Stoch Environ Res Risk Assess
DOI 10.1007/s00477-012-0617-6

ORIGINAL PAPER

Geometric harnessing of precipitation records: reexamining four storms from Iowa City

Huai-Hsien Huang · Carlos E. Puente · Andrea Cortis

Abstract Complex geometries often present in hydrologic data sets such as precipitation records have been difficult to model in their totality using classical stochastic methods. In recent years, we have developed extensions of a deterministic procedure, the fractal-multifractal (FM) method, whose patterns share fine details and textures of individual data sets in addition to the usual key statistical properties. This work discusses our latest efforts at encoding four geometrically distinct storms gathered in Iowa City with parameters found running a modified particle swarm optimization procedure. The results reaffirm the capabilities of the FM method as all storms are closely fitted within measurement errors. All sets may be encoded with a compression ratio exceeding 350:1, have a maximum error in cumulative distribution less than 2.5 %, and closely preserve the autocorrelation, power spectrum, and multifractal spectrum of the records.

Keywords Rainfall in time · Fractals · Multifractals · Inverse problem · Particle swarm optimization · Fractal-multifractal approach

1 Introduction

The development of sophisticated mathematical techniques that has been witnessed in the past few decades, such as models based on stochastic theories and fractal geometry, has allowed substantial progress of rainfall complexity representation (e.g., Schertzer and Lovejoy 1987; Tessier et al. 1996; Marsan et al. 1996; Menabde and Sivapalan 2000; Deidda et al. 1999, 2004, 2006; Deidda 2000; Veneziano et al. 2006, 2007, 2009; Veneziano and Langousis 2010). Although these ideas provide a suitable language for describing and simulating these intricate data sets, oftentimes they are still inadequate to study, on a storm by storm basis, the immense variety of rainfall time series produced by nature.

As rainfall sets are generally erratic, noisy, intermittent, complex, or in short, "seemingly random," it has become quite natural to represent them using stochastic (fractal) theories (e.g., Rodríguez-Iturbe 1986; Rodríguez-Iturbe et al. 1989; Lovejoy and Schertzer 1990; Menabde et al. 1997; Langousis and Veneziano 2007; Veneziano and Langousis 2005a, b; Langousis et al. 2009). This has inspired a wide array of approaches that, while retaining relevant statistical (physical) attributes (e.g., moments, autocorrelation, power spectrum, distribution of extremes, etc.), yield realizations that often fail to replicate specific details (e.g., positions of major peaks) and relevant textures (e.g., periods of no activity) of distinct storms.

Intrinsic limitations of stochastic approaches have led us to propose a fractal geometric methodology (e.g., Puente 1996) using fractal interpolating functions (e.g., Barnsley 1988) to transform multifractal measures in the hopes of capturing the overall complexity of rainfall patterns beyond key statistical features. Our "fractal-multifractal" (FM) approach can indeed generate a vast class of relevant patterns, over one or more dimensions, that, as the stochastic methods, preserve distinct characteristics of rainfall sets, but in addition, allow describing entire sets deterministically with a small number of parameters (e.g., Puente 1996, 2004; Puente and Obregón 1996; Obregón et al. 2002a, b).

In this article, we use a modified particle swarm optimization method (Fernández Martínez et al. 2010) and

H.-H. Huang · C. E. Puente (✉)
Land, Air & Water Resources, University of California, Davis, 127 Veihmeyer Hall, One Shields Ave., Davis, CA 95616, USA
e-mail: cepuente@ucdavis.edu

A. Cortis
Lawrence Berkeley National Laboratory, One Cyclotron Rd., Berkeley, CA 94710, USA

Published online: 30 August 2012

extensions of the FM approach to reexamine the encodings of four high-resolution rainfall data sets with distinct geometries as previously gathered by Georgakakos et al. (1994) and reported in Obregón et al. (2002a). As shall be seen, the search procedure results in suitable solutions of the intrinsic inverse problem, leading to FM representations that, within measurement errors, bear impressive resemblance to the original.

2 Materials and methods

In this section, we summarize the constructions of four FM alternative approaches and outline the strategy for the solution of the key inverse problem.

2.1 The original FM approach

In its original and simplest form, a FM pattern is obtained in a Platonic way, as it were, as the projection of the graph of a fractal interpolating function "illuminated" by a multifractal measure. The graph $G = \{(x,f(x))|x \in [0,1]\}$ of a fractal interpolating function f: $x \to y$ passing by $N+1$ ordered points along x, $\{(x_n, y_n)|x_0 < \cdots < x_N, \quad n = 0,1,\ldots,N\}$ is defined as the unique attractor of N affine maps: (Barnsley 1988)

$$w_n\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a_n & 0 \\ c_n & d_n \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e_n \\ f_n \end{pmatrix}, \quad n = 1,\ldots,N, \tag{1}$$

where the vertical scaling parameters d_n satisfy $|d_n| < 1$ and the other parameters a_n, c_n, e_n, and f_n are based on the contracting initial conditions, once again along x.

$$w_n\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} x_{n-1} \\ y_{n-1} \end{pmatrix} \tag{2}$$

and

$$w_n\begin{pmatrix} x_N \\ y_N \end{pmatrix} = \begin{pmatrix} x_n \\ y_n \end{pmatrix} \tag{3}$$

As may be easily verified, Eqs. (2) and (3) map the first and last interpolating points onto values defined in interior subintervals. In virtue of a fixed-point theorem, the attractor G exists and contains the $N+1$ interpolating points (Barnsley 1988). Substitution of Eqs. (1) into (2) and (3) uniquely determines the values of a_n, c_n, e_n, and f_n in terms of d_n, the vertical scalings, and (x_n, y_n), the coordinates of the interpolating points. Ultimately, the iteration of the affine mappings in Eq. (1) yields a unique (and hence deterministic) set G having a fractal dimension $1 \le D < 2$.

In a practical setting, the graph of a fractal function, typically shaped as a convoluted wire, is obtained by a pointwise sampling of the attractor through a procedure also known as the chaos game (Barnsley 1988). The idea is to start the process at a given (x_n, y_n) already in G and progressively iterate the N maps w_n according to, for example, the outcomes of independent coin tosses. As this process is performed, a unique invariant measure is also induced over G that reflects how the attractor is filled up. The existence of such a measure (akin to a histogram) allows computing unique—and again, fully deterministic—projections over the coordinates x and y (denoted by dx and dy)—and at any resolution—that turn out to display irregular shapes as found in a variety of geophysical applications and beyond (see, e.g., Puente 2004).

As will be fully appreciated later on, Fig. 1a in Appendix 2 shows an example of a fractal wire passing through the four points $\{(0,0),(0.36,-0.44),(0.81,-1.33),(1,1)\}$ as generated by 10^6 iterations of the three maps

$$w_1\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0.36 & 0 \\ -0.28 & -0.16 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{4}$$

$$w_2\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0.46 & 0 \\ -1.28 & 0.39 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0.36 \\ -0.44 \end{pmatrix}, \tag{5}$$

and

$$w_3\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0.19 & 0 \\ 1.85 & 0.48 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0.81 \\ -1.33 \end{pmatrix} \tag{6}$$

In other words, the three maps w_1, w_2, and w_3 have vertical scaling parameters $d_1 = -0.16$, $d_2 = 0.39$, and $d_3 = 0.48$ and operate into the intervals [0,0.36], [0.36,0.81], and [0.81,1] respectively. In addition to the graph G, the figure also displays the projections (histograms) dx and dy induced while carrying the previously mentioned chaos game according to a biased 26–31–43 % proportion on w_1, w_2, and w_3.

