

A deterministic geometric representation of temporal rainfall: Results for a storm in Boston

Carlos E. Puente and Nelson Obregón

Hydrologic Sciences, Department of Land, Air and Water Resources, University of California, Davis

Abstract. The use of a deterministic fractal-multifractal (FM) representation to model high-resolution rainfall time series via projections of fractal interpolating functions weighed by multifractal measures is reported. It is shown that the intrinsic shape and variability of an 8-hour Boston storm recorded every 15 s on October 25, 1980, may be encoded wholistically, employing the fractal geometric methodology. It is illustrated that the FM methodology provides very faithful descriptions of both major trends and small (noisy) fluctuations for this storm, resulting in preservation of not only classical statistical characteristics of the records but also multifractal and chaotic properties present in them. These results, and those for other storms, suggest that a stochastic framework for rainfall may be bypassed in favor of a deterministic representation based on projections.

1. Introduction

Modeling the structure of temporal rainfall has attracted much attention in the literature for the last 30 years. Although several sophisticated (deterministic and stochastic) rainfall models exist, they are built to preserve some characteristics of the records and consequently do not capture all the variability observed at a fixed location when a storm passes by. Typically, these models preserve some statistics of the available data set but do not keep the intermittent details present in the records. Since knowing details may be important in predicting the rainfall process, an alternative approach to temporal rainfall should be considered.

Existing temporal rainfall representations may be classified depending on their basic building blocks, as follows: (1) physically based, (2) point process, (3) chaotic, and (4) fractal geometric. Physically based models approximate the physical processes giving rise to the observed rainfall [e.g., *Georgakakos and Bras*, 1984]. By employing basic thermodynamics, cloud microphysics principles, and integration of the cloud on top of a recording station, these representations produce intermittent outcomes which nonetheless do not capture the inherent details present in the records. This happens, in part, because of approximated parameterizations which reflect our lack of understanding of all the processes taking place within the atmosphere.

Typical stochastic point process models approximate the irregular and complex rain patterns (of a fractal and/or multifractal nature) by superimposing randomly arriving Euclidean objects (e.g., rectangular pulses). Models with Poisson and cluster-based arrival processes have been defined such that some statistical features of the rainfall time series (e.g., mean, variance, autocorrelations, etc.) are preserved (see, for instance, *Eagleson* [1978]; *Kavvas and Delleur* [1981]; *Smith and Karr* [1983]; and *Rodriguez-Iturbe* [1986], among others). Although it has been found that cluster-based procedures are better than Poisson-based ones [e.g., *Valdes et al.*, 1985; *Rodriguez-Iturbe et al.*, 1987], inconsistencies have been ob-

served in cluster models regarding their ability to have a unique set of parameters irrespective of the data's degree of aggregation [*Foufoula-Georgiou and Guttorp*, 1986]. Despite providing a theoretical framework which includes beautiful ideas and models, these representations are typically limited by their analytical tractability due to the simplifying assumptions which they require, for example, stationarity, ergodicity, etc. Also, the geometric outcomes they produce do not fully capture the very complex intermittency patterns observed in rainfall.

The possibility that rainfall could be understood as a low-dimensional chaotic process was introduced by *Rodriguez-Iturbe et al.* [1989]. Even though determinism becomes an important component of this approach, actual dynamic modeling (i.e., finding the equations of motion from a time series) is quite difficult [e.g., *Crutchfield and McNamara*, 1987; *Casdagli*, 1989] and limits the practical applicability of the otherwise beautiful methodology. The typical length of available storm records also precludes a complete analysis following such ideas. In any event, whether or not there is a climatic attractor remains a relevant topic of research [e.g., *Tsonis and Elsner*, 1989].

A stochastic fractal representation of rainfall was introduced by *Lovejoy and Schertzer* [1985, 1990] and *Schertzer and Lovejoy* [1987] via the notion of universal multifractals. Their idea is to represent rainfall as a realization of a Levy process and parameterize it via its codimension function (basically the left portion of the multifractal spectrum [e.g., *Feder*, 1988; C. E. Puente et al., Deterministic multifractals with negative dimensions?, submitted to *Physical Review E*, 1996]). Success has been attained in characterizing such a function for alternative data sets [e.g., *Lovejoy and Schertzer*, 1990]. Even though reasonable looking simulations, having intermittency as found in rainfall, may be obtained, it is difficult to find conditional simulations with such an approach. It is also pertinent to note that the codimension function (even when compactly parameterized), as well as other classical statistical characteristics like mean, variance, and autocorrelation, do not really characterize a given data set. In fact, having the codimension function preserved does not imply the capturing of the actual locations of the details present in the records.

Copyright 1996 by the American Geophysical Union.

Paper number 96WR01466.
0043-1397/96/96WR-01466\$09.00

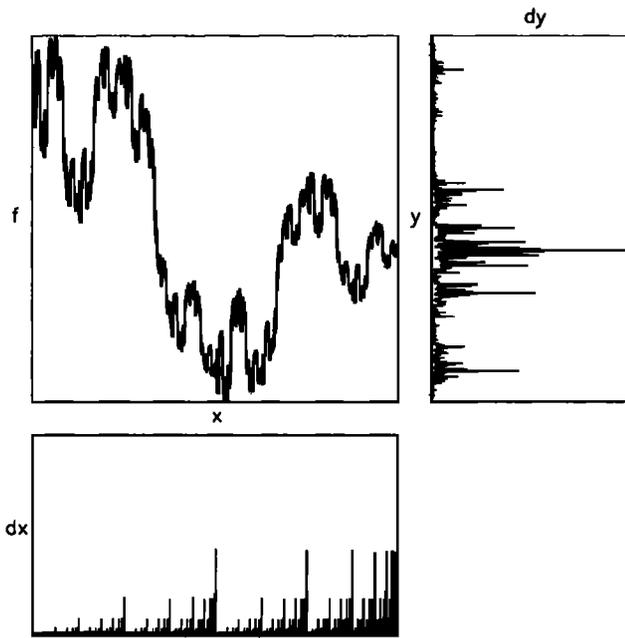


Figure 1. The fractal-multifractal framework in two dimensions; $dy = dx \circ f^{-1}$.

In this work, a new procedure for the quantification of hydrologic (geophysical) phenomena is reviewed and a particular application to temporal rainfall for a high-resolution storm is given. The procedure to be used is the fractal-multifractal representation (FM), as introduced by Puente [1992, 1994]. The basis for developing such an approach is the fact that latest developments in physics recognize (1) the relevance of details in our ability to predict and (2) the possibility of describing apparently random sets by means of simple deterministic rules [e.g., Moon, 1987].

The idea behind the FM approach is to think of the complex, jagged, and intricate hydrologic patterns as projections of fractal functions which are “illuminated” via simple multifractal measures. An important trait of the FM approach is that it is entirely deterministic. Also, it does not require any statistical assumptions such as stationarity or ergodicity or a minimal length for the records under study. As will be illustrated, the FM representation results in “random-looking” outcomes which are not random at all and which resemble actual rainfall (geophysical) records, not only in their appearance but also in their statistical, multifractal, and chaotic properties. This last point will be shown finding a FM approximation of an 8-hour storm recorded in Boston every 15 s.

2. The Fractal-Multifractal Approach

A large number of deterministic measures may be obtained using the FM methodology [Puente, 1992, 1994]. These measures dy are defined following a classical derived distributions approach, using a generic multifractal dx as the parent distribution and a fractal interpolating function f as the transformation ($dy = dx \circ f^{-1}$). The two components that make up the construction, dx and f , and how they are combined are reviewed next.

The classical cascade model in turbulence gives rise to the most generic multifractal measures. The binomial multiplicative process results in a binomial multifractal measure as fol-

lows. Begin with a uniform distribution over an interval I , say, $[0, 1]$, and select an intermittency (redistribution) parameter p , $0 < p < 1$. Then, redistribute the mass such that in the interval $[0, 1/2]$, $p\%$ of the mass is uniformly distributed and likewise the remaining $(1 - p)\%$ in the interval $[1/2, 1]$. Repeat this same process on each of the uniform pieces, ad infinitum, to arrive at a deterministic binomial multifractal measure with parameter p . A sketch of such a measure, dx , is shown in the bottom of Figure 1 for $p = 0.3$ and for 13 stages in the cascade. As is seen, the measure dx is multifractal, i.e., it is singular and contains (in the limit) infinitely many layers which correspond to intertwined Cantor sets in x . General deterministic multifractals may be constructed by splitting the mass into more than two pieces and by using different length scales [Mandelbrot, 1989].

The set of transformations used in the context of derived distributions are the fractal interpolation functions introduced by Barnsley [1986, 1988]. These are continuous functions f that interpolate a given set of $N + 1$ points in the plane $\{(x_0, y_0), (x_1, y_1), \dots, (x_N, y_N)\}$; $x_0 < x_1 < \dots < x_N$ and whose graphs may be fractal. They are obtained iterating N contractile affine mappings w_n such that $G = \cup_{n=1}^N w_n(G)$, where $G = \{(x, f(x)): x \in [x_0, x_N]\}$ is the graph of f . The following are the mappings used and the conditions that guarantee the existence of a unique (and hence deterministic) set G :

$$w_n \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a_n & 0 \\ c_n & d_n \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e_n \\ f_n \end{pmatrix}, \quad (1)$$

such that

$$w_n \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} x_{n-1} \\ y_{n-1} \end{pmatrix}, \quad w_n \begin{pmatrix} x_N \\ y_N \end{pmatrix} = \begin{pmatrix} x_n \\ y_n \end{pmatrix}, \quad (2)$$

and $0 \leq |d_n| < 1$, for $n = 1, \dots, N$.

