. The pollen was counted at a magnification of $\times 400$ for the first year cycle (August 1987-1988) and at $\times 200$ for the second (August 1988-1989), and corresponding to 13.5 and 27 min of sampling every 2 h, respectively. The method of counting pollen follows that of Käpylä and Penttinen [18]. Hourly counts were stored in a database file for further analysis. Statistics of hourly counts may be seen in Tables 1 and 2 of Ref. [5]. The concentration values correspond to total pollen grains. The main species found were Cupressus, Gramineae, Eucalyptus, Pinace, Chenopodiineae, Plantago, Cyperaceae, Betula, Cruciferae, compositae Tueulflorae, Ambrosia, Ulmus, Umbelliferae, Platanus, and Fraxinus.

III. MULTIFRACTAL FORMALISM

The aim of this formalism is to determinate the $f(\alpha)$ singularity spectrum of a measure μ . It associates the Hausdorff dimension of each point with the singularity exponent α, which gives us an idea of the strength of the singularity,

$$N_\alpha(\epsilon) \sim \epsilon^{-f(\alpha)}, \tag{1}$$

where N_ϵ is the number of boxes needed to cover the measure, and ϵ is the size of each box [19].

A partition function Z can be defined from this spectrum (it is the same model as the thermodynamic one),

$$Z(q,\epsilon) = \sum_{i=1}^{N(\epsilon)} \mu_i^q(\epsilon) \sim \epsilon^{\tau(q)} \quad \text{for} \quad \epsilon \to 0, \tag{2}$$

where $\tau(q)$ is a spectrum which arouses by Legendre transforming the $f(\alpha)$ singularity spectrum.

The spectrum of *generalized fractal dimensions* D_q is obtained from the spectrum $\tau(q)$

$$D_q = \frac{\tau(q)}{(q-1)}. \tag{3}$$

The capacity or box dimension of the support of the distribution is given by $D_0 = f(\alpha(0)) = -\tau(0)$. $D_1 = f(\alpha(1)) = \alpha(1)$ corresponds to the scaling behavior of the information, and is called the *information dimension*. For $q \geqslant 2$, D_q and the *q-point correlation integrals* are related. As we will show in Sec. IV, the wavelet transform is specially suited to analyze a time series as a multifractal.

IV. WAVELET TRANSFORM

The wavelet transform (WT) [20,21] of a signal $s(t)$ consists of decomposing it into frequency and time coefficients, asociated to the wavelets. The analyzing wavelet ψ, by means of translations and dilations, generates the so-called family of wavelets.

The wavelet transform turns the signal $s(t)$ into a function $T_\psi[s](a,b)$:

$$T_\psi[s](a,b) = \frac{1}{a} \int \psi^* \left(\frac{t-b}{a} \right) s(t)\, dt, \tag{4}$$

where ψ^* is the complex conjugate of ψ, a the frequency dilation factor, and b the time translation parameter.

The wavelet to apply must be chosen with the condition

$$\int \psi(t)\, dt = 0, \tag{5}$$

and to be orthogonal to lower-order polynomials,

$$\int t^m \psi(t)\, dt = 0, \quad 0 \leqslant m \leqslant n, \tag{6}$$

where m is the order of the polynomial. In other words, lower- order polynomial behavior is eliminated, and we can detect and characterize singularities even if they are masked by a smooth behavior.

