

APLICACION DE DOS MODELOS NO LINEALES A LA SIMULACION DE SERIES HIDROLOGICAS

José E. Salazar Velásquez, Oscar J. Mesa Sánchez

Posgrado en Aprovechamiento de Recursos Hidráulicos,
Facultad de Minas, Universidad Nacional, Sede Medellín¹

RESUMEN

Los modelos autorregresivos lineales han sido ampliamente utilizados en el análisis de las series caudales. Sin embargo, la estrecha relación entre la hidrología colombiana y los procesos climáticos globales que son no lineales, justifican el estudio de otros modelos alternativos que reproduzcan algunas características no lineales. Con este fin se estudian dos modelos: *Ecuación diferencial estocástica con dos estados estables (EDE)*, un modelo markoviano continuo no lineal en el cual se considera que la variable de interés se mueve entre dos estados estables, y *Modelo Autorregresivo dependiente del régimen (RAR)*, combinación de varios modelos autorregresivos lineales que se seleccionan dependiendo de una serie índice que determina las características estadísticas de la variable.

Los modelos se estudian en cuanto a su estructura, propiedades y manera de estimar sus parámetros y se aplican a la serie de caudales mensuales de la estación RG8 (Riogrande) de las Empresas Públicas de Medellín, y a otras series hidrológicas. Los resultados indican que los dos modelos reproducen aceptablemente las principales características estadísticas de las series de lluvias y caudales incluyendo estadísticos de eventos extremos. En ambos se obtienen valores del rango ajustado más cercanos a las observaciones y mayores que los encontrados utilizando el método autorregresivo lineal usual con transformación logarítmica a los datos. Estas características de los modelos son particularmente importantes para los estudios de planeación de almacenamientos de agua.

1 Este trabajo fue posible gracias al patrocinio de las Empresas Públicas de Medellín.

ABSTRACT

Linear autoregressive models have been used traditionally in stream-flow simulation. But there is a close relation between colombian hydrology and the global ocean-atmosphere interaction in the tropics which is known to be non-linear. This interaction justifies a search for other models that may incorporate this observed non-linear behavior. To that end two models are considered: *Non-linear Stochastic Differential Equations* with a potential formed by two stable wells separated by an unstable state. The two stable states represent the two states of the El Niño-Southern Oscillation (ENSO). The other model is called *Regime dependent Autoregressive (RAR) Model* that combines different linear autoregressive components selectively according to either an external time series or the process itself.

This work undertakes the study of the structure, properties, parameter estimation and application of those two models to the Riogrande river in northwestern Colombia. Both models reproduce the same statistic parameters of the observed time series that are normally reproduced with the traditional linear autoregressive models. In addition, both models have a better representation of the statistics of extremes. The improvement is particularly important for the planning and operation of reservoirs.

1. INTRODUCCION

El intenso racionamiento de energía eléctrica ocurrido en el año 1992, que tuvo como una de sus principales causas la ocurrencia de un período anormalmente seco; sumado a una serie de investigaciones que muestran la estrecha relación entre la climatología colombiana y variables climáticas globales que exhiben un comportamiento no lineal, han despertado interés sobre la capacidad de los diferentes modelos para representar adecuadamente las series hidrológicas.

En Colombia, los modelos autorregresivos lineales han sido los más utilizados para el estudio de series temporales de lluvias y caudales. Sin embargo, estas series son no lineales, y los modelos no reproducen ciertos comportamientos propios de los datos históricos que son de gran importancia para el aprovechamiento del recurso hidráulico.

La naturaleza no es lineal y el uso de modelos no lineales en diversas áreas de la ciencia ha demostrado claramente la posibilidad de

conseguir mejores resultados que los obtenidos mediante los modelos lineales.

En este trabajo se hace el planteamiento y aplicación de dos modelos no lineales al estudio de series temporales en hidrología. Ellos son:

1. *Ecuación diferencial estocástica con dos estados estables.*

Consistente en un modelo markoviano continuo no lineal en el que se considera que la variable de interés se mueve entre dos estados estables.

2. *Modelo Autorregresivo dependiente del régimen.*

En el cual se combinan modelos autorregresivos lineales y el efecto de una serie índice determinante de las características estadísticas de la variable.

Los modelos se aplican a la serie de caudal mensual de la estación RG8 de las Empresas Públicas de Medellín y a algunas otras series hidrológicas y se estudia su capacidad de reproducir las características de interés de dichas series.

De esta manera se da un primer paso hacia el estudio de modelos no lineales, que pueden permitir una mejor comprensión de la evolución de los procesos hidrológicos y plantean la posibilidad de lograr un aprovechamiento más adecuado de estos recursos.

2. MODELO DE ECUACION DIFERENCIAL ESTOCASTICA NO LINEAL

2.1 Introducción

Existen fenómenos representados por una serie temporal en la cual se observa la exis-

tencia de varios estados estables coexistentes. Esto significa que la variable de interés adopta características estadísticas diferentes en cada uno de los estados. El paso de la variable de un estado al otro se puede considerar debido a una fluctuación ocasionada por una fuerza externa o interna. El tiempo de residencia en cada estado es a su vez una variable aleatoria cuyas propiedades dependen del fenómeno en consideración.

Los modelos autorregresivos lineales no permiten analizar series de datos que presenten estas características. Un modelo conceptual que lo permite, está dado por un *modelo de regresión no lineal* en el cual la evolución de la variable de interés cambia entre dos estados estables debido a la acción de una perturbación que representa la fluctuación interna o externa. Dicha fuerza, que posee un comportamiento aleatorio, hace que la variable se mueva de un estado a otro.

2.2 Descripción del modelo

2.2.1 Procesos Autorregresivos Continuos de Primer Orden

El modelo continuo AR(1) puede escribirse de la siguiente forma:

$$\frac{dx}{dt} + \alpha x = F(t) \quad (2.1)$$

Esta ecuación es denominada *Ecuación Diferencial Estocástica Lineal*. El proceso estocástico $F(t)$ se denomina perturbación estocástica o fuerza estocástica y puede ser definido formalmente como la derivada de un proceso de Wiener (Priestley, 1992). Normalmente $F(t)$ se modela como un ruido blanco:

$$E[F(t)] = 0 \quad E[F(t)F(t')] = q^2 \delta(t-t') \quad (2.2)$$

(donde q^2 es la varianza y $\delta(t)$ es la función delta de Dirac).

Una extensión de los procesos autorregresivos lineales continuos son los modelos no lineales que presentan la siguiente forma:

$$\frac{dx}{dt} = f(x) + F(t) \quad (2.3)$$

A diferencia de la ecuación lineal, la función $f(x)$ es no lineal e incluye los aspectos no lineales de la evolución de la variable.

El ajuste de un modelo de este tipo puede ser considerado como una *regresión no lineal* para la interpretación de los datos (Demarée 1990).

2.2.2 Dinámica de los Procesos Markovianos Continuos

Con el fin de introducir algunos de los principales conceptos de la dinámica de los procesos continuos markovianos, se hará uso de un ejemplo de la mecánica: el oscilador armónico. Considérese el caso de una partícula con masa m pequeña unida a un resorte y sobre la cual actúa una fuerza de amortiguación γ muy grande. Llamando k a la constante de Hooke, la ecuación de movimiento de la partícula toma la forma:

$$\frac{dz}{dt} = -\frac{k}{\gamma} z \quad (2.4)$$

donde z es la posición de la partícula.

