Stoch Environ Res Risk Assess
DOI 10.1007/s00477-009-0349-4

ORIGINAL PAPER

Encoding hydrologic information via a fractal geometric approach and its extensions

Andrea Cortis · Carlos E. Puente · Bellie Sivakumar

© The Author(s) 2009. This article is published with open access at Springerlink.com

Abstract We present the application of a deterministic fractal geometric approach—the so-called fractal-multifractal procedure, FMFP—to the modeling of hydrologic data at different resolutions. The FMFP can generate a wide range of complex patterns that are virtually indistinguishable from observed hydrologic data sets (e.g., rainfall series, radar images, clouds, contamination plumes, width functions). We illustrate the use of the FMFP for hydrologic data encoding and model simplification by comparing a few representative rainfall time series to FMFP-generated patterns. We also present the time evolution of two-dimensional FMFP-patterns reminiscent of rainfall-radar images. As the deterministic FMFP-generated patterns are completely characterized by a small number of geometric parameters, we discuss the prospect of compact descriptions of hydrologic data sets. We also discuss how this parsimonious deterministic parameterization may eventually lead to the classification of patterns and simplification in the records' parameter space. Finally, we highlight some connections between the FMFP and nonlinear and chaotic dynamics.

Keywords Hydrologic modeling · Rainfall time series · Radar images · Scale · Fractals · Multifractals · Chaos

A. Cortis
Earth Sciences Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA
e-mail: acortis@lbl.gov

C. E. Puente (✉) · B. Sivakumar
Department of Land, Air and Water Resources, University of California, Davis, CA 95616, USA
e-mail: cepuente@ucdavis.edu

B. Sivakumar
e-mail: sbellie@ucdavis.edu

1 Introduction

Modeling the dynamics of hydrologic information, rainfall in particular, constitutes one of the most important unsolved problems in hydrology today. During the past few decades, a great deal of research has been devoted to gathering vast sets of hydrologic records and to developing a host of hydrologic models. Rainfall models, specifically, may be roughly divided into four classes: (i) physically-based models (e.g. Georgakakos and Bras 1984), (ii) stochastic point process models (e.g. Rodríguez-Iturbe 1986; Morrissey 2009), (iii) stochastic cascade models (e.g. Lovejoy and Schertzer 1990; Gaume et al. 2007), and (iv) low-dimensional chaotic models (e.g. Sivakumar et al. 2001). Despite considerable progress, however, these representations do not account for all relevant information and, hence, predictions having non-trivial errors still remain. Stochastic and chaotic models of temporal rainfall reproduce, for instance, important characteristics of the records, such as moments, correlations, and/or fractal and chaotic properties, but typically fail to capture the specific details of rainfall series, such as the position of the peak with respect to the beginning of the event, or periods of no activity during the event itself. The incorrect prediction of these important features often results in the underestimation of observed extremes.

In the last few years, we have developed a novel geometric modeling approach aimed at representing complex hydrologic (and geophysical) data sets as projections of fractal functions illuminated by multifractal sets, the fractal-multifractal procedure (FMFP) (e.g. Puente 1992, 1996, 2004) and its extensions (e.g. Cortis et al. 2009). Contrary to the most common approaches mentioned before, the fractal-multifractal approach (and its extensions) is an entirely deterministic procedure that aims at capturing not only the

Published online: 11 October 2009

important statistics of the records but also their entire geometric details (e.g. Puente and Sivakumar 2007).

To date, the FMFP has been applied in the context of hydrologic applications, to: (i) one-dimensional data sets (e.g. rainfall, turbulence, width function, chaotic signals) as measured in time at a given site (e.g. Puente and Obregón 1996, 1999; Puente et al. 2002; Puente and Sivakumar 2003), and (ii) two-dimensional spatial patterns (radar images, contamination plumes) as gathered over a region (e.g. Puente et al. 2001a, b).

In this work, we first review the fundamentals of the fractal-multifractal method and its possible extensions, and then illustrate its application to the hydrologic sciences showing a series of synthetic rainfall data series (at different resolutions) and synthetic radar-rainfall images.

2 The fractal-multifractal procedure (FMFP) and its extensions

The key elements of the original FMFP are illustrated in Fig. 1a. A parent multifractal distribution dx (bottom left) is transformed by a fractal interpolating function f, from x to y, (upper left) to generate the derived distribution dy (upper right).