These previous equations allow solving for the parameters $a_n, c_n, e_n,$ and f_n in terms of the interpolating points' coordinates and the “free” scaling parameters d_n . The fractal dimension D of the graph of an interpolating function is (1) $D \geq 1$, the solution of $\sum |d_n| a_n^{D-1} = 1$, if $\sum |d_n| > 1$, and (2) 1 if $\sum |d_n| \leq 1$ [Barnsley, 1988]. Figure 1 includes an example of a

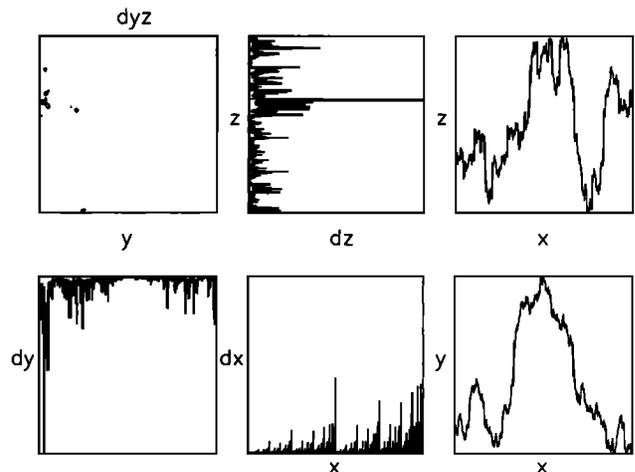


Figure 2. Three-dimensional fractal interpolator projections and derived measures: $\{(0, 0.1, 0), (0.5, 1, 0.4), (1, 0, 0.2)\}$. Here $r_1^{(1)} = r_1^{(2)} = -0.6$, $\theta_1^{(1)} = \theta_1^{(2)} = 45$, $r_2^{(1)} = 0.6$, $r_2^{(2)} = -0.6$, $\theta_2^{(1)} = \theta_2^{(2)} = 45$, $p_1 = 0.3$, $p_2 = 0.7$. Angles are in degrees.

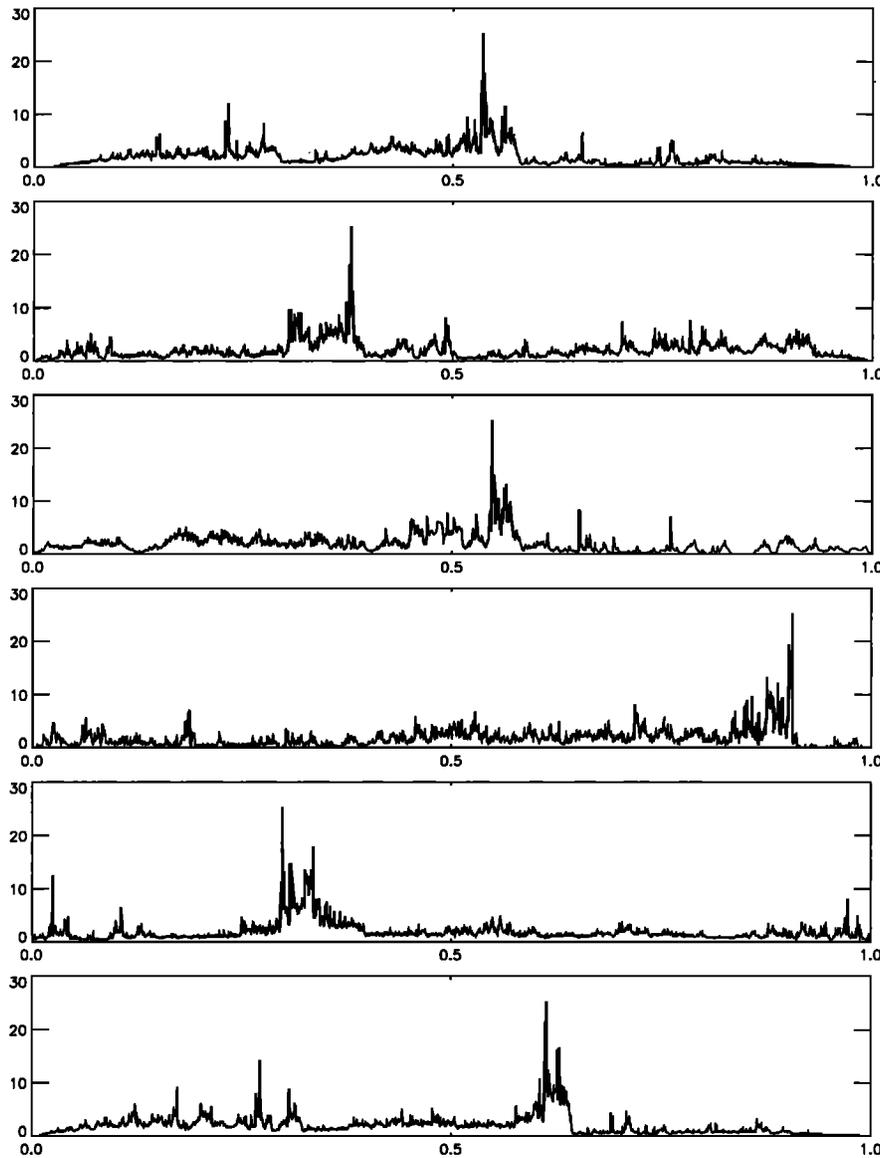


Figure 3. Projections from two-dimensional fractal interpolators and a storm in Boston.

fractal interpolating function f that passes by the three data points $\{(0, 0), (0.5, -0.35), (1, -0.2)\}$ shown by the solid dot and which has $d_1 = -0.8$ and $d_2 = -0.6$. Observe that the graph G is a self-affine fractal set with dimension $D = 1.48$. For details on how to generate the fractal interpolating functions in practice, see *Barnsley [1988]* and *Puente [1994]*.

Figure 1 also illustrates how the derived distributions approach is used to transform a binomial multifractal measure dx via a fractal interpolating function f , in order to determine a derived distribution dy . Mathematically, dy is defined considering all relevant parent events in x and adding their measures, i.e., $dy(B) = dx(f^{-1}(B)) = dx\{x: f(x) \in B\}$, for a Borel subset B . The measure dy can be interpreted as a weighted projection of the function f , with the weights given by dx .

A great variety of derived measures are obtained by varying the parameters of f and dx [*Puente, 1992, 1994*]. Depending on the nature of the fractal interpolating function, the following overall behavior is found. When the fractal dimension D is close to one, the derived measures are (1) singular (i.e., mul-

tifractal) and (2) not strictly self-similar nor self-affine. As D grows from one to two, the measures (1) progressively become absolutely continuous (i.e., have a density) and (2) in the limit become Gaussian.

Given the relevance of multinomial multifractals to represent intermittent natural phenomena, [e.g., *Meneveau and Sreenivasan, 1987; Sreenivasan, 1991*], the derived measures have a physical interpretation: They could be thought of as images or “projections” of turbulence. The seemingly random appearance and the complex intermittency of the measures dy (e.g., see Figure 1) represents the basis for trying to use the FM approach to model complex series, as the ones given by rainfall. It is worth emphasizing that having dy not strictly self-similar or self-affine is a welcomed property, since a common objection against the use of fractal geometry has been that natural objects do not exhibit such geometric behavior ad infinitum.

It is important to stress some merits that the FM approach may have over other methods currently in use. First, the FM

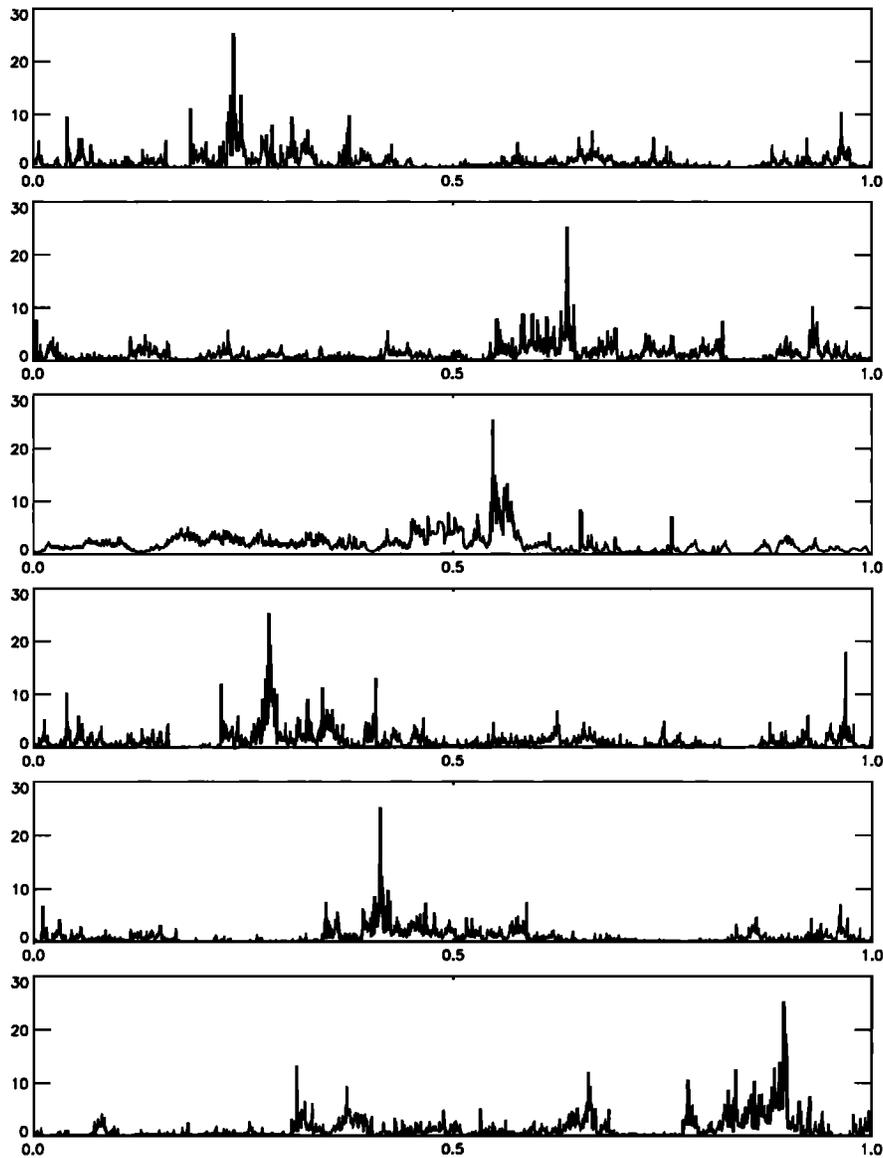


Figure 4. Projections from three-dimensional fractal interpolators and a storm in Boston.

approach is entirely deterministic: Both the parent multifractal measure and the transforming mapping may be uniquely obtained via simple recursive procedures [Puente, 1994]. The data at hand are interpreted as a normalized distribution, a probability measure, which is encoded via a parent multifractal measure and a unique fractal interpolating function. Second, instead of concentrating on the statistics of the actual

realization(s), the FM approach focuses on a wholistic description of geophysical patterns. Rather than concentrating on describing or characterizing the distribution of the data, the FM approach uses derived distributions to describe the data. It appears that the FM procedure, or others based on a similar idea, may provide a very parsimonious representation of natural data sets.

