

## Predicción de caudales medios mensuales en ríos de Antioquia

Joany Sánchez Molina

Ingeniero , Estudiante de Maestría en Ingeniería- Recursos Hidráulicos  
jsanche@unal.edu.co

Germán Poveda Jaramillo

Ingeniero Civil, M. Sc., Ph.D.  
Profesor Asociado escuela de Geociencias y Medio Ambiente  
Universidad Nacional de Colombia Medellín  
gpoveda@unal.edu.co

**Resumen:** Se implementan tres metodologías para el pronóstico de caudales medios mensuales en cuatro ríos de Antioquia (Nare, Guatapé, Guadalupe y Río Grande II): (i) el método no paramétrico MARS ("Multivariate Adaptive Regression Splines"), (ii) un método ARIMA generalizado no paramétrico, y (iii) el método paramétrico de "Holt-Winters" durante el período 1990-2004. El método MARS permite incorporar la persistencia hidrológica y la influencia de fenómenos macroclimáticos como el ENSO. Se evalúan los métodos de predicción usando intervalos de pronóstico de 1 mes, para horizontes de pronóstico de 12 meses. La bondad del pronóstico se cuantifica a partir del error cuadrático medio y del porcentaje de aciertos por terciles, dependiendo del mes en el cual se inicia el pronóstico. Los mejores resultados según el criterio del error cuadrático medio los presenta MARS. Se evidencian importantes ganancias en la capacidad de pronóstico al usar la información del Chorro del Chocó (Poveda, 1998) como variable predictora.

### 1. Introducción

Las metodologías de regresión son ampliamente utilizadas en todos los campos de la ciencia como herramientas de análisis de datos; fundamentalmente en el modelamiento y predicción de series. En general una regresión se utiliza para desarrollar un modelo matemático que relaciona variables dependientes ó respuesta con variables explicativas o de entrada. Estos métodos cobran gran importancia en la predicción de caudales medios mensuales, en países como Colombia, donde la mayor parte de la energía generada es de carácter hídrico, lo cual pone de presente la implementación de modelos de predicción hidrológica.. En el Posgrado en Aprovechamiento de los Recursos Hidráulicos se han llevado cabo diversas investigaciones en los que se implementan varias metodologías de predicción (ver Poveda et al., 2002; Poveda et al., 2003)

Este trabajo tien por objeto implementar tres metodologías de predicción de caudales medios mensuales en ríos de alta importancia para la generación hidroeléctrica de Colombia, incluyendo los ríos

Nare, Guatapé, Guadalupe y Grande I. El Método MARS, no paramétrico, y metodologías de Holt-Winters y ARIMA Generalizado. En el numeral 2 se describen los datos usados, en el numeral 3 se describen los métodos de pronóstico y en el numeral 4 se avalúan los resultados obtenidos por las diversas metodologías, usando diversos criterios de estimación de los errores de pronóstico, tales como el error medio cuadrático, "RMS", y el criterio de acierto por terciles. Finalmente, se hace una discusión sobre la bondad de los métodos de pronóstico usados.

## 2. Información utilizada

Las metodologías de predicción fueron evaluadas usando los registros de caudales medios mensuales de los ríos Nare, Guatapé, Guadalupe y Grande II. Los registros mensuales para los ríos Guadalupe y Grande II corresponden al período común entre 1951-2004, mientras que para los ríos Nare y Guatapé corresponden al período entre 1959-2004. El período de calibración de los modelos se seleccionó desde el inicio de las series hasta diciembre de 1989; los siguientes 15 años fueron utilizados en la validación de los modelos

Las variables macroclimáticas usadas como variables predictoras son: Índice Multivariado del ENSO (MEI), el Índice de Oscilación del Sur (SOI) y la Temperatura Superficial del Mar en las regiones NIÑO 3 y 4 y la variable Choco Zonal. En el cálculo del Índice Multivariado del ENSO (Wolter, 1987 en Poveda *et al*, 2002), se utilizan 6 variables en las cuales tiene gran influencia el fenómeno ENSO (presión a nivel del mar, componentes zonales y meridionales del viento superficial, SST, temperatura del aire en la superficie y nubosidad total), mientras que el Índice de Oscilación del Sur se calcula como la diferencia estandarizada de presiones atmosféricas entre Tahití y Darwin (Trenberth, 1976 en Poveda *et al.*, 2002). El chorro del oeste de Colombia ó Chorro del CHOCÓ (Chorro del Occidente Colombiano) es una corriente de vientos que ejerce una fuerte influencia sobre la climatología de Colombia particularmente en las regiones occidental y central del país (Poveda, 1998; Poveda y Mesa, 2000).

