

# CODIFICACIÓN, SIMULACIÓN, DESAGREGACIÓN Y PREDICCIÓN DE CAUDALES EMPLEANDO UN MODELO GEOMÉTRICO FRACTAL-MULTIFRACTAL

Carlos E. Puente, Mahesh L. Maskey, Bellie Sivakumar y Mario A. Diaz-Granados

Department of Land, Air and Water Resources, Universidad de California, Davis, Estados Unidos & Universidad de los Andes, Colombia.  
E-mail: cepuente@ucdavis.edu, mlmaskey@ucdavis.edu, s.bellie@unsw.edu.au, mdiazgra@uniandes.edu.co

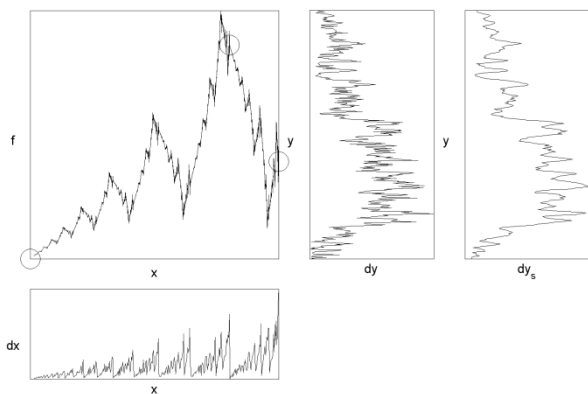
## Introducción

Este trabajo ilustra la aplicabilidad de un procedimiento determinista geométrico, el método Fractal-Multifractal, FM, introducido por Puente (1996, 2004), para codificar, simular, desagregar y predecir caudales. Específicamente, se muestra cómo es posible: (a) codificar datos observados a nivel diario y durante un año, requiriendo solamente un máximo de 9 parámetros geométricos, (b) simular registros diarios durante un año, preservando todo el histograma y toda la función de auto-correlación de una serie de tiempo dada, (c) desagregar fielmente datos mensuales, quinquenales y semanales a información diaria y (d) predecir tendencias de caudales cinco años hacia adelante, empleando redes neuronales para pronosticar los parámetros geométricos de codificaciones FM de registros observados, cada cinco años, en el pasado. Dicha gama de aplicaciones se exhibe aquí empleando información del río Sacramento en California, Estados Unidos. Para efectos de comparación, la presentación también mostrará resultados obtenidos utilizando datos del río Magdalena en Colombia.

## Métodos

El método empleado para codificar la información de caudales diarios, observados durante uno o más años, es una adaptación del procedimiento Fractal-Multifractal FM (Puente et al., 2018). Como lo ilustra la Figura 1, una medida *multifractal*,  $dx$  sobre  $x$ , le sirve de *entrada* a un *sistema* dado por una función *fractal*  $f$  definida de  $x$  a  $y$ , y él produce como *salida*, transfiriendo  $dx$  por medio de la función, una distribución *derivada*,  $dy$  sobre  $y$ . Al suavizar localmente dicha salida, se obtiene un conjunto  $dy_s$  sobre  $y$ , que, al interpretarse la coordenada  $y$  como el tiempo, posee una apariencia similar a la de caudales (encima de su flujo base y normalizados), a la escala requerida por el estudio.

Los objetos  $dx$  y  $dy$  son *histogramas* estables de puntos  $(x, y)$ , sobre  $x$  y  $y$ , compuestos por 365(366) compartimentos. Los puntos se hallan iterando funciones sencillas con componentes lineales del plano al plano, las cuales contraen el espacio y definen la geometría de un *atravante* único—la función *fractal* (Barnsley, 1986)—que pasa por ciertos puntos, cual indicado en el ejemplo de la Figura 1 por las circunferencias mostradas.



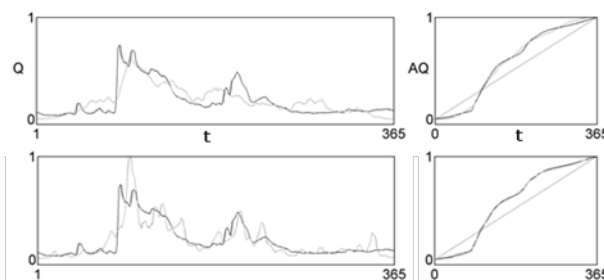
**Figura 1.-** Una adaptación del método FM: de una medida multifractal  $dx$ , a una medida derivada  $dy$ , mediante una función fractal de  $x$  a  $y$ . El objeto  $dy_s$  se halla suavizando  $dy$  sobre 365(366) valores.

La construcción empleada aquí requiere de dos o tres tales funciones y un máximo de 9 parámetros. Ellos corresponden a las coordenadas de los puntos que definen el atrayente, parámetros que escalan la coordenada y en cada una de las funciones y el sesgo de una “moneda” (de dos o tres lados) usada para guiar las iteraciones.

