

Research papers

A comparison of fractal-multifractal techniques for encoding streamflow records

Mahesh L. Maskey^a, Carlos E. Puente^{a,*}, Bellie Sivakumar^{a,b}^a Department of Land, Air and Water Resources, University of California, Davis, One Shields Ave., Davis, CA 95616, USA^b School of Civil and Environmental Engineering, The University of New South Wales, Sydney, NSW 2052, Australia

ARTICLE INFO

Article history:

Received 20 February 2016

Received in revised form 1 September 2016

Accepted 9 September 2016

Available online 10 September 2016

This manuscript was handled by K. Georgakakos, Editor-in-Chief, with the assistance of Alessio Domeneghetti, Associate Editor

Keywords:

Fractal-multifractal

Streamflow

Dynamics

Encoding

Climate change

ABSTRACT

This study introduces the deterministic fractal-multifractal (FM) approach and some of its variants for encoding, for the first time, streamflow data. For this purpose, daily streamflow records gathered over 65 years at the Sacramento River in California, USA are studied. The specific FM representations considered here include three main categories and six cases as follows: (i) the original FM approach, yielding a data set as a fractal transformation of a multifractal measure, and obtained by iterating two and three affine maps; (ii) an extension, defining records as a transformation of an input measure by a general attractor (not a function), also obtained by iterating two and three affine maps; and (iii) two procedures that loop two fractal functions or two general attractors, providing a data set as a weighted projection via the iteration of a total of four affine maps each. Highlighting the results obtained for six different years of records having distinct geometries, it is shown that all six FM variants yield faithful representations, as evidenced by the maximum errors in accumulated records never exceeding a mere 2%. Of all approaches, however, it is shown that the one based on three affine maps, leading to substantial compression ratios of 28:1, is generally the best, for such end up following the geometry of all sets rather closely. As illustrated for the whole 65-year period, the time evolution of the FM parameters may be useful to track and predict changes in streamflow over time.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

The quest for understanding complex natural records, such as rainfall and streamflow, has led to the development of deterministic and stochastic models since the 1960s. Various ideas have been tried in order to model natural phenomena. Such include usage of: (a) physically-based models, (b) classical stochastic models, and (c) fractal-stochastic models. While the first are based on well-accepted physical principles such as conservation laws, the others seek to preserve statistical/fractal properties in the records such as autocorrelation and co-dimension functions. The past decades have witnessed numerous applications of these ideas and models in hydrology; for details, see Eagleson (1970, 1978), Freeze and Cherry (1979), Kavvas and Delleur (1981), Mandelbrot (1982),

Rodríguez-Iturbe (1986), Sivakumar et al. (2001), Yang et al. (2007), Özger et al. (2010), Chen and Wu (2012), Tongal et al. (2013), and Poornima and Jothiprakash (2015), among others.

Although these approaches yield reasonable representations of the overall structure of the data, due to their underlying assumptions, they do not always capture the finer details and textures present in natural records. For instance, in the physical approaches, the lack of spatial records leads to discretization and computational issues, and in the classical stochastic methods the specific realizations are hard to condition to mimic the specific geometry of a given set. Even though the fractal-stochastic methods, with the addition of fractal and multifractal notions, have resulted in more realistic simulations and models (e.g., Lovejoy and Schertzer, 2013), the specific geometries of sets still require of a new “language” for describing their complexity.

These intrinsic limitations led Puente to think that perhaps the erratic, intermittent, complex, and altogether “seemingly random” nature of the records could be modeled as a projection of a fractal function. Inspired by the simplicity and depth of the work by Lorenz (1963), the discovery of a simple cascade to model fully-developed turbulence by Meneveau and Sreenivasan (1987), and the seminal work of Barnsley (1988), Puente (1996) developed

Abbreviations: RMSEAS, root mean square error in accumulated streamflow; maximum error in accumulated streamflow, MAXEAS; Nash-Sutcliffe efficiency of streamflow data, NSD; Nash-Sutcliffe efficiency of autocorrelation function, NSA; Nash-Sutcliffe efficiency of histogram, NSH; Nash-Sutcliffe efficiency of entropy function, NSE; percent of data points outside $\pm 10\%$ bands of scatter plots, PO₁₀; percent of data points outside $\pm 20\%$ bands of scatter plots, PO₂₀.

* Corresponding author.

E-mail address: cepuente@ucdavis.edu (C.E. Puente).

the fractal-multifractal method by means of which a given (normalized) natural set is thought of as the derived measure found by transforming a multifractal measure via a fractal interpolating function. Such an idea, known as the fractal-multifractal (FM) method, has turned out to be a fertile approach to parsimoniously and deterministically encode complex geophysical patterns in order to explore the complexity of nature. Indeed, the FM approach and its variants have been shown to work reasonably well in defining a vast class of patterns (Puente, 2004), over one and higher dimensions, preserving not only the key statistical indicators (e.g., moments, autocorrelation, power spectrum, multifractal spectrum) but also the overall geometric features and textures present in the data sets. While stochastic approaches are typically not capable of preserving finer details beyond the main statistical features, the FM methodology is found to yield indeed fruitful results in encoding a host of geophysical records, such as precipitation sets (e.g., Puente and Obregón, 1996; Cortis et al., 2009, 2013; Huang et al., 2012, 2013; Maskey et al., 2015), fluid turbulence (Puente and Obregón, 1999), river width function (Puente and Sivakumar, 2003), and groundwater contamination patterns (Puente et al., 2001a, 2001b).

The present study demonstrates, for the first time, the capability of six distinct variants of the FM approach to encode streamflow data. For this purpose, daily streamflow records gathered over a period of 65 years in the Sacramento River are studied and an attempt to understand the dynamics of streamflow based on FM geometric parameters is ascertained. For the purpose of illustration, six different years having distinct geometries are discussed in detail, comparing the performance of the different FM approaches in terms of a visual inspection as well as a number of statistical qualifiers. Having identified the best FM model, a novel visualization of the dynamics of the process is then advanced.

The organization of the rest of this manuscript is as follows. First, the original FM approach and the variants used in this study are reviewed. Second, the optimization procedure and strategy used to encode streamflow records are advanced. Third, the results from the different approaches for each of the six distinct years of streamflow records considered for illustration are reported. Fourth, the best FM encodings over 65 years of records and their implied dynamics are reported. Finally, a summary of the results and conclusions of the study are given.

2. The Fractal Multifractal (FM) approach

This section summarizes the original FM approach (Puente, 1996) and introduces the other variants of the procedure used in this study for encoding streamflow records.

2.1. The original method

Utilizing fractal interpolating functions, pioneered by Barnsley (1988), Puente (1996) introduced the FM procedure. The notions rely on finding the projection – over the range of a fractal function – of the unique multifractal measure induced by iterating simple maps that determine the graph of such a function. Specifically, a fractal function $f: x \rightarrow y$, passing through $N + 1$ non-aligned points $\{(x_n, y_n) \mid x_0 < \dots < x_N, n = 0, 1, \dots, N\}$ and resulting in a graph $G = \{(x, f(x)) \mid x \in [x_0, x_N]\}$, is defined from successive iterations of N simple affine maps:

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

where d_n , the vertical scaling parameters, satisfy $|d_n| < 1$ and the other parameters a_n , c_n , e_n , and f_n are evaluated, in terms of the interpolating points and the scaling parameters, from contracting initial conditions:

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

which guarantee the existence of a stable attractor G (Barnsley, 1988).

As the process is carried out, iterating the affine maps according to arbitrary and fixed “coin tosses,” it: (a) produces, via a point-wise sampling of the attractor G known as the “chaos game,” a convoluted “wire” function f , whose graph has a fractal dimension $D \in [1, 2]$; and (b) induces a unique invariant measure over G , whose projected histograms over the x - and y -coordinates are, respectively, deterministic multinomial multifractal measures (Mandelbrot, 1989) and derived measures (via function f) that resemble the intricate irregular shapes akin to natural patterns. At the end, the original FM representation transforms a multifractal measure dx into a derived measure dy via a fractal interpolating function f .

To further illustrate the ideas, consider the two affine maps:

$$w_1 \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0.66 & 0 \\ -0.38 & -0.47 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (3)$$

$$w_2 \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0.34 & 0 \\ 1.51 & 0.33 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0.66 \\ -0.85 \end{pmatrix}, \quad (4)$$

associated with the interpolating points $\{(0, 0), (0.66, -0.85), (1, 1)\}$, and iterated 2^{14} times according to a 56–44% proportion. As seen in Fig. 1, with interpolating points shown as bullets, these iterations induce a fractal function f , having a fractal dimension $D = 1$ (top left) and two unique invariant measures over the two coordinates: dx (below) over the x -axis and dy (to the right) over the y -axis. While dx comes from a multiplicative cascade and yields a multifractal measure exhibiting a rather repetitive appearance (Mandelbrot, 1989), the derived measure dy , found transforming (over y) the input dx via the function f , gives rise, after local smoothing (shown on the right as dy_s), to a set that resembles geometrically (after a 90 degree flip and interpreting y as time and dy as discharge), a daily streamflow time series, and other sets as encountered in nature (Obregón et al., 2002a, 2002b; Puente, 2004).

Requiring little effort in the implementation, which simply provides a collection of (x, y) points via the arbitrary iterations guided by a “coin,” i.e., the chaos game, – as included in the Matlab code in the Online Appendix <http://puente.lawr.ucdavis.edu/omake/jhydro2016/jhydro2016.html>, the measures dx and dy are simply the computed histograms over x and y that may be obtained at arbitrary resolutions. As shall be explained later on, the FM parameters in the maps in Eqs. (3) and (4) are obtained via an optimization exercise such that the shown object dy_s resembles the geometry of normalized streamflow (minus baseflow) as gathered at the Sacramento River in 1968, hence using a total of 365 or 366 bins for computing the output measure dy and a 5 day smoothing to define dy_s . The reader is encouraged to consider the Online Appendix, as it also includes an interactive demo, based on Fig. 1, of the patterns the FM method may produce.

To summarize, the original FM method requires, for the encoding of a given natural set, the solution of an inverse optimization problem that relies on the following parameters: (a) the fractal interpolation points, (b) the vertical scalings d_n , (c) the frequencies used in the iteration of the affine maps, and (d) a smoothing parameter as in Fig. 1, if needed. Ultimately, the FM approach requires six and ten geometric parameters when using two and three affine maps, respectively.

2.2. A generalization to leafy attractors

The original FM approach can be generalized replacing the contractile initial conditions in Eq. (2) by the more general conditions (Huang et al., 2013)

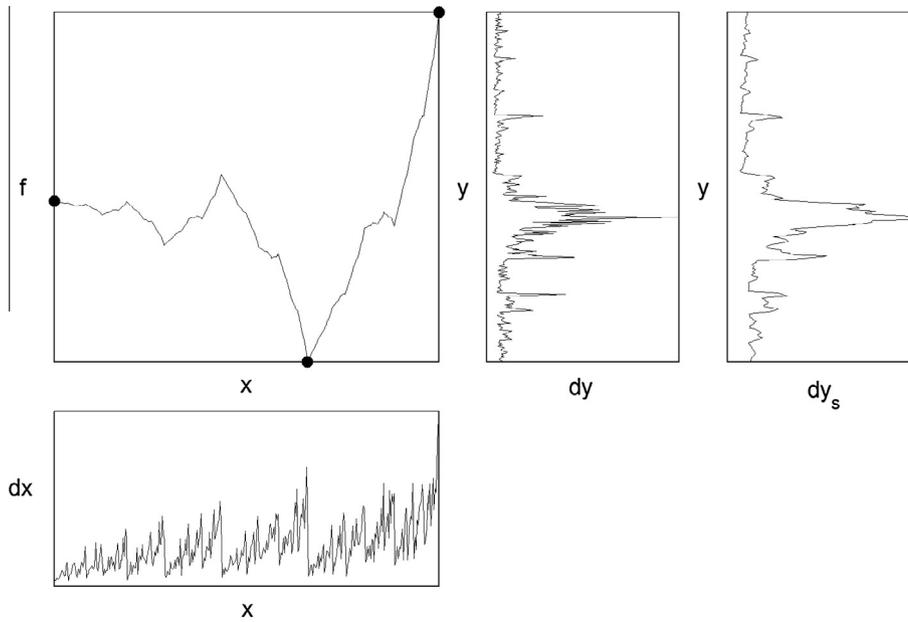


Fig. 1. The original FM approach: from a multifractal dx to a projection dy , via a fractal interpolating function $f: x \rightarrow y$. dy_s is a smoothed version of dy .

$$w_n \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} x_{2n} \\ y_{2n} \end{pmatrix}, \quad w_n \begin{pmatrix} x_{2n-1} \\ y_{2n-1} \end{pmatrix} = \begin{pmatrix} x_{2n+1} \\ y_{2n+1} \end{pmatrix}, \quad (5)$$

where each affine map, as in Eq. (1), is associated with possibly non-touching end-points $\{(x_0, y_0), (x_{2n}, y_{2n})\}$ and $\{(x_{2n+1}, y_{2n+1}), (x_{2N-1}, y_{2N-1})\}$, satisfying $x_0 \leq x_{2n} < x_{2n+1} \leq x_{2N-1}$, $n = 1, \dots, N$.

