

PREDICCIÓN NO LINEAL DE CAUDALES MEDIOS MENSUALES EN AMERICA DEL SUR

Modelos Matemáticos - Estadísticos

Julián D. Rojo H,

Universidad Nacional de Colombia - Sede Medellín
Escuela de Geociencias y Medio Ambiente

San José - Costa Rica
Septiembre 09 de 2012





A variabilidade interanual das vazões dos rios na América do Sul depende do comportamento do clima na escala global.



Entender essa variabilidade, suas causas, seus impactos e, sobretudo, sua previsibilidade é o que nos queremos .

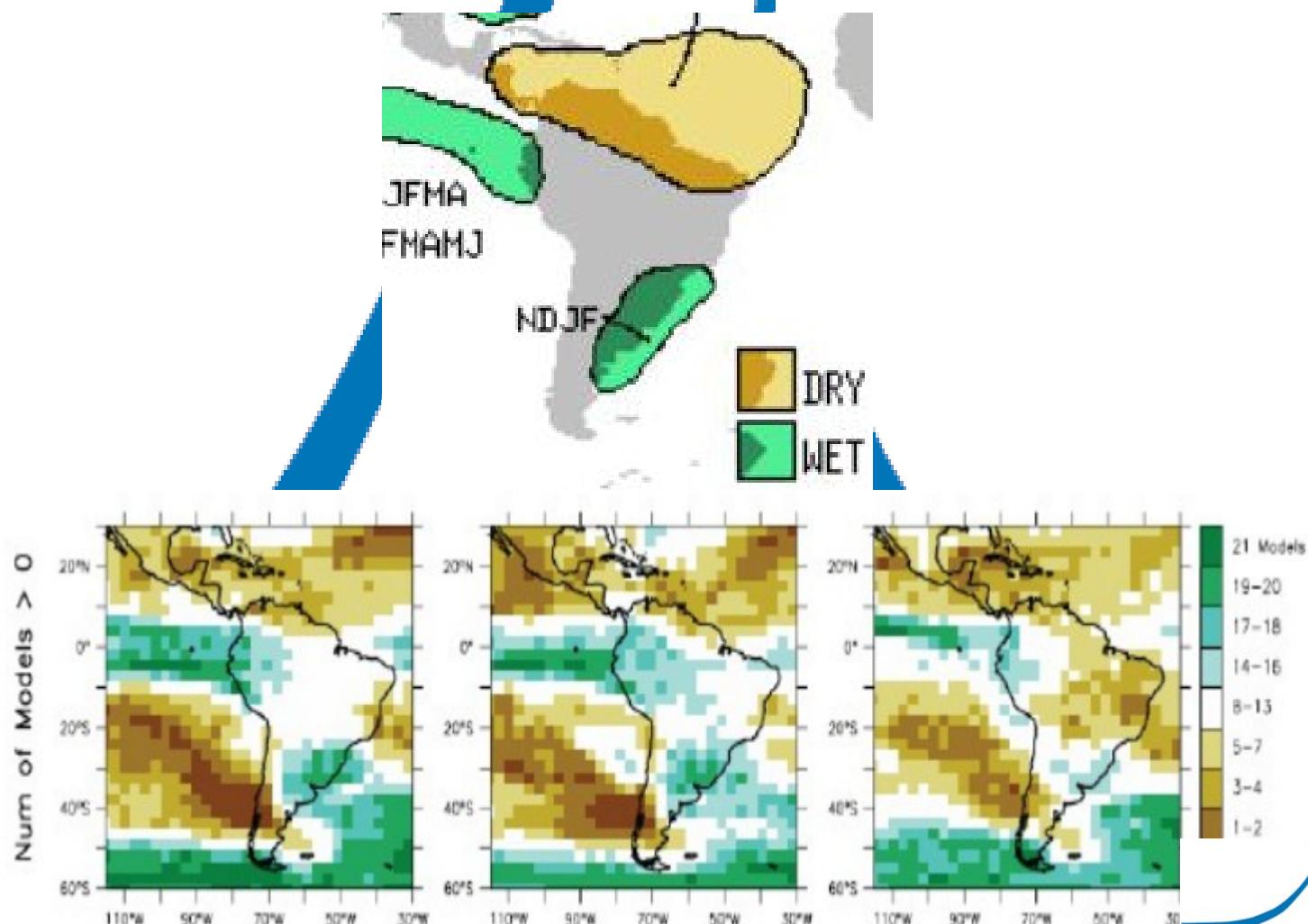


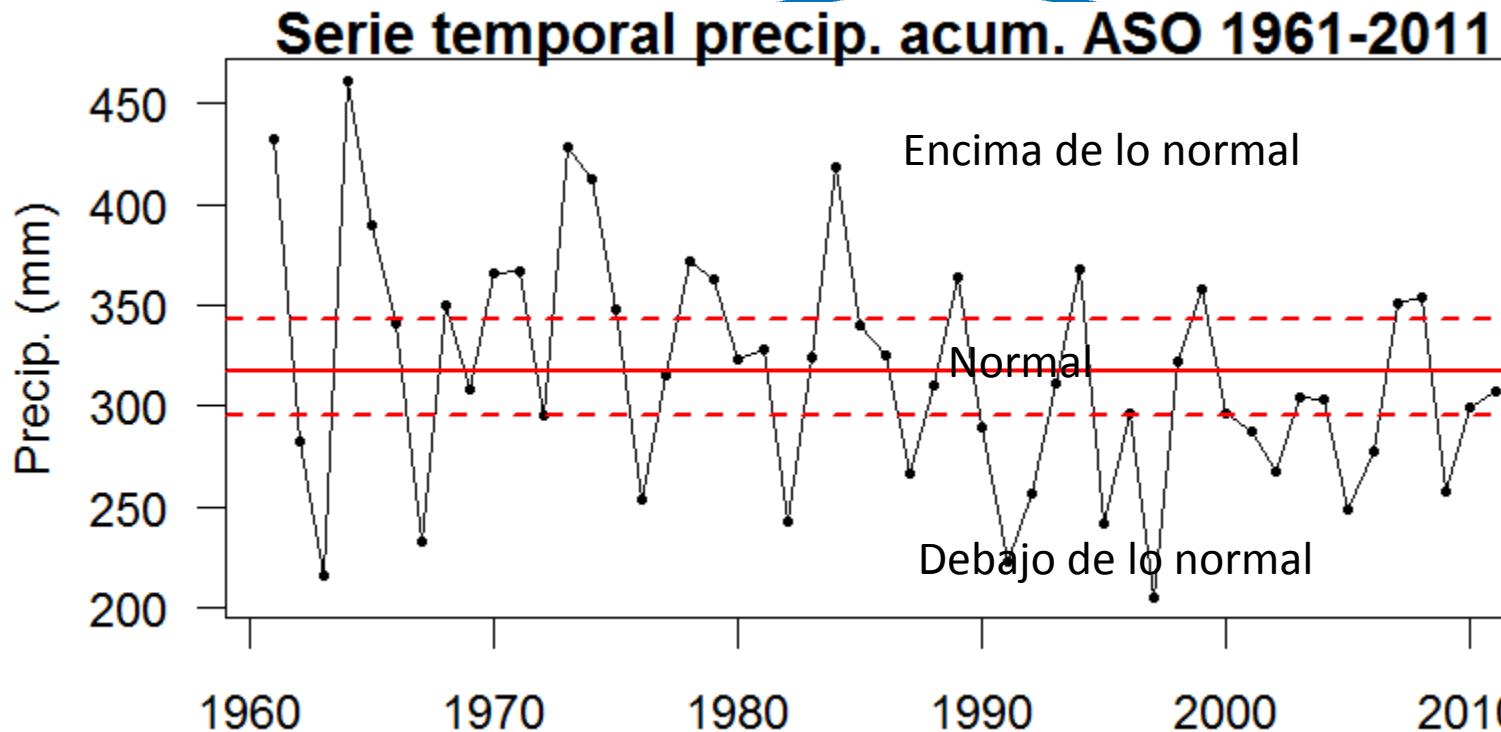
2. OBJETIVO DEL PRESENTE TRABAJO

- Series de caudales medios mensuales son:
 - No estacionales
 - Cuasi-Periódicas
 - Son el resultado de una interacción compleja entre muchas variables.
- Requerimientos en la formulación de modelos
 - Modelación no paramétrica y no lineal.
 - Involucrar los principales ciclos asociados a la variabilidad climática.
 - Manejo simultáneamente de variables explicativas.
 - Capacidad de ponderar la información disponible en el tiempo.

Desarrollar un modelo de pronóstico *multivariado, no paramétrico, y no lineal,* que tenga la capacidad de *ponderar la información disponible, involucrando las frecuencias mas relevantes de la serie de tiempo.*

EL ENSO Y LOS MODELOS CLIMÁTICOS





$n = 51$ años de datos observados

Media:

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i = 317.8 \text{ mm}$$

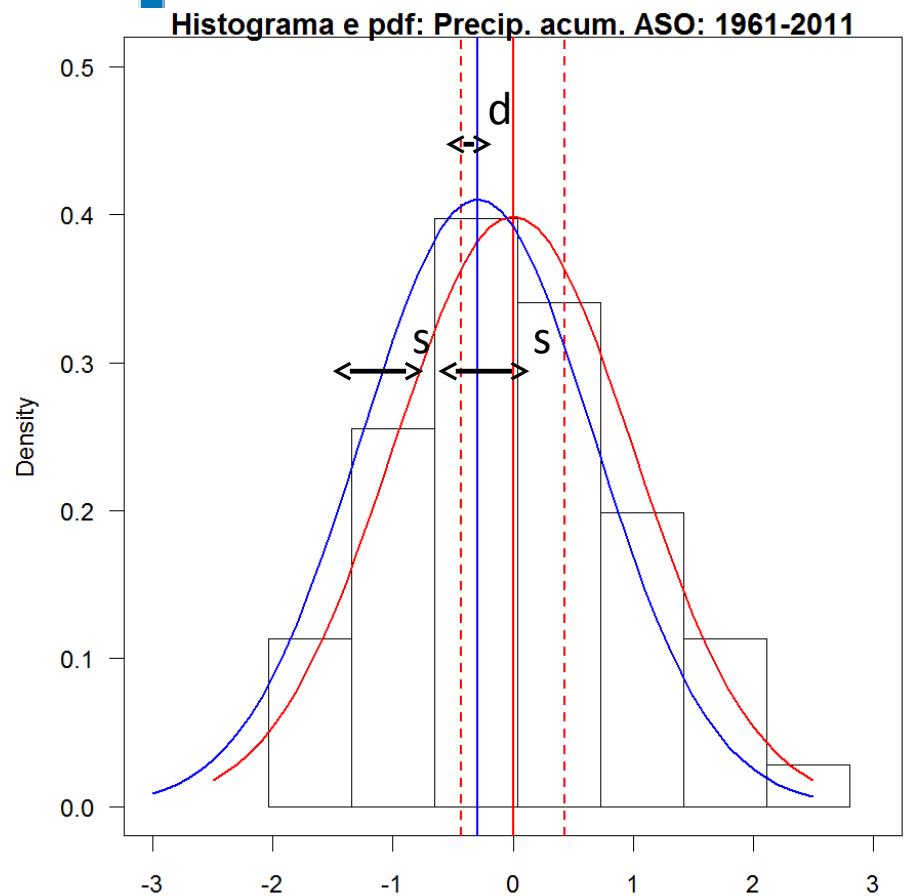
Desviación:

$$s_y = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2} = 57.9 \text{ mm}$$

Parámetros básicos de los pronósticos probabilísticos

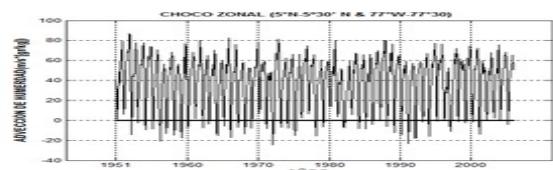
d: Anomalía, desplazamiento: esperanza condicional de un evento. Resultado numérico de los pronósticos.

s: desviación, amplitud de la fdp: asociada al desempeño del modelo, estimada a partir de los errores del modelo.

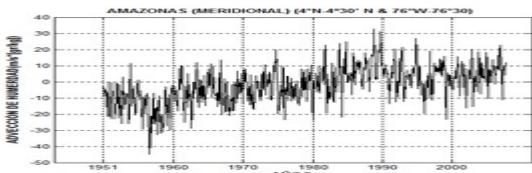


THE LONGER LEAD HYDROLOGY

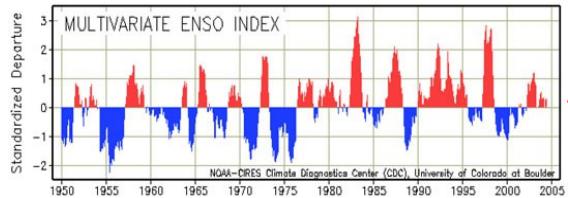
ZCIT



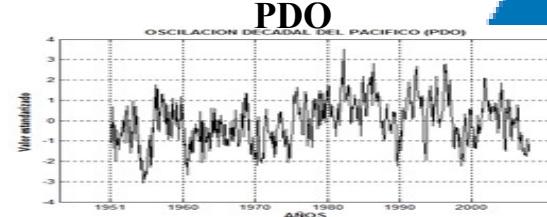
MDJO



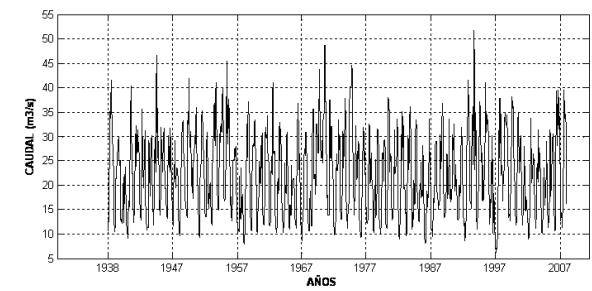
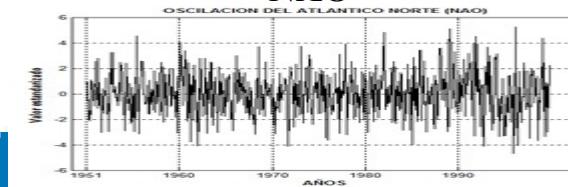
MEI



$$f(\dots) + \varepsilon =$$



NAO



$$\text{Caudales}_t = f(\text{Caudales}_{t-j}, \text{ENSO}_{t-k}, \text{CHOCO}_{t-l}, \dots) + \varepsilon$$



3.1 ANÁLISIS PREDICTIVO

- Por lo general, el análisis de regresión estima el valor esperado (esperanza condicional) de una variable dependiente dadas las variables independientes (X_1, X_2, \dots, X_n)
$$y = f(X_1, X_2, \dots, X_n) + \varepsilon$$
$$Caudales_t = f(Caudales_{t-j}, ENSO_{t-k}, PDO_{t-l}, \dots) + \varepsilon$$
- Dado que no se conoce a ciencia cierta el proceso de formación de la variable dinámica que se desea modelar, el análisis de regresión se fundamenta en muchas suposiciones que se hacen del proceso, las cuales pueden ser corroboradas si se tiene una cantidad suficiente de datos.