Para definir caudales—normalizados a una unidad—mediante el método FM, i.e. para hallar parámetros adecuados para series dadas, se empleó un algoritmo de optimización basado en un *enjambre de partículas* con miembros dinámicos (Fernández-Martínez et al., 2010). Mientras que para la codificación, desagregación y predicción de caudales se usó como función objetivo el error cuadrático medio, a nivel diario, de caudales acumulados durante el año (o cinco años) en cuestión, i.e. su *curva de masa*, para la simulación se utilizó también el error cuadrático medio, pero del histograma y la función de auto-correlación de datos a preservar. Los resultados mostrados a continuación corresponden a los mejores parámetros hallados, típicamente luego de 5 horas de CPU por año de datos, los cuales, dada la dimensión del espacio de búsqueda, no siempre corresponden a “óptimos globales”.

## Resultados

La Figura 2 muestra ejemplos típicos de codificaciones de caudales en el río Sacramento (del 1 de octubre al 30 de septiembre) empleando dos representaciones FM, una basada en la iteración de dos funciones (arriba) y otra que emplea tres (debajo), promediando localmente el histograma  $dy$  usando cinco compartimentos.



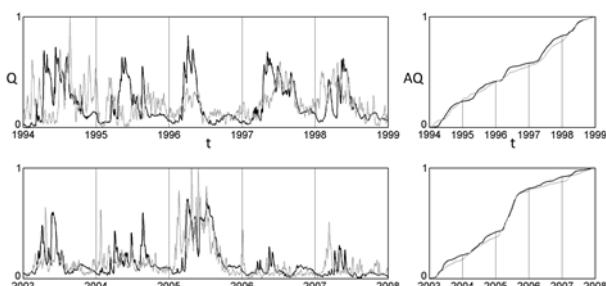
**Figura 2.-** Codificaciones FM típicas de caudales y curvas de masa en el río Sacramento iterando dos funciones (arriba) y tres (debajo).

Como se observa para el año 1965, con datos observados en negrita, aunque los ajustes FM muestran diferencias en textura y aproximan de diversa forma el primer pico, ellos siguen las tendencias observadas. Esto da lugar a una eficiencia de Nash-Sutcliffe comparando “caudales” que, para 65 años estudiados, supera un significativo 65% y a curvas de masa que aproximan fielmente lo observado, con errores, cuadrático medio y máximo, que no exceden un mero 2.1 y 4%. Basado en estos resultados se concluye que el método geométrico permite describir las series de caudales eficientemente, comprimiendo la información en un factor superior de 40:1 (Puente et al., 2018).

La secuencia de los parámetros FM puede emplearse para clasificar la geometría de los caudales de año en año (i.e. usando técnicas de agrupamiento), lo cual permite visualizar la dinámica de la información en función del tiempo. Dada la

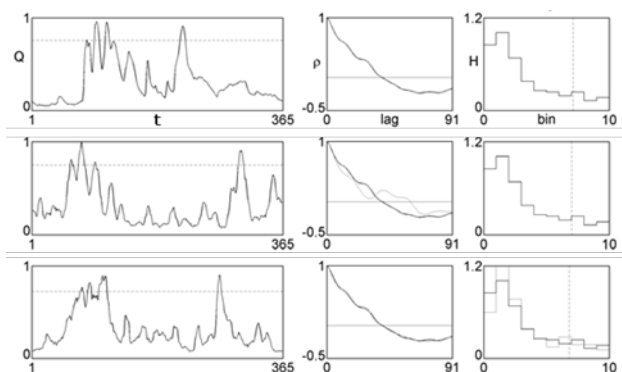
variabilidad intrínseca de la geometría observada en los caudales del río Sacramento, los parámetros FM hallados y sus implicadas clasificaciones exhiben una amplia variabilidad de año a año, y esto impide el poder obtener predicciones fieles de caudales futuros, un año hacia adelante (Puente et al., 2018).

Aunque esto es así, se ha establecido recientemente que si se codifican, no un año, sino cinco años de caudales seguidos, además de encontrarse resultados igualmente fidedignos y con un factor de compresión cinco veces mayor, se pueden, en efecto, hallar predicciones de parámetros FM futuros—ajustando redes neuronales basadas en comportamientos pasados—las cuales se traducen en útiles predicciones de la geometría del proceso (i.e. un año hacia adelante y conteniendo cinco años de caudales). Como se ilustra en la Figura 3, para ejemplos que corresponden a quinquenios que terminan en los años 1999 (arriba) y 2008 (debajo), las predicciones acumuladas (en gris) ajustan bien la distribución global de los datos (en negra). Para 21 casos considerados, finalizando con los cinco años que terminan en 2015, los errores cuadrático medio y máximo son de  $3.4\% \pm 0.7$  y  $8.8\% \pm 2.5$ , lo cual es notable.



**Figura 3.-** Cinco años seguidos de caudales, observados y predichos, y sus curvas de masa en el río Sacramento. Los parámetros FM futuros son calculados a partir del pasado usando redes neuronales.