This new setting typically generates a cloud of points rather than a single-valued function, a more sophisticated attractor – akin of a fractal “leaf”. This happens when the end-points overlap over x or when successive end-points in y are different. Fig. 2 exemplifies such a leafy attractor obtained by iterating the two affine maps:

$$w_1 \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0.70 & 0 \\ 5.07 & -0.61 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (6)$$

$$w_2 \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0.31 & 0 \\ 0.10 & -0.57 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0.68 \\ 1.47 \end{pmatrix}, \quad (7)$$

according to a 56–44% proportion.

As may be inferred, the maps have as end-points $\{(0, 0), (0.70, 4.45)\}$ and $\{(0.69, 1.47), (1, 1)\}$ (shown in circles in Fig. 2) and have as scaling parameters $d_1 = -0.61$ and $d_2 = -0.57$. Clearly, the two maps w_1 and w_2 contract the domain $[0, 1]$ into the sub-intervals $[0, 0.70]$ and $[0.69, 1]$, hence having an overlap of 0.01 between them, which, compounded by the fact that the end y -values differ (i.e., $4.45 \neq 1.47$), yields a leafy attractor having a repetitive structure.

As illustrated in Fig. 2, the projection over x , of the unique measure induced by playing the chaos game with these maps over the ultimate leaf, is, once again, a spiky measure dx , which, when transformed by the cloud of points, yields the derived projection over y called dy . The measure dy , which is simply the histogram over y of all points gathered via the chaos game (for a given number of equal bins spanning the minimum and the maximum of the

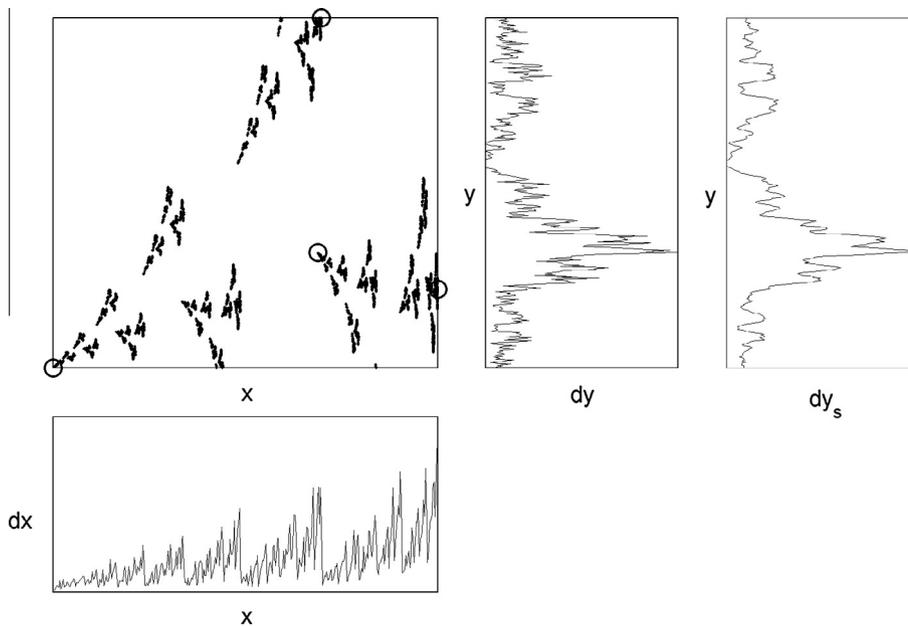


Fig. 2. The FM approach from maps with an overlap: from an input measure dx , to an output projection dy , via a “leafy” attractor. dy_s is a smoothed version of dy .

range of the attractor), results again in an irregular pattern that, after due smoothing (see dy_s on the right hand side), happens to resemble the geometry of geophysical records, such as a stream-flow set in the Sacramento River in water year 2008.

In the same vein as with the original FM approach, usage of these ideas on a practical setting requires the solution of an inverse problem for the generalized FM parameters: (a) the end-points of maps that define contracting placings of the generalized attractor, (b) the vertical scalings d_n , (c) the frequencies used in chaos game iterations, and (d) a smoothing parameter, if any. This FM variant ultimately requires eight and 14 parameters when iterating two and three affine maps, respectively. For more information on this extension, the reader may access the Online Appendix <http://puente.lawr.ucdavis.edu/omake/jhydro2016/jhydro2016.html>, which also includes a Matlab code for such a generalization and an interactive demo for the notions based on the pattern in Fig. 2.

2.3. Extensions closing a loop

The notions just explained may also be extended by using either: (a) two fractal interpolating functions having equal interpolating points over x but distinct intermediate vertical placements and vertical scalings, or (b) two leafy attractors having the same original end-points in x but with possibly different intermediate placements in y . By maintaining the same iteration frequencies in the respective chaos games, the weighing of the individual projections over y , via an additional parameter, also yields interesting projections dy (Huang et al., 2012). These ideas that essentially close a loop of fractal functions or leafy attractors are illustrated next.

Fig. 3 shows the construction of a loop of fractal functions f_1 (in black) and f_2 (in gray), which pass by the following sets of three interpolating points $\{(0, 0), (0.69, -4.95), (1, 1)\}$ (black dots) and $\{(0, 0), (0.69, -3.33), (1, 1)\}$, (gray dots) and associated, respectively, with maps w_1 and w_2 and \hat{w}_1 and \hat{w}_2 :

$$w_1 \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0.69 & 0 \\ -4.08 & -0.87 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \tag{8}$$

$$w_2 \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0.31 & 0 \\ 5.65 & 0.30 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0.69 \\ -4.95 \end{pmatrix}, \tag{9}$$

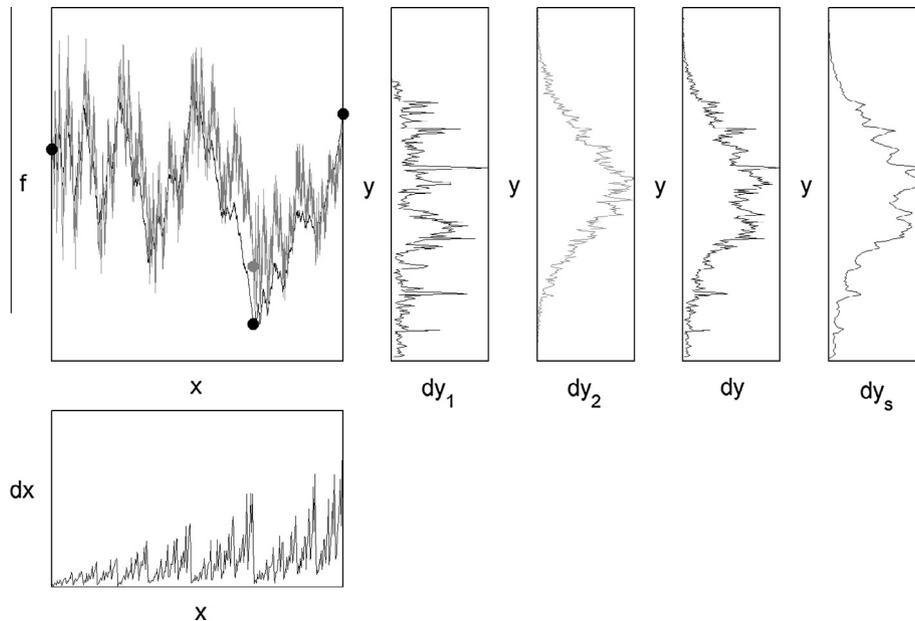


Fig. 3. Closing the loop using two FM wires: from a multinomial multifractal measure dx to derived measures dy_1 and dy_2 , via fractal wires f_1 (in black) and f_2 (in gray). dy is found weighing dy_1 and dy_2 and dy_s is a smoothed version of dy .

and

$$\hat{w}_1 \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0.69 & 0 \\ -2.33 & -0.99 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \tag{10}$$

$$\hat{w}_2 \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0.31 & 0 \\ 3.76 & 0.56 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0.69 \\ -3.33 \end{pmatrix}, \tag{11}$$

iterated, by couples, according to a 55–45% “coin.”

As shown, the derived measures based on each fractal function, dy_1 and dy_2 , when weighed according to a 33–67% split, yield the interesting set dy , which when smoothed gives another set dy_s that resembles geometrically geophysical (e.g., streamflow) data, in particular flows at the Sacramento River for water year 1958. As expected, this generalization results in additional parameters, which, for two maps, as will be used later on, require a total of ten parameters: (a) the intermediate interpolating points in x , (b) the intermediate interpolating points in y , (c) the vertical scalings, (d) the frequencies used in chaos game iterations, (e) the proportion by which the two projections are combined, and (f) a smoothing parameter, if necessary.

Fig. 4 presents the counterpart of Fig. 3, but based on looping two general attractors, each generated by two affine maps, as indicated by the black and gray circles. As seen, the combination of the resulting projections turn out to be also interesting, and such a representation relying on 13 FM parameters, that is, is similar in size as using a leafy attractor based on three maps, also generates, after due smoothing, sensible approximations of the geometry of streamflow data, this time a close rendering of river flows at the Sacramento River for water year 2008.

3. Methodology

Once a parameter set is given for any variant of the fractal-multifractal method, it is easy to calculate a given derived measure dy at any desired resolution, i.e., 365 or 366 bins as is appropriate for a given year. As explained in the aforementioned Online Appendix, the chaos game is certainly simple and its implementation efficient. However, finding the FM parameters for a given data set is not easy. This is because there are no analytical formulas for a derived measure dy in terms of FM parameters. As such, this

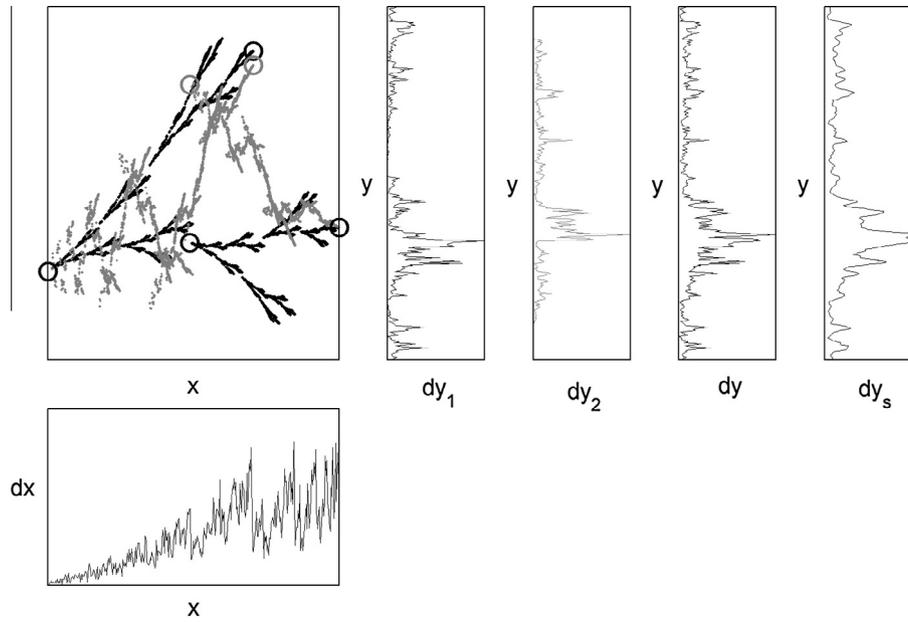


Fig. 4. Closing the loop using two FM leaves: from an input measure dx to output measures dy_1 and dy_2 , via fractal leaves (in black and in gray). dy is found weighing dy_1 and dy_2 and dy_s is a smoothed version of dy .

requires a search method over a higher-dimensional space, up to 14 coordinates, in order to find a numerical solution to an inverse problem. It turns out that searches are sensitive not only to the definition of a suitable objective function but also to the usage of a particular optimization algorithm. As a consequence, and based on our past experience, we herein employ a generalized particle swarm optimization algorithm (GPSO) (Fernández Martínez et al., 2010; Huang et al., 2013) to minimize the L^2 -norm of the differences between the accumulated streamflow records (given, and normalized to add up to one) and the accumulated FM sets:

$$RMSEAS = \sqrt{\frac{1}{N} \sum_{i=0}^N (r_i - \hat{r}_i)^2}, \quad (12)$$

where N is the number of data points (365 or 366 days depending upon the year considered), and r_i and \hat{r}_i are, respectively, the i th accumulated values of the observed records and of the FM set. In addition, and in order to ensure that FM “solutions” share similar geometrical features with the target set, various penalties are also imposed on the optimization exercise, such that the maximum cumulative deviations (MAXEAS) between records and FM sets would not exceed 10% and such that the length (over two dimensions) of an FM set would not exceed 10% of the length of the original records.