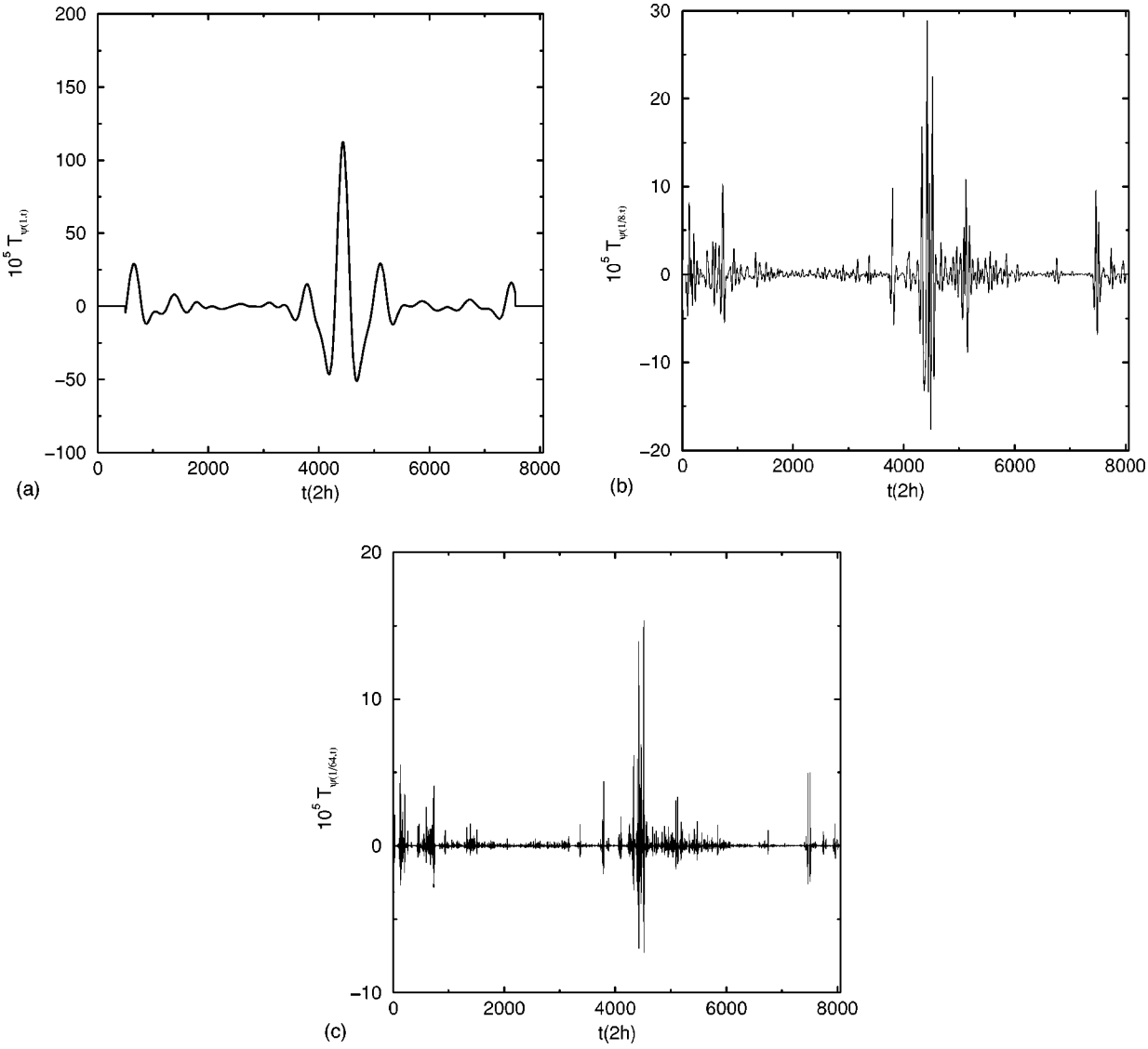

FIG. 3. Wavelet transform data of pollen time series distribution: (a) scale $a=1$, (b) scale $a=\frac{1}{8}$, and (c) scale $a=\frac{1}{64}$. The units are the same as the distribution function, pollen grains times hours.

The WT provides a useful tool in the detection of self-similarity or self-affinity in temporal series. For a value b in the domain of the signal, the modulus of the transform is maximized when the frequency a is of the same order of the characteristic frequency of the signal $s(t)$ in the neighborhood of b; this last one will have a local singularity exponent $\alpha(b) \in]n,n+1[$.

This means that, around b,

$$|s(t)-P_n(t)| \sim |t-b|^{\alpha(b)}, \tag{7}$$

where $P_n(t)$ is an n-order polynomial, and

$$T_\psi(a,b) \sim a^{\alpha(b)}, \tag{8}$$

provided the first $n+1$ moments are zero. If we have $\psi^{(N)} = d^{(N)}(e^{x^2/2})/dx^N$, the first N moments are vanishing.