As the x-coordinate is independent of y (as implied by the zero value in Eq. (1)), dx ends up as a recursively generated deterministic (multinomial) multifractal measure. In turn, dy is the derived measure of dx via f, and such a histogram is computed, for any given value of y, adding the corresponding "events" dx that satisfy $f(x) = y$ As can be seen, the geometric construction generate "random-looking" sets dy that resemble a rainfall time series (e.g., Puente 2004; Obregón et al. 2002b). Given that multifractals can be used to model energy distributions in turbulent atmospheric flows, we can physically interpret our outputs dy as "reflections" of turbulence or non-trivial (fractal) integrations of spiky multifractals. As these derived measures, for suitable sets of parameters, share the spectrum of singularities of general random cascades (Puente and Obregón 1999) also modeled via "universal multifractals" (Tessier et al. 1993), they may also be thought of as specific realizations of random cascades, which have the advantage of being fully

characterized, in their entirety, by a small set of parameters (Cortis et al. 2012).

2.2 Extension 1: introducing overlaps and holes

The procedure illustrated in Sect. 2.1 may be generalized so that the attractor G is no longer a function from x to y, but a "cloud" of points. This is accomplished by iterating N affine maps, as in Eq. (1), but replacing the contractive initial conditions by

$$w_n\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} x_{2n} \\ y_{2n} \end{pmatrix} \tag{7}$$

and

$$w_n\begin{pmatrix} x_{2N-1} \\ y_{2N-1} \end{pmatrix} = \begin{pmatrix} x_{2n+1} \\ y_{2n+1} \end{pmatrix} \tag{8}$$

$n = 0, 1, \ldots, N-1$ while satisfying $x_0 \leq x_{2n} < x_{2n+1} \leq x_{2N-1}$. Here, each one of the maps is associated with an arbitrary set of endpoints that mark some new subintervals in x, $[x_{2n}, x_{2n+1}]$ and its associated values in y, that is, $[y_{2n}, y_{2n+1}]$. When these endpoints, from mapping to mapping, match both in x and y, this setting gives the previously defined fractal interpolating function. However, a more sophisticated attractor not shaped as a single-valued function emerges when endpoints in x or y do not match (in other words, when $x_{2n-1} \neq x_{2n}$ or $y_{2n-1} \neq y_{2n}$).

Figure 1b in Appendix 2, to be fully understood later on, uses the three maps

$$w_1\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 2.11 & 0.36 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{9}$$

$$w_1\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0.16 & 0 \\ 6.89 & 0.47 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0.59 \\ -4.07 \end{pmatrix} \tag{10}$$

and

$$w_3\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0.20 & 0 \\ 5.55 & -0.71 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0.80 \\ -3.83 \end{pmatrix} \tag{11}$$

with w_1 (used 36 % of the time), w_2 (36 %), and w_3 (28 %) operating into three regions defined by the endpoints $\{(0,0),(0,2.48)\}$, $\{(0.59,-4.07),(0.75,3.30)\}$, and $\{(0.80,-3.83),(1,1)\}$. As can be seen, the first mapping concentrates on a point and the projection over x is no longer a spiky multifractal. However, the projection over y yields yet another set with the typical features of a natural precipitation pattern. Notice both in the attractor and the projection dx the presence of gaps, due to the unaccounted region in x from 0 to 0.59 and from 0.75 to 0.8, that propagate through the iterations, ultimately resulting in a less "filling" attractor but a projection over y that contains no holes.

2.3 Extension 2: nonlinear perturbations

Another way to extend the original procedure is the addition of a small perturbation on each of the y-components of the affine mappings: (Cortis et al. 2009)

$$w_n\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a_n & 0 \\ c_n & d_n \end{pmatrix}\begin{pmatrix} x \\ y + A_n\cos(\omega_n y) \end{pmatrix} + \begin{pmatrix} e_n \\ f_n \end{pmatrix} \tag{12}$$

while keeping the same contractive initial conditions as in Eqs. (7) and (8). Figure 1c in Appendix 2 shows one such nonlinear example for a setting with distinct endpoints as in Sect. 2.2. Here, there are three regions, defined by the endpoints $\{(0,0),(0.44,5)\}$, $\{(0.22,-1.03),(0.56,-2.29)\}$, and $\{(0.66,-1.31),(1,1)\}$, in which the three maps

$$w_1\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0.44 & 0 \\ 4.38 & 0.59 \end{pmatrix}\begin{pmatrix} x \\ y + 0.88 * \cos(0.30y) \end{pmatrix} + \begin{pmatrix} 0 \\ -0.53 \end{pmatrix} \tag{13}$$

$$w_2\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0.34 & 0 \\ -1.49 & 0.18 \end{pmatrix}\begin{pmatrix} x \\ y + 0.26 * \cos(-1.51y) \end{pmatrix} + \begin{pmatrix} 0.22 \\ -1.07 \end{pmatrix} \tag{14}$$

and

$$w_3\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0.34 & 0 \\ 3.10 & -0.70 \end{pmatrix}\begin{pmatrix} x \\ y + 1.00 * \cos(-0.51y) \end{pmatrix} + \begin{pmatrix} 0.66 \\ -0.61 \end{pmatrix} \tag{15}$$

operate according to the proportions 22, 31, and 47 %, respectively. The attractor G yields a projection patterns over y that displays, once again, the typical features of natural precipitation as mentioned earlier and as will be appreciated later on.

2.4 Extension 3: higher dimensions

The examples thus far have yielded attractors in two dimensions, but the geometric procedure can be extended so that the attractor G is embedded in higher dimensional space. Introducing a z-coordinate into the first generalization of the wire (that is, Extension 1 from Sect. 2.2) results in a map in the form of

$$w_n\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a_n & 0 & 0 \\ d_n & r_n^{(1)}\cos\theta_n^{(1)} & -r_n^{(2)}\sin\theta_n^{(2)} \\ g_n & r_n^{(1)}\sin\theta_n^{(1)} & r_n^{(2)}\cos\theta_n^{(2)} \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} l_n \\ m_n \\ n_n \end{pmatrix} \tag{16}$$

which, coupled with three-dimensional analogues of Eqs. (7) and (8), yields marginal projections over y (and also z) that still resemble precipitation patterns. The example in Fig. 1d in Appendix 2, to be appreciated later, corresponds to such a construct when the three simple maps

$$w_1\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}0.21 & 0 & 0\\ -1.33 & 0.64\cos(1.55) & -0.69\sin(1.57)\\ 3.07 & 0.64\sin(1.55) & 0.69\cos(1.57)\end{pmatrix} \times \begin{pmatrix}x\\y\\z\end{pmatrix} + \begin{pmatrix}0\\0\\0\end{pmatrix} \quad (17)$$

$$w_2\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}0.32 & 0 & 0\\ -0.13 & 0.95\cos(1.54) & -0.53\sin(1.57)\\ -7.20 & 0.95\sin(1.54) & 0.53\cos(1.57)\end{pmatrix} \times \begin{pmatrix}x\\y\\z\end{pmatrix} + \begin{pmatrix}0.66\\5\\4.71\end{pmatrix} \quad (18)$$

and

$$w_3\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}0 & 0 & 0\\ -3.89 & 0.02\cos(-1.57) & -0.14\sin(1.21)\\ -0.77 & 0.02\sin(-1.57) & 0.14\cos(1.21)\end{pmatrix} \times \begin{pmatrix}x\\y\\z\end{pmatrix} + \begin{pmatrix}1\\5\\1.57\end{pmatrix} \quad (19)$$

are employed.

As may be appreciated, an increase in the dimensionality of the construction results in an increase in the number of parameters that specify the affine mappings (the vertical scaling parameters now become matrices, shown here in polar coordinates), but such does not alter the deterministic nature of the associated derived measures. As before, a complex-looking shape reminiscent of rainfall records can be seen by plotting only the results from x to y.

2.5 Solving the inverse problem

While it is straightforward to produce FM patterns, solving the corresponding inverse problem, that is, finding the mappings and associated FM parameters that reasonably encodes a given data set, has been far more challenging, due to several factors. Among them: (i) high dimensionality of the search space, (ii) ambiguity in the choice of the objective function to be minimized, (iii) complexity of the corresponding objective function landscape, and (iv) choice of the minimization algorithm, its parameters and search strategy.