CLIMATE PREDICTABILITY TOOL

$$\hat{y} = Ax + b$$

Copyright 2003

IRIINTERNATIONAL RESEARCH INSTITUTE
FOR CLIMATE AND SOCIETY

REGRESSION ANALYSIS METHODS

- ARIMA
- Autoregressive integrated moving average Models.

$$\varphi(B)\nabla^d z_t = \theta(B)a_t$$

- RLM

Multiple Linear Regression

$$\hat{y} = C_1 X_1 + C_2 X_2 + \dots + C_n X_n$$

- POLPON

Radial Basis Functions

$$\hat{y} = \mu + \sum_{i=1}^N \lambda_i \phi(\|x - x_i\|)$$

- RNA

Artificial Neural Network

$$\hat{y} = \sum_{j=1}^H w_{i,j} \cdot g\left(\sum_{i=0}^{m_l-1} w_{i,j} \cdot X_i\right)$$

- MARS

Multivariate Adaptive Regression Splines

$$\hat{y} = a_0 + \sum_{k_m=1} f_i(X_i) + \sum_{k_m=2} f_{ij}(X_i, X_j) + \sum_{k_m=3} f_{ijk}(X_i, X_j, X_k) + \dots$$

- RNPOL

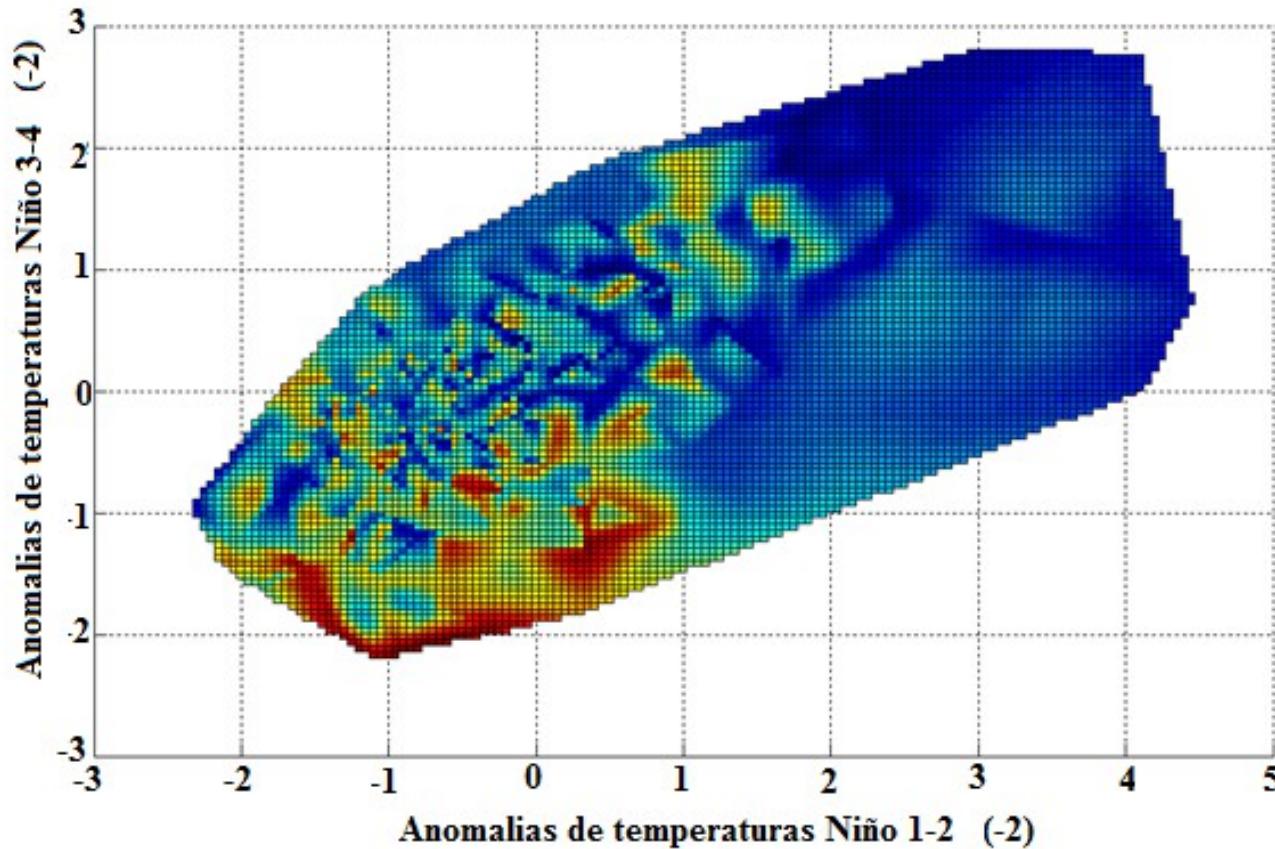
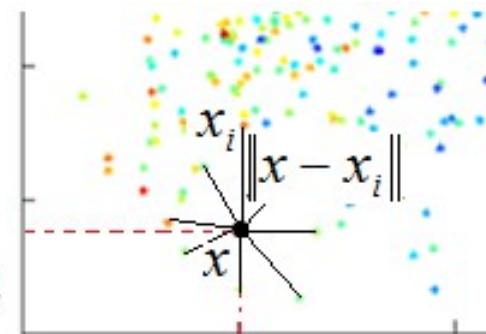
Polynomial Neural Networks

$$\hat{y} = a + \sum_{i=1}^m b_i X_i + \sum_{i=1}^m \sum_{j=1}^m c_{ij} X_i X_j + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^m c_{ijk} X_i X_j X_k + \dots$$

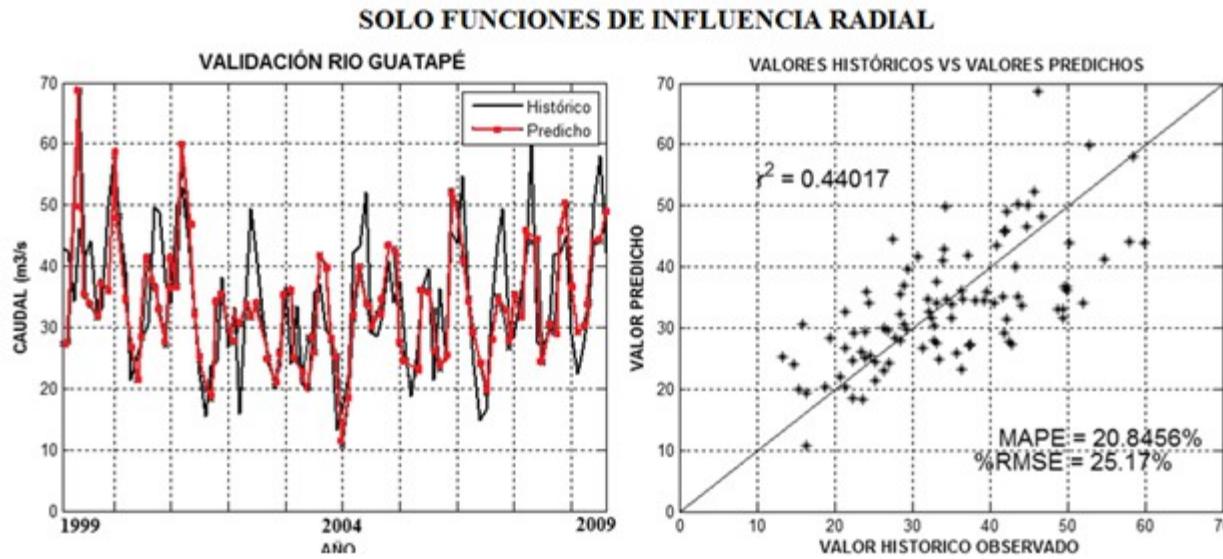
5. FUNCIONES DE INFLUENCIA RADIAL (RBF)

6. APLICACIÓN DE LAS RBF AL PRONÓSTICO DE CAUDALES

CAUDALES ESTANDARIZADOS RECONSTRUIDOS DEL RÍO GUATAPÉ



7. ESTRATEGIA DE VALIDACIÓN



Validación Retroactiva

AÑO	1951-1990	1991	1992	1993	1994	...	2000
1991	Calibración 1951-1990	Predicción 1991	Información omitida				
1992	Calibración 1951-1990	Información omitida	Predicción 1992	Información omitida			
1993	Calibración 1951-1990	Información omitida		Predicción 1993	Información omitida		
1994	Calibración 1951-1990	Información omitida			Predicción 1994	Información omitida	
...	Calibración 1951-1990	Información omitida			Predicción	Información omitida	
2000	Calibración 1951-1990	Información omitida				Predicción 2000	

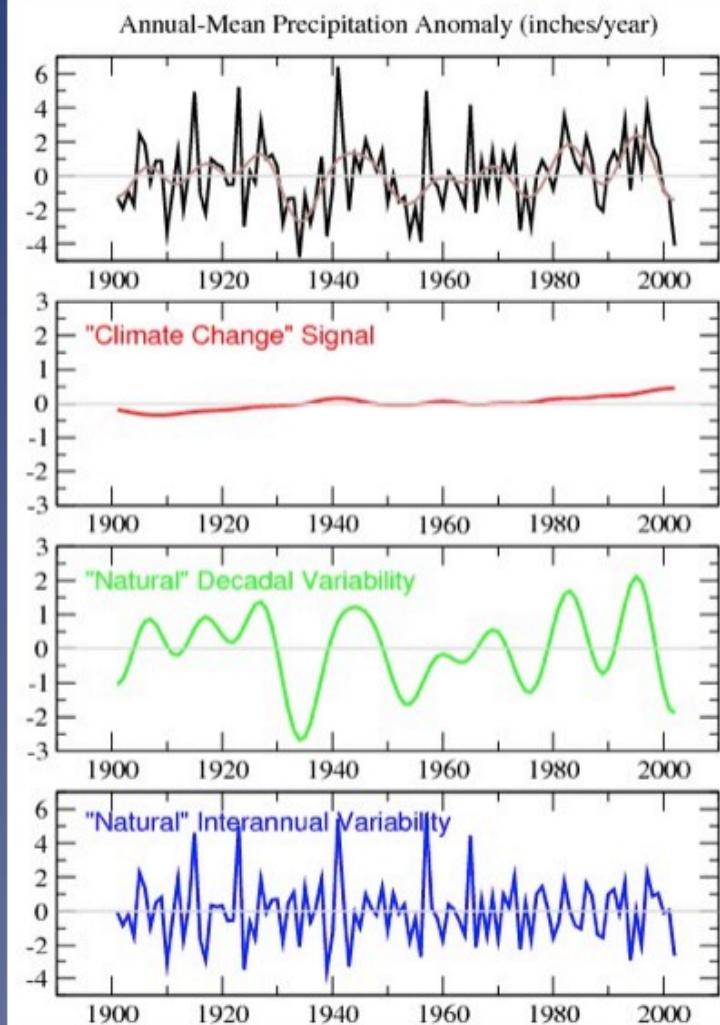
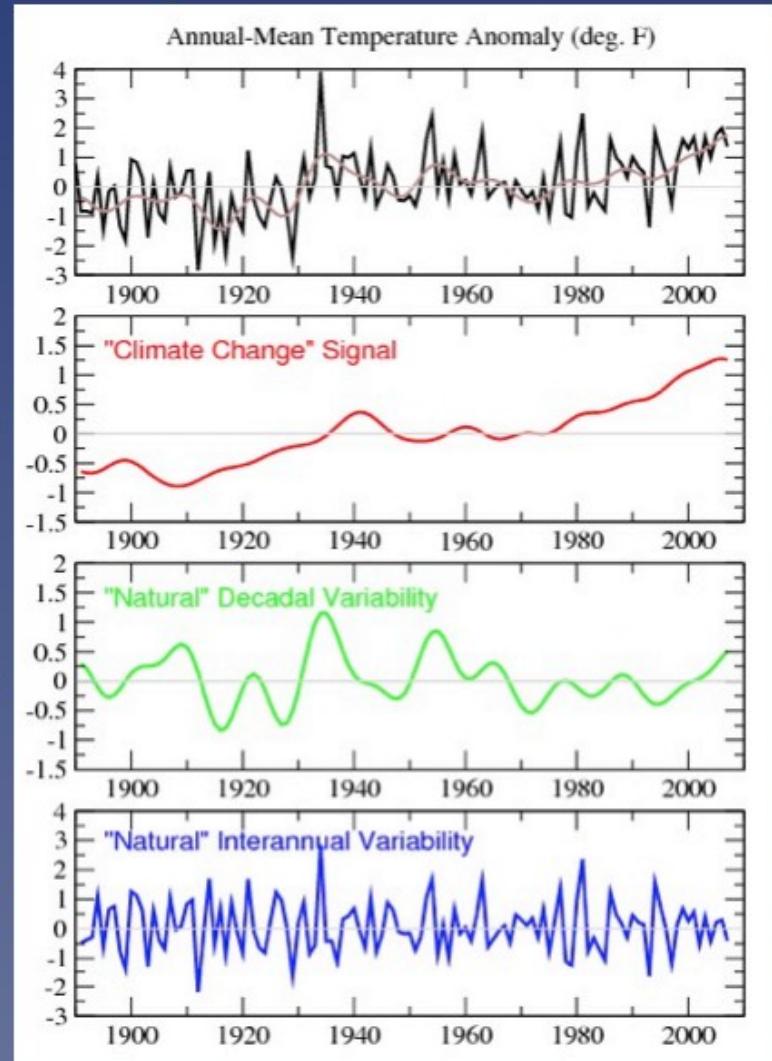
$$e_t = y_t - \hat{y}_t$$

$$SSE = \sum_{t=1}^T e_t^2$$

$$MSE = \frac{SSE}{T}$$

$$RMSE = \sqrt{MSE}$$

$$MAPE = \frac{1}{T} \sum_{i=1}^T |p_i|$$



RI

8. TÉCNICAS DE DESCOMPOSICIÓN ESPECTRAL

- AES

Singular Spectrum Analysis

$$X_{i,j} = \sum_{k=1}^{cp} a^k i E_j^k \quad 1 \leq j \leq M$$

- OND

Discrete Wavelet Transform

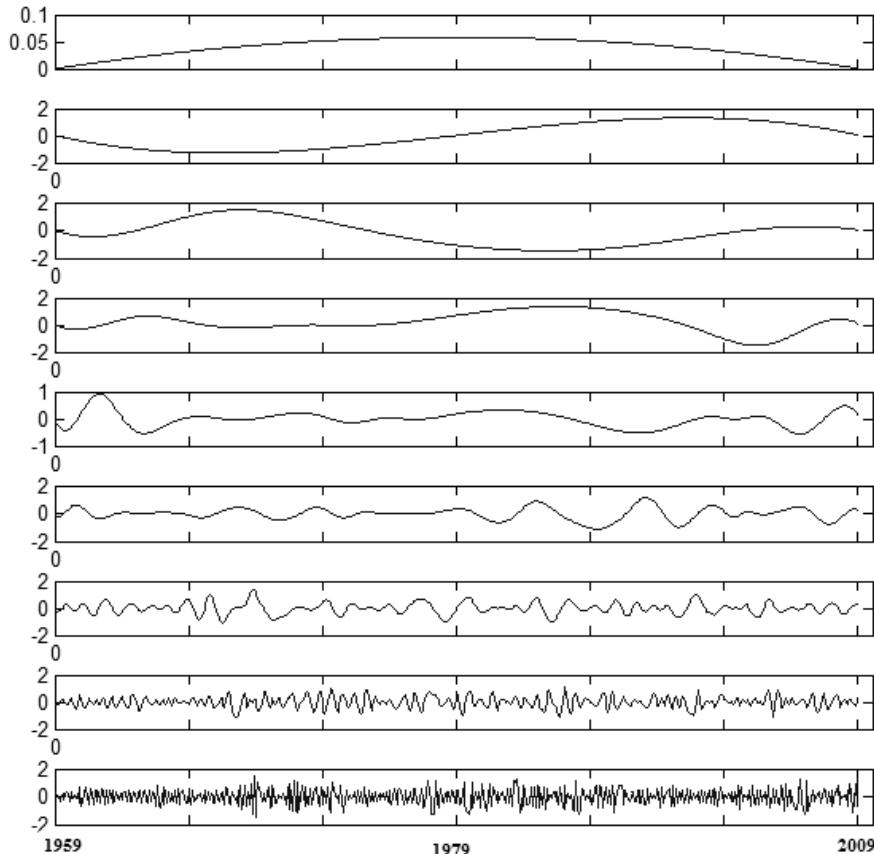
$$X(t) = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W(\tau, s) \frac{1}{s^2} \psi\left(\frac{\tau - t}{s}\right) d\tau ds$$

