

DINÁMICA DE LOS APORTES TOTALES AL SISTEMA HIDROELÉCTRICO INTERCONECTADO NACIONAL COLOMBIANO

Juan David Velásquez H. y Isaac Dyner R.

Grupo de Energía

Universidad Nacional de Colombia, Sede Medellín

jdvelasq@unalmed.edu.co; idyner@unalmed.edu.co

Reinaldo C. Souza

Faculdade de Engenharia Eletrica

Pontificia Universidade Católica do Rio de Janeiro

reinaldo@ele.puc-rio.br

Recibido para evaluación: 15 de junio de 2005/ Aceptación: 25 de agosto de 2005/ Recibida versión final: 30 de agosto de 2005

RESUMEN

En este artículo se analiza la dinámica de los aportes totales al sistema hidroeléctrico colombiano de generación desde la creación de la Bolsa de Energía, y se presenta un modelo autoregresivo de transición logística para su modelación. El modelo propuesto pasó un conjunto de pruebas estándar de diagnóstico. El análisis realizado indica que un modelo lineal autoregresivo no representa completamente las características no lineales intrínsecas a la dinámica de la serie. El modelo no lineal propuesto consta de dos modelos lineales, que representan la dinámica de los aportes "altos" y "bajos". Aunque el modelo obtenido es caótico localmente, permite realizar la predicción de los aportes mensuales totales varios meses adelante.

PALABRAS CLAVE: Hidrología, predicción, modelos STR

ABSTRACT

In this paper the dynamic of the total flows to the Colombian hydro electrical generation system since the creation of the spot market is analyzed, and a logistic smooth transition autoregressive model is used for its representation. The proposed model passes a battery of diagnostic standard tests. The results of the realized analysis indicate that a linear autoregressive model is not being able for representing completely the intrinsic non-linear characteristics of the time series dynamic. The proposed model is formed by two linear autoregressive models representing the dynamics of the high and low flows. Although the obtained model is locally chaotic, it allows to realize the prediction for the total monthly flows several periods ahead.

KEYWORDS: Hydrology, Prediction, STR models

1 INTRODUCCIÓN

La privatización del sector eléctrico colombiano, ocurrida a partir de la entrada en vigencia de la Ley de Servicios Públicos Domiciliarios (Ley 142) y la Ley Eléctrica (Ley 143), ha buscado fortalecer al sistema hidrotérmico de generación ante su vulnerabilidad frente a eventos hidrológicos extremos, tal como ya se evidenció durante los racionamientos de electricidad ocurridos a raíz del fenómeno de El Niño ocurrido entre 1991 y 1992. No obstante, el país se sigue basando fundamentalmente en la generación hidráulica durante periodos hidrológicos normales o húmedos, mientras que la generación térmica es requerida para mantener la estabilidad y seguridad del sistema de transmisión nacional. Consecuentemente, los precios de la electricidad en la Bolsa de Energía, dependen en gran medida de la evolución de los niveles de los embalses y los aportes hidrológicos, tal como lo indican Velásquez y Dyner (2001), y Pulgarín, Smith y Poveda (2001), así como de otras variables físicas que describen el estado del sistema de generación.

Bajo esta perspectiva, el recurso agua sigue siendo vital para garantizar la atención de la demanda de electricidad en el país, de tal suerte, que los generadores hidráulicos deben basar sus decisiones en las predicciones de las afluencias hidrológicas que llegarán a sus plantas de generación durante los próximos meses, de tal suerte, que los estudios realizados que se han reportado en la literatura, están orientados a aumentar la capacidad de predicción de los caudales de los ríos que llegan a las distintas centrales de generación [véase por ejemplo los trabajos de Smith, Velez, Velásquez, Ceballos, Correa, Goez, Salazar y Zapata (2004) y Zapata, Velásquez y Smith (2004)]. No obstante, es también importante entender la dinámica global de los aportes que llegan al sistema interconectado permitiendo una mayor comprensión tanto de la hidrología colombiana como del funcionamiento del mercado eléctrico. En consecuencia, este estudio está encaminado a la formulación de modelos autoregresivos de transición suave (STAR) que representen su dinámica, de tal forma que existan más elementos que permitan su mejor comprensión.

La sección 2 revisa la modelación no lineal de series temporales y el ciclo de formulación de modelos STAR. Posteriormente, los datos utilizados son presentados en la sección 3. Seguidamente en la sección 4, se reportan los resultados obtenidos; finalmente, en la sección 5 se concluye.

2 MODELOS DE TRANSICIÓN SUAVE Y SU ESPECIFICACIÓN

Desde un ámbito matemático, una serie temporal se define como un conjunto de observaciones ordenadas en el tiempo, $\{y_t\}_{t=1}^T$, las cuales son consideradas en conjunto, como la realización de una variable estocástica Y_t , cuya distribución de probabilidades está dada por:

$$Y_t \sim (f(\Psi; \mathbf{x}_t), \epsilon_t \sqrt{g(\Theta; \mathbf{x}_t)}) \quad (1)$$

donde $\mathbf{x}_t$ representa el conjunto de variables explicativas del comportamiento de y_t , que puede estar formado tanto por observaciones rezagadas de y_t como por otras variables exógenas. En la ecuación anterior, el valor esperado de la variable está determinado por $f(\Psi; \mathbf{x}_t)$, mientras que su varianza $g(\cdot)$ representa la incertidumbre inherente en la obtención del valor real y_t aunque se conozcan los respectivos regresores. Igualmente, se asume que ϵ_t es una variable aleatoria con media cero y varianza unitaria. Las funciones $f(\cdot)$ y $g(\cdot)$, así como el vector de regresores $\mathbf{x}_t$ se consideran desconocidos.

En el modelado y la predicción de series temporales se desea construir funciones $\hat{f}(\cdot)$ y $\hat{g}(\cdot)$ que aproximen con suficiente precisión a las funciones desconocidas $f(\cdot)$ y $g(\cdot)$, tal que el proceso generador de las observaciones es idealizado como:

$$\hat{Y}_t = \hat{f}(\hat{\Psi}; \mathbf{x}_t) + \hat{\epsilon}_t \sqrt{\hat{g}(\hat{\Theta}; \mathbf{x}_t)} \quad (2)$$

cuyo estudio permite entender e interpretar mejor la estructura de causalidad de la variable real estudiada, confrontando evidencias estadísticas basadas en los datos disponibles, con el conocimiento subjetivo que se posee. El proceso de construcción y especificación del modelo anterior es dependiente del objetivo final perseguido: en la caracterización, el modelo es usado como un medio que permite entender e interpretar mejor el mecanismo que genera la serie temporal; en la predicción de corto plazo, el modelo anterior es usado para calcular $\{\hat{Y}_t\}_{t=T}^{T+k}$ aplicando las funciones $\hat{f}(\cdot)$ y $\hat{g}(\cdot)$ al conjunto de regresores $\{\mathbf{x}_t\}_{t=T}^{T+k}$ en un horizonte de predicción de k períodos de tiempo, y prima como objetivo alcanzar la mayor precisión entre la predicción y el valor real observado; finalmente, en la predicción de largo plazo se busca determinar las grandes tendencias en el comportamiento de Y_t .

