

DESARROLLO DE UN SISTEMA EXPERTO PARA LA PREDICCIÓN DE CAUDALES MEDIOS MENSUALES EN COLOMBIA

Julián David Rojo Hernández, I.C

Director

Luis Fernando Carvajal Serna, I.C, MSc.

Universidad Nacional de Colombia-sede Medellín
Escuela de Geociencias y Medio Ambiente -Facultad de Minas
Medellín- Julio 29 de 2011



UNIVERSIDAD
NACIONAL
DE COLOMBIA
SEDE MEDELLÍN



CONTENIDO

1. PROBLEMÁTICA DE LA PREDICCIÓN DE CAUDALES
2. RELACIONES ENTRE LOS CAUDALES DE COLOMBIA Y LA CLIMATOLOGÍA GLOBAL
3. MÉTODOS ACTUALES PARA EL PRONÓSTICO DE CAUDALES
4. INCORPORACIÓN DE NUEVAS TÉCNICAS DE PRONÓSTICO
5. COMPARACIÓN Y COMBINACIÓN DE PRONÓSTICOS HIDROLÓGICOS
6. PREDICCIÓN BASADA EN REGLAS DE JUICIO
7. PROTOTIPO DE SOFTWARE
8. CONCLUSIONES Y RECOMENDACIONES



CONTENIDO

- 1. PROBLEMÁTICA DE LA PREDICCIÓN DE CAUDALES**
2. RELACIONES ENTRE LOS CAUDALES DE COLOMBIA Y LA CLIMATOLOGÍA GLOBAL
3. MÉTODOS ACTUALES PARA EL PRONÓSTICO DE CAUDALES
4. INCORPORACIÓN DE NUEVAS TÉCNICAS DE PRONÓSTICO
5. COMPARACIÓN Y COMBINACIÓN DE PRONÓSTICOS HIDROLÓGICOS
6. PREDICCIÓN BASADA EN REGLAS DE JUICIO
7. PROTOTIPO DE SOFTWARE
8. CONCLUSIONES Y RECOMENDACIONES



1.1 SOBRE LA NECESIDAD DEL PRONÓSTICO DE CAUDALES

De todos los elementos del ciclo hidrológico los caudales de los ríos pueden ser los que mayores consecuencias tienen sobre la vida de los seres humanos.

Seguridad Alimentaria

Generación de Hidroelectricidad

Consumo Doméstico

Abastecimiento de industrias

Ríos como vía de comunicación

Hábitat de especies Animales

La oferta hídrica provista por los ríos es variable y depende del comportamiento climático a escala global, en consecuencia conocer a futuro la cantidad de agua disponible, es una labor estratégica y necesaria en la planeación y manejo del recurso hídrico.



1.2 EXPERIENCIAS EN LA PREDICCIÓN DE CAUDALES MEDIOS MENSUALES EN COLOMBIA.

Años 70-80:

AR, ARMA, ARIMA

Años 90:

RAR (Salazar, 1994), AES (Carvajal, 1994), RNA (Carvajal, 1994), MARS (Rendón, 1997), PRBEO (Hoyos, 1998).

Años 2000-2010:

Regresiones Kernel (Poveda et al, 2002), ANFIS (Smith et al, 2004), Modelos Lineales Inversos (Ceballos, 2005), Funciones Holt – Winters (Poveda et al. 2006).



1.4 BARRERAS QUE DIFICULTAN LA PREDICCIÓN DE CAUDALES MEDIOS MENSUALES EN COLOMBIA. (II)

- Las características particulares de las series de caudales medios mensuales (no estacionalidad, cuasi-periodicidad, etc.), la complejidad y el gran número de factores que los afectan, imponen los siguientes requerimientos (Velásquez, 2008):
- Modelación no lineal.
- Capacidad para manejar simultáneamente una gran cantidad de variables explicativas.
- Capacidad de ponderar la información disponible en el tiempo.



1.5 OBJETIVOS

Objetivo general:

- Mejorar la predicción de caudales medios mensuales mediante la utilización de un prototipo basado en modelos matemáticos y juicio experto.

Objetivos específicos:

- Establecer una base sobre el estado actual de predicción de caudales medios mensuales en el País.
- Incorporar en la construcción del modelo el uso de variables explicativas que reflejen el comportamiento macro-climático
- Revisar los últimos avances en temas relacionados con la predicción en hidrología e incorporar aquellos que sean relevantes para caso Colombiano.
- Ajustar los pronósticos mediante reglas de juicio.
- Desarrollar un Prototipo de Software para la predicción de Caudales Medios Mensuales Basado en modelos matemáticos y reglas de juicio.



1.7 CASOS DE APLICACIÓN

Horizontes
de
pronóstico

1 mes
3 meses
6 meses
12 meses



UNIVERSIDAD
NACIONAL
DE COLOMBIA
SEDE MEDELLÍN

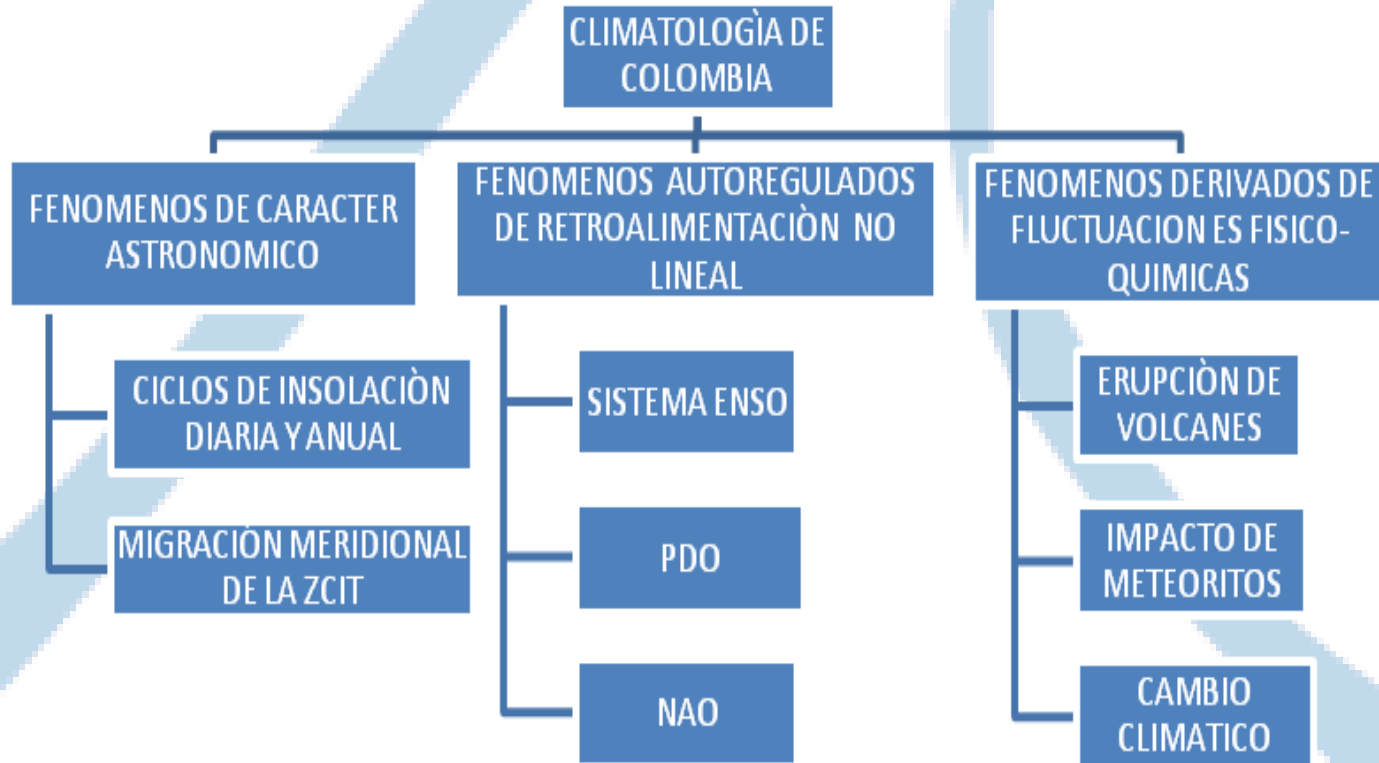
POSGRADO EN APROVECHAMIENTO
DE RECURSOS HIDRÁULICOS

CONTENIDO

1. PROBLEMÁTICA DE LA PREDICCIÓN DE CAUDALES
- 2. RELACIONES ENTRE LOS CAUDALES DE COLOMBIA Y LA CLIMATOLOGÍA GLOBAL**
3. MÉTODOS ACTUALES PARA EL PRONÓSTICO DE CAUDALES
4. INCORPORACIÓN DE NUEVAS TÉCNICAS DE PRONÓSTICO
5. COMPARACIÓN Y COMBINACIÓN DE PRONÓSTICOS HIDROLÓGICOS
6. PREDICCIÓN BASADA EN REGLAS DE JUICIO
7. PROTOTIPO DE SOFTWARE
8. CONCLUSIONES Y RECOMENDACIONES



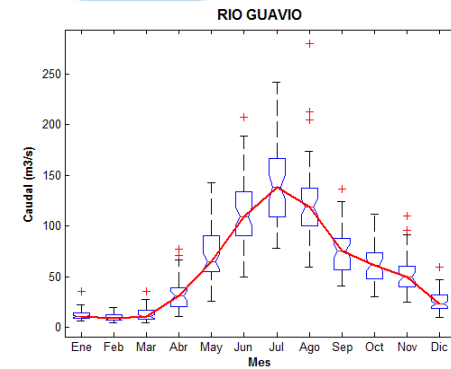
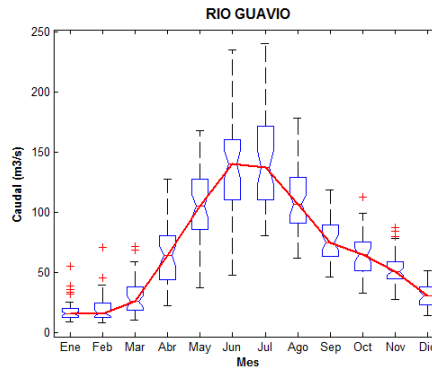
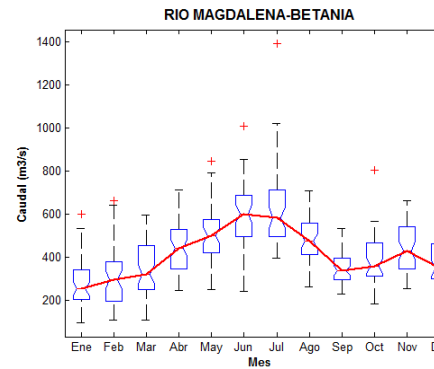
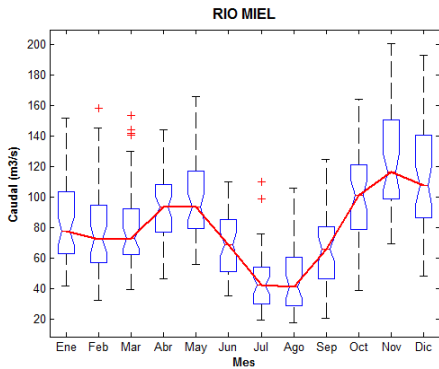
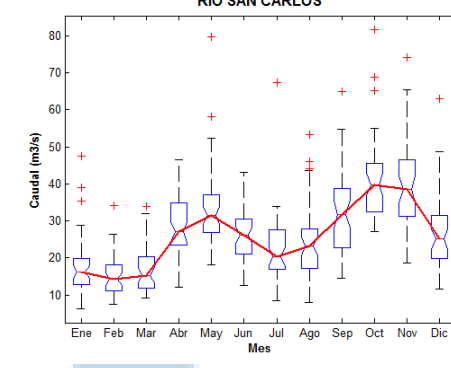
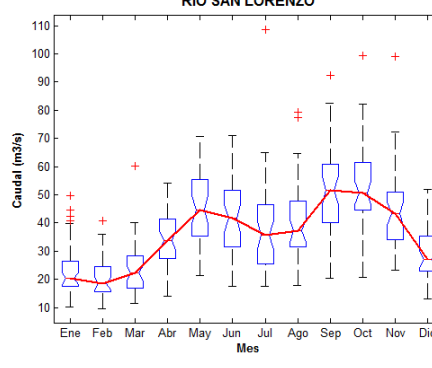
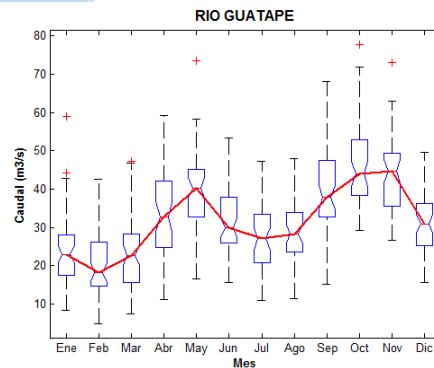
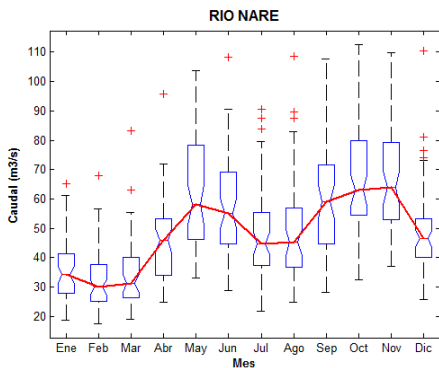
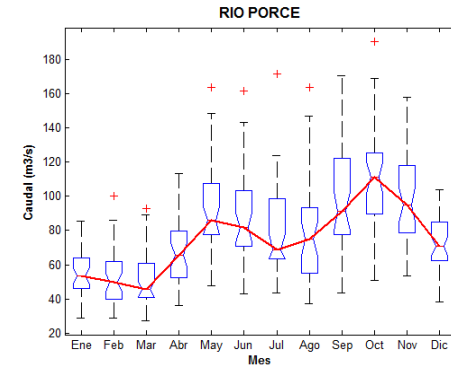
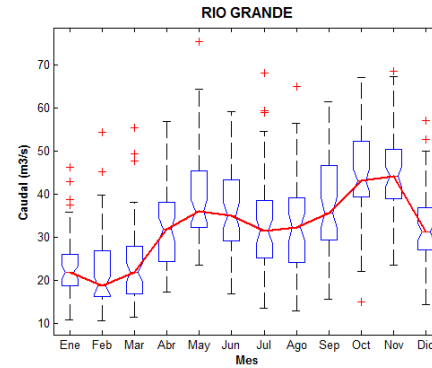
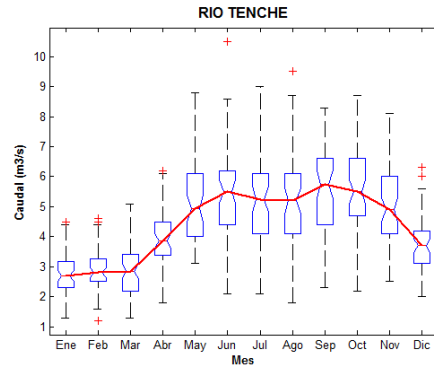
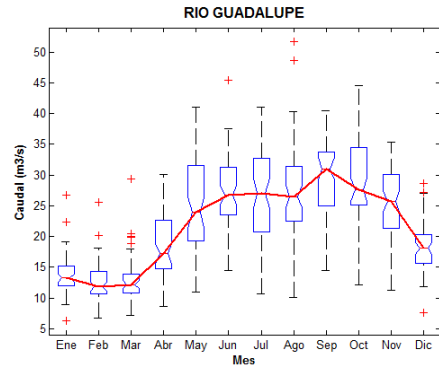
2.1 HIDROCLIMATOLOGÍA DE COLOMBIA A DIFERENTES ESCALAS ESPACIO-TEMPORALES