Table 1. Surrogate Parameters for “Storms” in Figure 3

Storm ^a	Localization					Regularity				Intermittency								
	x_0	x_1	x_2	x_3	x_4	y_0	y_1	y_2	y_3	y_4	d_1	d_2	d_3	d_4	p_1	p_2	p_3	p_4
1	-0.055	0.158	0.324	0.753	0.936	0.771	0.077	0.241	-0.337	-1.142	-0.747	-0.082	0.482	0.744	0.331	0.160	0.135	0.374
2	-0.050	0.160	0.320	0.750	0.940	0.770	-0.500	-0.500	0.500	-0.800	0.750	-0.080	0.480	-0.500	0.330	0.160	0.140	0.370
4	-0.047	0.149	0.418	0.750	0.949	-0.853	0.566	0.598	-0.530	0.004	0.745	0.058	-0.172	-0.475	0.308	0.173	0.168	0.351
5	-0.062	0.162	0.312	0.757	0.940	0.780	-0.021	0.129	-0.355	0.883	-0.739	0.073	-0.500	0.598	0.322	0.191	0.142	0.345
6	-0.060	0.161	0.339	0.771	0.903	0.768	0.324	0.416	-0.338	-1.108	-0.759	-0.062	0.500	0.723	0.313	0.149	0.114	0.424

^aNumbers correlate with panels in Figure 3, starting at the top.

Alternative derived distributions may be obtained by considering fractal interpolating functions over higher dimensions [Barnsley, 1988; Barnsley et al., 1989]. Analogous to the two-dimensional case, one may consider $N + 1$ data points, $\{(x_n, y_n, z_n): x_0 < \dots < x_N; n = 0, 1, \dots, N\}$, and a set of N contractile maps of the form

$$w_n \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a_n & 0 & 0 \\ c_n & d_n & h_n \\ k_n & l_n & m_n \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} e_n \\ f_n \\ g_n \end{pmatrix}, \quad (3)$$

such that

$$A_n = \begin{pmatrix} d_n & h_n \\ l_n & m_n \end{pmatrix} = \begin{pmatrix} r_n^{(1)} \cos \theta_n^{(1)} & -r_n^{(2)} \sin \theta_n^{(2)} \\ r_n^{(1)} \sin \theta_n^{(1)} & r_n^{(2)} \cos \theta_n^{(2)} \end{pmatrix} \quad (4)$$

has norm less than 1 (i.e., square root of maximum eigenvalue of $A_n^T A_n < 1$) and

$$w_n \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} = \begin{pmatrix} x_{n-1} \\ y_{n-1} \\ z_{n-1} \end{pmatrix}, \quad w_n \begin{pmatrix} x_N \\ y_N \\ z_N \end{pmatrix} = \begin{pmatrix} x_n \\ y_n \\ z_n \end{pmatrix}, \quad (5)$$

for $n = 1, \dots, N$. Now a three-dimensional “wire” $G = w_1(G) \cup \dots \cup w_N(G)$ appears, the graph of a continuous deterministic function f from $[x_0, x_N]$ to the yz plane. This function perfectly interpolates the $N + 1$ data points, provided that all scaling parameters $r_n^{(i)}$ have magnitudes less than 1. The construction for a three-dimensional fractal interpolating function leaves now four free parameters per map, with $a_n, c_n, e_n, f_n, g_n,$ and k_n all determined in terms of $r_n^{(1)}, r_n^{(2)}, \theta_n^{(1)}, \theta_n^{(2)}$, and the $N + 1$ data points. As in two dimensions, G becomes fractal as the magnitude of $r_n^{(i)}$ increases toward 1 [Barnsley, 1988; Puente and Klebanoff, 1994].

Figure 2 is analogous to Figure 1 for the three-dimensional case. In addition to a parent measure in x , a fractal interpolating function from x to y , and a derived measure in y , now there are also an interpolating function from x to z (a projection of the unique wire in the $x - z$ plane), a derived measure in z , and a joint derived measure in yz . The advantage in going to higher dimensions is that “hidden” relationships between processes in y and in z can now be incorporated. As may be seen, rainfall in time may be modeled by considering either dy or dz . Notice that as in the two-dimensional case, these measures possess intermittency properties typically observed in nature.

In summary, the FM approach relies on the description of intermittent (normalized) data sets as weighted projections of fractal interpolating functions. The relevant parameters that need to be specified for a given data set are (1) the points by which the fractal interpolation function passes (localization parameters); (2) the scalings (and rotations) (regularity param-

Table 3. Relevant Statistics for Real and Fractal-Multifractal Fitted Storms in Boston

Statistic	Order	Real value	Fitted value
$M(y)$	1	0.445	0.439
	2	0.237	0.215
	3	0.265	0.211
	4	2.502	2.412
$M(y)^*$	1	0.547	0.535
	2	0.258	0.235
	3	-0.854	-0.925
	4	2.272	2.273
$M(dy)$	1	0.075	0.077
	2	0.074	0.072
	3	3.643	3.169
	4	28.76	24.65
$\tau(f m)$		23	16
β		2.711	2.686
$\tau(q)$	-1.0	2.034	2.139
	-0.2	1.204	1.202
	0.6	0.396	0.396
	1.4	-0.391	-0.392
	2.2	-1.161	-1.153
	3.0	-1.908	-1.873
	3.8	-2.629	-2.545
	4.6	-3.330	-3.184
	5.4	-4.014	-3.803
	6.2	-4.688	-4.411
$D(1)$		0.982	0.984

eters), and (3) the parent multifractal redistributions (intermittency parameters).

3. Preservation of Overall Features

It is shown in this section that it is possible to find parameter combinations, both for two- and three-dimensional fractal interpolating functions, such that the derived distributions they produce resemble actual rainfall records. In order to illustrate this point, a variety of derived measures are shown in Figures 3 and 4 for two- and three-dimensional fractal interpolating functions, respectively. So that comparisons could readily be made, these measures were generated having the same size of an October 25, 1980, storm gathered in Boston every 15 s (i.e., 1990 data points) and were scaled so that they all have the same intensity range as the actual storm (in $100 \times$ millimeters per 15 s). In both figures, made of six series, the data set depicting the Boston storm is included in the third panel from the top.

As may be seen, all “storms” have features which are similar to those of the Boston data: They have a large peak, few intermediate ones, and low intensities which appear to contain noise. Observe that the deterministic FM representations result in measures which resemble the details present on actual records at a wide range of scales. In fact, it is not easy to discriminate between the real series and those that were generated, as these latter representations do share similar statistical and multifractal characteristics. A complete comparison between the first deterministic storm in Figure 3, which captures well the timing of the major peak, and the actual records from the Boston storm will be given in the next section.

All projections in Figure 3 were obtained from a two-dimensional fractal interpolating function that passed by five data points. Table 1 includes all the FM parameters that were used. The projections in Figure 4, on the other hand, came from three-dimensional fractal interpolating functions which

Table 2. Relevant Surrogate Parameters for “Storms” in Figure 4

Storm ^a	Localization		
	y_2	z_1	z_2
1	0.0	0.90	-0.20
2	0.1	0.40	0.20
4	0.0	1.00	0.00
5	0.1	0.50	0.00
6	0.1	0.20	0.40

^aNumbers correlate with panels in Figure 4, starting at the top.

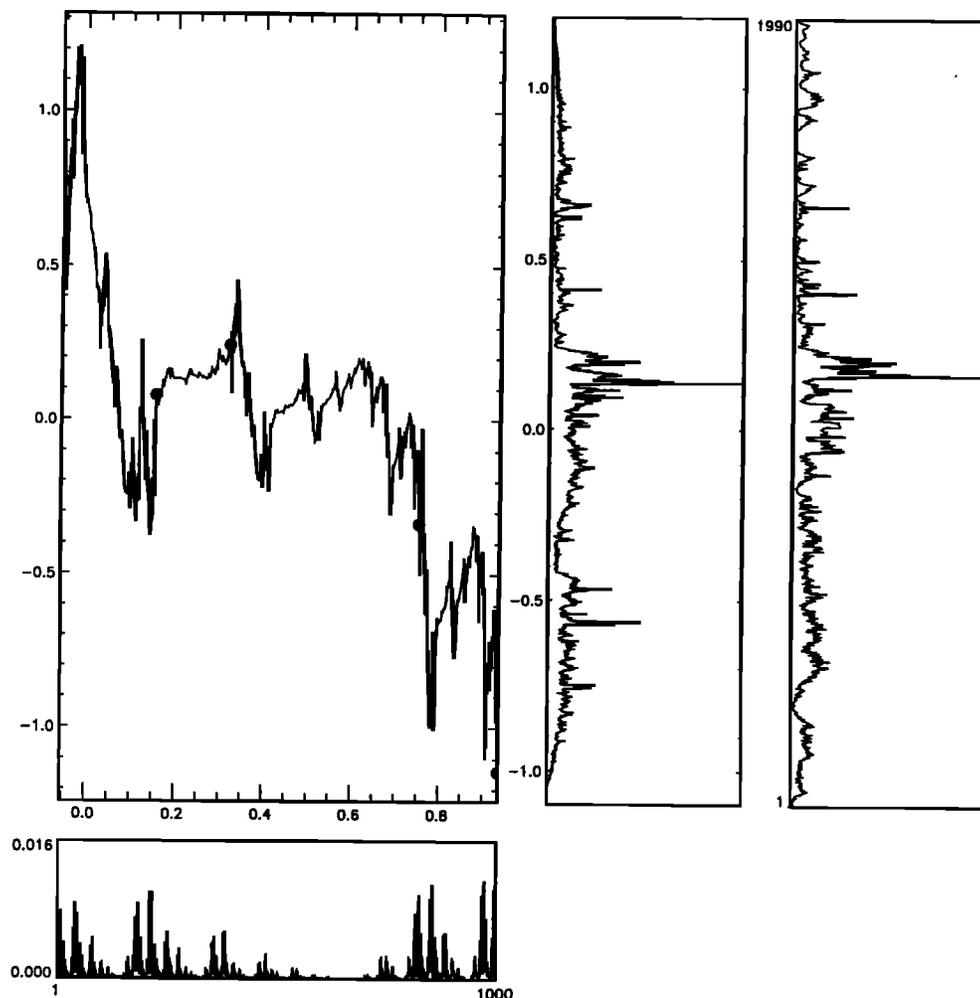


Figure 5. Fitted and real rainfall records for the storm in Boston. The FM parameters are those of the first storm in Table 1.

passed by three data points. Table 2 includes the varying FM parameters for these cases, as they always share $x_0 = 0, x_1 = 0.5, x_2 = 1, y_1 = 1, y_2 = 0, z_0 = 0; r_1^{(1)} = r_1^{(2)} = -0.6, \theta_1^{(1)} = \theta_1^{(2)} = 45, r_2^{(1)} = 0.6, r_2^{(2)} = -0.6, \theta_2^{(1)} = \theta_2^{(2)} = 45;$ and $p_1 = 0.3, p_2 = 0.7$. The angles above are in degrees.