Obsérvese que la forma de la ecuación (2.4) es idéntica a la parte determinística de la

ecuación lineal (2.1). La fuerza f depende de la posición z de la partícula y realiza un trabajo W total:

$$W = \int_{z_0}^{z_1} f(z) dz \quad (2.5)$$

El negativo de W es llamado *Potencial* U . Usando la ecuación (2.5) se obtiene:

$$f(z) = -\frac{dU}{dz} \quad (2.6)$$

Esta curva se puede interpretar como la pendiente de una ladera. En el sitio de pendiente cero $f(z)$ se anula, la partícula está en un punto de equilibrio estable.

Si se incluye ahora en la ecuación de movimiento de la partícula la fuerza aleatoria $F(t)$, se obtiene una ecuación idéntica a la (2.3). Esta fuerza tiene como efecto, que en su movimiento, la partícula será constantemente 'golpeada' en forma aleatoria, lo que la hará desplazarse de su posición de equilibrio.

Considérese ahora una función no lineal para la fuerza $F(z)$. Por ejemplo supóngase que $F(z)$ contiene un término cúbico además del lineal, la ecuación de movimiento es

$$\frac{dz}{dt} = -kz - k_1 z^3 \quad (2.7)$$

El potencial para el caso en el que $k < 0$ y $k_1 > 0$ aparece dibujado en la figura 1. Los puntos de equilibrio están determinados por la condición $dz/dt = 0$, con lo cual se encuentran tres soluciones, una inestable y dos estables.

2.2.3 La Función de Distribución de Probabilidad y los Momentos del Proceso

La ecuación (2.3), se puede reescribir en términos del potencial así:

$$\frac{dx}{dt} = -\frac{\partial U}{\partial x} + F(t) \quad (2.8)$$

Debido a la componente aleatoria en la ecuación (2.8) es imposible conocer la posición exacta de la variable x en un tiempo t . En cambio es posible determinar la función de probabilidad. Así, la densidad de probabilidad $P(x,t)$ de tener un valor x de la variable de estado en un tiempo t está dada por la ecuación de Fokker-Planck (véase Haken, 1983).

$$\frac{\partial P(x,t)}{\partial t} = \frac{\partial (U'P(x,t))}{\partial x} + \frac{q^2}{2} \frac{\partial^2 P(x,t)}{\partial x^2} \quad (2.9)$$

$$\text{con } U' = \frac{\partial U}{\partial x}$$

La solución de estado estable de esta ecuación, $P_e(x)$, es

$$P_e(x) = \frac{1}{Z} \exp\left[-\frac{2}{q^2} U(x)\right], \quad (2.10)$$

donde el factor de normalización Z está dado por

$$Z = \int_D \exp\left[-\frac{2}{q^2} U(x)\right] dx$$

donde D es el dominio de variación de x .

Es interesante resaltar el hecho de que en el caso de $f(x)$ lineal la función de distribución

de probabilidad es una función de probabilidad normal. En la figura 1 se puede observar el potencial y la función de probabilidad para $f(x)$ cúbica.

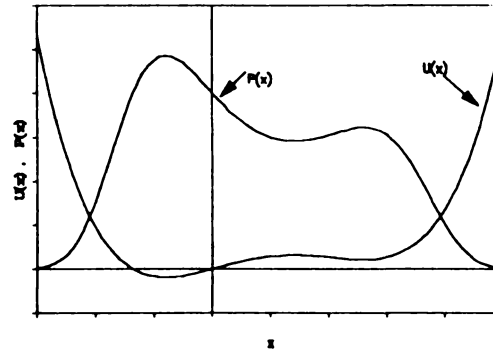


FIGURA 1.
Potencial y Función de Prob.
para $f(x)$ cúbica.

Al contar con la función de distribución de probabilidad de estado estable de la variable $P_e(x)$ es posible determinar el valor de los momentos mediante integración, por ejemplo para el primer momento se tiene

$$E[x] = \int_D x P_e(x) \quad (2.11)$$

donde D es el dominio de variación de x .

2.2.4 Tiempo de Permanencia de la Variable en un Estado

Es de interés conocer por cuanto tiempo una partícula cuya posición es descrita por la ecuación de Fokker-Planck permanece dentro de una cierta región de x . Este problema ha sido denominado tiempo de escape sobre una barrera de potencial.

El tiempo medio necesario para que la variable pase desde el punto a hasta un punto x_0 está dado por (véase Gardiner, 1985)

$$\pi(a \rightarrow x_0) = \frac{1}{q^2} \int_a^{x_0} \exp\left[\frac{U(y)}{q^2}\right] \int_{-\infty}^y \exp\left[-\frac{U(z)}{q^2}\right] dz dy \quad (2.12)$$

El tiempo medio necesario para que la variable pase desde el punto c hasta un punto x_0 está dado por:

$$\pi(c \rightarrow x_0) = \frac{1}{q^2} \int_c^{x_0} \exp\left[\frac{U(y)}{q^2}\right] \int_y^{\infty} \exp\left[-\frac{U(z)}{q^2}\right] dz dy \quad (2.13)$$

Utilizando las anteriores ecuaciones es fácil determinar el tiempo medio necesario para que la variable pase del valor a al valor c y viceversa (figura 2).

De igual manera es posible determinar el tiempo medio necesario para que la variable, teniendo un valor dado, alcance otro valor cualquiera.

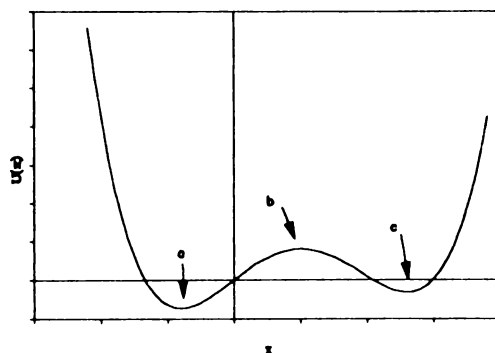


FIGURA 2.
Potencial con dos modas.

2.3. Aplicación del modelo a las series hidrológicas

Los procesos climáticos de gran escala están constantemente afectados por perturbaciones debidas a procesos de una escala más pequeña, que pueden ser consideradas fuer-

zas aleatorias tanto en el tiempo como en el espacio. Esto ha incentivado a varios investigadores a usar el modelo de ecuación diferencial estocástica en el estudio de los procesos climáticos. (Véase: Hasselman 1976, Sutera 1979, Nicolis y Nicolis 1981, Rodríguez-Iturbe et al. 1990, Demarée y Nicolis 1990).

2.3.1 Elección del Modelo

Algunos argumentos útiles para la escogencia del modelo adecuado en este caso, han sido dados por anteriores estudios sobre el comportamiento climático global.

Enfield (1989) y otros investigadores han indicado que es posible considerar que la Oscilación del Sur (que refleja la interacción entre el océano y la atmósfera en el Pacífico) se mueve entre dos estados contrastantes: El Niño y La Niña.

La estrecha relación entre la Oscilación del Sur y el régimen de caudales y lluvias en el país (Hastenrath 1990, Aceituno 1988, Poveda y Mesa 1993) induce a pensar en que también se puede considerar que estos últimos varían entre dos estados coexistentes.

La función cúbica, como se ilustró antes, es la forma más simple de incluir en la ecuación de evolución de un proceso la existencia de dos estados estables coexistentes; entonces, se utilizará esta función para el presente caso.