The wavelet modulus function $|T_\psi[s](a,t)|$ will have a local maximum around the points where the signal is singular. These local maxima points make a geometric place called modulus maxima line $\mathcal{L}$.

$$|T_\psi[s](a,b_l(a))| \sim a^\alpha(b_l(a)) \quad \text{for} \quad a \to 0, \tag{9}$$

where $b_l(a)$ is the position at the scale a of the maximum belonging to the the line $\mathcal{L}$.

The wavelet transform modulus maxima method consists of an analysis of the scaling behavior of some partition functions $Z(q,a)$ that can be defined as

$$Z(q,a) = \sum |T_\psi[s](a,b_l(a))|^q, \tag{10}$$

and will scale as $a^{\tau(q)}$ [12,13]. This partition function works like the previously defined partition function for singular measures. For $q>0$ the most pronounced modulus maxima will prevail, and, on the other hand, for $q<0$ the lower ones will survive. The most pronounced modulus take place when very deep singularities are detected, while the others correspond to smoother singularities. We can obtain $\tau(q)$ [Eq. (2)] and $f(\alpha)$ and D_q spectra, as explained previously. The shape of $f(\alpha)$ is a hump that has a maximum value. The generalized fractal dimensions D_q are meaningful for mea-

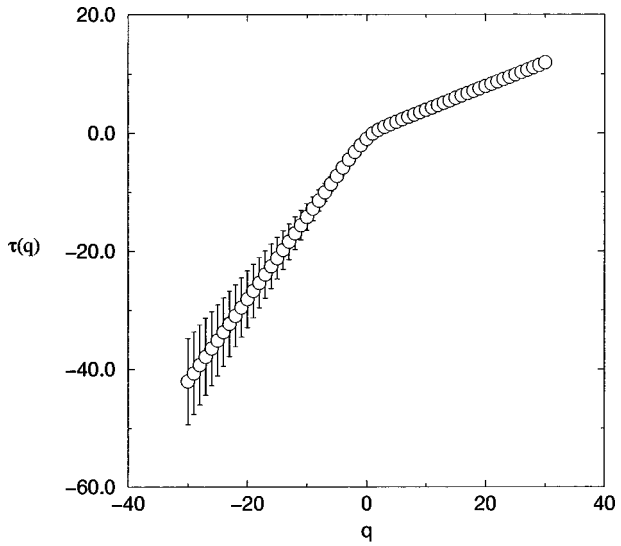

FIG. 4. $\tau(q)$ spectrum of a pollen time series.

sures only. They do not have any meaning for general functions. The pollen time series is a singular measure.

V. APPLICATION OF WTMM METHOD TO THE POLLEN TIME SERIES

The airborne pollen concentration time series may be seen in Fig. 1. The way to deal with singular measures μ as the pollen time series is to work with its corresponding distribution functions [i.e., $f(x)=\mu([0,x])$], because their singular behavior is given by the singularities of their associated singular measures [12]. The distribution function associated with the pollen time series is shown in Fig. 2.

The third derivative of the Gaussian function was chosen as the analyzing wavelet,

$$\psi^{(3)}(t)=\frac{d^3}{dt^3}(e^{t^2/2}), \tag{11}$$

Twelve wavelet transform data files were obtained by applying the wavelet transform with $\psi^{(3)}$, ranging the scaling factor a from $a_{\min}=1/256$ to $a_{\max}=8$ in steps of 2^n. To give an idea of the effect of the change of scale on the wavelet transform of the pollen time series, three of them are shown in Fig. 3.

We computed the partition function $Z(q,a)$ for $-30\leqslant q\leqslant 30$ and $1/256\leqslant a\leqslant 8$, obtaining $\tau(q)$, as shown in Fig. 4. $\tau(q)$ is a nonlinear convex increasing function with $\tau(0)=-0.97\pm 0.15$ and two asymptotic slopes which are $\alpha_{\min}=0.40\pm 0.12$ for $q>0$ and $\alpha_{\max}=1.39\pm 0.33$ for $q<0$.