## 3. Métodos de Pronóstico

### 3.1 Método de Holt-Winters

El método de Holt-Winters es básicamente un procedimiento de suavización exponencial. Este tipo de procedimientos facilitan los cálculos y reducen los requerimientos de almacenamiento en las bases de datos, lo cual cobra importancia cuando se están prediciendo muchas series de tiempo.

Los modelos de suavización exponencial están basados en la actualización, para cada período, de hasta tres parámetros:

- media (modelo de suavización simple).
- Media y tendencia.

- Media, tendencia y estacionalidad (modelo de Holt-Winters).

Los modelos anteriormente mencionados son conocidos en la literatura como de suavización exponencial de uno dos y tres parámetros respectivamente (Segura y Vercher, 2001)..

### 3.2 Polinomios de Regresión Adaptivos Multivariados (MARS)

MARS ("Multivariate Adaptive Regression Splines ") es una implementación de las técnicas propuestas por Friedman (1991) para resolver problemas de regresión múltiple no paramétrica a priori. Tales técnicas tienen como objetivo el predecir valores de una variable continua dependiente o de salida, a partir de un grupo de variables independientes llamadas comúnmente variables predictoras. El método MARS es un método de regresión no paramétrica y no lineal, que está basado en una generalización del particionamiento recursivo, planteando algunas generalizaciones de los procedimientos, garantizando modelos continuos con derivadas continuas. En Friedman (1991) se encuentran descritos los principales aspectos de la modelación no paramétrica y de la computación adaptativa. Además, allí se presenta completamente el algoritmo de ajuste de MARS.

### 3.3 Método ARIMA Generalizado

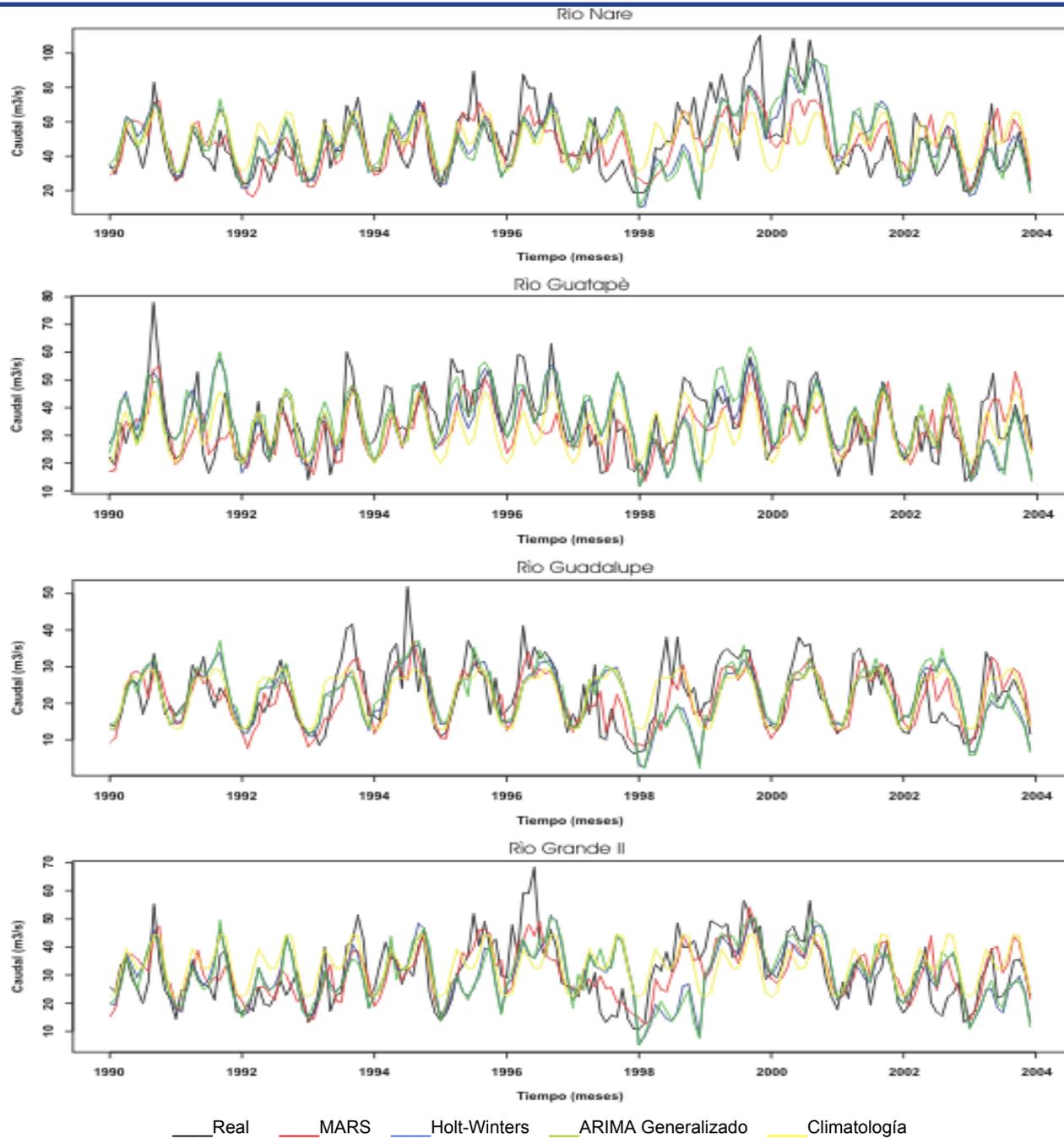
Está basado en la descomposición de la serie de tiempo en un determinado número de componentes, los cuales se especifican por un sistema de varianzas en el error (dar una referencia). Algunos de ellos pueden llegar a ser cero. El método permite elegir entre un modelo ARIMA (0,1,1), y el software usado tiene incorporado un modelo que permite modelar la posible existencia de tendencias lineales en el tiempo. El tercer modelo incorporado es el llamado modelo básico estructural, el cual presenta tendencia local y una componente estacional adicional.