In the remainder of the article, the goodness of the alternative FM approaches is quantified in terms of various statistical attributes beyond those used in the optimization process. These attributes help further compare differences between the observed records (i.e., the data) and the FM-fits and include: (a) the aforementioned maximum deviations on accumulated sets; (b) scatter plots between data and FM-fitted sets; and (c) the Nash-Sutcliffe efficiency indices associated with the data records (NSD), its histogram (NSH), its autocorrelation function (NSA), and its Rényi entropy function (NSE), as given in their most general form by (Nash and Sutcliffe, 1970),

$$NSI = 1 - \frac{\sum_{i=1}^S (s_i - \hat{s}_i)^2}{\sum_{i=1}^S (s_i - \bar{s})^2} \quad (13)$$

where s_i is the i th statistical value associated with the observed set, \bar{s} is the mean of such a statistic over a total of S values, and \hat{s}_i is the i th

statistical value obtained from the FM fit. For the records, histogram, autocorrelation function, and entropy function, the total number of statistical attributes S is, in order, 365 or 366, 10 (number of bins), 91 (number of lags equal to a quarter of the size of the records), and 101 (for equally-sized exponents varying from -5 to 5).

4. Comparison of distinct FM streamflow encodings

This section includes a comparison of the various FM approaches introduced earlier in Section 2, when using the search methodology and statistical analysis presented in Section 3. For this purpose, six interesting and distinct sets of daily streamflow records from the Sacramento River near Freeport (out of the 65 years of data considered in this study), California (USGS station 11447650) and each spanning a water year (from October 1st to September 30th) are analyzed in detail: 1958, 1968, 1975, 1984, 1998 and 2008, with the numbers indicating the ending year. As the raw data are not fully amenable for encoding via a FM variant due to elevated values, a constant base flow (groundwater component) is subtracted from each of them (possibly not equal) and then the data sets are normalized so that the accumulated volume becomes unity.

For each of the six sets to be illustrated herein, encodings via six specific FM representations are found. Such include derived measures obtained from the aforementioned variants of the fractal-multifractal methodology, as follows: (a) two obtained from “wires” calculated via two and three affine maps (Section 2.1), (b) two found from “leaves” based on two and three affine maps (Section 2.2), and (c) two defined “closing a loop” using either two wires or two leaves and defined via two affine maps each (Section 2.3). Table 1 lists some basic details of these six FM representations. The representations are denoted as A to F, for the sake of brevity and convenience in the presentation of the results. Table 1 shows that while the number of parameters of these six representations varies from 5 to 13, the corresponding compression ratios span from a “high” 73:1 to a “low” 28:1. Notice that the local smoothing used on a given derived set dy , in order to define an output measure dy_s to approximate a given streamflow set, is not included in the table, for such is kept constant in all sets at a value of 5 days.

Table 1
Information about the different FM representations.

Approach	Fit	Number of maps	Number of parameters	Compression ratio
Fractal wire	A	2	5	73:1
	B	3	9	41:1
Fractal leaf	C	2	7	52:1
	D	3	13	28:1
Loop of wire	E	4	9	41:1
Loop of leaf	F	4	12	30:1

The results that follow correspond to the “best” FM encodings. It is important to note that these “best” encodings are not identified from the best attribute optimized plus constraints, i.e., the root mean square error of accumulated streamflow (RMSEAS) subject to penalties (Section 3). Rather, they are identified from the least maximum error in accumulated streamflow (MAXEAS), as found over all the swarms calculated and guided by RMSEAS: i.e., 300 randomly defined swarms, each iterated for a total of 40 cycles. Results from the best RMSEAS, which after a few swarm iterations always meet the proposed constraints, are qualitatively similar, even if they typically correspond to distinct FM parameter values. The results herein, inspired by the usage of MAXEAS in

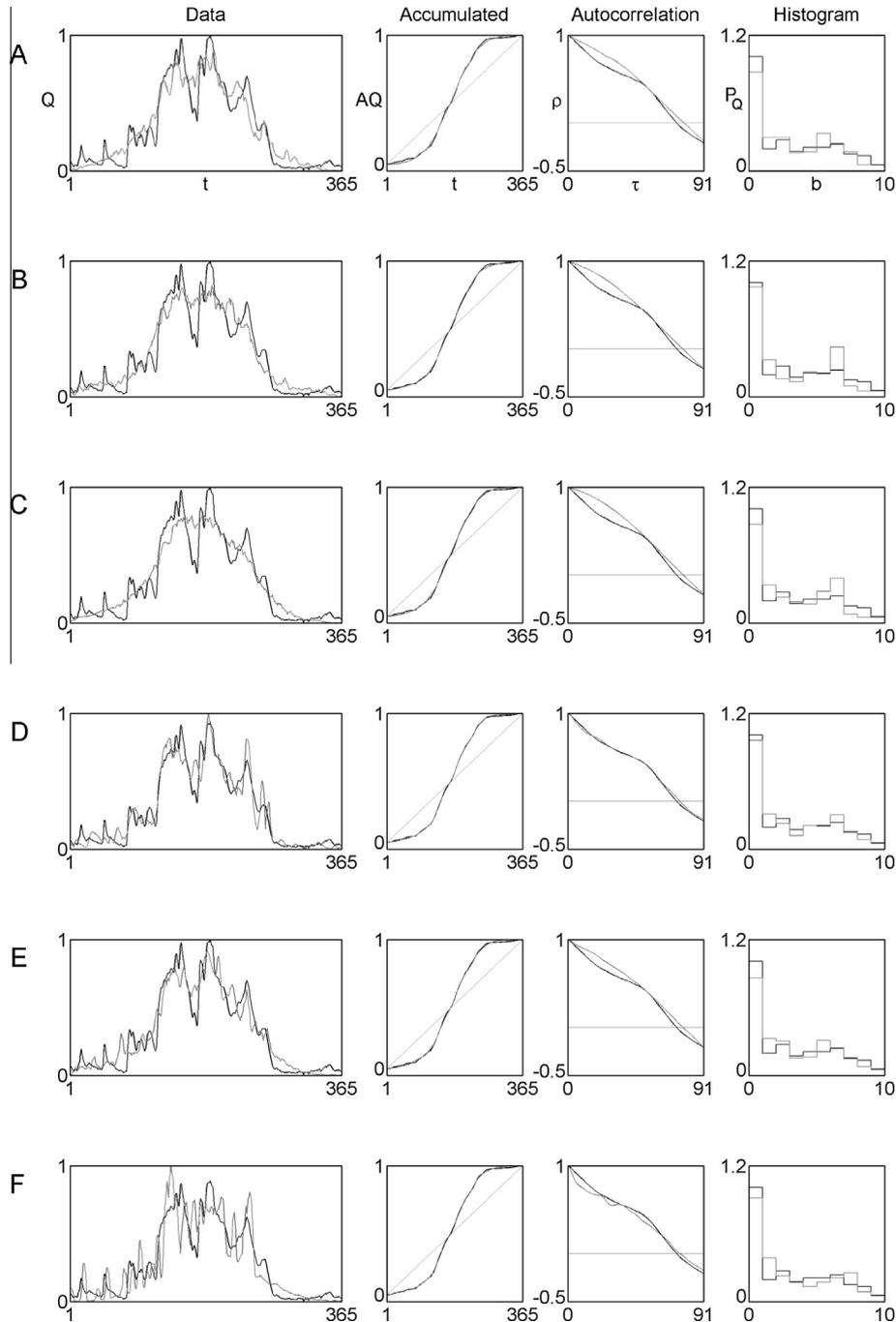


Fig. 5. Daily streamflow records at the Sacramento River for water year 1958 (black) and six FM representations A to F (gray), followed by their accumulated sets, autocorrelation functions, and histograms. A and B are defined via “wires” corresponding to two and three maps. C and D are obtained from “leaves” based on two and three maps. E closes a loop of wires each using two maps and F combines two leaves found via two maps each.

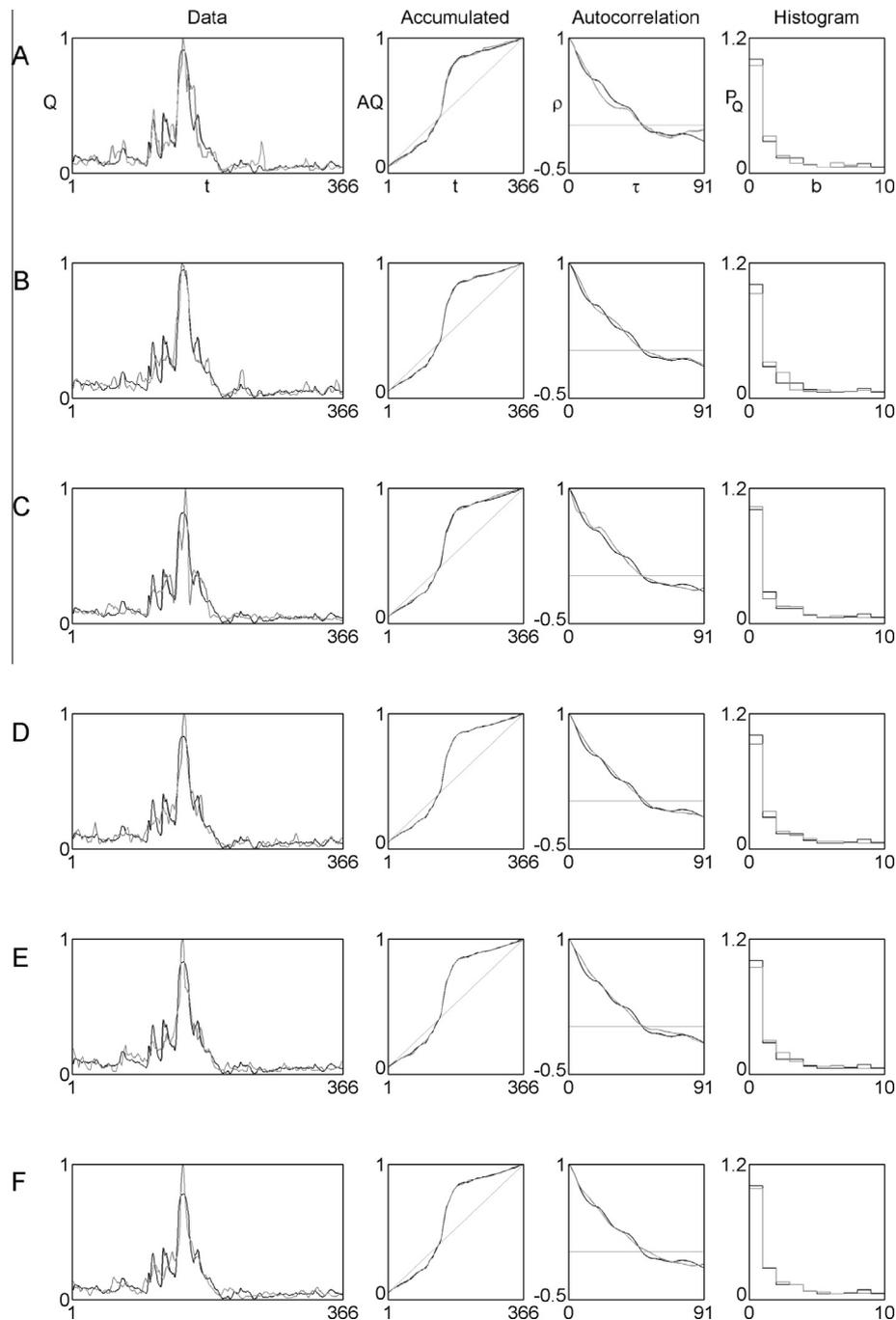


Fig. 6. Daily streamflow records at the Sacramento River for water year 1968 (black) and six FM representations A to F (gray), followed by their accumulated sets, autocorrelation functions, and histograms. A and B are defined via “wires” corresponding to two and three maps. C and D are obtained from “leaves” based on two and three maps. E closes a loop of wires each using two maps and F combines two leaves found via two maps each.

