

# **ANALISIS ESPECTRAL SINGULAR Y REDES \* NEURONALES APLICADAS EN HIDROLOGIA**

**Oscar Mesa S., Luis F. Carvajal S.**  
**Posgrado en Aprovechamiento de Recursos Hidráulicos,**  
**Facultad de Minas, Universidad Nacional de Colombia,**  
**Sede Medellín**

## **RESUMEN**

El presente trabajo considera dos metodologías que utilizan la dinámica no lineal (caos) y su aplicación en la predicción de lluvias y caudales. La motivación de estos trabajos está en la oscilación internaual del clima tropical conocida como El Niño-Oscilación del Sur (ENOS) es un fenómeno no lineal de clara influencia en la hidrología de Colombia.

La primera es el *Análisis Espectral Singular* (AES), metodología útil en el filtrado de señales, con una amplia aplicación en el campo de la electrónica y recientemente en la hidrometeorología. Se busca capturar la componente no lineal presente en la dinámica de las series temporales naturales mediante el estudio del atractor en el espacio fásico y la utilización de métodos tradicionales de predicción en modelos autorregresivos AR(P) para extrapolar el comportamiento dinámico futuro. Este procedimiento fue aplicado a la modelación de lluvias y caudales en diversas cuencas del departamento de Antioquia, así como caudales en el territorio nacional.

La otra metodología es la de *las redes neuronales* (modelo de retropropagación), que han tomado gran auge en diferentes campos de la actividad investigativa, presentando nuevos caminos para el estudio de fenómenos no lineales y que es aplicada a la predicción de lluvias y caudales, considerando información de precipitación, caudales, manchas solares, temperatura superficial del mar (TSM), el Índice de Oscilación del Sur (IOS) y datos de la oscilación quasi-bienal del viento (OQB). Los resultados plantean la posibilidad de considerar las redes neuronales como métodos de predicción con un cierto éxito.

---

\* Trabajo realizado con el apoyo de Empresas Públicas de Medellín.

## Abstract

The El Niño-Southern Oscillation (ENOS) influence on colombian hydrology motivated the study of two non-linear methodologies for rainfall and streamflow prediction.

The first methodology is *Singular spectral analysis* which is a very powerful method in nonlinear (chaos) analysis of time series. The main idea is to filter the deterministic nonlinear component of a time series and to model it by means of traditional time series methods (autoregressive methods). The original series is embedded in a higher dimensional space for the reconstruction of the attractor of the dynamic system producing the series. The method is equivalent to the expansion in Fourier series but instead of using prescribed sine and cosine functions empirical eigen functions are used. The eigen functions come from the eigen vectors of the covariance matrix of the series in the phase space. Predictions are made extrapolating tendencies in the attractor. The method is applied to rainfall data in Colombia.

The second method is *Neural Network Analysis*, currently a very popular tool for modeling a great variety of nonlinear phenomena. This work presents the application of neural networks to the prediction of streamflows and rainfall in Colombia, using the southern oscillation index, QBO index (Quasibiennial stratospheric wind tropical velocity) and Pacific sea surface temperatures as predictors.

## 1. Introducción

La crisis del sector eléctrico en el año de 1992 y principios de 1993, sirvió de trampolín para cuestionar las metodologías de predicción utilizadas en el sector eléctrico, donde una de las causas fue atribuida a la presencia de una onda cálida en el Océano Pacífico (El Niño), que altera el ciclo interanual del frente intertropical de convergencia (FIC), produciendo una disminución de la precipitación y de los caudales, especialmente en la zona andina y la vertiente del Pacífico, zona donde se encuentran los principales desarrollos hidroeléctricos que conforman el sistema eléctrico interconectado colombiano.

El AES es usado en recientes estudios climáticos para referirse a la aplicación univariada del Análisis de Componentes Principales (ACP) en el tiempo y es equivalente a aplicar las Funciones Ortogonales Empíricas (FOE), a series de tiempo univariadas. Este ha sido un método usado por años en procesamiento de señales digitales. Fue introducido en oceanografía por Colebrook (1978) y en dinámica no lineal por Broomhead y King (1986) y por Fraedrich (1986). Algunos de los más recientes trabajos en AES son referenciados a continuación.

Rasmusson, Wang y Ropelewski (1990), mostraron que el fenómeno El Niño-Oscilación del Sur (ENOS) presenta una señal

quasi-bianual bastante regular modulada por una frecuencia baja, y una oscilación menos regular de 4 a 5 años.

Ghil, Yiou y Vautard (1992) aplicaron el AES a series de temperatura superficial de la tierra de 135 años de longitud, encontrando evidencias de oscilaciones en años y decenios, confirmado por Allen et al (1992).

Ghil y Mo (1991) aplicaron el AES de múltiples canales a datos de altura geopotencial en el Hemisferio Norte extratropical, el cual reveló ciclos con períodos de 40 a 50 días, 20-25 días y 70 días.

Penland, Ghil y Weickman (1991), mostraron que el prefiltrado por AES, permite usar el Método de Máxima Entropía (MME) con modelos autorregresivos (AR) de bajo orden en estimación espectral.

Basado en esta combinación Keppenne y Ghil (1992) hicieron predicciones del Índice de Oscilación del Sur (IOS) para el Niño-Oscilación del Sur (ENOS) con cierta habilidad en períodos de 30 meses.

En cuanto a las redes neuronales, asemejan su funcionamiento al cerebro humano; una red neuronal está conformada por neuronas y enlaces entre estas, cada neurona de entrada recibe una señal que lleva 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. Los desarrollos más recientes de redes neuronales han provocado un auge importante en los últimos años en cuanto a su aplicación en diversos campos algo disímiles, economía, óptica, electrónica, el campo militar, ciencias de la tierra, etc. Desde la segunda mitad

del siglo XX diversos desarrollos han caracterizado el avance de las redes neuronales.

Rosenblatt en 1958 desarrolló el Perceptrón: un modelo con tres capas, la del medio conocida como la capa de asociación. Había entre las unidades de salida procesos competitivos que resultan de conexiones inhibitorias entre las unidades de salida que ayudaron a la clasificación de patrones. Este sistema de tres capas pudo aprender a conectar o asociar una entrada a una unidad de salida aleatoria. El Perceptrón fue computacionalmente preciso y fue una máquina de aprendizaje verdadera. El sistema por su comportamiento adaptivo complejo tiene limitaciones.

Luego apareció el ADALINE (Adaptive Linear Element, Widrow y Hoff, 1960) y el MADALINE (muchos ADALINES), los cuales son aparatos electrónicos análogos que emplean la regla Delta de aprendizaje, la cual minimiza la diferencia entre la salida observada y la deseada.

Posteriormente aparecieron una serie de investigadores que enriquecieron el campo de las redes neuronales; en 1972 Shun-Ichi Amari presentó la red auto-organizativa, como un modelo de memoria asociativa. James Anderson (1977) comenzó a desarrollar en los 70 el modelo llamado cerebro-estado-en-una-caja, de tipo memoria asociativa; en esta categoría Teuvo Kohonen, desarrolló en 1972 las redes Cuantización del vector adaptivo auto-organizado y la Cuantización del vector de aprendizaje.

Werbos en 1974 presentó la red de retropropagación, la cual es un Perceptrón con múltiples capas, una función umbral diferente y una regla de aprendizaje robusta y aplicada.

Gail Carpenter y Steve Grossberg (1988) desarrollaron redes de teoría de resonancia

adaptiva basados en modelos plausibles biológicoicamente.

En 1975 Kunihiko Fukushima presentó el modelo Cognitron, sistema neuronal multicapa que interpreta caracteres de escritura. En los años 80 se presentó el modelo mejorado Neocognitron.

En la década de los 80, las redes mas populares fueron: la red de retropropagación (Le-cun, 1986; Parker, 1985; Rumelhart y McClelland, 1986), la red Hopfield (Hop-field, 1982, 1986; Hopfield y Tank, 1986; Tank y Hopfield, 1987) dentro del tipo de redes autoasociativas es ampliamente conocida.

Bart Kosko (1986, 1987a, 1987b, 1987c), ha desarrollado redes que utilizan la memoria asociativa bidireccional, dentro de una dinámica llamada resonancia, en la que los patrones en la primera y segunda capa estimulan la una a la otra hasta que el patrón en cada capa presenta un estado estable.