### 3.4 Climatología

Es el más simple de todos los métodos de predicción y consiste en tomar como valor de predicción para cada mes el promedio mensual multianual.

## 4. Resultados y análisis.

Se hacen predicciones ciegas con un mes de inicio determinado, y con una ventana de hasta doce meses. En la figura 1 se muestran las series de tiempo predichas por los diferentes métodos y las series observadas. Sólo se muestran los resultados de las predicciones con una ventana de 12 meses, iniciando siempre en el mes de enero.

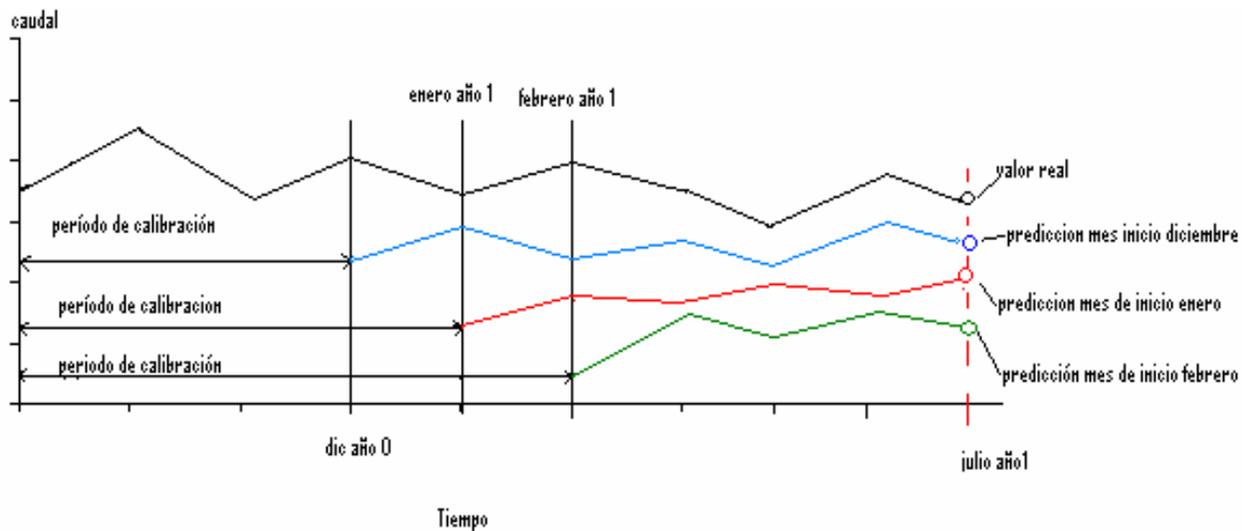


**Figura 1. Series de caudales medios mensuales en el período de validación.**

El desempeño de los modelos es evaluado para todas las series mediante el error cuadrático medio (RMS), el cual está definido como se muestra a continuación:

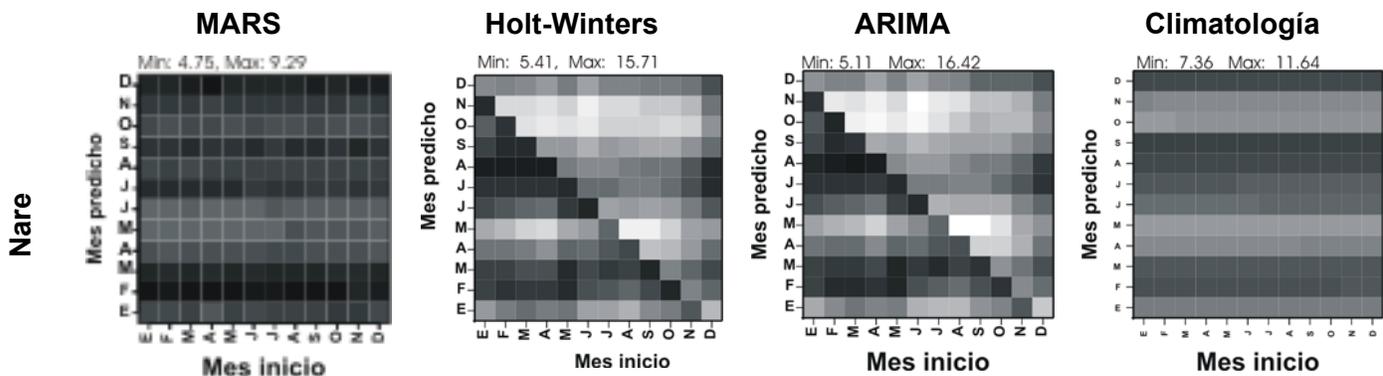
$$\%RMS = \frac{\sqrt{\frac{1}{n} \sum (Q_{real} - Q_{predicho})^2}}{\bar{Q}_{real}} \times 100 \quad (17).$$

Así por ejemplo se quiere evaluar el error medio cuadrático (RMS) de predecir julio de cualquier año comenzando en diciembre del año anterior, el criterio lo que hace es establecer una diferencia entre el valor real del caudal de julio de ese año y el valor predicho para julio de ese mismo, los valores de las predicciones de todos los Julios con mes de inicio diciembre en todo el período de validación son almacenados en un vector. La n de la fórmula lo que representa es el numero de años de validación de los modelos que para nuestro caso es igual a 15, que es igual al numero de Julios predichos con mes de inicio diciembre. El procedimiento anterior es análogo para todas las predicciones con mes de inicio enero, febrero, marzo, etc.. Es importante tener en cuenta que el período de calibración se va actualizando a medida que se va pasa de un mes de inicio a otro. En la figura 2 se ilustra lo planteado anteriormente.