Kolmogorov-Smirnoff tests for assessing differences between two distributions, turn out to demonstrate the goodness of some FM representations over the others.

Figs. 5–10 illustrate the results obtained for the six water years considered here, once again: 1958, 1968, 1975, 1984, 1998, and 2008. In these figures, the observed data are plotted in black and the best FM fits in gray. Each of these six figures contains six rows, one for each of the FM procedures (A to F), and include the following information (in columns), in order: (a) the observed streamflow data and the best FM fit as a function of time (i.e., Q vs. t); (b) the subsequent accumulated sets themselves (i.e., AQ vs. t); (c) the autocorrelation functions for the data and the best FM fit (i.e., ρ

vs. τ); and, for variety, (d) either the corresponding histograms per bin (i.e., P_Q vs. b) (Figs. 5 and 6), the scatter plot between the data and the FM fit (Figs. 7 and 8), or the implied Rényi entropy functions (i.e., H vs. q) (Figs. 9 and 10). Tables 2–7 present a summary of a host of statistical qualifiers corresponding to each of the figures; see below for details. The relevant FM parameters associated with all the figures are included in Appendix A.

Fig. 5 depicts the six FM representations for the streamflow data for water year 1958. As seen on the first column, such a year contains three major peaks that build up over periods of about a month, with most of the mass distributed within the center of this time span. As observed, the distinct FM representations are capable

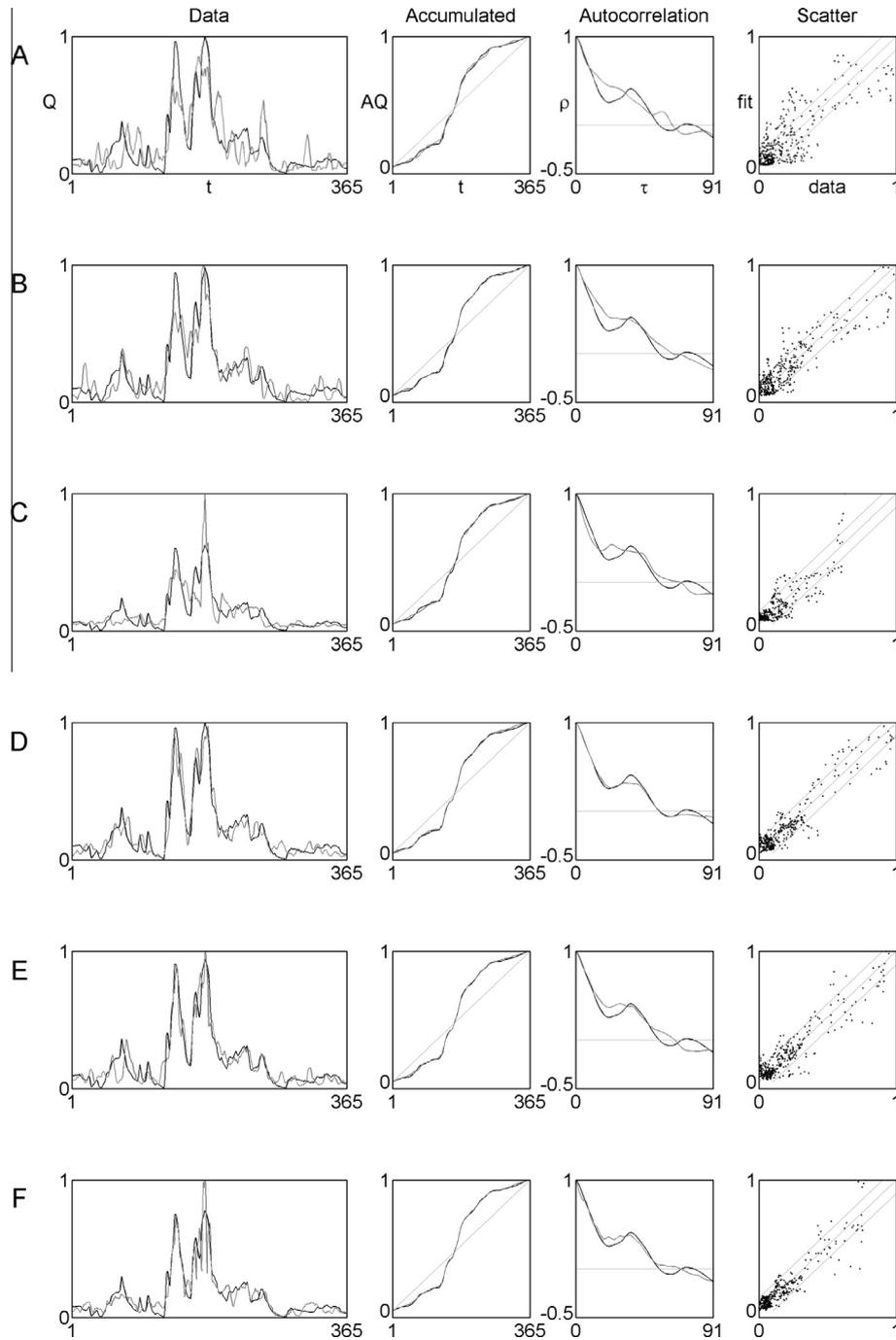


Fig. 7. Daily streamflow records at the Sacramento River for water year 1975 (black) and six FM representations A to F (gray), followed by their accumulated sets, autocorrelation functions, and scatter plot. A and B are defined via “wires” corresponding to two and three maps. C and D are obtained from “leaves” based on two and three maps. E closes a loop of wires each using two maps and F combines two leaves found via two maps each.

of representing those peaks with varying success, including model E that corresponds to the construction shown in Fig. 3. As may be appreciated from the second column of the figure, the modeled accumulated fits for all FM variants (in gray) are rather close to the observed set (in black), indicating that the optimization algorithm does yield suitable solutions. Although, to the naked eye, there appear to be no major differences among the distinct FM sets, the variants plotted for the data itself (first column) show clear differences. While all representations do capture the overall shape of the hydrograph, the ones labeled D and E turn out to be quite remarkable, as they also preserve the overall rising and falling around the peaks. Models A, B, and C are clearly smoother than

the records, while fit F is too erratic. The nature of these encodings may be further appreciated in the last two columns of Fig. 5 that compare the autocorrelation functions and histograms of all FM variants with those of the records. As seen, model D, coming from the projection of a leaf defined via three maps, preserves the autocorrelation function almost precisely and also the histogram rather very well. The other four variants also capture the variability and properties of the streamflow data reasonably well.

Table 2 includes various statistical attributes associated with Fig. 5 that qualify the trends seen in the analysis. First, all FM variants have rather low values for the root mean square in accumulated streamflow (RMSEAS) and maximum error in accumulated

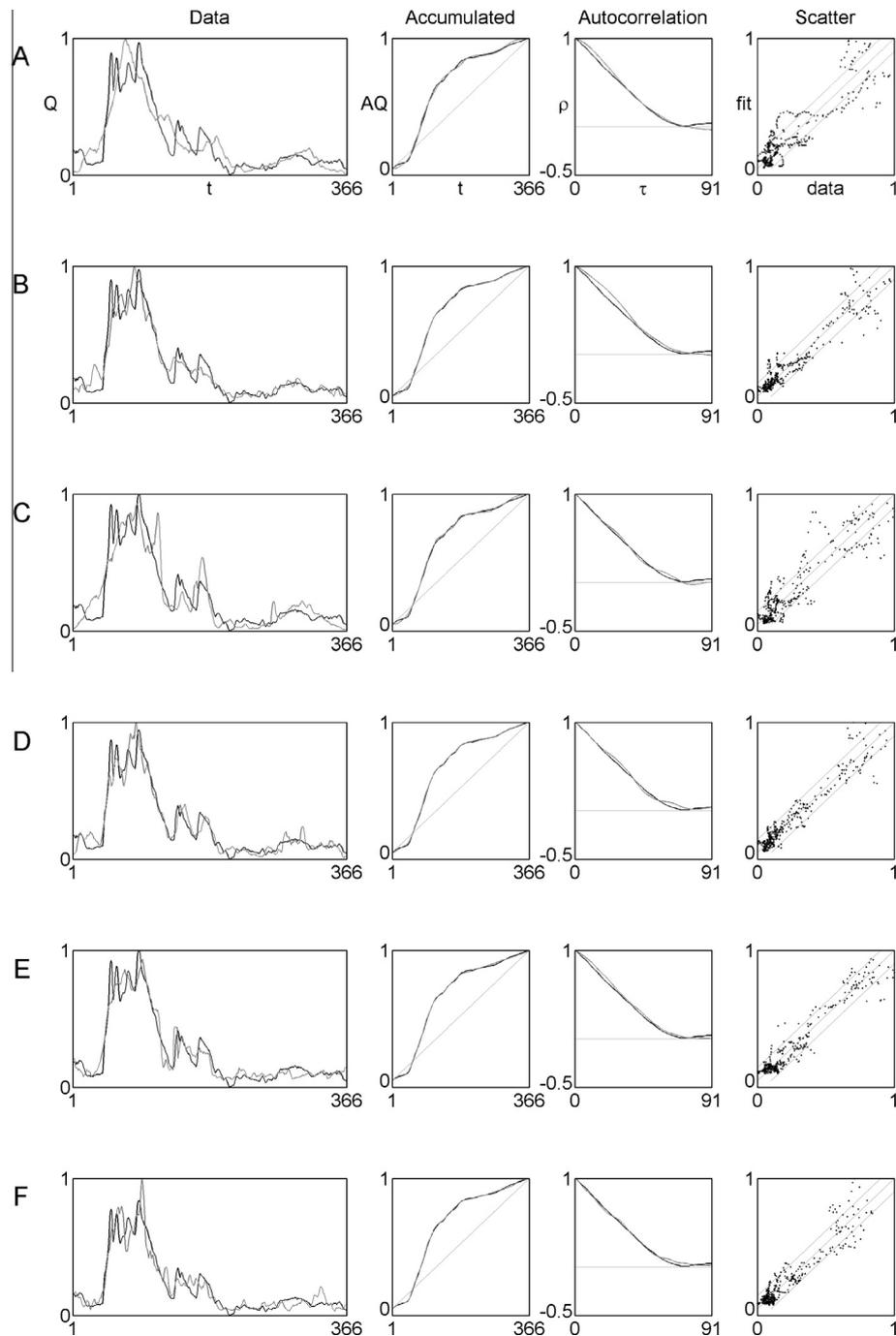


Fig. 8. Daily streamflow records at the Sacramento River for water year 1984 (black) and six FM representations A to F (gray), followed by their accumulated sets, autocorrelation functions, and scatter plot. A and B are defined via “wires” corresponding to two and three maps. C and D are obtained from “leaves” based on two and three maps. E closes a loop of wires each using two maps and F combines two leaves found via two maps each.

streamflow (MAXEAS), with numbers that are always less than 1.0 and 1.9%, respectively. Second, all representations have high Nash-Sutcliffe statistics for all attributes considered: data (NSD), autocorrelation (NSA), histogram (NSH), and entropy (NSE), as obtained values are always higher than, in order, 77.5, 93.3, 87.0, and 98.9%. Notice that while FM fit D is confirmed as the best, with remarkable Nash-Sutcliffe values for the aforementioned attributes of 91.4, 99.7, 95.9, and 99.9%, the other variants, with the exception of model F, also provide rather reasonable results, as these numbers are always larger than 87.6, 93.3, 87.0, and 98.9%. Table 2 also includes information regarding the actual timing of records when compared to FM fits via deviations from a 45-degree line on a

scatter plot having scales from zero to one, i.e., as the sets appear on the first column of Fig. 5. As such, the number of data vs. FM fit points landing outside the 20% ($\pm 10\%$) bands relative to the line of perfection (PO_{10}) and the ones found outside a wider 40% ($\pm 20\%$) bands (PO_{20}) corroborate the goodness of the FM fit labeled as D – having a mere 3.6% of the records outside the wide band – and with all others (except fit F) following at values less than 7.9%.