## 2. Análisis Espectral Singular (AES) y Ajuste de Autorregresivos AR(P) por Máxima Entropía

El AES es la expansión de un campo discreto ( $X_i, 1 \leq i \leq N$ ), en sus componentes principales (CP) y Funciones Ortogonales Empíricas (FOE) considerando una longitud de ventana M:

$$X_{i,j} = \sum_{k=1}^{cp} a_i^k E_j^k \quad 1 \leq j \leq M \quad (2.1)$$

Los coeficientes de proyección  $a_i^k$  son las llamadas CP y los vectores  $E_j^k$ , las FOE.

Para un análisis espectral simple la expansión sería:

$$x_{i+j} = \sum_{k=1}^{cp} a_i^k E_j^k \quad 1 \leq i \leq N \\ 0 \leq j \leq M-1 \quad (2.2)$$

En la expresión 2.1 y 2.2, M es llamada la longitud de ventana, la cual es fijada por el analista.

Para un análisis espectral multivariado el cual considere L sitios de información  $X_{l,i}$ ,  $1 \leq l \leq L, 1 \leq i \leq N$ , la expansión en las CP y FOE es:

$$X_{l,i+j} = \sum_{k=1}^{L \times M} a_i^k E_{l,j}^k \quad 1 \leq l \leq L, \\ 1 \leq j \leq M \quad (2.3)$$

El problema tiene una longitud de ventana M, pero los autovectores tendrán una dimensión igual a  $L \times M$ .

Las FOE son los autovectores de la matriz Toeplitz,  $T_x$  que contiene los coeficientes de covarianza cruzada de los diferentes L vectores para rezagos de 0 a M-1. Las ecuaciones 2.1 a 2.3 son la aplicación de la expansión biortogonal de Karhunen-Loëve (1962), muy usada en procesamiento de señales digitales.

La ortogonalidad en tiempo (covarianza cruzada igual a cero para dos CP en el rezago cero) y el espacio (ortogonalidad de las FOE), implica que  $\lambda_k$  (autovalor k de la matriz Toeplitz) representa la varianza de la k-ésima CP. Si se trunca la ecuación 2.1 en la CP p, se reduce la información a las primeras p CP, esto se hace en una forma óptima ya que las primeras p componentes describen la mayor varianza de la varianza total que se puede obtener de una proyección en p vectores ortogonales.

Vautard y Ghil (1989) aplicaron el AES a datos paleoclimáticos y encontraron que cuando un par de autovalores están muy cerca se tiene actividad periódica. El AES

descompone la señal en la parte significativa y el ruido.

El AES se ha aplicado a una docena de datos geofísicos, en escalas de tiempo de días a milenios, de varias longitudes y con extensiones a procesos multivariados.

A continuación se presentan las ideas básicas del AES. Para estimar los elementos de la matriz de covarianza,  $T_x$ , se tiene la expresión:

$$c(j) = \frac{1}{N-j} \sum_{i=1}^{N-j} x_i x_{i+j} \quad (2.4)$$

Alternativamente, el algoritmo de Burg (1967, 1968) el cual estima los coeficientes autorregresivos asociados con el método de máxima entropía, también estima implícitamente los  $c(j)$ . La expresión usada por Box y Jenkins (1970) para calcular los  $c(j)$ , la cual divide la ecuación 2.4 por N es fuertemente sesgada cuando N es pequeño, mientras que la ecuación 2.4 tiene mayor varianza pero es menos sesgada, cuando N es pequeño. El método de Burg puede ser sesgado para los grandes rezagos, si hay picos en períodos más grandes que N (Vautard et al, 1992).

La matriz  $T_x$  es simétrica y no negativa con autovalores  $\lambda_k$  positivos (Vautard et al, 1992). Los autovalores son ordenados en orden decreciente y los correspondientes autovectores están normalizados.

Un problema importante en el AES es la escogencia de la ventana M, según el número de datos disponibles N, si se considera que la capacidad de resolución del AES es igual a la longitud de la ventana M y se desea reconstruir un atractor extraño, cuyo espectro incluye períodos de longitud arbitraria, el mayor M es el mejor, pero para que los errores estadísticos no afecten los últimos

valores de la función de autocovarianza, M no debe ser mayor que  $(1/3)N$ .

La k-ésima CP es el coeficiente de la proyección ortogonal de la serie original en el k-ésimo vector propio:

$$a_i^k = \sum_{j=1}^M x_{i+j} E_j^k, \quad 0 \leq i \leq N-M \quad (2.5)$$

Las CP son versiones filtradas de la serie original y se pueden considerar como un promedio móvil de la información en la longitud de ventana M considerada. En la ecuación 2.2 los términos individuales dependen del índice j (variando de 1 a M), por lo tanto hay M formas diferentes de reconstruir los componentes de la señal, que en general no dan los mismos resultados. Otro inconveniente es que las series resultantes de usar la ecuación 2.2 son de longitud  $N-M+1$  y no de longitud N. Las CR permiten extraer series de longitud N, correspondientes a un conjunto de autovalores ya seleccionados.

Se busca una serie y de longitud N,  $y=R_Ax$ , tal que la diferencia entre y y x al cuadrado sea mínima, siendo A un subconjunto de k autovalores sobre los cuales se va a realizar la reconstrucción.

La solución de  $y=R_Ax$ , en forma óptima de acuerdo al criterio de mínimos cuadrados es (Vautard, 1992, ecs 2.17a-2.17c):

$$(R_Ax)_i = \frac{1}{M} \sum_{j=1}^M \sum_{k \in A} a_{i+j}^k E_j^k, \quad M \leq i \leq N-M+1 \quad (2.6)$$

$$(R_Ax)_i = \frac{1}{i} \sum_{j=1}^i \sum_{k \in A} a_{i-j}^k E_j^k, \quad 1 \leq i \leq M-1 \quad (2.7)$$

$$(R_A x)_i = \frac{1}{N-i+1} \sum_{j=i-N+M}^M \sum_{k \in A} a_{i-j}^k E_j^k, \quad (2.8)$$

$N-M \leq i \leq N$

Las CR son las  $R_A x$  (denotadas por  $x^k$ ), y tienen propiedades aditivas, como:

$$R_A x = \sum_{k \in A} x^k \quad (2.9)$$

Adicionalmente las series  $x$  pueden ser expandidas como la suma de sus CR.

No obstante su aspecto lineal, la transformación de  $x$  en  $y$ , es no lineal, pues los autovectores  $E^k$  dependen no linealmente de  $x$ . Una desventaja de los CR es que son correlacionados aún en el rezago cero. La ventaja de los componentes reconstruidos es que si hay períodos cortos de oscilaciones en la señal, se pueden localizar en forma precisa. Una desventaja es que  $x$  y  $x^k$  no están en fase excepto al final de la serie.

Una vez se ha realizado el AES, el utilizar el MME es útil por su alta resolución espectral con la desventaja de que pueden aparecer picos falsos según la resolución, incrementando el orden del modelo. Las series espaciadas un intervalo de tiempo  $t$ , se les puede ajustar un modelo autorregresivo de orden  $P$ .

$$x(t) = \sum_{k=1}^P b_k x[t - k] + \epsilon \quad (2.10)$$

donde  $\epsilon$  es un ruido blanco.

La densidad espectral de potencia de Fourier correspondiente a un proceso AR es:

$$P(f) = \frac{b_0}{\left[ 1 + \sum_{k=1}^P b_k e^{2\pi i k f \Delta} \right]^2} \quad (2.11)$$

donde las constantes  $\{b_k; k=1,2,\dots,P\}$  son los coeficientes del proceso de la ecuación 2.14 y el valor  $b_0$  es la varianza del ruido,  $b_0 = \langle \epsilon^2 \rangle$ .