- HH

Hilbert-Huang Transform

$$x(t) = \sum_{j=1}^n c_j(t) + r_n(t)$$

DESCOMPOSICIÓN EN MODOS EMPIRICOS PARA LA SERIE ESTANDARIZADA DEL RÍO GUATAPÉ



CPT PROGRAM

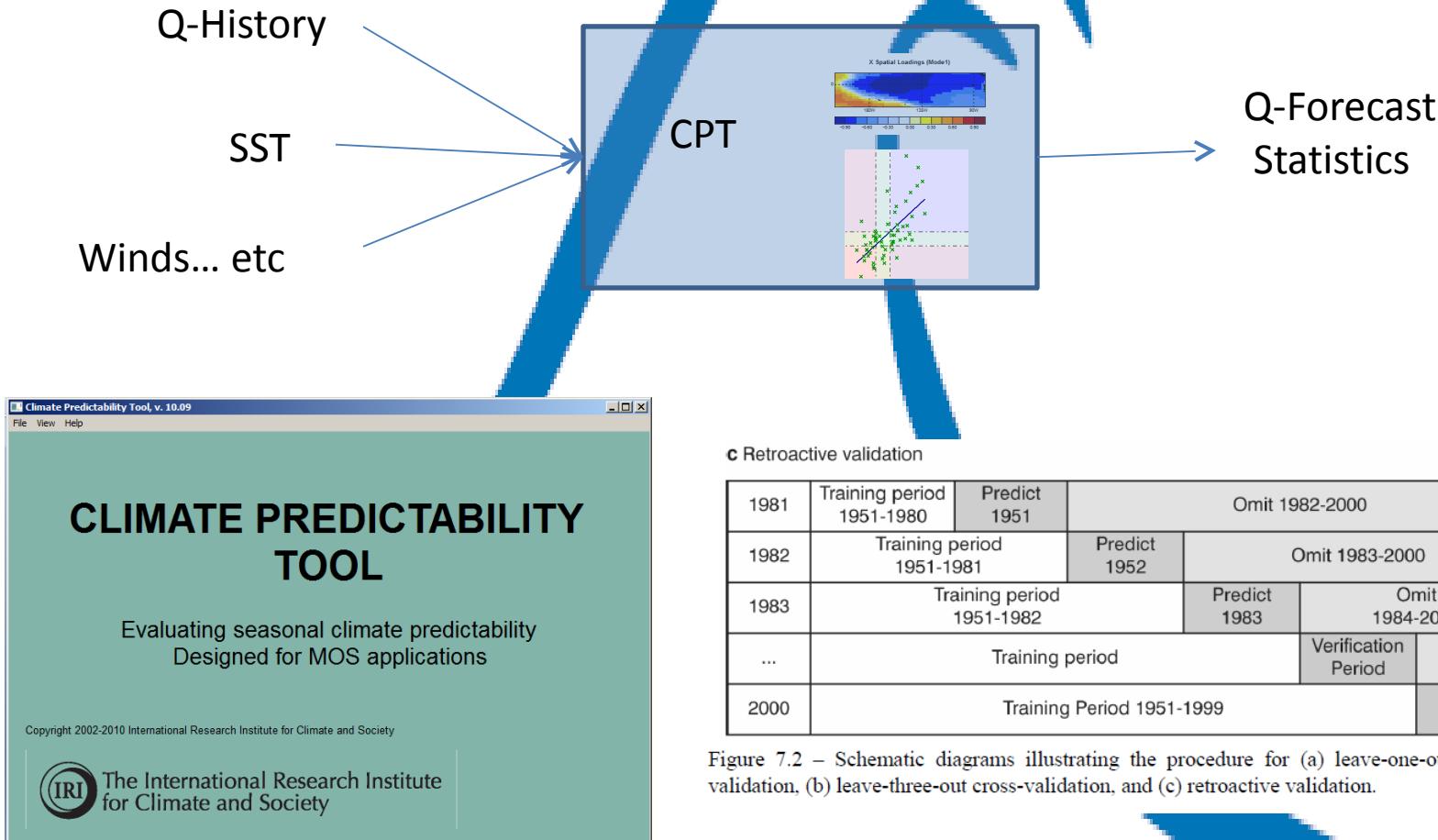
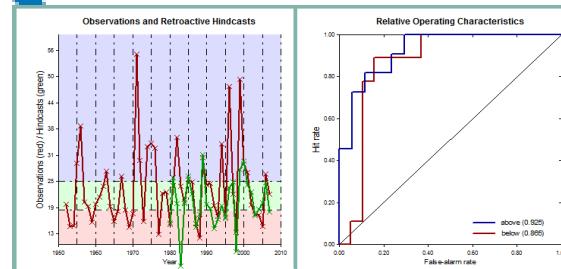
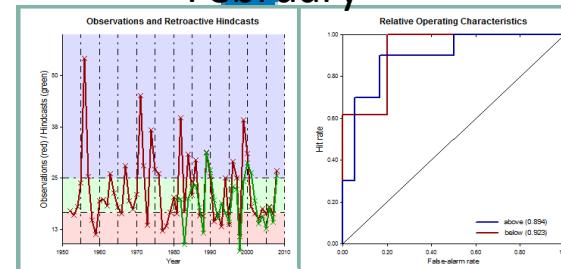
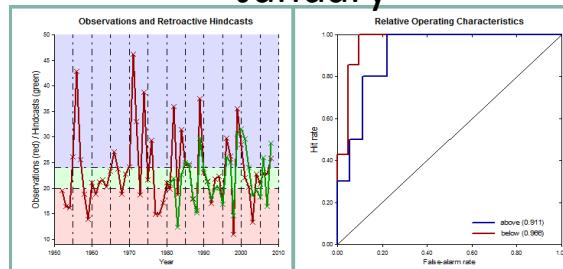


Figure 7.2 – Schematic diagrams illustrating the procedure for (a) leave-one-out cross-validation, (b) leave-three-out cross-validation, and (c) retroactive validation.

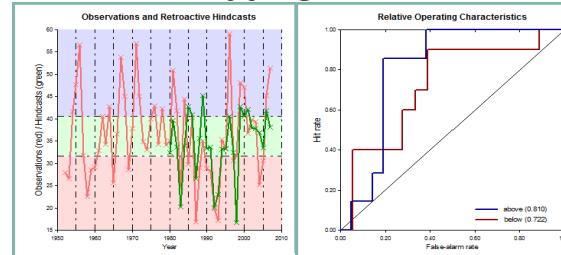
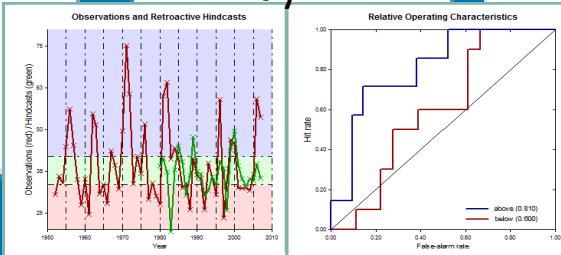
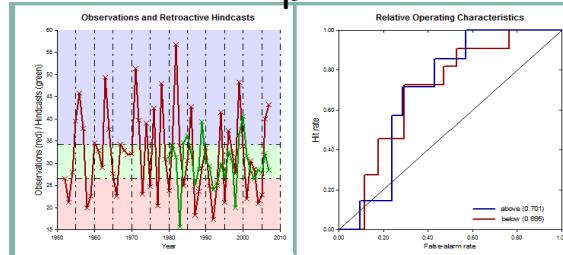
Mason & Baddour (2007)

RESULTS

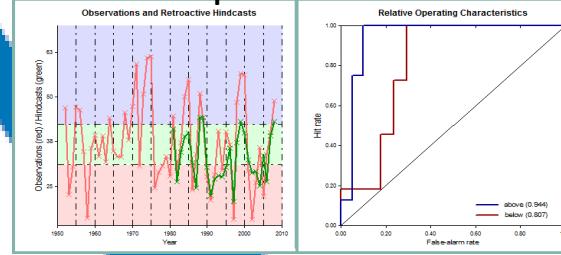
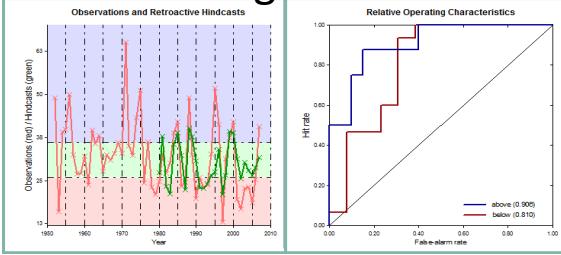
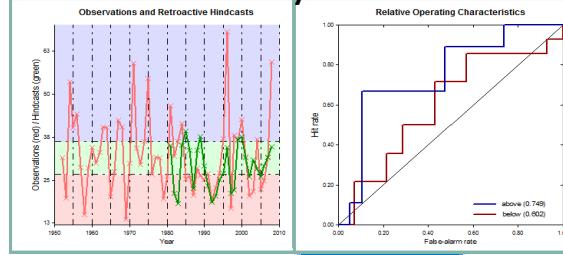
February



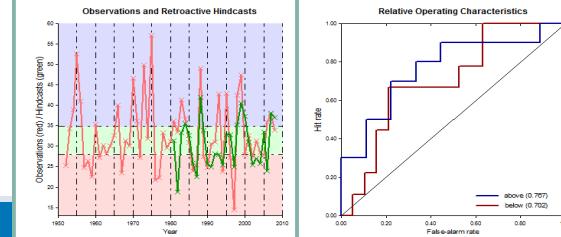
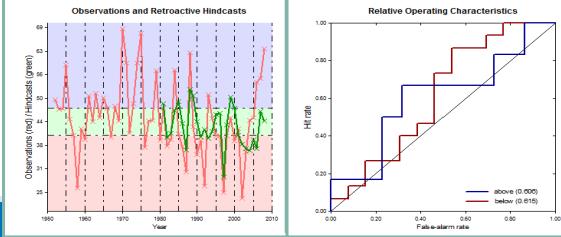
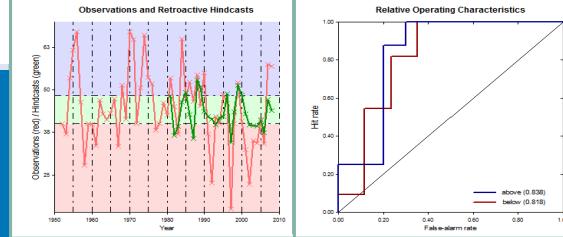
April