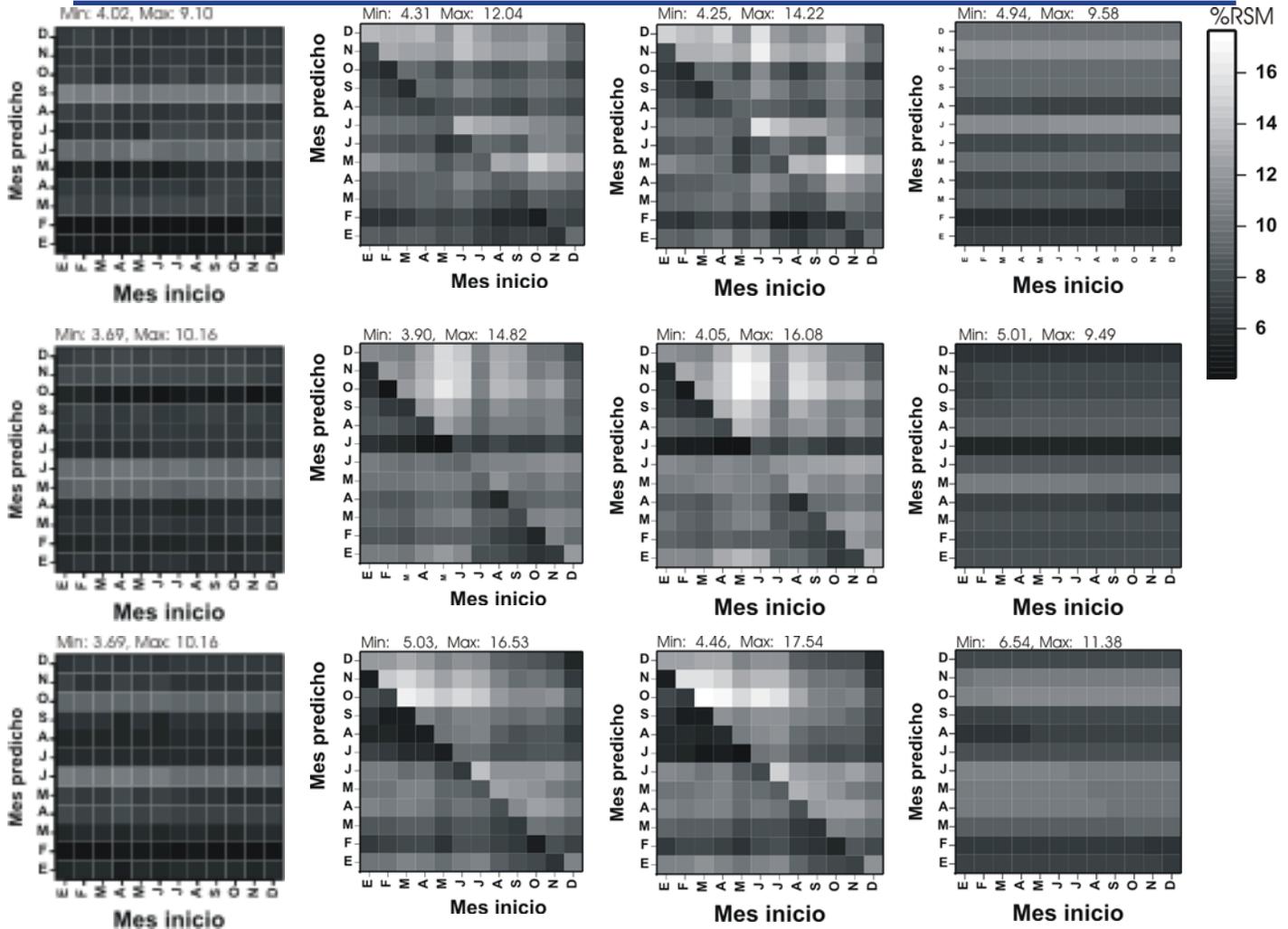


**Figura 2. Actualización de los periodos de calibración para cada mes de inicio**

En la figura 3 se muestran los resultados obtenidos por esta metodología. Donde para los diferentes ríos se compara el RMS promedio de las predicciones hechas con inicio en cada uno de los meses de todos los años en el período de validación de las metodologías, con una ventana de predicción de doce meses. En estos gráficos se presenta el mes de inicio de la predicción en las abscisas y el mes que se va a predecir en las ordenadas.



Gutapé



**Figura 3. Porcentaje de RMS de las predicciones en el período de validación**

Los valores más altos del RMS se presentan para el río Grande II, mientras que el río Guatapé se muestra como el de mejor predictibilidad. El resultado obtenido es de gran importancia pues El río Guatapé al igual que el Nare alimentan el embalse de El Peñol que es el de mayor capacidad de regulación del país, esta información puede ser consultada en la página web (<http://www.eppm.com/epmcom/contenido/acercade/infraestructura/generacion/Centraleshidroelectricas/cenralguatape.htm>).