A pictorial representation of the building blocks of the second graph from the top in Figure 4 was already given in Figure 2. In fact, the second storm corresponds to the measure dz in Figure 2 (be aware of proper orientation).

4. A Fitted Storm for the Boston Records

It is easy to verify that the FM approach is continuous with respect to its parameters, i.e., small changes in either dx or f yield a small change in the outcome dy . Unfortunately, there is no simple analytical formula that gives the derived measure dy , or its most common statistics, in terms of these parameters. This implies that the inverse problem of finding the FM parameters for a given data set cannot be obtained analytically but rather requires a numerical solution. At the end, this inverse problem becomes nontrivial due to (1) the large number of combinations of surrogate parameters (even with few interpolating points) and (2) the practically infinite number of derived measures that may be generated, many sharing common statistical and multifractal features. It has been our experience

that a cataloging exercise is a must before attempting any sophisticated search algorithm.

Besides interacting with a catalog, it is clear that a multidimensional optimization procedure and a properly defined objective function are required to fine-tune the FM parameters corresponding to a data set. This section reports the results obtained with a heuristic two-step optimization procedure which was found reliable after extensive experimentation. The method relied on the multidimensional simplex method [Press *et al.*, 1989] to obtain preliminary FM parameters and on simulated annealing [Otten *et al.*, 1989] and sequential quadratic programming [Zhou and Tits, 1993] for fine tuning.

The multidimensional simplex method was used from an initial simplex defined around a set of FM parameters from the catalog, which resulted in "reasonable" visual agreement with the records at hand. In order to specify the procedure, an objective function was defined accounting for weighted sums of squared differences between qualifiers of the real and FM records. These attributes included classical statistical indicators and multifractal characteristics: (1) the first ten moments around the mean in the time axis, (2) the first ten moments around the mode in the time axis, (3) the first ten moments around the mean in the intensity axis, and (4) the mass exponents $\tau(q)$ ($-2 \leq q \leq 7$, at 0.2 increments), which

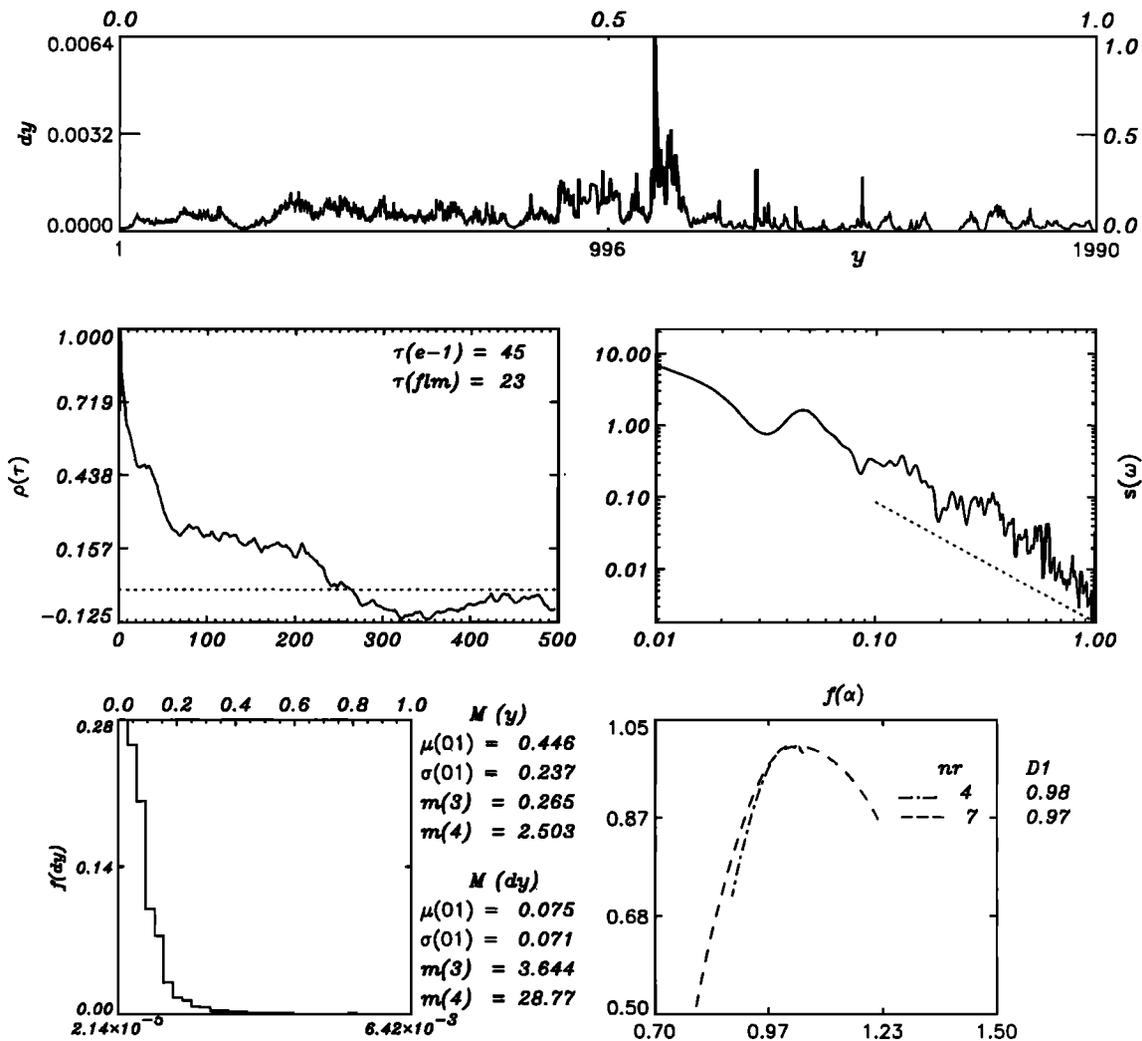


Figure 6. Statistics for the observed storm in Boston: autocorrelation ($\rho(\tau)$), power spectrum ($S(\omega)$), data histogram ($f(dy)$), and multifractal spectrum ($f(\alpha)$).

account for the scaling laws of the q th-order moments of the data sets (μ):

$$\sum_i [\mu_i(\delta)]^q \approx \delta^{-\tau(q)}, \tag{6}$$

where $\mu_i(\delta)$ is the total measure (mass) in the i th piece of resolution δ .

Each one of the entries on the four sets of attributes above had a component in the objective function having as general structure $w_i(1 - q_i/\hat{q}_i)^2$, where w_i is the weight of the i th attribute, with q_i and \hat{q}_i representing the FM and real attributes, respectively. The weights on each of the sets were determined as follows: (1) For the moments they decrease linearly as the order of the moment increases and (2) for the mass exponents they are constant. In addition to the individual weights just explained, a set of four extra weights was defined in order to properly give emphasis to the four classes of attributes considered. These were determined using the initial FM parameters around which the initial simplex was built, in such a way that at that point in space, all four sets of attributes contribute to the objective function equally. In order to have a dynamic procedure that builds on current knowledge, these four weights were modified before starting the second stage in

the optimization procedure. The best parameter values found via the multidimensional simplex method were used to re-define such quantities.

Clearly, the procedure used for assigning weights is by no means unique, as it may be modified in a number of ways. It is worth remarking, however, that the objective function selected did not account explicitly for the autocorrelation or power spectra of the real and FM records. In regards to usage of the optimization algorithms, it should be added that no particular care must be exercised other than demanding that the mathematical structure of the model be preserved, for example, $|d_n| < 1$, $\sum p_n = 1$, etc.

The "optimal" fit obtained for the storm in Boston was already included as the top storm in Figure 3. Figure 5 shows the whole FM construction and the Boston records, for the best fitted parameters as reported in the first row of Table 1. Table 3 provides relevant statistical information for the observed and fitted data sets. These statistics include the following: central moments for the records seen from the time axis, $M(y)$; moments around the mode for records seen from the time axis, $M(y)^*$; central moments computed from the rainfall intensity axis, $M(dy)$; time lag where first local minimum of autocorrelation function happens, $\tau(flm)$; the scaling expo-