(Poveda, 2004)

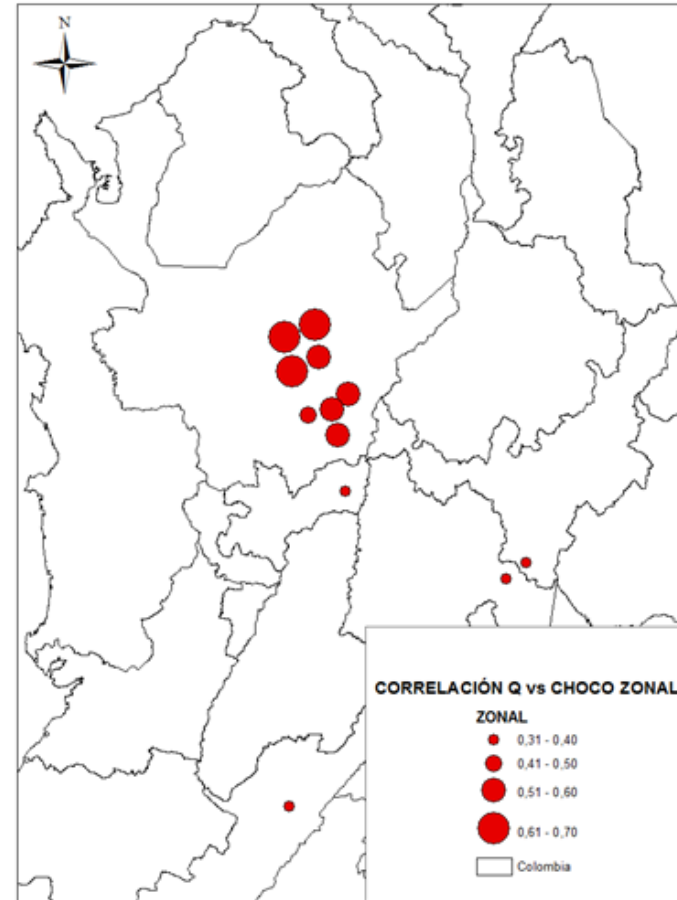
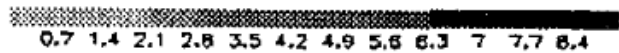
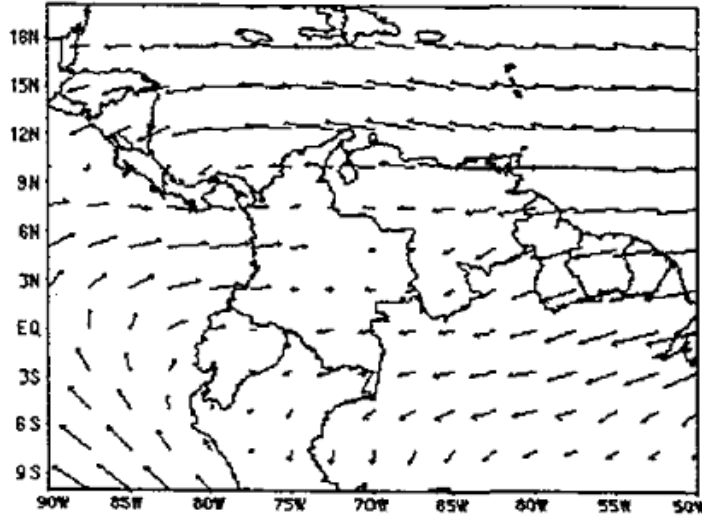


2.2 VARIABILIDAD ANUAL DE LOS CAUDALES EN COLOMBIA (II)-CICLO ANUAL

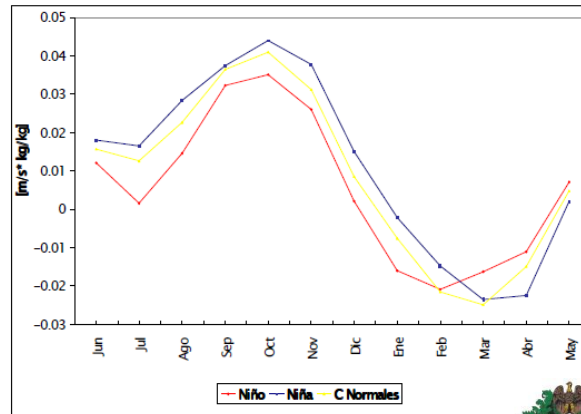
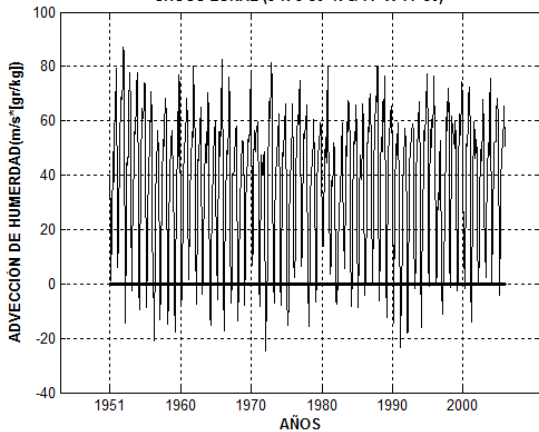


2.2 VARIABILIDAD ANUAL DE LOS CAUDALES EN COLOMBIA (III)-CHORRO DEL CHOCO

SON



CHOCO ZONAL (5°N-5°30' N & 77°W-77°30')



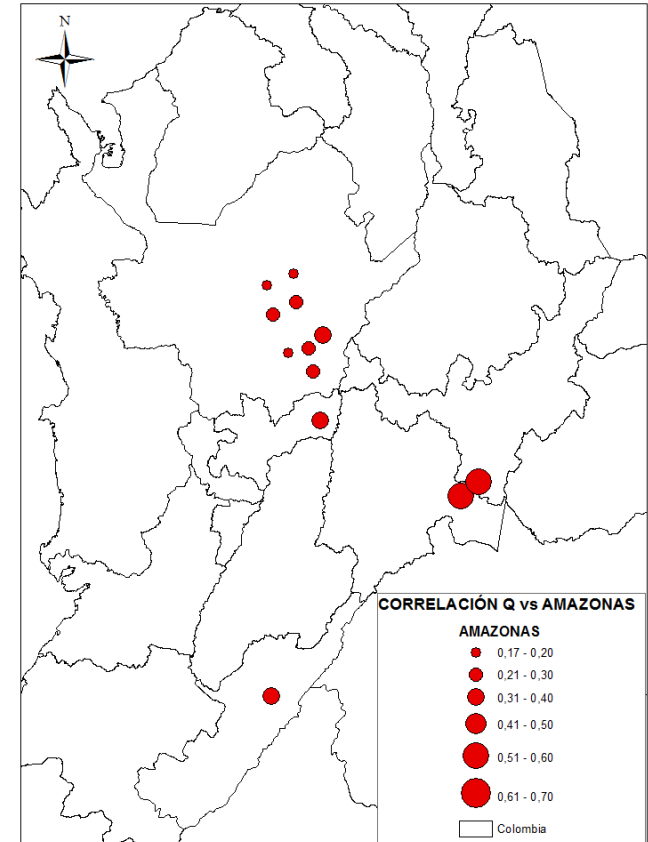
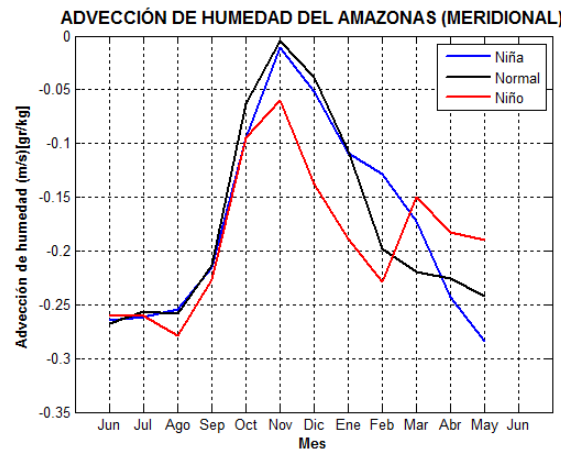
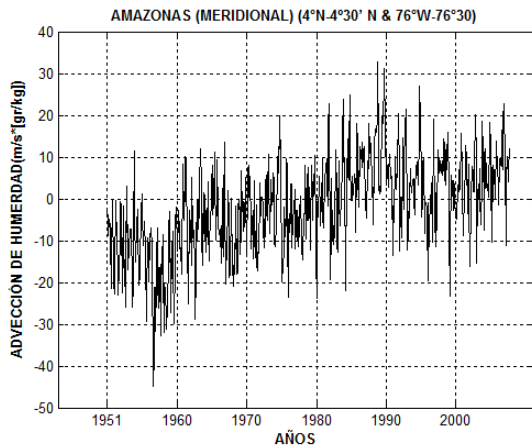
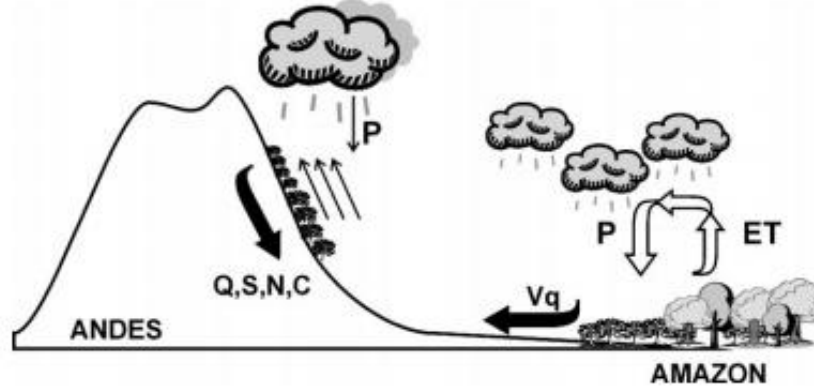
(Rendón, 2001), Poveda et al (2006)



UNIVERSIDAD
NACIONAL
DE COLOMBIA
SEDE MEDELLÍN

POSGRADO EN APROVECHAMIENTO
DE RECURSOS HIDRÁULICOS

2.2 VARIABILIDAD ANUAL DE LOS CAUDALES EN COLOMBIA (IV)- ADVECCIÓN DE HUMEDAD DESDE EL AMAZONAS



Poveda et al. (2006)

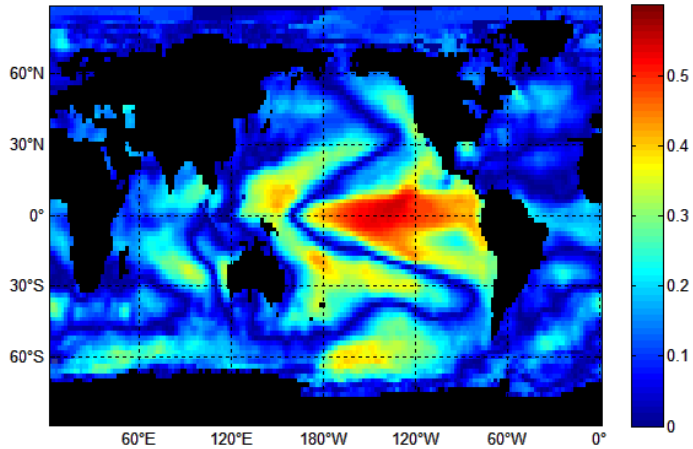


UNIVERSIDAD
NACIONAL
DE COLOMBIA
SEDE MEDELLÍN

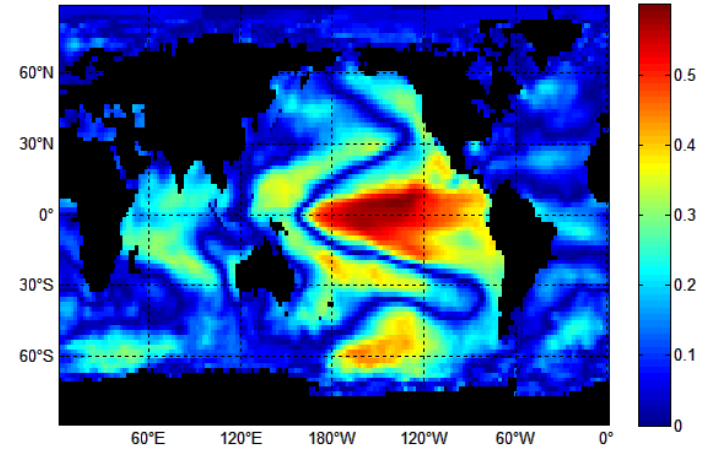
POSGRADO EN APROVECHAMIENTO
DE RECURSOS HIDRÁULICOS

2.3 VARIABILIDAD DE LOS CAUDALES MEDIOS MENSUALES A ESCALA INTER-ANUAL (I)

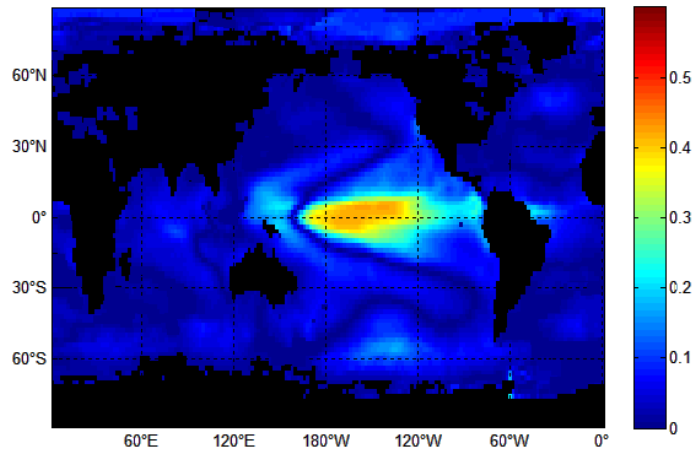
ICP RVP vs ANOMALIAS SST(-1)



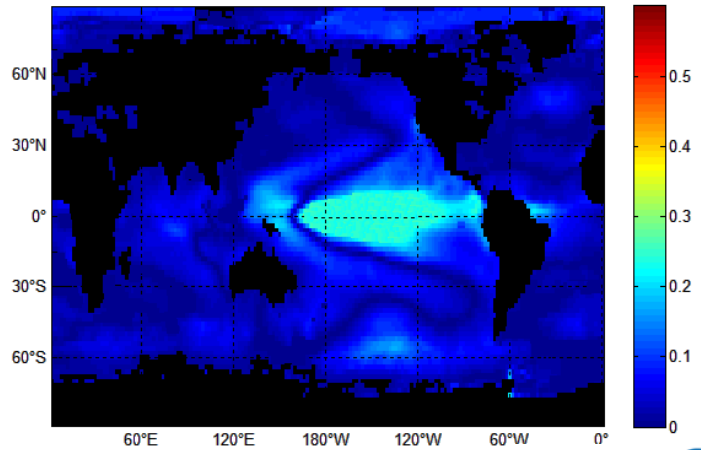
ICP RVN vs ANOMALIAS SST(-2)



