

SIMULACIÓN NO LINEAL DE CAUDALES USANDO LA TRASFORMADA
EN ONDITAS

Julián David Rojo H., Luís Fernando Carvajal S.

*Posgrado en aprovechamiento de los recursos hidráulicos, Universidad Nacional de Colombia, Colombia
jdrojoh@unal.edu.co, lfcarvaj@unal.edu.co*

RESUMEN:

En este artículo se analiza la representación de la dinámica de una serie de caudales usando un modelo no lineal basado en la Transformada de Onditas. El modelo propuesto utiliza una combinación aleatoria de subseries obtenidas a partir de los modos de oscilación que componen la serie original mediante la transformada de onditas.

El método presentado se compara con los modelos estocásticos lineales a partir de la simulación de la serie de caudales medios mensuales de la estación RG8 sobre el río Grande en Colombia. Los resultados muestran la superioridad del modelo basado en la transformada en onditas para reproducir a largo plazo las principales características de la serie de tiempo analizada.

ABSTRACT:

In this paper a method based on the wavelet transform is used to simulate a non linear streamflow time series. The proposed method uses a periodic random combinatorial reconstruction to simulate the periodic hydrologic time series.

The presented method is compared with the linear stochastic models to simulate the monthly streamflow of the RG8 station on the Rio-Grande River in Colombia. The results show the superiority of the model based on the wavelet transformed to reproduce the long term main features of the time series analysis.

PALABRAS CLAVES: (1^{era}. hoja)

Transformada en onditas, Simulación de caudales, descomposición en frecuencias de oscilación, Hidrología estocástica.

INTRODUCCIÓN

Desde finales de los años 60 diferentes técnicas y modelos de simulación estocástica han sido sugeridos en la literatura para el análisis de series de tiempo hidrológicas y su aplicación en la generación de información que permita la planificación, operación y manejo de los recursos hidráulicos de una manera mas racional (Smith, 1983). Durante las últimas tres décadas el desarrollo de la simulación de series hidrológicas estuvo dominada por modelos auto-regresivos estocásticos cuyo supuestos, basados en una estructura de autocorrelación, proponen una relación de tipo lineal y Gaussiana a través de la persistencia de los caudales. Desde entonces, en la Hidrología se ha hecho un gran esfuerzo para tratar de mejorar los modelos inicialmente propuestos, a fin de desarrollar nuevos enfoques que se justifiquen físicamente y permitan estudiar su impacto en el planeamiento, manejo y diseño de sistemas de recursos hidráulicos.

En términos generales, todos los modelos propuestos para ser usados en la simulación de series hidrológicas tienen por objeto reproducir a largo plazo las principales características de las secuencias históricas, y aunque todos tienen sus meritos han sido criticados por los siguientes aspectos: por acomodar las series hidrológicas a supuestos estadísticos que no obedecen a la naturaleza de los caudales (estacionalidad, linealidad y normalidad), por no reproducir persistencias de largo y corto plazo, por la dificultad para estimar los parámetros, por la cantidad de parámetros necesarios, por las dificultades para generar muestras sintéticas de gran tamaño con pocos datos y por la ausencia de explicación física dentro de muchos de los modelos propuestos.

El problema de la simulación se vuelve aún mas complejo cuando se analiza físicamente la producción de escorrentía en las latitudes tropicales, por ejemplo en Colombia la variabilidad de la hidrología es altamente no lineal y ocurre en un amplio rango de escalas espacio-temporales (Poveda, 2004). Los ciclos decadales e interanuales (5-3 años) de la hidrología corresponden a fenómenos autorregulados de retroalimentación no lineal tales como la PDO, el sistema ENSO y la NAO entre otros; los ciclos anuales y semianuales responden a forzamientos atmosféricos de gran escala, entre otros, la migración latitudinal de la zona de convergencia intertropical (ZCIT) y la fluctuación e intensidad de transporte de humedad por los vientos alisios del este desde el mar Caribe, la cuenca amazónica y el Océano Pacífico. A escalas de tiempo menores la influencia de las oscilaciones Madden-Julian y las denominadas ondas del este controlan gran parte de los ciclos de precipitación en el país, y a escala diurna los grandes cambios de temperatura derivados del ciclo de insolación (también astronómico) representan un factor determinante en la formación, desarrollo e intensidad de la lluvia tropical. Físicamente los caudales de los ríos Colombianos deben ser tratados como un sistema dinámico, no lineal y no estacionario cuya modelación requiere del estudio y de la incorporación de los principales ciclos que afectan la producción de escorrentía en cuencas tropicales.

El presente estudio utilizará una técnica espectral conocida como la transformada en Onditas (Wavelet) para la generación de series sintéticas. Dicha técnica permite descomponer las series de caudales en subseries de diferentes periodos y a partir de dicho análisis se elige un periodo representativo que permita construir una combinación aleatoria de las señales para producir una serie sintética de caudales. Algunos estudios como el de Feng (2002) han demostrado las bondades y ventajas del método para la simulación de series periódicas y no lineales. El presente trabajo muestra inicialmente la teoría básica de la transformada en onditas y su aplicación a la simulación de series periódicas, luego se expone brevemente la teoría de los modelos estocásticos lineales y se presentan algunos criterios para la evaluación de las simulaciones con los diferentes modelos. Por último, las metodologías presentadas serán aplicadas a la serie de caudales de la estación RG8 sobre el río Grande en Colombia y al final se hace una comparación de los resultados.