July

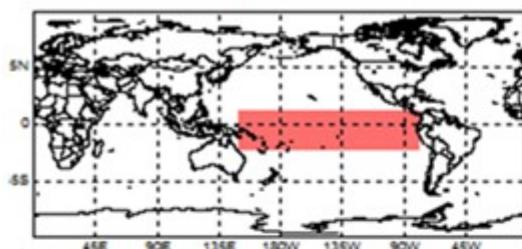


October

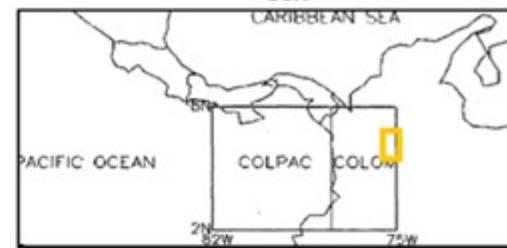


CPT - Validation results

SST REGION

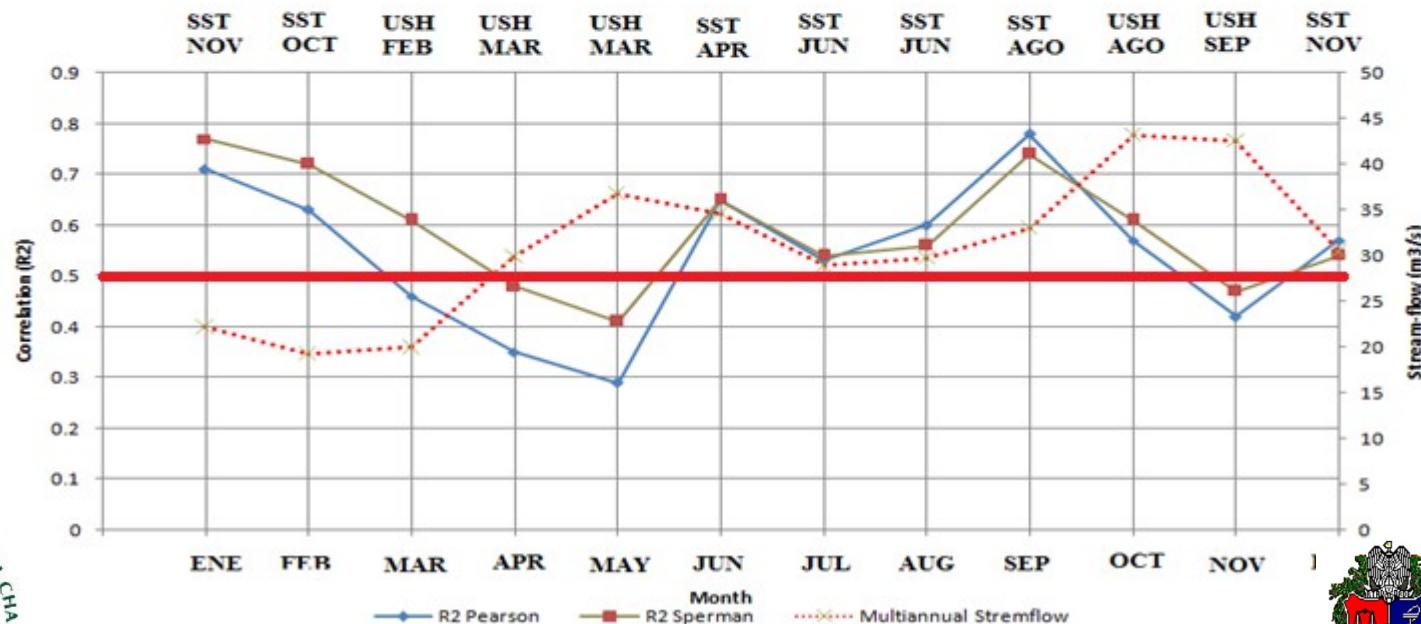


Covar. u and specific hum REGION
USH



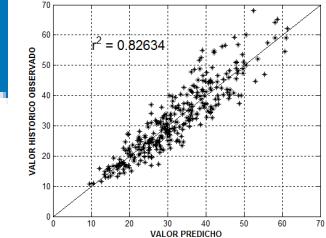
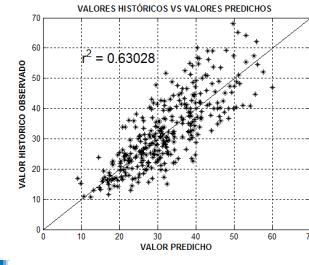
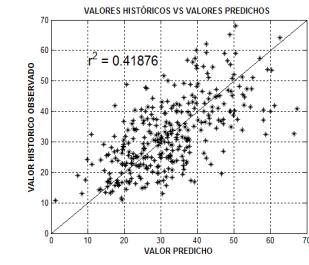
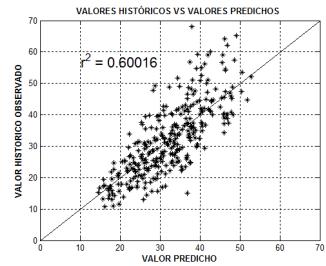
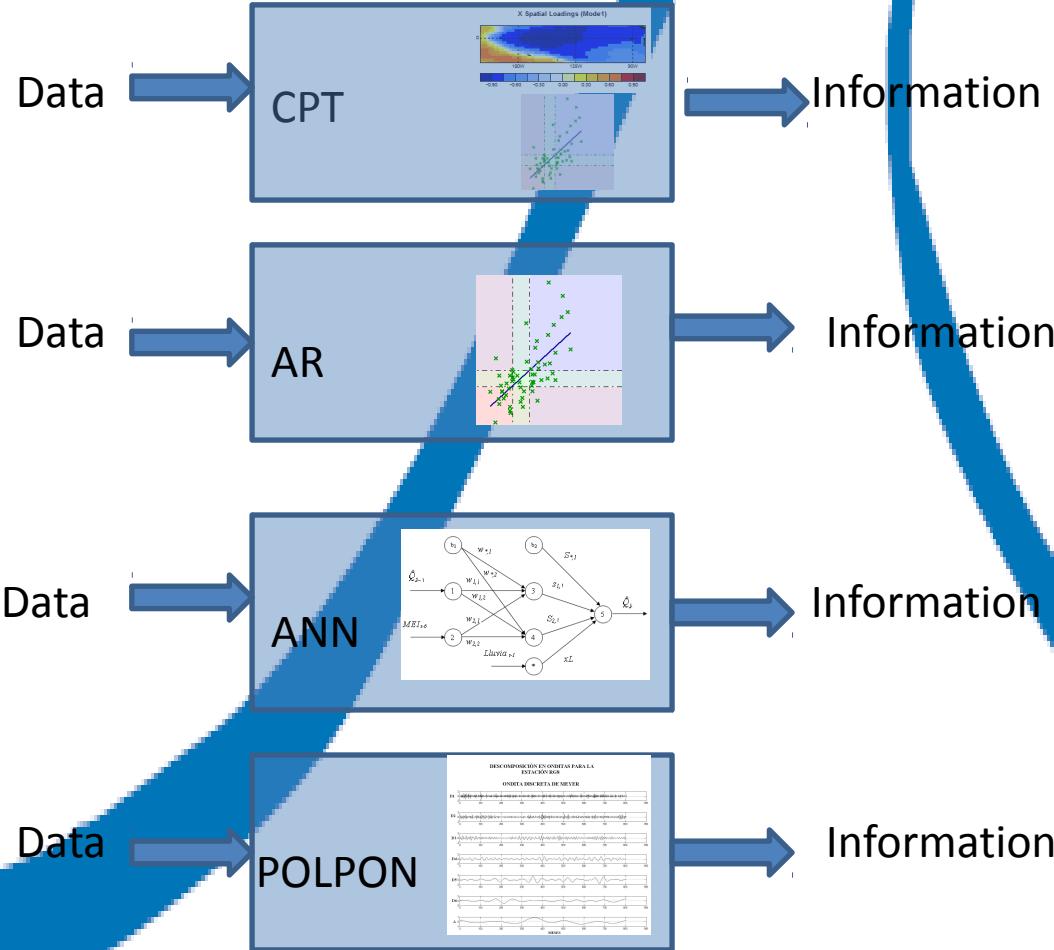
Correlation Coeficient for each month

VALIDATION 1980-2007

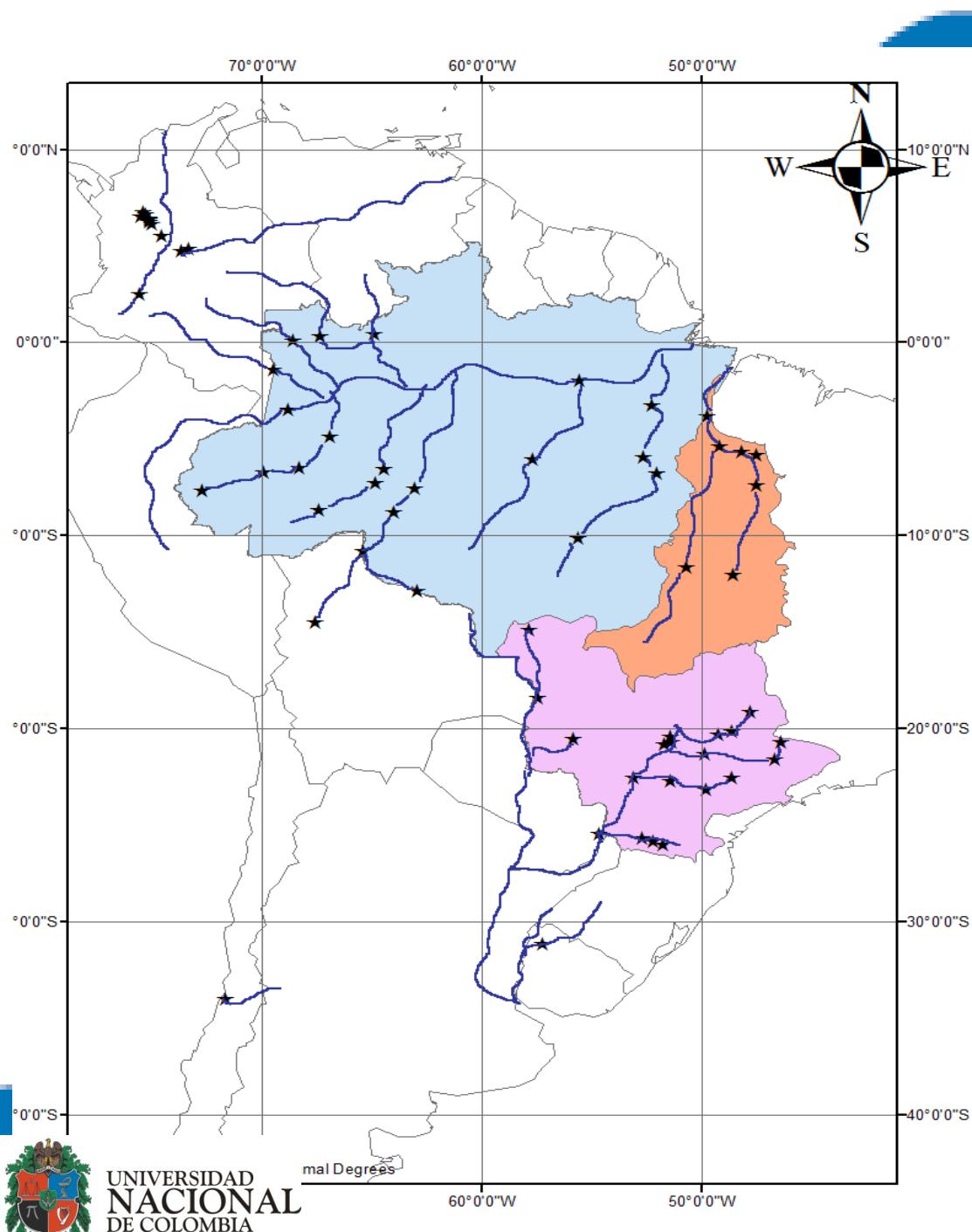


CPT vs OTERS MODELS

Riogrande-River



CPT: Climatic predictability Tool; **AR:** Autoregressive Model; **ANN:** Artificial neural Network ; **AESMP:** Periodic Model based on Singular spectral Analysis.



Informação hidrológica:

69 estações de vazão na Colômbia, Brasil, Peru, Bolívia, Paraguai e Uruguai.

Serries de vazões mensais com dados **desde janeiro de 1980 até dezembro de 2005** (312 dados).

Agência Nacional de Águas (ANA)

Centro de Ciência do Sistema Terrestre CCST/INPE.