Fig. 6 illustrates the FM encodings obtained for streamflow from water year 1968, which has a geometry dominated by a single peak. As seen, all FM variants, including the one labeled A which is associated with the construction in Fig. 1, closely follow the accumulated records and all of them yield fitted sets that, unlike

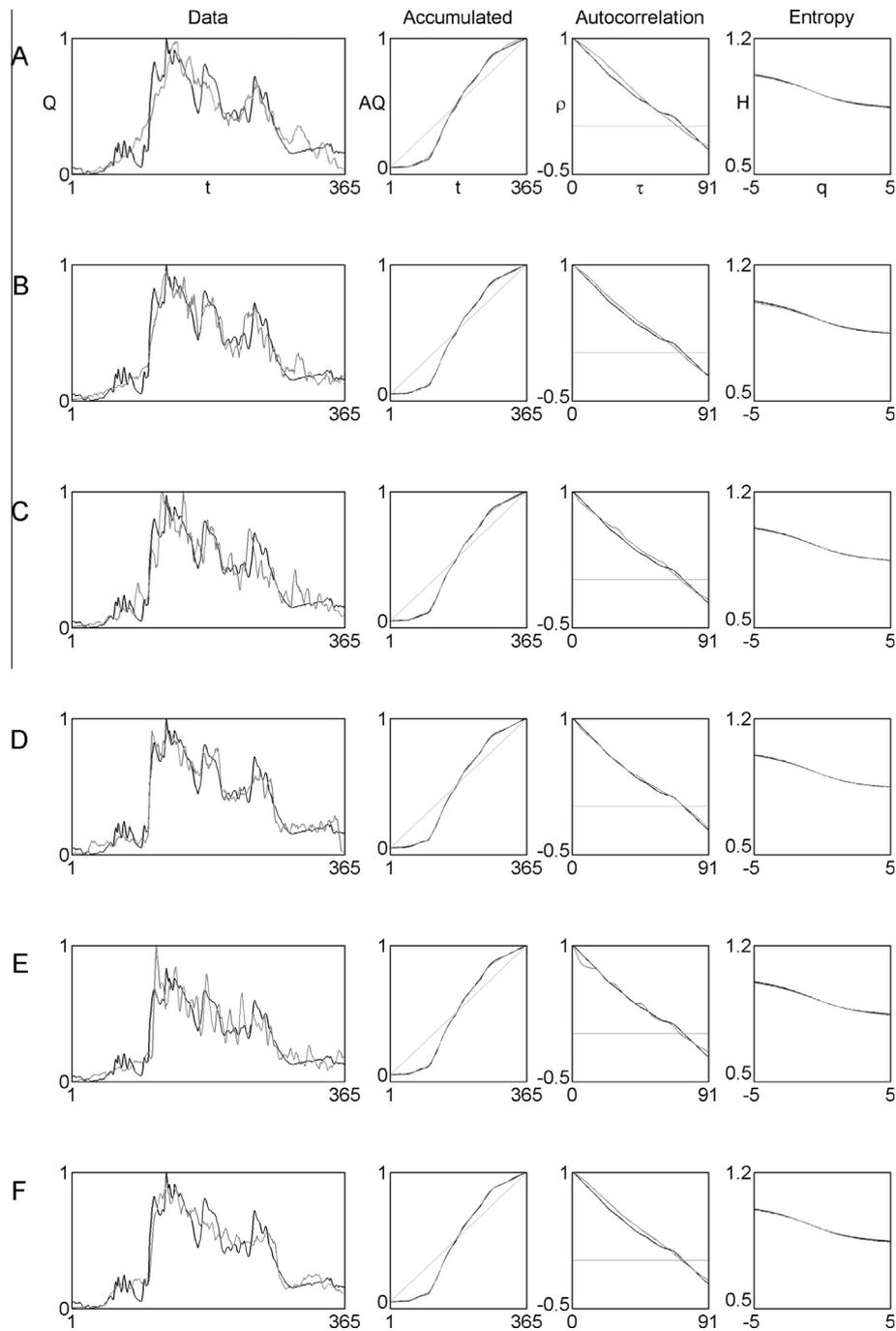


Fig. 9. Daily streamflow records at the Sacramento River for water year 1998 (black) and six FM representations A to F (gray), followed by their accumulated sets, autocorrelation functions, and Rényi entropy functions. A and B are defined via “wires” corresponding to two and three maps. C and D are obtained from “leaves” based on two and three maps. E closes a loop of wires each using two maps and F combines two leaves found via two maps each.

the previously seen results in Fig. 5 for water year 1958, closely approximate the salient features of the data set, including the timing and duration of the single peak. All the FM variants also turn out to give representations that closely follow the autocorrelation and histogram of the records. These general observations may also be ascertained by the statistical information presented in Table 3: (a) all fits have RMSEAS and MAXEAS values that are less than 1 and 2.4%, respectively, (b) all Nash-Sutcliffe efficiency values are rather high and comparable across the FM fits, and (c) the number of points outside of scatter plot bounds are similar and rather small. Although the results indicate that any of the six FM variants used here may be suitable to approximate the streamflow in water

year 1968, it is worthy to highlight that case D, defined as the projection from a leaf generated via three maps, performs the best or close to the best in all the statistical attributes, similar to what was found earlier for water year 1958 (Fig. 5).

Fig. 7 shows the six alternative FM fits found for streamflow from water year 1975, which contains two main peaks that come and go in a sharper way than those already seen for year 1958 (Fig. 5). As the graphs on the second column indicate, the optimization process produces, once again, suitable FM representations, whose RMSEAS and MAXEAS values are all below 1.3 and 2.9%, respectively (see Table 4). When the time series (first column) is examined, however, it is seen that there are some FM representations that are clearly better

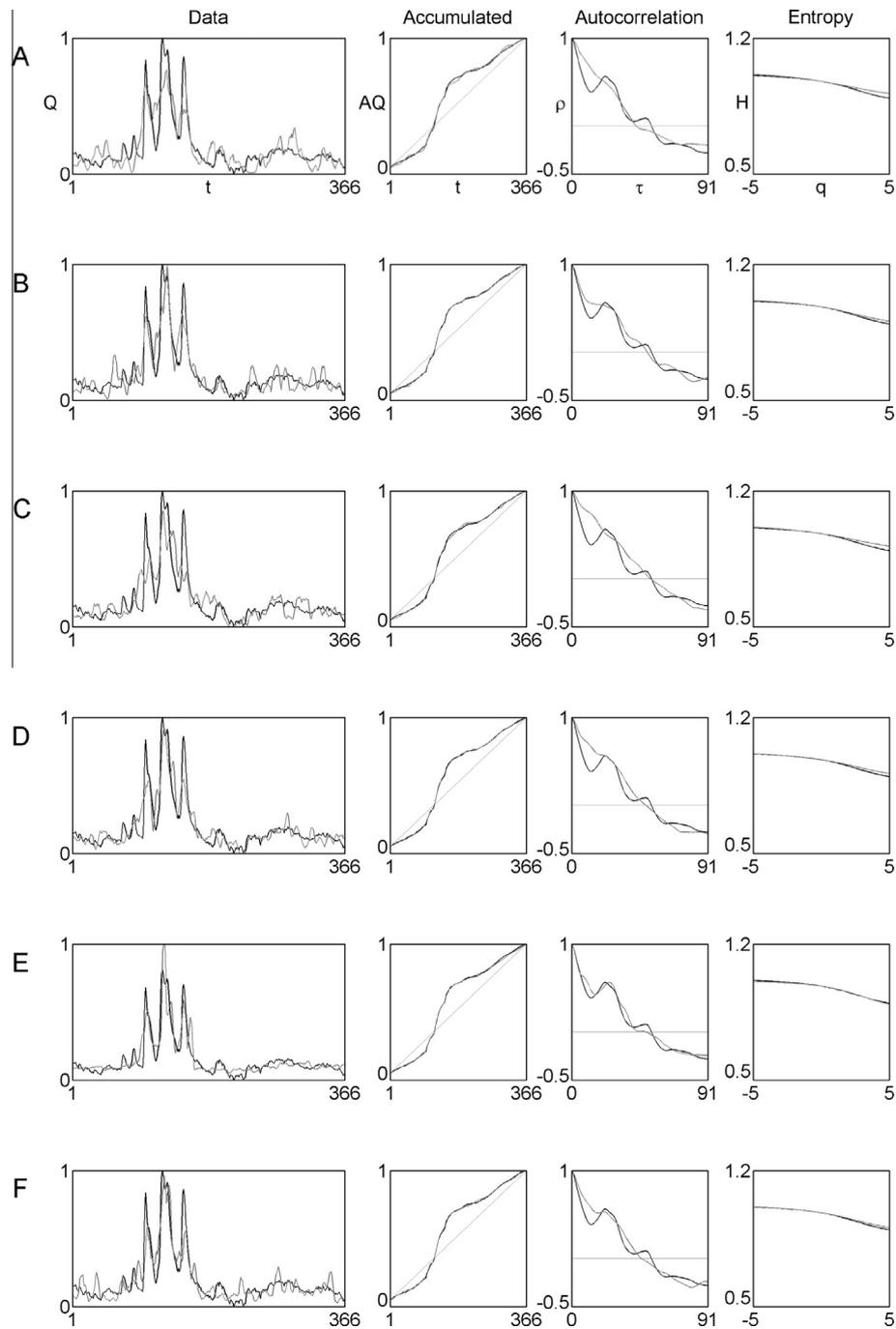


Fig. 10. Daily streamflow records at the Sacramento River for water year 2008 (black) and six FM representations A to F (gray), followed by their accumulated sets, autocorrelation functions, and Rényi entropy functions. A and B are defined via “wires” corresponding to two and three maps. C and D are obtained from “leaves” based on two and three maps. E closes a loop of wires each using two maps and F combines two leaves found via two maps each.

Table 2
Error statistics related to Fig. 5.

Fit	RMSEAS (%)	MAXEAS (%)	NSD (%)	NSA (%)	NSH (%)	NSE (%)	PO ₁₀	PO ₂₀
A	1.0	1.8	87.6	95.5	91.9	99.3	27.9	7.1
B	0.8	1.8	87.8	94.8	87.3	98.9	26.8	7.9
C	0.9	1.9	87.8	93.3	87.0	99.2	27.9	7.9
D	0.5	1.1	91.4	99.7	95.9	99.9	17.3	3.6
E	0.8	1.5	89.2	96.7	91.6	99.6	26.3	6.0
F	0.7	1.4	77.5	97.9	89.6	99.6	33.2	12.3

Table 3
Error statistics related to Fig. 6.