LA TRASFORMADA EN ONDITAS

La transformada en onditas (Wavelet) es un tipo especial de transformada de Fourier que representa una señal en términos de versiones trasladadas y dilatadas de una onda finita denominada onda madre. El término original francés es ondelette, introducido por Jean Morlet y Alex Grossmann, ha sido traducido al inglés como wavelet, y también al castellano como, óndulas y onditas.

Una ondita es una función que satisface ciertas condiciones matemáticas pudiendo representar series de tiempo y otro tipo de funciones; todas las onditas están basadas en una función llamada Onda Madre u onda de análisis $\psi(t)$ la cual posee unas propiedades especiales y satisface la condición $\psi(0) = 0$. Bajo dicha condición la transformada en onditas para una señal unidimensional $x(t)$ es definida como:

$$W(\tau, s) = \int_{-\infty}^{\infty} x(t) \frac{1}{\sqrt{|s|}} \psi^* \left(\frac{\tau - t}{s} \right) dt, \quad \tau \in \mathfrak{R} \wedge s \in \mathfrak{R} - 0 \quad [1]$$

Donde $W(\tau, s)$ es el coeficiente de onda; las series de la señal descompuesta, $\psi^*(t)$ es la conjugada compleja de $\psi(t)$; τ es un parámetro de posición y s es el parámetro de escala. La ecuación inversa de [1], o reconstrucción de $x(t)$ puede ser representada como:

$$x(t) = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W(\tau, s) \frac{1}{s^2} \psi \left(\frac{\tau - t}{s} \right) d\tau ds \quad [2]$$

Siendo C_ψ una constante de admisibilidad la cual depende de la ondita usada y satisface la condición de admisibilidad:

$$C_\psi = \int_{-\infty}^{\infty} \frac{|\mathcal{F}(\psi)|}{|\omega|} d\omega \leq \infty \quad [3]$$

Donde $\mathcal{F}(\psi)$ representa la transformada de Fourier de $\psi(t)$; En la práctica es necesario discretizar la ondita y sus transformaciones para poderlas aplicar a las series de tiempo, Mediante discretización la ecuación [2] se representa como:

$$x(t) = \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} \psi_{jk} \psi_{jk} \quad [4]$$

Con ψ_{jk} como coeficiente de ondita $\psi_{jk} = W(\tau, s) = W(k, 2^j)$ y

$$\psi_{jk} = \frac{1}{\sqrt{2^j}} \psi \left(\frac{t - 2^j k}{2^j} \right) \quad [5]$$

El algoritmo empleado para la multi-resolución, y reconstrucción de las series de tiempo usando la transformada en onditas de onditas fue presentado por Mallat (1989).

Existen diversas familias de onditas cuyas cualidades varían según múltiples criterios, sin embargo todas cumplen las siguientes restricciones:

Las ondas $\psi(t)$ convergen a cero cuando el tiempo o la frecuencia tiende a infinito. Las ondas poseen algún tipo de simetría, ya sea con alguno de los ejes, con el origen o con un eje definido. Además poseen una amplia regularidad que les permite conservar sus propiedades a lo largo del tiempo.

- Una onda $\psi(t)$ converge a cero cuando el tiempo o la frecuencia tiende a infinito.
- Las ondas poseen algún tipo de simetría, ya sea con alguno de los ejes, con el origen o con un eje definido.
- Las ondas poseen una amplia regularidad.

Típicamente existe una cantidad importante de ondas propuestas y generalmente son agrupadas según sus propiedades y características principales. Muchas de las familias de ondas reciben sus nombres de acuerdo a los autores que propusieron su formulación, algunos ejemplos de ondas se presentan en la Figura 1.

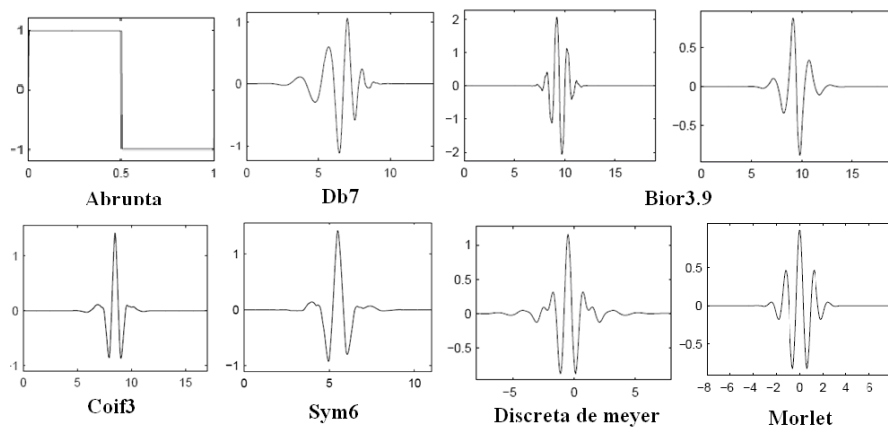


Figura 1.- Algunos tipos de ondas

SIMULACIÓN DE SERIES PERIÓDICAS USANDO LA TRANSFORMADA EN ONDITAS

La serie original puede ser descompuesta en varias componentes mediante el análisis armónico provisto por la transformada en ondas y para abordar el problema de simulación la pregunta básica es como reconstruir la serie luego de la descomposición, una posible aproximación para resolver el problema fue propuesta por Feng (2002) y es la siguiente:

- Elegir una onda madre.
- Se descompone la serie periódica en sus armónicos mediante la transformada en ondas.
- Se divide las series obtenidas de la descomposición en pequeñas subseries (o tramos) cuyo tamaño sea de un periodo definido.
- Se utiliza una combinación aleatoria de tramos de para generar la serie sintética.

