

MODELOS DE PREDICCIÓN DE CAUDALES PARA EL SECTOR ELÉCTRICO COLOMBIANO. PARTE I: MODELOS MENSUALES Y MODELO SEMANAL.

Oscar J. Mesa, Ricardo A. Smith, José E. Salazar y Luis F. Carvajal
Posgrado en Aprovechamiento de los Recursos Hidráulicos
Facultad de Minas, Universidad Nacional de Colombia

1. INTRODUCCION

Interconexión Eléctrica S.A. contrató con el Posgrado en Aprovechamiento de Recursos Hidráulicos de la Facultad de Minas de la Universidad Nacional el proyecto de investigación "Pronóstico Hidrológico" con el objeto de desarrollar e implementar modelos de predicción en las cuencas de los embalses de San Lorenzo, Chivor y Salvajina a diferentes niveles temporales, incluyendo, de ser necesario, la influencia de variables climáticas de gran escala. Este artículo presenta el resumen de los trabajos realizados en el marco de este proyecto a nivel mensual y semanal.

La enseñanza de la crisis de abastecimiento sufrida por el sector eléctrico colombiano en 1992, con respecto a la hidrología es clara. Los modelos autorregresivos que se usaron durante mucho tiempo para la planeación de la operación y expansión cumplieron su ciclo y requieren ser reemplazados por una nueva generación que tenga en cuenta las relaciones causales existentes entre el clima colombiano y ciertas variables macroclimáticas como el fenómeno del

Niño- Oscilación del Sur (ENOS). Esta relación abre la esperanza de poder disponer de mejores herramientas predictivas lo que a su vez redundará en mejores modelos de operación y de planeación. Este trabajo es un paso de los muchos que hay que dar en esa dirección.

En la época inicial del uso de la hidrología estocástica para el planeamiento de los recursos hidráulicos se dieron discusiones relativas a la duración e intensidad de los períodos críticos que hoy son pertinentes. La práctica anterior usaba el llamado período histórico crítico como elemento de diseño. Las trazas sintéticas generadas por los modelos markovianos no necesariamente reproducen las características de estos eventos extremos (Millán, 1972). Algo semejante se puede decir con relación al fenómeno de Hurst (Mesa y Poveda, 1993). Sin embargo estas discusiones se resolvieron más que todo con criterios pragmáticos que con claridad teórica (Klemes, 1974; Pereira et al., 1984).

En estudios anteriores (EPM - U NAL, 1994) se ha confirmado la relación causal

entre la hidrología colombiana y los eventos ENOS. Tanto las observaciones disponibles como el análisis meteorológico y climático soportan esta conclusión. Sin embargo, esa relación no es lineal y simple; tal vez una afirmación más correcta sea decir que la variabilidad del clima colombiano es parte integral de un ciclo aperiódico en la interacción océano-atmósfera-continente en la zona ecuatorial, caracterizada por una escala espacial global y una baja frecuencia (2 a 5 años) y cuya principal manifestación es el ENOS. Aunque es bien reconocido el efecto de la oscilación ENOS sobre las lluvias, no se conoce el hecho de que esta relación sea recíproca y que los efectos energéticos de la evapotranspiración, la condensación de vapor y la precipitación tropical son a su vez retroalimentadas y tiene efectos no despreciables sobre el resto del clima (Simpson, 1988).

La ilusión de una relación determinística simple que permita excelentes predicciones de las lluvias y los caudales en Colombia con base en observaciones meteorológicas u oceanográficas es muy optimista. La estadística tradicional del análisis lineal múltiple da resultados sólo aceptables (correlaciones del orden 0.7 o menos) y que no concuerdan con las expectativas tradicionales de una relación causal. La relación de esta dificultad con las predicciones reside el llamado caos. Tanto en meteorología como en climatología hay, hoy en día, suficiente claridad y aceptación sobre los límites de modelos dinámicos, con buena base física, matemática, computacional y observacional. El clima es un sistema dinámico altamente no lineal, prototipo de la turbulencia y de la incapacidad predictiva. En meteorología, en los países más avanzados, las predicciones son confiables para períodos muy cortos, del orden de una semana. Entre todas las variables, la precipitación es la más difícil de predecir (Lorenz, 1969).

Sin embargo, la posibilidad de pasar de modelos totalmente estocásticos, cuyo único contacto con la realidad hidrológica es a través de los valores medios, sus varianzas y covarianzas, a modelos con base física, con ecuaciones bien definidas y con el apoyo de una amplia red de observación, abre perspectivas importantísimas. De manera inmediata se puede acceder a los modelos océano-atmósfera del Pacífico, que de manera simplificada (Cane et al., 1986) o más detallada (Lattif et al, 1993) producen predicciones a mediano plazo (3 meses) y largo plazo (hasta 18 meses) que sin ser perfectos son superiores al mero azar y que compiten con las herramientas estadísticas más sofisticadas (Lattif et al., 1994). Estos modelos sirven de manera inmediata como alertas que señalan el comportamiento futuro de las lluvias y los caudales en Colombia. En etapas futuras hay que desarrollar modelos, para estudiar directamente el clima colombiano.

A nivel mundial, la predicción a mediano y largo plazo (variabilidad interestacional e interanual) está apenas iniciándose. Los fenómenos tropicales, en particular los fenómenos ENOS, son mucho más promisorios que los correspondientes a las latitudes medias. Esto es paradójico si se mira el desarrollo de la medición, y de la comunidad científica y se constata la falta de un marco conceptual, análogo a la teoría cuasi-geostrofica, que pueda ser usado para desarrollar una comprensión completa del movimiento de la atmósfera tropical (Holton, 1992; p. 368).

La razón se deriva de que sólo cerca al Ecuador son posibles las ondas Kelvin que producen oscilaciones de gran escala en el tiempo (varios años) y en el espacio. Esto ocurre tanto en la atmósfera como en el océano. En la interacción de estas ondas con las ondas Rossby y en la geografía de las

cuencas oceánicas y los continentes está, a grandes rasgos la explicación de la oscilación ENOS (Bjerknes, 1969). Sin embargo, el hecho de que la interacción océano-atmósfera sea tan primordial, y de que las comunidades científicas meteorológicas y oceanográficas hayan trabajado de manera independiente por muchos años, contribuyen a que el desarrollo de una teoría sobre el ENOS sea relativamente tardía y que todavía haya mucho que aprender. Una clara ilustración de las limitaciones en la capacidad predictiva actual del fenómeno ENOS se puede apreciar del seguimiento de las predicciones en los últimos años.