Para escoger el orden del modelo se usa el criterio de información de Akaike (1969), así para cada  $P$  la varianza de los residuos  $S_P$  se calcula con la expresión:

$$S_P = \frac{1}{N-P} \sum_{t=P+1}^N \left[ x(t) - \sum_{k=1}^P b_k x[t-k] \right]^2 \quad (2.12)$$

Generalmente  $S_P$  decrece cuando  $P$  crece. Para tener en cuenta la importancia asociada del ajuste de los datos con el número de parámetros, Akaike (1969) introdujo el FPE, el cual para una serie de tiempo centrada de longitud  $N$  está dado por:

$$FPE_P = \left( \frac{N+P+1}{N-P-1} \right) S_P \quad (2.13)$$

El criterio de selección del orden  $P$  del modelo será aquel para el cual el FPE sea mínimo. El AES puede ser usado para retener los CP significativos estadísticamente de una serie de tiempo. Esto permite obtener modelos AR de bajo orden, sin la presencia de picos falsos.

La predicción lineal (PL) usando los coeficientes AR dados por el MME ha mostrado ser especialmente exitoso en extrapolación de señales suaves y oscilatorias, no necesariamente periódicas.

### **3. Aplicación AES con AR(p) a precipitación y caudales**

#### **3.1 Predicción de precipitación en las cuencas de los embalses del Peñol, Troneras y Piedras Blancas**

Información: se consideraron tres cuencas, la del embalse del Peñol, embalse de Troneras y embalse de Piedras Blancas. Se utilizaron los datos promedios mensuales de precipitación de cada cuenca con registros históricos mensuales de 1966-1992, 27 años.

Para este caso se obtuvo una matriz de covarianzas cruzadas de orden 72x72 (longitud de ventana 24 meses). Se consideran para el análisis los primeros 23 autovalores y autovectores en orden decreciente. El total de la varianza en los 23 autovalores es 81.10%.

En la Figura 1 se pueden ver los autovectores 1 a 3, el autovector 1 presenta ciclos de 24 meses. El primer autovector tiene componentes positivas mientras que los restantes dos autovectores tienen componentes positivas y negativas.

Se determinó la correlación entre el IOS y las primeras cuatro componentes principales y se graficó el IOS y la componente principal con mayor coeficiente de correlación, ver Figura 2, en la cual se observan picos opuestos en todo momento.

**TABLA 1**  
**Correlación IOS y componentes principales 1, 2, 3 y 4, precipitación en las cuencas de los embalses el Peñol, Troneras y Piedras Blancas**

<b>ios vs cp</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
<b>Correlación</b>	-0.068	-0.521	-0.245	-0.014

El análisis para el ajuste de parámetros de los modelos a cada componente principal por el criterio de información de Akaike, dio el resultado que se muestra en la Tabla 2.

**TABLA 2**  
**Número de orden de los AR para las primeras 23 componentes principales, precipitación en las cuencas de los embalses el Peñol, Troneras y Piedras Blancas**

<b>cp</b>	<b>Orden AR</b>	<b>cp</b>	<b>Orden AR</b>
1	85	13	85
2	100	14	150
3	110	15	125
4	100	16	95
5	85	17	75
6	105	18	55
7	85	19	110
8	100	20	80
9	85	21	75
10	100	22	85
11	110	23	80
12	95		

Se hicieron predicciones para diferentes períodos durante los años 1991-1992, se consideraron períodos de verano (diciembre-mayo) e invierno (junio-noviembre) siempre teniendo en cuenta actualizar los parámetros del modelo AR con las componentes principales del período o períodos predichos anteriormente. La calibración de los modelos se hizo con los datos que van de 1966 a 1990, dejando los últimos dos años de información para validar la predicción.

Para determinar la bondad de la predicción se calculó el error cuadrático medio, definiendo como la raíz cuadrada del promedio de la sumatoria al cuadrado de la diferencia entre

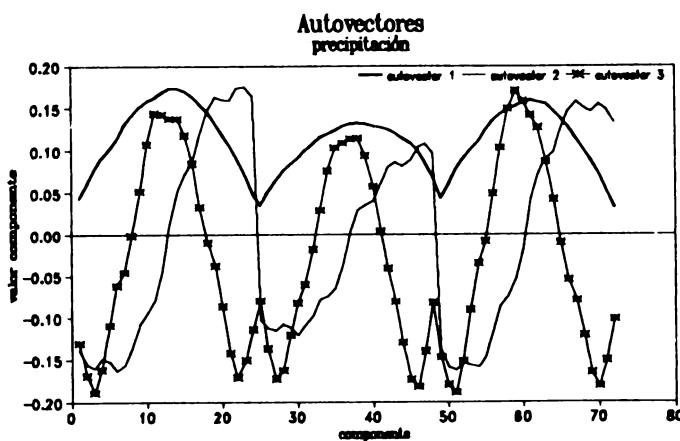


FIGURA 1. Autovectores 1, 2 y 3, caso de precipitación en las cuencas de los embalses del Peñol, Troneras y Piedras Blancas

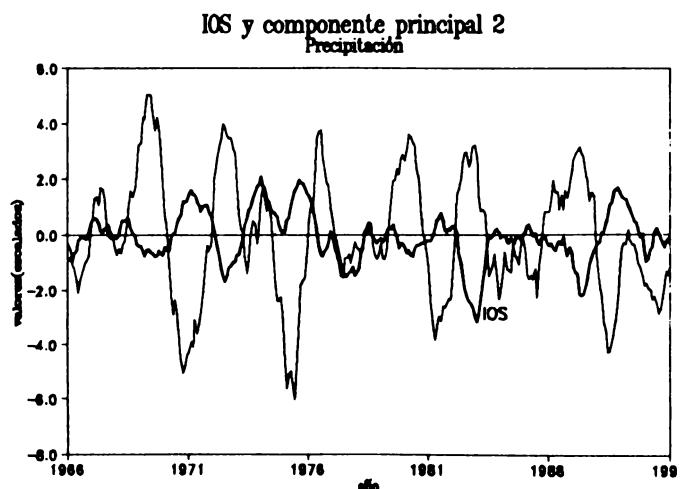
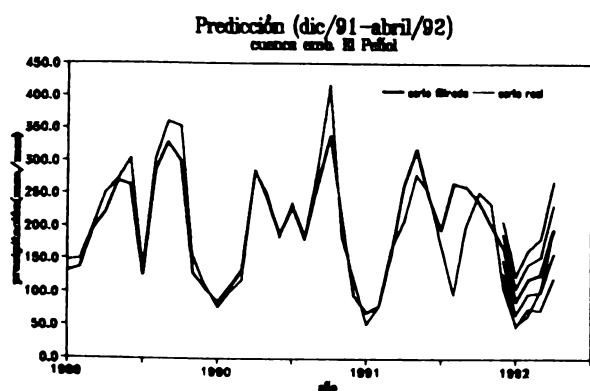


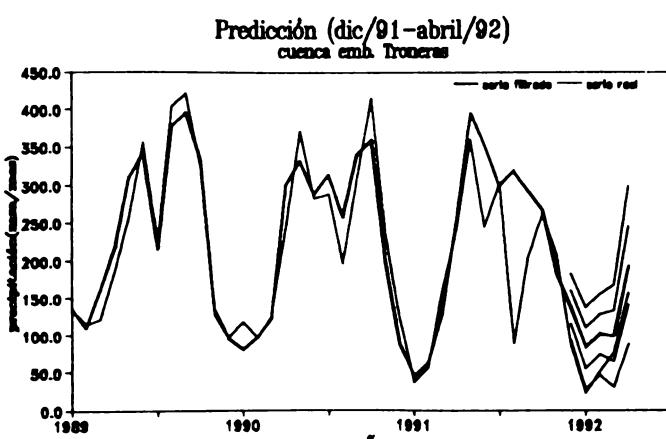
FIGURA 2. IOS y componente principal 2, caso de precipitación en las cuencas de los embalses del Peñol, Troneras y Piedras Blancas.

el valor predicho y el valor real para el período correspondiente, ver Tabla 3. Adicionalmente se calcularon y graficaron bandas de error de una y dos desviaciones a cada lado de la predicción.

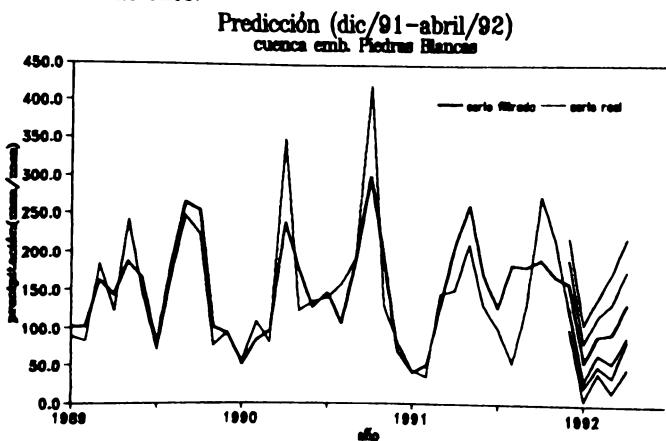
En las Figuras 3, 4 y 5 se presentan las predicciones obtenidas para los períodos de diciembre a abril de 1992 en las tres cuencas.