El método de combinación aleatoria de los tramos de cada subserie es el usado en el presente trabajo para la simulación. La Figura 2 ilustra el concepto con el cual la serie puede ser reconstruida por descomposición de señales utilizando una combinación arbitraria de periodos.

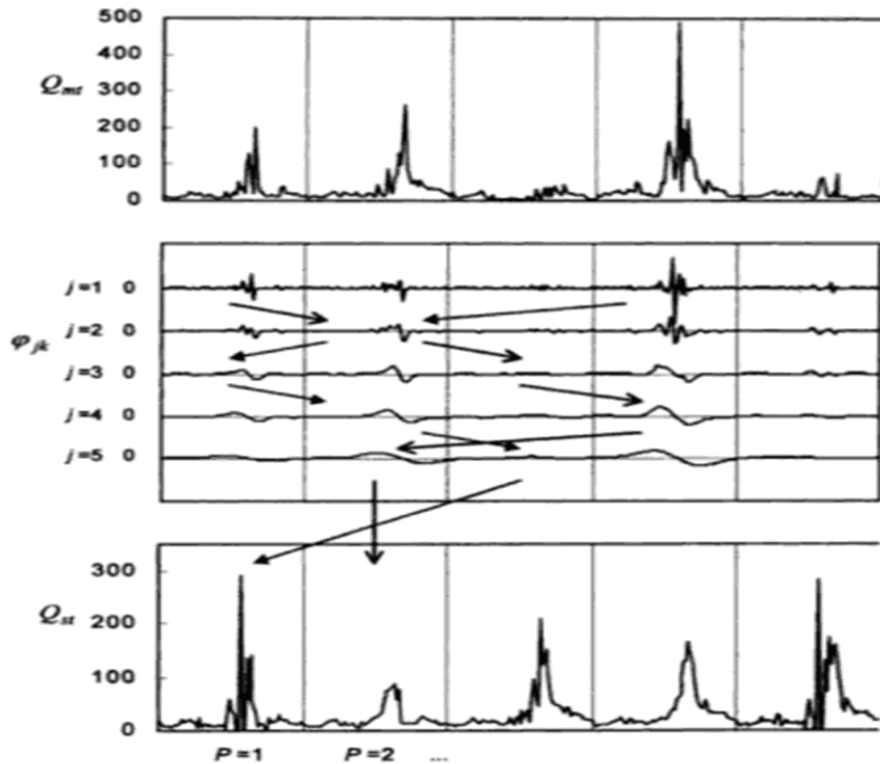


Figura 2.- Simulación de series utilizando la descomposición en onditas (Feng, 2002)

Así entonces $z(t=1,2,\dots,n)$ es una serie con P periodos significantes de longitud T cada periodo, por ejemplo, $T_j (j=1,2,\dots,P)$. La longitud de cada periodo T puede considerarse como el periodo predominante de la serie que depende de los intervalos de tiempo, el número de estaciones en el año y la relación de los caudales con el macro-clima. Para el caso de series mensuales podría ser $T=12$. Si se desea descomponer la serie original en j armónicos usando la transformada de onditas entonces la longitud de la serie simulada podrá tener un tamaño máximo de N datos, donde N está dado por:

$$N = TP^j - n \quad [5]$$

Y tendrá un número de periodos simulados

$$M = \frac{N}{T} \quad [6]$$

MODELOS ESTOCÁSTICOS USADOS EN LA SIMULACIÓN DE SERIES HIDROLÓGICAS

Con el objeto de comparar el modelo propuesto con los esquemas comúnmente usados para la simulación de caudales, se presenta a continuación una breve descripción de los modelos estocásticos lineales tradicionalmente usados para la modelación hidrológica. El lector podrá encontrar mayor información en los trabajos de Box & Jenkins (1976); Peña (1994); Box et al. (1994).

Un proceso estocástico, $x(t, \omega); t \in I, \omega \in \Omega$ es una función de dos variables t, ω . Para un momento $t = t^*$ específico, se tendrá $x(\omega)$ es una función solo de ω , por tanto se comporta como una variable aleatoria. En el caso de fijar ω , $x(t)$ se convierte en una función del tiempo y se denomina una "trayectoria muestral" de un proceso estocástico o serie de tiempo. Un proceso estocástico es

En 1938 Wold demostró que un proceso estocástico débilmente estacionario puede escribirse como una función lineal de variables aleatorias no correlacionadas:

$$Z_t = \varepsilon_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad [9]$$

Donde los ε_t corresponden a variables aleatorias no correlacionadas y Z_t tiene distribución normal. De esta forma, la estacionalidad débil coincide con la estricta y los parámetros de este modelo corresponden a los representados por θ_k y se obtienen mediante el sistema:

$$\begin{aligned} \theta_0 &= (1 + \theta_1^2 + \dots + \theta_q^2)^{-1/2} \\ \theta_k &= (-\theta_k + \theta_1 \theta_{k+1} + \dots + \theta_{q-k} \theta_q) \theta_0^{-1} \quad \text{para } k = 1, 2, \dots, q \\ \theta_k &= 0 \quad \text{para } k > q \end{aligned} \quad [10]$$

Donde los θ_i corresponden a los coeficientes de correlación parcial de rezago i . El valor del argumento q representa el máximo rezago hasta el cual el coeficiente de autocorrelación parcial es significativamente distinto de cero y el modelo resultante se considera un modelo de media móvil (MA).

Modelo ARMA(p,q):

Los modelos ARMA son modelos estocásticos que combinan los modelos autorregresivos (AR) con los modelos de media móvil (MA). La forma general de estos modelos está dada por:

$$Z_t = \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad [11]$$

Donde los parámetros ϕ_i , θ_i y ε_t tienen el mismo significado que para los modelos AR(p) y MA(q).