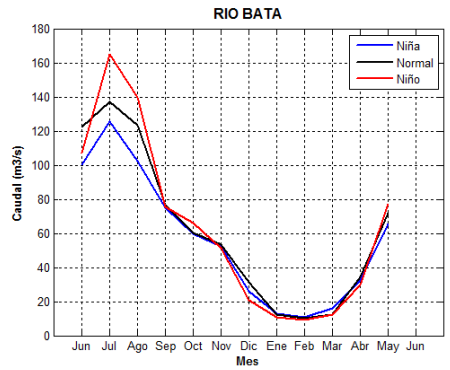
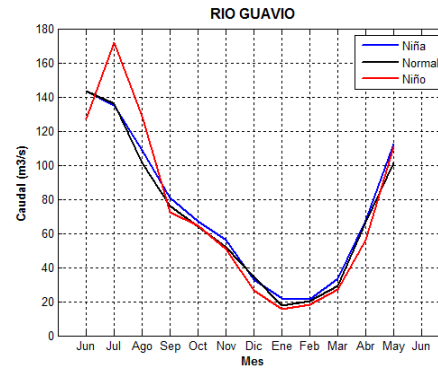
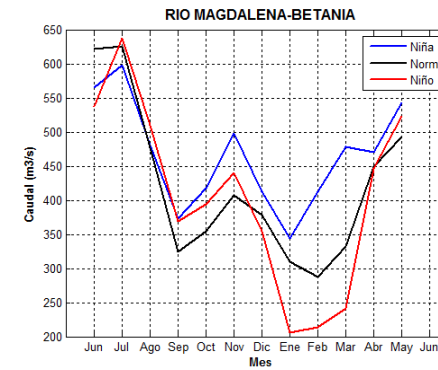
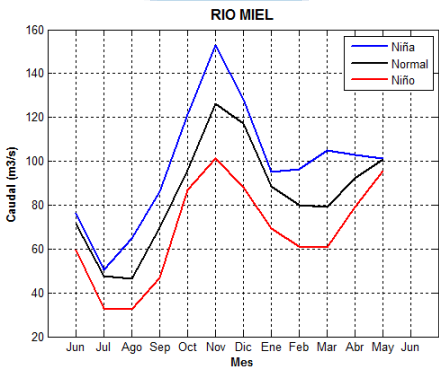
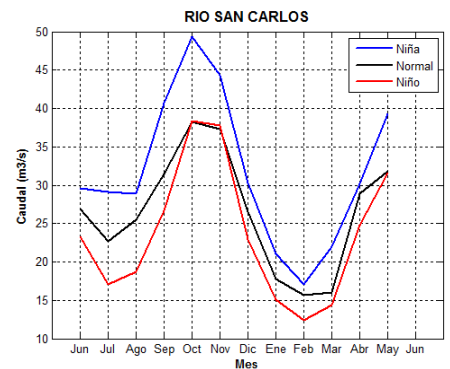
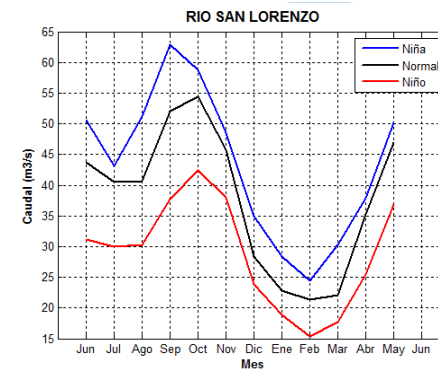
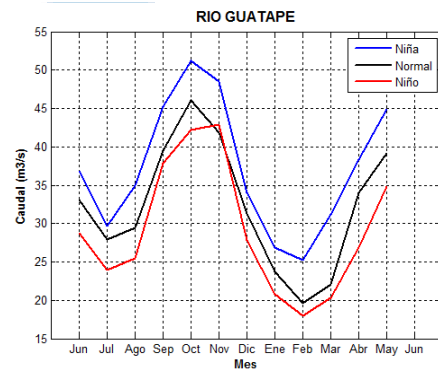
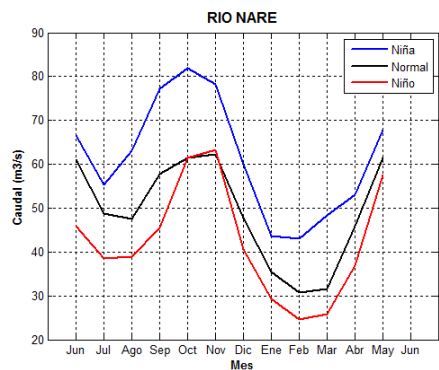
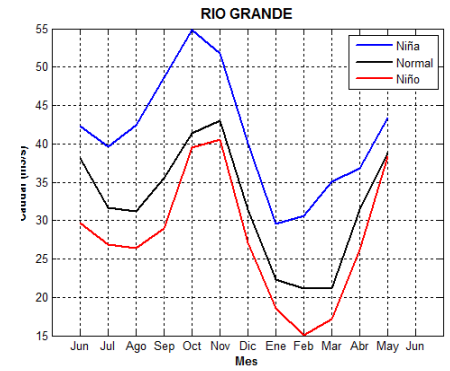
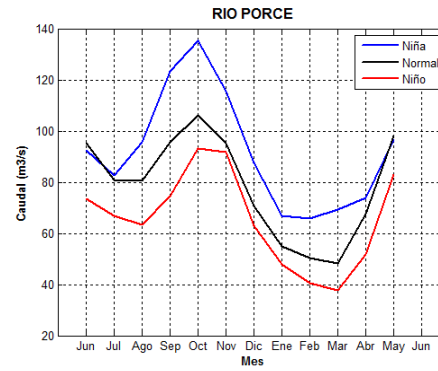
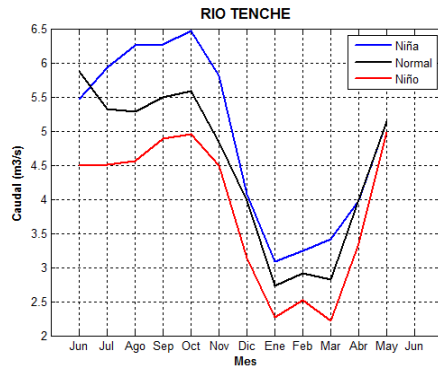
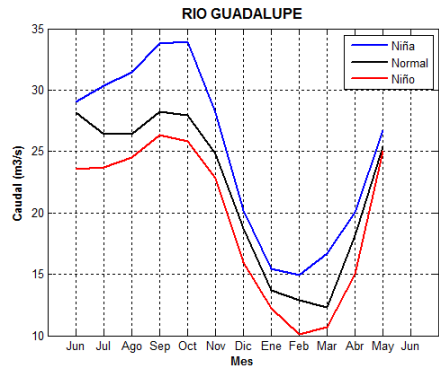
ICP RVM vs ANOMALIAS SST(-3)



ICP ROC vs ANOMALIAS SST(-6)

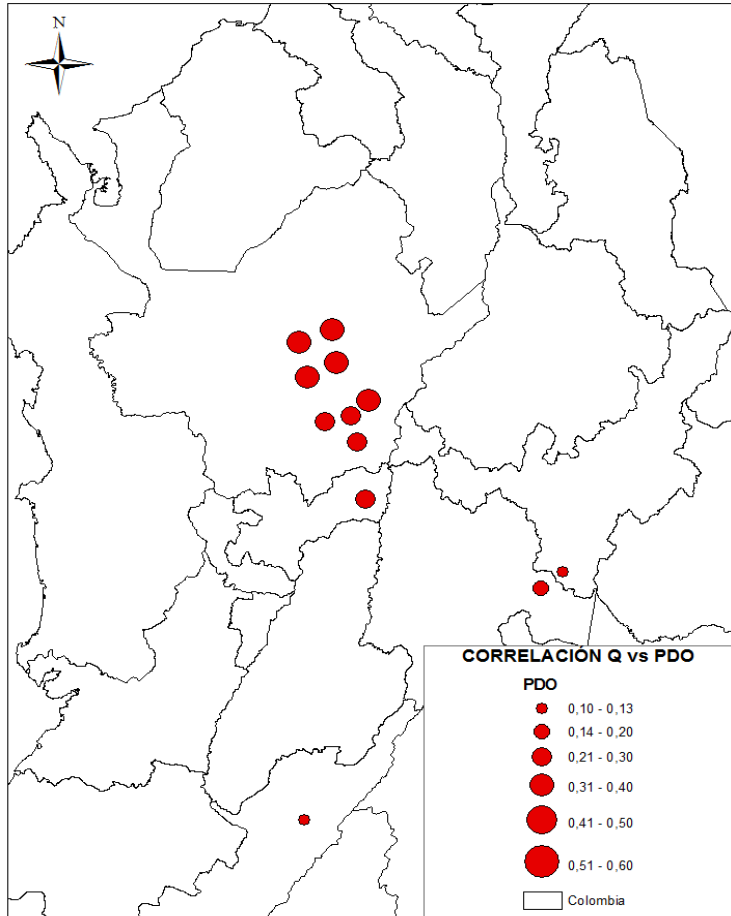


2.3 VARIABILIDAD DE LOS CAUDALES MEDIOS MENSUALES A ESCALA INTER-ANUAL (II)

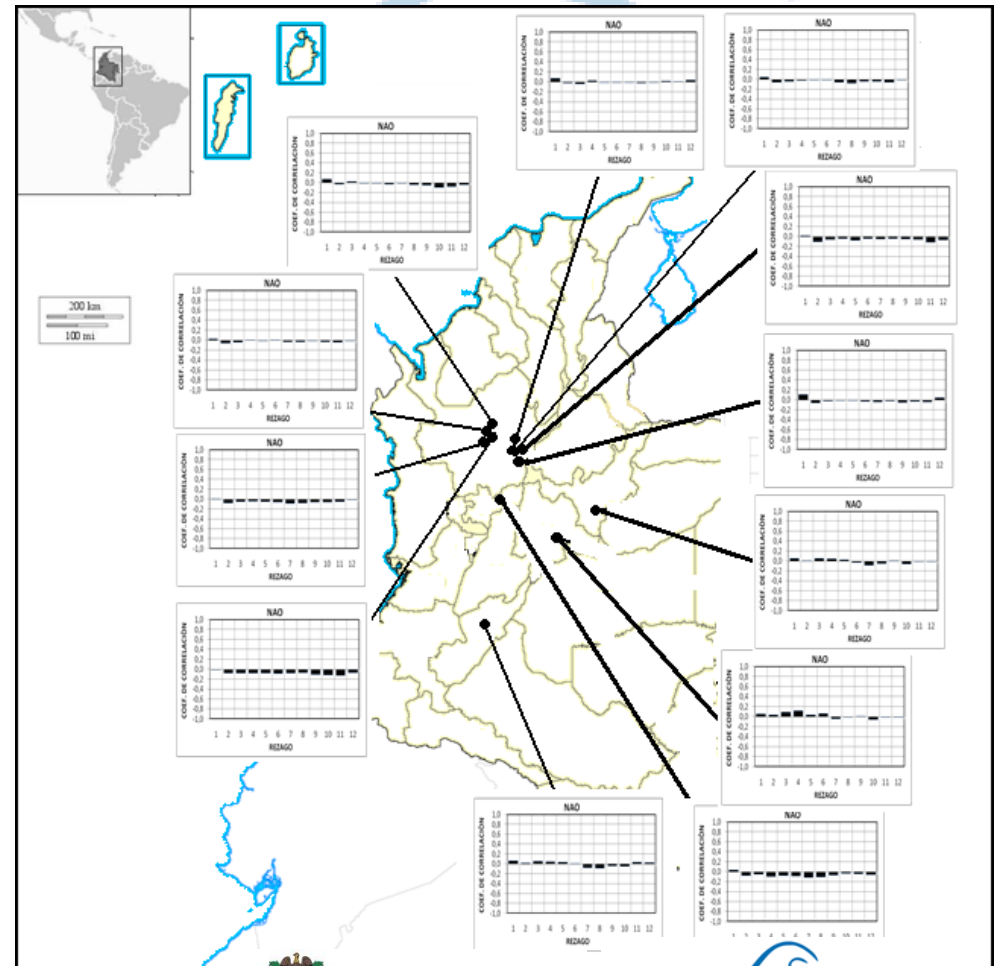


2.4 VARIABILIDAD DE LOS CAUDALES MEDIOS MENSUALES A ESCALA INTER-DECADAL

PDO



NAO



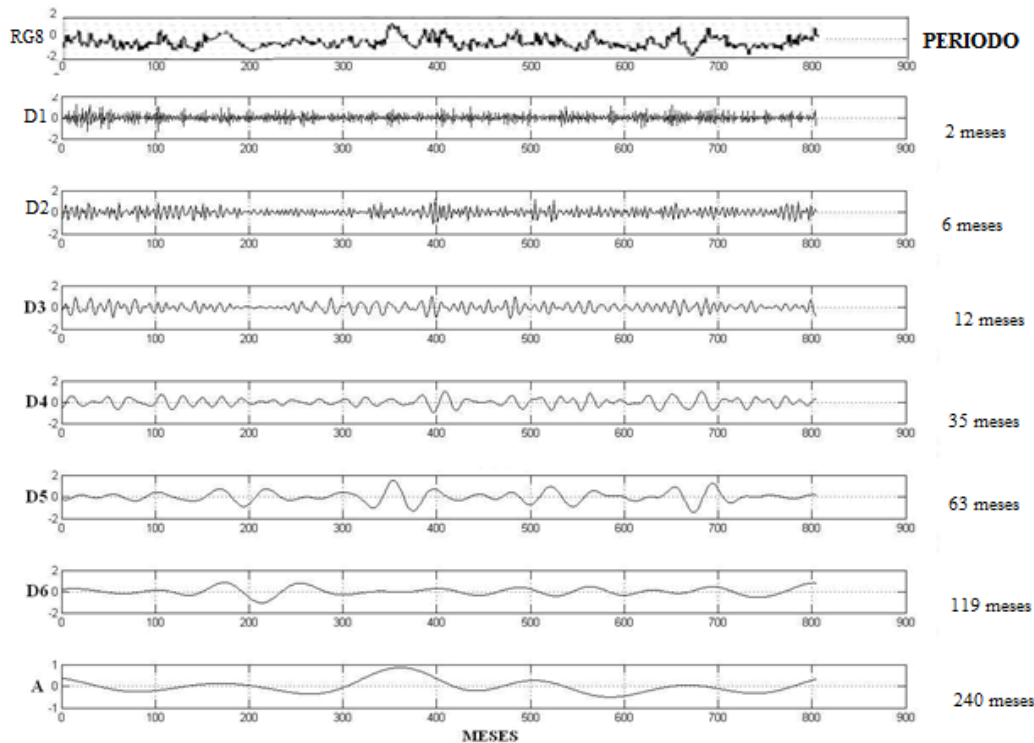
2.5 OTROS ANÁLISIS

Análisis de Componentes principales para ríos de Colombia, Rojo y Carvajal (2010)

	1CP	2CP	3CP	4CP	5CP	6CP	7CP	8CP	9CP	10CP	11CP	12CP
RIO SAN CARLOS	128-256	32	14	8	4	8-3	5	4	3	4-2	3	4
RIO GRANDE	128-64	32-64	16	8	6	5	4	3	2	3	3	3
RIO NARE	128-64	32	16	8	8	6	5	4	4	3	3	2
RIO GUAVIO	128-64	16	64-12	6	5	4	4	2	3	2	2	2
RIO BETANIA	128-64	32-12	32-16	8	6	4	5	8-4-2	2	4	3	3

DESCOMPOSICIÓN EN ONDITAS PARA LA ESTACIÓN RG8

ONDITA DISCRETA DE MEYER



Rojo y Carvajal
(2010)



UNIVERSIDAD
NACIONAL
DE COLOMBIA
SEDE MEDELLÍN

POSGRADO EN APROVECHAMIENTO
DE RECURSOS HIDRÁULICOS

CONTENIDO

1. PROBLEMÁTICA DE LA PREDICCIÓN DE CAUDALES
2. RELACIONES ENTRE LOS CAUDALES DE COLOMBIA Y LA CLIMATOLOGÍA GLOBAL
- 3. MÉTODOS ACTUALES PARA EL PRONÓSTICO DE CAUDALES**
4. INCORPORACIÓN DE NUEVAS TÉCNICAS DE PRONÓSTICO
5. COMPARACIÓN Y COMBINACIÓN DE PRONÓSTICOS HIDROLÓGICOS
6. PREDICCIÓN BASADA EN REGLAS DE JUICIO
7. PROTOTIPO DE SOFTWARE
8. CONCLUSIONES Y RECOMENDACIONES



3.1 ANÁLISIS PREDICTIVO

- Por lo general, el análisis de regresión estima el valor esperado (esperanza condicional) de una variable dependiente y dadas las variables independientes (X_1, X_2, \dots, X_n)

$$y = f(X_1, X_2, \dots, X_n) + \varepsilon$$

$$Caudales_t = f(Caudales_{t-j}, ENSO_{t-k}, CHOCO_{t-l}, \dots) + \varepsilon$$

- 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.