A fractal interpolating function is constructed by iterating at least two contractive affine maps of the form (Barnsley 1988):

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

The particular function shown in Fig. 1a is obtained from two such maps, setting the parameters $d_1 = -0.8$, $d_2 = -0.6$, and finding all the others, that is, a_n, c_n, e_n, and f_n, from the boundary conditions:

$$w_1\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, w_1\begin{pmatrix} 1 \\ -0.2 \end{pmatrix} = \begin{pmatrix} 1/2 \\ -0.35 \end{pmatrix},$$
$$w_2\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/2 \\ -0.35 \end{pmatrix}, w_2\begin{pmatrix} 1 \\ -0.2 \end{pmatrix} = \begin{pmatrix} 1 \\ -0.2 \end{pmatrix} \tag{2}$$

that ensure that the resulting attractor passes by the three points marked as circles, namely {(0, 0), (1/2, −0.35), (1, −0.2)}.

By iterating such two maps in a random fashion according to a 0.3–0.7 proportion, say 2^{18} times, and by keeping track of the points obtained, one may define the graphs in Fig. 1a as follows. The distribution dx is found projecting the set of points over the x-axis and finding its histogram, which, given the decoupled nature of the maps in x, corresponds to a binomial multifractal measure as obtained from a deterministic multiplicative cascade (Mandelbrot 1989). The distribution dy is similarly obtained as a histogram of projected values over the y-axis, that is hence defined as a (fractional) integration of the measure in x by adding contributions associated with the crossings of the function f, or in other words, dy is the derived distribution of dx as defined via the function f, which itself is obtained just by plotting the 2^{18} (x, y) points.

The set of points in the x–y plane define indeed a non-trivial function that is typically non-monotonic and not one-to-one (Barnsley 1988). However, such a set is a stable attractor for the affine maps, irrespective of the random order for the iterations; and, hence, the geometric construction just illustrated turns out to be fully deterministic. The parameters a_n and d_n determine the fractal dimension of the graph of the function f (Barnsley 1988), which for the example has a value of 1.48. As may be appreciated, the parameters for Fig. 1a were chosen to let the distribution dy resemble a rainfall time series, and other interesting sets may be found from other parameter values and from the iteration of more than two affine maps (Puente 2004).

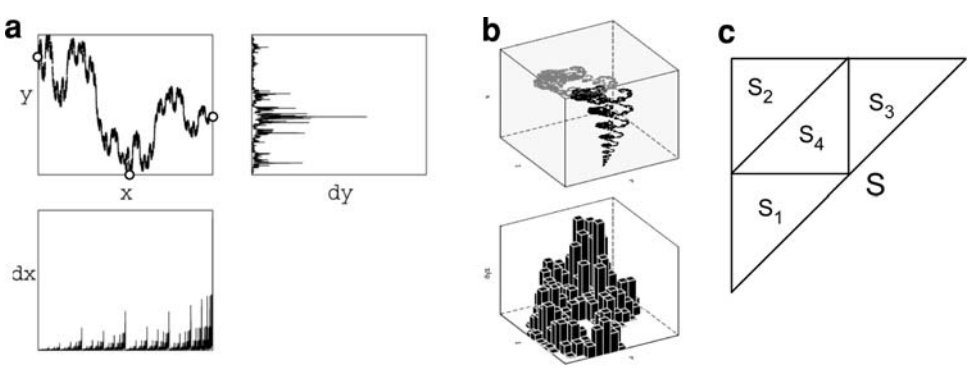

Fig. 1 The FMFP and its extensions. **a** The original framework in two dimensions: from a parent distribution dx to a derived distribution dy via a fractal interpolating function f from x to y. **b** The FMFP framework in three dimensions: a fractal interpolating function f from x to the plane y–z (*top*), and a derived joint distribution over the plane y–z (*bottom*). **c** Domain partition for a four-dimensional extension

Multifractal distributions, in particular the one shown as dx in Fig. 1a, are ubiquitous in turbulent phenomena (e.g. Meneveau and Sreenivasan 1987). For this reason, the distribution dy shown may be given an informal "physical" interpretation as a "projection" or "reflection" of turbulence. In this light, one may interpret the fractal function f as a summary of the travel times it takes rainfall-producing particles to wander about in the atmosphere until they land at the site under consideration. As the transformation on our geometric procedure conserves mass, such raindrops (originally distributed as in the generic turbulent cascade) 'walk' and 'coalesce' due to turbulent eddies and finally arrive at the raingage in distinct travel paths reflecting the specific complex geometries observed.

The FMFP may be extended in a variety of ways in order to define yet other interesting derived patterns. A suitable way to do this is to add a bounded nonlinear perturbation $g(y)$ on the y-component of the affine maps (Cortis et al. 2009):

$$w_n\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a_n & 0 \\ c_n & d_n \end{pmatrix}\begin{pmatrix} x \\ y + g(y) \end{pmatrix} + \begin{pmatrix} e_n \\ f_n \end{pmatrix}, \tag{3}$$

say $g(y) = A\ cos(\omega\ y)$, while keeping the boundary conditions, as in Eq. 2, fixed. As shall be illustrated later on, such an extension yields other interesting distributions that may also be used to model hydrologic phenomena, and in particular rainfall time series.