Las limitaciones de los modelos dinámicos abren posibilidades para modelos estocásticos y estadísticos. Hoy por ejemplo, para predicción del ENOS se emplean modelos de componentes principales, análisis espectral singular y análisis canónico con relativo éxito (Latif et al., 1994). Una característica clara de este tipo de modelos que los distingue de los modelos tradicionales usados en hidrología es su escala espacial y temporal. De manera semejante, en este trabajo se proponen modelos estocásticos que sin embargo representan de alguna manera la nueva concepción de la dinámica hidrológica. Estos modelos tienen una mejor base física que los autorregresivos tradicionales y son desde el punto de vista práctico tan versátiles como ellos para su empleo en predicciones hidrológicas en Colombia.

2. INFORMACION HIDROMETEOROLOGICA UTILIZADA

2.1 INFORMACION DE PRECIPITACION Y CAUDALES

Para la aplicación de los diferentes modelos se utilizó información de precipitación y

caudales a nivel diario con su correspondiente agregación a nivel mensual, especialmente para los modelos de resolución mensual. Esta información corresponde a los proyectos hidroeléctricos de las cuencas de Chivor, Salvajina y Jaguas, que son las aplicaciones objeto de este estudio.

En el proceso de selección de la información a utilizar en los diferentes modelos, se utilizaron las estaciones de precipitación que permiten un cubrimiento de la cuenca en toda su extensión, con datos de caudales a la salida de la cuenca. Las estaciones de precipitación seleccionadas se definieron con base en los criterios de longitud de registro, homogeneidad, cercanía entre estaciones y cubrimiento. Se destaca en la información de caudal el corto registro de la cuenca Jaguas y la consideración en la cuenca de Chivor de los caudales a la entrada del embalse la Esmeralda, conformados por la suma de los ríos Garagoa y Somondoco, sin tener en cuenta la quebrada Chivor y las desviaciones de los ríos Rucio, Negro y Tunjita.

2.2 INFORMACION CLIMATICA

La información climática utilizada proviene de la base de datos COADS (Climate Oceanic Atmospheric Data Set), la cual contiene las variables climáticas que a continuación se describen.

De las variables consideradas en la Tabla 1, se posee información desde enero de 1940 a diciembre de 1991 a nivel mensual, monitoreadas en diferentes regiones del Océano Pacífico y Atlántico, las cuales son definidas conjuntamente con su ubicación en la Tabla 2. Algunas de estas regiones (las más importantes) se presentan en la Figura 1.

Figura 1. Esquema de ubicación de las regiones N1, N3 y N4 (tomado de EPM-U NAL, 1994)

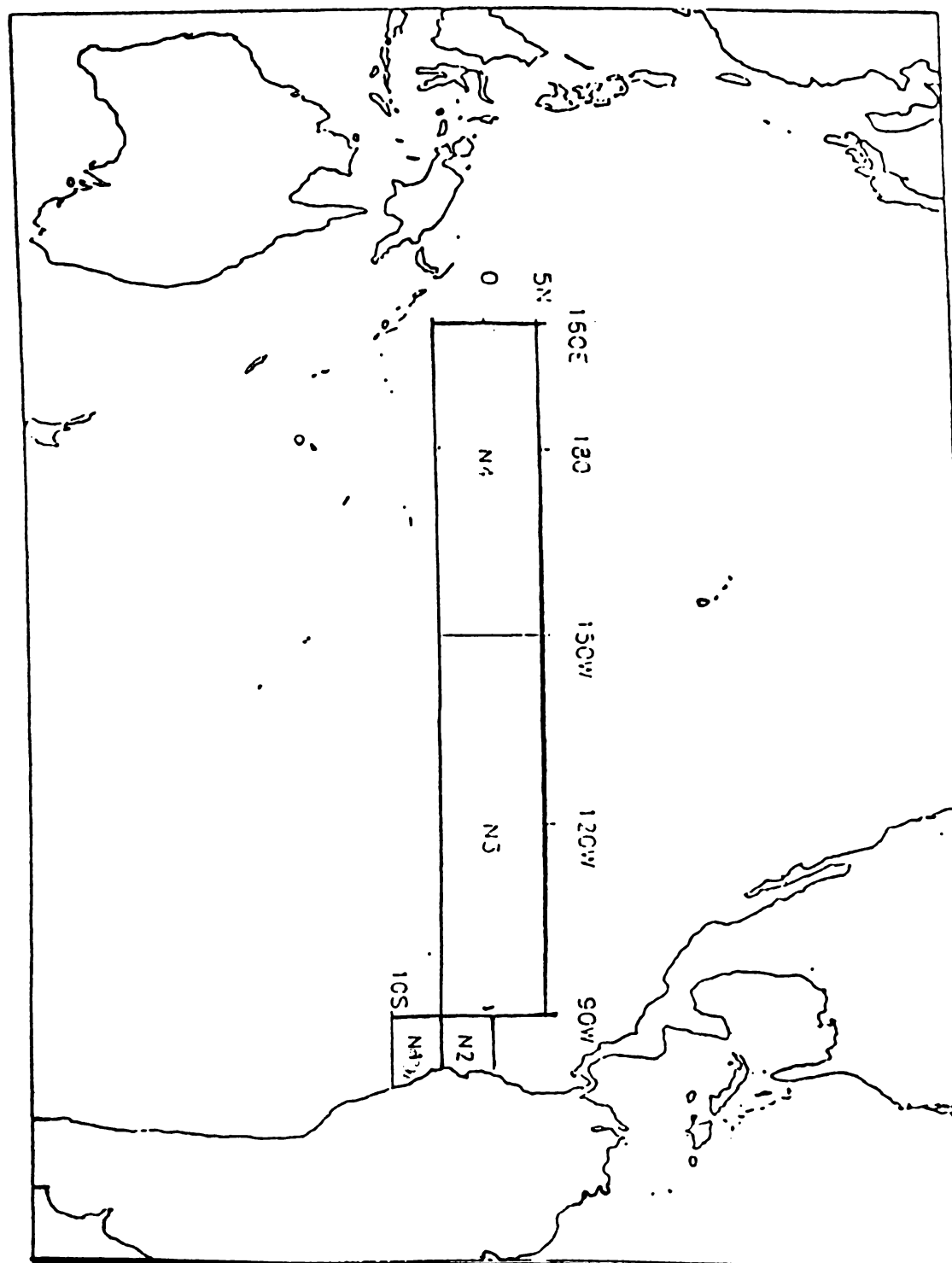


Tabla 1.
Variables meteorológicas empleadas (Base de datos COADS) (EPM-U NAL, 1994)