More recently, we have found that a generalized particle swarm optimization algorithm (Fernández Martínez et al. 2010) yields excellent results for a storm in Boston (Huang et al. 2012), when minimizing for the ℓ^2-norm of averaged cumulative distribution deviations,

$$\text{err} = \sqrt{\frac{1}{N_0}\sum_{n=1}^{N_0}(c_n - \hat{c}_n)^2} \quad (20)$$

where N_0 is the number of data points (256 in this case), and c_n and $\hat{c}_n$ denote the nth value of the cumulative distribution of the original record and the generated fit, respectively. To ensure geometric similarity between plausible solutions and the original data set, we imposed penalties on the objective function: the maximum deviations in cumulative distributions to not exceed 10 % anywhere, the difference in arc lengths of the data must be less than 30 %, and the difference in arc lengths of the cumulative distribution must be less than 20 %.

3 Results and discussion

The data sets to be encoded in this section are part of the seven storms gathered by Georgakakos et al. (1994) using an optical rain gauge. The four storms used in this paper, labeled Iowa A through D per increasing number of major peaks, took place on November 30th, October 3rd, November 1st, and May 3rd of 1990 and correspond to events that lasted 12.2, 9.8, 9.3, and 9.3 h, respectively. These sets, originally gathered every five seconds, were slightly truncated to a multiple of 256, downscaled to such a resolution, and finally normalized to unity. The new time intervals for the data points are, in order, every 2 min 40 s, 2 min 15 s, 2 min 10 s, and 2 min 10 s.

In what follows, we discuss the encoding results for each of the four storms.

3.1 Iowa A

Figure 2 in Appendix 2 shows the results for the Iowa A data set using the four FM procedures in the order described in Sect. 2 (from top to bottom). The plots on the left are the best fits for the procedure (black) imposed on the data set (gray). For presentational purposes, the plot has been re-scaled such that the maximum value of the original data is 1. In the middle are plots of the cumulative distributions, with the data set shown as a black line and deviations shown as gray areas. On the right are scatterplots of the data (x-axis) versus the fits (y-axis), and the two gray bands denote ± 10 % deviation from the one-to-one line.

The most obvious first observation is that all FM sets, for all variants, tightly fit the flow of the original data and moreover, preserve the position and magnitudes of the

major peaks as desired. Such is confirmed by the extremely close following of the records' cumulative distribution by all FM algorithms and the very few points outside the 10 % bands in all scatterplots. The goodness of these results is confirmed through a host of statistical qualifiers as reported in Table 1 in Appendix 1.

As seen on the second row of the table, the search using a construction based on a wire with 3 maps (for a total of 9 parameters) produced a projection, already shown as in Fig. 1a in Appendix 2, that yields an optimized average error in cumulative distribution (cf. Eq. (20)) equal to 0.69 % that corresponds to a maximum deviation in cumulative distribution of 1.44 % and an average error in the data itself (not optimized) of 0.26 %. These numbers, clearly small, represent a significant improvement over our previous efforts (Obregón et al. 2002a) where, despite the visual similarities between the data and the FM representation, the maximum deviation exceeded 8 %.

The Nash–Sutcliffe model efficiency coefficients (1 minus the ratio of the sum of squared deviations between data and model to the sum of squared deviations of the data relative to its mean) for the data and its autocorrelation range between 0.71 and 0.88, values that reflect the close fits of the FM procedure. These are remarkable values when considering that a particular representation can achieved with as little as 9 parameters. The rest of the table includes a comparison of the fractal dimension, entropy dimension, power spectrum exponent, and autocorrelation lag times at $1/e$ and 0, which shows close agreement in all attributes.

The third to fifth rows in Table 1 in Appendix 1 include the same information for the three extensions of the FM approach, which require additional parameters. As surmised from Fig. 2 in Appendix 2, the statistics for these extensions are also excellent, with Extension 1 yielding the lowest average cumulative error (0.58 %), the lowest individual data error (0.23 %), the highest Nash–Sutcliffe model efficiency coefficient for both the data and the autocorrelation (0.76 and 0.96, respectively), and perfectly matching both autocorrelation lag times (9 and 18, respectively, for $1/e$ and 0).

Examination of the searches reveals that there are other solutions for the same storm that share very similar physical features despite having wildly different parameters, suggesting that there is no unique encoding for a single data set. Though omitted from the paper for the sake of brevity, the interested reader can obtain all parameters for the FM representations in this paper along with other information of interest in the online appendix at http://puente.lawr.ucdavis.edu/omake/serra2012.html.

3.2 Iowa B

In Fig. 3 in Appendix 2, we see fits for another storm with two major peaks, presented in the same format as Fig. 2 in Appendix 2, with the second row being the overlap/holes extension shown in Fig. 1b in Appendix 2. As before, while all shown representations closely retrace the path of the original data through periods of high and low activity before fading out, the last one (corresponding to Extension 3) has the best cumulative distribution fit and the least number of points (33 of 256, or 13 %) falling outside the 10 % bands.

As attested by the numbers in Table 2 in Appendix 1, all FM representations demonstrate excellent fidelity toward the original data. Notably, all average errors in cumulative distribution are less than 0.60 %, all average errors in individual data less than 0.20 %, and all maximum deviations in cumulative distribution less than 1.70 %. The Nash–Sutcliffe coefficients for data and autocorrelation are impressively high—more than 0.73 and as high as 0.84 for the former, and very close to 1 for the latter. To cap it up, all the other statistics presented in the table are tremendously close to that of the original data.

Again, there are additional suitable representations in other regions of parameter space not shown in this paper but available in the online appendix for Extension 1.

3.3 Iowa C

Figure 4 in Appendix 2 presents the results for a storm that opens with high activity followed by a secondary peak and contains noticeable periods of oscillatory behavior and gradually fades out with decreasing bumps, and the third row is the nonlinear extension shown in Fig. 1c in Appendix 2. Although the fits here may not appear visually as good as those for Iowa B, a close examination of the scatterplots reveals that for all representations, the number of points falling outside the 10 % bands are still less than the best one from Iowa B—in other words, even if the values are not precisely traced over, the representation is still in the vicinity of the original.

As seen in Table 3 in Appendix 1, while the statistics from fractal dimension to lag times of correlations show impressive agreement, the Nash–Sutcliffe coefficients for these fits are definitely not as good as the ones before, due to the increasing variability present in this data set not perfectly captured by the FM representation. Notice, however, that autocorrelation functions are properly fitted, as given by the NS numbers that range from 0.83 to 0.92.

Altogether, the objective function for this storm falls in between 0.61 and 0.66 %, which is within the range of such established by Iowa A and B. Note in particular that the maximum deviation in cumulative distribution for the higher dimensional storm (2.31 %) drastically increased the misfit in the objective function (0.66 %), as examination of the cumulative distribution plot (bottom of Fig. 4 in Appendix 2) suggests that it is otherwise well-fitted. It should also be noted that the higher dimensional fit uses only 2 maps (16 parameters), while all other fits described in the paper uses 3 or 4.

As previously noted, there exist more suitable fits in other regions of parameter space not shown here; for this data set, the online appendix includes additional examples using the nonlinear extension of the FM approach.

3.4 Iowa D

In Fig. 5 in Appendix 2, we have fits for a storm that contains significantly more activity than all the previous ones, having more than five noticeable peaks and a cumulative distribution that is closer to the one-to-one line. While the bottom row, corresponding to the higher dimensional extension shown in Fig. 1d in Appendix 2, exhibits the best visual fit in cumulative distribution, the scatterplots immediately reveal that this highly erratic storm is not as well-captured as the others. Even though the FM sets seem to follow in various degrees the overall trends and textures of the storm and the related cumulative distributions, it can be seen that the model based on a wire consistently underestimates the largest peaks yet its scatterplot exhibit less dispersion, while the other representations show more points away from the 10 % band. Overall, the representation based on Extension 1 appears to be best as it preserves the largest peaks better than the others.