**FIGURA 3.** Predicción de precip. por AES, cuenca embalse del Pefiol, en noviembre para un horizonte de predicción de cinco meses, bandas de error de una y dos desviaciones.



**FIGURA 4.** Predicción de precip. por AES, cuenca embalse Troneras, en noviembre para un horizonte de predicción de cinco meses, bandas de error de una y dos desviaciones.



**TABLA 3**  
**Raíz del error cuadrático medio, precipitación en las cuencas**  
**de los embalses el Peñol, Troneras y Piedras Blancas**

<b>Raíz del error cuadrático medio (mm/mes)</b>			
<b>Período</b>	<b>Peñol</b>	<b>Troneras</b>	<b>Piedras Blancas</b>
dic - abril / 91	115.21	147.94	86.80
mayo - nov / 91	76.69	111.34	78.14
dic - abril / 92	41.00	46.10	44.29
mayo - nov / 92	75.19	88.73	78.56
dic/90 - dic/92	62.14	77.66	63.65
dic - jun / 91	22.93	49.06	28.02
jul - dic / 91	85.28	108.19	79.11
ene - jun / 92	47.86	69.09	60.33
jul - dic / 92	69.68	73.47	68.63
dic/90 - dic / 91	60.26	81.98	59.62
ene - dic / 92	55.74	66.86	61.43

### **3.2 Predicción de caudales en los ríos Guatapé, Nare, Riogrande, Guadalupe y Piedras**

Información: se dispone de 34 años de registros históricos mensuales de 1959-1992. Se poseen las series completas exceptuando para los ríos Guatapé y Piedras, donde se hizo la siguiente reconstrucción de datos. Para el río Guatapé los datos faltantes se completaron con los datos respectivos de la serie de Guatapé que posee ISA, la cual se hizo con un modelo multivariado. Para el río Piedras los datos faltantes se reconstruyeron con regresiones periódicas entre los meses.

Considerando los cinco ríos se obtiene una matriz de covarianzas cruzadas del orden de 120x120, con una longitud de ventana de 24 meses. La matriz tiene 120 autovalores y autovectores de los cuales se consideran para el análisis los primeros 20 autovalores

y autovectores en orden decreciente. El total de la varianza en los 20 primeros autovalores es 77.36%.

Como en el caso anterior, el autovector 1 presenta una ciclicidad de 24 meses (igual a la longitud de ventana). Los autovectores 2 y 3 están desfasados pero son muy similares. Los tres primeros autovectores están asociados con los mayores autovalores, 25.76, 18.80 y 8.65. En los restantes autovectores las frecuencias son cada vez más altas a medida que disminuye el autovalor que representa la varianza.

Se determinó la correlación entre el IOS y las primeras cuatro componentes principales. De los resultados se observa la relación entre el IOS y la componente principal 2, con coeficiente de correlación igual a -0.624 donde las dos variables son opuestas especialmente en el período de 1959 a 1969 (ver Carvajal, 1994).

**TABLA 4**  
**Correlación IOS y componentes**  
**principales 1, 2, 3 y 4, caudales en**  
**Guatapé, Nare, Riogrande,**  
**Guadalupe, Piedras**

<b>ios vs cp</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
<b>Correlación</b>	0.213	-0.624	0.255	0.011

Una vez seleccionados los primeros 20 autovalores, se obtienen las 20 componentes principales correspondientes. A cada componente se le ajusta un modelo AR, cuyo

orden se obtiene con el estadístico FPE o el criterio de información de Akaike (AIC). Los órdenes obtenidos son:

**TABLA 5**  
**Número de orden de los AR para las primeras**  
**20 componentes principales,**  
**caudales en Guatapé, Nare, Riogrande,**  
**Guadalupe, Piedras**

<b>cp</b>	<b>Orden AR</b>	<b>cp</b>	<b>Orden AR</b>
1	75	11	30
2	50	12	55
3	130	13	50
4	15	14	45
5	85	15	35
6	100	16	45
7	95	17	40
8	130	18	40
9	45	19	50
10	130	20	20

Se hicieron predicciones para diferentes períodos durante los años 1991-1992, siempre teniendo en cuenta actualizar los parámetros del modelo AR con las componentes principales del período o períodos predichos anteriormente. Las predicciones se hicieron considerando veranos de 5 meses e inviernos de 7 meses a partir de diciembre de 1990, así como predicciones durante 6 meses y 12 meses.

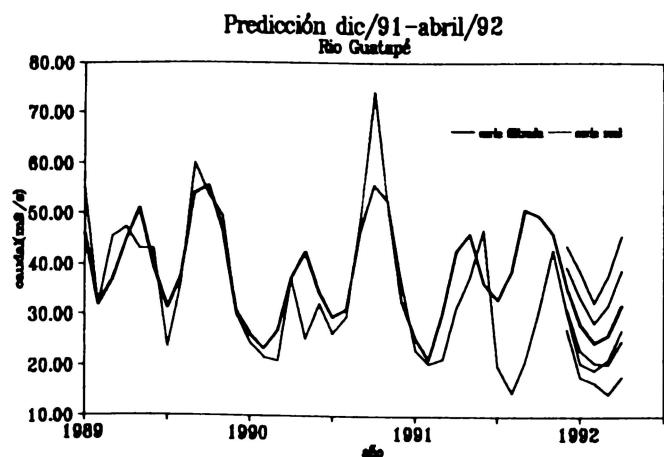
Para determinar la bondad de la predicción se calculó el error cuadrático medio (Tabla

6) y bandas de error de una y dos desviaciones a cada lado de la predicción, en las Figuras 6 a 10 se presentan las predicciones en los diferentes ríos para el período noviembre de 1991 a diciembre de 1992.

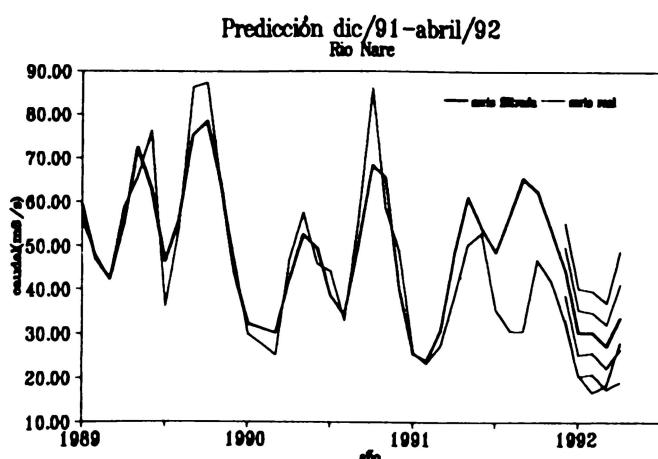
Adicionalmente a esta aplicación se han realizado otras que comprenden: caudales en el río Nare y precipitación en la cuenca del embalse del Peñol y caudales en los ríos Batá, Magdalena (Betania), Gaudalupe, Nare y Salvajina (Carvajal, 1994).

**TABLA 6**  
**Raíz del error cuadrático medio,**  
**caudales en Guatapé, Nare,**  
**Riogrande, Guadalupe, Piedras**

Raíz del error cuadrático medio ( $m^3/\text{mes}$ )	Guatapé	Nare	Riogrande	Guadalupe	Piedras Blancas
Período					
dic - abril / 91	8.73	9.12	5.40	3.49	1.53
mayo - nov / 91	19.01	22.7	9.08	6.46	2.93
dic - abril / 92	5.64	10.3	3.31	0.84	0.88
mayo - nov / 92	11.54	20.8	12.90	5.52	2.69
dic/90 - dic/92	12.75	18.2	9.27	4.81	2.67
dic - jun / 91	9.33	9.9	5.55	3.25	1.49
jul - dic / 91	20.14	24.6	9.28	6.61	3.08
dic - jun / 92	6.32	13.5	7.34	4.72	1.60
jul - dic / 92	11.61	21.4	13.18	4.63	2.71
dic - dic / 91	15.33	18.0	7.35	5.71	2.35
ene - dic / 92	9.06	17.9	10.80	4.57	2.21



**FIGURA 6.** Predicción de caudales por AES, río Guatapé, en noviembre, para un horizonte de predicción de cinco meses, bandas de error de una y dos desviaciones.