3.2 FUNCIONES DE REGRESIÓN

- ARIMA**

Auto-regresivos Integrados de Media Móvil

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

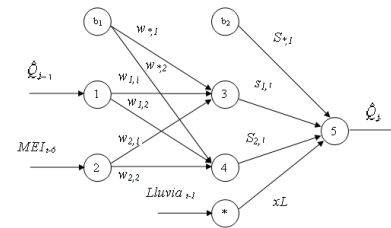
- RLM**

Regresión Lineal Múltiple

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

- RNA**

Redes Neuronales Artificiales



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

- MARS**

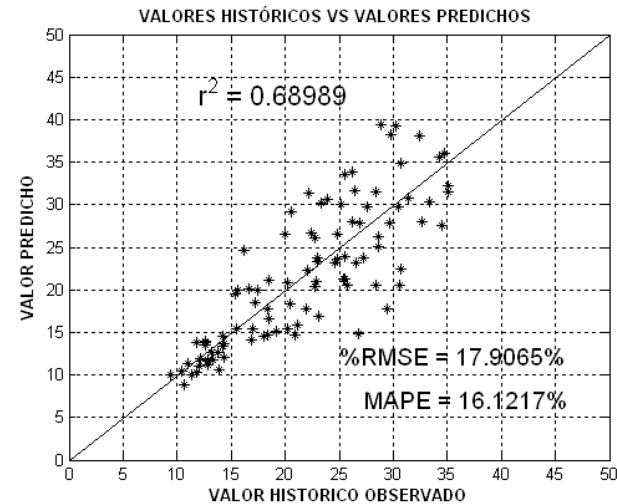
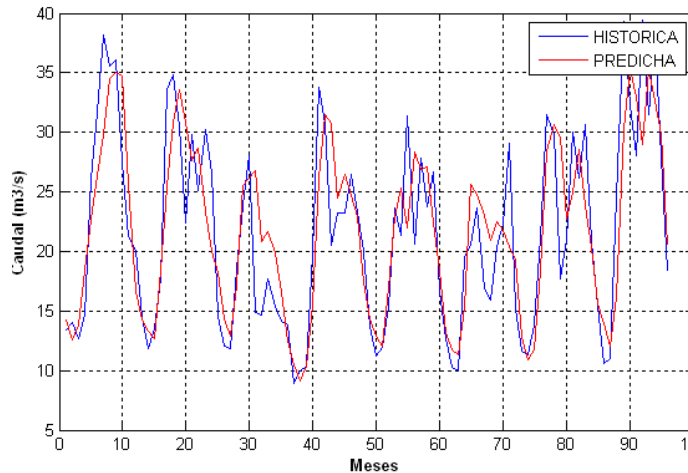
Polinomios Adaptivos de Regresión Multivariada

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



3.3 ESTRATEGIA DE VALIDACIÓN

Predicción de caudales del río Guadalupe utilizando
Regresión Lineal Múltiple



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

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

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

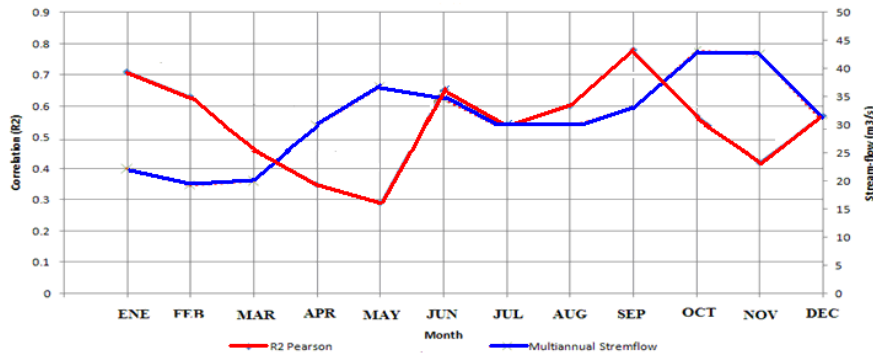
$$MSE = \frac{SSE}{T}$$

$$RMSE = \sqrt{MSE}$$

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

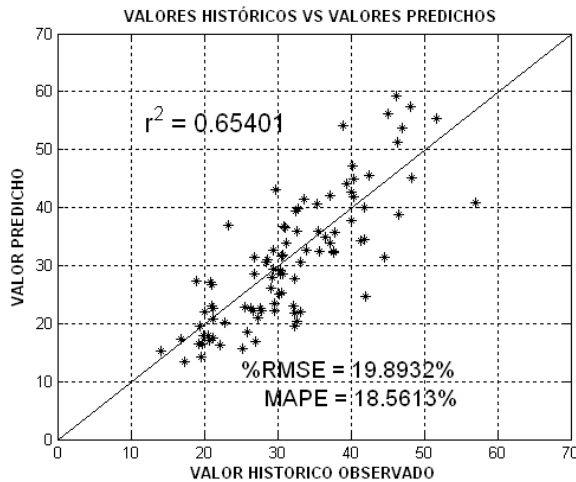
3.5 MODELOS PERIÓDICOS DE PRONÓSTICO

Correlacion de Pearson Vs mes para el rio Grande
Validación 2000–2007 (ventana de 1 mes)

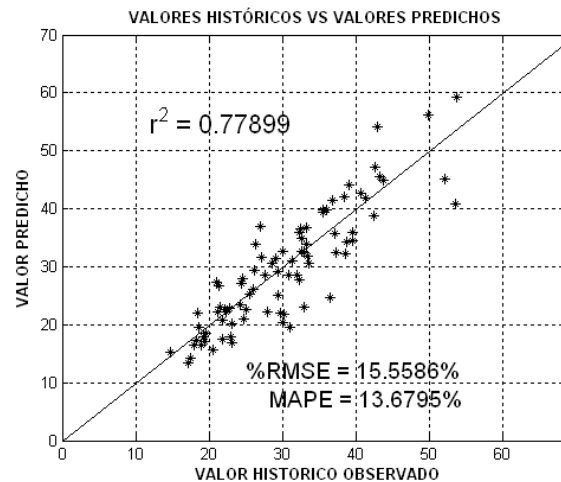


$$\hat{Q}_{v\tau} = \hat{f}(X, \tau) = \begin{cases} \hat{f}_1(X_{1,1}, X_{2,1} \dots X_{n,1}) \Rightarrow \tau = 1 \\ \hat{f}_2(X_{1,2}, X_{2,2} \dots X_{n,2}) \Rightarrow \tau = 2 \\ \vdots \\ \hat{f}_\tau(X_{1,\tau}, X_{2,\tau} \dots X_{n,\tau}) \\ \vdots \\ \hat{f}_{12}(X_{1,12}, X_{2,12} \dots X_{n,12}) \Rightarrow \tau = 12 \end{cases}$$

Predicción de Caudales medios del Rio Grande con ventana de un mes



Predicción con modelo Lineal

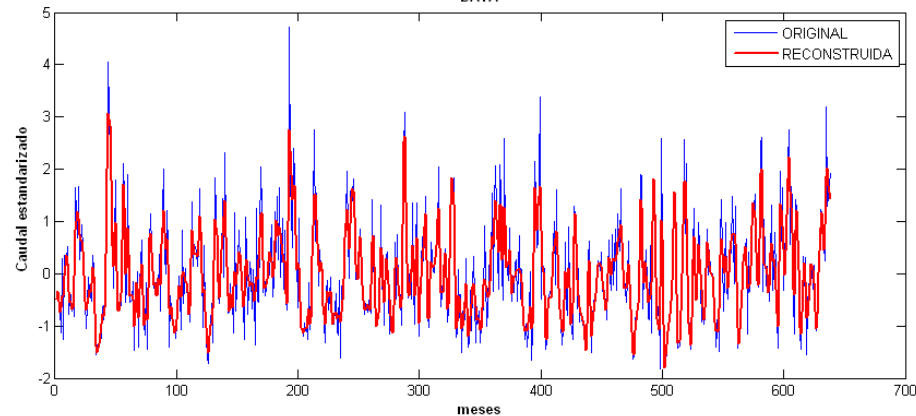


Predicción con modelo Lineal periódico

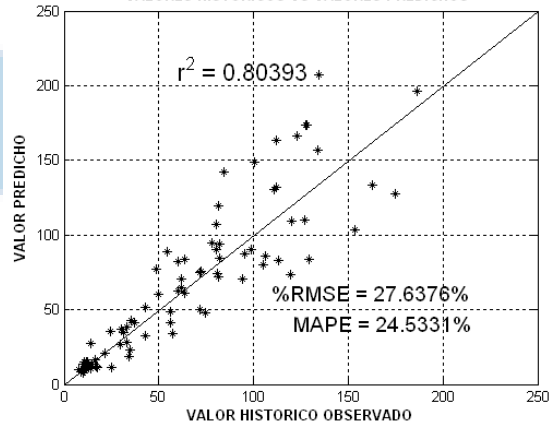


3.6 INCORPORACIÓN DE MÉTODOS ESPECTRALES

SERIE ORIGINAL VS SERIE RECONSTRUIDA RIO BATA

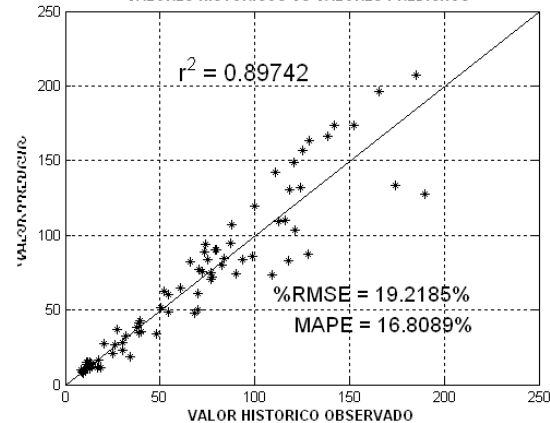


VALORES HISTÓRICOS VS VALORES PREDICHOS



Predicción con modelo lineal periódico

VALORES HISTÓRICOS VS VALORES PREDICHOS



Predicción con modelo lineal periódico & AES



3.8 CONSTRUCTOR DE MODELOS DE PRONÓSTICO



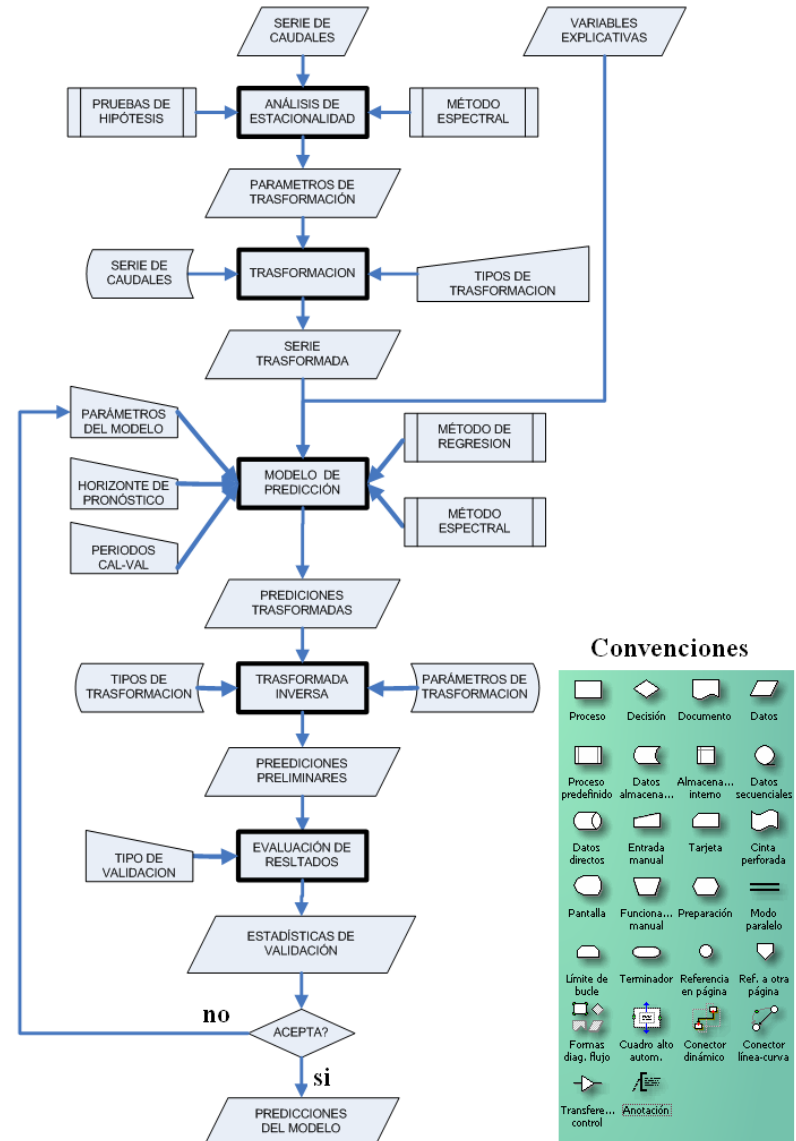
UNIVERSIDAD
NACIONAL
DE COLOMBIA
SEDE MEDELLÍN



POSGRADO EN APROVECHAMIENTO
DE RECURSOS HIDRÁULICOS

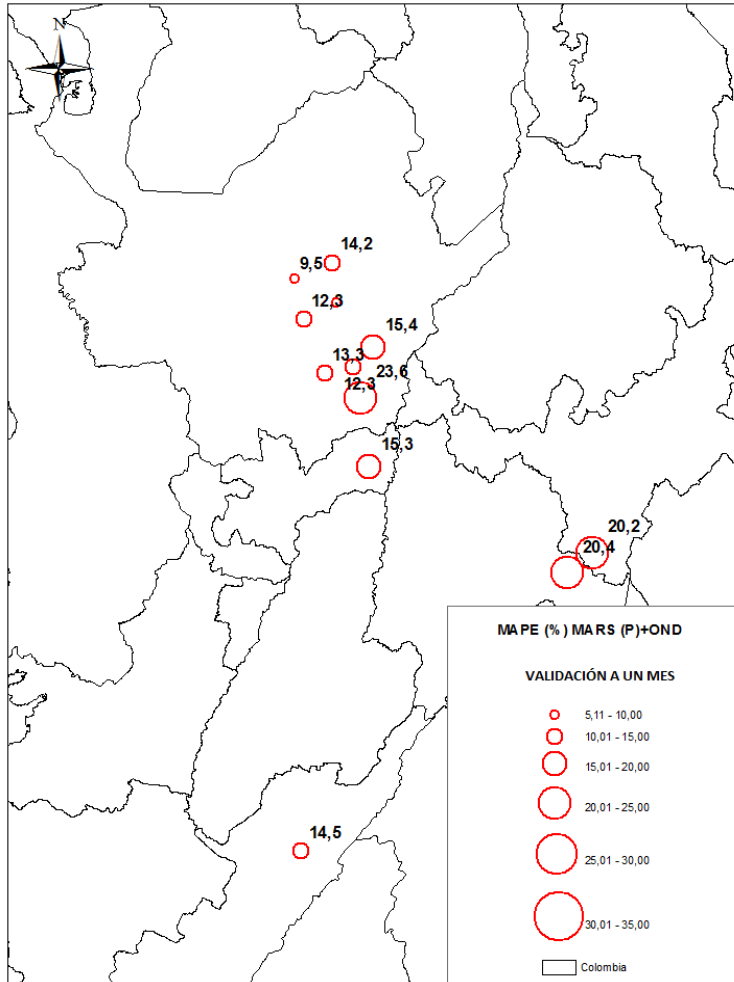
3.9 MODELOS SELECCIONADOS

- **AR (2)**
Modelo auto-regresivo de orden 2.
- **RLM**
Regresión lineal múltiple.
- **RNA (P) + HH**
Modelo periódico de regresión basado RNA y HH.
- **RLM (P) + AES**
Modelo periódico de regresión basado RLM y AES.
- **MARS (P) + OND**
Modelo periódico de regresión basado MARS y ONI
- **PERÍODO DE CALIBRACIÓN**
Inicio de la serie de tiempo hasta 1989
- **PERÍODO DE VALIDACIÓN**
1990-2007