2.3.2 El Modelo Cúbico

Supóngase que la ecuación de evolución del proceso es

$$\frac{dx}{dt} = -x^3 + \lambda x^2 - \rho x + v + F(t) \quad (2.14)$$

Los parámetros λ , ρ , v están relacionados con los tres puntos de equilibrio (a, b, c) mediante

$$\begin{aligned} \lambda &= a + b + c, & \rho &= bc + ca + ab, & (2.15) \\ v &= abc \end{aligned}$$

Se encuentra que el potencial $U(x)$ está dado por:

$$U(x) = \frac{x^4}{4} - \lambda \frac{x^3}{3} + \rho \frac{x^2}{2} - vx \quad (2.16)$$

Las soluciones estables (a, c) corresponden a los mínimos de U , mientras que la solución inestable b corresponde al máximo.

2.3.2.1 Determinación de los Parámetros de la Ecuación

El problema ahora, consiste en determinar los parámetros de la ecuación de evolución partiendo de los datos históricos. Haciendo uso de las ecuaciones (2.15) se concluye que al determinar los parámetros a, b, c, q^2 , la ecuación del proceso queda definida. Demaree y Nicolis (1990) en la aplicación de este modelo de función cúbica a una serie de datos anuales de precipitación en el Sahel, determinaron los parámetros a y c como los promedios de la serie dividida en dos grupos de datos y encontraron los valores de los parámetros b , q^2 , y el parámetro de evolución temporal mediante linealizaciones de la ecuación de evolución del proceso alrededor de los puntos a y c .

Este procedimiento sin embargo, tiene el inconveniente de no preservar la media, la

varianza, ni el coeficiente de correlación de rezago uno de la serie original.

Por la razón anterior, en el presente trabajo se adoptó un método diferente para la determinación de los parámetros de la serie, método que a continuación se describe.

Si se parte de la hipótesis de que existen dos estados estables, la función de densidad de probabilidad de la variable presentará dos modas, y el histograma de frecuencias de la serie histórica, que es la versión discreta de la función de distribución de probabilidad, debe mostrar a su vez la presencia de dichas modas (véase la figura 1). Estos valores corresponden a la ubicación de los puntos a y c . Existe un método cuantitativo para la determinación de las modas de una serie descrito por Press (1989) en el cual se hace uso de la teoría de los procesos de Poisson. Se remite al lector interesado en los detalles a dicha referencia.

Para la determinación de los parámetros b (punto de equilibrio inestable) y q^2 (varianza de la fuerza aleatoria) se utilizará la condición de que con estos parámetros el modelo preserve la media y la varianza de la serie histórica. Para ello se evaluarán numéricamente las ecuaciones de la media y la varianza y variando los valores de b y q^2 (en un proceso de ensayo y error) se hallan los valores adecuados con los cuales se obtiene la media y la varianza de la serie histórica.

Estas ecuaciones no son lineales, por lo cual es posible que existan varias soluciones para b y q^2 . Un criterio para el ajuste es que la función de probabilidad del modelo teórico debe ser 'parecida' al histograma de frecuencias, o lo que es igual (pero que evita el problema de la discretización) que la función acumulada de probabilidad del modelo sea 'parecida' al histograma acumulado de frecuencias.

Este modelo, con la función $f(x)$ cúbica, se trató de aplicar a la serie de caudales mensuales estandarizados de la estación RG8 de las EE.PP.MM. y a otras series de caudales y lluvias. Los resultados obtenidos no son buenos con respecto a las 'colas' de la distribución por la siguiente razón: La ecuación del potencial $U(x)$ es de cuarto orden. De acuerdo a la ecuación (2.10) esta función de densidad de probabilidad tiene un decaimiento de las colas muy rápido debido al orden de $U(x)$, característica que está en desacuerdo con el comportamiento de las series de datos históricos estudiados en este caso. Por esta razón se decidió tomar una nueva forma para la función $f(x)$.

2.3.3 La Función $f(x)$ Compuesta

Si se adopta la suposición de que las colas de la distribución de los datos decaen según una función de distribución normal, esto significaría que en esas zonas el potencial $U(x)$ es de orden dos y por lo tanto la función $f(x)$ es de orden uno.

Se quiere responder entonces a la siguiente pregunta: ¿Cómo conseguir un potencial que presente dos estados estables y que tenga además una función de probabilidad con decrecimiento normal en las colas, sin necesidad de añadir nuevos parámetros?

La respuesta más simple está dada por un potencial como el mostrado en la figura 3., en el cual $U(x)$ es de orden dos para valores menores que a , $U(x)$ es de orden cuatro para valores entre a y c , y $U(x)$ es de orden dos para valores mayores a c .

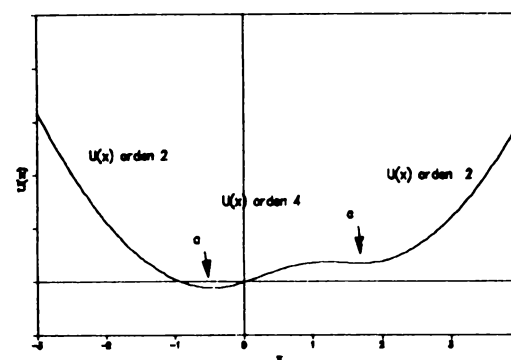


FIGURA 3.
Potencial Compuesto

Los parámetros adicionales se determinan exigiendo la condición de que el valor del potencial y su derivada sean continuos.

Sobre la naturaleza del decaimiento de las colas de la densidad de probabilidad en las series hidrológicas ha existido una gran discusión por parte de los investigadores (Mandelbrot 1971). En este trabajo se hizo la suposición más 'simple' asumiendo que este decaimiento ocurre de acuerdo a una distribución normal. Sin embargo es claro que es muy fácil incluir en el modelo cualquier otra alternativa, simplemente variando el potencial $U(x)$.

Los parámetros a, b, c, q^2 que definen la función $f(x)$ compuesta son los mismos que los de la ecuación cúbica explicados en el numeral anterior y se determinan de igual forma.

2.3.4 El Parámetro de Evolución Temporal

La estructura de evolución temporal del proceso gobernado por la ecuación diferencial estocástica que se está estudiando está dado por la función de autocovarianza. La solución numérica de esta ecuación reviste una

relativa dificultad y por ello se hará uso de la simulación para lograr que el modelo preserve el coeficiente de autocorrelación de rezago uno a una escala de tiempo determinada.

Es decir, que el tiempo t del proceso de la ecuación 2.2 es una unidad que debe ser reescalada para que corresponda al tiempo t' del proceso que se quiere modelar.

Esto es : $\omega t = t'$

lo que significa que ω unidades de tiempo del proceso original corresponden a una unidad de tiempo del proceso que se modela.

La determinación del parámetro ω se hace generando series con diferente valor de t' , hasta encontrar que los datos generados presenten un valor de autocorrelación de rezago uno igual al histórico.

2.3.5 *Uso del modelo para generación de series*

Para generar series de datos con el modelo se requiere resolver numéricamente la ecuación de evolución del proceso. Se utilizó el método de Euler, que además de ser el más simple, permite encontrar resultados adecuados si el intervalo de solución es pequeño.

La ecuación anterior en diferencias queda

$$X_{t+\Delta t} = X_t + \Delta t f(x) + \Delta t \varepsilon(t) \quad (2.17)$$

donde $\varepsilon(t) = \Delta t^{0.5} q W(t)$ con $W(t)$ un ruido blanco con media cero y varianza 1. La 'lectura' de los datos se hace cada ωt . Es

decir que si bien la ecuación (2.17) se resuelve cada Δt , sólo interesan los valores de la variable cada ωt .

