Predicting Streamflow from Fractal Geometric Encodings of Yearly and Decadal Records

AGU 2015, San Francisco
Paper Number: H33E-1670

Carlos E. Puente,[1*] Mahesh L. Maskey,[1] Bellie Sivakumar[2,1] and Andrea Cortis[3]
[1]University of California, Davis, CA; Department of Land, Air, and Water Resources
[2] University of New South Wales, Sydney, Australia; Department of Civil and Environmental Engineering
[3]ION GX Technology, Houston, TX
* Corresponding Author; E-mail: cepuente@ucdavis.edu

Abstract

In an attempt to model the specific complexity of geophysical (hydrological) records --beyond some key statistical qualifiers--, in the past we have developed a deterministic geometric procedure: the fractal-multifractal (FM) method and have introduced some promising variants. We show here that the FM ideas are indeed capable of faithfully encoding both yearly and decadal runoff sets gathered at the Sacramento River on a daily basis, with maximum cumulative errors that are, for a sixty year period, always less than a mere 2.5%. Then, we explain how the time variation of FM parameters at such scales allow us to: (a) closely predict whole decadal information in a rather precise manner, and (b) issue sensible forecasts of yearly information based on the transfer of FM decadal to FM yearly parameters via k-nearest neighbor algorithms. It is also illustrated how the notions provide relevant classifications (for yearly and decadal sets) that allow following the evolution of the geometry of runoff, with potential applications to global climate change.

Introduction

The understanding, archiving, and forecasting of information is undoubtedly one of the most important problems faced in geophysics in this twenty-first century. Despite clear advances, we are constantly confronted with rather complex data sets containing intricate details that demand our attention, as it happens with a host of hydrological patterns.

In the past years, we have developed a fractal geometric methodology, the fractal-multifractal FM approach, Puente (1996), aimed at capturing the complexity of observed patterns, as they are. Instead of thinking that given data set is a realization of an stochastic process, whose histogram, autocorrelation function, and distribution of inter-arrival times may be fitted, it is assumed that such records are instead a fractal transformation of a multifractal measure. We have shown that these deterministic ideas are indeed useful for encoding rainfall events measured every few seconds (e.g., Puente, 1996; Obregón et al., 2002, Huang et al., 2012) and also streamflow patterns gathered daily (Puente et al., 2013).

The present work shows faithful FM encodings of yearly and decadal runoff time series at the daily scale and presents a methodology for predicting yearly records at once, based on the FM parameters of encodings. For this purpose a total of 64 years gathered at **Sacramento River** near Freeport (USGS station 11447650) were analyzed.

The Fractal-Multifractal Approach

Figure 1 shows the construction of a FM pattern dy (adding up to one) transforming a *multifractal* measure dx (also adding up to one) via a *fractal* interpolating function $f: x \to y$, Barnsley (1988). For such an example, the function passes by {(0,0), (**0.31**, **1.86**), (**0.52**,**−0.85**), (1,1)} and it is defined iterating three simple affine maps:

$$w_1\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}0.31 & 0\\1.42 & \mathbf{0.44}\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix},$$

$$w_2\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}0.21 & 0\\-2.22 & \mathbf{-0.48}\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix} + \begin{pmatrix}0.31\\1.86\end{pmatrix},$$

$$w_3\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}0.49 & 0\\1.23 & \mathbf{0.62}\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix} + \begin{pmatrix}0.52\\-0.85\end{pmatrix},$$

where values in **bold** are FM parameters, *mid-interpolating points* and *vertical scalings*, together with the *proportions* used in doing the iterations, **20**-**10**-70%.

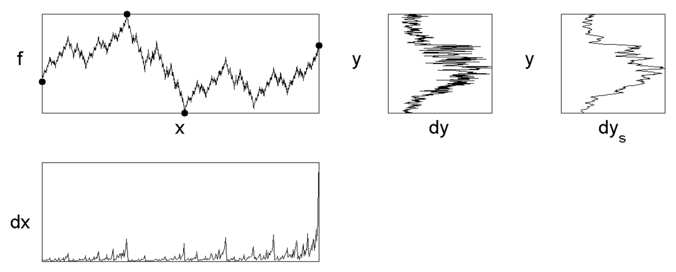


Figure 1. The FM method: from a multifractal dx, to a projection dy, via a fractal interpolating function f, a "wire" from x to y, followed by an smoothed version of set dy_s.