Variable	Unidades	Código
Temperatura superficial del mar	°C	SST
Humedad Específica	g/Kg	SHq
Índice de Oscilación del Sur (IOS: Presión atmosférica al nivel del mar Tahití-Darwin)	mb	SLP
Componente este de la velocidad del viento superficial	m/s	WEU
Componente norte de la velocidad del viento superficial	m/s	WNV

Tabla 2.
Definición de regiones (EPM-U NAL, 1994)

Región	Latitud	Longitud	Referencia
Niño 1,2 (N1)	0-10S	80W-90W	Rasmusson et al, 1982
Niño 3 (N3)	5N-5S	90-150W	Rasmusson et al, 1982
Niño 4 (N4)	5N-5S	150W- 160E	Rasmusson et al, 1982
Niño 5 (N4)	2-6N	90W- 170W	Hastenrath, 1990
Niño 6 (N6)	2N-6S	90-180W	Hastenrath, 1990
Niño 7 (N7)	6S-10S	110W- 150W	Hastenrath, 1990
Niño 8 (N8)	0-10N	78W-90W	Mesa, 1994
Atlánt. N Trop. (AN)	0-24N	20W-50W	Hastenrath, 1990
Atlánt. S Trop. (AS)	0-24S	10W-40W	Hastenrath, 1990
Caribe (CA)	12-16N	60W-80W	Mesa, 1994

3. PREDICCIÓN A NIVEL MENSUAL (TRES MESES)

Para la predicción a nivel mensual se proponen cuatro herramientas: el modelo de regresión lineal usando variables macroclimáticas, el modelo de Análisis Espectral Singular (AES), el modelo de redes neuronales de Retropropagación y el modelo autorregresivo dependiente del régimen (RAR). Los tres últimos métodos tienen la no linealidad como característica fundamental, con lo cual se busca capturar esta propiedad de la naturaleza.

Los modelos lineales autorregresivos (AR) han sido, hasta hace poco, los más ampliamente utilizados para la predicción de caudales. Una extensión de estos modelos son los modelos RAR (autorregresivos dependientes del régimen) consistentes en distintas ecuaciones AR cuyos parámetros dependen del régimen que ocupe una variable índice, en este caso una variable climática. Según se comprobó en recientes estudios (EPM-U.NAL 1994) en el país el uso del modelo RAR y modelos de regresión que incluyen variables macroclimáticas brinda más información que el uso de un modelo autorregresivo simple.

A continuación se presenta una descripción de cada uno de los métodos propuestos para la predicción mensual, su aplicación y algunas consideraciones de interés.

4. MODELOS DE REGRESION LINEAL UTILIZANDO VARIABLES MACROCLIMATICAS

El modelo de relación más simple entre variables es el lineal. Los procesos de selección de las variables independientes a incluir en el modelo de regresión y la estimación de los parámetros del modelo se hace con procedimientos estadísticos estándares y ampliamente conocidos.

Si bien es claro que los caudales que se presentan en una cuenca no son una función lineal de las variables macroclimáticas, ellas contienen información de la evolución del sistema físico global que puede ser parcialmente capturada por un modelo lineal. Se ha comprobado que las ecuaciones lineales que incluyen variables macroclimáticas son en muchos casos mejores para la predicción de caudales en algunas cuencas del país que los modelos autorregresivos simples (EPM-UNAL, 1994).

Debido a la regulación que realiza la cuenca, normalmente los caudales mensuales exhiben coeficientes de autocorrelación importantes, por ello se incluyó en la ecuación de regresión el caudal ocurrido en la cuenca en el mes anterior al pronóstico. De esta mane-

ra se conforma un modelo lineal con una componente autorregresiva, y otras componentes que dependen de variables macroclimáticas. No se consideró necesario incluir en las ecuaciones la precipitación promedio mensual en la cuenca ya que, debido al tamaño de las mismas, la precipitación se refleja rápidamente en los valores de caudales.

En la actualidad y gracias al creciente interés de la comunidad científica internacional en el comportamiento climático global y en particular de su evolución en el Océano Pacífico, se cuenta con información de distintas variables climáticas, entre ellas: temperatura superficial del mar, humedad específica, velocidad del viento, presión a nivel del mar, etc. Algunas de ellas son monitoreadas en todo el océano y promediadas en distintas zonas, algunas de las cuales se pueden observar en la figura 1.

Para la calibración del modelo, en primer lugar se determinó la correlación lineal que presentan los caudales de los sitios de interés (Chivor, Salvajina y Jaguas) con un grupo de variables macroclimáticas. Estas variables fueron seleccionadas entre todas las contenidas en la base de datos COADS de acuerdo a los resultados obtenidos en el estudio EPM-U. NAL, 1994 en el cual se hizo un análisis detallado de todas las series y se seleccionaron las más representativas en la evolución del clima en el país. Se remite al lector interesado a dicho estudio. El listado de las variables utilizadas se presenta en la Tabla 3.

Tabla 3.
Variables Macroclimáticas Utilizadas en el estudio

SOI:	Indice de Oscilación del Sur
NI4-SST:	Temperatura Superficial del Mar en la zona niño 4.
ATN-SST:	Temperatura Superficial del Mar en el Atlántico Norte.
CAB-SST:	Temperatura Superficial del Mar en el Caribe.
NI1-SHQ:	Humedad Específica en la zona niño 1.
NI7-SLP:	Presión a nivel del Mar en la zona niño 7.
ATN-SLP:	Presión a nivel del Mar en el Atlántico Norte.
ATS-SLP:	Presión a nivel del Mar en el Atlántico Sur.
NI4-WEU:	Velocidad del Viento Superficial en la zona niño 4.
NI8-WEU:	Velocidad del Viento Superficial en la zona niño 8.
NI7-WNV:	Componente Norte de la Velocidad del Viento Superficial en la zona niño 7.
NI8-WNV:	Componente Norte de la Velocidad del Viento Superficial en la zona niño 8.