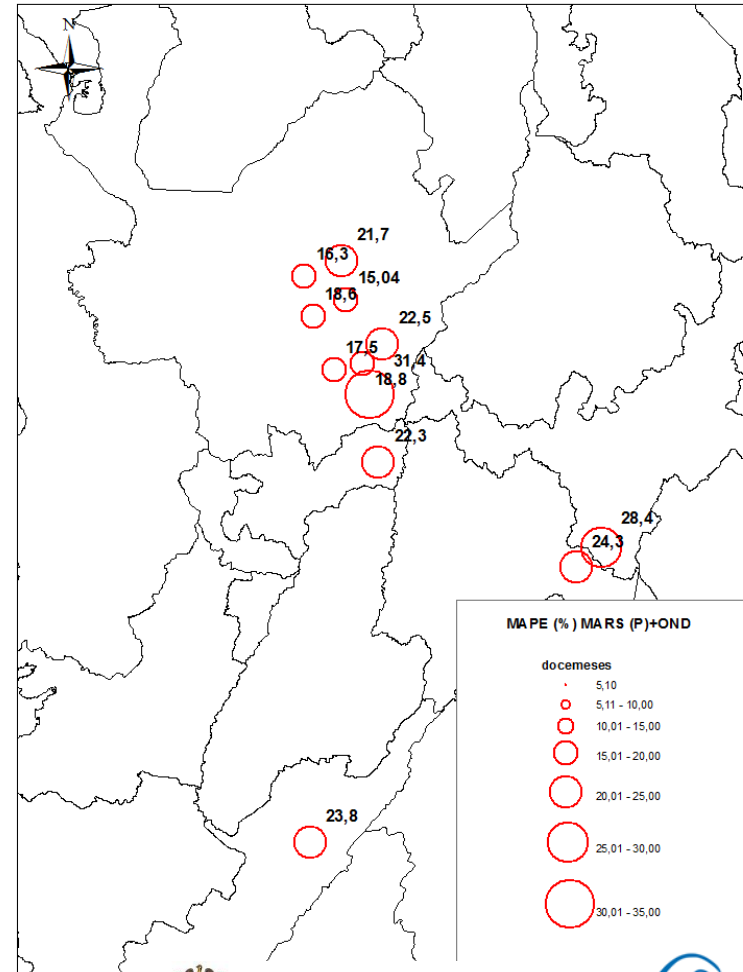


3.10 RESULTADOS

VALIDACIÓN A UN MES



VALIDACIÓN A 12 MESES



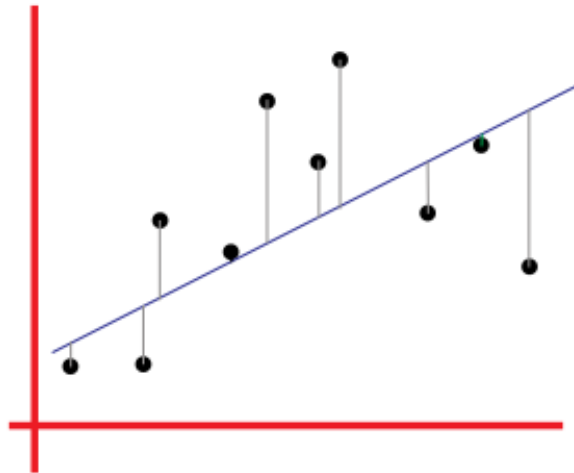
CONTENIDO

1. PROBLEMÁTICA DE LA PREDICCIÓN DE CAUDALES
2. RELACIONES ENTRE LOS CAUDALES DE COLOMBIA Y LA CLIMATOLOGÍA GLOBAL
3. MÉTODOS ACTUALES PARA EL PRONÓSTICO DE CAUDALES
- 4. INCORPORACIÓN DE NUEVAS TÉCNICAS DE PRONÓSTICO**
5. COMPARACIÓN Y COMBINACIÓN DE PRONÓSTICOS HIDROLÓGICOS
6. PREDICCIÓN BASADA EN REGLAS DE JUICIO
7. PROTOTIPO DE SOFTWARE
8. CONCLUSIONES Y RECOMENDACIONES

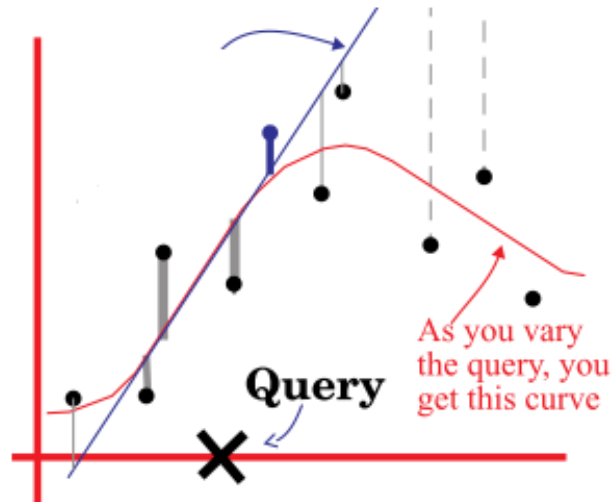


4.1 REGRESIONES LOCALES (I)

Regresion Global



Regresión Local



$$\hat{y}(x) = \beta_0 + \beta_1 x$$

$$\beta_1 = \min \sum_{k=1}^N w(x_{query}, x_k) (y_k - \hat{y}(x_k))^2 = \min \sum_{k=1}^N w(x_{query}, x_k) (y_k - \beta_0 + \beta_1 x_k)^2$$

$$w(x_{query}, x_i) = \exp(-\alpha \mu_i^2)$$

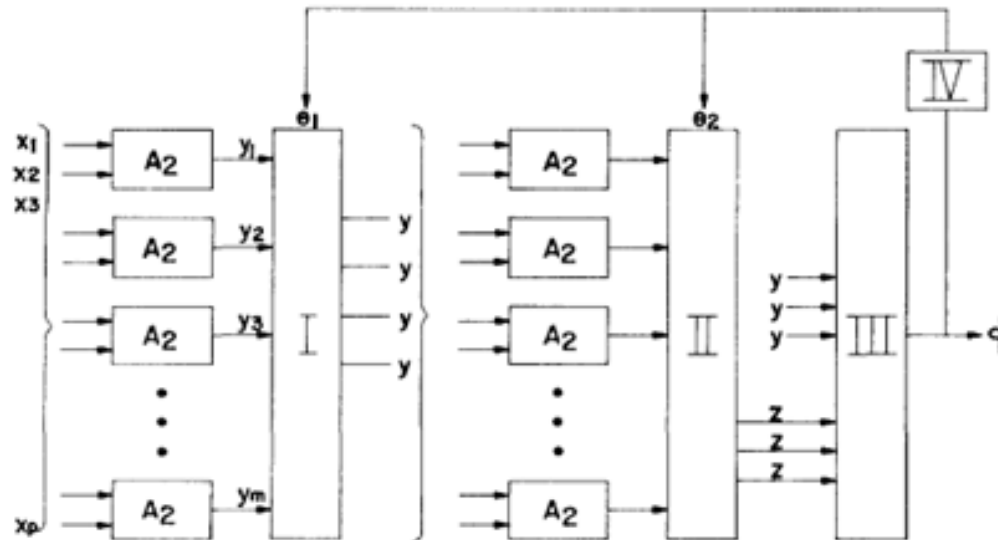
$$\mu_i = \left\| x_{query}, x_{farthest} \right\|$$



4.2 FUNCIONES DE INFLUENCIA RADIAL (I)



4.3 REDES POLINÓMICAS (I)



Algoritmo de agrupación para el tratamiento de datos. I: primer capa de auto-selecciones, II segunda capa de autoselección, III: selección de todas las soluciones, IV: optimización.

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



4.3 REDES POLINÓMICAS (II)

$$E_j < \theta = E^* + \delta$$

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

FUNCIONES DE REGRESIÓN

- **ARIMA**

Auto-regresivos Integrados de Media Móvil

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

- **RLM**

Regresión Lineal Múltiple

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

- **POLPON**

Polinomios ponderados

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

- **RNA**

Redes Neuronales Artificiales

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

- **MARS**

Polinomios Adaptivos de Regresión Multivariada

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

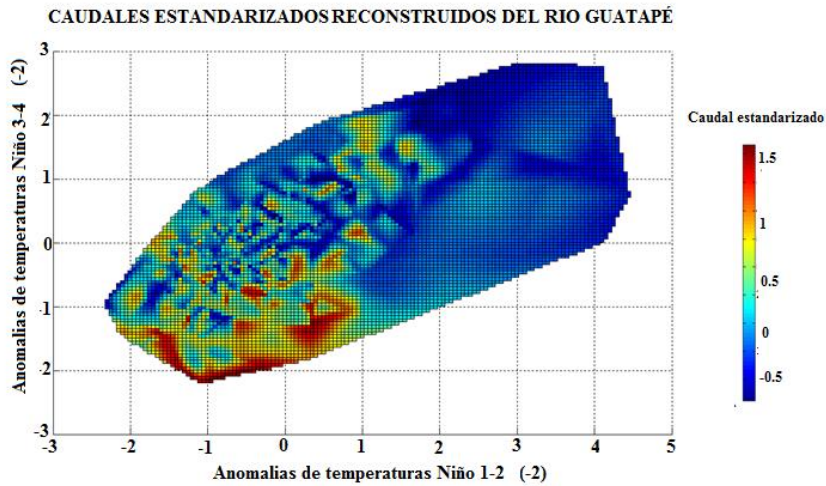
- **RNPOL**

Red Neuronal polinómica

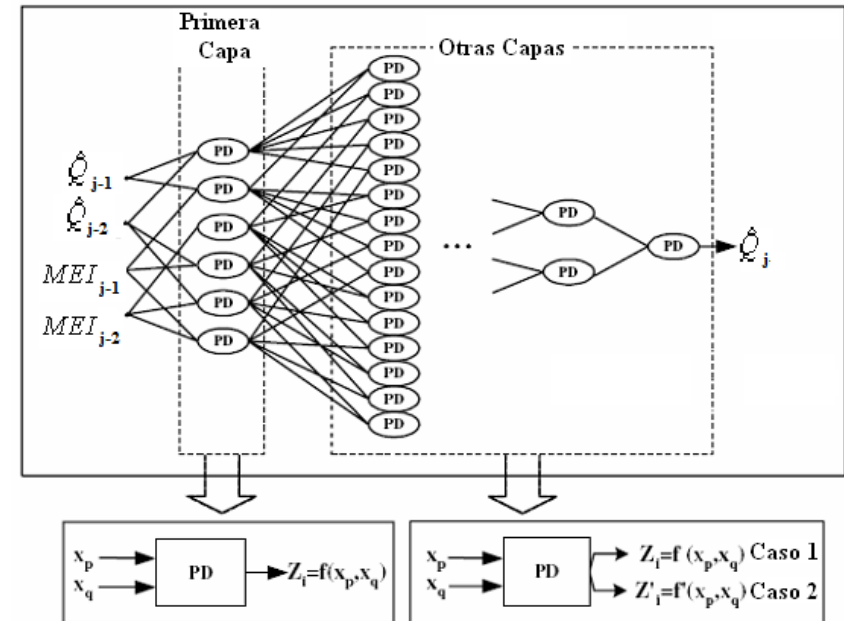
$$\hat{y} = a + \sum_{i=1}^m b_i X_i + \sum_{i=1}^m \sum_{j=1}^m c_{ij} X_i X_j + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^m c_{ijk} X_i X_j X_k + \dots$$