This lays the corresponding $f(\alpha)$ singularity spectrum obtained by Legendre transforming $\tau(q)$ for $-2\leqslant q\leqslant 4$, that is displayed in Fig. 5. The single humped shape with a nonunique Hölder exponent obtained characterizes a multifractal.

The D_q spectrum obtained from $\tau(q)$ can be seen in Fig. 6. The support dimension $D_o=D_{\max}=-\tau(0)=0.97\pm 0.15$, which implies that the capacity of the support is approximately 1; i.e., the support is not a fractal. D_q converges asymptotically to $D_\infty=0.40\pm 0.04$ for $q_{\max}$, and to $D_{-\infty}=1.38\pm 0.24$ for $q_{\min}$. The minimum value D_∞ corresponds to the strongest singularity which characterizes the most rarified zone, whereas higher values exhibit weaker singularities until $D_{-\infty}$, i.e., the weakest singularity corresponds to the densest zone.

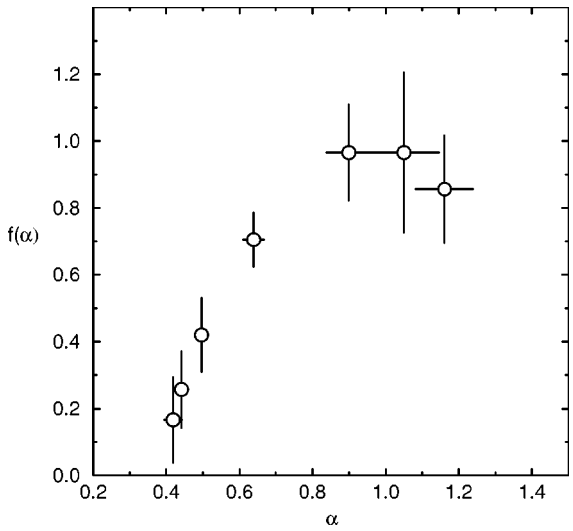

FIG. 5. $f(\alpha)$ spectrum of a pollen time series $-2\leqslant q\leqslant 4$.

The information dimension is $D_1=f(\alpha(1))=f(0.68)=0.68\pm 0.08$, which features the scaling behavior of the information. It plays an important role in the analysis of nonlinear dynamic systems, especially in describing the loss of information as chaotic system evolves in time [22]. $D_1=0.68\pm 0.08$ implies that we are in the presence of a chaotic system. The correlation dimension is $D_2=0.57\pm 0.12$, which characterizes a chaotic attractor and is very close to the value obtained previously with the Grassberger-Procaccia method [5].

VI. CONCLUSION

The wavelet transform modulus maxima method was applied to study the multifractal characteristics of an airborne pollen time series. Previous analyses of the dynamic charac-

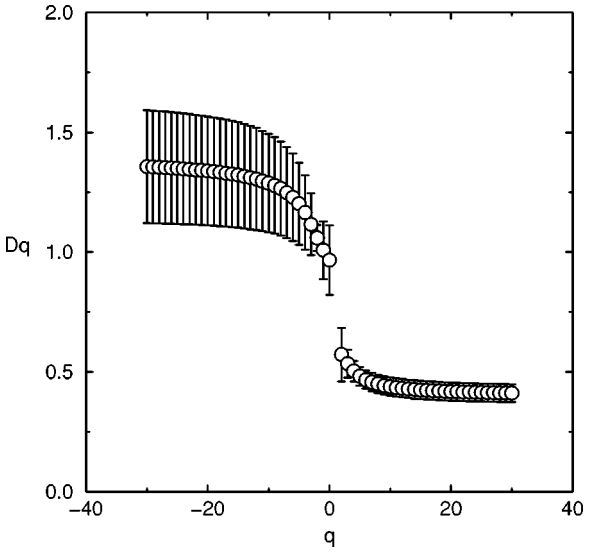

FIG. 6. D_q spectrum of a pollen time series. The support dimension $D_o=D(q=0)\sim 1$. D_q converges asymptotically to $D_{-\infty}=0.40\pm 0.04$ for $q_{\max}$, and to $D_\infty=1.38\pm 0.24$ for $q_{\min}$.