Another way to extend the FMFP is to use affine maps defined over more than two dimensions such that they generate fractal interpolating functions in higher dimensions (e.g. Barnsley 1988; Puente and Klebanoff 1994). As illustrated in Fig. 1b, a construct in three dimensions involves a wire-like fractal function f (from x in the vertical to the horizontal plane y–z) that allows the calculation of marginal measures over the y and z axes and also joint derived distributions over the y–z plane. Application of these notions to model two-dimensional rainfall patterns will be discussed in the next section.

Yet another plausible extension of the ideas is to replace the domain of the fractal functions from one to two dimensions, so that instead of generating wires one obtains hyper-surfaces in four dimensions, from say the plane x–y to the plane z–t. As illustrated in Fig. 1c, the simplest of such representations entails the usage of four maps that generalize the boundary conditions given in Eq. 2 for two maps, that is, such that mapping a large triangle S into the four sub-triangles S_k leads to a hyper-surface that passes by six points in a four-dimensional space. At the end, this extension yields joint derived distributions over the plane z–t and also marginal derived distributions over the z- and t-axis that, as shall be seen later on, may also be used to model, in particular, rainfall sets, and surely others, in one and two dimensions.

3 FMFP-generated rainfall time series and radar images

Analytical derivations of the derived distributions (and of their statistical-multifractal properties) are extremely difficult to obtain, even for the original FMFP (Puente 1992). Numerical experimentation is thus essential to study the many patterns that can be obtained by varying the geometric parameters of the FMFP. In this section, we illustrate how the FMFP and its variants can be used to simulate rainfall time series at a variety of resolutions. To this end, examples over one and two dimensions are shown next.

Figures 2 and 3 present typical rainfall time series when sampled every few seconds and every few hours, respectively, together with their corresponding autocorrelation functions, power spectra, and multifractal spectra.

While Fig. 2a corresponds to an actual rainfall event recorded in Boston every 15 s and for a total of about 8.3 h of observations, the other three sets in Fig. 2 are FMFP-generated sets that, as may be readily appreciated, have similar textures and intermittencies as the Boston storm. These three FMFP "storms" are generated as follows. The one labeled **b** is simply the one reported in Fig. 1a via the original FMFP whose fractal function passes by three points. The one labeled **c** is obtained via the FMFP also defined from an attractor passing by three points, but adding a nonlinear cosine perturbation with magnitude $A = 0.5$ and frequency $\omega = 1$. And the one labeled **d** corresponds to a one-dimensional marginal projection as generated extending the FMFP, so that it creates a hyper-surface passing by six points. Due to space limitations, the final parameters used to generate these patterns are not reported here but can be found at http://puente.lawr.ucdavis.edu/serra_09.htm.

The signals in Fig. 2 exhibit all the distinctive features of rainfall time series, and clearly it is not possible to discern the synthetic from the measured. This is also evident from a comparison of their statistical and multifractal characteristics. The four sets share in fact similar decays in their autocorrelation functions and reasonable correlation scales τ_e (the lag where $1/e$ is reached) ranging from 80 to 200 lags; all signals exhibit power-law scaling in their power spectrum ($S(w) \sim w^{-\beta}$) with spectral exponents β that include common values ranging from 1.15 to 1.37 (computed up to a logarithmic frequency scale of 2.2); all sets possess similar multifractal properties as reflected by their parabolic multifractal spectra (f vs. α curve), leading to similar entropy dimensions D_1 ranging from 0.85 to 0.96, as defined by the intersection of the parabola and the $f = \alpha$ line; and also it can be shown (not included) that all signals define low-dimensional chaotic dynamic systems of similar dimensions (Puente and Obregón 1996; Puente et al. 2002). As the sets in Fig. 2 cannot be distinguished from one another regarding