4.4 APLICACIÓN DE NUEVOS MODELOS

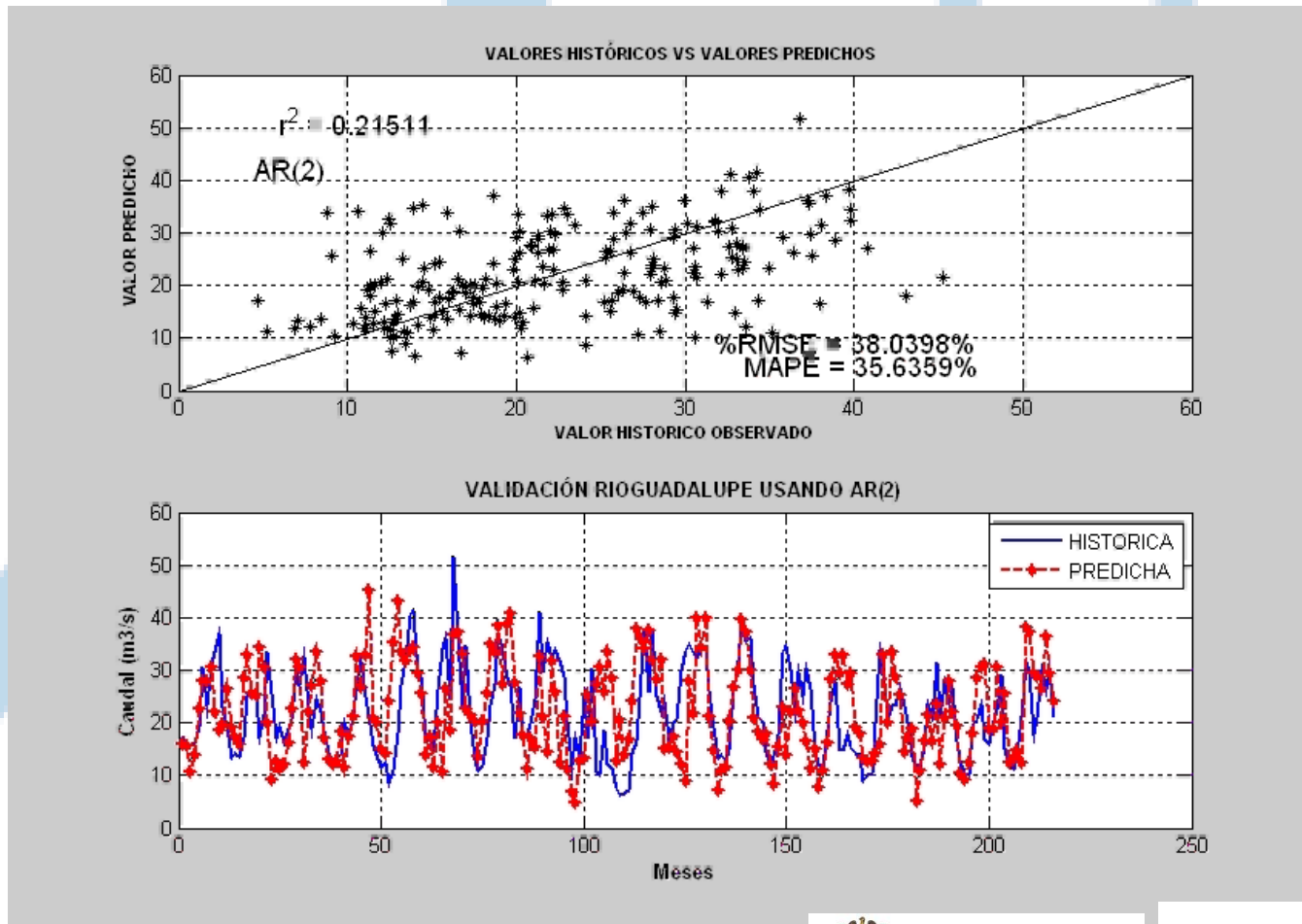
Aplicación de RBF



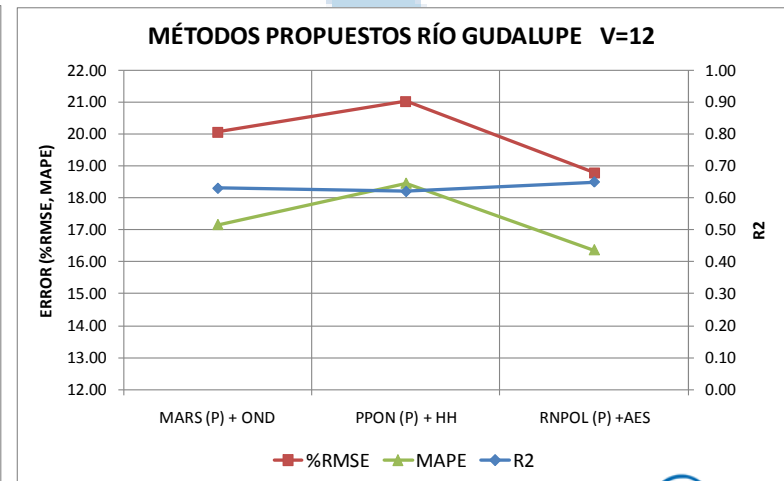
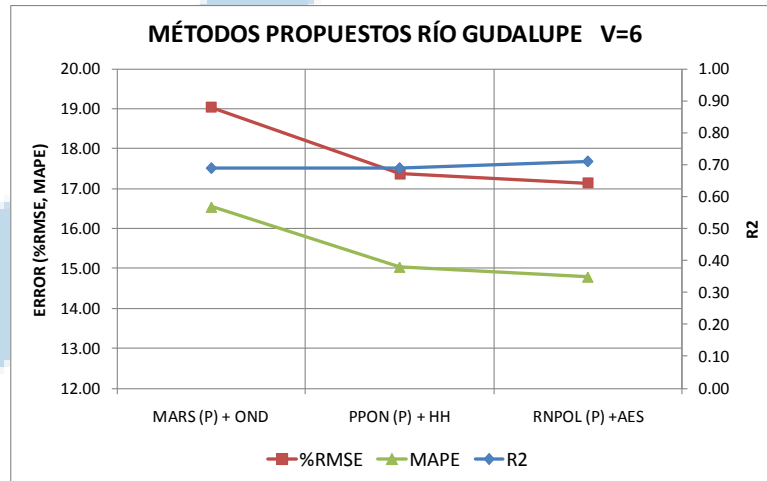
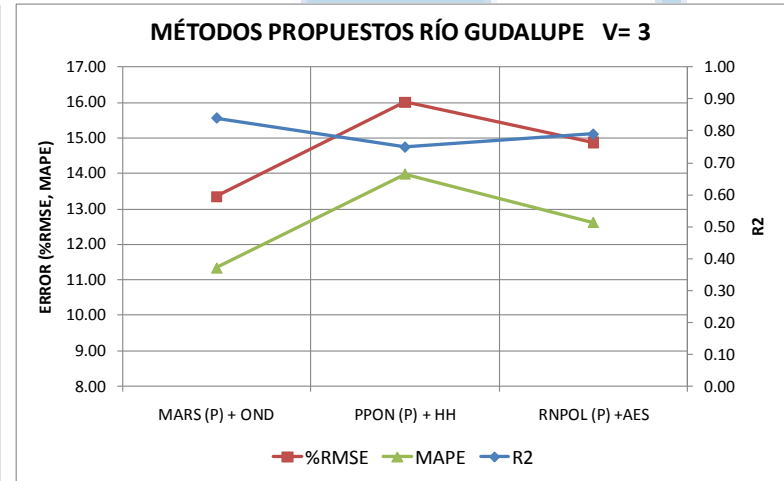
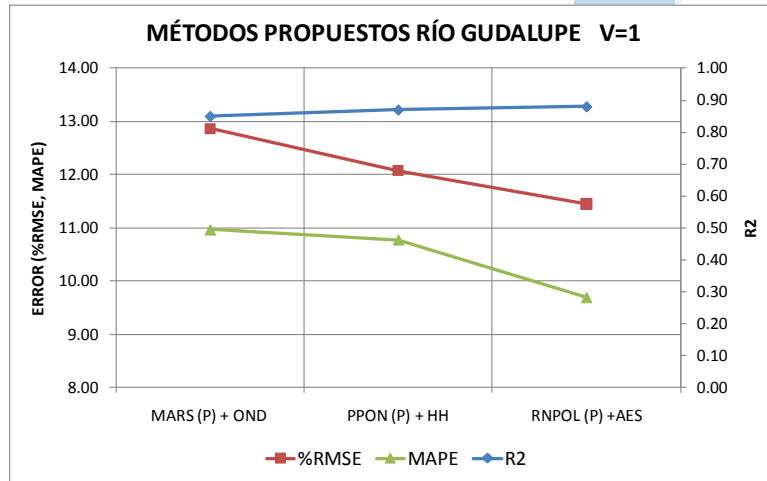
Aplicación de Redes Polinómicas



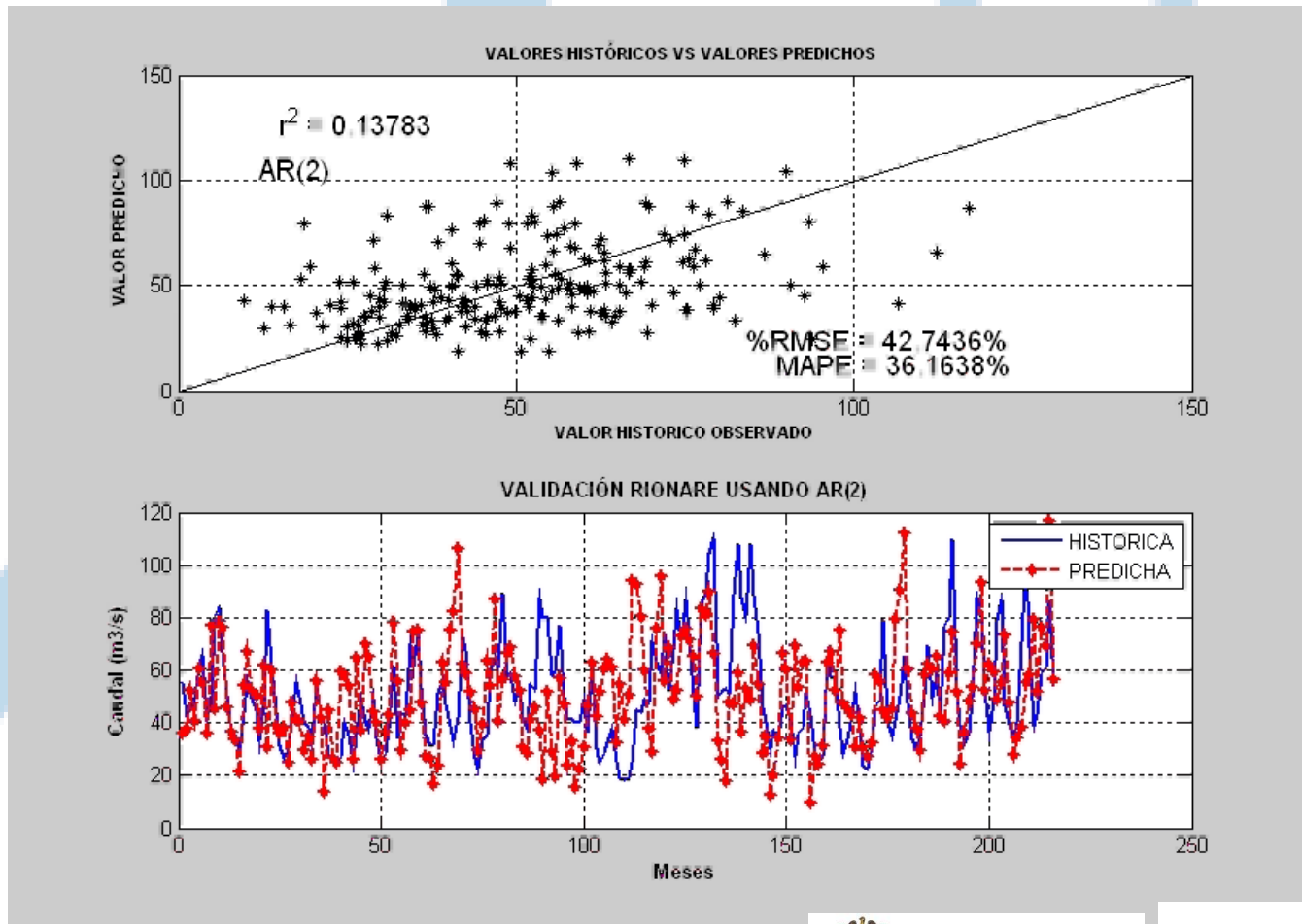
APLICACIÓN AL RÍO GUADALUPE



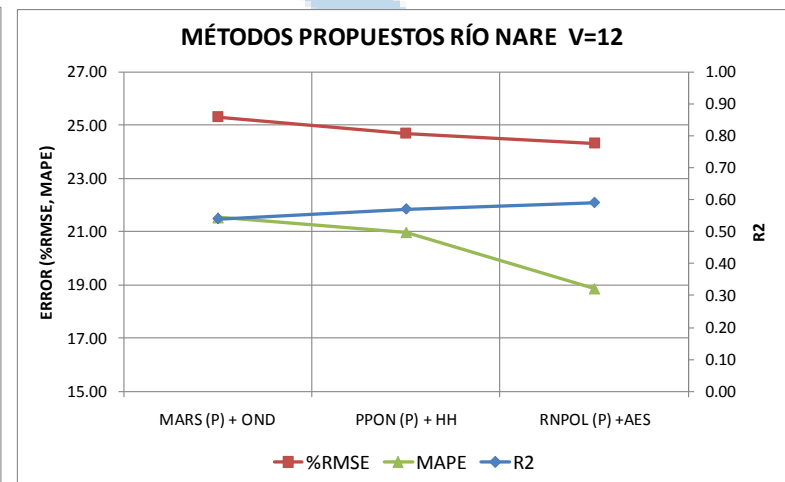
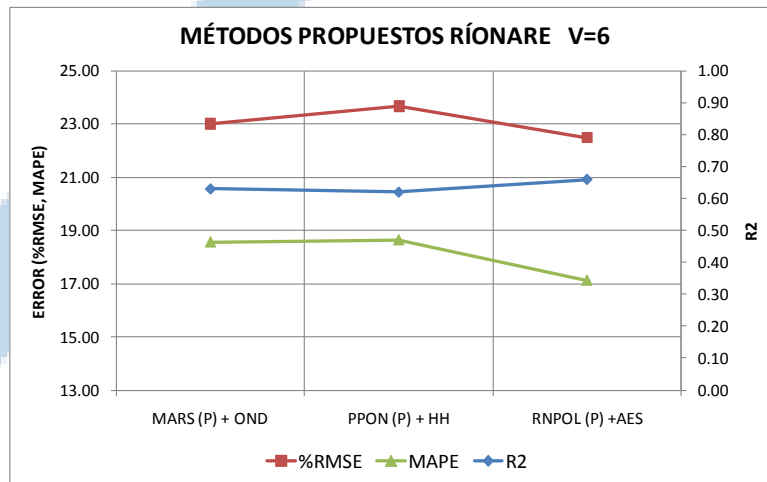
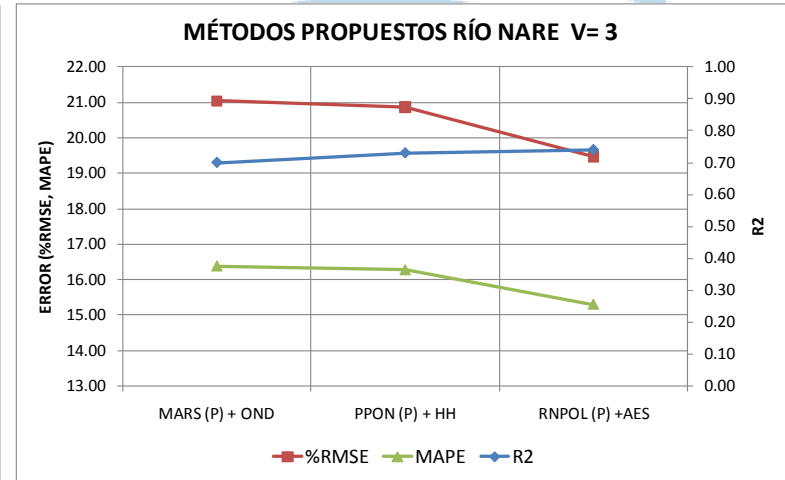
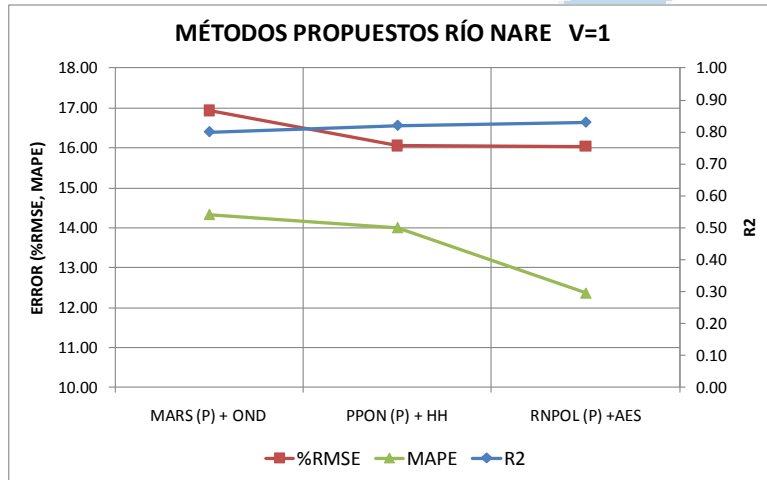
APLICACIÓN AL RÍO GUADALUPE