teristics of atmospheric pollen time series indicate that the system can be described by a low dimensional chaotic map. The full complexity of the scaling structure of the strange attractor associated with the airborne pollen dynamics is more conveniently reflected by the spectrum of generalized dimensions D_q. The information and the correlation dimensions correspond to a chaotic system showing a loss of information with time evolution. The characterization of the airborne pollen dynamics is important in order to improve airborne pollen forecasting.

ACKNOWLEDGMENTS

C.M.A. would like to thank Alain Arneodo for introducing him to wavelet transform multifractal analysis. This work was partially supported by a grant from the Universidad Nacional de Mar del Plata.

[1] R. Leuschner, J. Palyno **27**, 305 (1991).
[2] P. Comtois, G. Batchelder, and D. Sherknies, in *Aerobiology Health Environment: A Symposium*, edited by P. Comtois (University of Montreal Press, Montreal, 1989).
[3] M. O'Rourke, in *Aerobiology Health Environment: A Symposium* (Ref. [2]); L. Moseholm, E. Weeke, and B. Petersen, Pollen et Spores **29**, 305 (1987).
[4] C. Goldberg, H. Buch, L. Moseholm, and E. Weeke, Grana **27**, 209 (1988).
[5] M. M. Bianchi, C. M. Arizmendi, and J. R. Sanchez, Int. J. Biometeorol. **36**, 172 (1992).
[6] P. Grassberger and I. Proccacia, Phys. Rev. Lett. **50**, 346 (1983).
[7] N. B. Abraham, A. M. Albano, B. Das, G. De Guzman, S. Young, R. S. Giorggia, G. P. Puccioni, and J. R. Tredicce, Phys. Lett. **114A**, 217 (1986).
[8] J. D. Farmer and J. J. Sidorowich, Phys. Rev. Lett. **59**, 845 (1987).
[9] A. Lapedes and R. Farber (unpublished).
[10] C. M. Arizmendi, J. R. Sanchez, N. E. Ramos, and G. I. Ramos, Int. J. Biometeorol. **37**, 139 (1993).
[11] C. M. Arizmendi, J. R. Sanchez, and M. A. Foti, Fractals **3**, 155 (1995).
[12] J. F. Muzy, E. Bacry, and A. Arneodo, Int. J. Bifurcation Chaos Appl. Sci. Eng. **4**, 245 (1994).
[13] J. F. Muzy, E. Bacry, and A. Arneodo, Phys. Rev. E **47**, 875 (1993).
[14] A. Arneodo, Y. d'Aubenton-Carafa, E. Bacry, P. V. Graves, J. F. Muzy, and C. Thermes, Physica D **96**, 291 (1996).
[15] A. Arneodo, E. Bacry, P. V. Graves, and J. F. Muzy, Phys. Rev. Lett. **74**, 3293 (1995).
[16] A. Arneodo, E. Bacry, and J. F. Muzy, Physica A **213**, 232 (1995).
[17] M. Nicollet, A. Lemarchand, and G. M. L. Dumas, Fractals **5**, 35 (1997).
[18] M. Käpylä and A. Penttinen, Grana **20**, 131 (1981).
[19] Hëinz-Otto Peitgen, H. Jurgens, and D. Saupe, *Chaos and Fractals, New Frontiers of Science* (Springer-Verlag, New York, 1992).
[20] P. Goupillaud, A. Grossman, and J. Morlet, Geoexploration **23**, 85 (1984).
[21] A. Grossman and J. Morlet, SIAM (Soc. Ind. Appl. Math.) J. Math. Anal. **15**, 723 (1984); and in *Mathematics and Physics, Lectures on Recent Results*, edited by L. Streit (World Scientific, Singapore, 1985).
[22] M. Schröeder, *Fractal, Chaos, Power Laws* (Freeman, San Francisco, 1991).