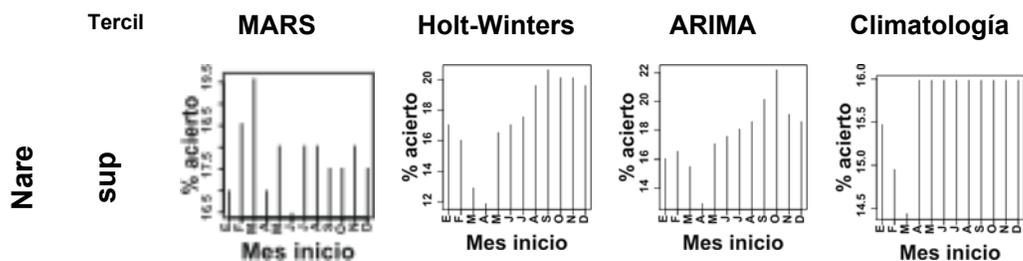
Por medio del análisis de los errores es posible concluir, en general, que MARS presenta los menores errores de pronóstico, mientras que los métodos ARIMA Generalizado y Holt-Winters presentan los más altos. También es posible afirmar que el método MARS presenta una gran mejoría en el caso del río Nare al compararse los resultados con los obtenidos por Poveda et al. (2002), debido posiblemente a la mayor longitud de las series de calibración que se han utilizado en este trabajo y que además en este ejercicio se está haciendo uso de la variable que representa el Chorro del Chocó, la cual es de gran influencia en el clima del occidente del país (Poveda, 1998).

En general, las mayores dificultades en pronóstico se presentan en la predicción del trimestre Mayo-Junio-Julio. Todas las metodologías, con excepción de MARS tienen problemas en la predicción para el trimestre Noviembre-Diciembre-Enero.

Todos los ríos en estudio presentan una dificultad parecida en lo que se refiere a la predicción de sus series de caudales medios mensuales. Sin embargo solo el método MARS se comporta de manera coherente en períodos con presencia del fenómeno de El NIÑO o de La NIÑA, tal efecto se puede observar claramente durante el periodo de validación comprendido entre los años 1998 y 2000, que corresponden a un periodo del fenómeno de la NIÑA, durante el cual los métodos de Holt-Winters y ARIMA hacen estimaciones muy pobres de los caudales medios mensuales, mientras que MARS se aproxima bastante bien a los valores reales de los caudales durante dicho período, lo anteriormente expuesto se puede apreciar en la figura 1.

Para la estimación de los errores de pronóstico por terciles, se puede comparar si el valor de la variable pronosticada corresponde al tercil observado. En caso afirmativo, se considera como un acierto. Un valor de 100% implica que todos los valores predichos para un tercil determinado, realmente se observaron en ese tercil. Para realizar este análisis es necesario conocer los límites que permiten discriminar si un valor determinado está en el tercil superior, medio o inferior (Poveda *et al.*, 2002). Al aplicar este criterio de error a las series obtenidas para todos los ríos, y para cada método se obtienen comportamientos muy diversos de la bondad del método bajo este criterio de error. Así para el río Nare el método que mejor se comporta según este criterio es la climatología pues muestra los porcentajes de acierto más altos para todos los terciles y en la mayoría de los meses, entre tanto para el río Guatapé, los mejores resultados para este criterio los muestra la climatología, solo es sobre pasada en el tercil inferior, en el cual los mejores resultados los presentan los métodos de Holt-Winters y ARIMA Generalizado.

Para el río Grande II y el río Guadalupe los resultados son más variados pues no hay un método que se mejor que otro, dado que para algunos meses y de acuerdo al tercil hay métodos que son superiores a los otros. Los resultados obtenidos para este criterio se muestran en la figura 4.