**FIGURA 7.** Predicción de caudales por AES, río Nare, en noviembre para un horizonte de predicción de cinco meses, las bandas de error corresponden a una y dos desviaciones.

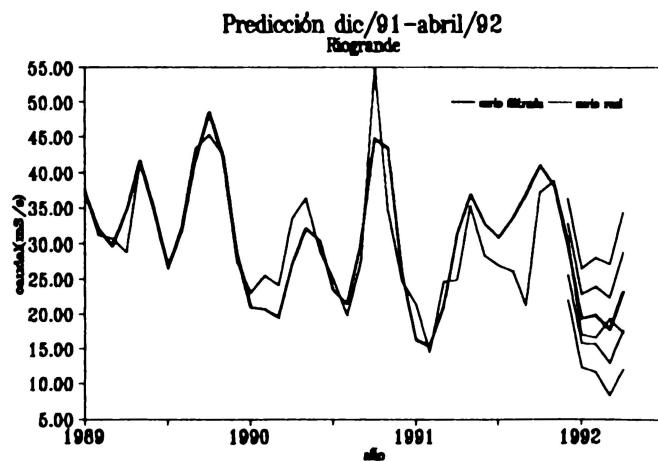


FIGURA 8.

Predicción de caudales por AES, Riogrande, en noviembre para un horizonte de predicción de cinco meses, las bandas de error corresponden a una y dos desviaciones.

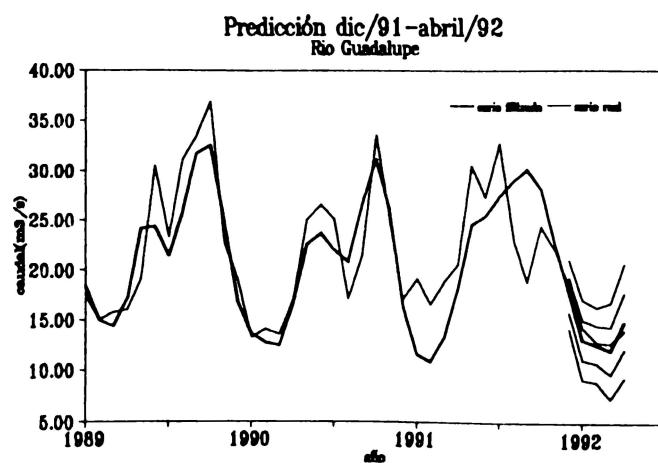


FIGURA 9.

Predicción de caudales por AES, río Guadalupe, en noviembre para un horizonte de predicción de cinco meses, las bandas de error corresponden a una y dos desviaciones.

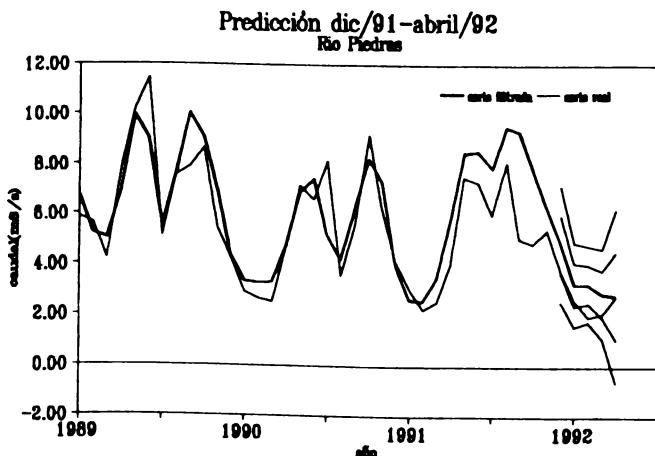


FIGURA 10.

Predicción de caudales por AES, río Piedras, en noviembre para un horizonte de predicción de cinco meses, las bandas de error corresponden a una y dos desviaciones.

#### 4. Red Neuronal, Modelo de retropropagación

Para implementar una red neuronal de retropropagación la configuración representada en la Figura 11 es la más común (Eberhart et al., 1990).

Los subíndices de las variables representan:

i: se refiere a la entrada de la red.

j: se refiere a las capas escondidas.

l: se refiere a la salida de la red.

Las variables representan:

$x_i$ : dato de entrada a la red

$y_l$ : valor de salida de la red

$n$ : nodo de una capa, ya sea de entrada i, escondida h o de salida l.

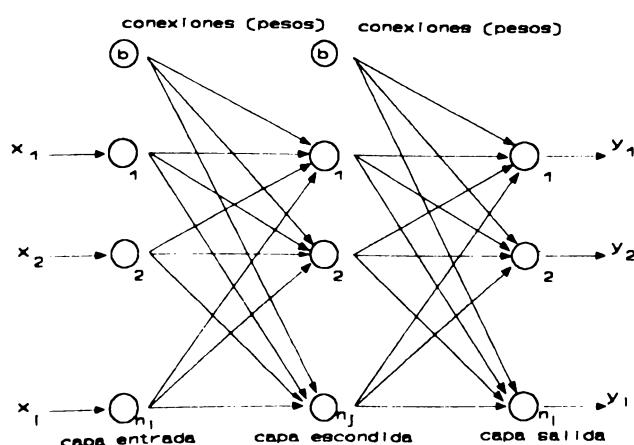


FIGURA 11.

Esquema general del modelo de retropropagación.

b: neuronas sesgadas

w<sub>ij</sub>: peso de la conexión del nodo i al nodo

j.

#### 4.1 Entrada a la red

La entrada a la red puede ser un conjunto de datos, de parámetros o un patrón singular de algún tipo, con la característica de que la entrada debe tener un valor entre cero y uno. Entonces, se puede normalizar las n entradas juntas, normalizar cada canal ( flujo de información a cada neurona de entrada) separadamente o grupos de canales.

Si la entrada consiste en datos puntuales, probablemente se normaliza todos los canales juntos, si la entrada consiste de parámetros, se puede normalizar cada canal o grupos de canales que representen parámetros similares.

En hidrología probablemente la mejor normalización consiste en estandarizar el dato de acuerdo a la media y la desviación estándar (anomalía) y aplicarle luego la transformación sigmoidal.

#### 4.2 Cálculos de alimentación hacia adelante

Una vez transformados los datos, las neuronas de entrada distribuyen la señal en múltiples caminos a las capas escondidas. La salida de una neurona de entrada también está en el rango de cero a uno. Un peso está asociado con cada conexión a una neurona escondida y cada neurona escondida está conectada a una neurona de salida. La información siempre va de izquierda a derecha, como se ve en la Figura 11, de ahí la carac-

terística de que la alimentación de la información sea hacia adelante.

El procesamiento de la señal que entra a las neuronas depende de la función de transferencia usada, generalmente en este tipo de algoritmo se usa la función sigmoidea. En los nodos de la capa de entrada no hay transformación de la señal de la información.

La entrada a una neurona escondida es la suma ponderada por los pesos de las conexiones de las señales que llegan de las neuronas de la capa anterior, ya sea otra capa escondida o la capa de entrada.

$$\text{entrada neta}_j = h_j = \sum_i W_{ji} Y_i \quad (4.1)$$

La salida de una neurona es el resultado de una función sigmoidea que transforma la señal de entrada y da un valor en un rango de cero a uno.

$$\text{salida}_j = y_j = \frac{1}{1 + \exp(-x_j)} \quad (4.2)$$

Para valores grandes positivos de entrada el resultado de la neurona se aproxima a uno y para valores pequeños negativos de entrada el valor de salida se aproxima a cero.

El número de neuronas escondidas  $h_j$  y el valor de  $\alpha$  (coeficiente de aprendizaje) y  $\eta$  (factor de momentum) son los parámetros a calibrar en el entrenamiento de la red. El cálculo de la entrada y salida para una neurona en la capa de salida es igual al de una neurona en la capa escondida.