Some FM Encodings of Streamflow at Sacramento River

Figure 2 demonstrates the ability of the FM approach to encode streamflow normalized records (adding up to one), both at the yearly (left) and decadal scales (right), and after removal of a constant base flow. Such were found solving an inverse problem aimed at fitting the accumulated records. Figure 3 shows the implied monthly flows (measured and encoded) for a total of 64 years and 55 decades, and Figure 4 presents the subsequent spring flows.

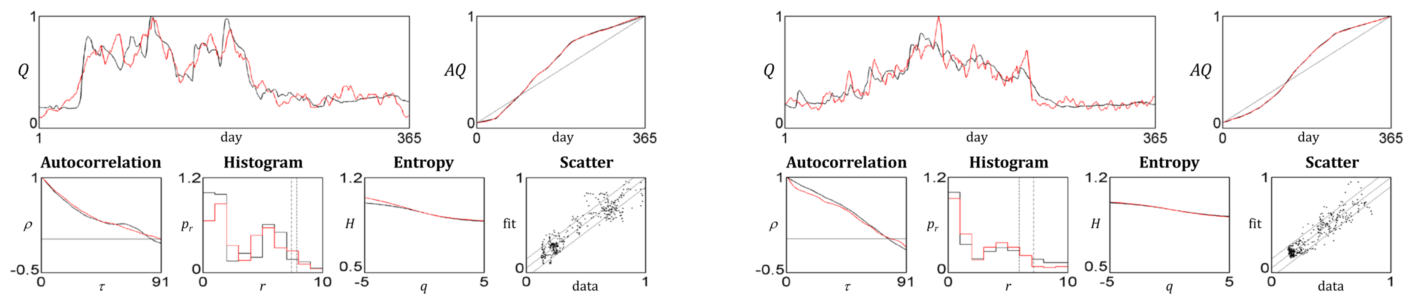


Figure 2. Normalized **yearly** and **decadal** sets (black), for Sept 1973-Oct 1974 and Sept 1956-Oct 1966, and FM representations (red), followed by accumulated runoff (top) and statistical attributes of the fits (bottom).

Figure 3. **Yearly** and **decadal** sets with base flow (black), for Sept 1950-Oct 2014 and Sept 1959-Oct 2014, and FM representations (red), in 10,000 cfs.

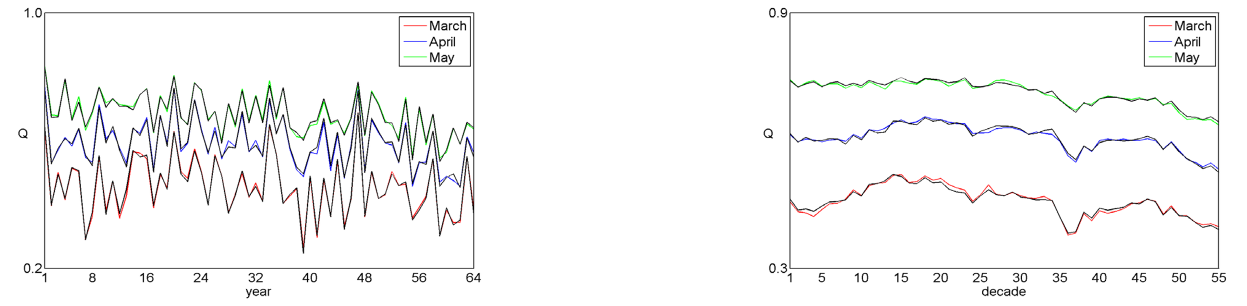


Figure 4. **Yearly** and **decadal** Spring flows for dates in Figure 3 (black) and respective FM fits (in color).

Overall FM Fittings and Parameters

The FM approach successfully encoded all streamflow records at the yearly and decadal scales. The root mean square errors in accumulated records **RMSEAR** (optimized) and maximum error in accumulated records **MAXEAR** over the years and decades are, in percent: *yearly* (0.8 ± 0.3, 1.8 ± 0.5); *decadal* (0.5 ± 0.4, 1.2 ± 0.7). While the goodness of the FM encodings is further appreciated in Figure 5, the corresponding evolution of FM parameters is shown in Figure 6.