The statistics in Table 4 in Appendix 1 confirm the nature of the results obtained: the Nash–Sutcliffe coefficients are now considerably smaller (with none of them exceeding 0.50 for the data and 0.90 for the autocorrelation), despite having maximum errors in cumulative deviation all less than 2 % and very reasonable results for the other statistics as well.

In the end, it can be argued that the FM models for this storm may still be considered as reasonable simulations for the storm, which is not free of measurement errors, as shown by Lanza et al. (2010).

As previously noted, there exist more suitable fits in other regions of parameter space not shown here; for this data set, the online appendix includes additional examples using the higher dimensional extension of the FM approach.

3.5 Scaling back to the original resolution

Figure 6 in Appendix 2 shows the cumulative distribution, autocorrelation, and log–log power spectrum of the overall best visual results in the 256-point resolution (corresponding to the fits using Extension 1) when upscaled to match the original data resolution, though without any additional optimization. In all plots, the gray indicate that of the original data, and the black denote the FM fit. Although the power spectrums of the storms are not perfectly matched, notice the closeness in cumulative distribution, and also how the autocorrelation function, the lag associated with the first crossing of zero correlation, and the positive/negative cycles are all extremely close to that of the original, despite it not being present in the objective function.

The FM fits shown iterate 4, 3, 3, and 4 maps respectively; the associated parameter counts are 19, 13, 13, and 19 parameters, which then translates to faithful encodings demonstrating compression levels of 431:1, 476:1, 512:1, and 350:1, and such ratios naturally increase with larger data sets. In conclusion, the good fits produced by the FM approach preserve the textures of typical rainfall sets with incredible parsimony even after scaled back to the original resolution. At the very least, the FM approach is a suitable tool for simulation of records (say fitting statistics of the records and not the records themselves) without the need of invoking the concept of "randomness." Certainly, the FM approach yields a multitude of possible solutions as attested for all storms in the online appendix, and we envision that analysis of additional sets may provide a classification scheme of patterns based on FM parameters.

4 Summary and conclusions

We have shown how the FM approaches, in conjunction with an effective optimization procedure for the inverse problem of a given set, can efficiently encode rainfall data. As anticipated in previous works (e.g., Puente 1996; Puente and Sivakumar 2007), this geometric procedure provides a practical alternative for handling erratic data, one that hints at the possibility of hidden determinism in natural complexity and one that surely adds to the view provided by stochastic approaches in order to best represent rainfall. The encoding procedures described herein may be used to understand rainfall dynamics in a reduced geometric space, potentially opening the way to new geometric (and perhaps even physical) descriptions of precipitation. Clearly, the FM approach may be safely used in simulation studies of watersheds just by varying certain parameters.

Furthermore, as the number of parameters for the FM approaches is few—ranging from 9 for a wire with three maps to 27 for Extension 3 with three maps—our 256-value representations result in compression ratios as high as 28:1, and moreover, since equally faithful solutions exist for the same data sets of higher resolution (as discussed in Sect. 3.5), compression ratios can easily exceed 300:1. These highly promising results represent a solid step forward with respect to past attempts (e.g., Obregón et al. 2002a) and substantiate the notion that complex rainfall patterns may be wholly characterized deterministically. Acknowledging the fact that precipitation measurements are inevitably plagued with nontrivial errors (Lanza et al. 2010), it can also be argued that all FM representations herein, with maximum deviation from cumulative distributions (at the original resolution) of no more than 2 % and

statistics that closely resemble those of the actual records, are acceptable renderings of what may happen in nature, for the overall trends and major peaks in all sets considered are sufficiently preserved.

Acknowledgments A. Cortis' work was supported in part by the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. The reviews we received led to relevant improvements to the manuscript and they are rightly acknowledged.

Appendix 1

See Tables 1, 2, 3, 4.

Table 1 Results of searching for a fit for the set Iowa A: ℓ^2-norm of averaged error on data, ℓ^2-norm of averaged cumulative distribution deviations, maximum deviation between cumulative distributions, Nash–Sutcliffe model efficiency coefficient for the data, Nash–Sutcliffe coefficient for the autocorrelation, fractal dimension, multifractal spectrum's entropy dimension, power spectrum exponent, autocorrelation lag at $1/e$, and autocorrelation lag at 0

Procedure	ErrData	ErrCuml	MaxDev	NS-Data	NS-Auto	FractalD	EntropyD	PowerSpc	Lag $1/e$	Lag 0
(Data)	n/a	n/a	n/a	n/a	n/a	1.33	0.91	–1.08	9	18
Original FM	0.26 %	0.69 %	1.44 %	0.71	0.88	1.39	0.90	–0.98	6	20
Extension 1	0.23 %	0.58 %	1.50 %	0.76	0.96	1.37	0.90	–1.36	9	18
Extension 2	0.25 %	0.71 %	2.16 %	0.73	0.95	1.35	0.91	–1.15	10	19
Extension 3	0.31 %	0.67 %	1.96 %	0.56	0.93	1.35	0.92	–1.13	6	18

The representations used 9, 19, 19, and 27 parameters, respectively

Table 2 Results of searching for a fit for the set Iowa B

Procedure	ErrData	ErrCuml	MaxDev	NS-Data	NS-Auto	FractalD	EntropyD	PowerSpc	Lag $1/e$	Lag 0
(Data)	n/a	n/a	n/a	n/a	n/a	1.36	0.94	–1.65	10	31
Original FM	0.16 %	0.58 %	1.34 %	0.83	0.96	1.37	0.95	–1.89	13	31
Extension 1	0.20 %	0.59 %	1.46 %	0.75	0.97	1.42	0.94	–1.44	10	34
Extension 2	0.20 %	0.58 %	1.69 %	0.73	0.97	1.44	0.94	–1.62	13	32
Extension 3	0.16 %	0.51 %	1.51 %	0.84	0.97	1.38	0.95	–1.72	13	36

Column descriptions are as in Table 1. The representations used 13, 13, 19, and 27 parameters, respectively

Table 3 Results of searching for a fit for the set Iowa C

Procedure	ErrData	ErrCuml	MaxDev	NS-Data	NS-Auto	FractalD	EntropyD	PowerSpc	Lag $1/e$	Lag 0
(Data)	n/a	n/a	n/a	n/a	n/a	1.37	0.93	–0.89	8	46
Original FM	0.26 %	0.62 %	1.74 %	0.62	0.90	1.43	0.93	–1.00	9	35
Extension 1	0.23 %	0.61 %	1.48 %	0.69	0.92	1.38	0.93	–0.94	8	49
Extension 2	0.31 %	0.66 %	2.09 %	0.46	0.92	1.44	0.92	–1.04	6	45
Extension 3	0.27 %	0.66 %	2.31 %	0.58	0.83	1.44	0.93	–1.64	9	17

Column descriptions are as in Table 1. The representations used 13, 13, 19, and 16 parameters, respectively

Table 4 Results of searching for a fit for the set Iowa D

Procedure	ErrData	ErrCuml	MaxDev	NS-Data	NS-Auto	FractalD	EntropyD	PowerSpc	Lag $1/e$	Lag 0
(Data)	n/a	n/a	n/a	n/a	n/a	1.48	0.96	–0.96	5	9
Original FM	0.21 %	0.68 %	1.80 %	0.49	0.64	1.56	0.97	–1.15	6	14
Extension 1	0.24 %	0.64 %	1.84 %	0.33	0.83	1.56	0.96	–1.00	5	11
Extension 2	0.24 %	0.76 %	1.95 %	0.34	0.82	1.49	0.96	–1.03	5	11
Extension 3	0.24 %	0.69 %	1.70 %	0.35	0.72	1.52	0.96	–0.93	6	12

Column descriptions are as in Table 1 in Appendix 1. The representations used 13, 19, 19, and 27 parameters, respectively

Appendix 2

See Figs. 1, 2, 3, 4, 5, 6.