Cuando se tienen neuronas sesgadas el cálculo es igual, pero su valor de salida siempre

será uno. El cálculo de la entrada y salida para cada neurona, en la fase de entrenamiento como para la fase de ensayo-corrida de la red son iguales.

#### 4.3 Entrenamiento por retropropagación del error

El proceso de aprendizaje involucra el concepto de error de la red que permite cuantificar su aprendizaje. La ecuación 4.3 define el error como:

$$E_p = 0,5 \sum_{j=1}^{n_p} (t_{pj} - y_{pj})^2 \quad (4.3)$$

El error depende del valor de salida calculado  $y_{pj}$  y del valor que supuestamente se debe obtener  $t_{pj}$  al cuadrado, promediado sobre todas las neuronas de la capa de salida. El subíndice  $p$  denota que el valor es para un patrón dado. El valor 0,5 se puede despreciar en los cálculos; esta ecuación fue presentada por Rumelhart y McClelland (1986). La meta del entrenamiento es minimizar  $E_p$  sobre todos los patrones de entrenamiento.

El error  $\delta_i$  para las neuronas de la capa de salida está definido como:

$$\delta_i = y_i(1-y_i)(t_i - y_i) \quad (4.4)$$

Se quiere propagar este error atrás y hacer un ajuste de pesos, acumulando los  $\delta$ s de cada neurona para el entrenamiento completo, adicionándolos, y propagando atrás el error basado en el  $\delta$  total; este procedimiento se llama agrupamiento.

El proceso de actualización de los pesos es el siguiente: Generalmente se inicializan los pesos aleatoriamente en un rango de -0,3 a

0,3; el por qué de este rango lo ha dado la experiencia de las personas que han trabajado al respecto, a veces este rango no funciona y hay que escoger un rango más amplio o más pequeño, dependiendo de la aplicación.

Para actualizar los pesos que alimentan la capa de salida,  $w_{lj}$ , se utiliza la ecuación 4.5, donde  $\eta$  es el coeficiente de aprendizaje,  $\alpha$  es el factor de momentum que ayuda al algoritmo a llegar al mínimo global (solución óptima global), cuyos parámetros varían entre cero y uno.

$$w_{lj}(\text{nuevo}) = w_{lj}(\text{anterior}) + \eta \delta_i y_j \\ + \alpha [\Delta w_{lj}(\text{anterior})] \quad (4.5)$$

Los pesos de las neuronas sesgadas son actualizados en igual forma, como en la ecuación 4.5, ellas sirven como unidades umbrales para las capas a las cuales ellas están conectadas.

Pasando a la actualización de los pesos de las conexiones de las neuronas escondidas, se utiliza el mismo concepto de error empleado en las neuronas de la capa de salida.

Rumelhart y McClelland (1986), mostraron que el error para una neurona escondida es:

$$\delta_i = y_j(1-y_j) \sum_{l=0}^{n_l} w_{lj} \delta_l \quad (4.6)$$

El ajuste de pesos de las conexiones de la capa de entrada a la capa escondida es de forma análoga a la ecuación 4.5.

$$w_{ji}(\text{nuevo}) = w_{ji}(\text{anterior}) + \eta \delta_j y_i \\ + \alpha [\Delta w_{ji}(\text{anterior})] \quad (4.7)$$

En el proceso de ajuste de pesos, primero se calcula el error para cada neurona de salida usando la ecuación 4.4 y para cada neurona escondida usando la ecuación 4.6, para cada patrón en el entrenamiento. Se suman los errores y después que los patrones han sido presentados una vez, se ajustan los pesos con las ecuaciones 4.5 y 4.7.

En el proceso de entrenamiento hay ciertos aspectos que hay que considerar:

1. Los valores de  $\eta$  y  $\alpha$  se pueden asignar capa por capa o neurona por neurona, generalmente un solo valor es seleccionado para cada parámetro en una aplicación. Estos valores son ajustados en el proceso de entrenamiento pero una vez escogidos, no se varían.
2. Para calcular los  $\delta$ s con la ecuación 4 y 6 se usan los pesos anteriores de la capa escondida a la capa de salida y de la capa de entrada a la escondida. Los pesos no son actualizados hasta que todos los patrones han sido presentados.

La selección de los valores de  $\eta$  y  $\alpha$  dependen de la aplicación, otro parámetro a tener en cuenta es el número de iteraciones en el entrenamiento para dar un buen error cuadrático medio.

La red de retropropagación ha sido aplicada en una gran variedad de tareas de reconocimiento de patrones, incluyendo identificación de fenómenos, lectura de textos, reconocimiento de caracteres de letras, compresión de imágenes, diagnóstico médico y selección de técnicas de combate. Es útil también para filtrar una señal y detectar discontinuidades en señales de tiempo variable.

## 5. Aplicacion Redes Neuronales

**5.1 Predicción de caudales en los ríos Guadalupe y Riogrande utilizando variables meteorológicas como: IOS, TSM en la región Niño 3, TSM en la región Niño 8, manchas solares, TSM Atlántico Norte y la oscilación quasi-bienal (OQB).**

Información: se utilizó información mensual de 1950 a 1991. La cual se normalizó por la respectiva media y desviación estándar mensual y se le aplicó la función sigmoidal para obtener la información en el rango de cero a uno.

En este caso se hicieron corridas teniendo en cuenta el rezago de las variables meteorológicas respecto de los caudales 2, 3 y 6 meses. El número de nodos en la capa de entrada es de 6, los nodos en la capa de salida son 2 y en la capa escondida son 4, 2 y 3 respectivamente.

**TABLA 7.**  
**Parámetros de entrenamiento de la red de retropropagación para predecir caudales en Riogrande y Guadalupe**

Rezago (meses)	Tasa de aprendizaje	Factor de Momentun
2	0.04	0.001
3	0.05	0.003
6	0.04	0.001

Para la predicción se consideró en el entrenamiento la información hasta 1990 inclusive, dejando los datos reales de 1991 para comprobar con la predicción.

**TABLA 8.**  
**Error de ajuste en el entrenamiento para**  
**predecir caudales en Riogrande y Guadalupe**

Error de Ajuste			
Iteraciones	Rezago 2	Rezago 3	Rezago 6
100	0.0340	0.0375	0.03540
200	0.0314	0.0319	0.03230
300	0.0311	0.0312	0.03214
400	0.0310	0.0311	0.03209
500	0.0307	0.0310	0.03206

**TABLA 9.**  
**Error cuadrático medio para predicción**  
**de caudales en Riogrande y Guadalupe**

Raíz del error cuadrático medio ( $m^3/s$ )		
Predicción 1991	Guadalupe	Riogrande
Rezago 2	5.36	4.47
Rezago 3	5.09	5.53
Rezago 6	5.01	8.01

El error en la predicción para este caso con la información rezagada 2 meses es de 0.183168 para Guadalupe y de 0.167298 para Riogrande, en el caso de la información rezagada 3 meses el error para Guadalupe es de 0.183485 y para Riogrande es de 0.168434 y para el caso de rezago de la información de 6 meses el error en Guada-

lupe es de 0.183532 y en Riogrande es de 0.174451.

Se presenta las Figuras 12 y 13 de las predicciones contra los valores reales para el caso de rezago de 3 meses, conjuntamente con bandas de error de una desviación a cada lado de la predicción.

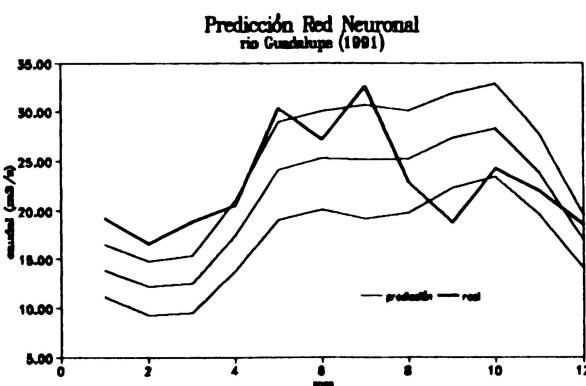


FIGURA 12.

Predicción de caudales por red neuronal, río Guadalupe, en dic. para un horizonte de 12 meses (rezago 3 meses), banda de error corresponde a una desviación.