Normalización y homogenización de series hidrológicas:

La modelación de series hidrológicas usando la teoría lineal de variables aleatorias implica que se asuman como ciertas las hipótesis de normalidad y estacionalidad de la serie de tiempo, lo cual no se cumple en las series de caudales. Las hipótesis asumidas obligan a adoptar unas transformaciones que distorsionan el comportamiento original de las series para cumplir con los rigores estadísticos del método.

Para cumplir con la condición de estacionalidad en sentido estricto se propone una transformación que remueve el coeficiente de asimetría de la serie para ajustar los datos a la distribución normal utilizando la siguiente expresión:

$$X_v = \log(Q_v - c \bar{Q}) \quad [15]$$

Donde

$$c_{\bar{N}} = \frac{a}{g_{\bar{N}}^2} \quad [16]$$

donde $Q_{v?}$ es el caudal medio para el mes $?^*$ ($= 1, \dots, 12$) y año v ($v=1, \dots, N_a$) siendo N_a el número de años de la series; $\bar{Q}_?$ es el caudal medio mensual del mes $?$ y a es un parámetro adimensional cuyo valor es de 0.35 (Ochoa, 2002) el cual resulta de un análisis de regresión entre $g?$ y $c?$; $g?$ es el coeficiente de asimetría para el conjunto $Q_{1?}, Q_{s?}, \dots, Q_{Na?}$; y $X_{v?}$ es la serie normalizada de los caudales para el año v y el mes $?$.

Cuando la serie de datos no es estacionaria se recurre a la diferenciación y a las transformaciones Box-Cox (Peña, 1992) para eliminar los cambios graduales y/o abruptos de los diferentes momentos de la serie hidrológica.

CRITERIOS PARA LA EVALUACIÓN DE LAS SIMULACIONES

Sobre series estandarizadas (con media cero y varianza 1) de caudales simulados se utilizarán los siguientes criterios para evaluar las capacidades del modelo propuesto:

Preservación de los momentos estadísticos

Se espera que las series simuladas reproduzcan a largo plazo las principales características estadísticas de la serie de datos modelada, entre ellas la media, la varianza, la asimetría, y la kurtosis.

Reproducción de eventos extremos

En el estudio de series de lluvias y caudales es de especial interés el análisis de las características relacionadas con las sequías y los excesos, para tal efecto se utiliza el concepto de excursiones del proceso aleatorio por encima y por debajo de un valor límite. Los desarrollos teóricos reciben el nombre de teoría de cruces (Salazar, 1994, Bras 1985). Los parámetros se conocen como excursiones o cruces positivos y excursiones o cruces negativos. Se determinan el número y la longitud de las excursiones por encima ($N+$, $L+$) y por debajo ($N-$, $L-$) de un valor de referencia además del área por encima ($A+$) y por debajo ($A-$). Para series estandarizadas con media cero y varianza 1 el valor de referencia o umbral para medir los cruces es $u=1$.

Almacenamiento y persistencia

Otra característica de interés de las series de caudales es aquella relacionada con el almacenamiento de embalses a largo plazo. Los primeros análisis fueron desarrollados por Hurst (1951) empleando un estadístico conocido con el nombre de rango reajustado a las desviaciones acumuladas o rango de fluctuación, el cual se define como:

$$R^*(n) = \text{Sup}S^*(i) - \text{Inf}S^*(j) \quad i \in (0, 1, 2, \dots, n) \quad j \in (0, 1, 2, \dots, n) \quad [17]$$

Donde Sup e Inf corresponden al mayor valor positivo y el menor valor negativo respectivamente de la función S^* definida por:

$$S^*(k) = \sum_{i=1}^k X_i - \frac{k}{n} \sum_{i=1}^n X_i \quad [18]$$

El rango de fluctuación puede considerarse como la capacidad que un embalse debería tener para suplir una demanda igual a la media sin incurrir en vertimientos o déficit durante un periodo de n años, por ello es un indicativo de la intensidad y duración de las sequías que presenta la serie de caudales. En sus investigaciones Hurst (1957) calculó el valor de $R^*(n)$ para múltiples series de los mas diversos tamaños y cada valor de $R^*(n)$ lo dividió por la desviación típica $S(n)$ de la serie de n observaciones, obteniendo el estadístico $R^*(n)/S(n)$ conocido con el nombre del rango ajustado reescalado. Hurst encontró que este estadístico variaba con el tamaño de la muestra n como:

$$\frac{R^*(n)}{S(n)} = Cn^h \quad [19]$$

El hecho de que el exponente h no sea igual a 0.5 es lo que se conoce como fenómeno de Hurst y dicho exponente se le relaciona con la persistencia en la serie de la siguiente manera: Las series con un $h > 0.5$ son series persistentes donde la auto dependencia juega un rol importante en el proceso, así por ejemplo $0.5 < h < 0.7$ implica persistencia de corto plazo y $h > 0.7$ implica persistencia de largo plazo. $h = 0.5$ Correspondería a series con distribución normal no persistentes, lo que equivale a un ruido blanco, y $h < 0.5$ corresponde a series antipersistentes.

También se utilizará como criterio para evaluar la persistencia en las series simuladas la reproducibilidad del coeficiente de auto-correlación de rezago 1 y las principales frecuencias de la serie original.

Las series simuladas deben poseer un tamaño igual al de la serie histórica para que parámetros como las excusiones y rango reescalado puedan ser comprobables.