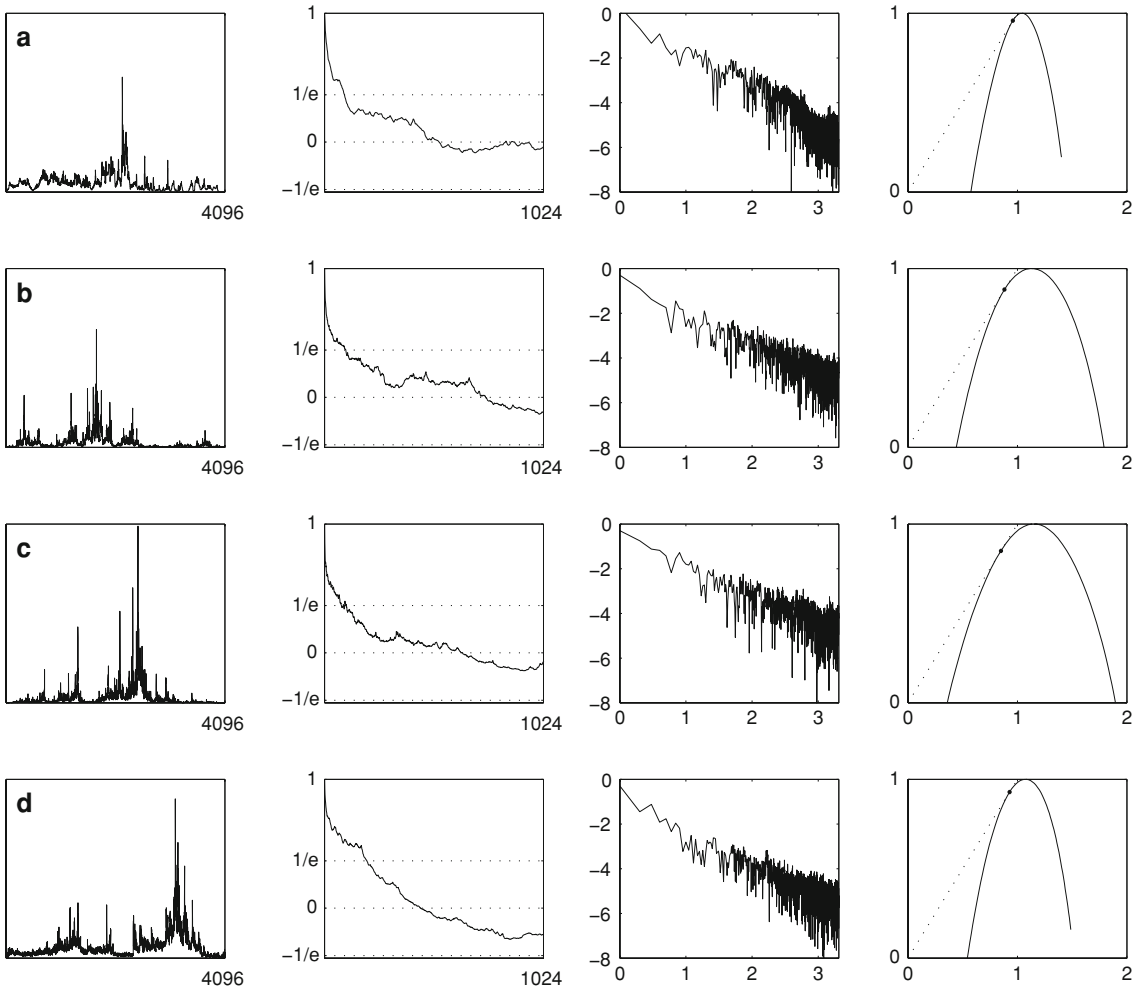

Fig. 2 One-dimensional signals and their autocorrelations, power spectra (log–log) and multifractal spectra. **a** Real rainfall measured in Boston every few seconds; **b** A FMFP pattern via the original formulation (Eqs. 1, 2); **c** A FMFP pattern via a nonlinear cosine perturbation extension; **d** A FMFP pattern from a hyper-surface extension. Key statistics (*top to bottom*): $\tau_e = 95,\ 101,\ 80,\ 200$; $\beta = 1.33,\ 1.15,\ 1.36,\ 1.37$; $D_1 = 0.96,\ 0.88,\ 0.85,\ 0.93$. The parameters used to generate these patterns may be found at the web page http://puente.lawr.ucdavis.edu/serra_09.htm

the overall features on such relevant qualifiers, we conclude that the FMFP framework can be used to simulate and encode detailed hydrologic information, as found in rainfall data sets gathered every few seconds.

When rainfall is sampled at a finer (i.e. coarser) resolution, say hours or more, periods of no rain activity are typically observed. It is, therefore, legitimate to ask if the FMFP can reproduce these characteristics of rainfall as well. This can, in fact, be achieved by defining fractal functions that interpolate four or more points (i.e., by iterating three or more maps) while at the same time quenching the relative impact of a selected subset of these maps (e.g., iterating three maps according to frequencies, say 0.65-0-0.35). By means of this generalization, the parent distribution dx so constructed contains a sequence of cascading holes, yielding a Cantorian distribution in x. Depending on the character of the function f, the derived distribution dy can also exhibit holes (not necessarily of a Cantorian nature) that may be used explicitly to model periods of rainfall inactivity.