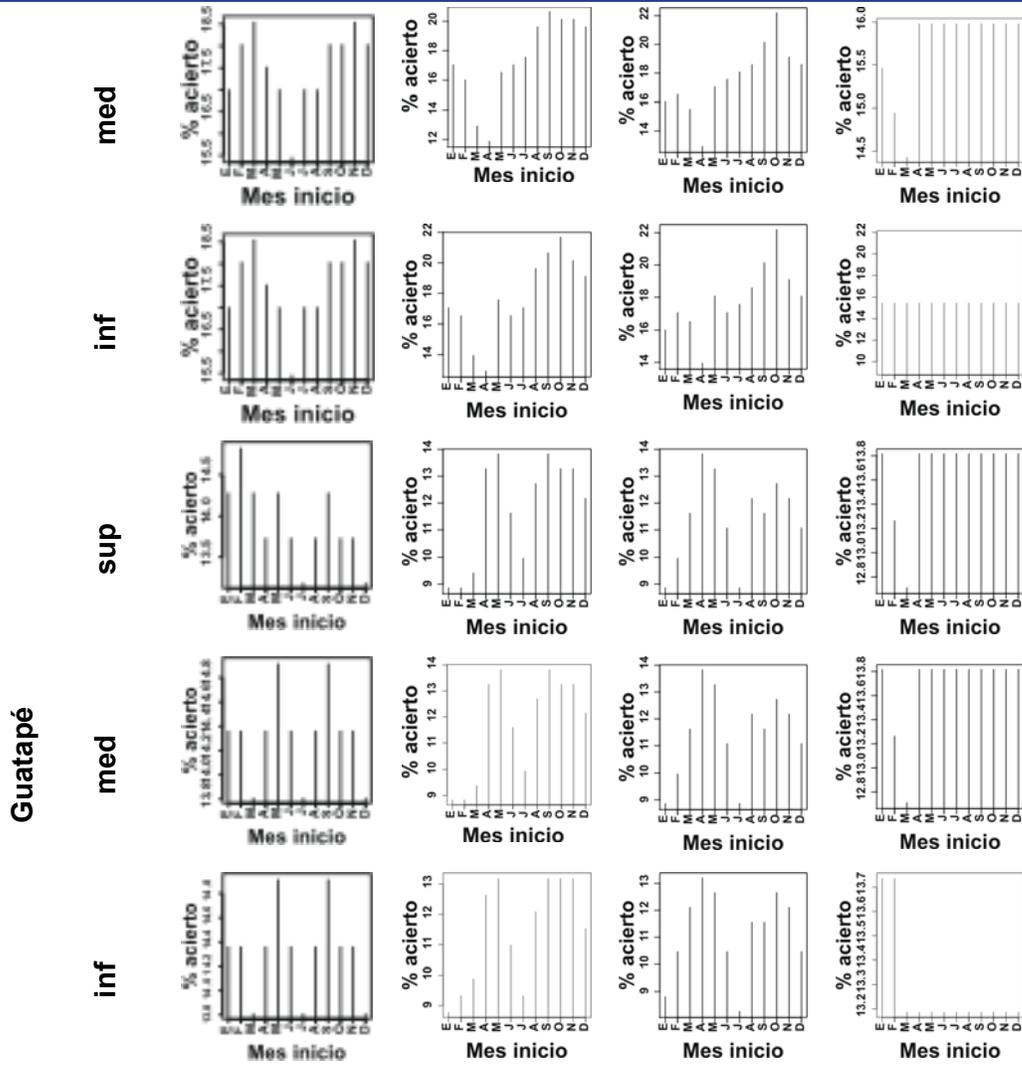


Figura 4 Porcentaje de acierto por terciles

	Tercil	MARS	Holt-Winters	ARIMA	Climatología
Guadalupe	sup				
	med				

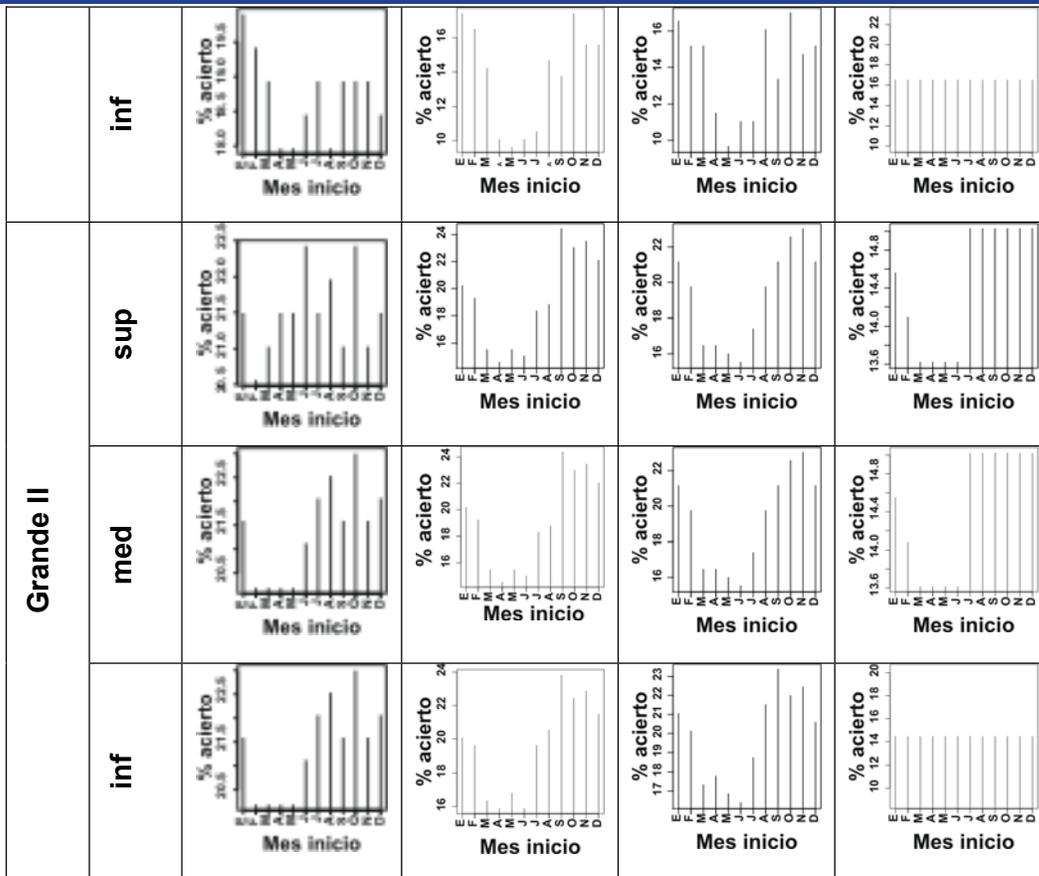


Figura 4 Porcentaje de acierto por terciles

**Conclusiones**

En este trabajo se han comparado diferentes metodologías de predicción, entre las cuales se encuentran las de tipo paramétrico y no paramétrico. MARS que corresponde a las metodologías de tipo no paramétrico es capaz de incorporar la influencia de fenómenos macroclimáticos como el ENSO.

Los resultados muestran que el mejor método de predicción es MARS, Pues para este se presenta el menor RSM, en cada uno de los ríos, y además presenta resultados parecidos en el acierto por terciles a las otras metodologías. Los resultados también indican que en un clima como el de Colombia, que es altamente influenciado por fenómenos macroclimáticos como el ENSO, se deben utilizar metodologías de predicción no paramétricas tales como MARS, las cuales incluyen la influencia de estos fenómenos en sus predicciones de caudales medios mensuales.

El método de predicción MARS presenta notables mejorías al incorporarse la variable macroclimática Chocó Zonal, la cual introduce en las predicciones la influencia del chorro del chocó, que es de gran importancia en la hidroclimatología de la zona occidental del país. De este modo se obtienen reducciones importantes del error cuadrático medio en comparación con los reportados por Poveda et al. , 2002