DATOS Y RESULTADOS

Descripción de los datos

Los modelos expuestos en el presente trabajo son aplicados a la serie de caudales de la estación RG8 sobre el río Grande de Empresas Públicas de Medellín. Dicha estación cuenta con registros desde el año 1942 a 2008 (67 años) y puede considerarse como una de las mejores series de caudales del país en cuanto a calidad de la información y extensión del periodo de registro (Salazar 1994).

Para remover el ciclo anual se estandarizó la serie restando a los datos las medias y dividiendo por las desviaciones estándar mensuales, de esta forma se obtiene una serie adimensional de media cero y varianza unitaria que se presenta en la en la Figura 3. Los principales parámetros estadísticos de la serie estandarizada son mostrados en la Tabla 1 y los resultados de varias pruebas para el análisis de normalidad y estacionalidad en la serie estandarizada con un nivel de significancia del 5% se exponen en la Tabla 2. Estas pruebas pueden ser consultadas en los trabajos de Chernoff S. and Hodges L. (1954), Mann, H. B., & Whitney D. R. (1947), Ansari A. R. & Bradley R. A. (1960).

Simulación de la serie RG8 usando la trasformada en onditas

A fin de aplicar la metodología basada en la transformada en onditas para simular la serie RG8 se evaluaron diferentes tipos de onditas siendo la ondita discreta de Meyer la de mejores resultados en la simulación. Un esquema de la ondita de Meyer fue presentado en la Figura 1 y las ecuaciones que describen su forma son las siguientes:

$$\varphi(\omega) = (2\gamma)^{-1/2} e^{i\omega/2} \operatorname{sen}\left(\frac{\gamma e}{2} \sqrt{\frac{3}{2\gamma} |\omega| - 1}\right) \Rightarrow \frac{2\gamma e}{3} \leq |\omega| \leq \frac{4\gamma e}{3} \quad [20]$$

$$\varphi(\omega) = (2\gamma)^{-1/2} e^{i\omega/2} \operatorname{sen}\left(\frac{\gamma E}{2} \sqrt{\frac{3}{4\gamma^{1/2}} |\omega| - 1}\right) \Rightarrow \frac{4\gamma E}{3} \leq |\omega| \leq \frac{8\gamma E}{3} \quad [21]$$

$$\varphi(\omega) = 0 \Rightarrow |\omega| \notin \left[\frac{2\gamma}{3}, \frac{8\gamma}{3}\right] \quad [22]$$

Donde $\nu(a) = a^4(35 - 84a + 70a^2 - 20a^3)$.