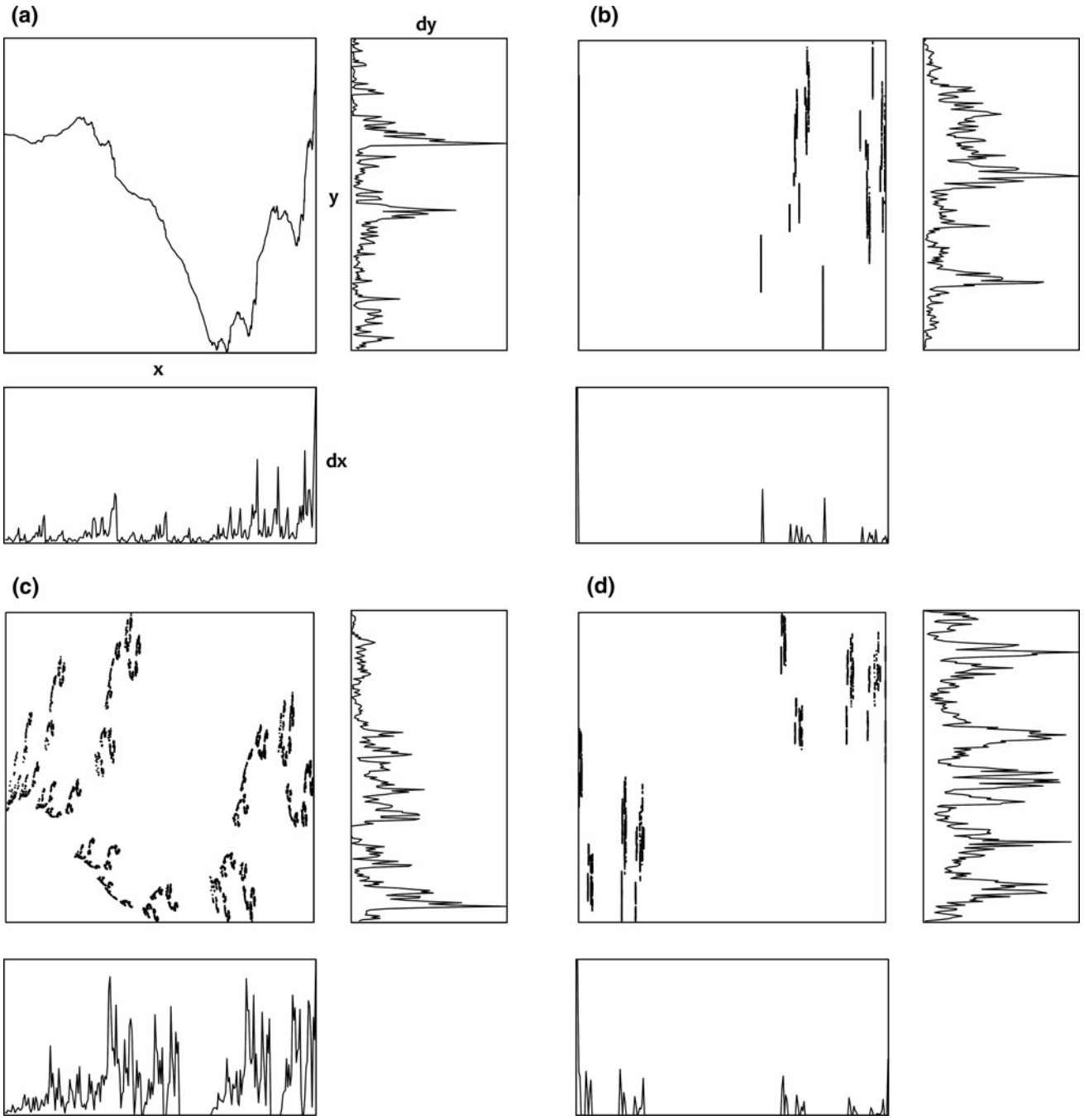

Fig. 1 The FM approach and three of its extensions. **a** *Top left* the original method: from a multifractal, *bottom*, to a projection, *right*, via a fractal interpolating function, center. **b** *Top right* an extension with overlaps and/or holes: from a multifractal, *bottom*, to a projection, *right*, via a "broken" attractor. **c** *Bottom left*, an extension with nonlinear perturbations: from a multifractal, *bottom*, to a projection, *right*, via an attractor with overlaps and holes generated with maps that contain nonlinear perturbations. **d** *Bottom right*, an extension to higher dimensions: from a multifractal, *bottom*, to a projection, *right*, of the y-marginal of an attractor in y and z