Fit	RMSEAS (%)	MAXEAS (%)	NSD (%)	NSA (%)	NSH (%)	NSE (%)	PO ₁₀	PO ₂₀
A	1.0	2.4	84.5	96.8	98.7	100.0	12.3	3.0
B	0.5	1.3	89.4	98.9	97.1	99.9	10.7	1.6
C	0.9	2.2	84.6	98.2	99.2	99.7	9.8	1.9
D	0.4	1.2	88.9	99.2	98.5	100.0	10.1	0.3
E	0.5	1.5	88.9	98.7	98.7	100.0	8.2	0.8
F	0.5	1.3	85.9	99.0	99.7	99.8	9.0	1.4

Table 4
Error statistics related to Fig. 7.

Fit	RMSEAS (%)	MAXEAS (%)	NSD (%)	NSA (%)	NSH (%)	NSE (%)	PO ₁₀	PO ₂₀
A	1.3	2.9	64.6	85.9	96.2	97.2	36.4	15.1
B	0.7	2.1	80.9	89.9	94.9	97.9	20.5	4.7
C	1.0	2.6	71.6	91.5	97.9	98.2	14.8	1.9
D	0.8	2.0	88.4	98.4	95.0	99.9	15.6	2.7
E	0.7	1.8	85.2	91.8	98.7	99.3	17.0	2.7
F	0.5	1.4	85.3	96.3	90.4	99.6	8.5	2.2

Table 5
Error statistics related to Fig. 8.

Fit	RMSEAS (%)	MAXEAS (%)	NSD (%)	NSA (%)	NSH (%)	NSE (%)	PO ₁₀	PO ₂₀
A	1.2	2.3	79.0	98.7	94.2	99.7	36.3	6.8
B	0.5	1.6	90.0	97.8	94.2	99.6	14.2	3.0
C	1.0	2.1	81.0	99.4	97.2	98.8	27.6	6.3
D	0.5	1.2	91.8	99.4	98.5	99.9	10.4	2.2
E	0.6	1.3	92.8	99.3	96.2	99.4	10.1	1.6
F	0.5	1.2	88.5	99.6	95.7	99.8	11.4	2.5

Table 6
Error statistics related to Fig. 9.

Fit	RMSEAS (%)	MAXEAS (%)	NSD (%)	NSA (%)	NSH (%)	NSE (%)	PO ₁₀	PO ₂₀
A	1.0	2.2	85.0	97.6	22.7	99.8	35.6	6.0
B	0.5	1.1	89.9	99.0	83.6	99.4	22.2	3.0
C	0.7	1.5	82.3	98.9	65.2	99.9	29.0	10.6
D	0.5	1.1	93.6	99.8	79.4	99.9	15.3	0.0
E	0.5	1.2	81.4	98.8	80.6	99.6	22.7	6.0
F	0.5	1.2	90.3	98.5	73.6	99.7	23.0	3.6

Table 7
Error statistics related to Fig. 10.

Fit	RMSEAS (%)	MAXEAS (%)	NSD (%)	NSA (%)	NSH (%)	NSE (%)	PO ₁₀	PO ₂₀
A	1.2	3.0	68.1	90.6	87.8	92.7	25.1	7.1
B	0.6	1.7	77.4	95.6	80.6	97.5	17.8	5.7
C	1.0	2.3	64.5	85.2	92.5	93.2	23.2	7.7
D	0.5	1.6	80.0	90.5	97.0	96.3	15.8	5.2
E	0.7	1.5	79.2	97.4	94.4	99.4	9.6	4.4
F	0.6	1.5	77.7	94.2	86.6	98.3	16.4	3.8

than the others. For example, fits D, E, and F do look better than the other three. This is also reflected by the rather high values of NSD of, in order, 88.4, 85.2, and 85.3%. The last two columns of this figure include the autocorrelation functions and the scatter plots for all the fits. As seen, the better fitting of records results in better fittings of autocorrelations, and, in this regard, the same aforementioned model D and also the closed loop F produce excellent followings of the decay of the undulating function, corresponding to remarkable NSA values of 98.4 and 96.3%, respectively (Table 4). As observed, the scatter plots, with the $\pm 10\%$ bands superimposed and defined by the largest scale of the data or fits, yield rather similar results for the models D to F, with statistics PO₁₀ and PO₂₀ that are less than 17.0 and 2.7%, respectively. Just as the case for the previous years

(1958 and 1968), all the six FM representations turn out to be reasonable, even if there is a noticeable overestimation of the second peak by variants labeled C and F. As seen in Table 4, all the models also approximate the histograms and the entropies rather well.

Continuing with the passage of time while recognizing the variability of observed streamflow, Fig. 8 presents the results for water year 1984 that concentrates its bulky mass at the beginning of the time period. As seen in the accumulated sets (second column) and as corroborated by the first two columns in Table 5, once again all the six FM models result in excellent fittings of the accumulated records, with RMSEAS and MAXEAS values less than 1.2 and 2.3%, respectively. Regarding the FM fits themselves (first column), a visual inspection reveals that those labeled B, D, and E are best,

which correspond to rather high NSD values of, in order, 90.0, 91.8, and 92.8% (see Table 5). While all FM models track closely the slowly decaying autocorrelation function (third column), i.e., NSA greater than 97.8% (Table 5), the scatter plots (fourth column) just reflect the visual quality of the fits, with the aforementioned B, D, and E again being best, as corroborated by the PO_{10} and PO_{20} values in Table 5.

Finally, and to conclude the analysis with yet distinct geometries, the goodness of the FM fits of streamflow for water years 1998 and 2008 is presented, respectively, in Figs. 9 and 10, with

the corresponding statistical information given in Tables 6 and 7. While streamflow in the water year 1998 remains close to zero for about three months and then abruptly rises to a high level for about six months (Fig. 9, first column), water year 2008 contains three distinct and sharp peaks in streamflow that happen during the first six months (Fig. 10, first column). This leads to markedly different accumulated records that, as has happened all along for the other four water years considered earlier, are very well reproduced by the FM approaches (see the first two columns of Tables 6 and 7 and notice that representations C and F in Fig. 10 correspond

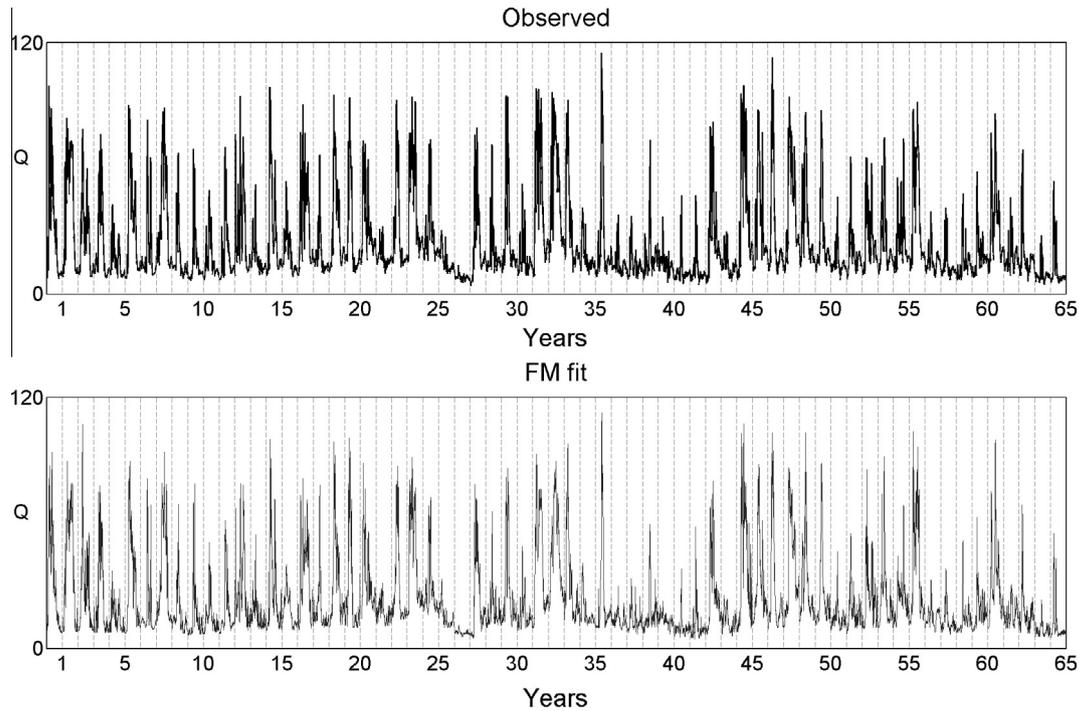


Fig. 11. Measured streamflow (in thousand cubic fit per second) gathered in the Sacramento River from water year 1951 until 2015 (top) and corresponding FM representations (bottom and found year by year) based on a fractal leaf representation employing three maps.

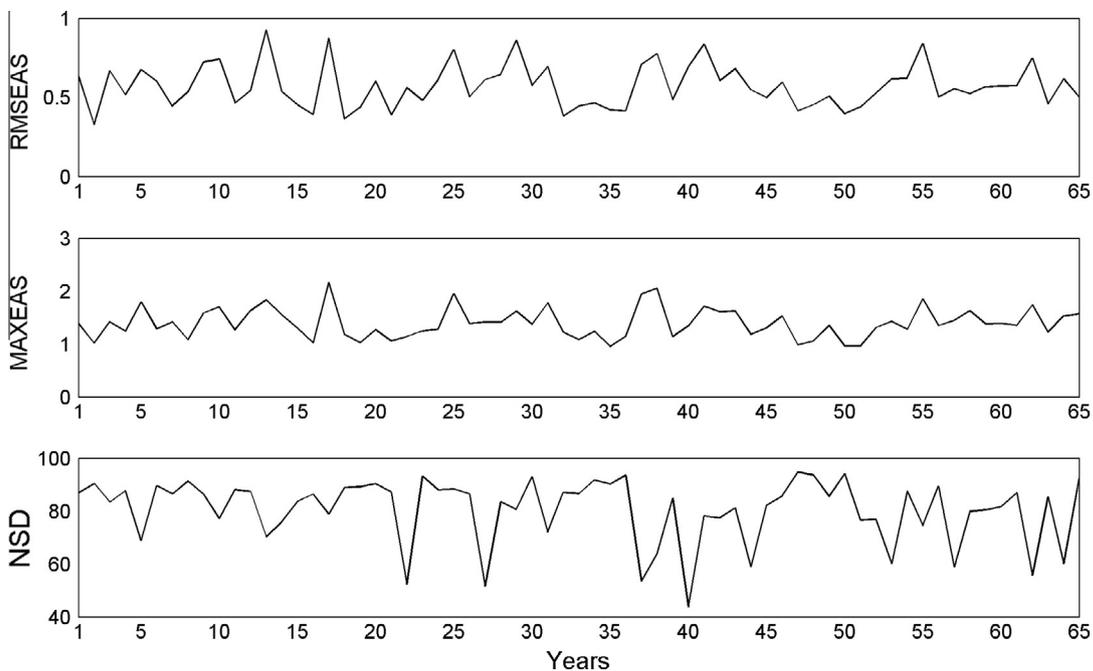


Fig. 12. Evolution of FM related statistics associated with Fig. 11: RMSEAS, MAXEAS and NSD.

to the constructions shown in Figs. 2 and 4, respectively). While the FM fits themselves for streamflow in the year 1998 appear to the eye to be reasonably good across the board (with NSD values exceeding a robust 81.4%, see Table 6), those for the more complex streamflow in the year 2008 have varying degrees of success. Although the best model all along, the one labeled D, remains the best in NSD at a value of 80.0 (Table 7), Fig. 10 also reveals that such a model does not capture, as well, the wavy nature of the autocorrelation function and that models B and E clearly perform better. As found for the other four water years above, and as illustrated in the last column of Figs. 9 and 10 and the attribute NSE in Tables 6 and 7, all the six FM models are capable of maintaining the Rényi entropy function of the records very well. As may be verified, Tables 6 and 7 do corroborate the goodness of the fits in terms of scatter and histogram information. However, a rather low value for NSH (Table 6) for the simplest model A and also, to a lesser

extent, for model C, reveals that such representations may not be suitable for streamflow from water year 1998, despite being reasonable in all other attributes. This result is certainly interesting, as it suggests that the used objective function may require yet an additional constraint in order to find even better FM solutions.