Se dice que una serie temporal es lineal cuando $f(\cdot)$ es especificada como una combinación lineal de los regresores $\mathbf{x}_t$, y $g(\cdot)$ es definida como una constante; cualquier desviación de las dos condiciones anteriores, hace que la dinámica de la serie sea considerada como no lineal, y más específicamente, como no lineal en media y/o no lineal en

varianza. El problema en la modelación de series no lineales, es que la forma de no linealidad es desconocida y debe ser aproximada empíricamente a partir de los datos, y para ello se han presentado diversas aproximaciones en la literatura.

La propuesta de Tong (1978) consiste en aproximar la dinámica de la serie temporal usando modelos autoregresivos que son validos en zonas de comportamiento local lineal, las cuales son obtenidas dividiendo el dominio de una función $F(z_t)$ en $\{R_k\}_{k=1}^K$ regiones disyuntas, tal que el modelo final es escrito como:

$$\text{if } F(z_t) \in R_k \Rightarrow y_t = \hat{a}_0^{(k)} + \sum_{j=1}^U \hat{a}_j^{(k)} \cdot y_{t-j} + \nu_t^{(k)} \quad (3)$$

el cual recibe el nombre de SETAR (Self Exciting Threshold Autoregressive) de k regímenes. z_t es una variable (de decisión) que determina en que regimen local se encuentra la serie temporal estudiada. La determinación de la cantidad de regiones K , sus límites, y los valores de los parámetros se hace a partir de la serie $\{y_t\}_{t=1}^T$. Algunas de sus propiedades estadísticas son discutidas por Tong (1990, Sección 3.3.1.1)] y Chan (1993). Tsay (1989) discute un procedimiento para su especificación en el caso univariado, mientras que Tsay (1998) lo hace para el caso multivariado. Recientemente estos modelos han sido usados en el modelado y predicción de series temporales no lineales: Ettestol (2002) pronostica la demanda residencial de electricidad en Noruega; Feng y Liu (2002) pronóstica el producto interno bruto canadiense; Holmes y Maghrebi (2002) analizan la presencia de cambio de régimen en las relaciones entre las economías asiáticas y los mercados de intercambio extranjeros; Medeiros y Veiga (2002) analizan la volatilidad financiera de diversos mercados; Tsay (2002, Sección 4.1.2) discute su aplicación en el modelado de volatilidad cambiante en el tiempo, mientras que Tsay (2002, Sección 8.6) discute la detección de oportunidades de arbitraje.

Para el caso del modelo SETAR de dos regiones, Baccon y Watts (1971) y Maddala (1977) incorporan una función de transición suave entre regímenes, $F(\cdot)$, cuyo rango está restringido al intervalo $[0, 1]$, tal que la dinámica de la serie es modelada como:

$$y_t = [1 - F(z_t)] \cdot \left(a_0^{(1)} + \sum_{j=1}^{P_1} a_j^{(1)} y_{t-j} \right) + F(z_t) \cdot \left(a_0^{(2)} + \sum_{j=1}^{P_2} a_j^{(2)} y_{t-j} \right) + \nu_t \quad (4)$$

Este modelo se conoce con el nombre de STR (smooth transition regression). El modelo LSTR (logistic smooth

transition regression) es obtenido cuando $F(\cdot)$ es especificada como la función logística o $S(\cdot)$:

$$F(u; \gamma, c) = \frac{1}{1 + e^{-\gamma(u-c)}}, \gamma > 0 \quad (5)$$

Cuando γ tiende a infinito, el modelo se reduce a un SETAR donde el parámetro c es el límite entre las dos regiones en que se divide el dominio de la función de transición $F(z_t)$. Si $F(\cdot)$ es especificada como la función exponencial:

$$F(u; \gamma, c) = 1 - \exp[-\gamma(u - c)^2], \gamma > 0 \quad (6)$$

se obtiene un modelo ESTR (exponential smooth transition regression). Un problema de esta función de transición es que converge a un valor constante cuando γ tiende a cero; mientras que cuando tiende a infinito se reduce a un modelo SETAR.

Algunas de las propiedades de los modelos STR son discutidas por Tong (1990, sec. 3.3.3); su proceso de especificación es presentado por Granger y Teräsvirta (1993, sec. 4.2), Teräsvirta (1994) y Eitrheim y Teräsvirta (1996), y se basa en los siguientes pasos:

1. Se especifica un modelo autoregresivo de orden p :

$$y_t = \alpha'_t x_t + \nu_t \quad (7)$$

donde $x_t = (1, \tilde{x}_t)'$ y $\tilde{x} = (y_{t-1}, \dots, y_{t-p})'$, a partir del cual se especifican los modelos lineales para cada una de las regiones del modelo STR. El proceso de selección entre diferentes modelos autoregresivos tentativos, los cuales difieren en los rezagos de la variable explicada usados como regresores, se realiza usando un criterio de información tal como los propuestos por Akaike (1973) o Hannan y Quinn (1979).

2. A partir del modelo anterior, se procede a demostrar la no linealidad de la serie. Bajo la hipótesis nula, la dinámica de y_t puede ser representada por el modelo anterior, tal que los residuales estimados serán ruido blanco. Para determinar si existen evidencias a favor de un modelo no lineal, se contrasta dicha hipótesis nula versus la alternativa de que existe una no linealidad remanente en los residuales la cual puede ser capturada por un modelo no lineal. Ramsey (1969), Tsay (1986) y Keenan (1985), entre otros, proponen pruebas estadísticas para contrastar la hipótesis de linealidad versus un modelo no lineal que no es especificado. No obstante, el interes en este paso es demostrar que la dinámica de la serie puede ser representada de una forma más adecuada por un modelo STR, para alguna variable de transición z_t . Teräsvirta (1994) indican que la regresión auxiliar

$$y_t = \beta_* + \beta_0 \tilde{x}_t + \beta_1 \tilde{x}_t z_t + \beta_2 \tilde{x}_t z_t^2 + \beta_3 \tilde{x}_t z_t^3 + \nu_t \quad (8)$$

permite determinar si hay evidencias a favor de un modelo STR (Prueba F_L). La serie es lineal si $\beta_1 = \beta_2 = \beta_3 = 0$, lo cual se demuestra a partir de la técnica de Multiplicadores de Lagrange.

3. Si en el paso anterior se demuestra la no linealidad, se debe determinar cual es la variable de decisión z_t y el tipo de función de transición (Exponencial o Logística). El proceso de selección usado por Granger y Teräsvirta (1993) se basa en realizar las siguientes pruebas secuenciales:

$$\begin{aligned} F_4 : \beta_3 &= 0 \\ F_3 : \beta_2 &= 0 | \beta_3 = 0 \\ F_2 : \beta_1 &= 0 | \beta_3 = \beta_2 = 0 \end{aligned} \quad (9)$$

a partir de las cuales se aplica la siguiente regla de decisión: si el valor crítico de F_3 es el menor de los tres, se selecciona un modelo ESTR; y un modelo LSTR en cualquier otro caso.

4. Finalmente, se realizan los estimados de máxima verosimilitud de los parámetros del modelo especificado en el punto anterior, así como las correspondientes pruebas de diagnóstico de los residuales.