Para la determinación de las correlaciones se usaron variables estandarizadas. Esta estandarización se hace con el fin de eliminar el ciclo anual que presentan los datos y de esta manera facilitar su análisis y tratamiento. El modelo de regresión se define y estima para la variable estandarizada. El modelo generará valores de caudal estandarizado.

De los análisis de correlación realizados es interesante resaltar que tanto para Jaguas como para Salvajina la variable más importante es la NI4-SST. Además de acuerdo a los rezagos, se observa que el calentamiento en el Pacífico tiene un efecto más rápido en la zona sur (indicado por el rezago 0) que en la zona de Jaguas (tres meses después). Se observa también que las correlaciones de las variables macroclimáticas con los caudales en Chivor son significativamente más bajas que para los otros dos embalses y su máximo valor es tan sólo de -0.235

El grupo de variables independientes se conformó con todas las variables macroclimáticas rezagadas según la mayor correlación y con el caudal del último mes. Las variables

climáticas presentadas en la Tabla 3 se encuentran disponibles para los distintos períodos de análisis. Sin embargo, sólo algunos de ellos se consiguen con cierta rapidez ya que son publicadas periódicamente (mensualmente) y las otras son publicadas algún tiempo después (hasta un año). Debido a esta situación y con el fin de facilitar la aplicación de estos modelos para la predicción se realizaron nuevamente los ajustes de las regresiones limitando las variables climáticas a aquellas cuya información se puede obtener con mayor rapidez (por medio de los boletines mensuales que publica la NOAA: National Oceanic and Atmospheric Administration de E.U.A.). Es decir con las variables climáticas: Temperatura Superficial del Mar (SST), Índice de Oscilación del Sur (IOS) y Presión a Nivel del Mar (SLP). Estas ecuaciones fueron posteriormente redefinidas utilizando la actualización de los índices de Temperatura Superficial del Mar realizados a finales de 1994 y que se basa en el método de interpolación óptima (Optimum Interpolation). Las ecuaciones finales se presentan a continuación.

Ecuaciones de regresión de Jaguas:

Predicción un mes hacia adelante:

$$Q_{t+1} = -0.002 + 0.706 Q_t - 0.1197 \text{ NI4-SST}_{t-1}, R = 0.761 \quad (1)$$

Predicción dos meses hacia adelante:

$$Q_{t+2} = -0.005 + 0.572 Q_t - 0.172 \text{ NI4-SST}_t, R = 0.658 \quad (2)$$

Predicción tres meses hacia adelante:

$$Q_{t+3} = -0.0045 + 0.4495 Q_t - 0.20 \text{ NI4-SST}_t, R = 0.56 \quad (3)$$

Ecuaciones de regresión de Salvajina:

Predicción un mes hacia adelante:

$$Q_{t+1} = 0.4498 Q_t - 0.2775 \text{ NI4-SST}_t, R = 0.633 \quad (4)$$

Predicción dos meses hacia adelante:

$$Q_{t+2} = -0.003 + 0.238 Q_t - 0.382 \text{ NI4-SST}_t, R = 0.539 \quad (5)$$

Predicción tres meses hacia adelante:

$$Q_{t+3} = -0.004 + 0.179 Q_t - 0.387 \text{ NI4-SST}_t, R = 0.497 \quad (6)$$

No se presentan ecuaciones para Chivor ya que no presento ninguna correlación significativa con las variables macroclimáticas. Es notorio que todo el grupo final de ecuaciones utiliza el valor estandarizado de NI4-SST. La correlación de esta variable (y de las demás) con los caudales en Chivor es pequeña y muestra que este embalse se localiza en una zona con régimen hidrológico posiblemente influenciada por factores climáticos distintos al de las otras dos cuencas; por ejemplo, por las condiciones en la cuenca amazónica.

Los valores del coeficiente de correlación de las regresiones de Jaguas y Salvajina son relativamente altos y señalan que las variables climáticas aportan información para la predicción. Esto se hace notorio si se observa que los valores del coeficiente de correlación decrecen lentamente de una ecuación a otra. Este decrecimiento es más lento que

el que produciría un modelo autorregresivo simple. Esto significa que las ecuaciones presentadas muestran más claramente su ventaja, comparadas con un autorregresivo simple, mientras mayor sea el rezago. Se observa que, aun siendo este modelo de regresión lineal bastante sencillo, se obtienen ecuaciones de predicción que superan claramente las que se pueden obtener con un modelo autorregresivo simple. Este hecho es notorio en los embalses de Jaguas y Salvajina.

5. ANALISIS ESPECTRAL SINGULAR (AES) (Vautard, Yiou y Ghil, 1992)

El AES ha sido usado en recientes estudios climáticos para referirse a la aplicación uni-

variada del Análisis de Componentes Principales (ACP) en el tiempo y es equivalente a aplicar las Funciones Ortogonales Empíricas (FOE), a series de tiempo univariadas. Se desarrolló e implementó un modelo basado en análisis espectral singular para el caso de los tres embalses. Los fundamentos teóricos de este modelo pueden ser revisados en publicaciones previas (Mesa et al, 1994).

Una observación importante con la aplicación de este tipo de modelos en predicción tiene que ver con el mes anterior al inicio de la predicción, la cual no arranca en el valor histórico sino desplazada debido al filtrado de la señal. Lo anterior debido a que la metodología solo trabaja con la varianza significativa despreciando el resto. Esta diferencia podría ser corregida usando algún procedimiento preestablecido pues de otra manera la predicción no tendría valor como tal.

6. MODELO DE RETROPROPAGACION (Eberhart y Dobbins, 1990)

Las redes neuronales asemejan su funcionamiento al cerebro humano, una red neuronal está conformada por neuronas y enlaces entre éstas, cada neurona de entrada recibe una señal que es la información, transforma la señal en la neurona y transmite la señal de respuesta a otra neurona, mediante un enlace ya predeterminado y ponderado por un peso dentro del proceso de transformación de la señal, hasta que se obtiene una respuesta de una neurona de salida.