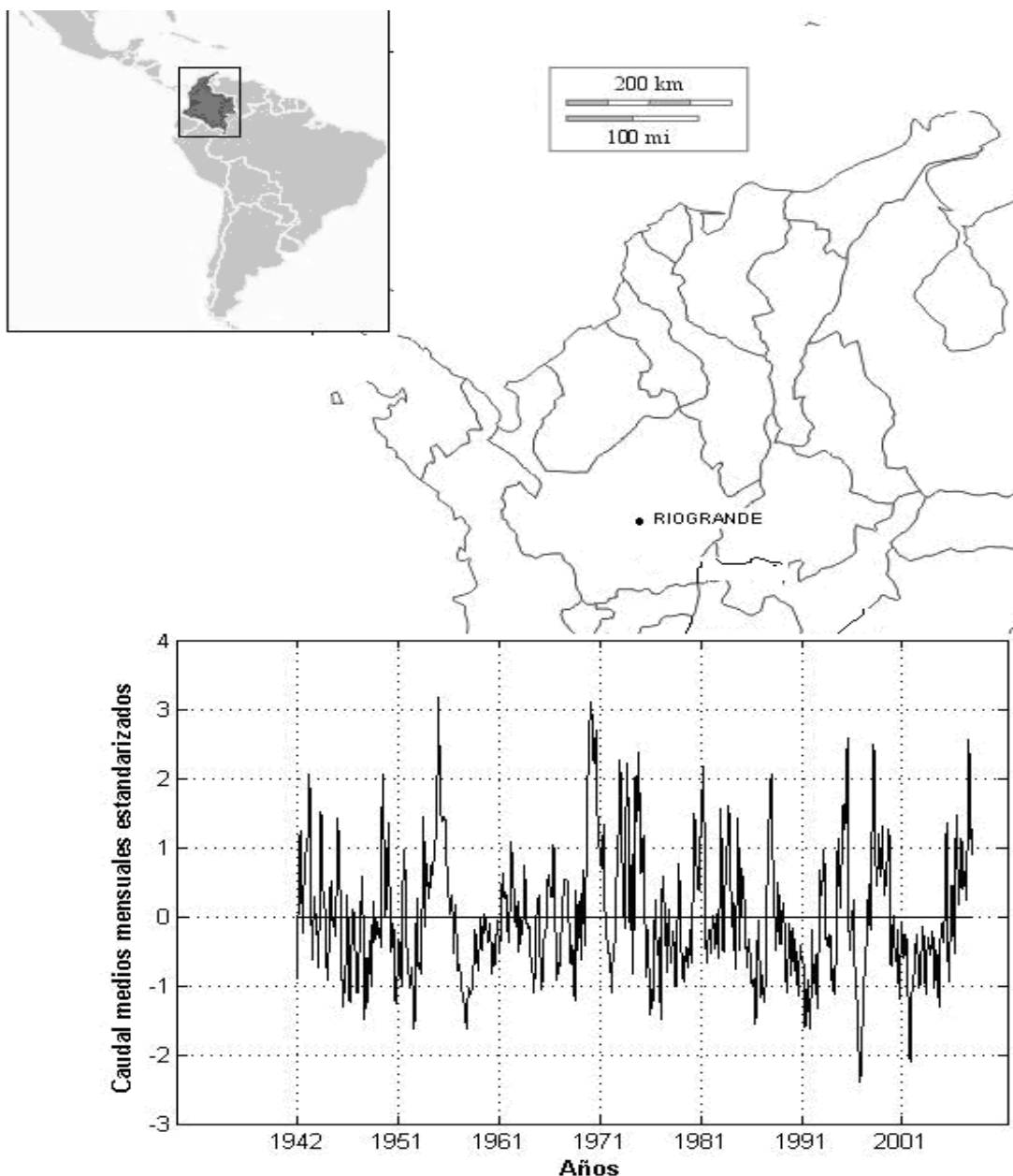


Figura 3.- Serie de caudales medios mensuales estandarizados de la estación RG8 (río Grande)

Tabla 1.- Parámetros estadísticos de la serie RG8 estandarizada

Parámetro	Valor
Media	0.00
Varianza	1.00
Asimetría	0.76
Kurtosis	3.67

Tabla 2.- Análisis de Normalidad y Estacionalidad de la serie RG8 estandarizada

Test estadístico	valor crítico	valor estimado	conclusión
Test de Normalidad Chi cuadrado de Pearson	28	50.11	La serie no es normal
Test de cambio en la media Mann-Whitney	1.96	3.66	Existe cambio en la media
Test de cambio en la varianza de Ansari-Bradley	3.75	7.58	Existe cambio en la varianza

DESCOMPOSICIÓN EN ONDITAS PARA LA ESTACIÓN RG8

ONDITA DISCRETA DE MEYER

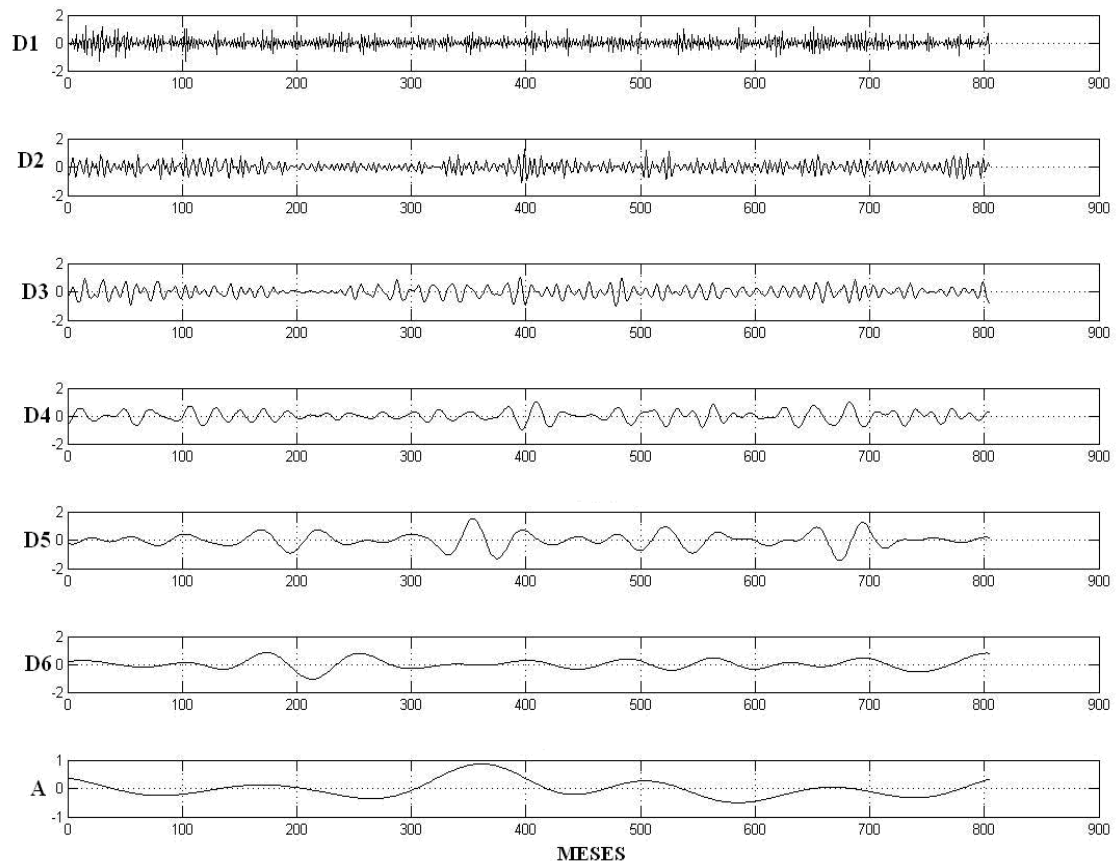


Figura 4.- Descomposición en onditas de la serie RG8 (río Grande) usando la ondita discreta de Meyer

La descomposición en onditas usando la ondita discreta de Meyer, para la estación RG8, se presenta en la figura 4. La serie de caudales estandarizados de la estación RG8 fue descompuesta en 7 armónicos (D1,..., D6, A) y para la combinación aleatoria de periodos cada armónico fue fraccionado en tramos de 12 meses. Se generaron 100 series sintéticas de 67 años a nivel mensual utilizando la metodología basada en la trasformada en onditas. Para cada serie simulada se estimaron los criterios de comparación de los resultados y el valor promedio de tales criterios para las 100 series simuladas se presenta en la Tabla 3.