3 DESCRIPCIÓN DE LOS DATOS

Los datos con que se elaboró este estudio corresponden a los aportes mensuales totales en energía que recibió el sistema hidroeléctrico de generación colombiano entre AGO95 y SEP04, los cuales están disponibles en el sistema Neón, y que recogen las características de la hidrología desde la creación de la Bolsa de Energía. Para garantizar que el modelo utilizado siempre entrega valores positivos, se realizó la modelación sobre el logaritmo natural de la serie real, cuya gráfica es presentada en la Figura 1. Las propiedades estadísticas de la serie son resumidas en sus autocorrelogramas simple y parcial, su espectro de energía y su histograma, los cuales son presentados en la Figura 2.

La serie de aportes está caracterizada por una componente estacional asociada al ciclo hidrológico intra-anual, que usualmente presenta su valor más bajo en los meses de enero o febrero, y su valor más alto en mayo, junio o julio. Se considera que la estación de verano tiene una duración de 5 meses y que se presenta entre el 1 de diciembre de cada año y el 31 de abril del año siguiente. Los meses restantes son considerados como pertenecientes a la estación de invierno.

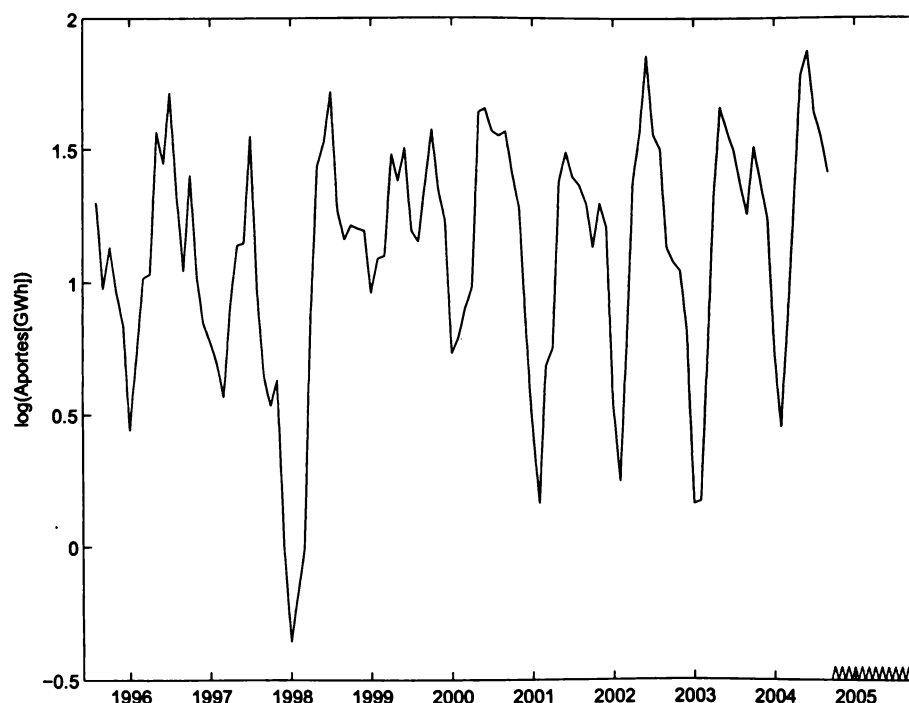


Figura 1: Aportes hidrológicos mensuales al SIN en GWh

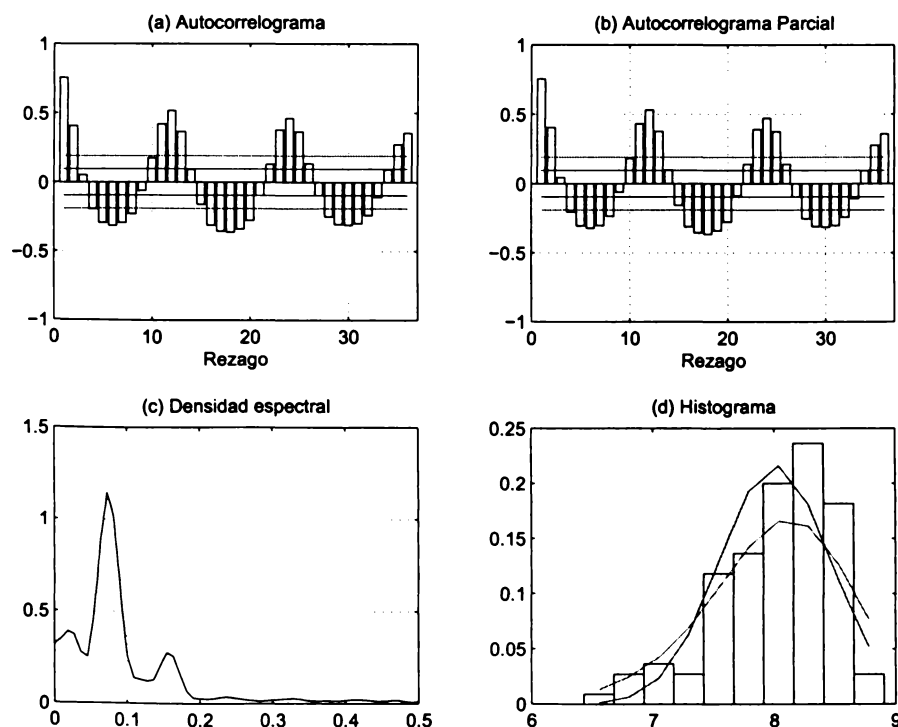


Figura 2: Propiedades estadísticas del logaritmo de la serie de aportes hidrológicos mensuales al SIN

Ha sido ampliamente aceptado, que la hidrología colombiana es afectada por el ENSO y sus fases extremas conocidas como los fenómenos de El Niño y La Niña, los cuales son cuantificados a través de las anomalías en la superficie del Océano Pacífico [véase a Poveda (2004) y las referencias dentro de él]; la NOAA ofrece la siguiente clasificación de dichos fenómenos, considerando como periodos de Niña aquellos cuya anomalía de temperatura calculada sobre un promedio móvil de tres meses esté por debajo de -0.5°C , y como periodos de Niño aquellos con una anomalía superior a los 0.5°C . Durante el periodo de análisis se registraron los siguientes eventos: La Niña SEP95-FEB96 (-0.7°C), El Niño MAY97-ABR98 (1.9°C), La Niña JUL98-JUN00 (-1.1°C) y OCT00-FEB01 (-0.6°C), El Niño MAY02-MAR03 (1.0°C) y JUL04-FEB05 (0.8°C). Los valores entre paréntesis corresponden al promedio de la anomalía registrada durante el fenómeno.

4 MODELADO E INTERPRETACIÓN DE LOS RESULTADOS

En esta sección se realiza la caracterización de la serie de aportes totales al sistema hidroeléctrico interconectado nacional colombiano, y se reportan los resultados obtenidos para las pruebas de no linealidad, la estimación del modelo final y el proceso de diagnóstico. Toda la información disponible es usada para la estimación de los modelos, ya que se busca encontrar una representación matemática que permita reproducir con la mayor fidelidad posibles las propiedades estadísticas de la serie real, y no se busca realizar un modelo de predicción, en donde prima la precisión alcanzada ¹.