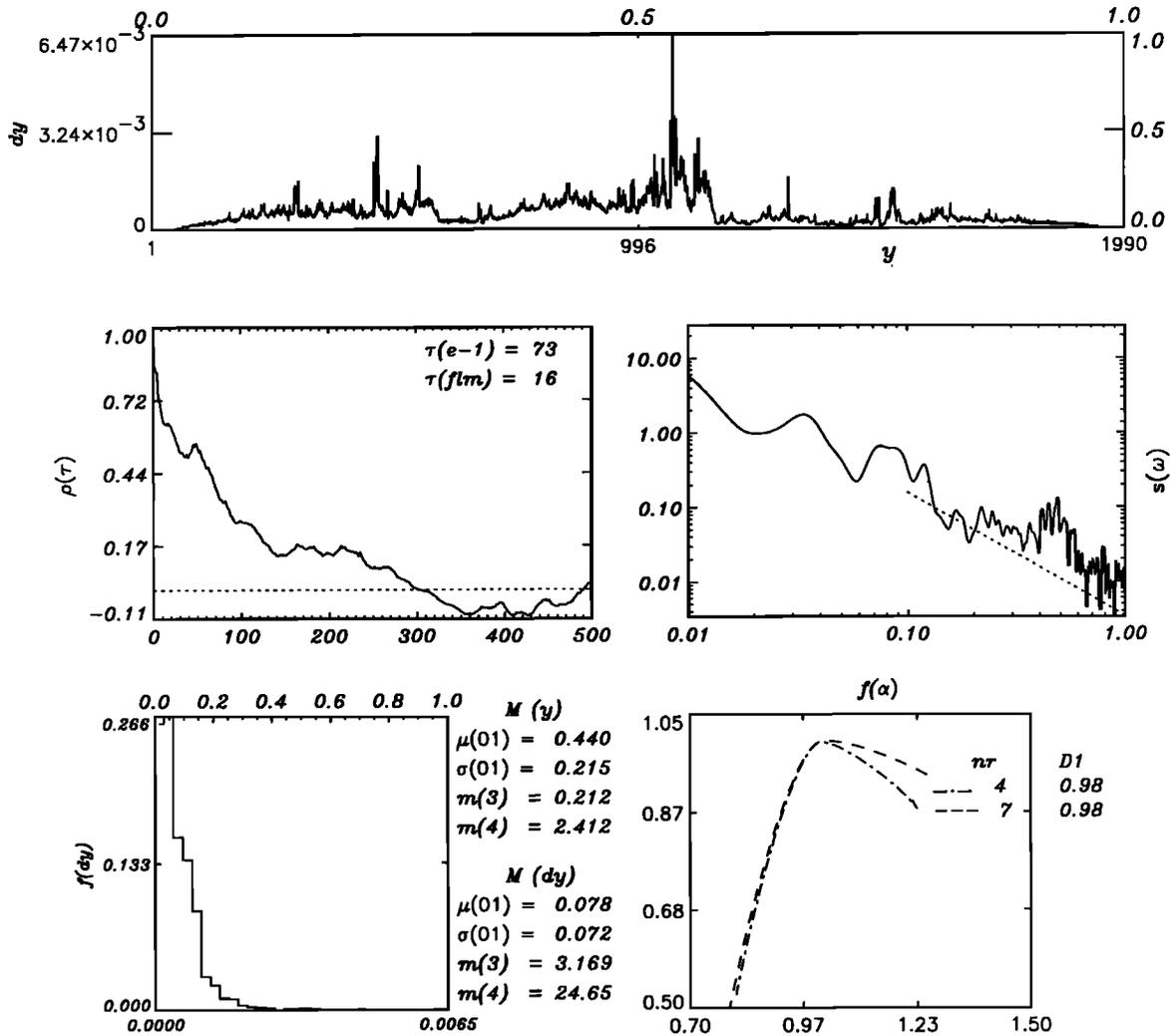


Figure 7. Statistics for the FM fitted storm in Boston: autocorrelation ($\rho(\tau)$), power spectrum ($S(\omega)$), data histogram ($f(dy)$), and multifractal spectrum ($f(\alpha)$).

ment β for the power spectrum at large frequencies, i.e., $S(\omega) \sim \omega^{-\beta}$; mass exponents $\tau(q)$ for the indicated orders q ; and the information dimension for the data, $D(1) = -d\tau(q)/dq$ at $q = 1$, which measures the dimension of a Cantor set where the measure concentrates [Feder, 1988]. The mass exponents above were computed fitting the best regression to equation (6) using four consecutive resolutions δ in powers of 2.

As may be seen, the FM description (having 17 parameters: 10 coordinates, 4 scalings, and 3 independent intermittencies) not only matches well the actual records pictorially, but it also preserves the optimized statistical and multifractal characteristics of the Boston storm. As observed in Table 3, excellent agreement is found for all moments of orders less than 4, with modal moments on the time axis being slightly better than those found around the mean. This fact is a clear indication that the FM representation nicely captures the timing and location of the largest peak observed. Even though no autocorrelation function attributes were accounted for in the optimization exercise, the first local minimum of the autocorrelation function (oftentimes used to define a scale for chaotic analysis; see next section) was properly preserved, i.e., 16 versus 23 lags are very close, considering that the time series studied are made of 1990 values. A close agreement on power

spectrum scaling (also not accounted for during the optimization) was also found, as evidenced by a FM fitted exponent β (computed for frequencies greater than 0.4) which differs by only 1% from the exponent of the real records. As is seen, the mass exponents function $\tau(q)$ is nicely captured, especially for values of the order q between -1 and 3 . The information dimension $D(1)$ is also fitted accurately.

In order to further verify how the FM fit behaves, Figures 6, 7, and 8 compare relevant functions of the observed and FM fitted records from Boston. Observe that despite the duration of 1990 data points for the records in Boston, the FM projection provides reasonably close fittings of both the autocorrelation function and the power spectrum of the actual data. Notice the similarity in the shapes of these functions, with the autocorrelation of the actual data being rougher than the one predicted by the FM approach, which qualitatively maintains the delays where the autocorrelation equals e^{-1} and 0, and with the power spectrum of the predicted records exhibiting a less stable power law behavior than the actual records (in terms of regression fit).

For the data histogram ($f(dy)$), there is indeed a good visual agreement and excellent fit of the moments as previously reported in Table 3. Overall, the moments in both time and

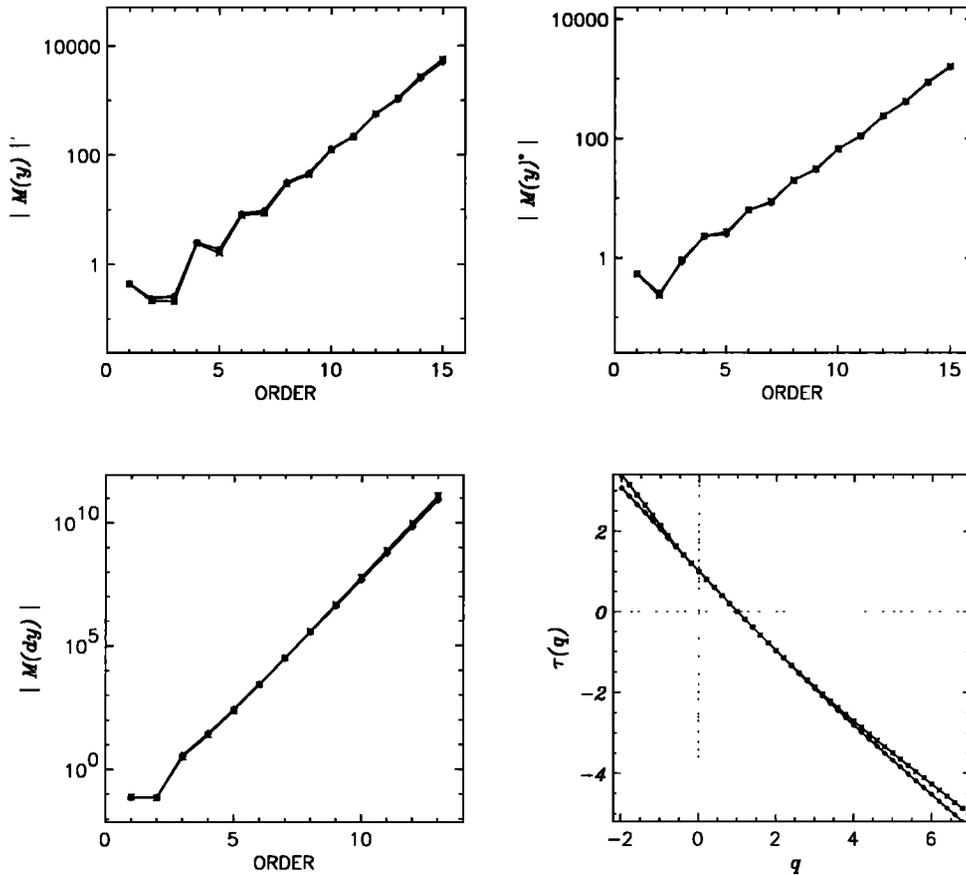


Figure 8. Moments for observed and FM fitted (*-*) storm in Boston: central moments in time ($M(y)$), modal moments in time ($M(y)^*$), central moments in intensity axis ($M(dy)$), and mass exponents ($\tau(q)$).

intensity axes are very well preserved for orders not considered during the optimization exercise. As seen in Figure 8, this happens for time moments of orders up to 15.

As previously included in Table 3 and as depicted in Figure 8, the mass exponents functions of the actual records are nicely preserved by the FM representation. This implies that the multifractal spectra, $f(\alpha)$ versus α , for the observed and fitted storms (found via Legendre transforms: $\alpha = -d\tau/dq$ and $f(\alpha) = \alpha q + \tau(q)$) should be close, especially for values of q between -1 and 3 . Indeed, the bottom right-hand corners of Figures 6 and 7 exhibit good agreement between real and FM fitted multifractal spectra, especially for its left-hand portion. This means that the FM approach preserves the codimension function of the data, i.e., the scaling structure which corresponds to large singularities. As real and fitted mass exponents deviate for large magnitudes of the exponent q , the multifractal spectra of both series differ in both their tails. It is clear that the right-hand portion of the predicted multifractal spectrum is not preserving the scaling on the very small values present in the data, even while changing the number of resolutions considered to define the number of regression points nr used in equation (6). This fact and the observation that many multifractal models exhibit instabilities on the right-hand side branch of the spectrum suggest that negative values of q may have been omitted from the optimization exercise. Overall, the entropy dimension, a stable qualifier of the spectrum, is well preserved by the FM measure. This is particularly true for regressions made with four and seven resolutions.

5. Are Rainfall Time Series Chaotic?

Whether rainfall may be viewed as a deterministic chaotic process has attracted attention in the literature [e.g., *Rodriguez-Iturbe et al.*, 1989]. Given that it is possible to model complex data sets by means of the FM approach, it is relevant to ask if the outcomes of the deterministic procedure lead to deterministic chaos. This section studies this question by analyzing the actual records and the “best” representation for the Boston storm.

The specific chaotic analysis carried for both series relies on standard tests as defined in the literature. The actual steps undertaken are as follows: (1) phase-space reconstruction of the time series using a time delay [e.g., *Packard et al.*, 1980]; (2) determination of minimal embedding dimension via correlation dimension stabilization [e.g., *Grassberger and Procaccia*, 1983; *Grassberger*, 1990]; (3) verification of chaotic behavior via K_2 entropy [e.g., *Provenzale et al.*, 1991]; (4) verification of minimal embedding dimension via the false neighbors algorithm [e.g., *Kennel et al.*, 1992]; and (5) verification of sensitivity to initial conditions via a positive Lyapunov exponent [e.g., *Wolf et al.*, 1985].

The results of the analysis for both of the Boston “records” are included in Figures 9–12. Figures 9 and 11 include the correlation dimension analysis and K_2 entropy calculations using a delay τ that corresponds to the first local minimum of the autocorrelation function of the real records (i.e., 23). These figures include the following information.