Figure 3 presents a few examples of signals that exhibit periods of no rain. While the pattern labeled **a** represents actual rainfall reported every hour at La Honda, California for a period of 171 days, sets labeled **b**, **c**, and **d** show FMFP-generated patterns of similar durations. Pattern **b** corresponds to the original FMFP with a fractal interpolating function passing by four points, and obtained by iterating only two out of the three corresponding affine maps. Patterns **c** and **d** are also obtained by iterating only two out of three maps, but are defined by the addition of a nonlinear cosine perturbation as used in pattern **c** of Fig. 2. Moreover, these last two projections have been computed at an angle not perpendicular to x, a feature that provides yet another generalization of our approach.

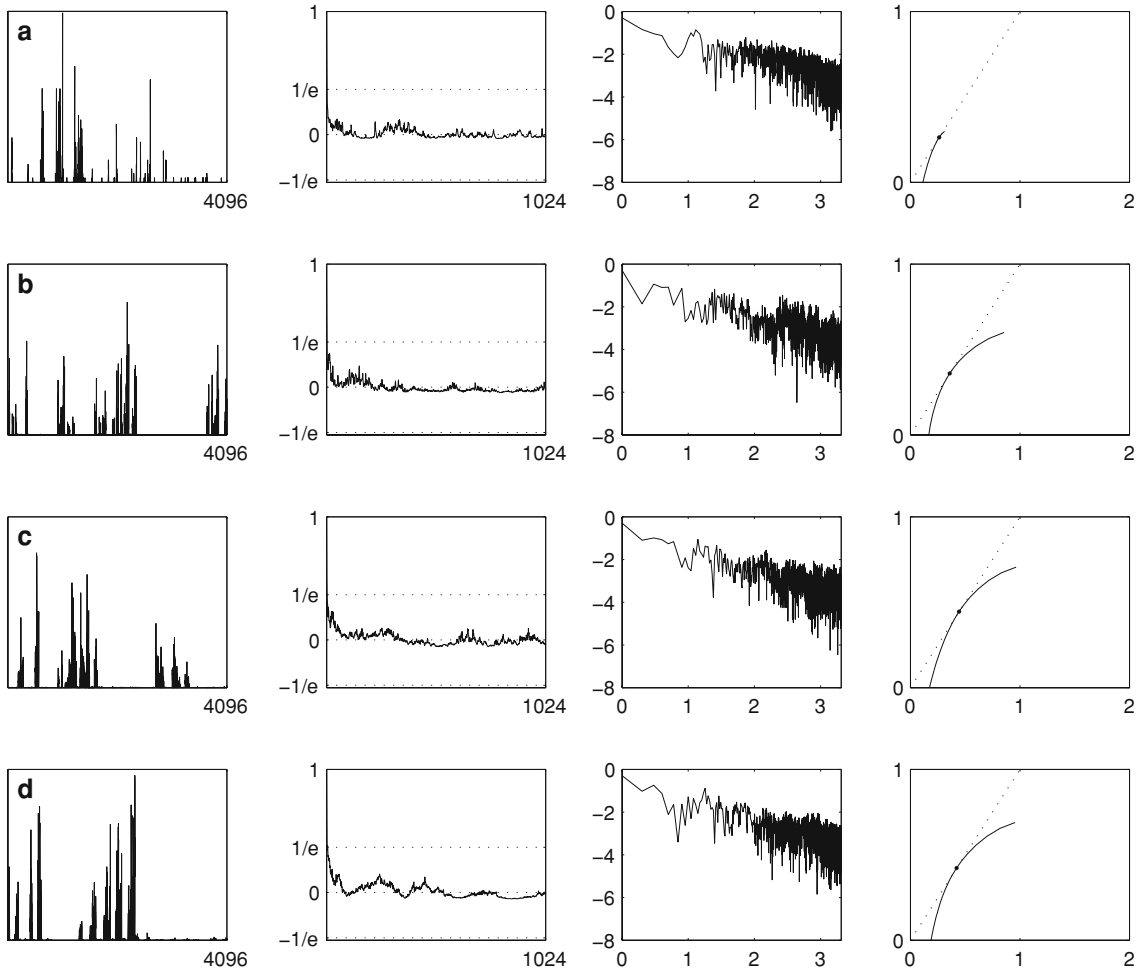

Fig. 3 One-dimensional intermittent signals and their autocorrelations, power spectra (log–log) and multifractal spectra. **a** Real rainfall data as measured in La Honda, California; **b** A FMFP pattern via the original formulation; **c** A FMFP pattern via a nonlinear cosine perturbation extension; **d** Another FMFP pattern via a nonlinear cosine perturbation extension. Key statistics (*top to bottom*): $\tau_e = 4$, 4, 4, 4; $\beta = 0.91$, 0.84, 0.71, 0.80; $D_1 = 0.26$, 0.36, 0.44, 0.42. Other statistics (*top to bottom*): number of events above 10% of maximum value in pattern **a** $N_e = 111$, 104, 95, 108; maximum number of consecutive times with no rain $N_{nr} = 1140$, 1330, 1116, 1719; lag-one events from rain to no rain $N_{rnr} = 53$, 57, 60, 60. The parameters used to generate these patterns may be found at http://puente.lawr.ucdavis.edu/serra_09.htm