El modelo autoregresivo inicial fue obtenido a partir de la búsqueda realizada sobre todos los modelos lineales autoregresivos que incluyen uno o más de los primeros 18 valores rezagados de la serie; la selección se realizó usando el criterio de Schwartz, obteniéndose como modelo

¹En este trabajo se realiza una comparación entre los resultados obtenidos usando un modelo lineal que usa los regresores que el modelo STR final; tal como lo indican Granger y Teräsvirta (1993), usar un periodo de validación para compararlos en términos de la precisión de su pronóstico extrapolativo no es adecuado, ya que el modelo STR podría tener un desempeño peor que el modelo lineal porque simplemente las características no lineales de la serie no están presentes durante dicho periodo.

²Los criterios de Akaike (1973) y Hannan y Quinn (1979) indican que los modelos iniciales serían un AR(11) y un AR(14) respectivamente; sin embargo, ello conduce a una gran cantidad de parámetros finales en relación al tamaño de la muestra de datos, y adicionalmente, todos los modelos STR obtenidos a partir de dicha especificación inicial fueron inestables por lo que debieron ser rechazados

más parsimonioso un $AR(3)^2$. No obstante, durante la fase posterior de diagnóstico del modelo no lineal se encontraron correlaciones residuales para los rezagos 11 y 14. Si se adiciona el rezago 11 dentro del conjunto de regresores, el modelo final obtenido es inestable (puede diverger a infinito) por lo que debe ser rechazado. El modelo STR final usa como regresores los rezagos 1-3 y 14, los cuales son considerados dentro del modelo lineal inicial para efectos de comparación posterior. Así el modelo lineal estimado fue:

$$y_t = \underset{(1.80)}{3.61} + \underset{(0.18)}{1.04}y_{t-1} - \underset{(0.19)}{0.18}y_{t-2} - \underset{(0.17)}{0.16}y_{t-3} - \underset{(0.10)}{0.15}y_{t-14} + \epsilon_t \tag{10}$$

$$\sigma_\epsilon^2 = 0.0645, \log L = -4.63, R^2 = 0.69, MAD = 0.20, LM_{ARCH} = 81.36 (0.00), BP = 40.11 (0.02), LB = 48.56 (0.00), ML = 15.89 (0.20), LJB = 2.42 (0.30), AK = 0.20, HQ = 0.25, SC=0.33, U=0.79$$

Los estadísticos reportados son los siguientes: σ_ϵ^2 es la varianza de los residuales; $\log L$ es el logaritmo de la función de verosimilitud; R^2 es el coeficiente de determinación; MAD es la desviación media absoluta; LM_{ARCH} es la prueba de heterocedasticidad de Engle (1982); BP , LB y ML son

los valores más críticos de los estadísticos de Box y Pierce (1970), Ljung y Box (1978) y McLeod y Li (1983). estimados sobre los primeros 24 rezagos de los residuales; LJB es la prueba de normalidad de los residuales de Lumnicki (1961) y Jarque y Bera (1980); AK , HQ y SC son los criterios de información de Akaike (1973), Hannan y Quinn (1979) y Schwartz respectivamente. U es el estadístico de Theil (1966). Los valores críticos de los estadísticos aparecen entre paréntesis. Las cantidades bajo los parámetros del modelo son las desviaciones estándar estimadas. El pronóstico un paso adelante, y las propiedades estadísticas de los residuales son presentados en las Figuras 3(a) y 4 respectivamente. Debe anotarse que los residuales del modelo presentan una correlación serial importante para el rezago 12, lo que es de esperarse dado el patrón estacional anual de la hidrología³.

Las pruebas de no linealidad fueron realizadas sobre el modelo anterior, considerando como variables de transición los rezagos 1 al 18; los resultados de las pruebas son presentados en la Tabla 1. Dicha Tabla contiene la variable de transición y el estadístico usado en la prueba de linealidad; cuando la hipótesis de linealidad es rechazada, la tabla también contiene los estadísticos para determinar el tipo de la función de transición. Los valores críticos para todos los estadísticos son presentados en paréntesis.

Tabla 1: Resultados de las pruebas de linealidad y de selección del modelo

Rezago	F_L	F_4	F_3	F_2	Tipo
1	1.40 (0.173)				
2	1.29 (0.232)				
3	1.25 (0.254)				
4	0.91 (0.540)				
5	1.00 (0.461)				
6	1.05 (0.414)				
7	1.51 (0.141)				
8	0.88 (0.567)				
9	1.19 (0.309)				
10	1.35 (0.209)				
11	2.92 (0.002)	0.47 (0.756)	0.48 (0.751)	8.24 (0.000)	L
12	3.85 (0.000)	2.16 (0.081)	1.13 (0.347)	7.68 (0.000)	L
13	0.97 (0.483)				
14	1.29 (0.233)				
15	0.70 (0.746)				
16	0.94 (0.509)				
17	1.98 (0.038)	1.90 (0.118)	1.12 (0.352)	2.71 (0.036)	L
18	1.59 (0.113)				

³No obstante, el modelo lineal obtenido no cumple con la metodología de Box y Jenkins (1970), ya que la serie presenta una integración estacional de periodo 12, la cual debería ser removida; sin embargo, este paso remueve los patrones cíclicos de la serie original, y lo que se pretende es que el modelo no lineal obtenido, capture dichos comportamientos