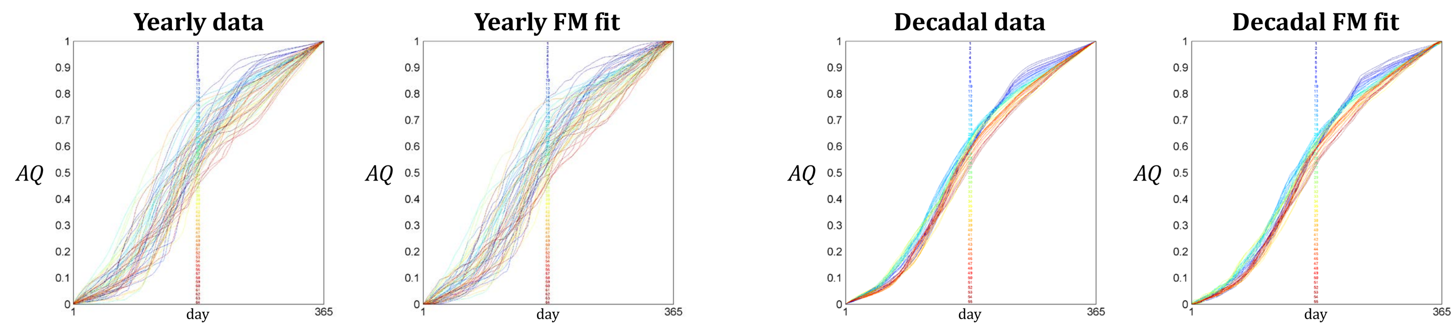


Figure 5. Accumulated sets for streamflow at Sacramento River: **yearly** and **decadal**, over 64 years and 55 decades, and respective FM encodings.

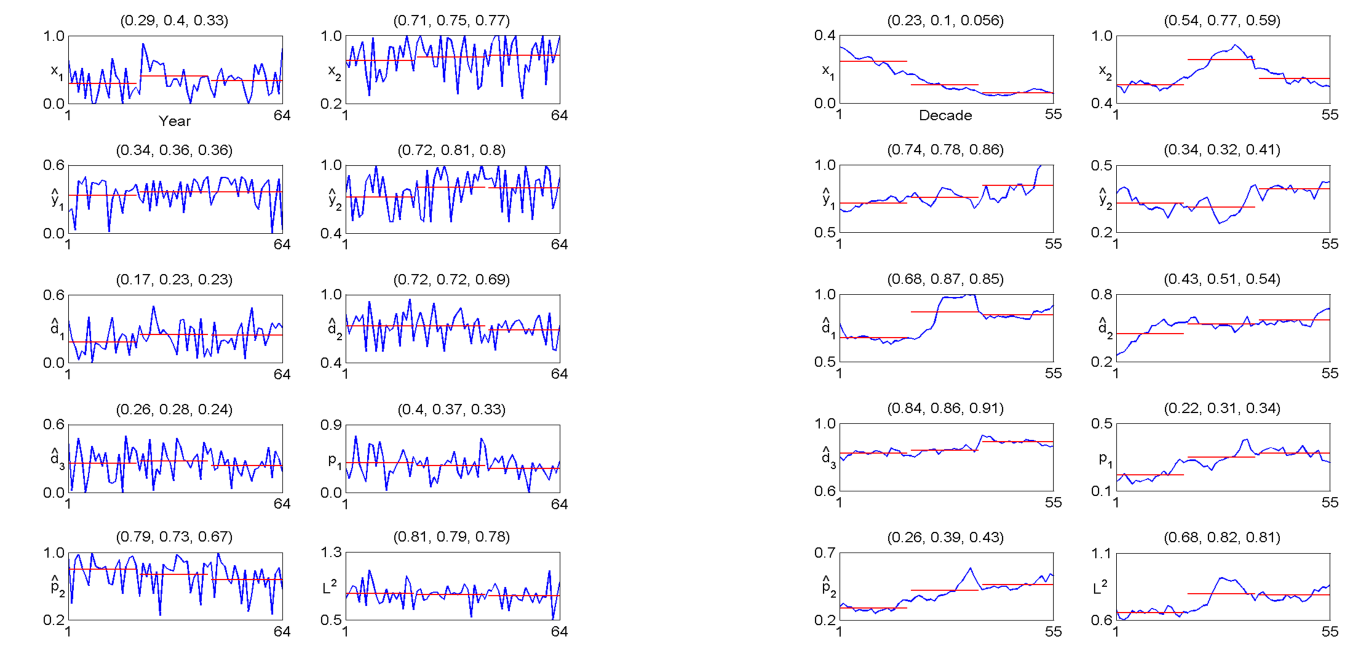


Figure 6. Evolution of FM parameters for **yearly** (left) and **decadal** (right) streamflow records. Values on top of each frame correspond to average parameter values over thirds of the sets, as plotted in red. The last block includes the L-2 norm of all parameters between 0 and 1.

Predicting a Year of Streamflow

The years and decades (and also pentades) of streamflow, up to a given year, may be classified based on k-means clustering of FM parameters in order to study the dynamics of the process and to establish transition matrices that may allow predicting streamflow in a holistic manner: the next decade, the next pentade and, ultimately, the next year. As an example up to water year 1998-99, Figure 7 shows the FM centroids for ten classes each at the scales considered, Figure 8 depicts the time evolution by classes and Figure 9 includes the obtained transition matrices between relevant scales.