2.4. *Aplicación del modelo*

2.4.1 *Criterios de Evaluación de los Resultados de la Aplicación del Modelo*

Es importante evaluar las series de datos obtenidas al aplicar el modelo.

Para ello se consideraron los siguientes aspectos:

Características Relacionadas con Eventos Extremos: Se determinaron los principales parámetros de las excursiones de la variable relacionadas con las sequías y con los excesos, o en forma más general, con las características de las excursiones del proceso aleatorio por encima o por debajo de un valor límite.

Características Relacionadas con el Almacenamiento: Otra característica de interés de las series es la relacionada con el almacenamiento. Los primeros trabajos en esta área fueron hechos por Hurst (1951) y sus resultados dieron lugar al hoy llamado fenómeno de Hurst.

Se determinó el rango reescalado de las desviaciones de la media acumuladas.

Espectro de Potencias: En el estudio de series de tiempo es muy importante la determinación de la presencia de comportamientos cíclicos. Para esto es útil el uso del espectro de potencias. En el presente trabajo

se determinó el espectro de potencias por el Método de Máxima Entropía (M.E.M.). (véase Press 1989).

2.4.2 Aplicación del Modelo a la Serie de Caudales de la Estación Riogrande-RG8.

Se aplicó el modelo de ecuación diferencial estocástica planteado, al estudio de la serie de caudales mensuales de la estación RG8 (rio Riogrande) de las Empresas Públicas de Medellín. Dicha estación cuenta con registros entre los años 1942 y 1993 y puede considerarse una de las mejores series de datos del país en cuanto a calidad de la información y extensión del periodo de registro. Con el fin de eliminar el ciclo anual se estandarizó la serie.

El histograma de frecuencias de la serie se muestra en la figura 4. La traza histórica se muestra en la figura 5.

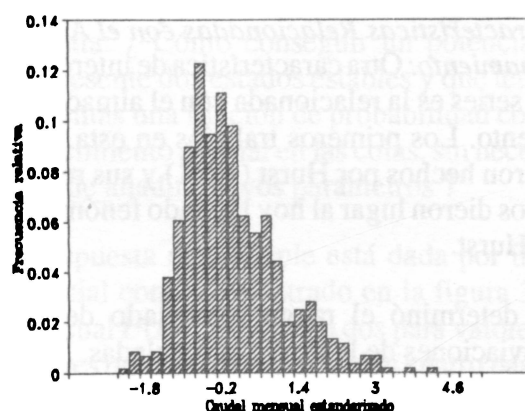


FIGURA 4. Histograma de frecuencias de la serie de caudal mensual estandarizado. Estación RG8.

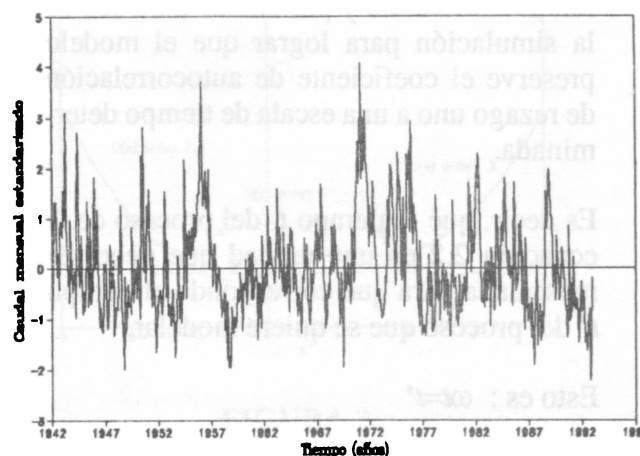


FIGURA 5. Serie histórica de caudal mensual estandarizado Estación RG8.

Los parámetros del modelo obtenidos para la serie de datos de la estación RG8 son los siguientes:

$$a = -0.4, b = 1.2, c = 1.7, q^2 = 1.23, \omega = 0.8$$

se encontraron los tiempos medios de transición entre estados con los siguientes resultados.

$$T'(a \rightarrow c) = 14.3 \text{ meses y } T'(c \rightarrow a) = 4.8 \text{ meses}$$

Al igual que se observa en la serie histórica, estos valores indican que estando en período de sequía el tiempo medio de transición hacia un período húmedo es mayor que el tiempo medio para pasar de una época húmeda a una seca.

El potencial y la función de probabilidad se muestran en la figura 6.

Una traza generada mediante el modelo se muestra en la figura 7.

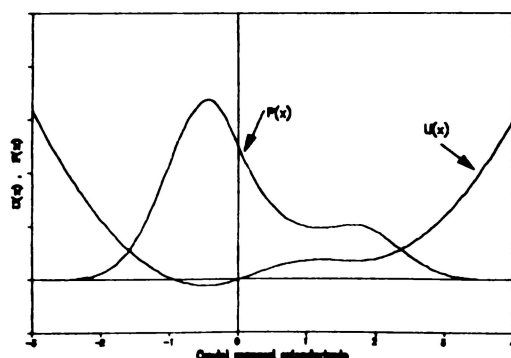


Figura 6. Potencial y Función de Prob. de los caudales de la estación RG8.

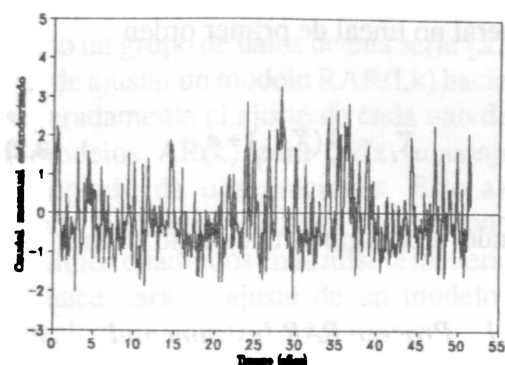


FIGURA 7. Serie generada con el modelo E.D.E. (parámetros de la estación RG8).

La aplicación señaló que los principales parámetros de eventos extremos (como valor esperado de la longitud de la excursión, volúmenes de excursión, número de excursiones, etc) son representados adecuadamente por el modelo.

Se encontró el rango reescalado de las series de igual longitud a la histórica generadas por el modelo. Los resultados se muestran a continuación al igual que los resultados obtenidos utilizando un modelo AR(1) discreto con transformación logarítmica (ARTL).

| Rango Escalado | | |
|---------------------|-------|-------|
| | EDE | ARTL |
| Promedio | 62.66 | 57.30 |
| Desviación estándar | 15.21 | 12.07 |

El valor de este parámetro en la serie histórica es 69.56, por lo cual se observa que el resultado del método de la EDE es preferible al del ARTL.

Aunque el modelo de EDE puede reproducir ciertas frecuencias en particular, el presente estudio no permitió concluir claramente las posibilidades del modelo en este sentido.

Se aplicó este modelo a otras series de caudales (por ejemplo al Río Tenche) y a series de precipitación (precipitación en el embalse el Peñol). Los resultados obtenidos concuerdan con los obtenidos para la serie de caudales de la estación RG8 y señalan que el modelo presenta ventajas sobre un modelo autorregresivo con transformación logarítmica.

3. MODELO AUTORREGRESIVO DEPENDIENTE DEL REGIMEN

3.1 INTRODUCCION

Los modelos autorregresivos lineales presentan diversas limitaciones para el estudio de ciertos comportamientos no lineales que se observan en los datos. Por ejemplo no permiten representar procesos en los cuales las propiedades estadísticas, dependen de un proceso externo que afecta la variable de interés.