Centro de Previsão de Tempo e Estudos Climáticos - CPTEC/INPE





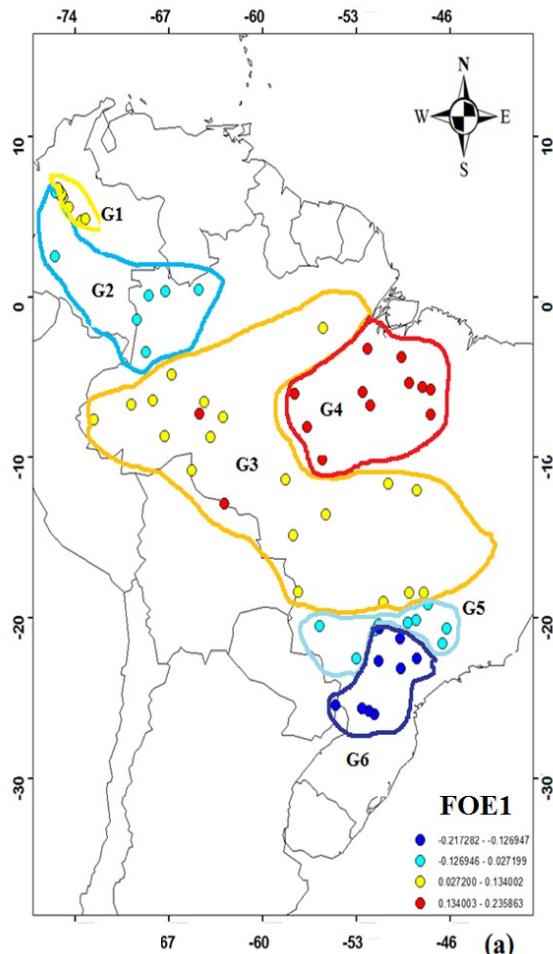
UNIVERSIDAD
NACIONAL
DE COLOMBIA
SEDE MEDELLÍN



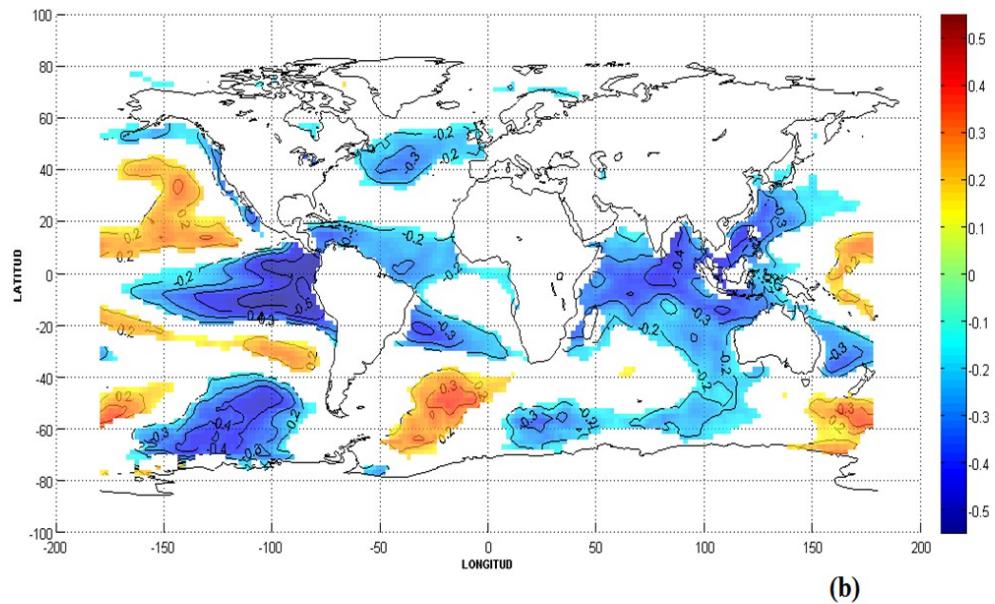
POSGRADO EN APROVECHAMIENTO
DE RECURSOS HIDRÁULICOS

Resultados (I): O primeiro modo de variabilidade dos rios na America do Sul

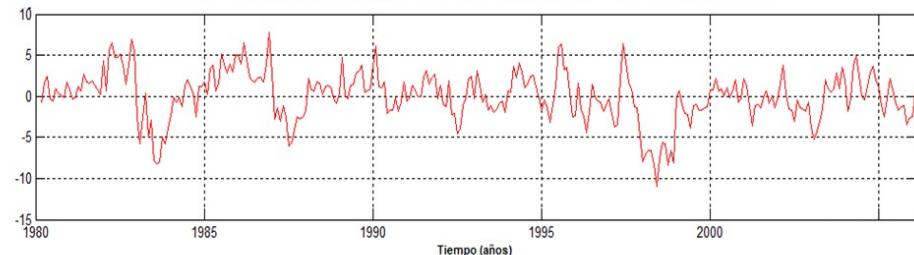
Função ortogonal empírica



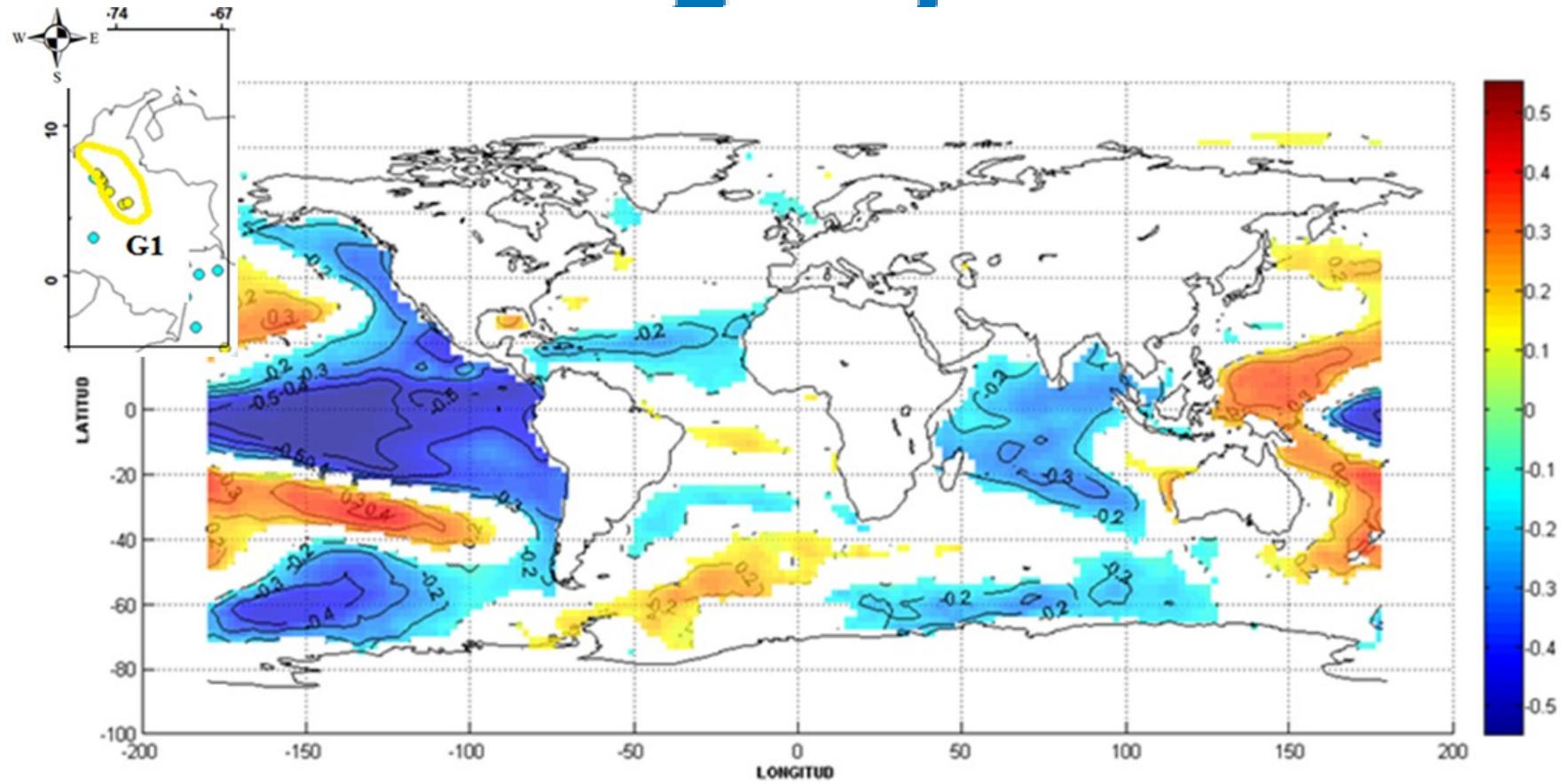
Mapa dos coeficientes de correlação entre sua serie temporal e a TSM



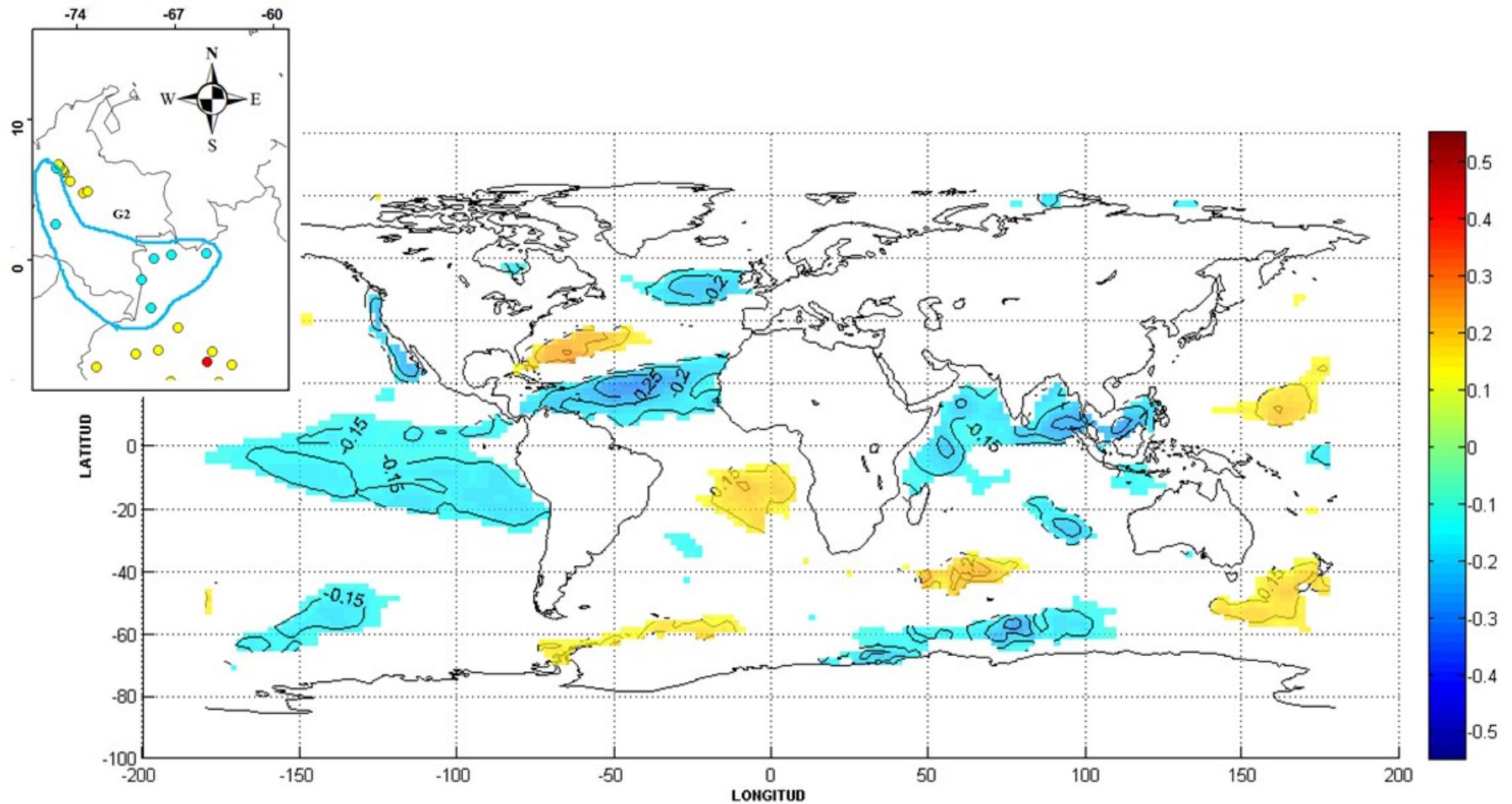
Serie temporal do primeiro modo de variabilidad.



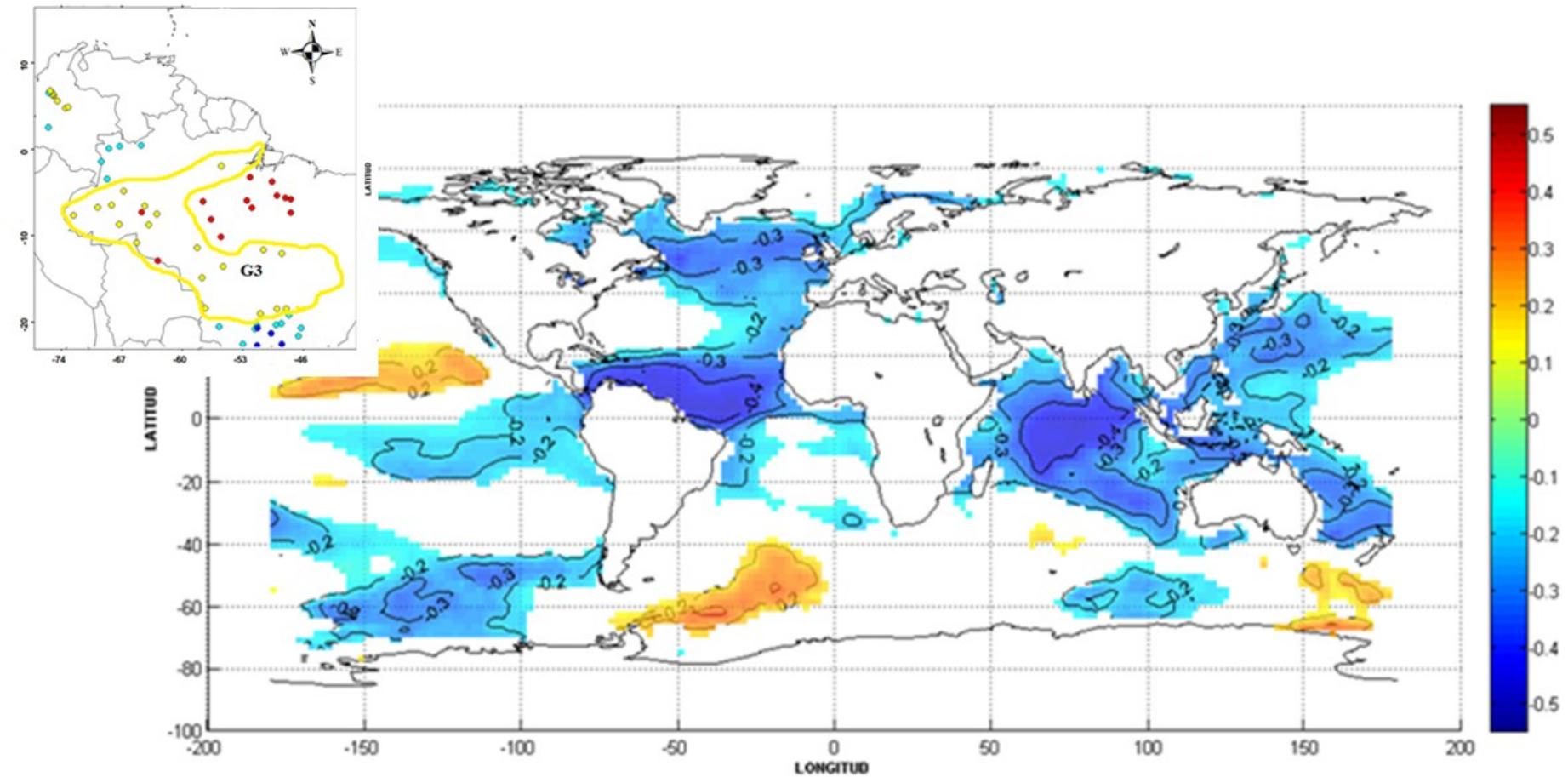
Resultados (II): Comportamento dos rios da região centro-oeste da Colômbia



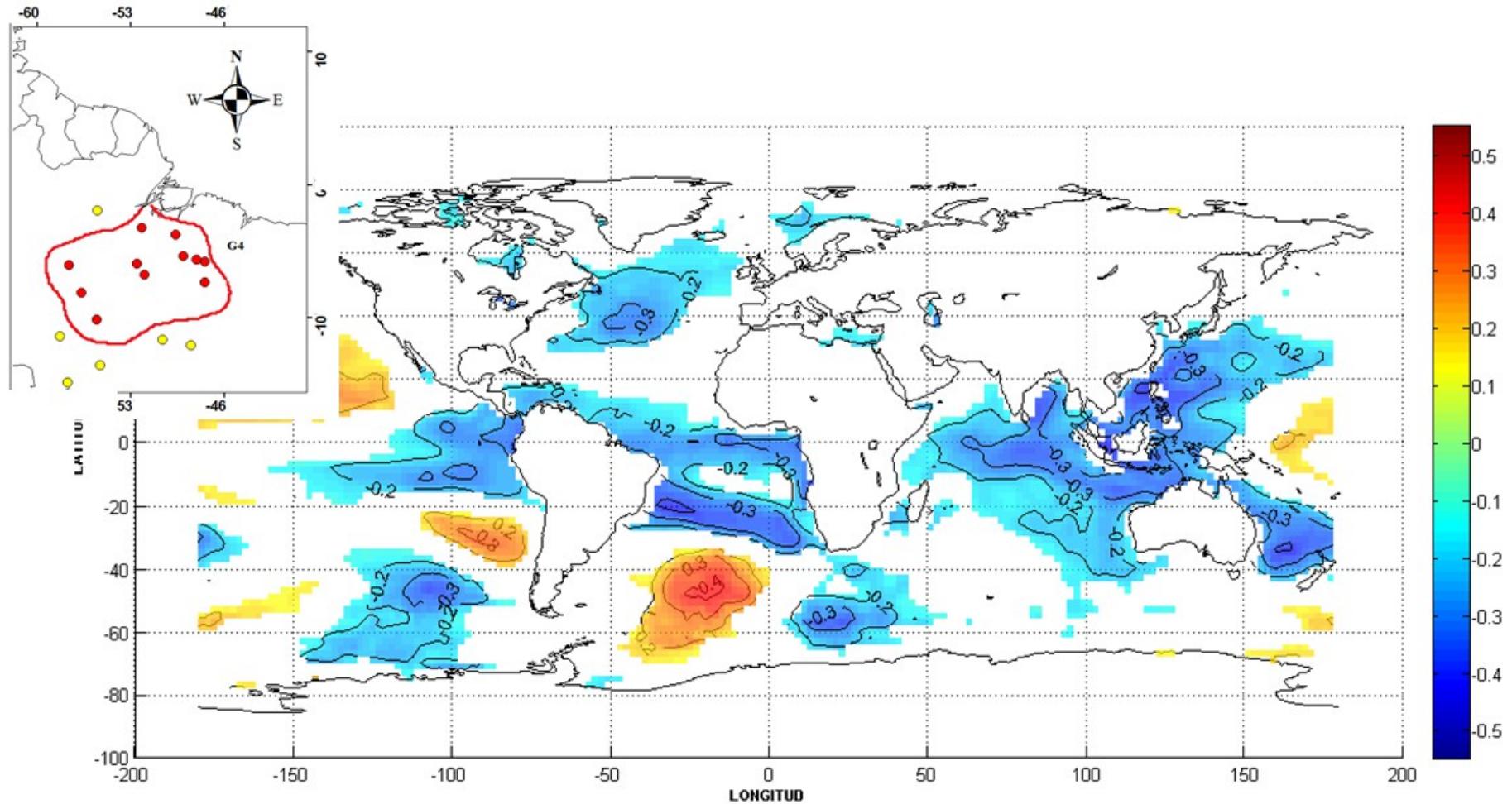
4. Resultados (III): Comportamento dos rios do centro da Colômbia e do norte da Amazônia