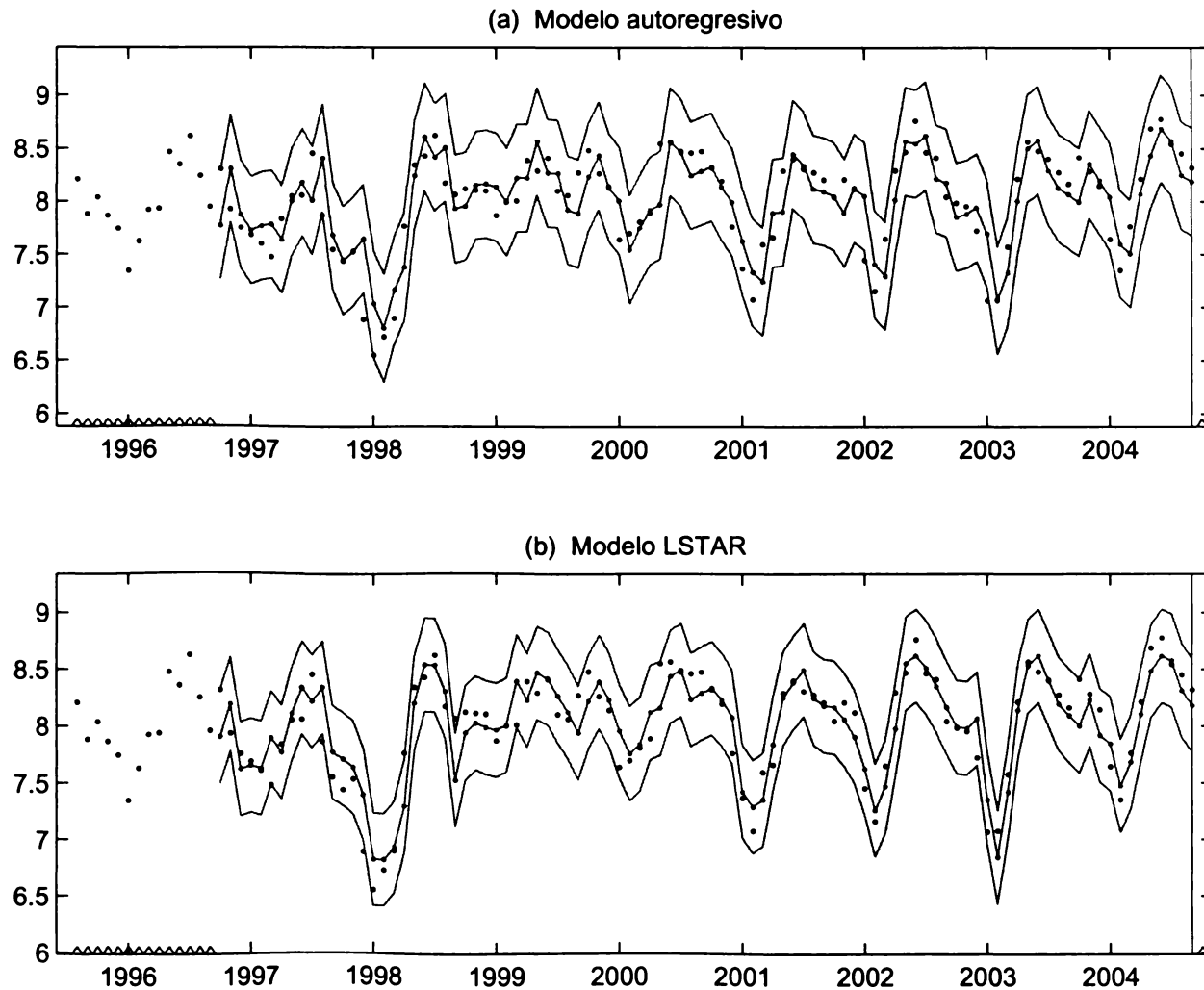


Figura 3: Pronóstico un paso adelante

Particularmente, la linealidad fue rechazada para todos los rezagos considerados, excepto para los 11, 12 y 17 con un nivel de confianza del 5%; en todos los casos se seleccionó la función logística para realizar la transición entre regímenes. El modelo final fue seleccionado a partir del valor más crítico (mayor evidencia) en contra de la hipótesis nula de linealidad:

$$y_t = [1 - F(y_{t-12})] \left[\begin{matrix} 4.53 \\ (0.48) \end{matrix} + \begin{matrix} 0.97 y_{t-1} \\ (0.15) \end{matrix} - \begin{matrix} 0.51 y_{t-2} \\ (0.52) \end{matrix} \right. \\ \left. + \begin{matrix} 0.36 y_{t-3} \\ (0.65) \end{matrix} - \begin{matrix} 0.43 y_{t-14} \\ (0.12) \end{matrix} \right] \\ + F(y_{t-12}) \left[\begin{matrix} 5.85 \\ (0.47) \end{matrix} + \begin{matrix} 0.51 y_{t-1} \\ (0.52) \end{matrix} + \begin{matrix} 0.22 y_{t-2} \\ (0.61) \end{matrix} \right. \\ \left. - \begin{matrix} 0.35 y_{t-3} \\ (0.38) \end{matrix} - \begin{matrix} 0.10 y_{t-14} \\ (0.67) \end{matrix} \right] + \epsilon_t \quad (11)$$

$$\sigma_{nl}/\sigma_l = 0.81, \sigma_\epsilon^2 = 0.04, \log L = 15.53, R^2 = 0.80, \text{MAD} = 0.16, LM_{ARCH} = 87.01 (0.00), \\ BP = 1.80 (0.40), LB = 1.87 (0.39), ML = 20.24 (0.09), LJB = 0.04 (0.98), AK = -0.12, \\ SC = 0.15, HQ = 0.00, U = 0.64$$

cuya función de transición es especificada como:

$$F(u) = (1 + \exp(-\frac{2.10(u - 7.87)}{(0.34)\sigma_u}))^{-1} \quad (12)$$

La relación σ_{nl}/σ_l muestra la reducción en la desviación estándar de los residuales entre el modelo LSTAR y un modelo AR que usa los mismos rezagos; este resultado indica que la capacidad no lineal del modelo permite caracterizar mejor regiones que su competidor lineal.

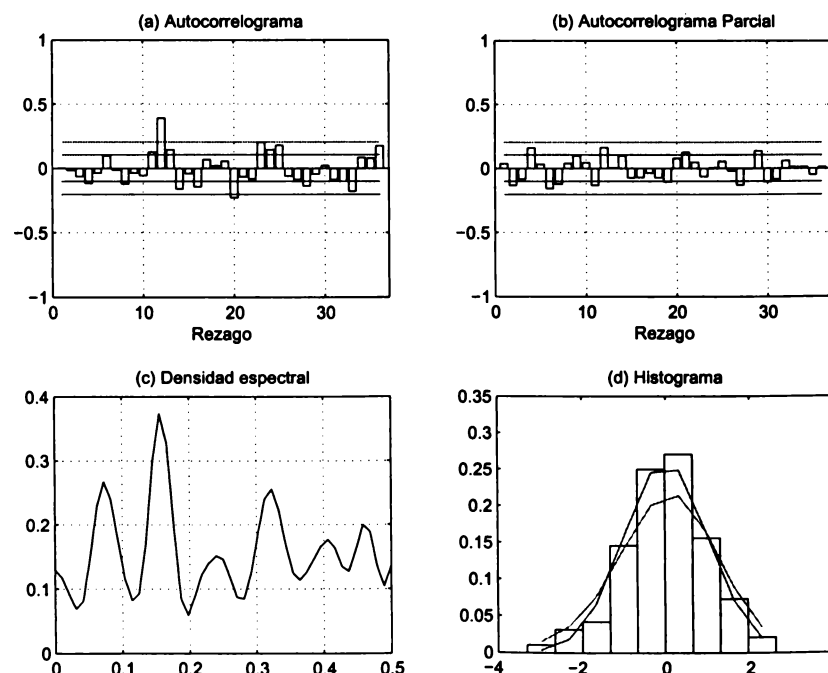


Figura 4: Propiedades estadísticas de los residuos del modelo AR

Las propiedades estadísticas de los residuos son resumidas en la Figura 5. La forma de la función de transición entre regímenes es presentada en la Figura 6, de donde puede deducirse que el modelo obtenido para $F = 0$ representa el comportamiento en la estación de verano del año anterior (dado que la variable de transición es y_{t-12}), mientras que $F = 1$ representa la estación de invierno. Es importante anotar, como puede observarse en la Figura 6a, que durante el invierno de 1998 hay un cambio muy rápido al régimen de verano (antes de finalizar el invierno), de tal forma, que los aportes son menores a lo esperado.

La inspección visual de la predicción un paso adelante durante el periodo de estimación (véase la Figura 3), indica que tanto el modelo lineal como el LSTAR tienen un comportamiento muy similar, aunque este último pronostica mejor los caudales bajos.