The four patterns in Fig. 3 not only share similar textures and intermittencies, but also comparable statistical-multifractal attributes as reflected in identical correlation scales τ_e, similar spectral exponents β that range from 0.71 to 0.91 (computed up to a logarithmic frequency scale of 3), and similarly low entropy dimensions D_1 ranging from 0.26 to 0.42. These patterns also share similar numbers of events (as defined above 10% of the maximum value in pattern **a**) ranging from 95 to 111, similar sizes on the maximum period with no rain that range from 1116 to 1719 values, and comparable number of events reflecting lag-one transitions from rain to no rain ranging from 53 to 60. Notice how, as previously found for high-resolution data, it is not possible to tell apart the real data set from the FMFP-generated patterns, as all of them share the same essential features, e.g., correlation scales, power spectrum scaling, entropy dimensions, number of events, maximum lengths of no rain (even at other thresholds not equal to 10%), and number of events from rain to no rain (even for lag separations other than one).

The FMFP framework may also be used to generate the two-dimensional complex-looking patterns observed, for instance, in a meteorological radar. A proof of concept of this approach is presented in Figs. 4 and 5, for a couple of pattern evolutions obtained via three-dimensional wires and four-dimensional hyper-surfaces, respectively. Observe how both evolutions, which interpolate linearly sets of parameters leading to the patterns on the top left and bottom right of both figures, produce rather complex images that indeed exhibit typical features of radar-rainfall spatial measurements. Notice, in particular, the non-trivial partition of a rainfall pattern into several disconnected sub-sets as

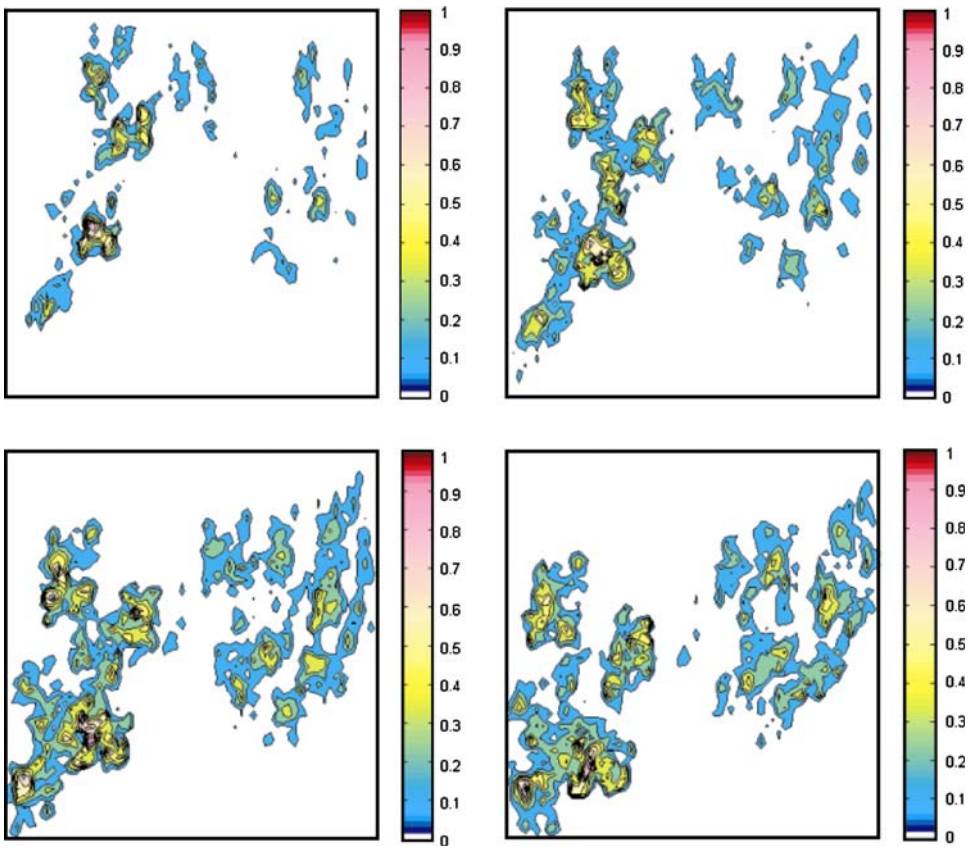

Fig. 4 Two-dimensional sets resembling complex spatial rainfall patterns obtained via the FMFP leading to a wire in three dimensions. For the specific parameters used in this graph, please visit http://puente.lawr.ucdavis.edu/serra_09.htm