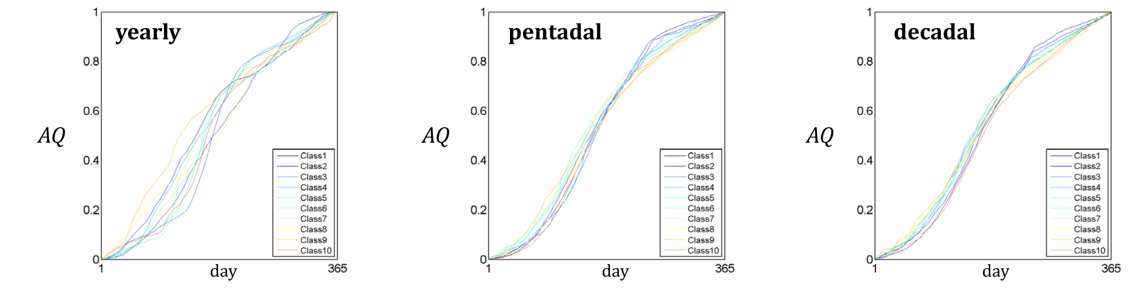


Figure 7. Centroids of FM classification of **yearly**, **pentadal** and **decadal** accumulated streamflow up to year 1998-99 at the Sacramento River.

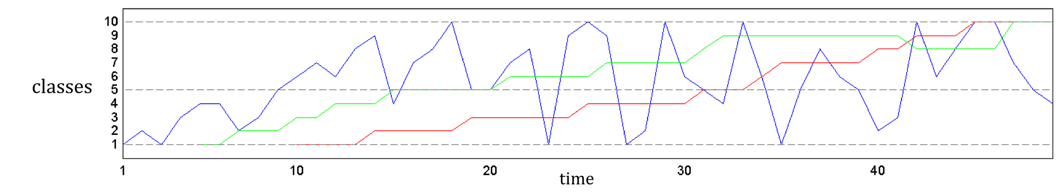


Figure 8. FM Class evolutions for **yearly** (blue), **pentadal** (green) and **decadal** (red) up to year 1998-99 at the Sacramento River.

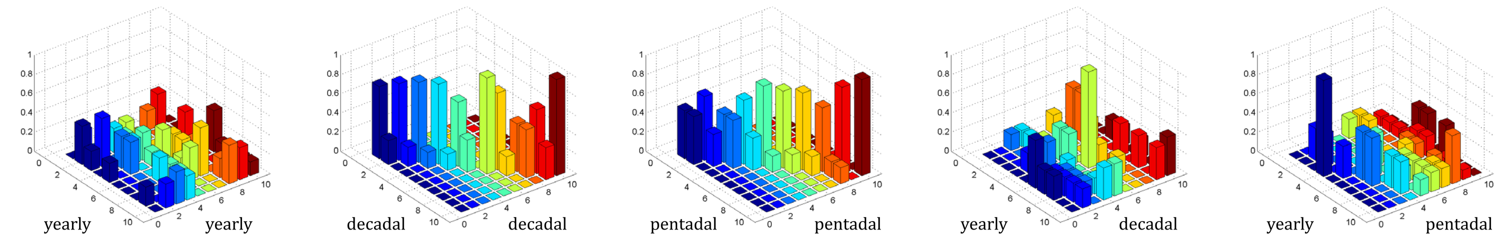


Figure 9. Transition matrices (from right to left) based on FM evolution of classes shown in Figure 8.

Predicting a Year of Streamflow (cont.)

Given the information so obtained, one may try to predict the next year 1999-00, decade 1990-00 and pentade 1995-00, using the transition matrices that summarize the evolution by classes. To this effect, Figure 10 shows predictions based on a time series model of FM parameters (like in Figure 6, right) and on class transitions for both decadal and pentadal scales, including also the average of the former two predictions. Similarly, but without a time series model, Figure 11 depicts the involved predictions at the yearly scale while using class transitions from year to year, from a decade to a year, and from a pentade to a year, together with their averages.