Un modelo de redes neuronales se desarrolló e implementó para los tres embalses. La teoría básica de redes neuronales puede ser

revisada en publicaciones anteriores (Mesa et al., 1994). Los resultados del entrenamiento del modelo de redes neuronales para los ríos de San Lorenzo, Garagoa más Somondoco y Salvajina, son los siguientes:

a. Río Salvajina

Estructura de la red: 3 nodos de entrada, 2 nodos capa escondida, 3 nodos capa de salida.

Información: caudales mensuales, índice de oscilación del sur (IOS) y la temperatura superficial del mar en la región Niño 4. La Información fue normalizada por su media y desviación estándar mensual y transformada por la función sigmoidea. La información va de enero de 1946 a diciembre de 1991.

Los resultados del entrenamiento fueron los siguientes:

error de ajuste: 0.027581

semilla: 3000

α : 0.015

η : 0.003

b. Río San Lorenzo

Estructura de la red: 3 nodos de entrada, 4 nodos capa escondida, 3 nodos capa de salida.

Información: caudales mensuales, índice de oscilación del sur (IOS) y la temperatura superficial del mar en la región Niño 4. La Información fue normalizada por su media y desviación estándar mensual y transformada por la función sigmoidea. La información va de enero de 1964 a diciembre de 1976.

Los resultados del entrenamiento fueron los siguientes:

error de ajuste: 0.0266
 semilla: 10000
 α : 0.006
 η : 0.002

c. Río Garagoa más Somondoco

Estructura de la red: 2 nodos de entrada, 4 nodos capa escondida, 3 nodos capa de salida.

Información: caudales mensuales y la temperatura superficial del mar en la región Niño 4. La Información fue normalizada por su media y desviación estándar mensual y transformada por la función sigmoidea. La información va de enero de 1973 a diciembre de 1990.

Los resultados del entrenamiento fueron los siguientes:

error de ajuste: 0.038163

semilla: 3500
 α : 0.13
 η : 0.007

7. PROCESOS AUTORREGRESIVOS DEPENDIENTES DEL REGIMEN

Los modelos autorregresivos dependientes del régimen (RAR) se pueden considerar extensiones de los modelos AR. Son una clase de modelos no lineales. Para un desarrollo detallado de la teoría de estos modelos el lector interesado puede revisar el Informe Final presentado a ISA (UN-ISA, 1995).

La idea básica es la siguiente. Supóngase que se tiene un modelo lineal para 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:

$$\begin{aligned} X_t &= a^{(1)} + a_1^{(1)} X_{t-1} + e_t^{(1)} \text{ si } X_{t-1} < r \\ X_t &= a^{(2)} + a_1^{(2)} X_{t-1} + e_t^{(2)} \text{ si } X_{t-1} \geq r \end{aligned} \quad (7)$$

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 RAR(1,1) el cual se define de la siguiente manera:

$$\begin{aligned} X_t &= a^{(i)} + a_1^{(i)} X_{t-1} + e_t^{(i)} \\ \text{si } X_{t-1} \in R^{(i)}, \quad i=1, \dots, l \end{aligned} \quad (8)$$

Donde $R^{(1)}, \dots, R^{(l)}$ son subconjuntos de la línea real R^1 los cuales definen una partición de R^1 en intervalos disyuntos $(-\infty, r_1]$, $(r_1, r_2], \dots, (r_{l-1}, \infty)$, en la cual $R^{(1)}$ denota el intervalo $(-\infty, r_1]$ y $R^{(l)}$ el intervalo (r_{l-1}, ∞) .

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

$$X_t = \lambda(X_{t-1}) + e_t \quad (9)$$

Donde λ es alguna función no lineal.

De manera similar se pueden construir modelos de mayor orden, por ejemplo un modelo RAR(1,k) está dado por

$$X_t = a_0^{(i)} + a_1^{(i)}X_{t-1}, \dots, a_k^{(i)}X_{t-k} + e_t^{(i)} \quad (10)$$

si $(X_{t-1}, \dots, X_{t-k}) \in R^{(i)}$ donde $R^{(i)}$ es una región dada del espacio Euclidiano k-dimensional $R^{(k)}$. Este modelo puede considerarse como una aproximación lineal al modelo general no lineal AR de orden k,

$$X_t = f(X_{t-1}, X_{t-2}, \dots, X_{t-k}) + e_t \quad (11)$$

En la práctica es difícil ajustar un modelo de la forma (11) a los datos cuando k es grande, ya que la determinación de las regiones límites requeriría una búsqueda sobre el espacio k-dimensional. Entonces, se ha restringido la atención a los casos donde los distintos grupos de valores de los parámetros se determinan de acuerdo a un único valor pasado, X_{t-r} .

Entonces, el modelo RAR(1,k) toma la forma,

$$X_t = a_0^{(j)} + \sum_{i=1}^k a_i^{(j)} X_{t-i} + e_t^{(j)} \quad \text{si } X_{t-d} \in R^{(j)}, \quad j = 1, 2, \dots, l \quad (12)$$

Donde $R^{(j)}$, es un subconjunto de la línea real R^1 y d es el rezago con relación al valor indicador de régimen.

En la ecuación anterior se ha asumido un orden común, k, para cada una de las autorregresiones. En la práctica, puede ser adecuado utilizar modelos de diferente orden según el régimen.

Con esta consideración la ecuación (12) toma la siguiente forma

$$X_t = a_0^{(j)} + \sum_{i=1}^{k_j} a_i^{(j)} X_{t-i} + e_t^{(j)}$$

$$\text{si } X_{t-d} \in R^{(j)}, \quad j = 1, 2, \dots, l \quad (13)$$

donde k_1, k_2, \dots, k_l denota el orden de las autorregresiones en los diferentes regímenes.

Una de las características interesantes de los modelos RAR, es que pueden mostrar un comportamiento cíclico, condición similar a la característica encontrada en las ecuaciones diferenciales no lineales.

Los modelos dependientes del régimen en donde los cambios entre los grupos de parámetros está determinado por un valor pasado del proceso asociado (variables macroclimáticas en este caso) reciben el nombre de modelos RAR instrumental. La forma es completamente similar a la anterior pero en este caso el cambio de régimen está dado por la variable asociada.