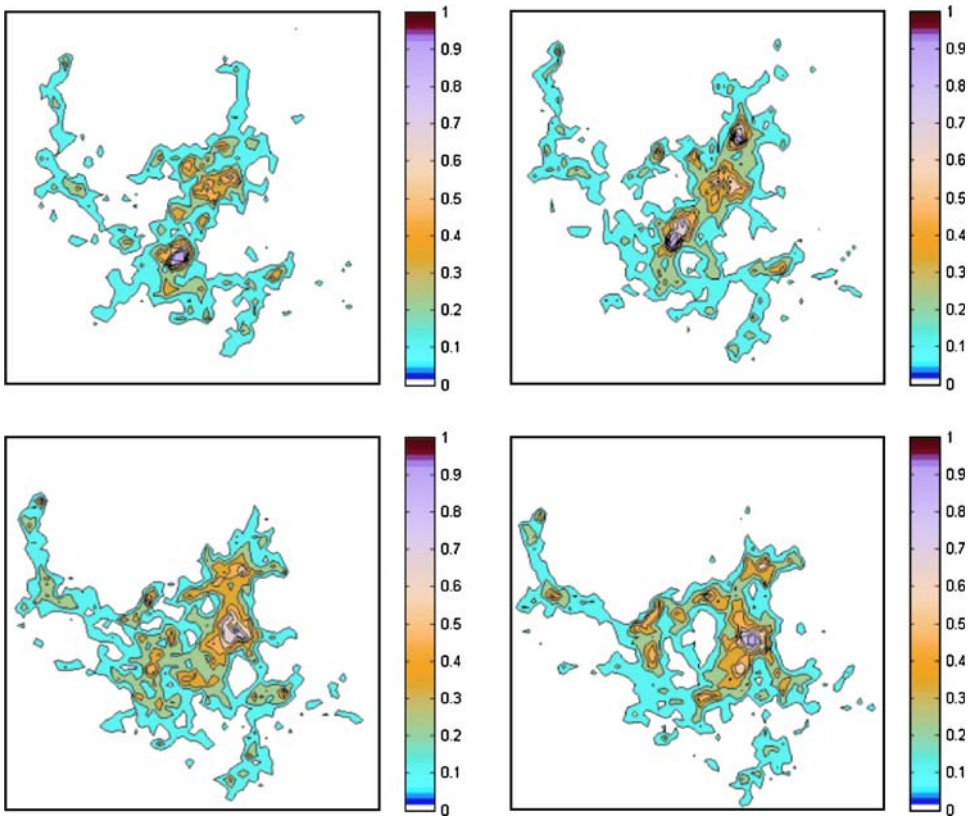

Fig. 5 Two-dimensional sets resembling complex spatial rainfall patterns obtained via the FMFP leading to a hyper-surface in four dimensions. For the specific parameters used in this graph, please visit http://puente.lawr.ucdavis.edu/serra_09.htm

depicted in Fig. 4 and the emergence of filaments and holes over a range of scales as shown in Fig. 5.

4 Implications of the FMFP for hydrologic modeling

That the FMFP, and its variants, may reproduce the features exhibited in Figs. 2–5 has indeed some relevant implications. First, our results imply that describing the complex rainfall patterns (and others) as "random," as it is customarily the case, may perhaps be unwarranted in each and every situation. Second, and pending the efficient resolution of the inverse FMFP problem, these results suggest the possibility of fully describing (encoding) hydrologic data sets by means of a small number of geometrical parameters. This fact has momentous consequences for data archiving and assimilation as the compression ratios are estimated to easily exceed 100:1. Third, and once again contingent on the aforementioned inverse problem solution, we foresee that our fractal-multifractal framework may be used to device a pattern classification scheme based on the FMFP geometric parameters (perhaps similar to a classification scheme based on attractor and dimensionality (e.g. Sivakumar et al. 2007), which may lead to an increased understanding of hydrologic and climatic regimes and the interrelation between rainfall and other relevant climatic variables. Finally, as the analysis of derived distributions leads to the identification of 'chaotic' and also 'random' dynamics for suitable regions in the FMFP parameter space (Puente et al. 2002) and as one may perhaps study the dynamics of rainfall via the successive FMFP parameters corresponding to successive sets, the present results also suggest that the general framework explained herein may indeed serve as a sensible 'middle-ground' approach to hydrologic modeling, especially in tandem with a chaotic dynamic framework [see also Sivakumar (2004, 2009) for some details], one not requiring stochastic partial differential equations but rather geometric trends.