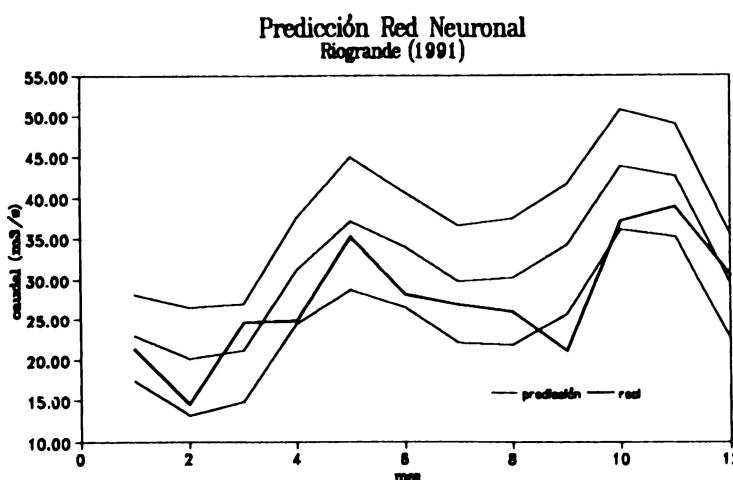


FIGURA 13.

Predicción de caudales por red neuronal, Riogrande, en dic. para un horizonte de 12 meses (rezago 3 meses), banda de error corresponde a una desviación.

## 5.2 Predicción de lluvia en la cuenca del embalse del Peñol

**Información:** El entrenamiento de la red se hizo con datos de precipitaciones promedias mensuales en la cuenca de 1959 a 1992. Como en los casos anteriores los registros mensuales se estandarizaron por su media y desviación estándar correspondiente y

transformados con la función sigmoidea para obtener los datos entre cero y uno.

La red se entrenó con 6 nodos en la capa de entrada, 3 nodos en la capa escondida y 6 nodos en la capa de salida, durante 500 iteraciones para los 65 patrones de entrenamiento disponibles. La tasa de aprendizaje fue de 0.25 y el factor de momentum de

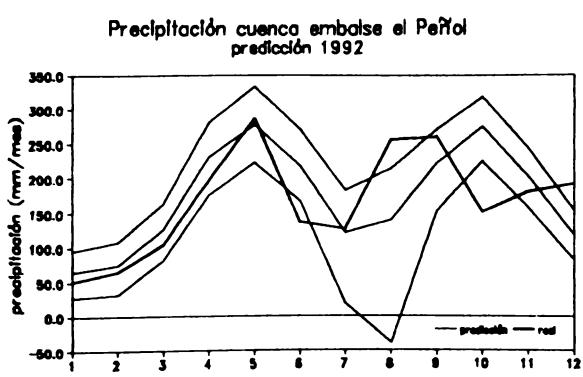
0.075. La información está organizada en ventanas de 6 meses, o sea que en la etapa de entrenamiento, 6 meses sirven para predecir los 6 meses siguientes y estos 6 meses los 6 meses que siguen y así hasta cubrir el registro total de la información.

**TABLA 10.**  
**Error de ajuste en el entrenamiento de la red de retropropagación para predecir lluvia en el embalse del Peñol**

Error de ajuste	
Iteraciones	Error
100	0.0414
200	0.0402
300	0.0397
400	0.0386
500	0.0381

El error en la predicción para este caso corresponde a valor del error de ajuste en la iteración 500 que fue de 0.19519.

Predicción: la predicción se hizo semestral para el año de 1992, evaluando la predicción con el error cuadrático medio que fue de 61.14 mm/mes. Igualmente que en los casos anteriores se graficó la predicción contra los valores reales, Figura 14.



**Figura 14.**

Predicción de lluvia por red neuronal, cuenca emb. del Peñol, en dic. para un horizonte de 12 meses (rezago un mes), banda de error corresponde a una desviación

## 6. Conclusiones y recomendaciones

Tratando de explorar nuevas alternativas de predicción, se han implementado la técnica de análisis espectral singular con AR(P) y las redes neuronales, obedeciendo a una situación coyuntural presentada durante el racionamiento de energía eléctrica en el año de 1992 y principios de 1993.

La expectativa de nuevos modelos busca predecir situaciones extremas en la hidrología colombiana, relacionadas en un alto grado con fenómenos externos a nivel de macroescala como lo es el fenómeno del niño.

Los modelos de AES y redes neuronales tienen como fundamento básico la dinámica no lineal, tratando de modelar el comportamiento de la naturaleza, la cual no obedece a leyes de tipo lineal.

El análisis espectral singular tiene como cualidad principal capturar la varianza representativa de la información, lo que permite quitarle ruido y dejar solo la señal filtrada, haciendo posible ajustar modelos de predicción más simples y de buen resultado. Filosóficamente también se puede interpretar esta metodología como el ajuste en el espacio de fase a las series de tiempo de un polinomio de orden alto para proceder a usar la extrapolación como predicción.

Según los dos casos presentados se puede concluir varias cosas. Los autovectores reflejan la importancia de estos, consecuentemente con los autovalores que representan la varianza del proceso y que están organizados de mayor a menor. El autovector 1 refleja una situación especial, ciclos de longitud igual a la ventana de rezago y que corresponde al mayor autovalor. Es el autovector que tiene mayor peso en las componentes principales por sus valores positivos,

los demás autovectores reflejan ciclos más cortos a medida que disminuye la varianza y con un peso promedio cercano a cero, ya que sus valores son positivos y negativos.

Otro aspecto importante es la correlación de -0.521 entre la componente principal 2 del caso de precipitación y el IOS, así como la correlación de -0.624 entre la componente principal 2 del caso de caudales y el IOS, lo cual deja un interrogante de la influencia del fenómeno del niño en la hidrología colombiana, dejando de presente la necesidad de incorporar nuevas metodologías que ya no miren solamente la parte de caudales, sino también que involucren fenómenos atmosféricos, con parámetros explícitos.

El orden de los AR(P) ajustados por el criterio de información de Akaike es muy grande, una de las causas puede ser el método seleccionado para el ajuste de los AR(P) por máxima entropía, el cual captura las frecuencias más importantes según el número de parámetros estimados, lo que permite capturar ciclos de períodos variables en las componentes principales a utilizar, dando la impresión de que para capturar las frecuencias se requiere de un número grande de parámetros para cada modelo.

Otro aspecto a considerar es el criterio de Akaike, el cual solo tiene en cuenta una sumatoria de residuos multiplicada por un factor que involucra el número de datos y de parámetros y no se hace la selección teniendo en cuenta la función de autocorrelación de las componentes principales.

Las predicciones presentan un comportamiento mejor para la precipitación que para los caudales, si se compara con las trazas reales para los dos años de predicción, con mejores resultados para el período diciembre-abril que para mayo-noviembre. Así como esta aplicación involucra AR(P) por má-

xima entropía, también se puede considerar ajuste de otros modelos por mínimos cuadrados, o por correlaciones parciales y con muchos menos parámetros. Las bandas de error involucraron un cálculo del error en la predicción de la parte filtrada y un cálculo proveniente del error en la predicción de los residuos entre la información filtrada y la real. También se observa en los resultados, que la predicción empieza un poco por encima del valor real, siendo posible mejorarla aunque en estos casos las series de residuos no permitieron tal ajuste ya que mostraron que eran aperiódicas.

Adicionalmente el explorar una longitud de ventana mayor a 24 meses (por ejemplo 60 meses), como posibilidad para considerar cambios hidrológicos con persistencia mayor a 2 años, está restringida por la parte computacional, representada en la exigencia de memoria y subrutinas rápidas que den soluciones estables de autovalores y autovectores de la matriz de covarianza.

La aplicación del modelo de retropropagación utilizando como datos de entrada variables meteorológicas para predecir caudales, arrojan una tasa de aprendizaje y factor de momentum bajos, lo que implica que el cambio de pesos en el proceso de entrenamiento ocurre en un rango pequeño, y que los pesos aleatorios iniciales entre 0.3 y -0.3 son una buena aproximación inicial para empezar a entrenar la red.