En este caso el proceso de estimación de parámetros incluye la definición del cambio de régimen, de la selección del orden del modelo autoregresivo en cada régimen y de los respectivos parámetros autoregresivos de cada modelo. En este caso se pueden adaptar procedimientos tradicionalmente utilizados para esos propósitos como son los criterios de parsimonia de Akaike y procedimientos de estimación de parámetros. El lector interesado puede referenciarse a trabajos anteriores (Mesa et al., 1994).

Se ajustó el modelo RAR instrumental (2,1) a los caudales mensuales de las cuencas de los embalses de Salvajina y Jaguas. Debido a la baja correlación lineal encontrada entre las variables climáticas y los caudales en el embalse de Chivor, no se consideró apropiado para esta cuenca el uso del modelo RAR instrumental basado en las variables climáticas analizadas y por ello, no se hizo el ajuste.

a. Embalse de Salvajina

Basándose en la información de correlaciones lineales se escogió como serie indicadora del régimen, a la variable climática Temperatura del Mar en la Zona Niño 4 por poseer los mayores valores de correlación

con los caudales en el embalse Salvajina. Aplicando la metodología señalada, es decir variando los rezagos de la variable indicadora y variando también el valor del criterio de separación entre regímenes se obtuvieron los parámetros del modelo RAR Instrumental (2,1).

REZAGO 1 mes

CRITERIO = -1.00

IND	No. DAT	MED	MED t-1	VAR.	VAR t-1	CORR t	VAR ERR
>= -1.00	421	-.223	-.223	.754	.739	.511	.558
< -1.00	94	1.001	1.002	.773	.840	.255	.722

AIC(1)= -235.966 AIC(2)= -20.561 TAIC= -256.527 TAICN= -.498

en donde IND es el límite del régimen de acuerdo a la variable índice, No. DAT representa el número de datos correspondiente a cada régimen, MED es la media en el tiempo t de la variable en el régimen, MED t-1 es la media en el tiempo t-1 de la variable en el régimen, VAR es la varianza de la variable en el régimen, CORR es el coeficiente de autocorrelación de rezago uno de la variable en el régimen, VAR ERROR es la varianza del error del ajuste del modelo, AIC() es el criterio de Akaike para el régimen, TAIC es el valor total del criterio de Akaike y TAICN es el valor total normalizado del criterio de Akaike.

REZAGO 2 meses

CRITERIO = .60

IND	No. DAT	MED	MED t-1	VAR.	VAR t-1	CORR t	VAR ERR
>= .60	156	-.567	-.481	.495	.561	.522	.360
< .60	358	.246	.211	.995	1.022	.529	.716

AIC(1)= -149.261 AIC(2)= -109.450 TAIC= -258.711 TAICN= -.503.

REZAGO 3 meses

CRITERIO = .80

IND	No. DAT	MED	MED t-1	VAR.	VAR t-1	CORR t	VAR ERR
>= .80	123	-.662	-.607	.450	.494	.505	.335
< .80	390	.209	.192	.972	.987	.526	.703

AIC(1)= -124.485 AIC(2)= -127.515 TAIC= -252.000 TAICN= -.491

b. Embalse de Jaguas

Se eligió como serie indicadora del régimen a la Temperatura del Mar en la Zona

Niño 4, luego del proceso de determinación de los parámetros se encontraron los siguientes resultados

REZAGO 1 mes

CRITERIO = .90

IND	No. DAT	MED	MED t-1	VAR.	VAR t-1	CORR t	VAR ERR
> = .90	73	-.735	-.707	.519	.575	.733	.240
< .90	262	.205	.201	.905	.898	.703	.458

AIC(1)= -94.056 AIC(2)= -194.417 TAIC= -288.473 TAICN= -.861.

REZAGO 2 meses

CRITERIO = .80

IND	No. DAT	MED	MED t-1	VAR.	VAR t-1	CORR t	VAR ERR
> = .80	84	-.708	-.653	.608	.573	.745	.270
< .80	250	.239	.224	.872	.911	.694	.452

AIC(1)= -99.837 AIC(2)= -188.623 TAIC= -288.460 TAICN= -.864

REZAGO 3 meses

CRITERIO = .90

IND	No. DAT	MED	MED t-1	VAR.	VAR t-1	CORR t	VAR ERR
> = .90	73	-.823	-.773	.515	.539	.614	.321
< .90	260	.235	.222	.860	.878	.708	.428

AIC(1)= -73.037 AIC(2)= -210.455 TAIC= -283.492 TAICN= -.851

El modelo RAR Instrumental se puede utilizar para predicción. De acuerdo al valor que presenta la serie índice, se cuenta con la ecuación de autorregresión condicionada a esta situación.

Para la predicción se puede hacer uso de varias ecuaciones de acuerdo al rezago. Por ejemplo utilizar las tres ecuaciones presentadas para cada cuenca. Otra forma es utili-

zar reiteradamente la ecuación con la cual se obtuvo el menor valor del TAICN. (por ejemplo la ecuación con los parámetros del rezago 2 en la cuenca de Jaguas). Nótese que en este último caso, la variable, por su carácter autorregresivo dentro del régimen, tenderá a la media condicionada.

En el proceso de determinación de los parámetros del modelo, se selecciona el modelo

con el menor valor del criterio de Akaike normalizado (TAICN). Este valor no siempre es monótonamente decreciente y por ello es posible que se encuentren mínimos locales. Para encontrar el mínimo global es necesario calcular el TAICN en todo el intervalo de variación del valor del criterio límite del régimen. Es de esta manera como se trabajó en este estudio para definir los diferentes modelos RAR.

8. DESAGREGACION MENSUAL A SEMANAL USANDO INTERPOLACION FRACTAL

Este modelo se utilizará para la desagregación de pronósticos de caudal mensual a caudales semanales. En primer lugar se hará una corta presentación sobre el desarrollo de esta teoría y su aplicación a las series hidrológicas, posteriormente se presenta la forma en que se utiliza el modelo.

En 1982 Mandelbrot introdujo el término 'Fractal' refiriéndose a ciertas estructuras geométricas que exhiben la propiedad de la invarianza en escala. Esto significa que en un amplio rango de escalas el objeto presenta básicamente las mismas características geométricas.