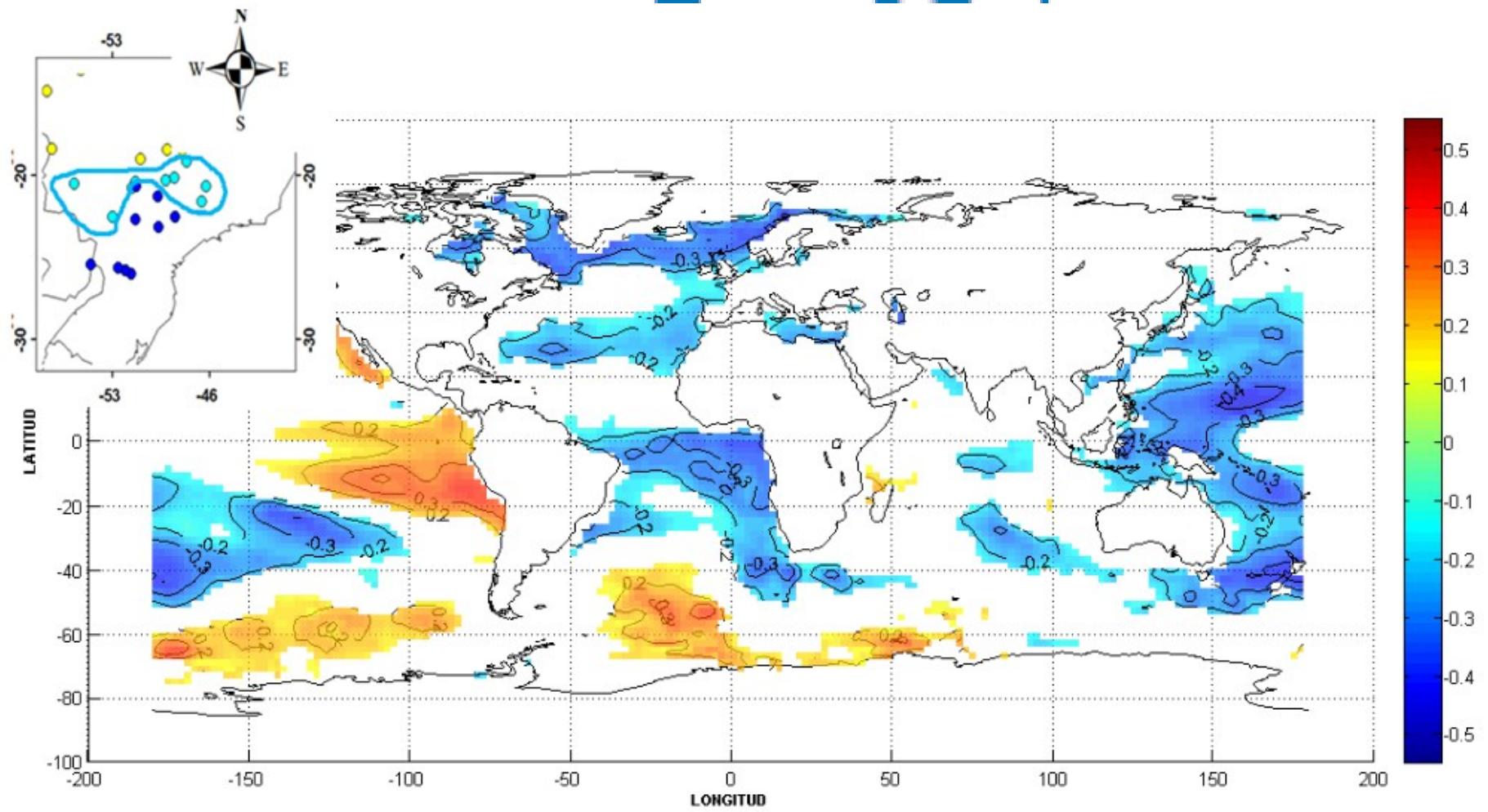
4. Resultados (IV): Rios da região de maior desenvolvimento do jato de baixos níveis ao longo dos Andes



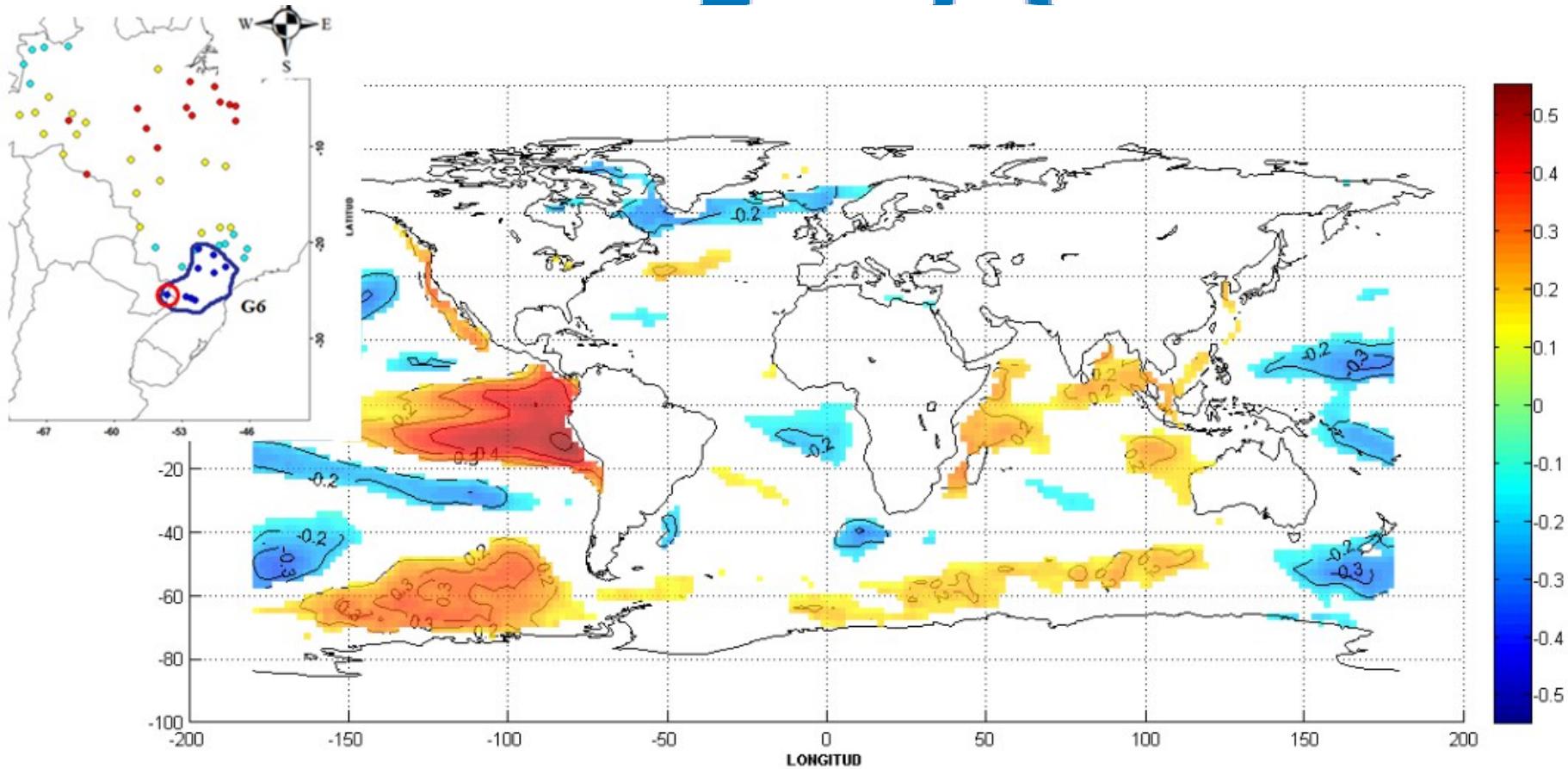
4. Resultados (V): Rios da região centro leste e nordeste do Brasil

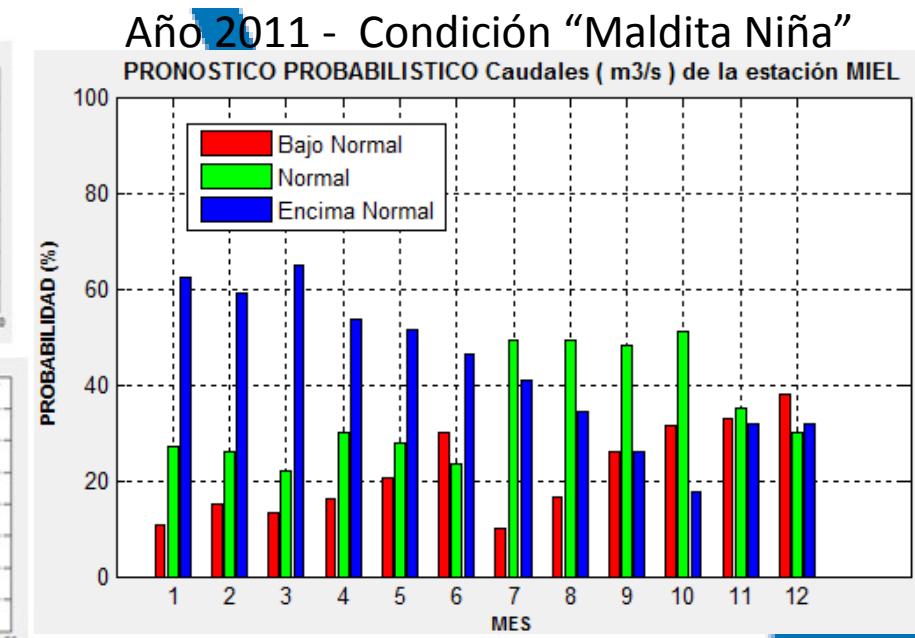
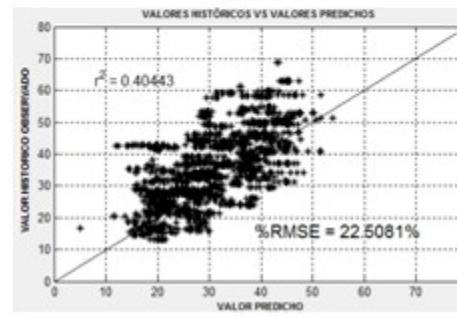
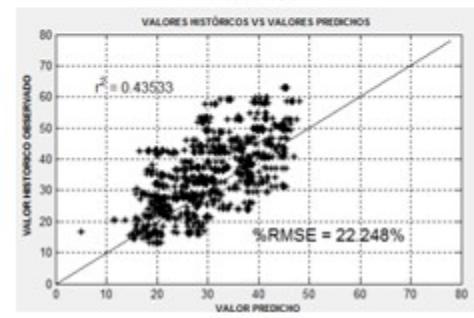
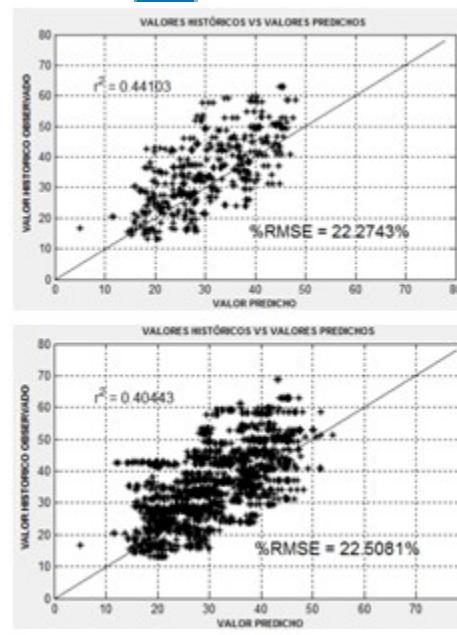
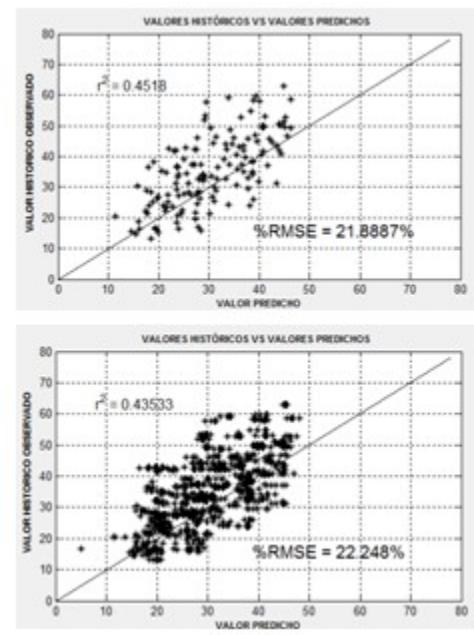
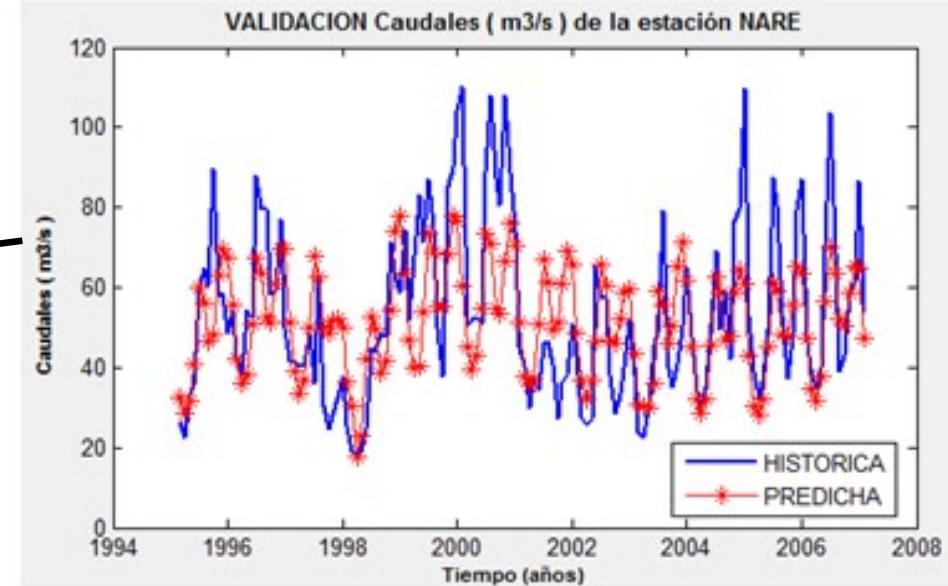
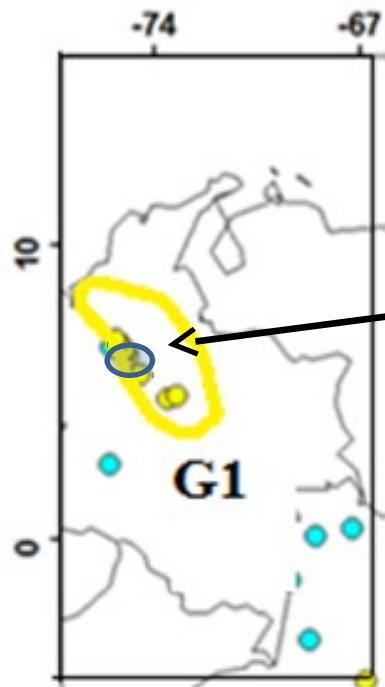