Simulación de la serie RG8 usando los métodos tradicionales

A la serie RG8 le fueron removidos los diferentes cambios en la media y la varianza con el objeto de obtener una serie homogénea y estacionaria; para garantizar estacionariedad en sentido estricto la serie fue normalizada y estandarizada utilizando la transformación [15]; finalmente se estandarizó la serie transformada restando a cada dato la media y dividiendo por la desviación estándar de cada mes. El autocorrelograma parcial de la serie RG8 normalizada y estandarizada se presenta en la Figura 5, en la cual se observa que los dos primeros rezagos representan significativamente la estructura de auto-correlación de la serie.

Los datos transformados de la estación RG8 son estacionarios en sentido estricto y por lo tanto no requieren ser diferenciados. Se propone entonces simular la serie utilizando los modelos AR(2) y ARMA(2,1). Se simularon 100 series de 67 años y en cada una de ellas se aplican las transformaciones inversas respectivas antes de hacer una evaluación de los resultados; para cada una de las 100 series simuladas se estimaron los criterios de evaluación y el valor promedio de las 100 series simuladas se presenta en la Tabla 3.

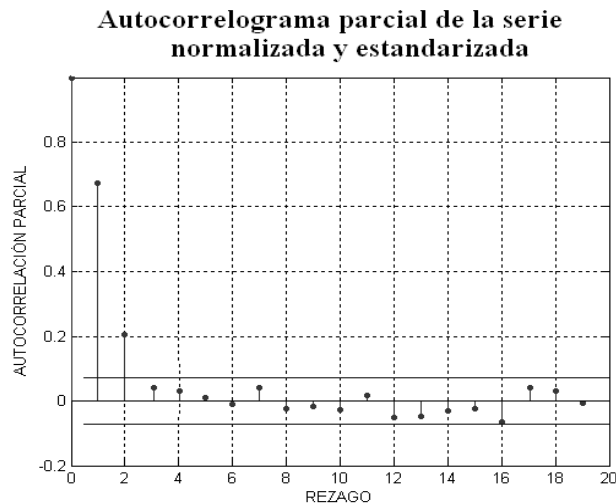


Figura 5.- Autocorrelograma parcial de los datos normalizados y estandarizados para la estación RG8

Tabla 3.-

	HISTÓRICA	ONDITAS	AR(2)	ARMA(2,1)
Criterios estadísticos				
Media	0	0	-0.02	-0.05
Desviación estándar	0.99	0.99	1.02	0.98
Asimetría	0.76	0.69	0.58	0.53
Kurtosis	3.67	3.48	3.02	3.30
Criterios relacionados con eventos extremos				
Número de excursiones positivas (NL+)	58	56.2	54	53
Longitud de excursiones Positivas (L+)	2.0	2.20	2.33	2.06
Área de excursiones positivas (A+)	3.53	3.51	3.56	3.13
Número de excursiones Negativas (NL-)	60	65.0	53	71
Longitud de excursiones Negativas (L-)	1.72	1.49	2.28	1.92
Área de excursiones Negativas (A-)	-2.28	-2.06	-3.54	-2.9
Criterios relacionados con la Persistencia				
Rango reescalado	72.32	72.50	67.38	78.80
Parámetro de Hurst (h)	0.75	0.74	0.72	0.78
Autocorrelación de rezago 1	0.67	0.65	0.7	0.63
Frecuencia dominante	0.025	0.023	0.015	0.013

Análisis de resultados

Los resultados expuesto en la Tabla 3 indican que las series de caudales simuladas por el método basado mediante la trasformada en onditas pueden preservar con mayor fidelidad las principales características estadísticas de la serie de tiempo en comparación con los modelos autoregresivos tradicionales. No existe una diferencia significativa entre la media y la varianza de los datos simulados usando la trasformada de onditas en comparación con los datos históricos; la asimetría se reproduce mucho mejor con el modelo basado en onditas pues el valor de los datos simulados utilizando dicho método es de 0.69, que es mas cercano al 0.76 de la serie histórica comparado con el 0.58 del AR(2) y el 0.53 del ARMA(2,1). Por ultimo, es de resaltar la capacidad del modelo basado en onditas para acercarse al cuarto momento de la serie histórica (la kurtosis), donde el valor arrojado por onditas es de 3.48 muy cercano al 3.67 de los datos observados.

El método basado en la trasformada en onditas reproduce con mayor fidelidad la cantidad de excursiones positivas y negativas cuando se simulan series generan serie con tamaño equivalente al de la serie histórica. 58 excursiones positivas tiene la serie histórica, 56 tienen en promedio las series simuladas con onditas, 54 la series simuladas con AR(2) y 53 las simuladas son ARMA(2,1), lo anterior significa que el modelo basado en la trasformada en onditas representa con mayor exactitud los periodos de excesos de caudal en el río Grande. Igual ocurre con las excursiones negativas donde el modelo a base de onditas tiene en promedio 65 excursiones negativas en comparación con las 60 observadas en la serie histórica, AR(2) tiene 53 y ARMA(2,1) 71, lo que se puede inferir una mayor capacidad de las series simuladas con la trasformada en onditas de representar eventos extremos como sequías en términos de su grado de ocurrencia e intensidad (véase la longitud de excursiones negativas).