El modelo obtenido presenta un ciclo límite de 12 meses, tal como es de esperarse dada la estacionalidad de la hidrología en Colombia, aunque tiene un comportamiento caótico debido a su dinámica interna, ya que no converge a valores estables aunque la componente aleatoria ha sido fijada en cero. Su comportamiento de largo plazo puede visualizarse en la Figura 8 donde se presenta el pronóstico determinístico para un horizonte de 100 meses. Esta característica puede observarse mejor en las trayectorias presentadas en la Figura 7, en donde puede apreciarse claramente el comportamiento caótico del modelo STAR obtenido, el cual converge a un ciclo límite, mientras que el modelo

lineal converge un punto de equilibrio estable, que corresponde aproximadamente al valor medio de la serie.

Para establecer el comportamiento dinámico del modelo, se estimaron las raíces de los polinomios característicos para $F = 0$ y $F = 1$, las cuales son presentadas en la Tabla 2. En todos los casos, los módulos de las raíces caen fuera del círculo unitario indicando inestabilidad de cada uno de los modelos lineales obtenidos.

Como una verificación final de la calidad del modelo para representar la serie analizada, se generó una serie sintética de 10.000 valores usando bootstrapping (Efron y Tibshirani, 1993), y posteriormente se estimaron sus principales momentos. Los resultados obtenidos aparecen en la Tabla 3.

La estabilidad local del modelo está determinada por el efecto que produce en el tiempo una perturbación aleatoria ocurrida en el instante t , y para su análisis se utiliza la Función Generalizada de Respuesta al Impulso (GIR) (Koop, Pesaran y Potter, 1996), la cual mide el cambio que se produce en el pronóstico determinístico para el periodo $t + k$, cuando se aplica una perturbación u_t conocida en el instante t al pronóstico de $\hat{Y}_t$ con la información conocida en $t - 1$, X_{t-1} ; dicha función es definida como:

$$GIR(k, \mathbf{x}_{t-1}, u_t) = E[\hat{Y}_{t+k} | X_t = \mathbf{x}_t, u_t] - E[\hat{Y}_{t+k} | X_t = \mathbf{x}_t, u_t = 0] \quad (13)$$

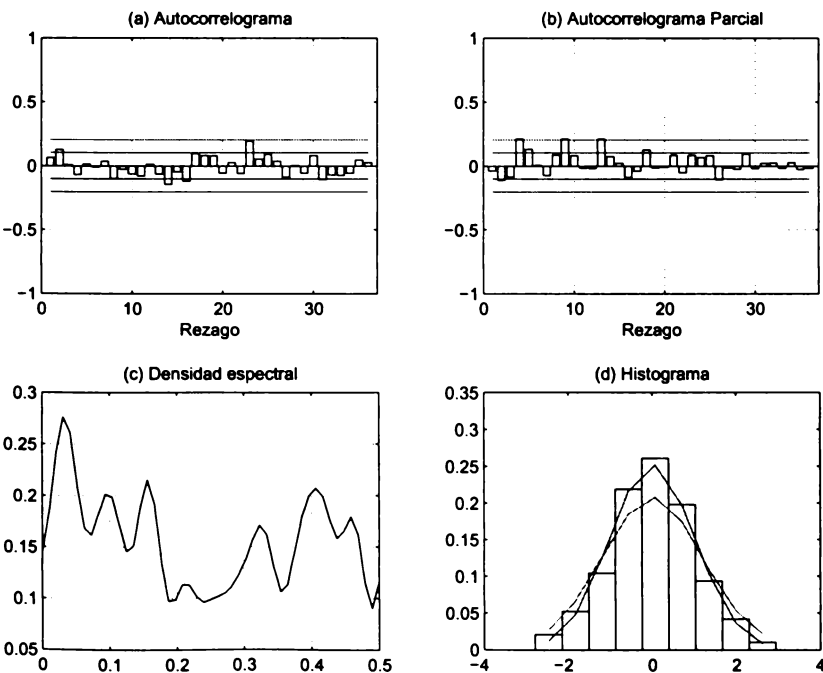


Figura 5: Propiedades estadísticas de los residuales del modelo LSTAR

Tabla 2: Raíces de los polinomios característicos

Regimen F = 0			Regimen F = 1		
Raíz	Módulo	Periodo	Raíz	Módulo	Periodo
1.2014	1.2014		1.3432	1.3432	
-1.1242	1.1242		-1.3495	1.3495	
$1.0740 \pm 0.5272i$	1.1964	13.77	$1.2178 \pm 0.5827i$	1.3500	14.08
$0.7354 \pm 0.9351i$	1.1896	6.95	$0.8396 \pm 1.0657i$	1.3567	6.95
$0.2600 \pm 1.1645i$	1.1931	4.65	$0.2766 \pm 1.3177i$	1.3464	4.61
$-0.2825 \pm 1.1651i$	1.1988	3.47	$-1.1989 \pm 0.5937i$	1.3379	2.34
$-0.7686 \pm 0.9071i$	1.1889	2.93	$-0.8175 \pm 1.0328i$	1.3172	2.80
$-1.0570 \pm 0.4637i$	1.1543	2.30	$-0.3143 \pm 1.2848i$	1.3227	3.47

y se considera que el modelo obtenido es inestable cuando la perturbación aleatoria se amplifica en el tiempo. En el caso no lineal, Koop et al. (1996) han mostrado como que la respuesta producida es dependiente del valor actual de la serie temporal, la magnitud de la perturbación (a la que se responde de forma no lineal), y a su signo; consecuentemente, el resultado de evaluar la función GIR es una cantidad aleatoria, por lo que para su representación

Tabla 3: Propiedades estadísticas de las distribuciones de probabilidades obtenidas usando Bootstrap

Momento	Serie	LSTAR	AR
Media	8.0080	7.9067	7.9939
Varianza	0.2226	0.1931	0.1888
Curtosis	3.4533	2.8112	3.0762
Asimetria	-0.9106	-0.2351	-0.0335
Máximo	8.7790	9.2715	9.6826
Mínimo	6.5512	6.0821	6.3972