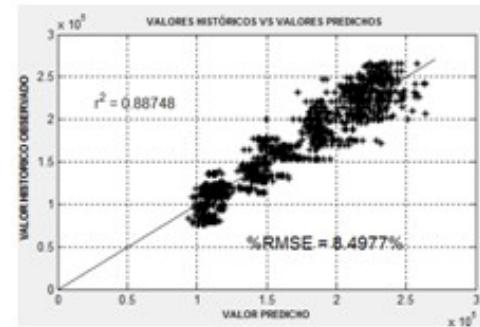
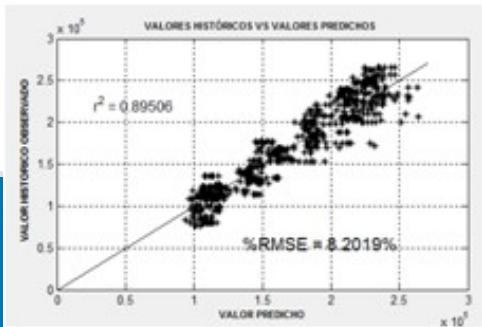
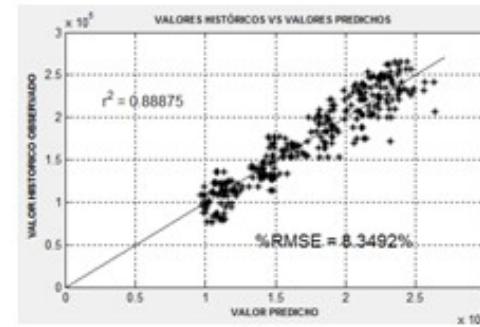
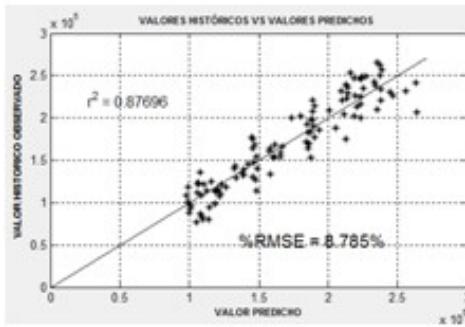
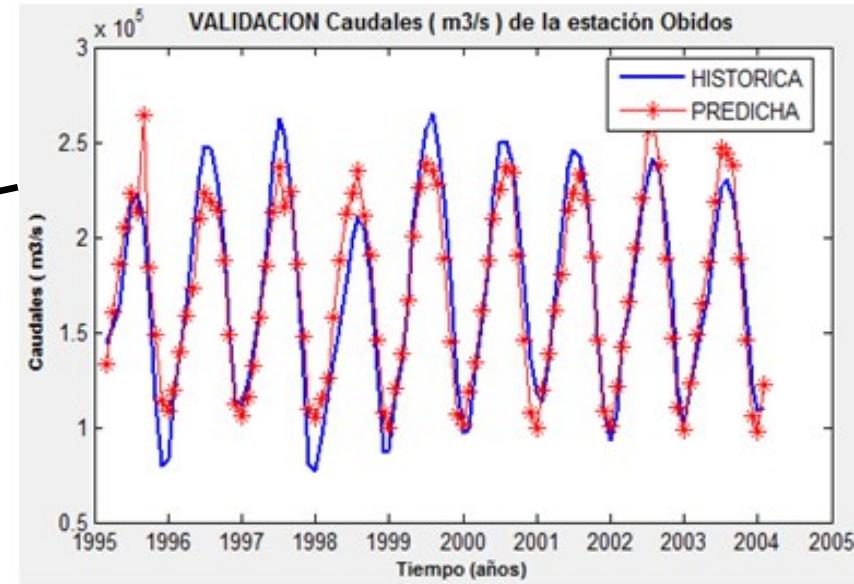
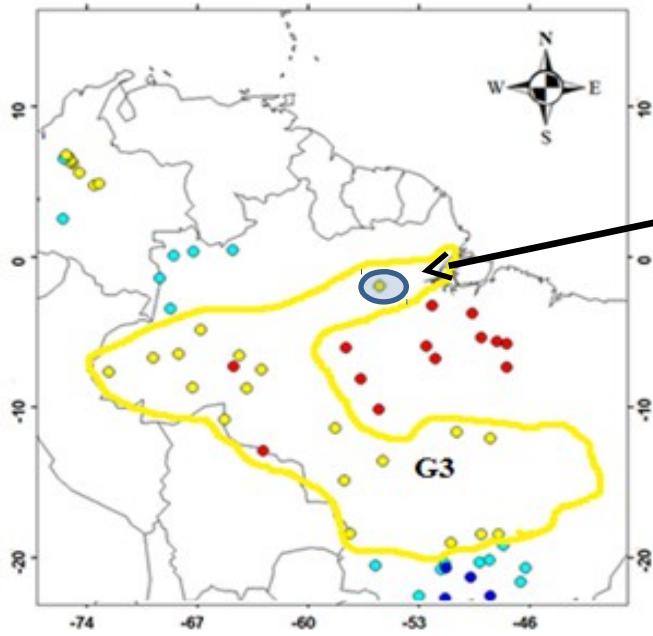
Resultados (VI): Ríos do sudeste do Brasil



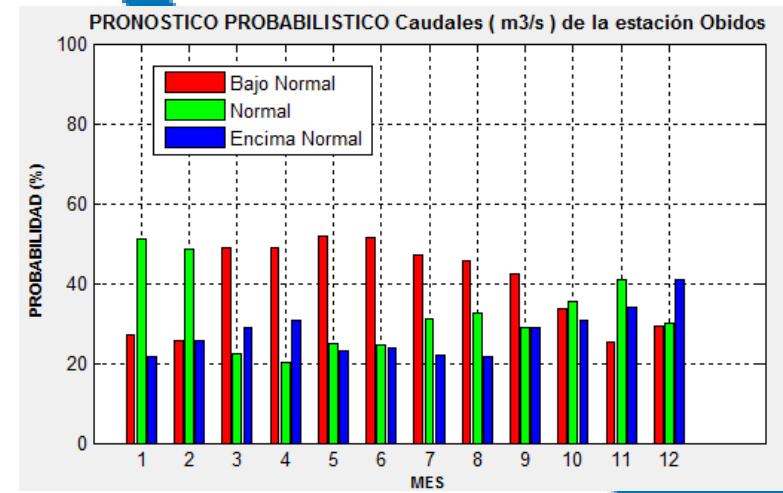
4. Resultados (VII): Ríos do Sul do Brasil

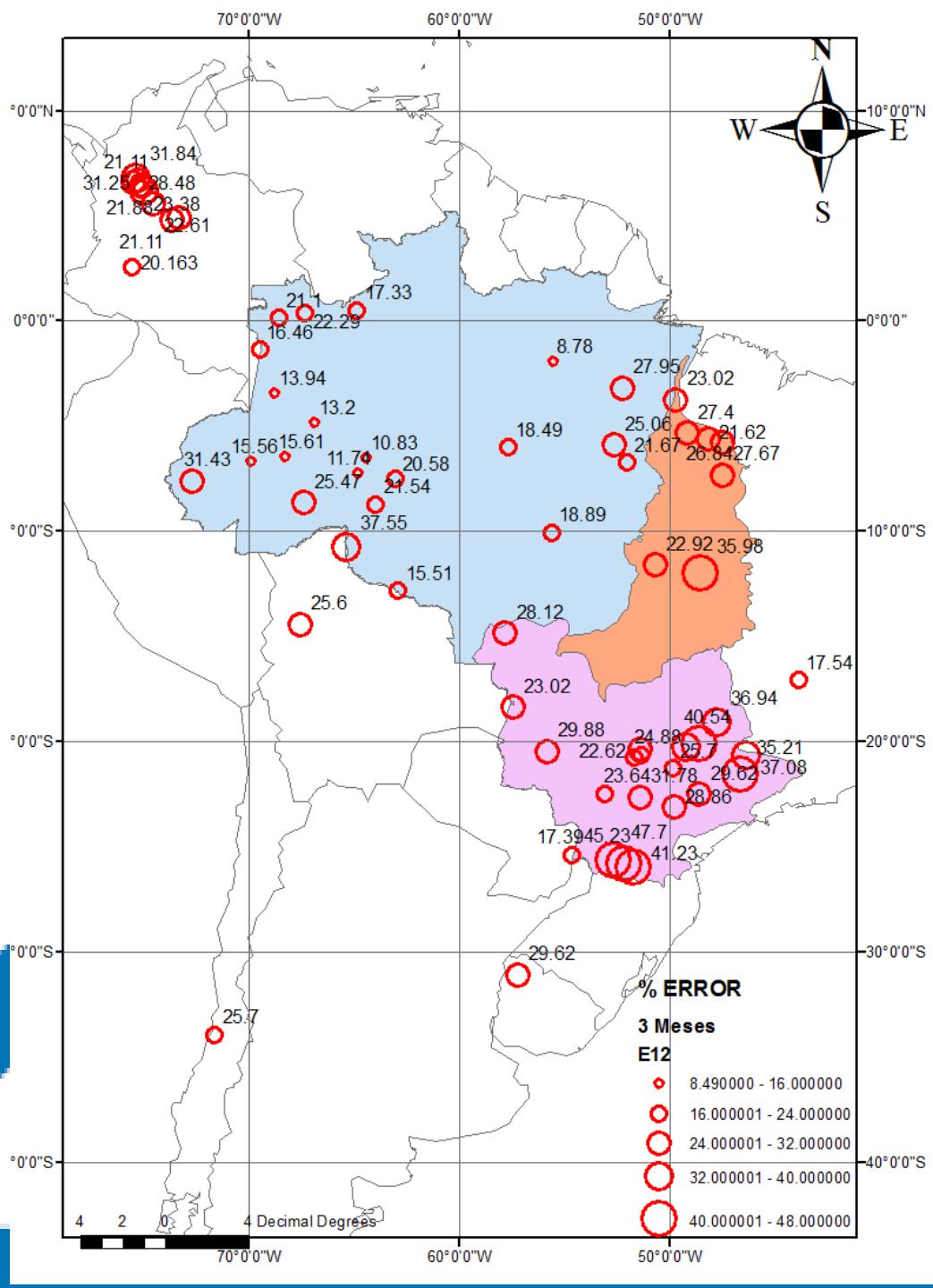




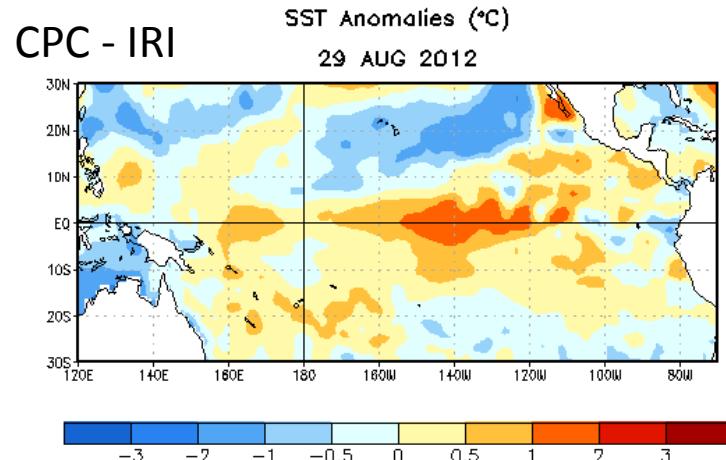


Sequía en el Amazonas año 2005

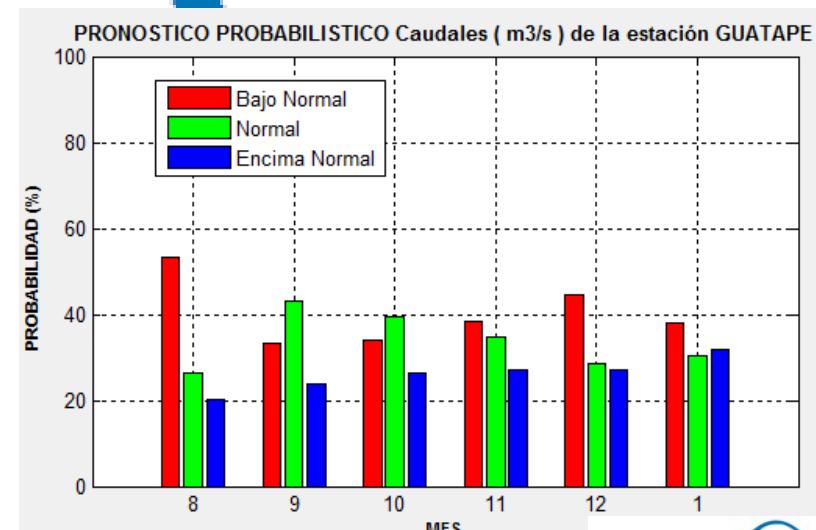
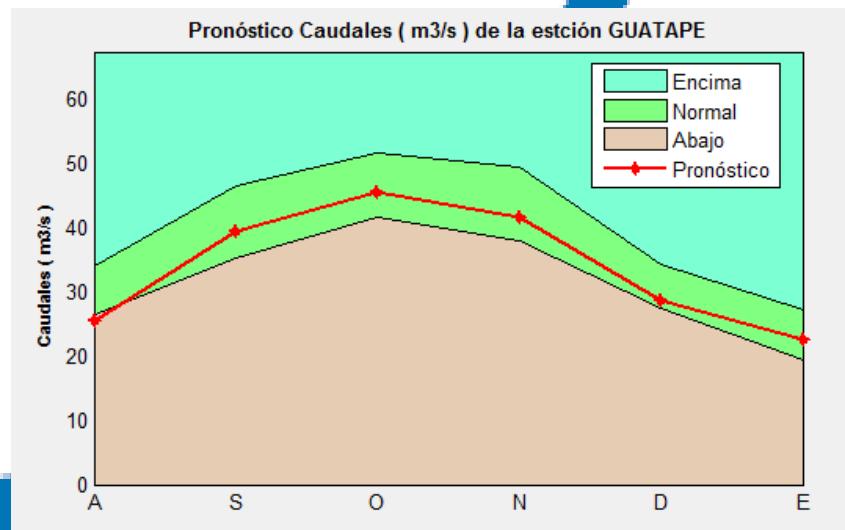