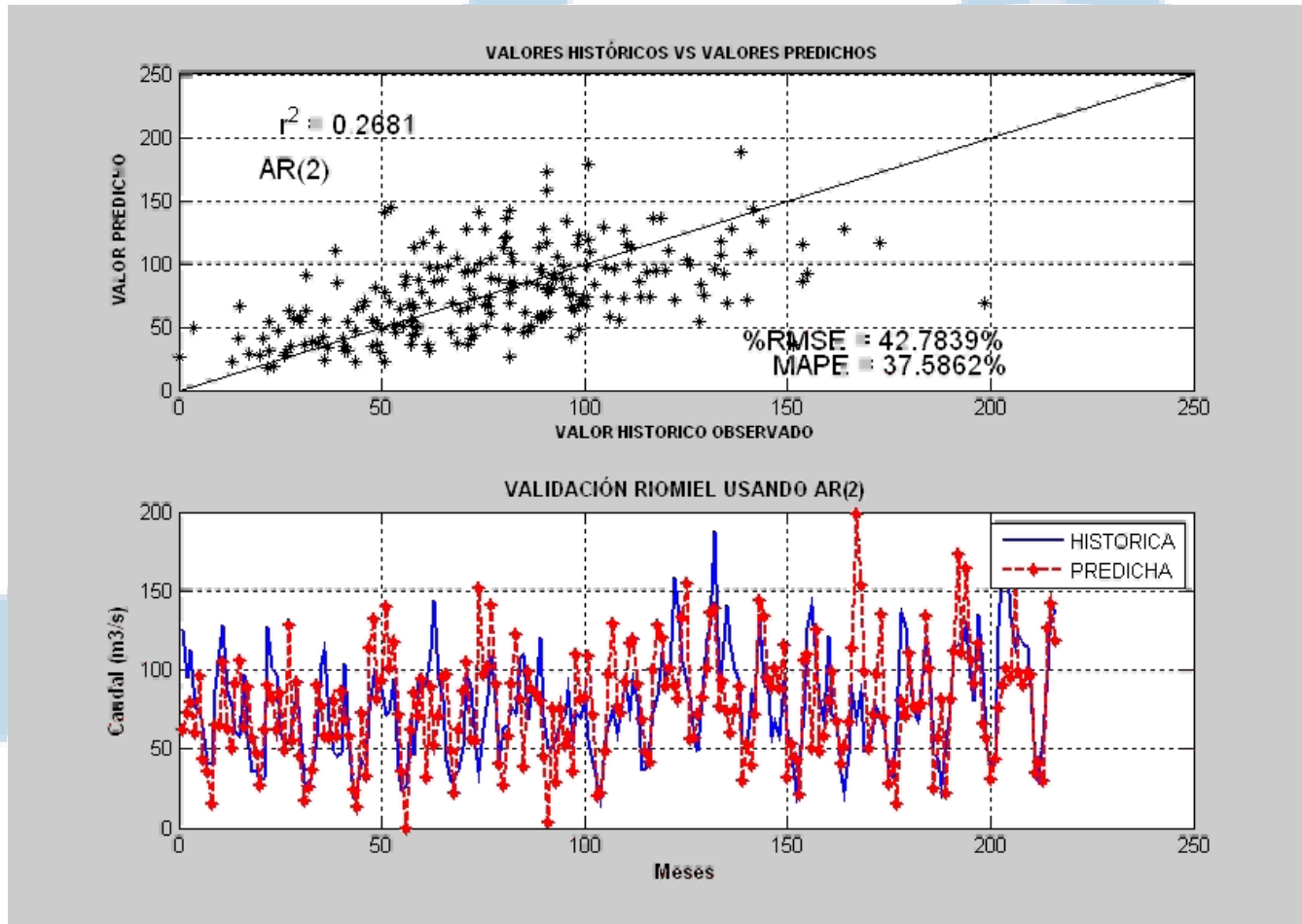
APLICACIÓN AL RÍO NARE



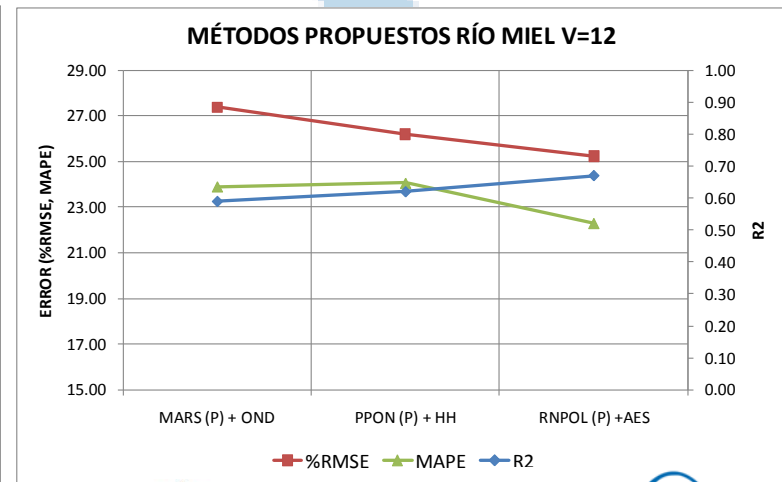
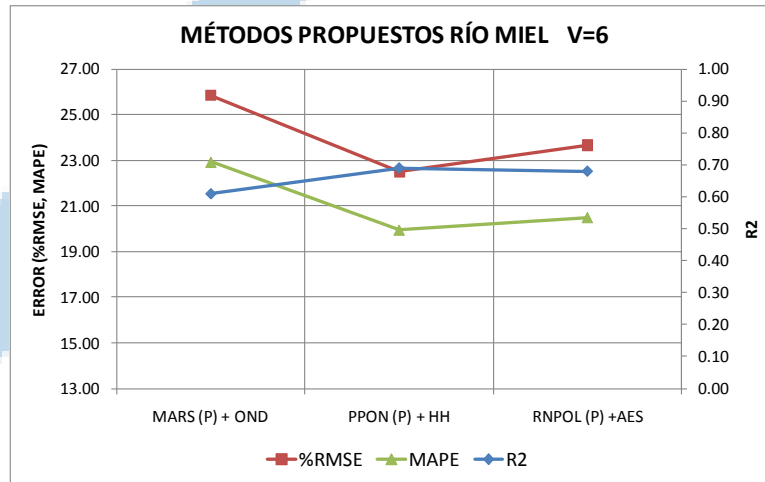
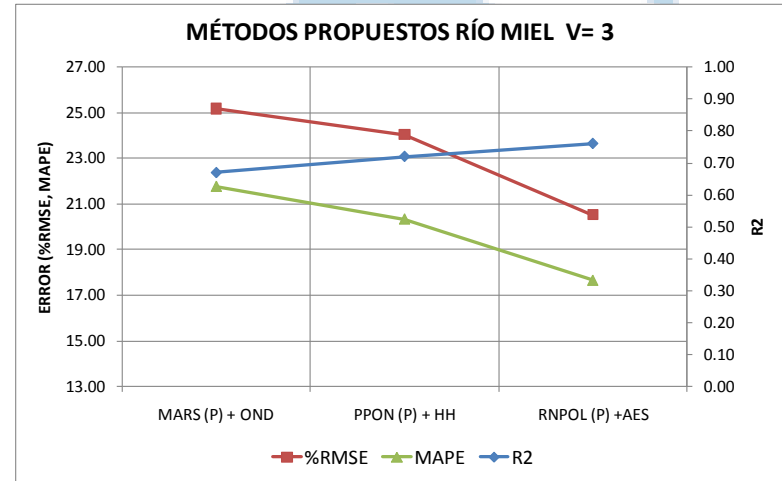
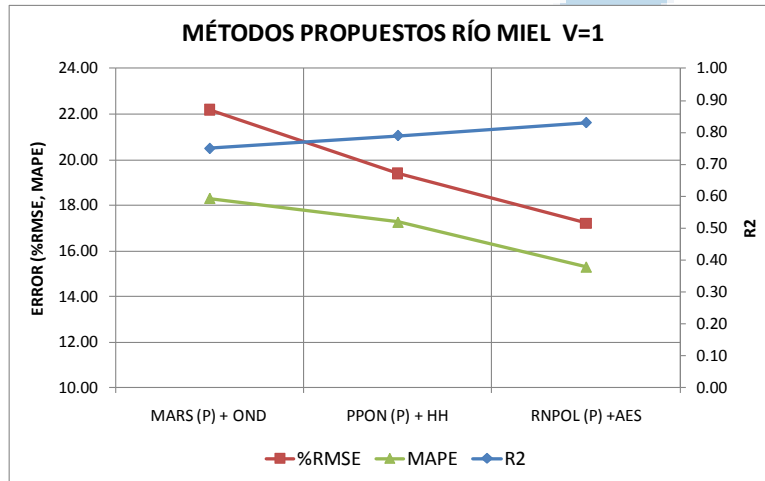
APLICACIÓN AL RÍO NARE



APLICACIÓN AL RÍO MIEL



APLICACIÓN AL RÍO MIEL



CONTENIDO

1. PROBLEMÁTICA DE LA PREDICCIÓN DE CAUDALES
2. RELACIONES ENTRE LOS CAUDALES DE COLOMBIA Y LA CLIMATOLOGÍA GLOBAL
3. METODOS ACTUALES PARA EL PRONÓSTICO DE CAUDALES
4. INCORPORACIÓN DE NUEVAS TÉCNICAS DE PRONÓSTICO
- 5. COMPARACIÓN Y COMBINACIÓN DE PRONÓSTICOS HIDROLÓGICOS**
6. PREDICCIÓN BASADA EN REGLAS DE JUICIO
7. PROTOTIPO DE SOFTWARE
8. CONCLUSIONES Y RECOMENDACIONES



5.1 COMPARACIÓN DE PRONÓSTICOS

Test de Diebold y Mariano

$$\varepsilon_{t+h|t}^1 = y_{t+h} - y_{t+h|t}^1$$

$$\varepsilon_{t+h|t}^2 = y_{t+h} - y_{t+h|t}^2$$

$$H_1 : E[L(\varepsilon_{t+h|t}^1)] \neq E[L(\varepsilon_{t+h|t}^2)]$$

$$L(\varepsilon_{t+h|t}^i) = (\varepsilon_{t+h|t}^i)^2$$

$$d_t = L(\varepsilon_{t+h|t}^1) - L(\varepsilon_{t+h|t}^2)$$

$$H_0 : E[d_t] = 0$$

$$S = \frac{\bar{d}}{[\text{avr}(\bar{d})]^{1/2}} = \frac{\bar{d}}{\left[\frac{LRV_{\bar{d}}}{T} \right]^{1/2}}$$

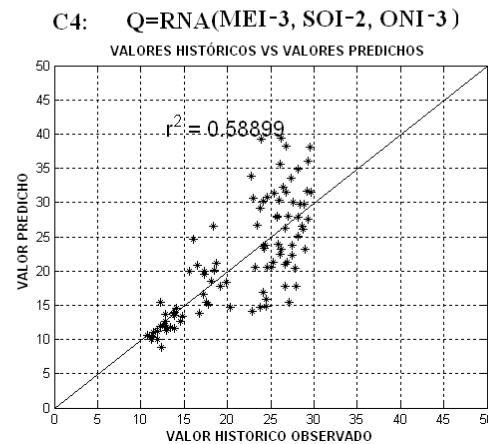
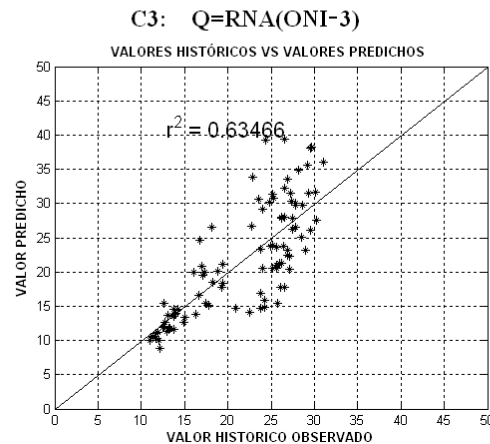
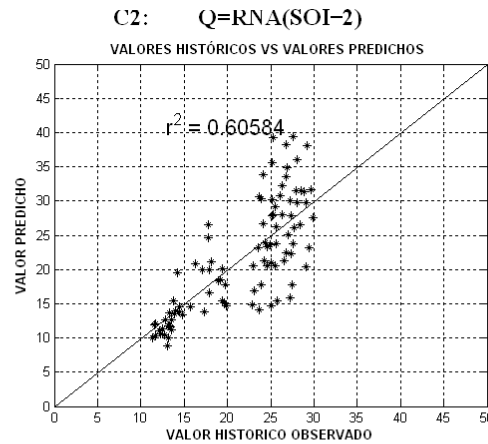
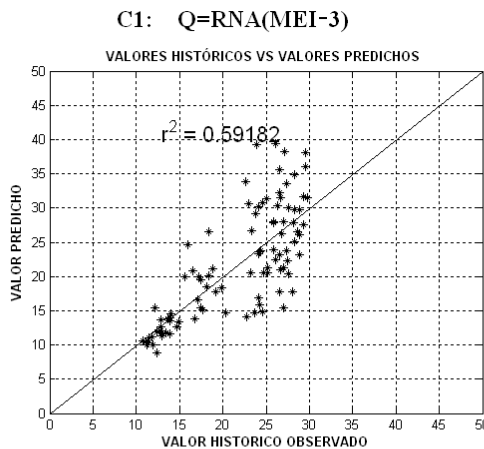
$$\bar{d} = \frac{1}{T} \sum_{t=1}^T d_t$$

$$LRV_{\bar{d}} = \gamma_0 + 2 \sum_{j=1}^{\infty} \gamma_j$$

$$|S| > 1.96$$



5.2 APLICACIONES DEL TEST DE DIEBOLD Y MARIANO (I)



Caso del río Guadalupe

	$ S > 1.96$			
	Q=RNA(MEI-3) C1	Q=RNA(SOI-2) C2	Q=RNA(ONI-3) C3	Q=RNA(ENSO) C4
Q=RNA(MEI-3) C1		0.6329	2.5991	1.4292
Q=RNA(SOI-2) C2	0.6329		1.6103	0.7345
Q=RNA(ONI-3) C3	2.5991	1.6103		2.9441
Q=RNA(ENSO) C4	1.4292	0.7345	2.9441	



5.2 APLICACIONES DEL TEST DE DIEBOLD Y MARIANO (II)

Caso del río Guadalupe

COMPARACIÓN DE MODELOS PREDICCIÓN A 12 MESES CAUDALES DEL RIO GUADALUPE							
$ S > 1.96$							
	AR(2)	RLM	RED (P)+HH	RLM (P) +AES	MARS (P) +OND	PPON (P)+ HH	RNPOL (P)+AES
AR(2)	NaN	4.2139	7.831	9.5291	8.861	7.6485	8.1394
RLM	4.2139	NaN	11.1204	19.2553	17.9237	10.3692	13.9627
RED (P)+HH	7.831	11.1204	NaN	2.5169	12.0237	0.3614	18.1542
RLM (P) +AES	9.5291	19.2553	2.5169	NaN	6.3334	3.4689	3.5708
MARS (P) +OND	8.861	17.9237	12.0237	6.3334	NaN	23.5782	1.5379
PPON (P)+ HH	7.6485	10.3692	0.3614	3.4689	23.5782	NaN	20.6276
RNPOL (P)+AES	8.1394	13.9627	18.1542	3.5708	1.5379	20.6276	NaN



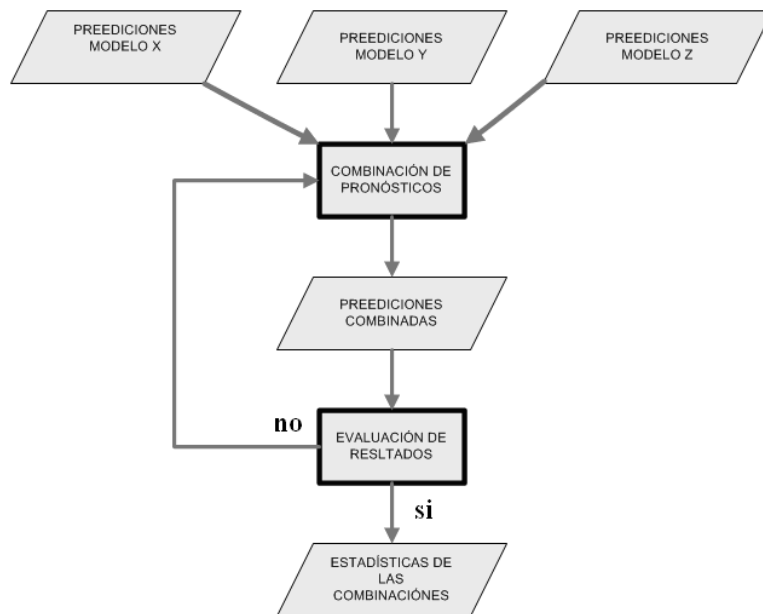
5.3 COMBINACIÓN DE PRONÓSTICOS (I)

- El investigador cuenta con varios pronósticos de caudales, *generalmente diferentes*.
- **PREGUNTA:** ¿Se debería escoger un sólo pronóstico o tomar de alguna forma un promedio de los resultados arrojados por los diferentes modelos?
- En todas las investigaciones realizadas se concluye que es mejor usar una mezcla de los pronósticos en lugar de usar sólo uno de ellos (Castaño y Melo, 2006).