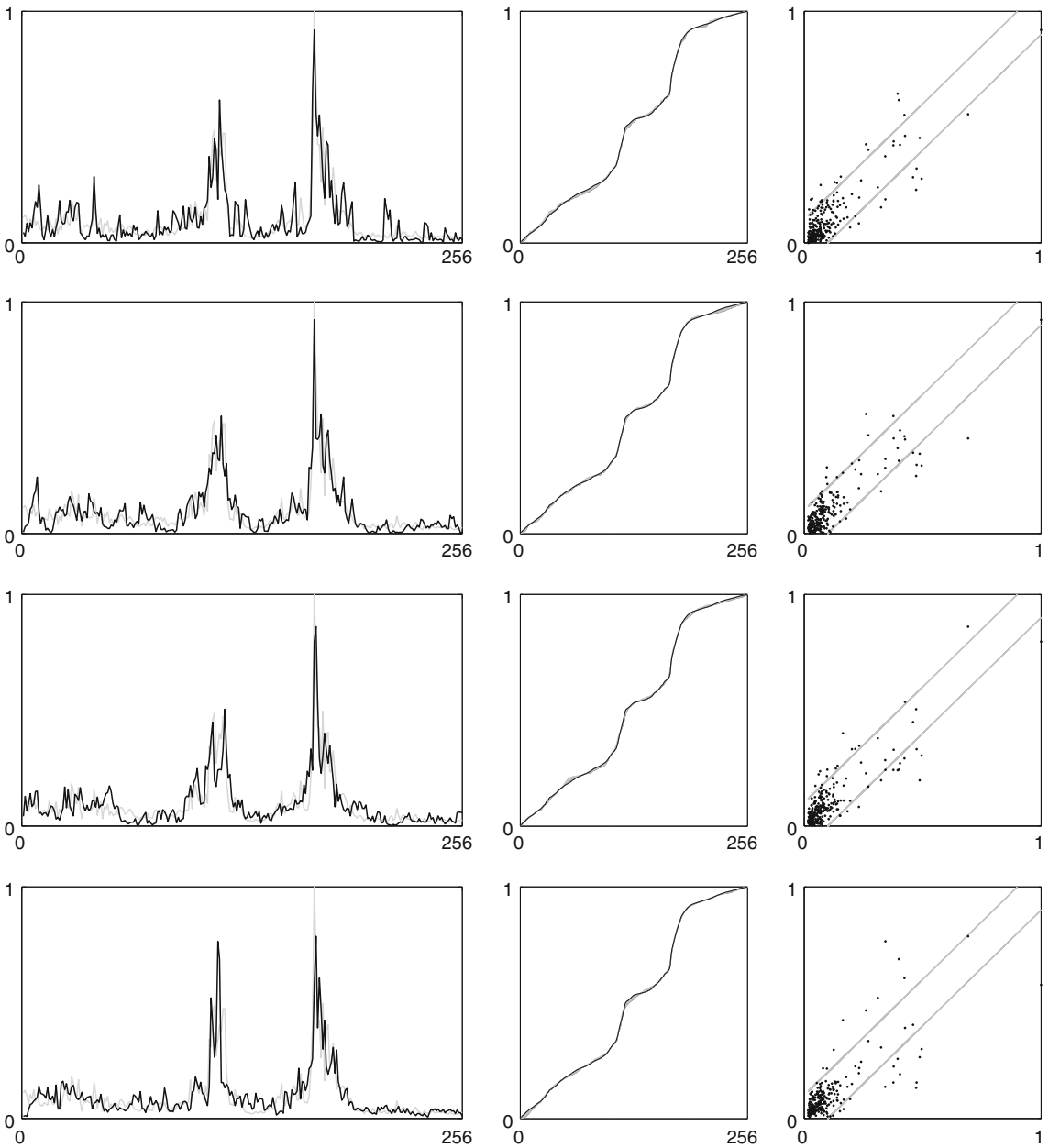

Fig. 2 Iowa A renderings. From *top to bottom*, representations using the original FM approach (using three maps), the extension with overlaps and holes (four maps), the nonlinear variant of the extension (three maps), and the higher dimensional extension (three maps). *Left column* the fit, in *black*, imposed on the original data, in *gray*. *Middle column* the cumulative distributions of the data and the fit, with errors *colored* as area between the *curves*. *Right column* scatterplot of the data (*horizontal*) versus the fit (*vertical*). *Diagonal lines* indicate ± 10 % lines. Readers interested in reconstructing the plots may navigate to the website, where all parameters may be found: http://puente.lawr.ucdavis.edu/omake/serra2012.html

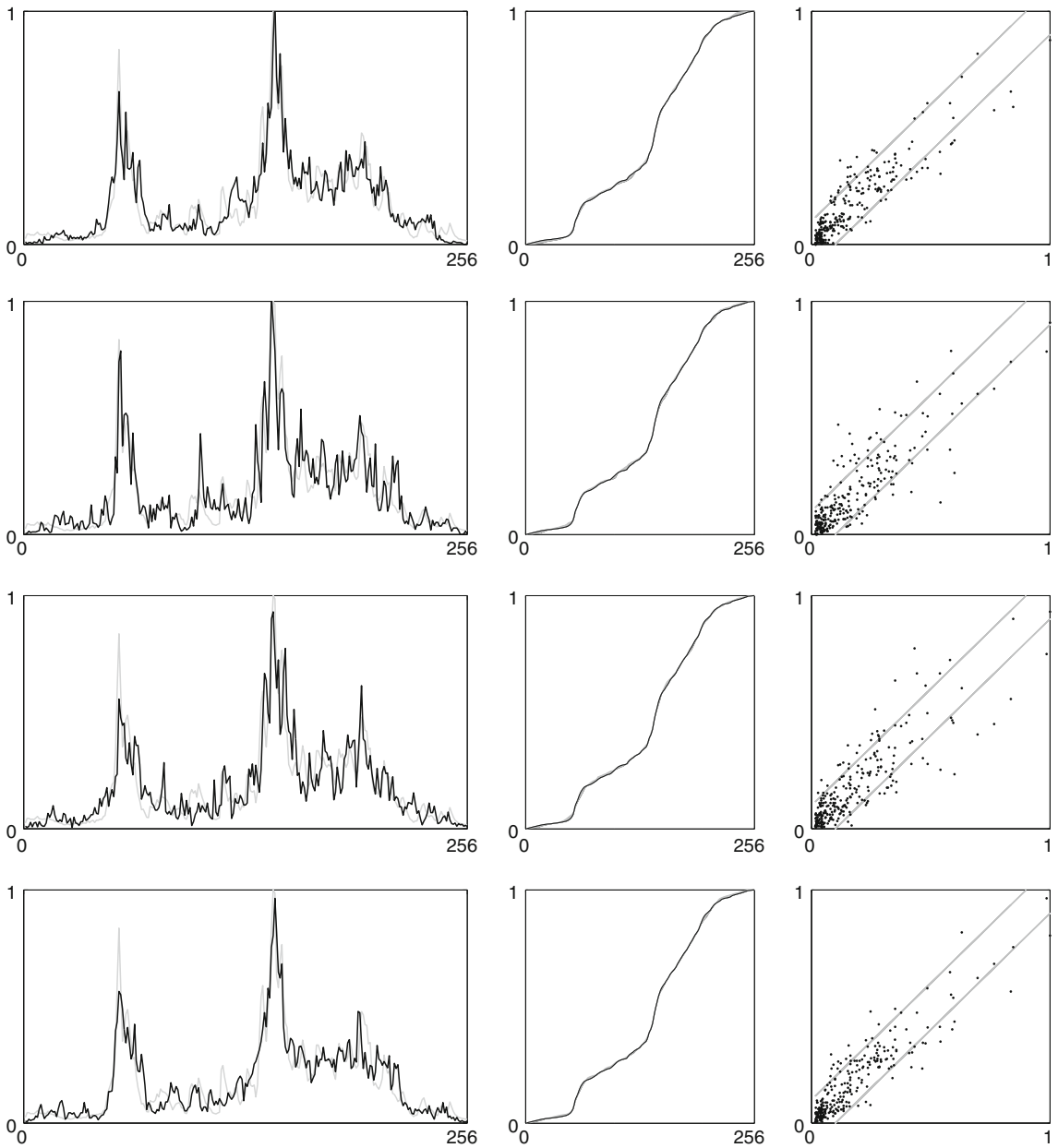

Fig. 3 Iowa B renderings. The *row* and *column* setup is as in Fig. 2 in Appendix 2. From *top to bottom*, the representations were constructed with four, three, three, and three maps