In summary, it may be argued that although all the six streamflow records illustrated herein require FM representations based on three or four maps for their encoding, attractors based on two maps may also be sufficient, in some cases, to represent the streamflow records. As has been illustrated, the goodness of a variant of the FM approach not only depends on the number of maps but also on the inherent complexity of the records being encoded. Even though closing the loops using “wires” or “leaves” requires a higher number of maps, the results they produce are, at the end, similar to those found from FM approaches based on three maps. In this regard, it may also be noted that leafy attractors with three

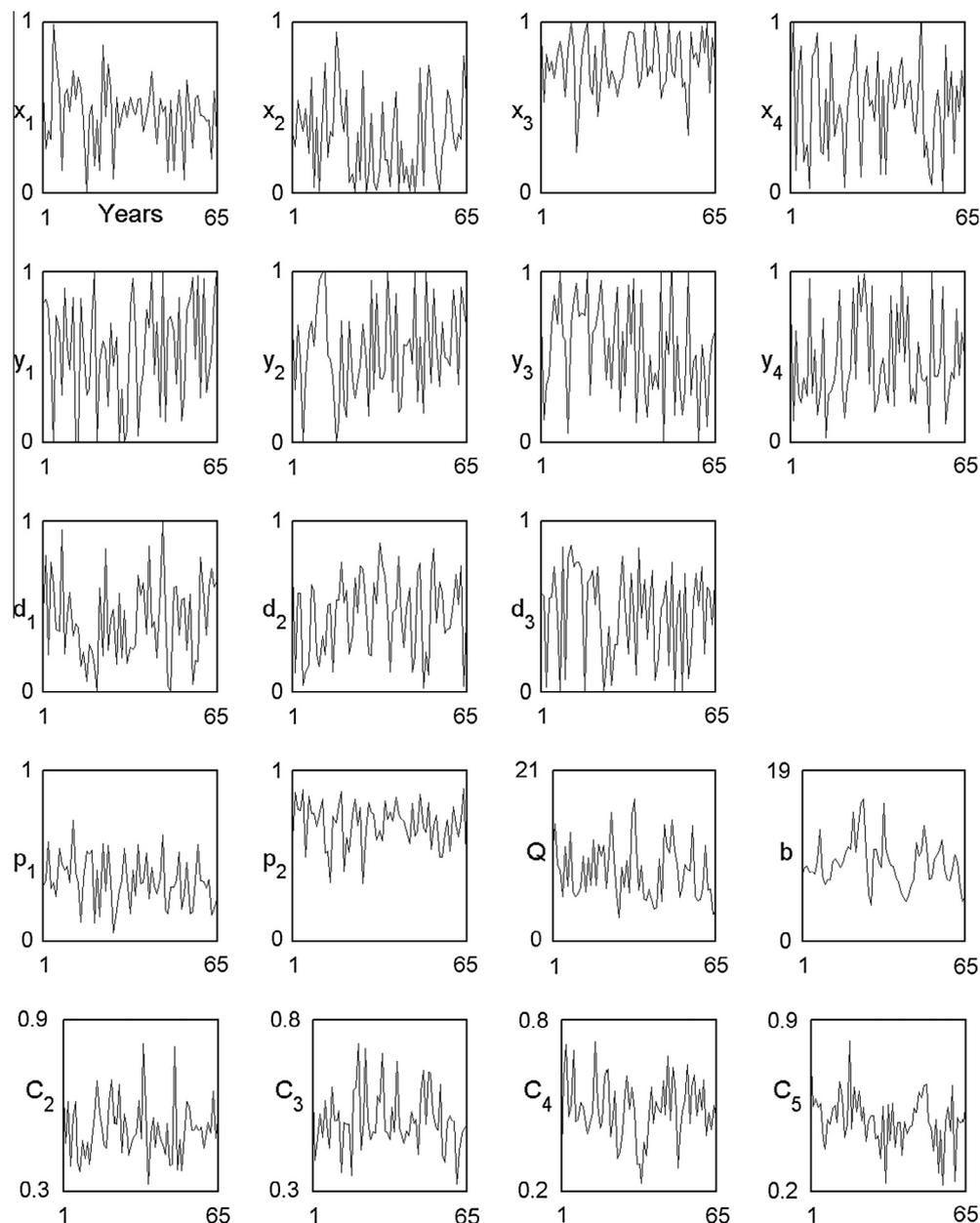


Fig. 13. Evolution of FM parameters for the representation based on a “leaf” and three maps over 65 years of streamflow records gathered in the Sacramento River, starting from water year 1951 till 2015. Shown y_i values are normalized from 0 to 1, but correspond to actual values between -5 and 5 . Shown d_i values are also normalized from 0 to 1, but correspond to actual values between -1 and 1 . The quantities Q and b , below, correspond to the total flow and the baseflow during a year, in million cubic feet per second and thousand cubic feet per second, respectively. C_2 to C_5 are the second to fifth normalized principal components associated with Q .

maps require just one more parameter than the closed loop using two leaves and four maps, but such representations have similar levels of quality.

Given the fact that different variants of the FM approach, using either the original FM formulation or generalized approaches based on overlaps on the inherent maps and also with or without loops, produce similar results, it is clear that there are also different ways in which a given set of records may be approximated. As illustrated in Figs. 5–10, only the original FM approach based on the iteration of two maps may be considered to perform consistently worse than the other five representations. On the other hand, overall, the usage of the leafy attractor based on three maps (case D in the figures and tables) may be considered the best, as it exhibits better performance in at least four major statistical qualifiers when compared to the other representations.

As mentioned before, the results presented here are chosen according to minimum values of maximum deviations in accumulated sets, MAXEAS. As such, there are several other “solutions” that have slightly lesser RMSEAS, which guided the optimization exercise. This fact further corroborates that the inverse problem, as it provides an approximation, does not give a unique solution, as it is usually the case in other disciplines (Hill and Tiedeman, 2007).

5. Encoding the Sacramento River over 65 years

Having illustrated the goodness of the FM model based on three maps generating a leafy attractor (i.e., Model D), this section includes results for streamflow encodings at the Sacramento River for the entire period of 65 years considered in this study, i.e., water years spanning 1951 to 2015. Fig. 11 compares the raw data with the obtained FM representations properly scaled, that is, adding a constant base flow over the year and upscaling by the total volume during the year. As seen, the overall encoding by FM model D is remarkable, as it follows the records in great detail. Similarly to what was reported in the previous section, such a geometric model is found to be consistently the best of the six variants and the corresponding RMSEAS, MAXEAS, and NSD values over the 65 years, as seen in Fig. 12, yield (with few exceptions on NSD) excellent performance, with averages plus or minus one standard deviation of, in order, 0.6 ± 0.1 , 1.4 ± 0.3 and $81 \pm 12.3\%$.

Fig. 13 shows the evolution of the key parameters of the FM encodings based on 14 geometric parameters, and also includes, after the weights used in the iterations p_1 and p_2 , the total flow Q and the baseflow b for each one of the years and, further down in the last row, the second to fifth normalized principal components for the total flows (as the first principal component follows closely the set Q). As readily seen, all FM parameters do exhibit a substantial amount of variation, akin of the one observed in the records themselves and in the principal components. When duly analyzed, all sets have correlations that decay rather rapidly and all of them share similar small scales – as defined by the first lag where correlation crosses zero – that are less than 2 years, and with autocorrelation functions that are always bounded by ± 0.2 bands. While the shown principal components are orthogonal and hence have zero cross-correlations, the FM parameters turn out to have also rather insignificant cross-correlations (less than 0.1), except between: p_1 and p_2 (0.53), x_3 and y_2 (–0.40), x_2 and x_3 (0.38), and x_1 and p_1 (0.36).

Certainly, and surely at a first glance, Fig. 13 does not show any obvious trends in FM parameters that may define nonstationarity in the mean for any of the graphs, but rather reflects ample variability and perhaps statistical stationarity altogether. Although the length of the sets limits what may be argued, the FM parameters for this specific example turn out to have similar

degrees of complexity as the annual records, base flows and principal components. Even though this fact is a clear reflection of the intrinsic variability of the natural processes at the annual scale, it should not be forgotten, however, that the FM parameters do allow visualizing the dynamics of the streamflow process year by year at the daily scale.

In an attempt to better understand relevant scale issues, a more complete analysis of FM parameter patterns for additional sites is being carried (the John Day River in Oregon, the Green River in Washington, the Mississippi River at various stations and at the Red River of the North), and also employing locally averaged streamflow sets every 5, 10 and 15 years. Certainly, the use of local averages – as done in climatic studies – results in successive annual sets at the daily scale that do look alike, which hence yields FM parameters that exhibit less variability than the ones in Fig. 13. It is envisioned that the structure in these parameter evolutions, or lack thereof, would be of relevance to compare and define the complexity of distinct sets at various sites and to assess if there are any trends in time, at different averaging scales, that may be related to climate change attributes or other effects such as the one produced by the construction of dams. Results of these analyses will be reported elsewhere.

6. Concluding remarks

This study has presented an application of a fractal-multifractal (FM) approach and some of its variants for encoding, for the first time, daily streamflow data observed in the Sacramento River in California, USA. The results indicate the appropriateness of all the six FM models considered in the study for encoding the daily streamflow geometrically and without the need for other physical (hydrological) information. The study clearly demonstrates how the notions that encompass the FM approach (i.e., projections of measures defined over fractal attractors) could be employed as a suitable tool to describe (encode) streamflow and other geophysical records that are complex and ‘seemingly random’ in nature. As illustrated herein, and elsewhere for rainfall sets (Maskey et al., 2015), such a geometric procedure could supplement and complement many of the existing stochastic procedures that aim at the modeling of complex sets. Certainly, the notions add to the idea that hidden determinism may be at the root of natural complexity (Puente, 1996; Puente and Sivakumar, 2007), as the study also suggests that understanding the dynamics of river flows may be attempted in a novel manner, by following the evolution of the FM parameters, as illustrated for one FM representation over 65 years of streamflow records.

In regards to the encoding of streamflow records, we have demonstrated not only the capability of the original and generalized FM approaches, based on the iteration of maps without or with overlaps (i.e. “wires” and “leaves”) but also the goodness of “closing the loop” notions that combine various attractors while sharing a common (multifractal) input measure. Given that the FM methods have been found useful in the modeling of rather complex rainfall patterns (Puente and Obregón, 1996; Cortis et al., 2009, 2013; Huang et al., 2012, 2013; Maskey et al., 2015), it is not surprising that the same notions would do very well with the generally less-complex streamflow sets. This work, however, shows, for the first time, that streamflow records may indeed be thought of as transformations of multifractal measures via fractal attractors, in a manner that possesses a physical interpretation as a special outcome of a non-trivial multiplicative cascade (Cortis et al., 2013). Certainly, and as presented in Table 1, the extensions of the original FM approach require a higher number of parameters, but the additional flexibility gained is noteworthy in capturing more details (and not just statistical information about

the records), while yielding still substantial compression ratios (28:1 at the daily scale for a “leaf” based on three maps). As streamflow are not measured without error (Hammel et al., 2006), the results shown here properly fit within the expected bands of streamflow measurement uncertainty.

As mentioned in the previous section, the FM method is currently being applied to various streamflow sets and employing local averages of the data that span 5, 10 and 15 years. Such studies, to be published elsewhere, shall include a more detailed assessment and comparison of the evolution of FM parameters using detrending techniques, including the fact that the FM approach is equifinal, as there are various distinct parameter sets that yield rather similar encodings. An examination of the FM parameter trends for long periods of streamflow records, with possible linkages to streamflow dynamics and climatic changes via plausible correlations of FM parameters and climatic attributes, is certainly a natural extension to the present study. We also envision that joint FM codings of rainfall (Maskey et al., 2015) and streamflow records (as done here), and within the same

catchment, may help define, via the intrinsic complexity of FM evolutions of rainfall and runoff, distinct climatic/geomorphic regimes.