El procedimiento geométrico también puede emplearse para simular registros que preservan características estadísticas relevantes, tales como el histograma y la función de auto-correlación de caudales, simplemente hallando parámetros FM que ajustan tal información. En dicho espíritu, la Figura 4 muestra una serie de tiempo (arriba) y dos simulaciones (debajo), una que ajusta fielmente el histograma de los registros y la otra su función de auto-correlación.

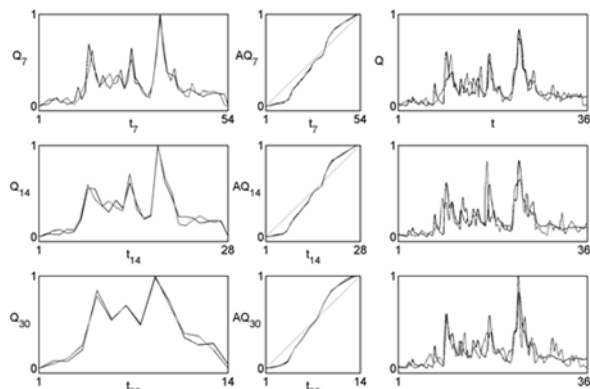


**Figura 4.-** Simulaciones FM que preservan el histograma (en medio) y la auto-correlación (debajo) de caudales en el río Sacramento (arriba).

Como se observa, las texturas de las simulaciones encontradas lucen, en efecto, similares a las del conjunto original y sus diversas geometrías representan escenarios alternativos que acaso pueden ser útiles en la práctica. Ciertamente, estas ideas generalizan lo que puede hacerse empleando técnicas clásicas de series de tiempo, las cuales, a lo sumo preservan, algunas características del histograma o de la función de auto-correlación, más no su totalidad, lo cual impide hallar texturas coherentes como las mostradas aquí (Maskey et al., 2017).

Finalmente, el procedimiento geométrico determinista también permite desagregar caudales, como sigue. Partiendo de información a una escala amplia, (i.e. acumulada en semanas, quincenas o meses), se optimiza el modelo FM para codificar dicha información ajustando la curva de masa a tal escala. Una vez encontrado un buen ajuste, algo que resulta ser más fácil dada la menor cantidad de sumandos en la función objetivo, se encuentra una desagregación estimada recalculando la única medida subyacente, o sea la salida del sistema, a la escala fina deseada (i.e. días) (Puente et al., 2018).

Tal y como se ilustra en la Figura 5, es posible obtener diversos conjuntos de caudales desagregados, los cuales provienen de ajustar la curva de masa de los datos a escalas amplias. Dichos conjuntos representan, nuevamente, escenarios alternativos que pueden emplearse en la modelación de recursos hidráulicos.



**Figura 5.-** Desagregaciones FM de caudales semanales, quinquenales y mensuales a diarios para un año de datos en el río Sacramento.

En resumen, el procedimiento determinista Fractal-Multifractal, FM, uno que además posee una interpretación física cual lo producido por una cascada multiplicativa generalizada (Cortis et al., 2013), puede emplearse para codificar, clasificar, simular, desagregar y también predecir caudales, tal y como se ilustró para el río Sacramento. Los resultados respaldan la noción que el modelo FM representa un lenguaje provechoso para estudiar la complejidad natural (Puente, 2004). Las ideas se ilustrarán empleando caudales medidos en el río Magdalena de Colombia.

## Referencias

- Barnsley, M.F. (1988). *Fractals Everywhere*, Academic Press, San Diego.
- Cortis, A., C.E. Puente, H.H. Huang, M.L. Maskey, B. Sivakumar & N. Obregón (2013). "A physical interpretation of the deterministic fractal-multifractal method as a realization of a generalized multiplicative cascade". *Stochastic Environmental Research and Risk Assessment*. Vol. 28, No. 6, pp. 421–429.
- Fernández-Martínez, J.L., E. García Gonzalo, J.P. Fernández Álvarez, H.A. Kuzma & C.O. Menéndez Pérez (2010). "PSO: A powerful algorithm to solve geophysical inverse problems Application to a 1D-DC resistivity case". *Journal of Applied Geophysics*, Vol. 71, No. 1, pp. 13–25.
- Maskey, M.L., C.E. Puente, B. Sivakumar & A. Cortis (2017) "Deterministic simulation of mildly intermittent hydrologic records". *Journal of Hydrologic Engineering*, Vol. 22, No. 8, doi.org/10.1061/(ASCE)HE.1943-5584.0001531.
- Puente, C.E. (1996). "A new approach to hydrologic modelling: derived distribution revisited". *Journal of Hydrology*, Vol. 187, pp. 65–80.
- Puente, C.E. (2004). "A universe of projections: may Plato be right?". *Chaos, Solitons and Fractals*, Vol. 19, No. 2, pp. 241–253.
- Puente, C.E., M.L. Maskey & B. Sivakumar (2018). "Combining fractals and multifractals to model geoscience records". En *Fractals: Concepts and Application in Geosciences*, Eds. B. Ghanbarian & A. Hunt, CRC Press, pp. 297–332.