El interés de utilizar este tipo de conceptualización a los fenómenos de lluvia y caudal se inicia con los trabajos de Lorentz (1963) acerca de la circulación del flujo convectivo en la atmósfera. El más importante de sus resultados fue mostrar la imposibilidad de la predicción atmosférica exacta, ya que las soluciones del modelo de circulación más simplificado muestran la presencia de un atractor extraño. Lovejoy y Mandelbrot (1985) establecieron que las nubes y la lluvia presentan no sólo invarianza de escala en el dominio espacial sino también tempo-

ral. En 1989 Rodríguez-Iturbe et al. analizando datos de tormentas encontraron evidencia de estructura fractal. Angel y Puente (1989) utilizaron la interpolación fractal para la generación de caudales diarios. Compararon este modelo con el modelo Gaussiano y encontraron un mejor comportamiento en la reproducción de los estadísticos de extremos. La aplicación que se hizo en este trabajo se basa en esta última referencia.

Con el fin de estimar la dimensión fractal, se construye para cada mes la serie conformada por los caudales diarios históricos. De esta manera se forman 12 nuevas series de datos. El cálculo de la dimensión fractal de cada una de las series se lleva a cabo por medio del algoritmo de Higuchi (1988). Este método es más estable y preciso que otros métodos existentes. Consiste en lo siguiente:

De la serie $X_{(1)}, X_{(2)}, \dots, X_{(N)}$ se construye una nueva serie seleccionando ciertos elementos:

$$X_k^M = X_{(m)}, X_{(m+k)}, X_{(m+2k)}, \dots, X_{(m+(N-m)/k)} \quad (14)$$

donde $[]$ denota la parte entera de la expresión encerrada. m es el tiempo inicial (entero) y k es el intervalo de tiempo.

La longitud de la curva está definida así:

$$L_m(k) = \frac{1}{k} \left(\sum_{i=1}^{\pi} |X_{(m+ik)} - X_{(m+(i-1)k)}| \right) \cdot (N-1) / (\pi K) \quad (15)$$

Donde $\pi = [(N-1)/(K)]$, el término $(N-1) / (\pi K)$ es un factor para la normalización de la longitud del subgrupo de la serie de tiempo. La longitud de la curva $L(k)$ para el intervalo de tiempo k es el promedio sobre los k grupos de $L_m(k)$. Si estos promedios

siguen la ley $L(k) \propto k^{-D}$, la curva es fractal con dimensión D.

Luego de que se ha medido la longitud de la curva para diferentes intervalos de tiempo, se calcula mediante mínimos cuadrados la pendiente del mejor ajuste entre $\log L(k)$ y $\log k$. La dimensión fractal es el negativo de esa pendiente.

Se tomaron valores de k hasta 5 ya que de esta forma se está teniendo en cuenta la variación en la escala entre 2 y 2^5 días, aproximadamente la longitud de un mes.

En la presente aplicación, el modelo se utilizará para realizar la desagregación semanal de los caudales basándose en caudales mensuales dados. Para ello inicialmente se realiza la desagregación a caudales diarios que luego se agregan para conseguir los caudales semanales.

Utilizando el programa DIMENS.FOR (Angel y Puente, 1989), que entrega los valores de $L(k)$ para distintos k, y haciendo la regresión lineal de los logaritmos se obtuvieron los valores de la dimensión fractal para las series de caudales diarios de Jaguas, Salvajina y Chivor los cuales se muestran en las tabla 4.

Tabla 4.
Dimensión fractal de los caudales.

mes	Jaguas	Salvajina	Chivor
enero	1.863	1.658	1.606
febrero	1.803	1.606	1.715
marzo	1.854	1.630	1.742
abril	1.957	1.693	1.819
mayo	1.982	1.716	1.842
junio	1.910	1.773	1.895
julio	1.837	1.850	1.880
agosto	1.909	1.863	1.908
septiembre	1.904	1.735	1.865
octubre	1.921	1.742	1.877
noviembre	1.886	1.729	1.810
diciembre	1.830	1.668	1.714

De acuerdo con lo anterior, los datos requeridos para el uso del modelo son:

- Caudal inicial: este valor corresponde al caudal inicial y desde el comenzará la interpolación de la serie de datos desagregados.
- Caudal final: valor de caudal en el que termina la traza de la serie desagregada. Este valor puede corresponder a la predicción del caudal medio del mes siguiente.
- Dimensión fractal del mes que se desagrega.
- Caudal promedio del mes que se desagrega. Este valor corresponderá a un resultado obtenido previamente de un modelo de predicción mensual de caudales.

Con estos datos el modelo realiza una interpolación entre los valores inicial y final preservando la media del mes que se desagrega y la dimensión fractal. La desagregación se lleva a cabo usando el modelo de interpolación FRAME (Fractal Mean) desarrollado por E. Angel y C. Puente (1989). Este programa comienza con un punto en el atractor, en este caso el valor del caudal inicial, escoge al azar una de las funciones e itera el punto inicial, se escoge de nuevo una función usando como punto inicial el obtenido en la anterior iteración. Se repite el proceso al infinito (en la práctica, un número grande de veces) y la función continua resultante es el atractor. (Para más detalles véase Angel y Puente 1989). Mediante esta desagregación se obtienen los datos diarios de caudal que son luego agregados para obtener los caudales semanales.

6. CONCLUSIONES Y RECOMENDACIONES

Se presentan ahora algunas conclusiones y recomendaciones sobre los modelos de predicción implementados.