Dada la similitud entre el rango reescalado de la serie histórica y el rango promedio de las series simuladas, de 72.32 y 72.50 respectivamente, se puede inferir que el método basado en la trasformada en onditas poseen mayor capacidad para simular eventos como excesos sequías en comparación con los modelos AR(2) y ARMA(2,1), lo que confirma las observaciones hechas con la teoría de cruces. Al respecto del parámetro de Hurst todos los modelos presentan una buena capacidad para representar la persistencia de largo plazo dados que el exponente h es mayor que 0.7 para todos los modelos. Sin embargo existe una mayor similitud entre h de los datos históricos y el obtenido por simulación usando la trasformada en onditas se puede argumentar que Este ultimo posee mas competencia para reproducir la persistencia de los datos observados. Los resultados de coeficiente de autocorrelación de rezago 1 y de frecuencias dominantes demuestran la idoneidad del modelo fundamentado en la trasformada en onditas a la hora de reproducir la autodependencia y los principales ciclos de la serie histórica.

Conclusiones

El modelo basado en la trasformada en onditas reproduce con mayor idoneidad las principales características estadísticas los datos históricos observados.

El modelo propuesto puede ser aplicado indistintamente en series estacionarias y no estacionarias, además posee la gran ventaja de requerir la estimación de parámetros.

La preservación de la media y varianza de la serie usando el modelo basado en la trasformada en onditas es casi perfecta y no se requiere distorsionar la serie normalizando los datos o removiendo los cambios en la media y la varianza, lo que representa una enorme ventaja en comparación con los métodos tradicionales.

El modelo propuesto reproduce la asimetría con mayor fidelidad que los métodos tradicionales AR(2) y ARMA(2,1). Lo que indica que a pesar de las transformaciones aplicadas, los modelos auto regresivos son muy dependientes de la condición de normalidad para operar con eficiencia.

La simulación de series periódicas usando la combinación aleatoria de la serie descompuesta mediante la transformada en ondas intenta representar el hecho de que las series de caudales son físicamente un sistema dinámico, no lineal y no estacionario producto de la superposición de los ciclos climáticos que afectan la producción de escorrentía en las cuencas tropicales.

Recomendaciones

La cantidad de armónicos elegidos para descomponer la serie de tiempo mediante la transformada en ondas puede variar de acuerdo a las condiciones específicas de cada serie a simular, en sus estudios FENA (2002) advierte que es imposible generar series sintéticas usando un solo nivel de descomposición ($j=1$), pues no existiría la combinación aleatoria de subtramos, sin embargo cuando se tienen dos ($j=2$), el modelo funciona perfectamente. Lo ideal sería escoger una cantidad de armónicos que puedan relacionarse físicamente con la serie histórica, ciclos decadales, de cinco años, de tres años, de un año pueden relacionarse pueden asociarse a eventos de la PDO, el ENSO, etc., por lo que el número de armónicos utilizados en la descomposición debe escogerse con al menos un mínimo conocimiento sobre la climatología de la zona de estudio.

El periodo T elegido para fraccionar los armónicos y crear los subtramos también debe ser elegido pensando en el macroclima y los periodos dominantes de la serie de tiempo, $T=12$ es un buen periodo para serie de tiempo mensuales, sin embargo pueden mejorarse los estimativos de criterios relacionados con la persistencia en la serie de tiempo si se toman valores mayores de T como 24, 36, 60 y hasta 120 meses, todo depende de la longitud del periodo de registro en los datos históricos.

Algunos estudios elaborados por el Posgrado en Aprovechamiento de los Recursos Hidráulicos (PARH) de la universidad Nacional de Colombia han demostrado la eficiencia del modelo cuando se cuenta con pocos datos históricos para generar serie sintéticas (ver [5] y [6]) lo que resulta muy útil a la hora de evaluar proyectos donde la información hidrológica es escasa.

Referencias Bibliográficas

- ANSARI A. R., BRADLEY R. A. (1960). "Rank sum tests for dispersion". *Annals of Mathematical Statistics*, 31, 1174-1189.
- Box, G.E.P., Jenkins, G.C., (1976). "Time Series Analysis: Forecasting and Control". Holden-day Inc., Calif, US.
- Box, G.E.P., G.M. Jenkins y G.C. Reinsel. (1994). "Time Series Analysis: Forecasting and Control". Prentice-Hall, Englewood Cliffs, New Jersey.
- Chernoff, H.; Lehmann E.L. (1954). "The use of maximum likelihood estimates in χ^2 tests for goodness-of-fit". *The Annals of Mathematical Statistics* 25: 579-586
- Feng, G. (2002). "A method of simulation of periodic hydrological time series using wavelet transforms". *Hydrological models for environment management*, 75-90.
- Hurst H. E. (1951). "Long Term Storage Capacity of Reservoirs" *Trans. Am. Soc. Civ. Eng.*, Vol. 116, paper 2447, pp: 770-808.

- Hurst H. E. (1957). "A Suggested Statistical Model for some Time Series Which Occurs in Nature", *Nature (London)*, Vol. 180,, pp. 494.
- Mallat, S.(1989) "Multiresolution approximation and wavelet orthonormal bases of L_2 ," *Trans. Amer. Math. Soc.*, June
- Mann, H. B., & Whitney, D. R. (1947). "On a test of whether one of two random variables is stochastically larger than the other". *Annals of Mathematical Statistics*, 18, 50–60
- Peña, D. (1994). "Estadística, Modelos y Métodos 2: Modelos Lineales y Series Temporales". *Alianza Universidad Textos*, Madrid.
- Poveda, G. (2004). "La Hidroclimatología de Colombia: una síntesis desde la escala inter-decadal hasta la escala diurna". *Rev. Acad. Colomb. Cien.*, 28 (107): 201-219.
- Smith, R. (1983). "Simulación de caudales con larga memoria", Experiences in the development and Application of Mathematical Models in Hydrology and Water Resources in Latin America (Proceedings of the Tegucigalpa Hydromath Symposium, September 1983). *IAHSPubl.No.152*. 267-287.
- .