Diversos estudios muestran que el régimen hidrológico Colombiano está claramente li-

gado a variaciones macroclimáticas y entre ellas principalmente con la Oscilación del Sur, de la cual el Índice de Oscilación del Sur (IOS) es un buen indicador (Hastenrath 1990, Aceituno 1988).

Por ello es interesante explorar la posibilidad de que las series de precipitación y caudal en Colombia tengan propiedades estadísticas dependientes del estado que presente esta variable macroclimática.

El propósito de este estudio es investigar la aplicación de una clase de modelos de series de tiempo no lineales que ajustan su media y sus propiedades estocásticas dependiendo del estado de una variable que se considere influyente en el comportamiento del fenómeno.

Esta clase de modelos llamados "Modelos autorregresivos dependientes del régimen (RAR)" son una extensión de los modelos autorregresivos lineales (AR).

3.2. DESCRIPCION DEL MODELO

Los modelos autorregresivos dependientes del régimen (RAR) son una clase de modelos no lineales, propuesto inicialmente por Tong (1983) quien los denominó '*threshold models*'.

La idea básica consiste en una serie $\{X_t\}$ cuyos parámetros varían de acuerdo a los valores de un número finito de valores pasados de X_t , o de acuerdo a un número finito de valores pasados de una serie asociada $\{Y_t\}$.

Por ejemplo el modelo RAR(2,1) está dado por:

$$X_t = a_0^{(1)} + a_1^{(1)}X_{t-1} + e_t^{(1)} \quad \text{si } X_{t-1} < r \quad (3.1)$$

$$X_t = a_0^{(2)} + a_1^{(2)}X_{t-1} + e_t^{(2)} \quad \text{si } X_{t-1} \geq r$$

Donde $[e_t^{(1)}], [e_t^{(2)}]$, son procesos de ruido blanco $a^{(1)}, a^{(2)}$, son constantes y r es una constante llamada 'límite del régimen'.

Este modelo es simplemente el uso de una de dos ecuaciones de procesos AR(1) dependiendo de si la variable X_t está por encima o por debajo del valor límite r , en la etapa anterior. Este modelo se puede extender a un modelo de l regímenes y k términos de regresión RAR(l, k).

El modelo RAR($l, 1$) puede considerarse una aproximación lineal por tramos al modelo general no lineal de primer orden

$$X_t = \lambda(X_{t-1}) + e_t \quad (3.2)$$

Donde λ es alguna función no lineal.

3.2.1 Proceso RAR Instrumental

Tong (1983) ha extendido los modelos dependientes del régimen para incluir los casos en donde el cambio entre los grupos de parámetros está determinado por un valor pasado de un proceso asociado $[Y_t]$, en vez de estarlo por un valor pasado de $[X_t]$.

Cuando sólo se tienen dos regímenes y a cada uno de ellos se ajusta un modelo AR(1), el modelo toma la forma

$$\begin{aligned} X_t &= a_0^{(1)} + a_1^{(1)}X_{t-1} + e_t^{(1)} \quad \text{si } Y_{t-d} < r \\ X_t &= a_0^{(2)} + a_1^{(2)}X_{t-1} + e_t^{(2)} \quad \text{si } Y_{t-d} \geq r \end{aligned} \quad (3.3)$$

La variable externa $\{Y_t\}$ es normalmente otra serie de tiempo estocástica; cuando esta variable índice cruza ciertos límites, los pa-

rámetros del modelo AR que gobiernan la evolución de la serie de tiempo modelada cambian. El cambio en los parámetros puede producirse luego de un cierto tiempo de rezago d .

La serie que se utiliza como índice debe estar ligada mediante algún mecanismo físico a la serie que se está modelando, de tal manera que se justifique suponer que en realidad su estado tenga un efecto importante sobre el comportamiento de la serie de interés.

3.2.2 Estimación de los parámetros

Dado un grupo de datos de una serie $\{X_t\}$, se puede ajustar un modelo RAR(1,k) haciendo separadamente el ajuste de cada uno de los l modelos AR(k) para cada subconjunto apropiado de observaciones. Este ajuste puede realizarse mediante el método de mínimos cuadrados en la misma manera que se hace para el ajuste de un modelo AR lineal.

Para la determinación de los parámetros estructurales del modelo, es decir el parámetro de rezago d , las regiones $\{R^{(j)}\}$ que señalan los distintos regímenes y el orden k_1, \dots, k_j , Tong propone un método que se basa en el Criterio de Información de Akaike (AIC) y que consiste en buscar la combinación de parámetros que produce el mínimo valor del AIC.

3.2.3 Determinación de los Coeficientes de las Ecuaciones

En este trabajo la atención se concentró en el modelo RAR definido por la ecuación (3.1) y en el modelo RAR Instrumental definido por la ecuación (3.3). Es decir los modelos que presentan dos regímenes en cada uno de los cuales la variable se modela

mediante un proceso AR(1) (procesos markovianos).

El interés ahora radica en la forma de definir los parámetros a de las ecuaciones (3.1) y (3.3) de tal manera que los modelos sean estacionarios de orden dos.

Para ello nótese la similitud entre los modelos RAR y los modelos estacionales. Thomas y Fiering (1971) plantearon un modelo AR(1) con coeficientes variables para considerar la estacionalidad de las series. Este modelo es conocido como modelo autorregresivo univariado estacional de rezago uno y tiene la siguiente forma

$$(X_{jt} - \mu_j) = \rho_j \frac{\sigma_j}{\sigma_{j-1}} (X_{j,t-1} - \mu_{j-1}) + \sigma_j (1 - \rho_j^2)^{1/2} W_{jt} \quad (3.4)$$

donde

X_{jt} es el valor del proceso en el tiempo t en la estación j .

ρ_j es el coeficiente de correlación de rezago 1 entre las estaciones j y $j-1$.

μ_j es el valor medio en la estación j .

σ_j es la desviación estándar en la estación j .

W_{jt} es una variable normal con media cero y varianza 1.

Aplicando este modelo a cada uno de los regímenes se encuentra que la ecuación (3.1) toma la forma

(3.5)

$$X_t = \mu_t^{(1)} + \rho_t^{(1)} \frac{\sigma_t^{(1)}}{\sigma_{t-1}^{(1)}} (X_{t-1} - \mu_{t-1}^{(1)}) + \sigma_t^{(1)} (1 - (\rho_t^{(1)})^2)^{1/2} e_t^{(1)} \quad \text{si } X_{t-1} < r$$

$$X_t = \mu_t^{(2)} + \rho_t^{(2)} \frac{\sigma_t^{(2)}}{\sigma_{t-1}^{(2)}} (X_{t-1} - \mu_{t-1}^{(2)}) + \sigma_t^{(2)} (1 - (\rho_t^{(2)})^2)^{1/2} e_t^{(2)} \quad \text{si } X_{t-1} \geq r$$

donde el superíndice (.) indica que estos parámetros se estiman para los valores condicionados al régimen.