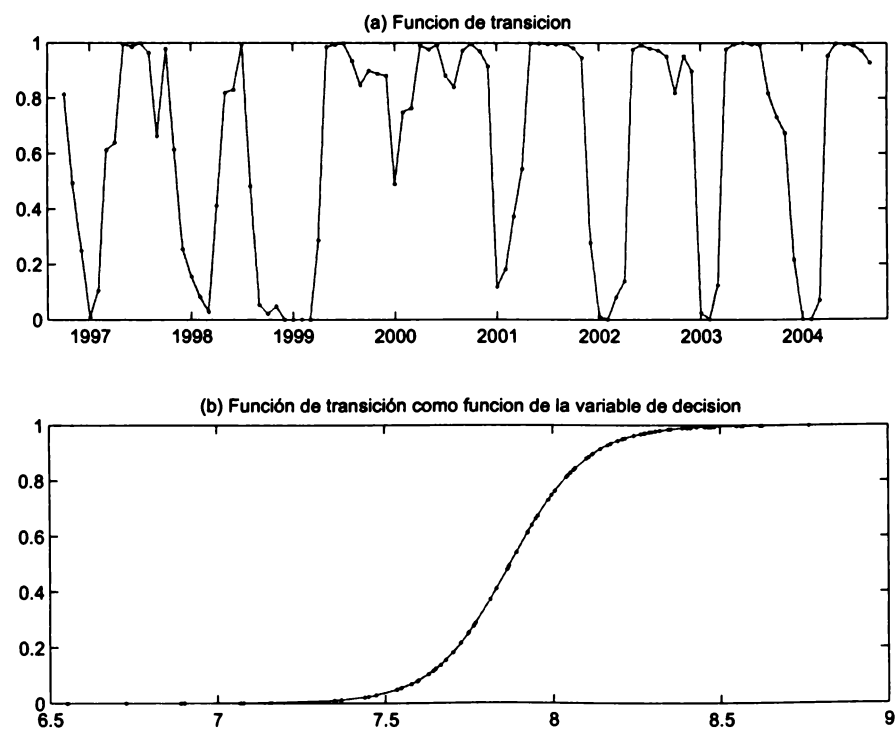


Figura 6: Función de transición entre regímenes

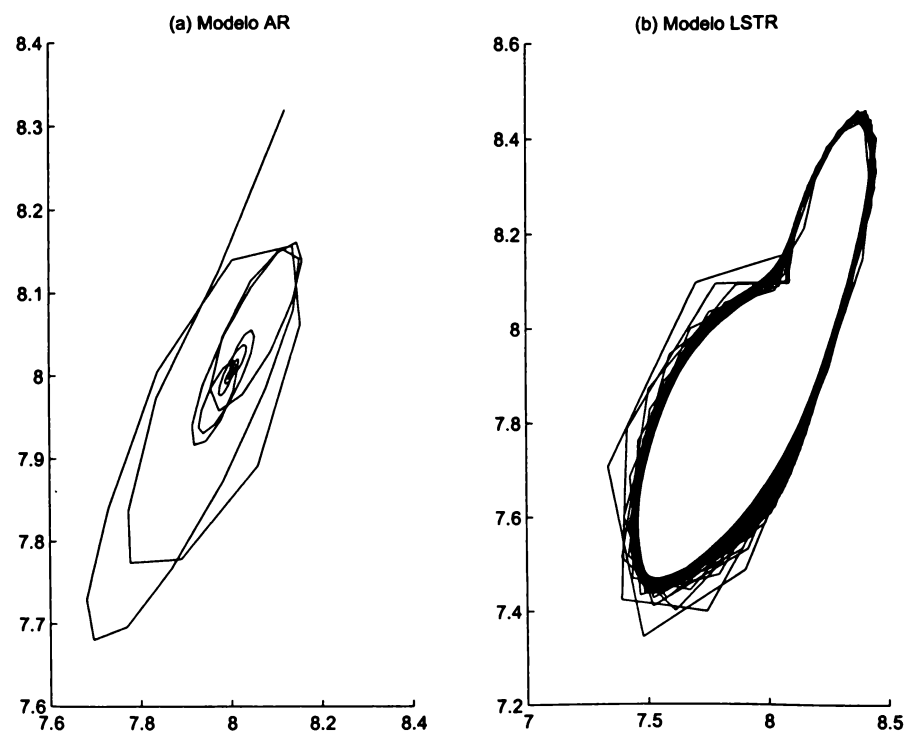


Figura 7: Trayectorias de convergencia para el pronóstico determinístico de los modelos analizados