**Agradecimientos.** A Jesús David Gómez por su extensa colaboración.

### REFERENCIAS BIBLIOGRÁFICAS.

Friedman, J. H. *Multivariate Adaptive Regression Splines*, Stanford University. Stanford, CA, 1991.

Harvey, A. *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press. Cambridge. 1992.

Poveda, G., Retroalimentación dinámica entre el fenómeno El Niño- Oscilación del Sur y la hidrología de Colombia, Tesis Ph.D., Universidad Nacional de Colombia, 1998.

Poveda, G., O. J. Mesa, L. F. Carvajal, C. D. Hoyos, J. F. Mejía, L. A. Cuartas y A. Pulgarín. Predicción de caudales medios mensuales en ríos colombianos usando métodos no lineales. *Meteorología Colombiana*, 6, 101-110, 2002.

Poveda, G., O. J. Mesa, and P. R. Waylen, Non-linear forecasting of river flows in Colombia based upon ENSO and its associated economic value for hydropower generation. En: *Climate and Water: Transboundary Challenges in the Americas*, H. Diaz and B. Morehouse (eds.), Kluwer Academic Publishers, Dordrecht, 351-371, 424 p., 2003.

Segura, J. V y Vercher, E., 2001. A spreadsheet modeling approach to the Holt-Winters optimal forecasting. *European Journal of Operational Research*. Vol 131, 375-388.

Salas, J.D. Delleur, J. Yevjevich, V. y Lane, W, *Applied Modeling of Hydrologic Time Series*. Water resources publications. Littleton, Colorado. USA. 1980. 484p.