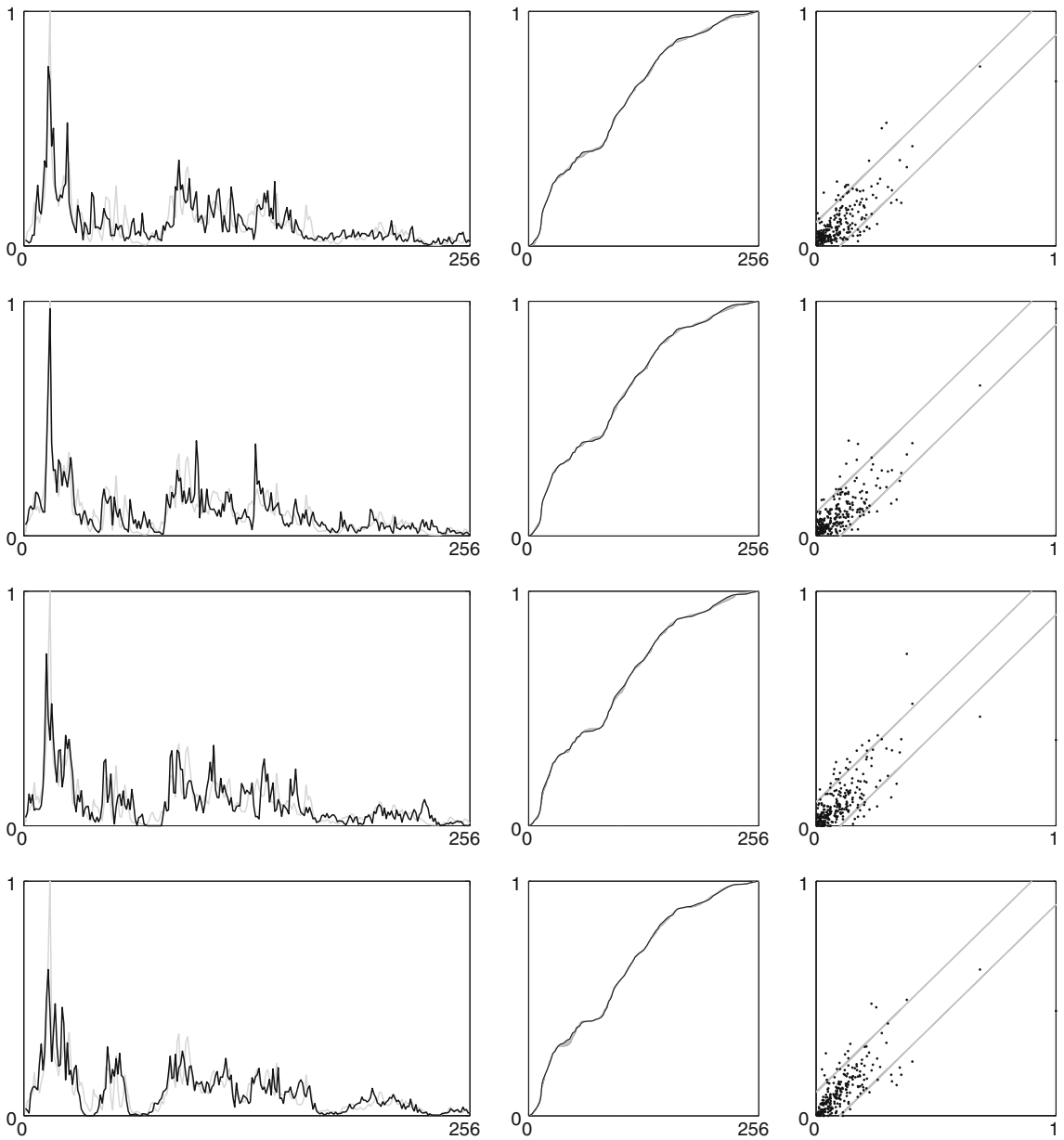

Fig. 4 Iowa C renderings. The *row* and *column* setup is as in Fig. 2 in Appendix 2. From *top to bottom*, the representations were constructed with four, three, three, and two maps

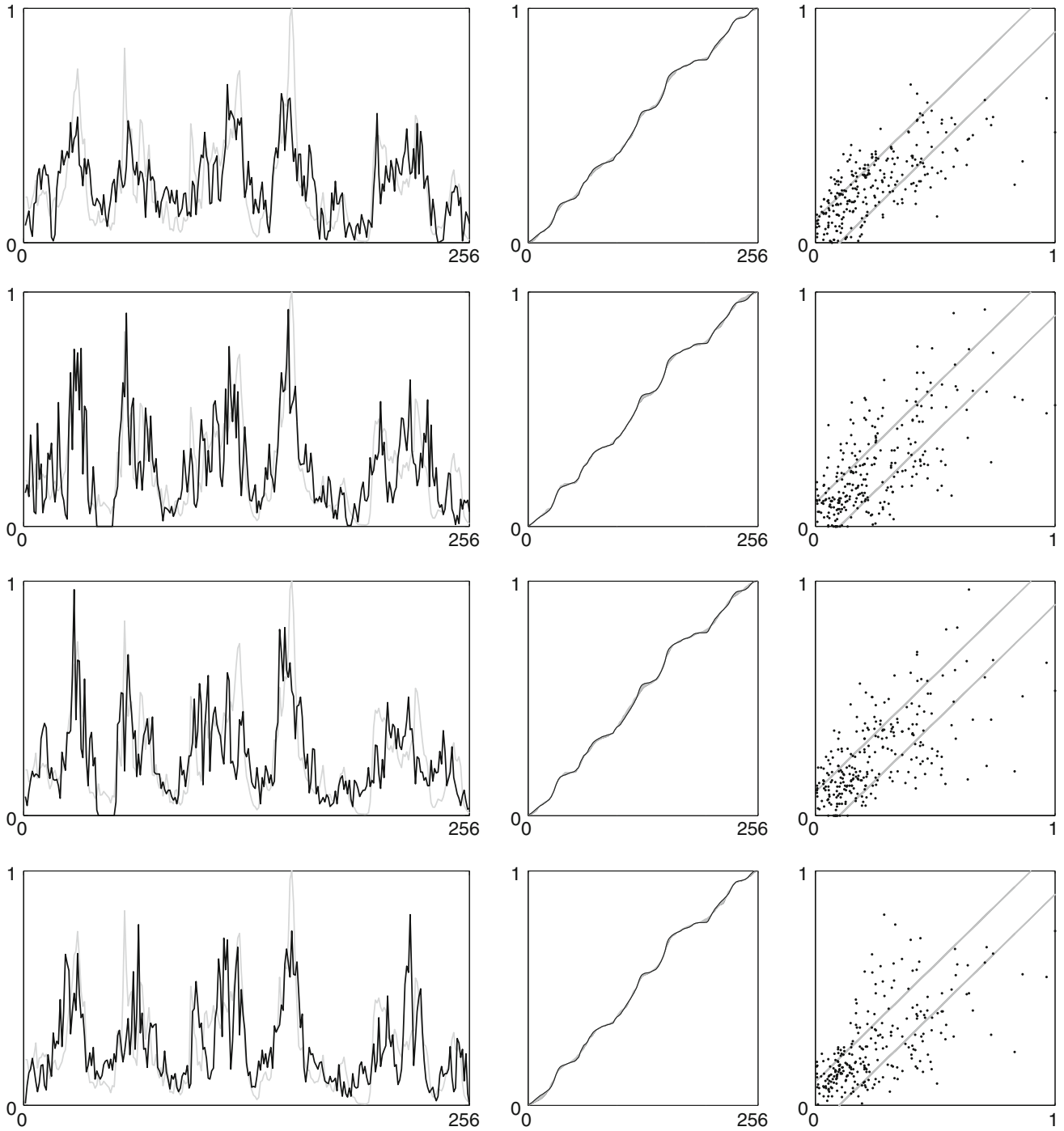

Fig. 5 Iowa D renderings. The *row* and *column* setup is as in Fig. 2 in Appendix 2. From *top to bottom*, the representations were constructed with four, four, three, and three maps