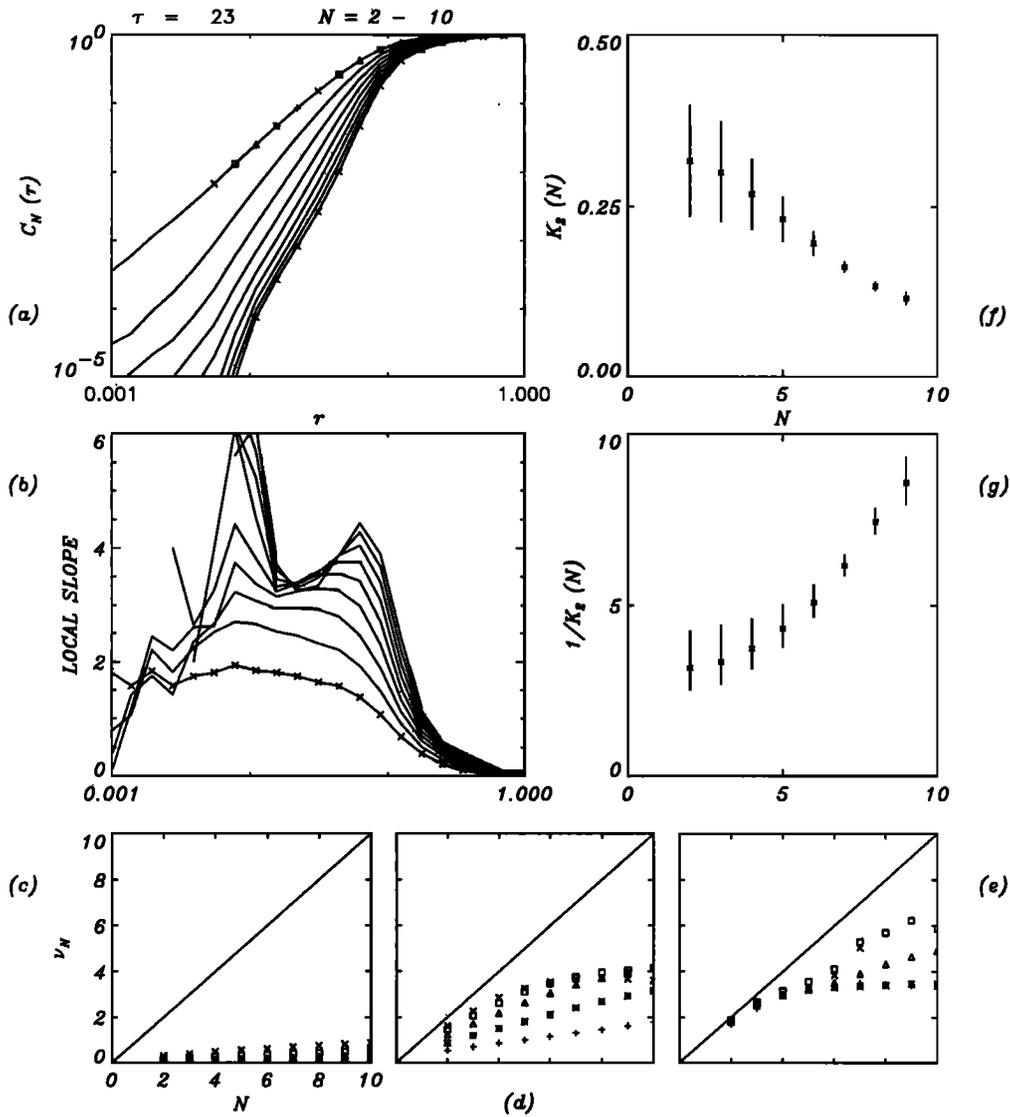


Figure 9. Observed storm in Boston correlation dimension and $K_2(N)$ entropy. Figure 9a shows phase-space correlation functions; Figure 9b shows local slopes of correlation functions; Figures 9c, 9d, and 9e show correlation dimensions for five large, five intermediate, and five small distances r , respectively, as indicated from right to left on the uppermost correlation function in Figure 9a by the plus signs, asterisks, open triangles, open squares and crosses; Figure 9f shows K_2 mean entropy and one standard deviation band; and Figure 9g shows $1/K_2$ and one standard deviation band.

Figures 9a and 11a show the phase-space correlation functions $C_N(r)$, which “count” the number of points, in a phase space with N coordinates, whose distances are less than r . Figures 9b and 11b show the local slope of $C_N(r)$ (computed via lag-one differences). Figures 9c and 11c, 9d and 11d, and 9e and 11e show the correlation dimensions ν_N , defined from the scaling relation $C_N(r) \sim r^{\nu_N}$, computed for five large, five intermediate, and five small values of the distance r , respectively. Such values are indicated on the upper correlation function curve (i.e., for two coordinates) and are found via a regression of three successive points ending with the symbol in question; for example, the values reported for ν_N and the plus sign are obtained employing $C_N(r)$ for distances indicated by the open triangle, the asterisk, and the plus sign. Figures 9f and 11f show the average $K_2(N)$ entropy, defined as the logarithm of the ratio of two consecutive phase-space correlation functions, found for the “stable” region in the correlation function

slope and the band of plus and minus one standard deviation, and Figures 9g and 11g show the average value of $1/K_2(N)$ and its corresponding band of one standard deviation.

As may be seen in Figures 9a and 11a, the computed correlation functions for both observed and fitted records, considering up to ten coordinates, are not perfect straight lines in log-log scale. This leads to stable and unstable local slopes (found from two successive points), as indicated in Figures 9b and 11b. While the actual data leads to a slope valley between two mounds, the local slopes on the FM fitted records show only one mound at small distances. Notice from Figures 9c, 9d, and 9e that the real data exhibit stable correlation dimension estimates (found via a regression of three consecutive points) in the intermediate zone, yielding a correlation dimension of 3.68; see the graph with the crosses in Figure 9d. The FM records, on the other hand, also lead to stable behavior and a close correlation dimension of 3.44, but now for the asterisk

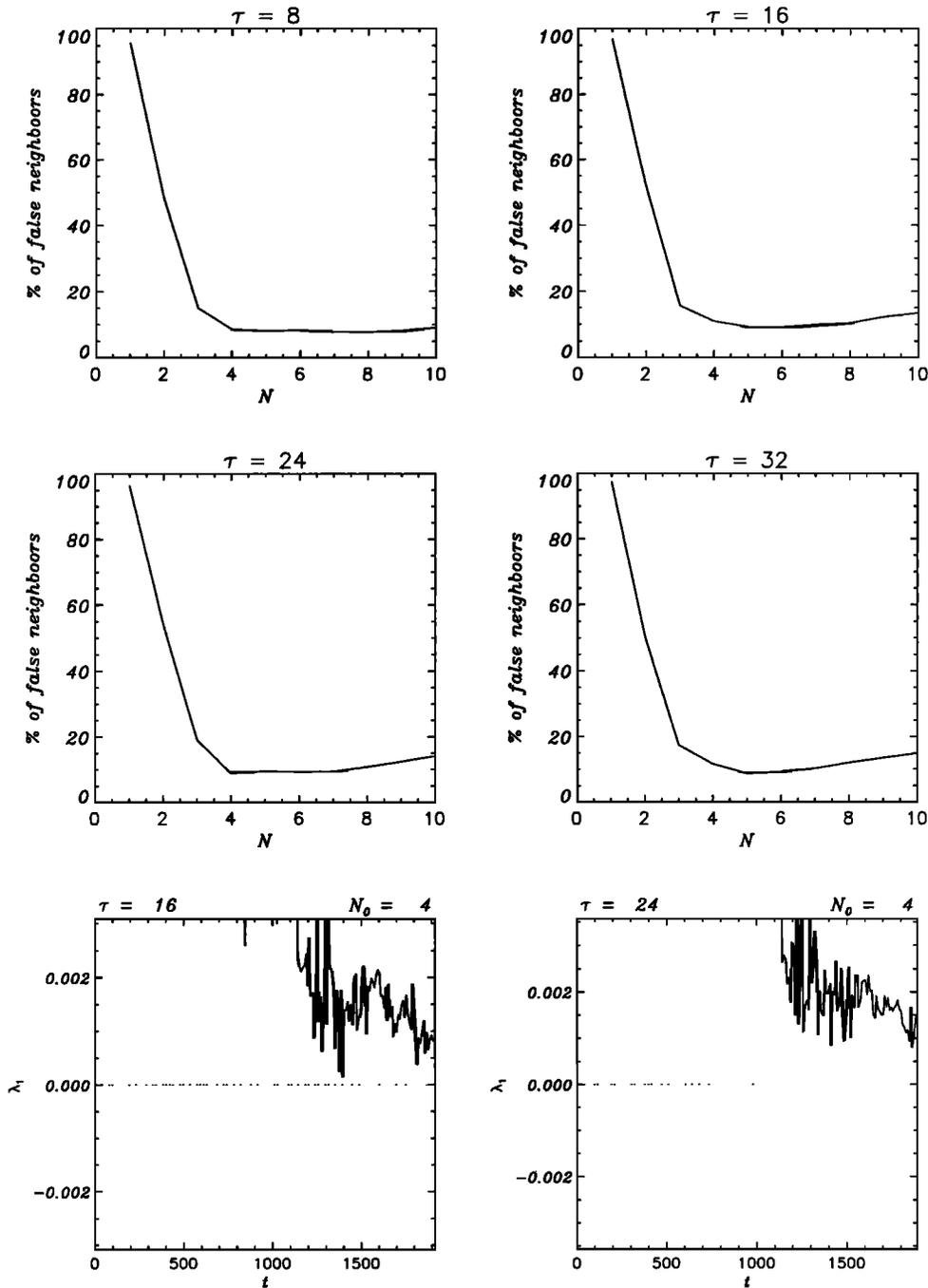


Figure 10. Observed storm in Boston: false neighbors for alternative delays (τ) and largest Lyapunov exponent for alternative dimensions (N_0).

corresponding to a small distance, i.e., Figure 11c. These correlation dimensions are quite close to the value of 3.78 reported for this storm by Rodriguez-Iturbe *et al.* [1989].

The $K_2(N)$ entropy analysis yields inconclusive results for both the real and FM fitted storms. As may be seen in Figures 9f and 9g and 11f and 11g, K_2 remains positive for all embedding dimensions N , but no apparent stabilization of either $K_2(N)$ or its inverse is attained. The short length of the time series precludes a more complete analysis regarding this attribute. Notice, however, that the behaviors of real and FM fitted $K_2(N)$ entropies are quite similar.