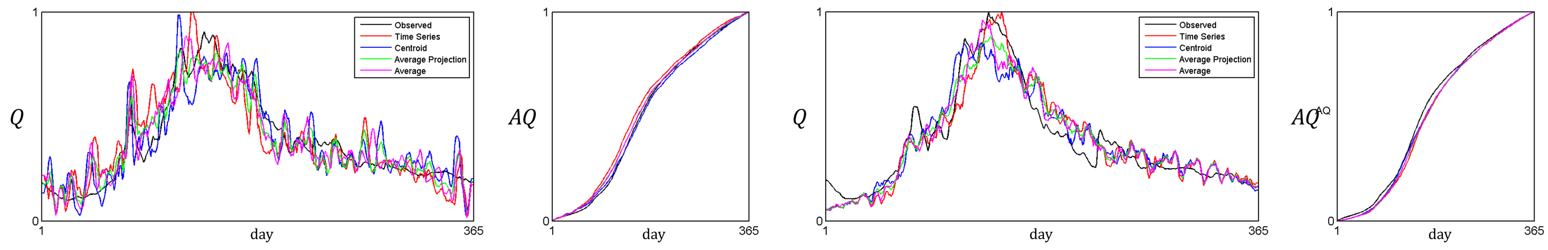


Figure 10. Predictions of **decade** 1990-00 and **pentade** 1995-00 based on time series model of FM parameters and class evolution as implied by second and third matrices in Figure 9.

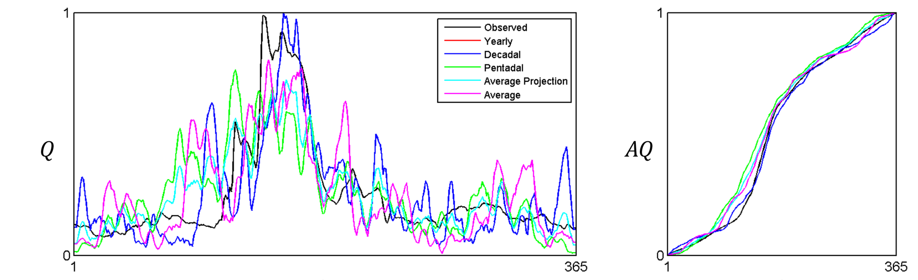
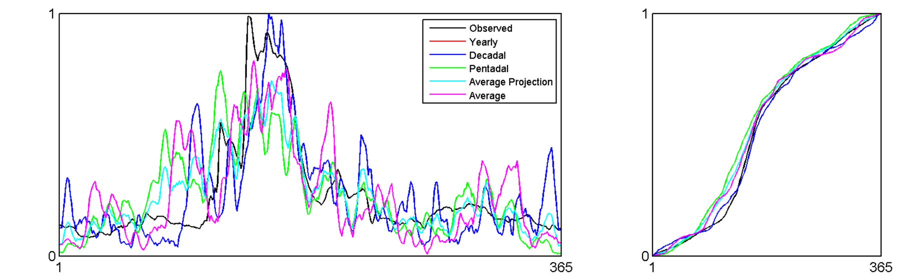


Figure 11. Predictions of **year** 1999-00 based on class evolution as implied by first, fourth and fifth matrices in Figure 9, followed by the accumulated streamflow set containing all years belonging to the class that provided best predictions, i.e., based on decadal information.

As may be seen, the performance obtained for year 1999-00 is quite good and such is confirmed by rather low values of **RMSEAR** and **MAXEAR** of 1.97 and 4.9%. A similar analysis has been carried out for 10 years at the Sacramento River, and as shown in Tables 1 and 2, excellent fittings at the decadal scale are always possible, and sensible predictions at the yearly scale are possible, with errors in the aforementioned attributes that are on the average 5.5 and 12.8%



Summary and Conclusions

The FM approach is capable of encoding the runoff data at the Sacramento river faithfully, with maximum errors in accumulated records which are less than 4% for the original approach having 5 parameters and less than 3% for the variant based on overlaps and defined based on 7 parameters. Despite these excellent fits---and also of the average records by both FM variants---patterns in parameters for subsequent years were not identified. For both FM approaches there is ample variability from year to year, a reflection of the distinct geometric patterns that were analyzed, which indeed vary from year to year.

References

Barnsley MF (1988) *Fractals Everywhere*, Academic Press, San Diego, California.
Huang H, A Cortis & CE Puente (2012) Geometric harnessing of precipitation records, *SERRA* DOI 10.1007/s00477-012-0617-6.
Obregón N, CE Puente & B Sivakumar (2002) Modeling high resolution rain rates via a deterministic approach, *Fractals,* 10(3), 387–394.
Puente CE (1996) A new approach to hydrologic modeling: derived distributions revisited, *Journal of Hydrology*, 187, 65-80.
Puente CE, ML Maskey, A Cortis & B Sivakumar (2013) Searching for trends in inter-annual runoff records via a fractal geometric approach. AGU Fall.