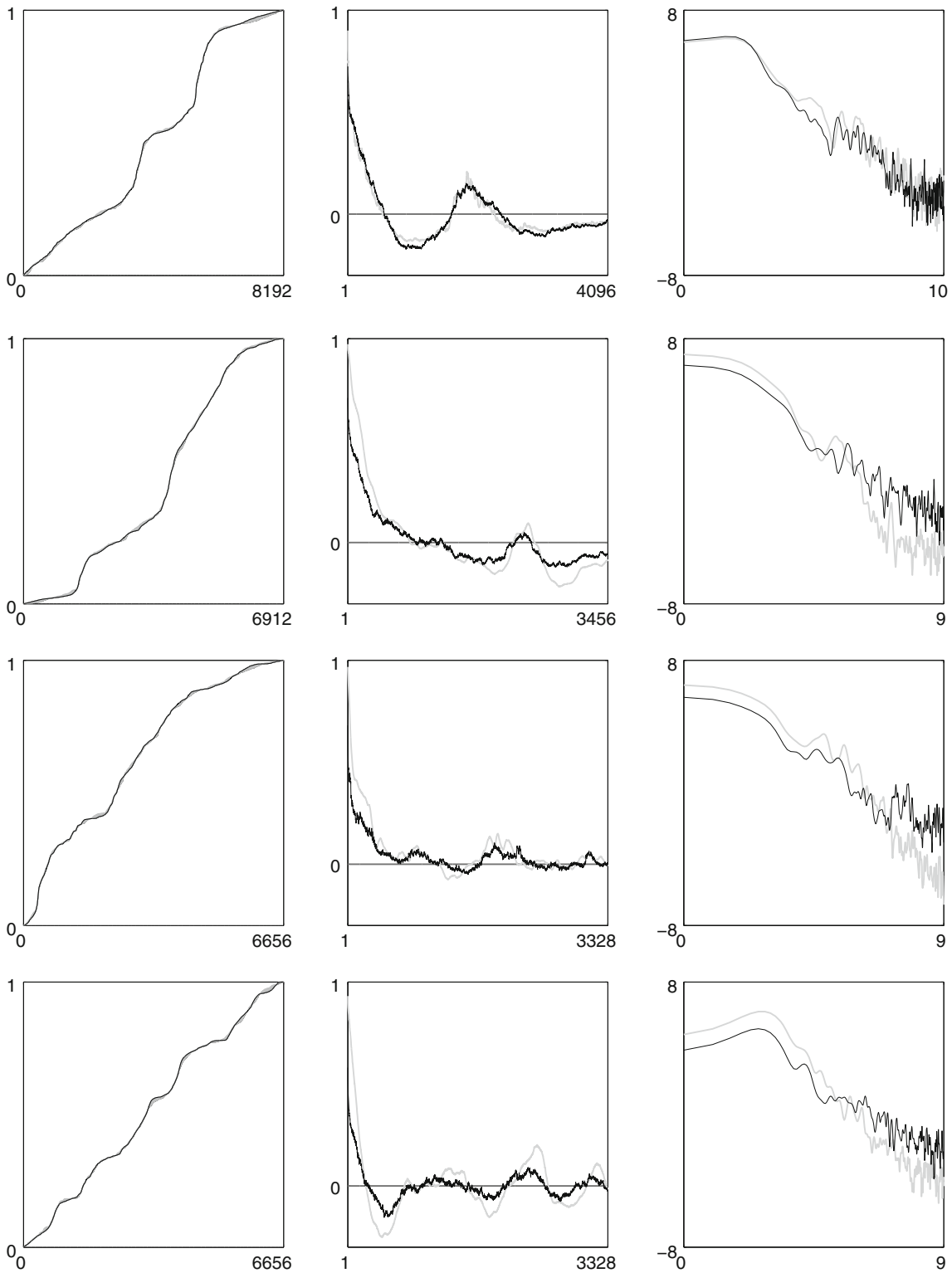

Fig. 6 High resolution renderings. Cumulative distributions, autocorrelation functions, and power spectra (plotted in log–log) for the four best renderings in *black*, imposed upon that of the original storms in *gray*. From *top to bottom*: Iowa A, B, C, D, all represented using the basic extension (Sect. 2.2) with four, three, three, and four maps

References

Barnsley MF (1988) Fractals everywhere. Academic Press, San Diego

Cortis A, Puente CE, Sivakumar B (2009) Nonlinear extensions of a fractal-multifractal approach for environmental modeling. Stoch Environ Res Risk Assess 23:890–906

Cortis A, Puente CE, Huang H (2012) A physical interpretation of the deterministic fractal-multifractal method as a realization of a generalized multiplicative cascade. Stoch Environ Res Risk Assess (in preparation)

Deidda R (2000) Rainfall downscaling in a space-time multifractal framework. Water Resour Res 36(7):1779–1794

Deidda R, Benzi R, Siccardi F (1999) Multifractal modeling of anomalous scaling laws in rainfall. Water Resour Res 35(6):1853–1867

Deidda R, Badas MG, Piga E (2004) Space–time scaling in high-intensity tropical ocean global atmosphere coupled ocean–atmosphere response experiment (TOGACOARE) storms. Water Resour Res 40. doi:10.1029/2003WR002574

Deidda R, Grazia-Badas M, Piga E (2006) Space–time multifractality of remotely sensed rainfall fields. J Hydrol 322:2–13. doi:10.1016/j.jhydrol.2005.02.036

Fernández Martínez JL, García Gonzalo E, Fernández Álvarez JP, Kuzma HA, Menéndez Pérez CO (2010) PSO: a powerful algorithm to solve geophysical inverse problems. Application to a 1D-DC resistivity case. J Appl Geophys 71(1):13–25

Georgakakos KP, Carteanu AA, Sturdevant PL, Cramer JA (1994) Observation and analysis of Midwestern rain rates. J Appl Meteorol 33(12):1433–1444

Huang H, Puente CE, Cortis A, Fernández Martínez JL (2012) An effective Inversion strategy for a fractal-multifractal encoding of a storm in Boston. Submitted to J Hydrol

Langousis A, Veneziano D (2007) Intensity–duration–frequency curves from scaling representations of rainfall. Water Resour Res 43. doi:10.1029/2006WR005245

Langousis A, Veneziano D, Furcolo P, Lepore C (2009) Multifractal rainfall extremes: theoretical analysis and practical estimation. Chaos Solitons Fractals 39:1182–1194. doi:10.1016/j.chaos.2007.06.004

Lanza LG, Vuerich E, Gnecco I (2010) Analysis of highly accurate rain intensity measurements from a field test site. Adv Geosci 25:37–44

Lovejoy S, Schertzer D (1990) Multifractals, universality classes and satellite and radar measurements of cloud and rain fields. J Geophys Res 95(D3):2021–2034

Marsan D, Schertzer D, Lovejoy S (1996) Causal space–time multifractal processes: predictability and forecasting of rain fields. J Geophys Res 101(D21):26333–26346

Menabde M, Sivapalan M (2000) Modelling of rainfall time series and extremes using bounded random cascades and Levy-stable distributions. Water Resour Res 36(11):3293–3300

Menabde M, Harris D, Seed A, Austin G, Stow D (1997) Multiscaling properties of rainfall and bounded random cascades. Water Resour Res 33(12):2823–2830

Obregón N, Puente CE, Sivakumar B (2002a) Modeling high resolution rain rates via a deterministic fractal-multifractal approach. Fractals 10(3):387–394

Obregón N, Sivakumar B, Puente CE (2002b) A deterministic geometric representation of temporal rainfall. Sensitivity analysis for a storm in Boston. J Hydrol 269(3–4):224–235

Puente CE (1996) A new approach to hydrologic modeling: derived distributions revisited. J Hydrol 187:65–80

Puente CE (2004) A universe of projections: may Plato be right? Chaos Solitons Fractals 19(2):241–253

Puente CE, Obregón N (1996) A deterministic geometric representation of temporal rainfall: results for a storm in Boston. Water Resour Res 32(9):2825–2839

Puente CE, Obregón N (1999) A geometric Platonic approach to multifractality and turbulence. Fractals 7(4):403–420

Puente CE, Sivakumar B (2007) Modeling hydrologic complexity: a case for geometric determinism. Hydrol Earth Syst Sci Discuss 11:721–724

Rodríguez-Iturbe I (1986) Scale of fluctuation of rainfall models. Water Resour Res 22(9):15S–37S

Rodríguez-Iturbe I, Febres de Power B, Sharifi M, Georgakakos KP (1989) Chaos in rainfall. Water Resour Res 25(7):1667–1675

Schertzer D, Lovejoy S (1987) Physical modeling and analysis of rain and clouds by anisotropic scaling of multiplicative processes. J Geophys Res 92:9693–9714

Tessier Y, Lovejoy S, Schertzer D (1993) Universal multifractal: theory and observations for rain and clouds. J Appl Meteorol 32:223–250

Tessier Y, Lovejoy S, Hubert P, Schertzer D, Pecknold S (1996) Multifractal analysis and modeling of rainfall and river flows and scaling, causal transfer functions. J Geophys Res 101(D21):26427–26440

Veneziano D, Langousis A (2005) The areal reduction factor a multifractal analysis. Water Resour Res 41. doi:10.1029/2004WR003765

Veneziano D, Langousis A (2005b) The maximum of multifractal cascades: exact distribution and approximations. Fractals 13(4):311–324

Veneziano D, Langousis A (2010) Scaling and fractals in hydrology. In: Sivakumar B, Berndtsson R (eds) Advances in data-based approaches for hydrologic modeling and forecasting. World Scientific, Singapore

Veneziano D, Langousis A, Furcolo P (2006) Multifractality and rainfall extremes: a review. Water Resour Res 42. doi:10.1029/2005WR004716

Veneziano D, Lepore C, Langousis A, Furcolo P (2007) Marginal methods of intensity–duration–frequency estimation in scaling and nonscaling rainfall. Water Resour Res 43. doi:10.1029/2007WR006040

Veneziano D, Langousis A, Lepore C (2009) New asymptotic and pre-asymptotic results on rainfall maxima from multifractal theory. Water Resour Res 45. doi:10.1029/2009WR008257