El planteamiento para el modelo RAR Instrumental de la ecuación (3.2) es idéntico, con lo cual se obtiene el modelo

$$X_t = \mu_t^{(1)} + \rho_t^{(1)} \frac{\sigma_t^{(1)}}{\sigma_{t-1}^{(1)}} (X_{t-1} - \mu_{t-1}^{(1)}) + \sigma_t^{(1)} (1 - (\rho_t^{(1)})^2)^{1/2} e_t^{(1)}$$

si $Y_{t-d} < r$

$$X_t = \mu_t^{(2)} + \rho_t^{(2)} \frac{\sigma_t^{(2)}}{\sigma_{t-1}^{(2)}} (X_{t-1} - \mu_{t-1}^{(2)}) + \sigma_t^{(2)} (1 - (\rho_t^{(2)})^2)^{1/2} e_t^{(2)}$$

si $Y_{t-d} \geq r$

3.3 RELACION ENTRE LOS MODELOS RAR Y LOS PROCESOS MARKOVIANOS CONTINUOS

El modelo RAR(2,1), dado por la ecuación (3.1) y el modelo RAR Instrumental(2,1), dado por la ecuación (3.3), son procesos en tiempo discreto. La variación de sus parámetros de acuerdo al valor de variable indicadora del régimen puede asimilarse con la existencia de dos condiciones de evolución coexistentes. (idea similar a la utilizada en el planteamiento del modelo de la Ecuación Diferencial Estocástica, véase el numeral 2).

Teniendo en cuenta la relación entre los procesos en tiempo discreto y en tiempo continuo, la versión continua del modelo RAR(1,2), dado por la ecuación (3.1), tendría un potencial con la forma mostrada en la figura 8.

El proceso (o la partícula) se mueve sobre el potencial de la izquierda si $X_{t-1} > r$ y sobre

el potencial de la derecha si $X_{t-1} \leq r$. Nótese que en este caso la escala de evolución temporal de la variable (determinada por el coeficiente de correlación) varía al pasar de un potencial al otro.

El modelo RAR Instrumental tendría, en su versión continua, el potencial mostrado en la figura 9.

El paso del proceso de un potencial al otro está determinado por la variable externa Y_{t-d} que hace 'saltar' la partícula (o el proceso) de un potencial al otro. Al igual que en el caso anterior la escala de evolución temporal varía en cada potencial.

Esta analogía ilustra la relación entre los procesos continuos descritos por la Ecuación Diferencial Estocástica (numeral 2) y los procesos RAR.

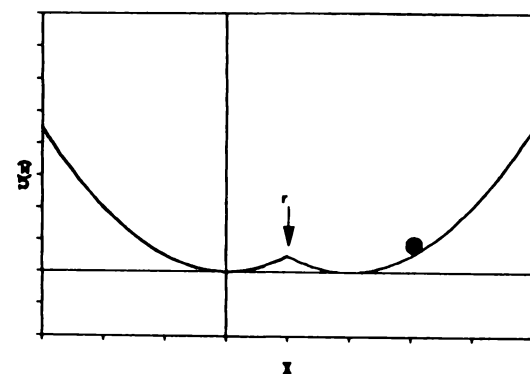


Figura 8. Analogía del potencial para un modelo RAR

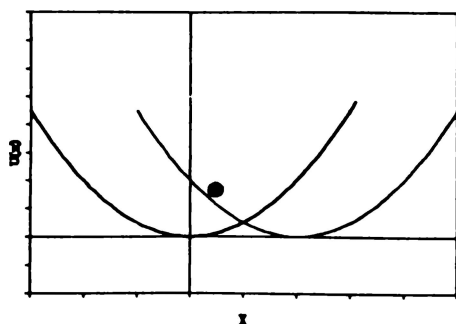


Figura 9 Analogía del potencial para un modelo RAR Instrumental

3.4. APLICACION DEL MODELO A ALGUNAS SERIES

Este tipo de modelos han sido utilizados para el estudio de diferentes clases de series. Tong (1980) las propuso para el análisis de series de caudales. Además, han sido utilizadas con algún éxito para modelar la serie de datos del número de manchas solares, y otras series que presentan un comportamiento no lineal.

El modelo RAR instrumental fue aplicado por Zwiers y Von Storch (1990) al modelamiento de la oscilación del sur.

Para la evaluación de los resultados de la aplicación del modelo se utilizaron los mismos criterios que en la evaluación del modelo de Ecuación Diferencial Estocástica:

Características Relacionadas con Eventos Extremos

Características Relacionadas con el Almacenamiento

Espectro de Potencias

Además se compararon los resultados con los obtenidos mediante un modelo autorregresivo de orden uno al cual se le hace la transformación logarítmica (que es el método más utilizado para el modelamiento de series de caudales y lluvias).

Aplicación del Modelo a la Serie de Datos de Caudal de la Estación RG8

Los modelos RAR Instrumental y RAR se aplicaron a la serie de caudales mensuales de la estación RG8 (Riogrande) de las Empresas Públicas de Medellín.

Aplicación del modelo RAR Instrumental

Según se mostró en la ecuación (3.3) el modelo RAR instrumental utiliza una variable externa para determinar el régimen en el cual se encuentra la variable en estudio.

En este caso se utilizó como variable externa la serie del IOS (Índice de Oscilación del Sur, definido como la diferencia de presiones normalizadas entre Tahití y Darwin) con promedio móvil 5 meses, que es un índice de la oscilación del sur, la cual ha mostrado tener una estrecha relación con los valores de los caudales y lluvias en el país (Hastenrath 1990, Aceituno 1988).

El espectro de potencias de la serie mensual del IOS, determinado mediante el método de máxima entropía, presenta los valores máximos (usando 100 polos) en las frecuencias $f=0.016$ (correspondiente a 5.2 años) y $f=0.034$ (correspondiente a 2.4 años). Utilizando 20 polos se obtiene un valor máximo en la frecuencia $f=0.015$ (correspondiente a 5.5 años). Existe una clara relación entre estas frecuencias y las de la serie de caudales históricos de la estación RG8 (figura 10).

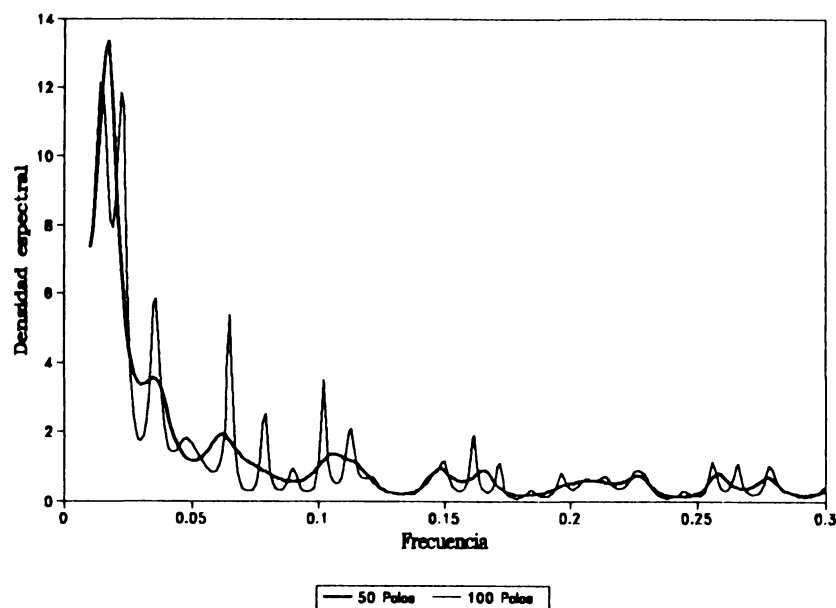


FIGURA 10. Espectro de potencias de la serie de caudal mensual estandarizado. Estación RG8

Se hizo la aplicación del modelo descrito por la ecuación (3.3) asumiendo que es suficiente el ajuste de modelos AR(1) para los diferentes regímenes. Además se supuso que solamente existen dos regímenes.