Further, it is worth noting that the complex patterns illustrated in this work can hardly be obtained by resorting to classical stochastic approaches. Moreover, the FMFP can generate an incredible variety of patterns that may cover, in principle, the whole variety of possible natural observations, thus suggesting a completeness of the framework in an algebraic sense.

5 Summary and conclusions

We have reviewed a general deterministic approach (and its extensions) aimed at the encoding of rainfall time series and rainfall spatial distributions. Our method can generate a vast array of patterns, which share the overall appearance and statistics of measured data sets. These patterns are obtained as projections of fractal functions illuminated by multifractal sets and are generated deterministically by iterating simple affine maps. The encoding of these patterns using this procedure is both efficient and parsimonious, as the number of coefficients that define the affine mappings is relatively small. This is a particularly desirable feature of a mathematical model as it implies, besides the simplicity of the record's description, also a significant compression ratio of the data set, which can exceed ratios of 100:1. While our method appears particularly suited to encode rainfall sets in one and two-dimensions, the procedure can be easily extended to simulate other climatic variables and other complex sets arising in hydrologic (geophysical) applications.

Further advancing the practical applicability of the fractal-multifractal procedure requires an efficient solution of the inverse problem, i.e., the problem of finding a set of suitable affine map's parameters yielding a projection that matches a given measured data series. Given the enormous variety of patterns that can be generated within the FMFP framework, finding the appropriate set of parameters is tantamount to finding a needle in a haystack. Efficient solution of such an inverse problem involves not only devising clever ways to minimize the number of generated patterns during a given search but also making critical advances in computational power. Definite advances have recently been made in this direction, and progress on this subject will be reported in a future manuscript, including the encoding of the time evolution of measured spatial patterns.

Acknowledgment This work was supported, in part, by the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

Barnsley MF (1988) Fractals everywhere. Academic Press, New York

Cortis A, Puente CE, Sivakumar B (2009) Nonlinear extensions of a fractal-multifractal approach for environmental modeling. Stoch Environ Res Risk Assess. doi:10.1007/s00477-008-0272-0

Gaume E, Mouhous N, Andrieu H (2007) Rainfall stochastic disaggregation models: calibration and validation of a multiplicative cascade model. Adv Water Resour 30:1301–1319

Georgakakos KP, Bras RL (1984) A hydrologically useful station precipitation model. 1. Formulation. Water Resour Res 20(11): 1585–1596

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

Mandelbrot BB (1989) Multifractal measures especially for the geophysicist. In: Scholz CH, Mandelbrot BB (eds) Fractals in geophysics. Birkhauser Verlag, Basel, pp 1–42

Meneveau C, Sreenivasan KR (1987) A simple multifractal cascade model for fully developed turbulence. Phys Rev Lett 59:1424–1427

Morrissey ML (2009) Superposition of the Neyman-Scott rectangular pulses model and the poisson white noise model for the representation of tropical rain rates. J Hydrometeorol 10:395–412

Puente CE (1992) Multinomial multifractals, fractal interpolators, and the Gaussian distribution. Phys Lett A 161:441–447

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, Klebanoff A (1994) Gaussians everywhere. Fractals 2(1):65–79

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):16–22

Puente CE, Sivakumar B (2003) A deterministic width function model. Nonlinear Processes Geophys 10:525–529

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

Puente CE, Robayo O, Díaz MC, Sivakumar B (2001a) A fractal-multifractal approach to groundwater contamination. 1. Modeling conservative tracers at the Borden site. Stoch Environ Res Risk Assess 15(5):357–371

Puente CE, Robayo O, Sivakumar B (2001b) A fractal-multifractal approach to groundwater contamination. 2. Predicting conservative tracers at the Borden site. Stoch Environ Res Risk Assess 15(5):372–383

Puente CE, Obregón N, Sivakumar B (2002) Chaos and stochasticity in deterministically generated multifractal measures. Fractals 10 (1):91–102

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

Sivakumar B (2004) Chaos theory in geophysics: past, present and future. Chaos Solitons Fractals 19(2):441–462

Sivakumar B (2009) Nonlinear dynamics and chaos in hydrologic systems: latest developments and a look forward. Stoch Environ Res Risk Assess. doi:10.1007/s00477-008-0265-z

Sivakumar B, Sorooshian S, Gupta HV, Gao X (2001) A chaotic approach to rainfall disaggregation. Water Resour Res 37(1):61–72

Sivakumar B, Jayawardena AW, Li WK (2007) Hydrologic complexity and classification: a simple data reconstruction approach. Hydrol Processes 21(20):2713–2728