Acknowledgements

A Henry A. Jastro Award at the University of California, Davis provided valuable support to the first author. Bellie Sivakumar acknowledges the financial support from the Australian Research Council (ARC) through the Future Fellowship grant awarded to him (FT110100328). The team of USGS service is also greatly appreciated for making the historical streamflow records available in its web portal. Comments and suggestions by anonymous reviewers helped improve the manuscript and are gratefully acknowledged.

Appendix A

Parameters for all FM representation in Figs. 5–10. FM parameters from “wires” and two maps.

Figure	x_1	y_1	d_1	d_2	p
5A	0.068	-2.068	0.703	-0.671	0.440
6A	0.659	-0.847	-0.465	0.333	0.559
7A	0.723	-1.160	-0.647	0.410	0.547
8A	0.623	-0.012	0.943	-0.549	0.868
9A	0.277	-0.032	-1.000	-0.714	0.444
10A	0.890	2.287	-0.646	-0.509	0.596

FM parameters from “wires” and three maps.

Figure	x_1	x_2	y_1	y_2	d_1	d_2	d_3	p_1	p_2
5B	0.348	0.931	-3.147	-0.819	1.000	0.961	0.000	0.694	0.128
6B	0.073	0.773	-0.288	-1.334	0.577	-0.468	0.283	0.124	0.628
7B	0.133	0.764	0.525	-0.954	-0.581	-0.394	0.599	0.441	0.406
8B	0.209	0.663	-2.120	4.637	0.397	-0.937	0.692	0.356	0.518
9B	0.529	0.997	1.645	5.000	0.865	0.098	0.406	0.660	0.308
10B	0.412	0.774	2.178	-1.182	-0.519	-0.353	0.726	0.374	0.389

FM parameters from “leaves” and two maps.

Figure	x_1	x_2	y_1	y_2	d_1	d_2	p
5C	0.825	0.111	1.897	3.475	0.668	0.560	0.214
6C	0.458	0.395	3.453	1.422	-0.175	0.169	0.458
7C	0.577	0.525	-0.535	-2.172	0.291	0.237	0.510
8C	0.463	0.078	1.020	-4.408	-0.965	0.527	0.190
9C	0.529	0.483	0.900	2.908	0.735	0.002	0.418
10C	0.698	0.688	4.455	1.467	-0.614	-0.566	0.564

FM parameters from “leaves” and three maps.

Figure	x_1	x_2	x_3	x_4	y_1	y_2	y_3	y_4	d_1	d_2	d_3	p_1	p_2
5D	0.134	0.680	0.887	0.025	-2.218	2.093	5.000	4.563	0.893	0.259	-1.000	0.453	0.294
6D	0.452	0.762	1.000	0.420	-1.839	-3.659	4.993	0.261	-0.445	0.078	0.289	0.511	0.286
7D	0.750	0.352	0.755	0.930	-2.870	-1.385	-1.316	-1.251	-0.510	-0.080	-0.671	0.562	0.115
8D	0.499	0.533	0.944	0.109	4.645	-1.291	0.781	-0.652	-0.507	0.482	0.385	0.418	0.168
9D	0.543	0.732	0.995	0.576	2.383	1.185	0.185	0.896	-0.985	0.504	-0.695	0.318	0.547
10D	0.573	0.601	0.823	0.868	4.755	-0.122	0.135	-3.927	-0.649	-0.261	0.386	0.567	0.144

FM parameters from combining two “wires” found via two maps each.

Figure	x_1	y_1	d_1	d_2	p	\hat{y}_1	\hat{d}_1	\hat{d}_2	w
5E	0.692	-4.952	-0.867	0.299	0.545	-3.326	-1.000	0.562	0.332
6E	0.660	-3.437	-0.762	-0.084	0.803	5.000	-0.041	0.256	0.759
7E	0.840	-1.275	0.263	0.168	0.659	-3.372	-0.516	-0.605	0.202
8E	0.380	0.875	0.966	-0.693	0.474	3.135	-0.075	-0.402	0.531
9E	0.757	4.993	-0.774	-0.372	0.780	4.440	-1.000	0.139	0.539
10E	0.538	3.983	-0.205	-0.305	0.513	1.591	0.456	0.290	0.771

FM parameters from combining two “leaves” found via two maps each.

Figure	x_1	x_2	y_1	y_2	d_1	d_2	p	\hat{y}_1	\hat{y}_2	\hat{d}_1	\hat{d}_2	w
5F	0.007	0.906	1.270	-0.508	0.142	0.782	0.214	0.858	1.005	0.850	-0.829	0.556
6F	0.530	0.475	1.161	-4.000	0.429	-0.130	0.610	0.620	4.956	-0.608	-0.692	0.565
7F	0.002	0.896	-2.077	-0.603	-0.169	0.890	0.545	-5.000	2.934	-0.455	-0.704	0.357
8F	0.544	0.322	-1.067	-0.773	0.079	-0.032	0.607	1.942	3.741	-0.604	-0.629	0.413
9F	0.817	0.550	-2.464	-4.061	0.078	0.039	0.623	-3.244	-1.393	-0.143	0.188	0.601
10F	0.705	0.487	5.000	0.649	-0.571	-0.436	0.448	4.680	4.239	-0.849	-0.343	0.554

References

- Barnsley, M.F., 1988. *Fractals Everywhere*. Academic Press, San Diego.
- Chen, J., Wu, Y., 2012. Advancing representation of hydrologic processes in the Soil and Water Assessment Tool (SWAT) through integration of the TOPographic MODEL (TOPMODEL) features. *J. Hydrol.* 420–421, 319–328.
- Cortis, A., Puente, C.E., Sivakumar, B., 2009. Nonlinear extensions of a fractal-multifractal approach for environmental modeling. *Stoch. Environ. Res. Risk Assess.* 23 (7), 897–906.
- Cortis, A., Puente, C.E., Huang, H.H., Maskey, M.L., Sivakumar, B., Obregón, N., 2013. A physical interpretation of the deterministic fractal-multifractal method as a realization of a generalized multiplicative cascade. *Stoch. Environ. Res. Risk Assess.* 28 (6), 1421–1429.
- Eagleson, P.S., 1970. *Dynamic Hydrology*. McGraw-Hill, New York.
- Eagleson, P.S., 1978. Climate, soil, and vegetation, 2. The distribution of annual precipitation derived from observed storm sequences. *Water Resour. Res.* 14 (5), 713–721.
- Fernández Martínez, J.L., García Gonzalo, E., Fernández-Álvarez, J.P., Kuzma, H.A., Menéndez Pérez, C.O., 2010. PSO: a powerful algorithm to solve geophysical inverse problems: application to a 1D-DC resistivity case. *J. Appl. Geophys.* 71 (1), 13–25.
- Freeze, R.A., Cherry, J.A., 1979. *Groundwater*. Prentice Hall, Englewood Cliffs.
- Hammel, R.D., Cooper, R.J., Slade, R.M., Haney, R.L., Arnold, J.G., 2006. Cumulative uncertainty in measured streamflow and water quality data for small watersheds. *Trans. Am. Soc. Agric. Eng.* 49 (3), 689–701.
- Hill, M.C., Tiedeman, C.R., 2007. *Effective Groundwater Model Calibration: With Analysis of Data, Sensitivities, Predictions and Uncertainty*. John Wiley & Sons, New Jersey.
- Huang, H.H., Puente, C.E., Cortis, A., Sivakumar, B., 2012. Closing the loop with fractal interpolating functions for geophysical encoding. *Fractals* 20 (3–4), 261–270.
- Huang, H.H., Puente, C.E., Cortis, A., Fernández Martínez, J.L., 2013. An effective inversion strategy for fractal-multifractal encoding of a storm in Boston. *J. Hydrol.* 496, 205–216.
- Kavvas, M.L., Delleur, J.W., 1981. A stochastic cluster model of daily rainfall sequences. *Water Resour. Res.* 17 (4), 1151–1160.
- Lorenz, E.N., 1963. Deterministic nonperiodic flow. *J. Atmos. Sci.* 20, 130–141.
- Lovejoy, S., Schertzer, D., 2013. *The Weather and Climate: Emergent Laws and Multifractal Cascades*. Cambridge University Press, Cambridge, pp. 505.
- Mandelbrot, B.B., 1982. *The Fractal Geometry of Nature*. H.B. Fenn and Company.
- Mandelbrot, B.B., 1989. Fractals in geophysics. In: Scholz, C.H., Mandelbrot, B.B. (Eds.), *Multifractal Measures Especially for the Geophysicist*. Birkhauser Verlag, Basel, pp. 1–42.
- Maskey, M.L., Puente, C.E., Sivakumar, B., Cortis, A., 2015. Encoding daily rainfall records via adaptations of the fractal multifractal method. *Stoch. Environ. Res. Risk Assess.* <http://dx.doi.org/10.1007/s00477-015-1201-7>.
- Meneveau, C., Sreenivasan, K.R., 1987. A simple multifractal cascade model for fully developed turbulence. *Phys. Rev. Lett.* 59, 424–427.
- Nash, J., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models, part I—a discussion of principles. *J. Hydrol.* 10 (3), 282–290.
- Obregón, N., Sivakumar, B., Puente, C.E., 2002a. Modelling high resolution rain rates via a deterministic fractal-multifractal approach. *Fractals* 10 (3), 387–394.
- Obregón, N., Sivakumar, B., Puente, C.E., 2002b. A deterministic geometric representation of temporal rainfall. Sensitivity analysis for a storm in Boston. *J. Hydrol.* 269 (3–4), 224–235.
- Özger, M., Mishra, A.K., Singh, V.P., 2010. Scaling characteristics of precipitation data in conjunction with wavelet analysis. *J. Hydrol.* 395, 279–288.
- Poornima, U., Jothiprakash, V., 2015. Extraction of trend of rainfall using singular spectrum analysis. *J. Hydrol. Eng.* 20 (12), 05015007.
- Puente, C.E., 1996. A new approach to hydrologic modelling: derived distribution revisited. *J. Hydrol.* 187, 65–80.
- Puente, C.E., 2004. A universe of projections: may Plato be right? *Chaos, Solitons Fractals* 19, 241–253.
- Puente, C.E., Obregón, N., 1996. A deterministic representation of temporal rainfall: result for a storm in Boston. *Water Resour. Res.* 32 (9), 2825–2839.
- Puente, C.E., Obregón, N., 1999. A geometric platonic approach to multifractality and turbulence. *Fractals* 7 (4), 403–420.
- Puente, C.E., Sivakumar, B., 2003. A deterministic width function model. *Nonlinear Process Geophys.* 10, 525–529.
- Puente, C.E., Sivakumar, B., 2007. Modeling hydrologic complexity: a case for geometric determinism. *Hydrol. Earth Syst. Sci.* 11, 21–724.
- Puente, C.E., Robayo, O., Díaz, M.C., 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, C.E., Robayo, O., Díaz, M.C., Sivakumar, B., 2001b. A fractal-multifractal approach to groundwater contamination. 2. Predicting conservative tracers at the Borden site. *Stoch. Environ. Res. Risk Assess.* 5 (5), 372–383.
- Rodríguez-Iturbe, I., 1986. Scale of fluctuation of rainfall models. *Water Resour. Res.* 22 (9), 155–375.
- Sivakumar, B., Sorooshian, S., Gupta, H.V., Gao, X., 2001. A chaotic approach to rainfall disaggregation. *Water Resour. Res.* 37, 61–72.
- Tongal, H., Demirel, M.C., Booi, M.J., 2013. Seasonality of low flows and dominant processes in the Rhine River. *Stoch. Environ. Res. Risk Assess.* 27, 489–503.
- Yang, D., Sun, F., Liu, Z., Cong, Z., Ni, G., Lei, Z., 2007. Analyzing spatial and temporal variability of annual water-energy balance in nonhumid regions of China using the Budyko hypothesis. *Water Resour. Res.* 43 (4), 1–12. <http://dx.doi.org/10.1029/2006WR005224>.