Figures 10 and 12 verify the correlation dimension results by

bounding the optimal embedding dimension using the method of false neighbors [Kennel *et al.*, 1992]. By following the evolution of points in phase-space and keeping track of points that remain close dynamically, a verification of a true attractor is made. These figures also include the largest Lyapunov exponents under alternative conditions, i.e., the growth in phase space of the largest principal axis in which an initial sphere progressively deforms [Wolf *et al.*, 1985]. Notice that the percentage of false neighbors behaves very similarly for both real and FM fitted records. As is seen, there is a marked decrease in false neighbors as the embedding dimension is increased from 2 to 4 for delays τ that range from 8 to 32. At this stage,

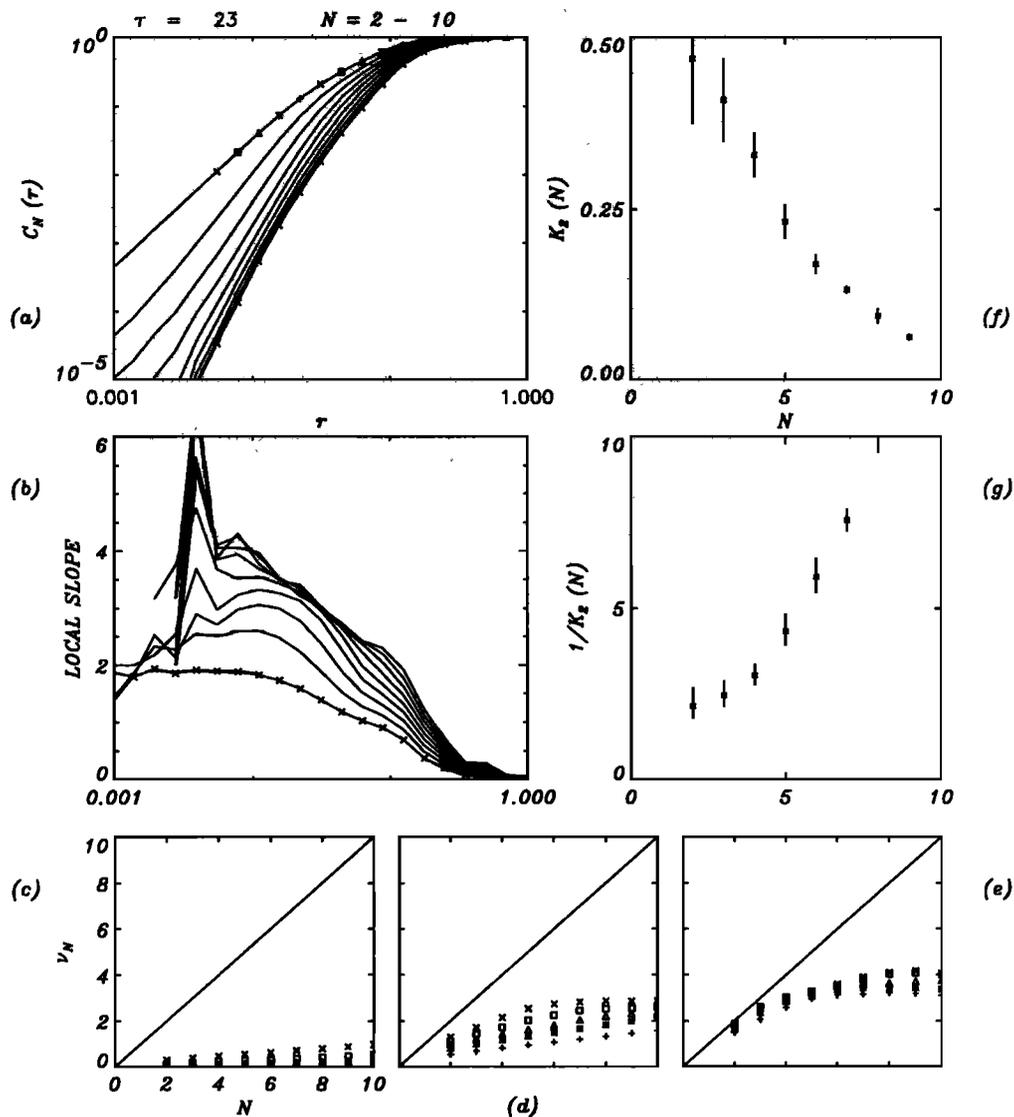


Figure 11. FM fitted storm in Boston correlation dimension and $K_2(N)$ entropy. Figure 11a shows phase-space correlation functions; Figure 11b shows local slopes of correlation functions; Figures 11c, 11d, and 11e show correlation dimensions for five large, five intermediate, and five small distances r , respectively, as indicated from right to left on the uppermost correlation function in Figure 11a by the plus signs, asterisks, open triangles, open squares, and crosses; Figure 11f shows K_2 mean entropy and one standard deviation band; and Figure 11g shows $1/K_2$ and one standard deviation band.

the percentage of false neighbors falls in all cases below 10%, and only a mild rise is observed thereafter. These results clearly support the validity of the correlation dimension values given before.

The largest Lyapunov exponent (λ_1) for both real and FM fitted series is given in Figures 10 and 12 for an optimal embedding of $N_0 = 4$ and for a couple of plausible delays τ . The results for the real data are clear. A positive Lyapunov exponent is found for both delays considered. For the predicted records the Lyapunov exponents vary in character. For most of the time, λ_1 gives positive values which signify chaotic behavior, but there are series durations for which negative values are found. Overall, the behavior of both real and predicted records is similar in relation to the largest Lyapunov exponent: Notice that the scales on both cases are quite close and that, indeed, positive values are found. Compare in particular the cases when $\tau = 16$.

In order to avoid issues related to the small length of the time series, the FM procedure was used to obtain a refinement of the Boston records over 2^{14} points. A complete analysis of these "records" confirmed the chaotic nature of the FM projection which nicely approximates the Boston data at the lower resolution. This time, a correlation dimension close to 3.2 was found, and the percentage of false neighbors also came below 10% for four embedding dimensions. This time a positive largest Lyapunov exponent was found, even for the duration of the series.

6. Summary and Conclusions

A deterministic geometric framework for the description of complex hydrologic (geophysical) time series has been reviewed. The approach relies on the use of fractal-geometric objects and in particular on the combination of fractal inter-

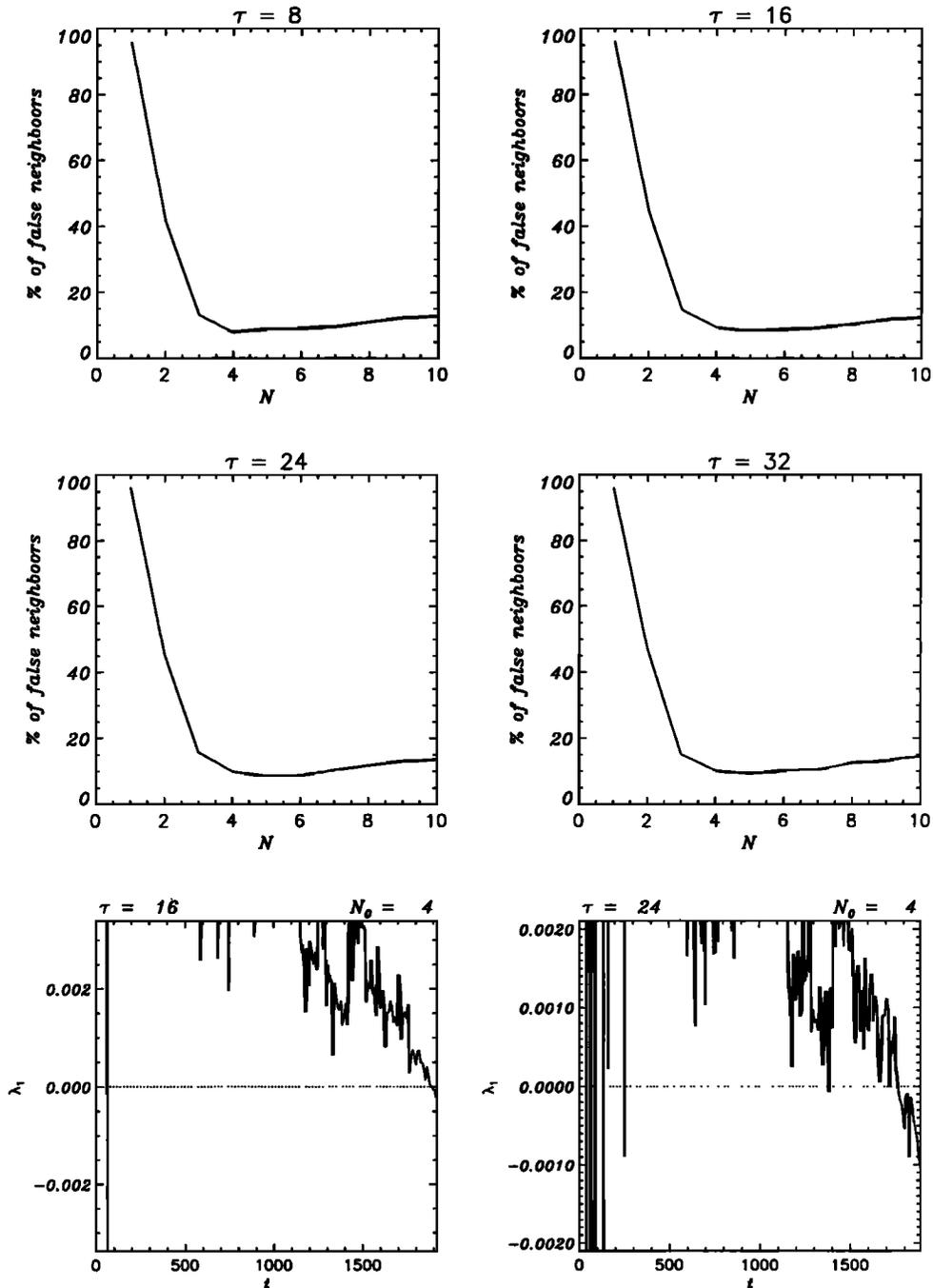


Figure 12. FM fitted storm in Boston: false neighbors for alternative delays (τ) and largest Lyapunov exponent for alternative dimensions (N_0).

polution functions and multinomial multifractal measures. It has been illustrated, by means of examples, that some derived measures produced by the fractal-multifractal (FM) procedure resemble high-resolution rainfall records, both for two- and three-dimensional fractal interpolating functions.