Entonces el modelo RAR Instrumental según la ecuación (3.6) tiene la siguiente forma:

$$X_t = 0.9 + 0.57 \left(\frac{1.05}{1.15} \right)^{0.5} (X_{t-1} - 0.84) + [1.05(1 - 0.57^2)]^{0.5} e_t$$

si $Y_{t-2} \geq 0.4$

$$X_t = -0.3 + 0.53 \left(\frac{0.6}{0.61} \right)^{0.5} (X_{t-1} + 0.28) + [0.61(1 - 0.53^2)]^{0.5} e_t$$

si $Y_{t-2} < 0.4$

donde $e(t)$ es un ruido blanco de media cero y varianza unitaria.

Con el modelo se generaron varias series de datos. Una traza de estas se muestra en la figura 11, El espectro de potencias de otra serie se muestra en la figura 12.

Aplicación del modelo RAR

El ajuste del modelo RAR a la serie de caudales de la estación RG8 tiene como se indica en las ecuaciones página siguiente.

Una traza generada con el modelo se presenta en la figura 13, el espectro de potencias de otra traza se muestra en la figura 14.

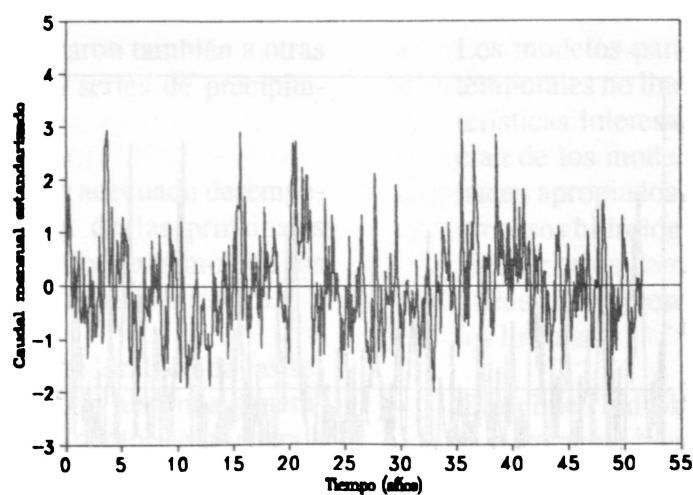


FIGURA 11. Serie generada con el modelo RAR Instrumental (parámetros estación RG8)

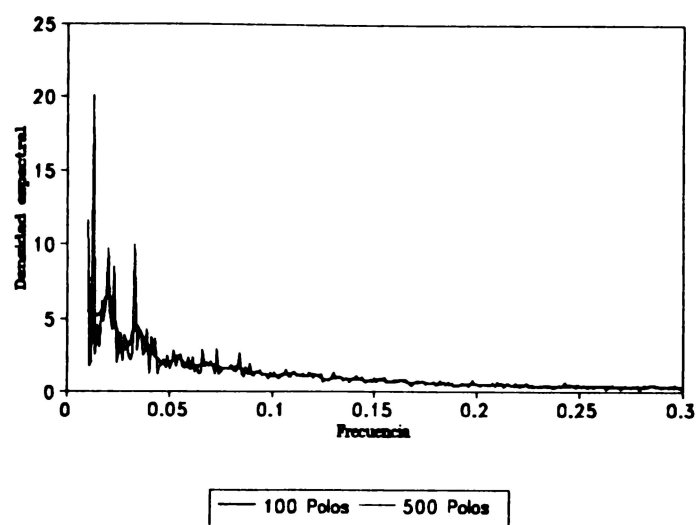


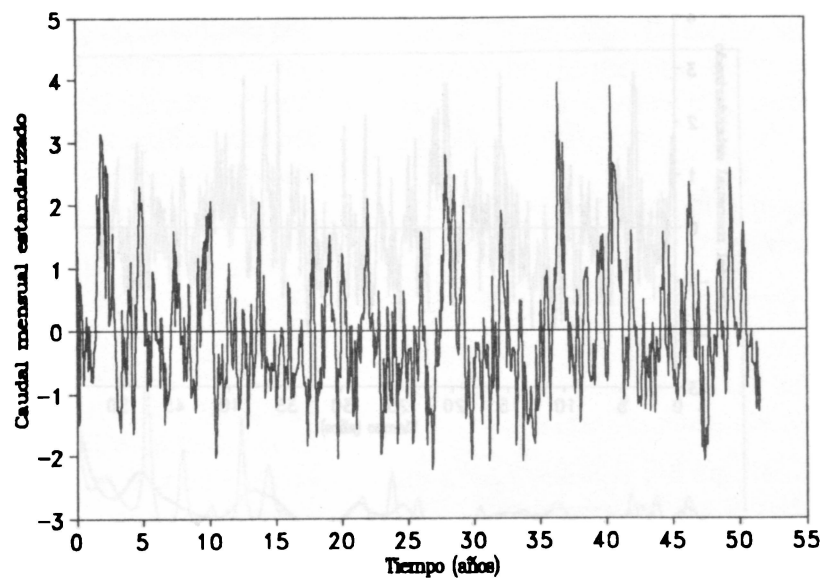
FIGURA 12. Espectro de potencias (M.M.E) de serie de 20000 datos generada con el modelo RAR Instrumental (parámetros estación RG8)

$$X_t = 0.83 + 0.46 \left(\frac{1.06}{0.54} \right)^{0.5} (X_{t-1} - 1.19) + [1.06(1 - 0.46^2)]^{0.5} e_t$$

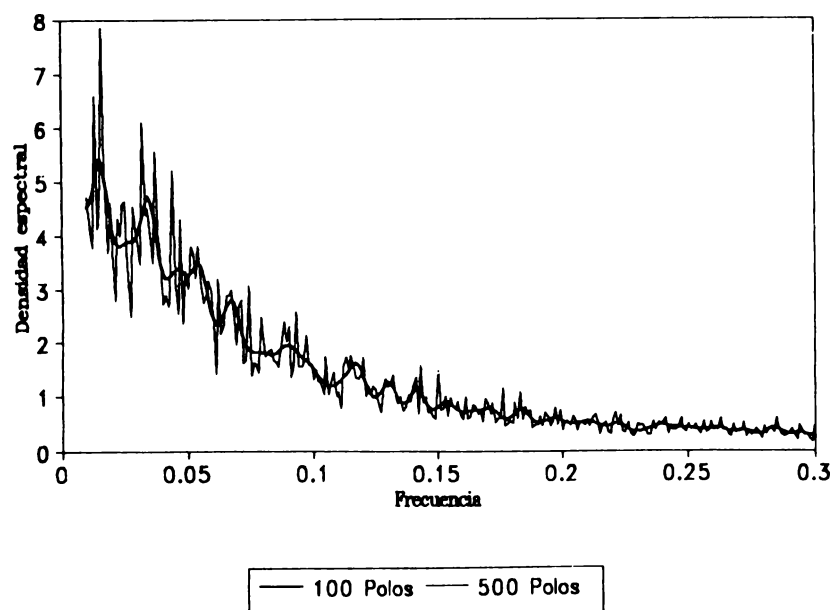
$$\text{si } X_{t-1} \geq 0.3$$

$$X_t = -0.37 + 0.39 \left(\frac{0.5}{0.25} \right)^{0.5} (X_{t-1} + 0.54) + [0.5(1 - 0.39^2)]^{0.5} e_t$$

$$\text{si } X_{t-1} < 0.3$$



**FIGURA 13. Serie generada con el modelo RAR
(parámetros estación RG8)**



**FIGURA 14. Espectro de potencias (M.M.E.) de serie de 20000 datos
generada con el modelo RAR (parámetros estación RG8)**

Estos modelos se aplicaron también a otras series de caudales y a series de precipitación.