La predicción de los ríos Guadalupe y Río Grande con variables meteorológicas es buena para Río Grande y no tanto para Guadalupe, sin tener en cuenta el rezago de la información. Para Río Grande la predicción está generalmente por encima de los valores reales notándose una gran diferencia en el mes de septiembre que es de bajo caudal real y que la red no predice bien. Para Guadalupe, la predicción está por debajo de los

valores reales, siendo buena para los últimos tres meses del año, aunque sigue la tendencia en el año deja de predecir los valores extremos de mayor caudal en los meses de mayo y julio, esta situación puede ser causada porque en el año específico de predicción 1991, se tuvo una hidrología atípica no presente en los registros históricos, especialmente en el mes de julio y por lo tanto no capturada por la red neuronal en su etapa de entrenamiento.

La Figura 14, presenta una predicción muy buena para los primeros cinco meses del año 1992, ya para el segundo semestre del año la predicción es muy mala, presentando un desfasamiento en el pico.

Lo anterior hace esta técnica de redes neuronales factible de usar para predicción, además la disponibilidad de información meteorológica es muy rápida y se puede obtener de diversas fuentes.

Para alimentar la red se debe disponer de un número de patrones grande, ya que la cantidad y calidad de la información es esencial para que la red pueda tener un proceso de aprendizaje satisfactorio.

La no linealidad representada en la red neuronal por una función de transformación sigmoidea, puede ser más evidente de aplicar a nivel de predicción que de simulación.

El modelo de red neuronal de tipo retropropagación, puede ser atractivo porque no requiere estimar ningún parámetro inicial a partir de los datos a excepción de los pesos de las conexiones entre nodos de las distintas capas, que son aleatorios y que el proceso de entrenamiento ajusta.

El modelo puede servir para predecir situaciones críticas a corto plazo de períodos no mayores a 6 meses.

La simulación con redes neuronales puede ser más complicada desde el punto de vista de la información para alimentar el modelo ya que se tendría que disponer de alguna metodología de simulación para alimentarlo.

El cálculo del margen de error para las predicciones con la red neuronal se basa en la raíz del error de ajuste durante la calibración, teniendo en cuenta la transformación inversa sigmoidea y la desestandarización.

Un aspecto importante en ambos modelos es el punto en el tiempo de iniciación de la predicción y la longitud de la ventana de predicción, lo cual involucra un buen conocimiento de la hidrología de la zona, para poder trabajar con los períodos intranuales críticos y con una ventana de predicción adecuada que cubra la longitud de esos períodos, y que los resultados obtenidos estén siempre sujetos a una mirada crítica.

## 7. Referencias

- Akaike, H. (1969). Fitting autoregressive models for prediction, Ann. Inst. Stat. Math., 21, 243-247.
- Allen, M. R., Read, P. L. y Smith, L. A. Bidecadal oscillations and the global temperature record, Nature 355 (1992) 686.
- Amari, S. (1972). "Learning patterns and patterns sequences by self organizing nets of threshold elements," IEEE Trans. Computers, C-21, 1197-1206.
- Anderson, J.A., Silverstein, J.W., Ritz, S.A., y Jones, R.S. (1977). "Distinctive features, categorial perception, and probability learning: some applications

- of a neural model, "Psych. Review, 84, 413-451, y reimpreso en J.A. Anderson y E. Rosenfeld (Eds.) (1998). Neurocomputing, MIT Press, Cambridge, MA.
- Box, G.E.P. y Jenkins, G.M. (1970). Time Series Analysis, Forecasting and Control, Holden-Day, San Francisco.
- Broomhead, D. S., y King, G. P. (1986). Extracting qualitative dynamics from experimental data, Physica D, 20, 217-236.
- Burg, J.P. (1967). Maximum entropy spectral analysis, Presentation at 37th annual meeting of the society of exploratory Geophysicists, Oklahoma City, reimpreso en Modern Spectrum Analysis (1968), D. G. Childers (ed.), 43-48.
- Carpenter, G.A., y Grossberg, S. (March, 1988). "The ART of adaptive pattern recognition," Computer, 77-88.
- Carvajal, L. F. Acerca de la Predicción No lineal en Hidrología. Tesis de Magíster en Aprov. de Rec. Hid., Facultad de Minas, Medellín, 1994. 136 p.
- Colebrook, J. M. Continuous plankton records: zooplankton and environment, North-East Atlantic and North Sea, 1948-1975, Oceanol. Acta, 1, 9-23, 1978.
- Eberhart, Russell C. y Dobbins, Roy W (1990). Neural Network PC Tools, A Practical Guide. Academic Press, San Diego, California, 35-49.
- Farley, B. y Clark, W. A. (1954). "Simulation of self-organizing systems by digital computer," IRE Transaction of information Theory, 4, 76-84.
- Fraedrich, k. (1986). Estimating the dimensions of weather and climate attractors, J. Atmos. Sci., 43, 419-432.
- Fukushima, K., Miyake, S. y Ito, T. (1975). "Neocognitron: a neural network model for a mechanism of visual pattern recognition," IEEE Transactions on Systems, Man, and Cybernetics, SMC-13, 826-834.
- Hopfield, J.J. (1982). "Neural network and physical systems with emergent collective computational abilities", Proc. of the National Academy of Sciences, USA, 79, 2554-2558.
- Hopfield, J.J., y Tank, D. (1986). "Computing with neural circuits: A model," Science, 233, 625-633.
- James, W. (1980). Psychology, Holt, New York, chapter XVI, 253-279.
- Keppenne, C.L. y Ghil, M. (1992). Adaptive filtering and prediction of the southern oscillation index. J. Geophys. Res., sub judice.
- Kohonen, T. (1972). "Correlation matrix memories," IEEE Trans. Computer C-21, 353-359.
- Kosko, B. (1987c). "Fuzzy Associative Memories," Fuzzy Expert Systems, Addison-Wesley, Reading, MA.
- Kosko, B. (1987a). "Adaptive bidirectional associative memories," Applied Optics, 26, 4947-4960.
- Kosko, B. (1987b). "Constructing and Associative Memory," BYTE, Sept., 137-144.

- Kosko, B. (1986). "Fuzzy cognitive maps," International Journal of Man-Machine Studies, 24, 65-75..
- LeCun, Y. (1986). "Learning processes in an asymmetric threshold network," Disordered Systems and Biological Organization, E. Beinenstock, Fogelman Souli, and G. weisbuch (Eds.), Springer, Berlin.
- Loève, M. (1962). Probability Theory, 3rd ed, Van Nostrand, Princeton.
- Maren, Alianna; Harston, Craig y Pap, Robert (1990). Handbook of Neural Computing Applications. Academic Press, San Diego, California, 13-28, 45-57.
- McCulloch, W.S. y Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity, "Bulletin of Mathematical Biophysics, 5, 115-133.
- Parker, D. (1985). "Learning logic," Technical report TR-87, Center for Computational Research in Economics and Management Science, MIT Press, Cambridge, MA.
- Penland, C., Ghil, M. y Weickmann, K. M. (1991). Adaptive filtering and maximum entropy spectra with application to changes in atmospheric angular momentum. J. Geoph. Res., 96, No. D12, 22,659-22,671.
- Press, William H. Numerical Recipes, The Art of Scientific Computing. Cambridge University Press, 1989.
- Rochester, N., Holland, J. H., Haibt, L. H. y Duda, W. L. (1956). "Test on a cell assembly theory of the action of the brain, using a large digital computer, "IRE Transactions of Information Theory, IT-2, 80-93.
- Rosenblatt, F. (1958). "The Perceptron: A probabilistic model for information storage and organization in the brain," Psych. Review, 65, 386-408, y reimpresso en Neurocomputing, J. Anderson y E. Rosenfeld (Eds.) (1988). MIT Press, Cambridge, MA., 92-114.
- Rumelhart, D. E. y Mc Clelland, "Parallel Distributed Processing, Explorations in the Microstructure of the Cognition. Vol. 1: Foundations." MIT Press, Cambridge, MA, 1986.
- Tank, D. W., y Hopfield, J.J. (1987). Collective computation in neuronlike circuits," Scientific American, 257, 104-114, 158.
- Vautard, R., Yiou, P. y Ghil, M. (1992). Singular-spectrum analysis: A toolkit for short, noisy chaotic signals. Physica D 58, 95-126.
- Vautard, R. y Ghil, M. (1989). Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series. Physica D 35, 395-424.
- Werbos, P. (1974). Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences. Ph.D. dissertation, Harvard University, Cambridge, MA.