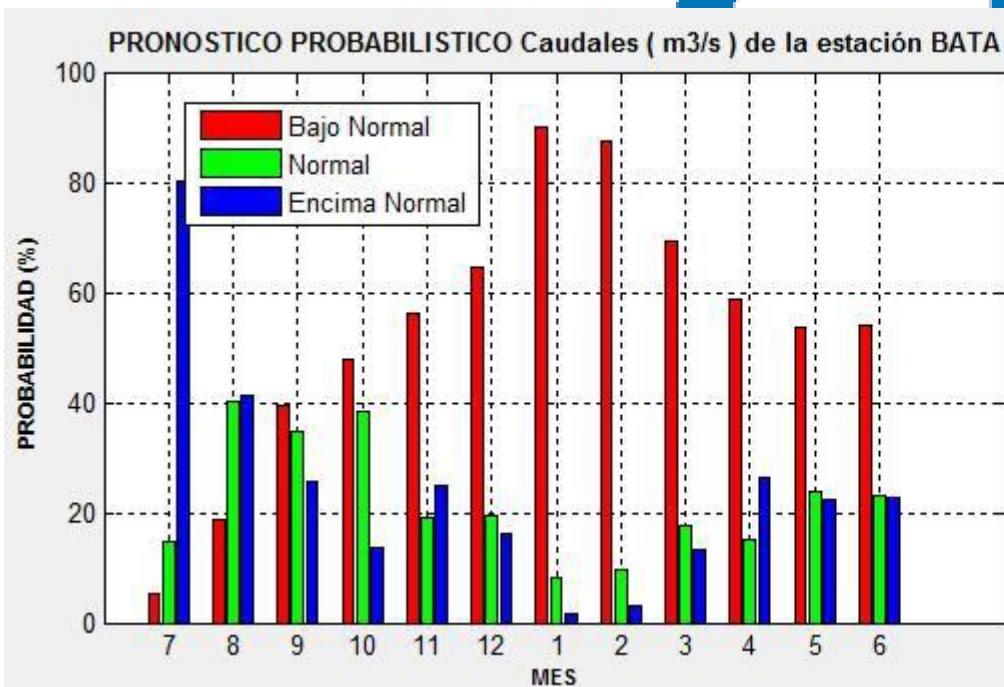
Pronóstico de Caudales en Colombia



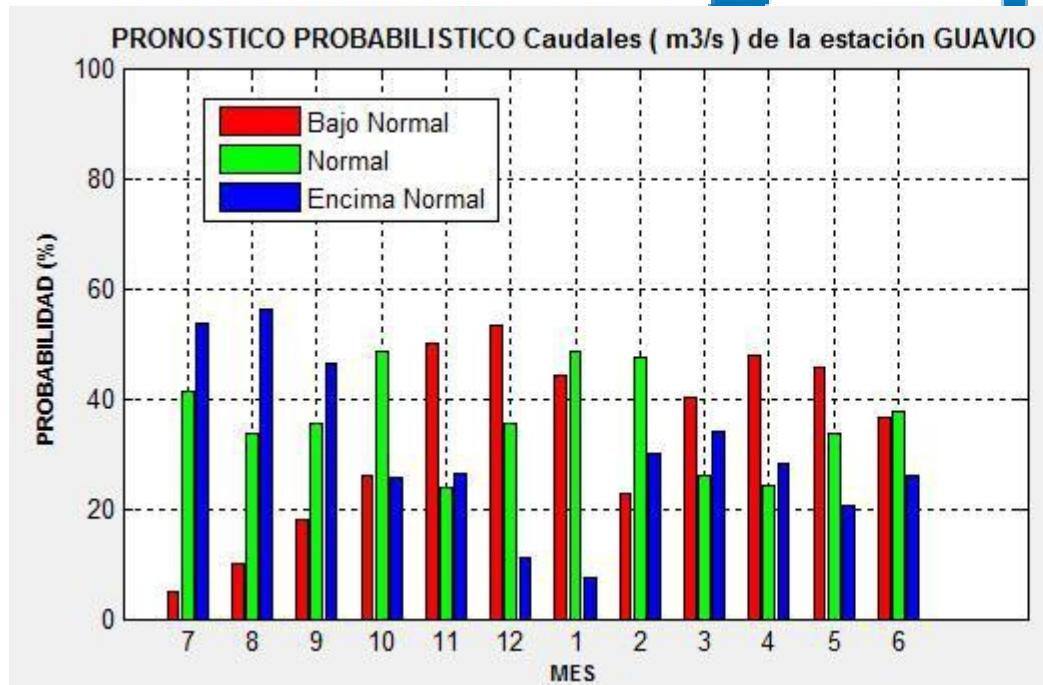
Pronóstico desde
Agosto de 2012
hasta Enero de
2013 (6 meses)



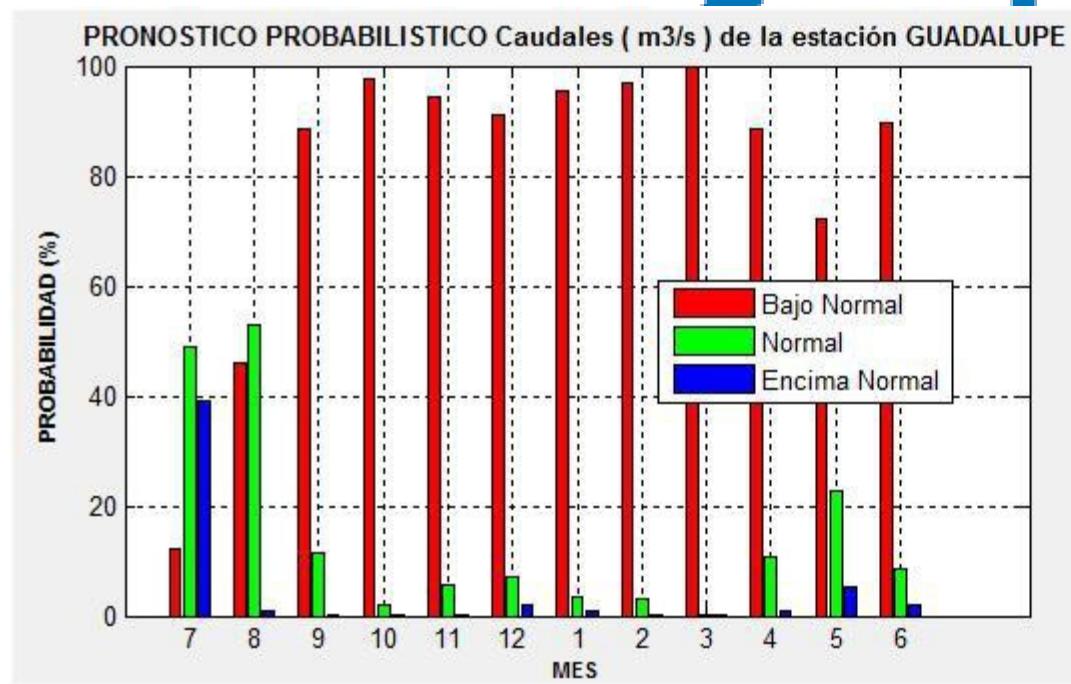
Rio Bata Julio 2012 – Junio 2013



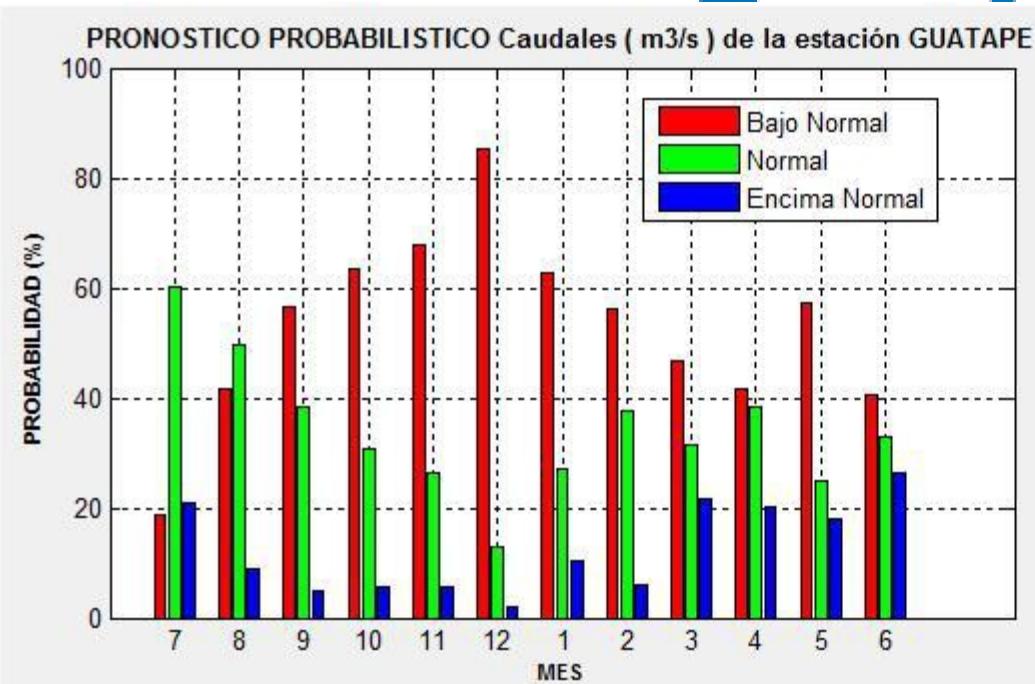
Rio Guavio Julio 2012 – Junio 2013



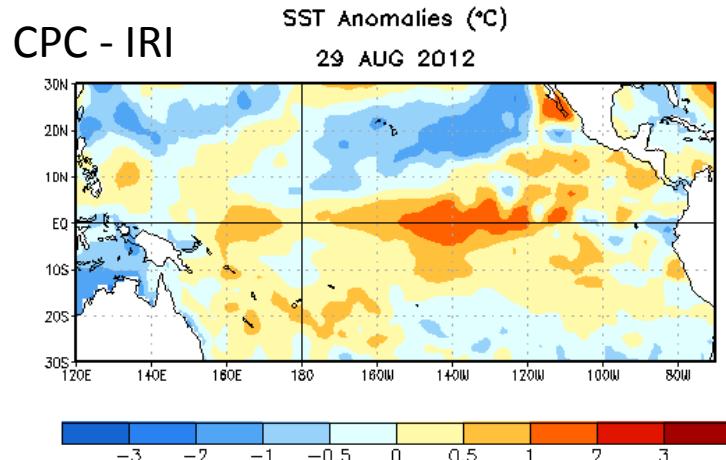
Rio Guadalupe Julio 2012 – Junio 2013



Rio Guatapé Julio 2012 – Junio 2013

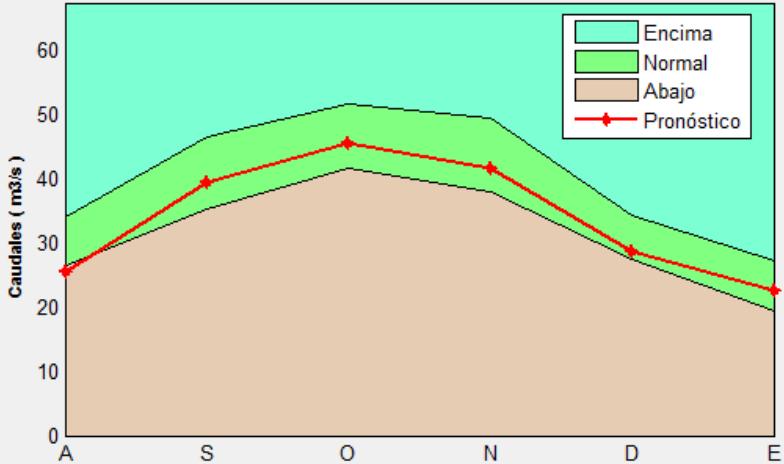


10. RESULTADOS (II)

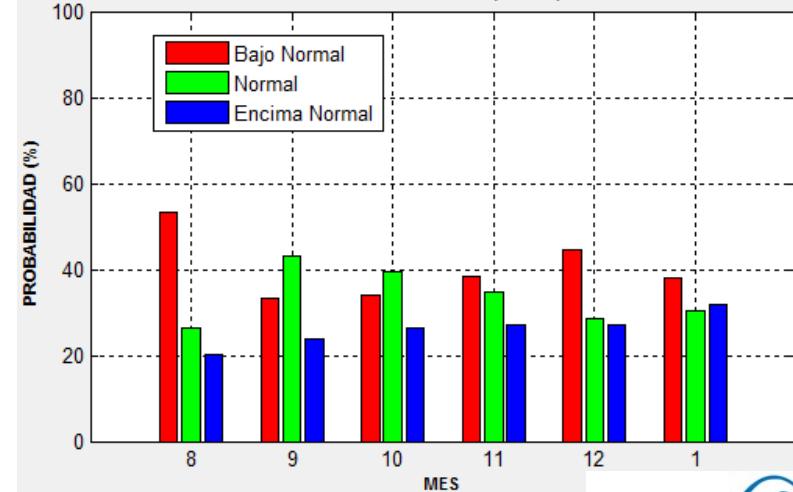


Pronóstico desde
Agosto de 2012
hasta Enero de
2013 (6 meses)

Pronóstico Caudales (m³/s) de la estación GUATAPE



PRONOSTICO PROBABILISTICO Caudales (m³/s) de la estación GUATAPE



11. CONCLUSIONES

- Para la predicción de caudales medios mensuales se ha propuesto un modelo híbrido basado en funciones de influencia radial (RBF) y técnicas espectrales.
- El método propuesto tienen la ventaja de ser no-paramétricas, por lo que el cálculo de los parámetros de calibración no depende de la distribución probabilística de los datos, además pondera la información en un campo discreto de datos sirviendo para la modelación de procesos no lineales y no estacionarios como los caudales.
- Con el modelo híbrido de predicción basado en la RBF multiquadrática y la transformada en onditas se obtiene mejores resultados para horizontes de predicción de 1 y 3 meses y un modelo de RBF+AES es el mejor para una ventana de 6 meses.

BIBLIOGRAFÍA

- Ghil M., Allen M., Dettinger M., Ide M., Kondrashov D., Mann M., Robertson A., Saunders A., Tian Y. y Varadi F. y Yiou P. (2002). "Advanced Spectral Methods for Climatic Time Series". *Review of Geophysics*. American Geophysical Union, Vol. 40, N° 1. pp. 1-41.
- H. Xie, Z. Wang. (2006). "Mean frequency derived via Hilbert.Huang transform with application to fatigue EMG signal analisys". *Comp. methods and programs in Biomed.* 82, pp. 114-120.
- Mallat, S. (1989) "Multiresolution approximation and wavelet orthonormal bases of L^2 ," *Trans. Amer. Math. Soc.*, June.
- N. E. Huang, Z. Shen, S. R. Long, M. L. C. Wu, H. H. Shin, Q. N. Zheng, N. C. Yen, C. C. Tung, H. H. Liu. (1998). "The empirical mode decomposition and the Hilbert spectrum for non-linear and non-stationary time series analysis". *Proc. Roy. Soc. Lond. Ser. A: Math. Phys. Eng. Sci.* 454, pp. 903-995,
- Priestley, M. B. y Chao, M. T. (1972): "Non-parametric function fitting". *J. Royal Stat. Soc. B*, 34, 385-392.
- Poveda G. (2006). "Aplicación de los métodos MARS, Holt-Winters y ARIMA generalizado en el pronóstico de caudales medios mensuales en ríos de Antioquia". *Rv Colombia Meteorología Colombiana*, ISSN: 0124-6984 ed: Gente Nueva v.10 fasc.1 p.36 – 46.
- Powell M.J.D. (1987). " Radial basis functions for multivariable interpolation: a review, Algorithms for Approximation", Mason J.C., Cox M.G. (eds.), London, Oxford University Press.
- Queipo NV, Haftka RT, Shyy W, Goel T, Vaidyanathan R, Tucker PK (2005). "Surrogate-based analysis and optimization". *Prog Aerosp Sci* 41:1–28
- Torrence, C., and G. Compo (1998). "A practical guide to wavelet analysis". *Bull. Am. Meteorol. Soc.*, 79, 61-- 78.
- Vautard, R. (1992). "Singular-spectrum analysis: A toolkit for shot, noisy chaotic signals". *Physica, D* 58'(1992): 95-126. North-Holland.

AGRADECIMIENTOS



Departamento Administrativo de
Ciencia, Tecnología e Innovación
Colciencias

República de Colombia



UNIVERSIDAD
NACIONAL
DE COLOMBIA
SEDE MEDELLÍN