En general muestran un adecuado desempeño en la reproducción de las principales características de eventos extremos de los valores observados en la serie histórica.

Los valores del rango reescalado de las series de longitud igual a la histórica generadas con los distintos modelos se muestra a continuación:

| | RAR In- strumen- tal | RAR | ARTL |
|------------------------|-------------------------------|-------|-------|
| Promedio | 60.76 | 61.81 | 57.30 |
| Desviación estándar | 15.84 | 17.90 | 12.07 |

Donde ARTL corresponde al modelo AR(1) con transformación logarítmica.

Teniendo en cuenta que el valor de este parámetro de la serie histórica es 69.56 se concluye que los modelos RAR producen valores más cercanos al histórico que el modelo ARTL.

Este resultado tiene importantes implicaciones en la determinación de los volúmenes necesarios para la regulación de los caudales.

Los modelos reproducen también adecuadamente la forma de la distribución del histograma de frecuencias de la serie histórica.

4. CONCLUSIONES

El desarrollo de este trabajo permite sacar las siguientes conclusiones de tipo general:

- Los modelos para el estudio de series temporales no lineales presentan características interesantes que los diferencian de los modelos lineales y que los hacen apropiados para el estudio de los procesos hidrológicos, ya que éstos exhiben ciertos comportamientos que no pueden ser representados por los modelos lineales.
- El régimen hidrológico colombiano está claramente ligado a ciertas variables macroclimáticas. De allí que la comprensión de esta relación sea de gran importancia para conseguir un adecuado aprovechamiento del recurso hídrico en el país.
- Los dos tipos de modelos acá tratados brindan una representación más ‘intuitiva’ de la evolución de los procesos climáticos que la obtenida al realizar una transformación logarítmica de los datos.

La existencia de dos estados coexistentes en la evolución de algunos procesos climáticos estrechamente relacionados con la climatología colombiana han sido señalados por varios investigadores.

- Los dos modelos planteados reproducen adecuadamente las principales características de las series históricas de lluvias y caudales. En especial las relacionadas con los eventos extremos y las características para almacenamiento de la serie de caudales.

Acerca del modelo de Ecuación Diferencial Estocástica se concluye:

- El modelo es útil para representar variables que evolucionan entre estados estables.

- Mediante él es posible preservar las principales características estadísticas de la serie (media, varianza, coeficiente de correlación de rezago uno, forma del histograma de frecuencias) y reproducir adecuadamente los eventos extremos. Permite encontrar, en términos estadísticos, el tiempo de transición de la variable de un valor a otro.
- Los datos de caudal generados mediante el modelo presentan condiciones más desfavorables para estudios de almacenamiento que los obtenidos con el método usual de realizar una transformación logarítmica de un autorregresivo lineal, lo que conduciría, en caso de ser utilizados, a realizar obras con un criterio más conservador.
- Si bien la estimación de los parámetros planteada, entraña elementos arbitrarios, la metodología supera las dificultades señaladas de la metodología de Demarée y Nicolis (1990).

Las conclusiones sobre el modelo RAR son:

- Los modelos RAR y RAR Instrumental reproducen aceptablemente las características estadísticas de interés de las series en cuestión. En especial las relacionadas con eventos extremos (número de eventos, longitud etc.).
- Las series generadas mediante estos modelos producen condiciones más críticas para efectos de almacenamiento que las obtenidas mediante un modelo autorregresivo lineal con transformación logarítmica.
- Los modelos RAR Instrumental reproducen en forma clara las frecuencias presentes en la serie histórica, como

consecuencia de la relación que se establece con la variable índice. Esta es una característica interesante que no pueden reproducir los modelos autorregresivos lineales.

- Este modelo se puede considerar una regresión lineal por tramos de la serie original. Se observó que reproduce bien la forma del histograma de frecuencias de la serie de datos original, mostrando la asimetría típica de los histogramas de las series de caudales y lluvias.

BIBLIOGRAFIA

- Aceituno, P. On the Functioning of the southern Oscillation in the south American Sector. Part I: Surface Climate. Monthly Weather Review, Vol 116. No 3, marzo 1988
- Akaike, H. Maximum Likelihood Identification of Gaussian Autorregressive Moving Average Models. Biometrika. 60, pp. 255-265. 1973
- Bras, R. y Rodríguez-Iturbe, I. Random Functions and Hydrology. Addison-Wesley, United States 1985
- Box, G. E. y G. Jenkins. Time Series Analysis, Forecasting and Control, Holden-Day, Oakland, Calif., 1970
- Demarée, G. R. y Nicolis C. Onset of sahelian drought viewed as a fluctuation-induced transition. Q.J.R. Meteorol. Soc. 116, 1990, pp. 221-238
- Enfield, D. B. El Niño, Past and Present. Reviews of Geophysics, 27,1 pp. 159-187. 1989

- Gardiner, C. W. Handbook of stochastic methods. Springer Verlag, Berlín 1983
- Haken, H. Synergetics. Springer Verlag, Berlín 1983
- Hasselmann, K. Stochastic Climate Models Part I. Theory. Tellus 28 PP. 473-485. 1976.
- Hastenrath, S. Diagnostics and Prediction of Anomalous River Discharge in Northern South America 1990.
- Millan, J. Drought Impact on Regional Economy. Hydrology papers No. 55, Colorado State University. Fort Collins, Colorado, 1972.
- Nicolis, C. y Nicolis, G. Stochastic Aspects of Climatic Transitions-Additive Fluctuations. Tellus 33 PP. 225-234 1981.
- Nicolis, C. Stochastic Aspects of Climatic Transitions-Response to a Periodic Forcing. Tellus 34 PP. 1-9 1982
- Press, W. H. et al. Numerical Recipes. Cambridge University Press, 1989.
- Priestley, M. B. Spectral Analysis and Time Series. Academic Press, Great Britain 1992
- Priestley, M. B. Non-linear and Non-stationary Time Series analysis. Academic Press, Great Britain 1992.
- Poveda G. y Mesa O. Metodologías de Predicción de la Hidrología Colombiana Considerando el Fenómeno El Niño/Oscilación del Sur (ENOS). Tercer Congreso Colombiano de Meteorología. 1993
- Rodriguez-Iturbe I. et al. Non-linear Dynamics of Soil Moisture at Climate Scales, Stochastic Analysis 1990
- Salazar, J. E. Aplicación de Dos Modelos No Lineales al Estudio de Series Temporales en Hidrología. Medellín 1994. Tesis de Magíster en aprovechamiento de Recursos Hidráulicos. Universidad Nacional, Facultad de Minas, Medellín.
- Salas, J. D. et al. Applied Modeling of Hydrologic Time Series. Water Resources Publications (U.S.A.) 1980.
- Sutera, A. On Stochastic Perturbation and Long-Term Climate Behaviour. Q.J.R. Meteorol. Soc. 107 pp. 137-153
- Vannitsem, S. y Nicolis, C. Detecting Climatic Transitions: Statistical and Dynamical Aspects. Beitr. Phys. Atmosph., Agosto 1991, pp. 245-254
- Zwiers F. y Von Storch H. Regime-dependent Autoregressive Time Series Modeling of The Southern Oscillation. Journal of Climate. Vol 3 pp.1347-1363. Diciembre 1990