A faithful FM representation of a high-resolution storm in Boston has been presented in this work. A detailed comparison of the real and FM fitted time series reveals that the geometric procedure not only captures the timing and size of the largest peak but also preserves the overall appearance of the actual records, including secondary peaks and small noisy fluctuations. The FM representation was reached minimizing a

weighted sum of squared differences between attributes of the real and FM outcomes of equal length. The obtained “optimal” solution indeed preserves the following attributes, explicitly accounted for in the objective function: (1) the first ten central and modal moments of the records when seen from the time axis; (2) the first ten central moments of the records as seen from the intensity axis; and (3) the mass exponents function of the data. The FM representation also captured (1) higher-order moments along both the time and intensity axes; (2) the overall shape of the autocorrelation and histogram functions for the records; (3) the scaling properties present in the power spectrum of the data; and (4) the chaotic nature of the set of observations.

Since similarly faithful FM representations (to be reported elsewhere) have also been found for a couple of the high-resolution storms (one having two peaks) gathered in Iowa City every 5 s by *Georgakakos et al.* [1994], these results imply that there may be no need in separating trends and (arguably unimportant) small fluctuations when dealing with rainfall records. In fact, this work suggests a new global perspective for understanding rainfall, a perspective in which major features and noisy details are captured jointly. Clearly, the presented analysis hints that a stochastic framework for rainfall modeling may not be necessary and reveals that the notion of projections may provide the proper alternative. The FM procedure is indeed quite general, as it may generate patterns of arbitrary lengths with a multitude of peaks and records which contain periods of no rain. This last point is easily made using a Cantorian measure to compute weighted projections.

There are of course unanswered questions regarding the FM methodology, which need to be studied in the future. These include (1) the complete understanding of the kinds of measures that may be generated by the FM procedure; (2) the search for efficient algorithms (including the proper objective function) to properly describe real hydrologic (geophysical) data sets like the one analyzed here; (3) the identification of the most parsimonious representation (with the least possible number of interpolating points) of a given data set; and (4) the determination of the physical meaning that the FM parameters may possess.

Clearly, the merit of the FM methodology to represent data sets will rest on our ability to solve the proper inverse problem in the least possible amount of time. As argued by *Puente* [1996], rainfall (geophysical) data sets are unique "signatures" of the physical, chemical, and biological processes taking place within the atmosphere, and understanding the geometry of these records may be very valuable for rainfall modeling and prediction. It is envisioned that once FM parameters are available for subsequent data sets, such surrogate geometric information may be useful to study the dynamics of rainfall. If trends in FM parameter space may be elucidated, this may lead to predictions of rainfall by chunks. Instead of allowing predictions few time steps ahead at a time (e.g., minutes, hours, or days ahead), the FM approach may result in wholistic representations of rainfall records at the same timescale as the input data used and with a size equal to the sizes of the records used to obtain the relevant trends on FM parameters. A plausible interpretation of the FM parameters should be elucidated when the ideas are tested under alternative climatic conditions.

It is important to emphasize that the present work, although dealing with the concept of multifractality in rainfall, is not fully comparable to the work on rainfall that relies on parameterizations of the record's multifractal spectrum in terms of stochastic cascades, for example, *Lovejoy and Schertzer* [1990], *Rajagopalan and Tarboton* [1993], and *Over and Gupta* [1994]. Although such models are more parsimonious than the FM procedure, they treat the data as a realization of a stochastic process whose prevalent characteristic is the multifractal spectrum and therefore cannot account for the uniqueness of the data at hand. The deterministic FM approach, on the other hand, seeks understanding of the whole and unique data set, leading to an approach which preserves the multifractal spectrum, other important qualifiers of the records, and the overall geometric appearance of the data.

Acknowledgments. The research leading to this article was supported by the University of California Water Resources Center, as part of Water Resources Center Project UCAL-WRC-W-804. Valuable comments by anonymous reviewers are acknowledged.

References

- Barnsley, M. F., Fractal functions and interpolation, *Constr. Approx.*, 2, 303–329, 1986.
- Barnsley, M. F., *Fractals Everywhere*, Academic, San Diego, Calif., 1988.
- Barnsley, M. F., J. Elton, D. Hardin, and P. Massopust, Hidden variable fractal interpolation functions, *SIAM J. Math. Anal.*, 20(5), 1218–1242, 1989.
- Casdagli, M., Nonlinear prediction of chaotic time series, *Physica D*, 35, 335–356, 1989.
- Crutchfield, J. P., and B. S. McNamara, Equations of motion from a data series, *Complex Syst.*, 1, 417–452, 1987.
- Eagleson, P. S., Climate, soil, and vegetation, 2, The distribution of annual precipitation derived from observed storm sequences, *Water Resour. Res.*, 14(5), 713–721, 1978.
- Feder, J., *Fractals*, Plenum, New York, 1988.
- Foufoula-Georgiou, E., and P. Guttorp, Compatibility of continuous rainfall occurrence models with discrete rainfall observations, *Water Resour. Res.*, 22(8), 1316–1322, 1986.
- Georgakakos, K. P., and R. L. Bras, A hydrologically useful station precipitation model, 1, Formulation, *Water Resour. Res.*, 20(11), 1585–1596, 1984.
- Georgakakos, K. P., A. A. Carsteanu, P. L. Sturdevant, and J. A. Cramer, Observation and analysis of midwestern rain rates, *J. Appl. Meteorol.*, 33(12), 1433–1444, 1994.
- Grassberger, P., An optimized box-assisted algorithm for fractal dimensions, *Phys. Lett. A*, 148, 63–68, 1990.
- Grassberger, P., and I. Procaccia, Measuring the strangeness of strange attractors, *Physica D*, 9, 189–208, 1983.
- Kavvas, M. L., and J. W. Delleur, A stochastic cluster model of daily rainfall sequences, *Water Resour. Res.*, 17(4), 1151–1160, 1981.
- Kennel, M. B., R. Brown, and H. D. I. Abarbanel, Determining embedding dimension for phase-space reconstruction using a geometrical construction, *Phys. Rev. A*, 45(6), 3403–3411, 1992.
- Lovejoy, S., and D. Schertzer, Generalized scale invariance in the atmosphere and fractal models of rain, *Water Resour. Res.*, 21(8), 1233–1250, 1985.
- Lovejoy, S., and D. Schertzer, Multifractals, universality classes, and satellite and radar measurements of cloud and rain fields, *J. Geophys. Res.*, 95(D3), 2021–2034, 1990.
- Mandelbrot, B. B., Multifractal measures especially for the geophysicist, in *Fractals in Geophysics*, edited by C. H. Scholz and B. B. Mandelbrot, pp. 1–42, Birkhauser Verlag, Basel, Switzerland, 1989.
- Meneveau, C., and K. R. Sreenivasan, Simple multifractal cascade model for fully developed turbulence, *Phys. Rev. Lett.*, 59, 1424–1427, 1987.
- Moon, F. C., *Chaotic Vibrations*, John Wiley, New York, 1987.
- Otten, R. H. J. M., and L. P. P. van Ginneken, *The Annealing Algorithm*, Kluwer Acad., Norwell, Mass., 1989.
- Over, T. M., and V. K. Gupta, Statistical analysis of mesoscale rainfall: Dependence of random cascade generator on large scale forcing, *J. Appl. Meteorol.*, 33, 1526–1542, 1994.
- Packard, N. H., J. P. Crutchfield, J. D. Farmer, and R. S. Shaw, Geometry from a time series, *Phys. Rev. Lett.*, 45, 712–716, 1980.
- Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes*, Cambridge Univ. Press, New York, 1989.
- Provenzale, A., A. R. Osborne, and R. Soj, Convergence of the K_2 entropy for random noises with power law spectra, *Physica D*, 47, 361–372, 1991.
- Puente, C. E., Multinomial multifractals, fractal interpolators, and the Gaussian distribution, *Phys. Lett. A*, 161, 441–447, 1992.
- Puente, C. E., Deterministic fractal geometry and probability, *Int. J. Bifurcations Chaos*, 4(6), 1613–1629, 1994.
- Puente, C. E., A new approach to hydrologic modeling: Derived distributions revisited, *J. Hydrol.*, in press, 1996.
- Puente, C. E., and A. Klebanoff, Gaussians everywhere, *Fractals*, 2(1), 65–79, 1994.
- Rajagopalan, B., and D. G. Tarboton, Understanding complexity in the structure of rainfall, *Fractals*, 1(3), 312–322, 1993.
- Rodriguez-Iturbe, I., Scale of fluctuation of rainfall models, *Water Resour. Res.*, 22(9), 15S–37S, 1986.

- Rodriguez-Iturbe, I., D. R. Cox, and V. Isham, Some models for rainfall based on stochastic point processes, *Proc. R. Soc. London A*, 410, 269–288, 1987.
- Rodriguez-Iturbe, I., B. Febres de Power, M. B. Sharifi, and K. P. Georgakakos, Chaos in rainfall, *Water Resour. Res.*, 25(7), 1667–1675, 1989.
- Schertzer, D., and S. Lovejoy, Physical modeling and analysis of rain and clouds by anisotropic scaling multiplicative processes, *J. Geophys. Res.*, 92(D8), 9693–9714, 1987.
- Smith, J. A., and A. F. Karr, A point process model of summer season rainfall occurrences, *Water Resour. Res.*, 19(1), 95–103, 1983.
- Sreenivasan, K. R., Fractals and multifractals in fluid turbulence, *Annu. Rev. Fluid Mech.*, 23, 539–600, 1991.
- Tsonis, A. A., and J. B. Elsner, Chaos, strange attractors, and weather, *Bull. Am. Meteorol. Soc.*, 70, 14–23, 1989.
- Valdes, J. B., I. Rodriguez-Iturbe, and V. K. Gupta, Approximations of temporal rainfall from a multidimensional model, *Water Resour. Res.*, 21(8), 1259–1270, 1985.
- Wolf, A., J. B. Swift, H. Swinney, and J. A. Vastano, Determining Lyapunov exponents from a time series, *Physica D*, 16, 286–317, 1985.
- Zhou, J. L., and A. L. Tits, Nonmonotone line search for minimax problems, *J. Optimization Theory Appl.*, 76, 455–476, 1993.
-
- N. Obregón and C. E. Puente, Hydrologic Sciences, Department of Land, Air and Water Resources, University of California, Davis, Davis, CA 95616.

(Received October 13, 1995; revised May 3, 1996; accepted May 9, 1996.)