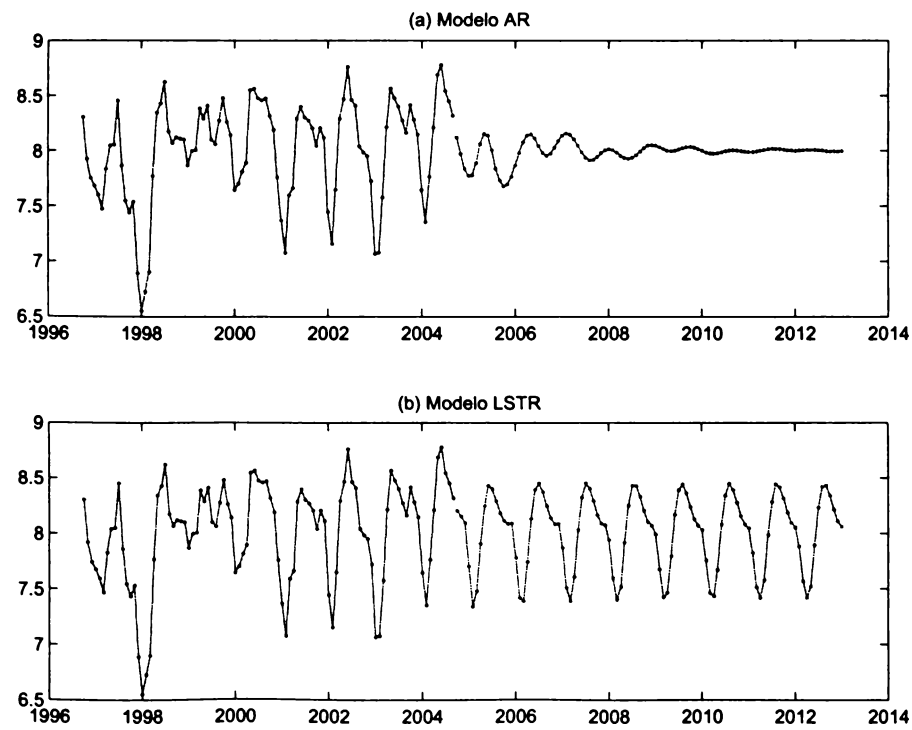


Figura 8: Pronóstico determinístico

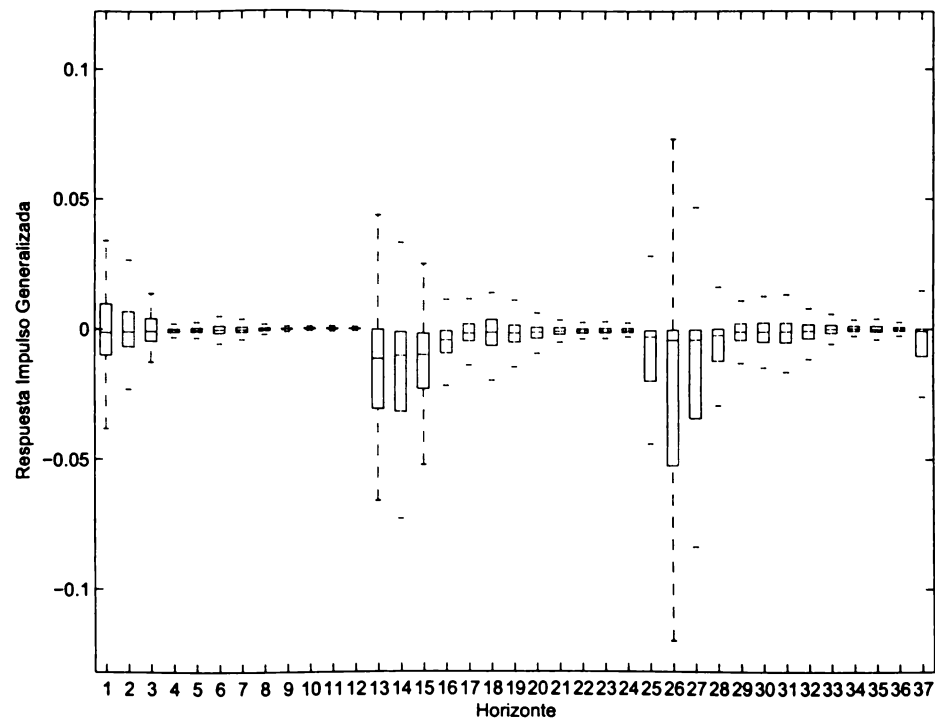


Figura 9: Función Generalizada de Respuesta al Impulso

se utilizan gráficos de caja. La función GIR obtenida es presentada en la Figura 9; ella indica que el impacto de una perturbación aleatoria es asimilado rápidamente en el corto plazo, mientras que su efecto revive 12 y 24 meses después, que corresponden a múltiplos del rezago de la variable de decisión. Ello indica que el modelo podría usarse en la predicción de corto plazo.

5 CONCLUSIONES

El análisis realizado de la serie de aportes hidrológicos al sistema hidrotérmico de generación, expresados en energía, indica un comportamiento no lineal inducido por la presencia de eventos hidrológicos extremos, que rompe con la continuidad del ciclo hidrológico anual. El modelo no lineal estimado presenta ventajas en este aspecto, ya que permite realizar una predicción más acertada durante dichos fenómenos. El modelo obtenido presenta un comportamiento caótico tal como se evidencia en la predicción determinística de largo plazo, donde no se presenta convergencia. No obstante, el modelo puede ser usado para construir una predicción de los aportes hidrológicos en un horizonte de varios meses.

BIBLIOGRAFÍA

- Akaike, H. (1973), Information theory and an extension of the maximum likelihood principle, in B. Petrov y F. Csaki, eds, '2nd International Symposium on Information Theory', Akademia Kiado: Budapest, pp. 267-281.
- Baccon, D. y Watts, D. (1971), 'Estimating the transition between two intersecting straight lines', *Biometrika* 58, 525-534.
- Box, G. E. P. y Jenkins, G. M. (1970), *Time Series Analysis: Forecasting and Control*, Holden-Day Inc.
- Box, G. y Pierce, D. (1970), 'Distribution of residual correlations in autoregressive-integrated moving average time series models', *Journal of American Statistical Association* (65), 1509-1526.
- Chan, K. (1993), 'Consistency and limiting distribution of the least squares estimator of a continuous autoregressive model', *The annals of statistics* (21), 520-533.
- Efron, B. y Tibshirani, R. (1993), *An Introduction to the Bootstrap*, Chapman & Hall.
- Eitrheim, O. y Teräsvirta, T. (1996), 'Testing the adequacy of smooth transition autoregressive models', *Journal of Econometrics* (74), 59-76.
- Engle, R. (1982), 'Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflations', *Econometrica* (50), 987-1007.
- Ettestol, I. (2002), Estimating residential demand for electricity with smooth transition regression. NTNU, Trondheim, Norway.
- Feng, H. y Liu, J. (2002), A setar model for canadian gdp: Non-linearities and forecast comparisons, Working paper - issn 1486-6441, University of Victoria.
- Granger, C. y Teräsvirta, T. (1993), *Modeling Nonlinear Economic Relationships*, Oxford University Press.
- Hannan, E. y Quinn, B. (1979), 'The determination of the order of an autoregression', *Journal of Royal Statistical Society Series B* (41), 190-195.
- Holmes, M. y Maghrebi, N. (2002), Non-linearities, regime switching and the relationships between asian equity and foreign exchange markets, Technical report.
- Jarque, C. y Bera, A. (1980), 'Efficient tests for normality, homoscedasticity and serial independence of regression residuals', *Economic Letters* (6), 255-259.
- Keenan, D. (1985), 'A turkey non-additivity-type test for time series nonlinearity', *Biometrika* (72), 39-44.
- Koop, G., Pesaran, M. y Potter, S. (1996), 'Impulse response analysis in nonlinear multivariate models', *Journal of Econometrics* (74), 119-147.
- Ljung, L. y Box, G. (1978), 'On a measure of lack of fit in time series models', *Biometrika* (66), 67-72.
- Lumnicki, Z. (1961), 'Test for departure from normality in the case of linear stochastic processes', *Metrika* (4), 37-62.
- Maddala, G. S. (1977), *Econometrics*, McGraw-Hill.
- McLeod, A. y Li, W. (1983), 'Diagnostic checking arma time series models using squared residual autocorrelations', *Journal of Time Series Analysis* (4), 269-273.
- Medeiros, M. y Veiga, A. (2002), Are there multiple regimes in financial volatility?, Technical report, Pontificia Universidad Catolica do Rio de Janeiro.
- Poveda, G. (2004), 'La hidroclimatología en colombia: una síntesis desde la escala inter-decanal hasta la escala diurna', *Revista de la Academia Colombiana de Ciencias* 28(107), 201-222.

- Pulgarín, A., Smith, R. y Poveda, G. (2001), Predicción del precio de la energía eléctrica con un modelo de redes neuronales y usando variables macroclimáticas, in 'XIV Seminario Nacional de Hidráulica e Hidrología'.
- Ramsey, J. (1969), 'Tests for specification error in classical linear least-squares regression analysis', *Journal of the Royal Statistical Society B* (31), 350–371.
- Smith, R., Velez, J., Velásquez, J., Ceballos, A., Correa, P., Goetz, C., Salazar, L. y Zapata, E. (2004), 'Modelos de predicción de caudales mensuales para el sector eléctrico colombiano', *Avances en Recursos Hidráulicos* (11).
- Teräsvirta, T. (1994), 'Specification, estimation, and evaluation of smooth transition autoregressive models', *Journal of the American Statistical Association* (89), 208–218.
- Theil, H. (1966), *Applied Economic Forecasting*, Rand McNally.
- Tong, G. (1978), On a threshold model, in C. Chen, ed., 'Pattern Recognition and Signal Processing', Sijhoff & Noordhoff, Amsterdam.
- Tong, H. (1990), *Non-linear Time Series, a dynamical system approach*, Oxford Statistical Science Series, Clarendon Press Oxford.
- Tsay, R. (1986), 'Nonlinearity tests for time series', *Biometrika* (73), 461–466.
- Tsay, R. (1989), 'Testing and modeling threshold autoregressive process', *Journal of the American Statistical Association* (84), 231–240.
- Tsay, R. (1998), 'Testing and modeling multivariate threshold models', *Journal of the American Statistical Association* (93), 1188–1202.
- Tsay, R. (2002), *Analysis of Financial Time Series*, John Wiley & Sons, Inc.
- Velásquez, J. D. y Dyner, I. (2001), Pronóstico de precios de bolsa de electricidad usando un modelo de redes neuronales artificiales, in 'EITI-2001', Universidad Nacional de Colombia.
- Zapata, E., Velásquez, J. y Smith, R. (2004), 'Modelamiento de series de caudal usando anfis', *Avances en Recursos Hidráulicos* (11).
-