- Las correlaciones obtenidas con los modelos de regresión lineal de caudales con variables macroclimáticas a nivel mensual no son muy significativas, llegando a ser del orden de 0,7 a 0,6 para Jaguas, de 0,45 para Salvajina y de 0,4 a 0,25 para Chivor. Sin embargo, ésta es una situación esperada pues tal como se discutió en la introducción las relaciones entre caudales y el clima no son simples ni lineales. A pesar de estos niveles de correlación, estos modelos muestran dependencias mucho más claras que modelos autorregresivos simples. Los modelos de regresión propuestos usan las mayores correlaciones entre las variables macroclimáticas y los caudales que en general tienen un rezago mayor a un mes. En el caso de los modelos autorregresivos simples, mientras mayor sea el rezago la autocorrelación cae muy rápidamente.
- En la definición de los modelos de regresión lineal a nivel mensual se usaron un grupo de variables macroclimáticas disponibles como variables independientes y se procedió a escoger las variables a ser incluidas en el modelo de regresión. Sin embargo, algunas de estas variables no se consiguen en tiempo real y por lo tanto se procedió a redefinir los modelos de regresión usando sólo las variables disponibles en tiempo real. La diferencia en la bondad del modelamiento no fue significativa.
- Los resultados del ajuste del modelo de regresión lineal que relaciona caudales con variables macroclimáticas a nivel mensual para Chivor no son aceptables, indicando que el régimen de caudales de ese embalse no está influenciado en forma directa por las variables macroclimáticas relacionadas con el ENSO. Sin embargo, pudiera estar afectada por el efecto de otros factores climáticos no estudiados como las condiciones de la cuenca amazónica, cuyas condiciones climáticas también afectan el clima en Colombia.
- El método de Análisis Espectral Singular para predicciones a nivel mensual tiene el problema de que la predicción no está amarrada al valor real histórico del mes anterior, sino que se encuentra desplazada debido al filtrado de la serie. Esto se debe a que la metodología sólo trabaja con la varianza significativa despreciando el resto. Este problema podría ser solucionado si se hace una corrección del punto de partida de la predicción por algún mecanismo que asocie ese punto con el valor histórico del mes anterior.
- Se presenta un procedimiento de desagregación de caudales mensuales a semanales basado en interpolación fractal. Este procedimiento permite predecir caudales a nivel semanal mediante la desagregación de la predicción a nivel mensual. El procedimiento realiza una interpolación entre los valores inicial y final del mes preservando la media del mes que se desagrega y la dimensión fractal.
- Se recomienda siempre en las predicciones a cualquier nivel de tiempo usar varios procedimientos. Las relaciones que se pretenden modelar para realizar las predicciones son tan complejas, que el uso de varios modelos permitiría tener un conocimiento de lo que pueda estar sucediendo. El involucrar las variables macroclimáticas en las predicciones de más largo plazo (3 meses) tiene mucho senti-

do en la representación de la dependencia del clima de nuestro país con los fenómenos tipo ENSO. El uso de modelos lluviaescorrentía en los procesos de predicción diaria reconoce el carácter cortoplacista de este tipo de predicciones, pero las amarra al comportamiento general del clima mediante una estructura de ponderación de las posibles realizaciones asociada con ese comportamiento.

7. REFERENCIAS

- Akaike, H. (1969). Fitting autoregressive models for prediction, *Ann. Inst. Stat. Math.*, 21, 243-247.
- Angel, E y Puente, C. Fractal Interpolators in Hydrology. Master's Thesis, Graduate Division, University of California at Davis.
- Angel, Enrique. Interpolación Fractal Aplicada a la Generación de Series de Caudales Diarios. X Seminario Nacional de Hidráulica e Hidrología, Manizales, 1990, 87-98.
- Bjerknes, J. (1969). Atmospheric teleconnections from the tropical Pacific. *Monthly Weather Review*, 97, 103-72.
- Cane, M. A., Zebiak, S. E. and Dolan, S. C., (1986). Experimental Forecasts of El Niño. *Nature*, 321, 827-832.
- Eberhart, Russell C y Dobbins, Roy W (1990). *Neural Network PC Tools, A Practical Guide*. Academic Press, San Diego, California, 35-49.
- EPM-U. Nal (1994). Metodología de Predicción Hidrológica. Estudio para Empresas Públicas de Medellín. Posgrado en Aprovechamiento de los Recursos Hidráulicos, Facultad de Minas.
- Higuchi, T. (1988). Approach to an irregular time series on the basis of the fractal theory. *Physica D*, 31: 277-283.
- Holton, J. R., 1992. *An Introduction to Dynamics Meteorology*. 3 ed. Academic Press. San Diego.
- Klemes, V. (1974). The Hurst phenomenon - a puzzle?, *Water Resour. Res.*, 10(4), 675-688.
- Lattif, M. et al (1993). Structure and Predictability of The El Niño/Southern Oscillation Phenomenon in a coupled Ocean Atmosphere General Circulation Model. *J. Climate* (Imprint)
- Lattif, M. et al (1994). A Review of ENSO prediction studies. *Climate Dynamics*, 9: 167-179.
- Lorenz, E. N. (1969). Atmospheric predictability as revealed by naturally occurring analogues. *J. Atmos. Sci.* 636-646.
- Lorenz, E. N. (1963). Deterministic non-periodic flow. *Journal of Atmospheric Sciences*, 20: 130-141.
- Lovejoy, S. and Mandelbrot, B. B. (1985). Fractal properties of rain and a fractal model. *Tellus*.
- Mesa, O. y Poveda, G. (1993). Metodologías de Predicción de la Hidrología Colombiana considerando el Fenómeno del Niño/Oscilación del Sur (ENOS). Tercer Congreso Colombiano de Meteorología.

- Mesa O. y Carvajal L.F. (1994). Análisis Espectral Singular y Redes Neuronales Aplicadas en Hidrología. Revista Avances en Recursos Hidráulicos No. 2, Facultad de Minas, Universidad Nacional de Colombia, Medellín.
- Salazar J.E. y Mesa O., (1994). Aplicación de dos Modelos no Lineales a la Simulación de Series Hidrológicas. Revista Avances en Recursos Hidráulicos No. 2, Facultad de Minas, Universidad Nacional de Colombia, Medellín.
- Millán, Jaime. Drought Impact on Regional Economic (1972), Hydrology papers No. 55, Colorado State University. Fort Collins, Colorado.
- Simpson, J. et al., (1988). A Proposed Tropical Rainfall Measuring Mission (TRMM) Satellite. Bull. Am. Met. Soc. 69(3), 278-295.
- Pereira, M. V. F., G. C. Oliveira, C. G. Costa, and J. Kelman, (1984). Stochastic streamflow modeling for hydroelectric systems, Water Resour. Res., 20(3), 379-390.
- Vautard, R., Yiou, P. y Ghil, M. (1992). Singular-spectrum analysis: A toolkit for short, noisy chaotic signals. Physica D 58, 95-126.