5.3 COMBINACIÓN DE PRONÓSTICOS (II)

Esquema general de combinación



Promedio simple y ponderado

$$\hat{Q}c_i = \sum_{j=1}^N a_j \hat{Q}_{j,i} + e_i$$

$$Q = PA + E$$

P : Matriz de pronósticos

$$A = [a_1, a_2, a_3, \dots, a_{N-1}, a_N]^T$$

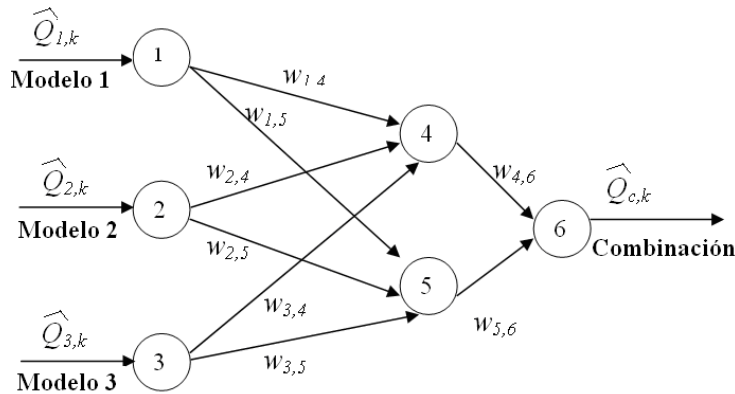
$$Q = [Q_1, Q_2, Q_3, \dots, Q_{k-1}, Q_k]^T$$

$$\hat{A} = [(P^T P)]^{-1} P^T Q$$



5.3 COMBINACIÓN DE PRONÓSTICOS (III)

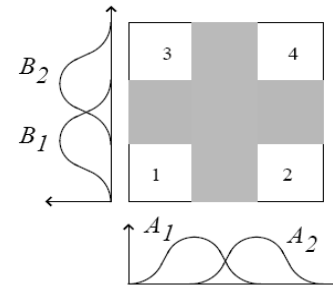
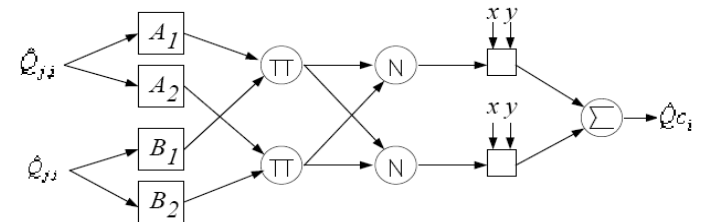
Redes Neuronales



$$\hat{Q}_{c_i} = \sum_{h=1}^m w_m \cdot f \left(\sum_{i=1}^n w_{m,i} \cdot \hat{Q}_{m,i} \right)$$

$$f(x) = \frac{2}{(1 + e^{-2x})} - 1$$

ANFIS



$$SI \hat{Q}_{j,i} \in A_1 \text{ ENTONCES } \hat{Q}_{c_i} = a^{(1)}_0 + \sum_{j=1}^N a^{(1)}_j \hat{Q}_{j,i} + e_i$$

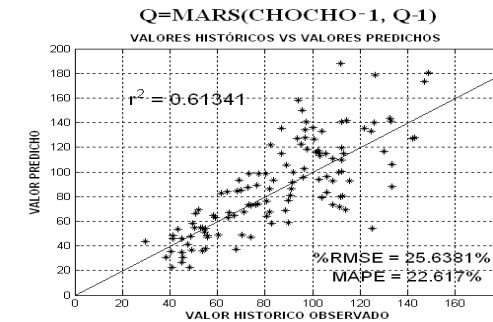
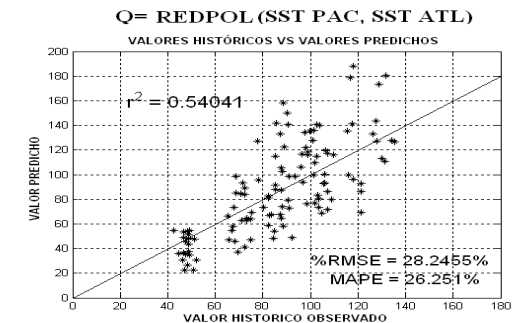
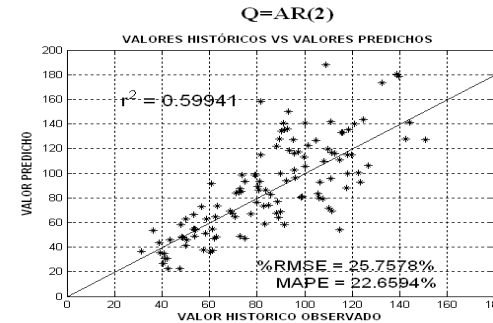
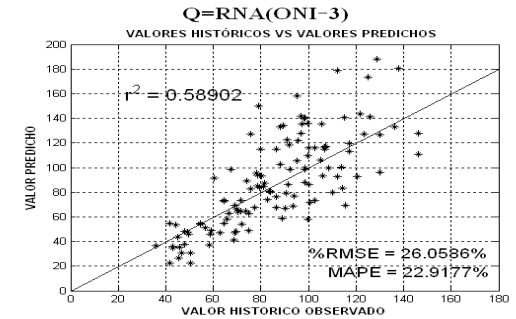
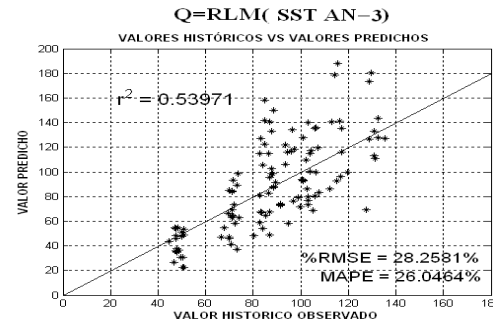
$$SI \hat{Q}_{j,i} \in A_2 \text{ ENTONCES } \hat{Q}_{c_i} = a^{(2)}_0 + \sum_{j=1}^N a^{(2)}_j \hat{Q}_{j,i} + e_i$$



5.3 APLICACIONES DE LA COMBINACION DE PRONÓSTICOS

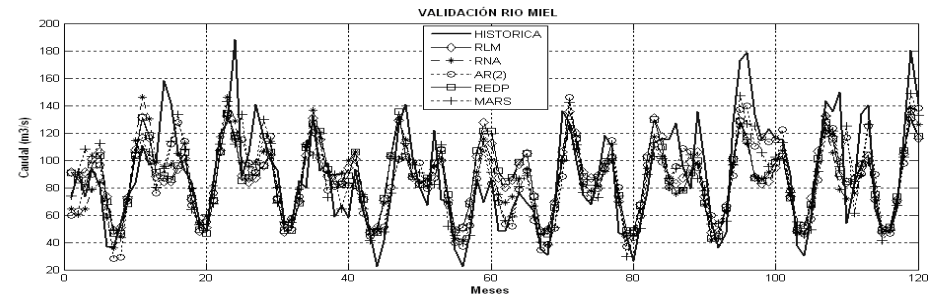
Calibración: 1990-1999
Validación: 2000-2007

PREDICCIÓN DE CAUDALES MEDIOS MENSUALES RIO MIEL



Matriz de comparación

	RLM	RNA	AR(2)	REDPOL	MARS
RLM	NaN	2.0091	3.7251	6.154	2.7977
RNA	2.0091	NaN	4.3045	4.3827	2.022
AR(2)	3.7251	4.3045	NaN	7.9769	1.9972
REDPOL	6.154	4.3827	7.9769	NaN	4.4545
MARS	2.7977	2.022	1.9972	4.4545	NaN

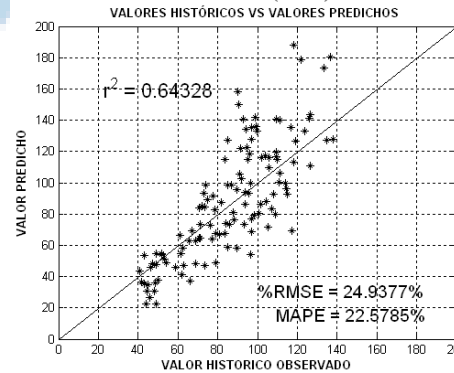


5.3 APLICACIONES DE LA COMBINACION DE PRONÓSTICOS

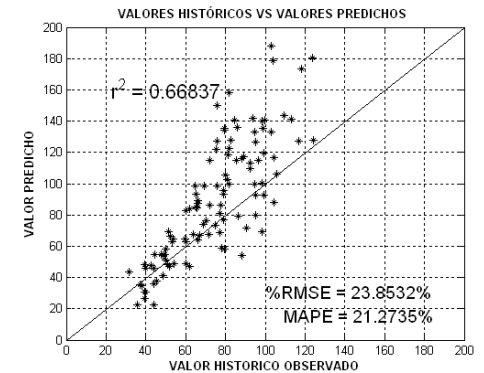
Calibración: 1990-1999
Validación: 2000-2007

COMBINACIÓN DE PRONÓSTICOS DE LOS CAUDALES MEDIOS MENSUALES DEL RIO LA MIEL

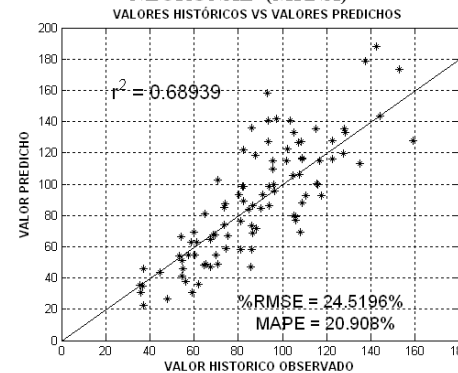
COMBINACIÓN MÉTODO DEL PROMEDIO SIMPLE (MPS)



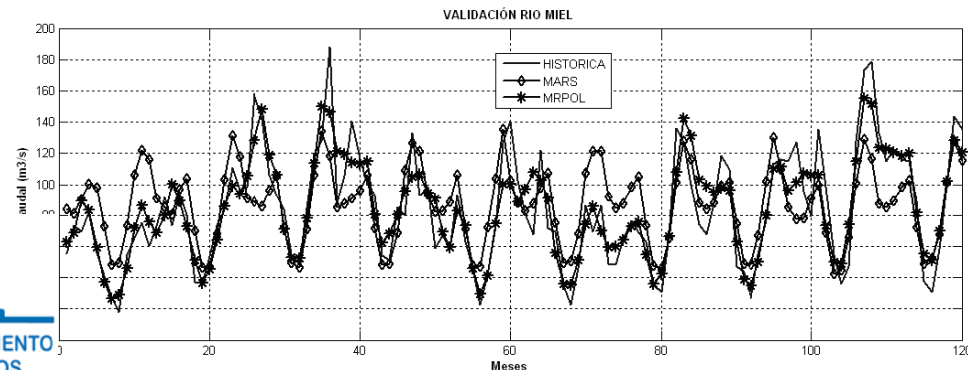
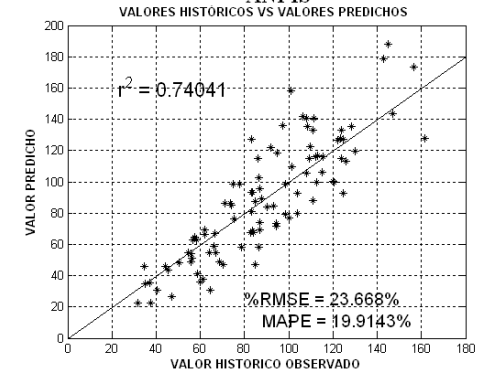
COMBINACIÓN MÉTODO DEL PROMEDIO PONDERADO (MPP)



COMBINACIÓN MÉTODO DE LA RED NEURONAL (MRNA)

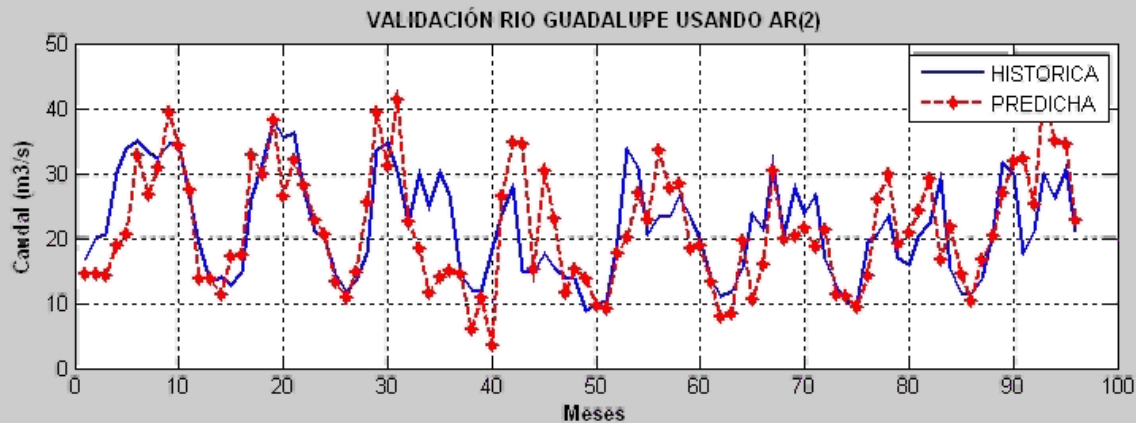
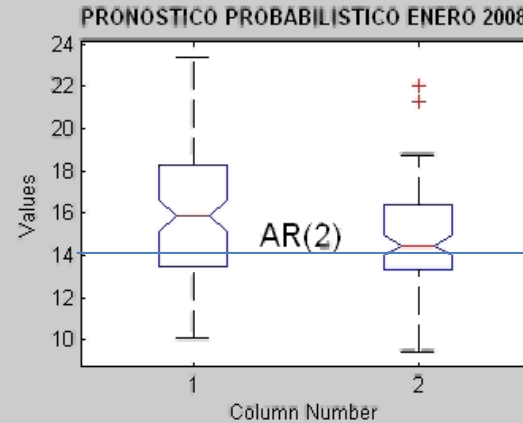
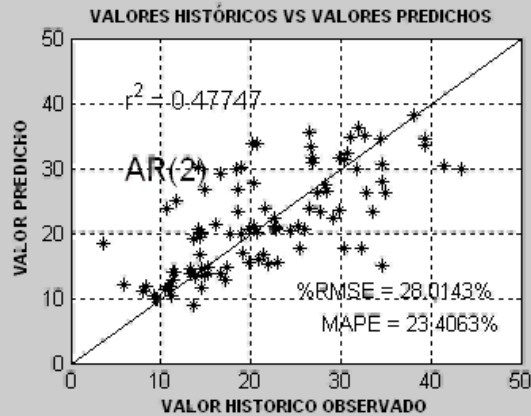


COMBINACIÓN DE PRONÓSTICOS CON ANFIS



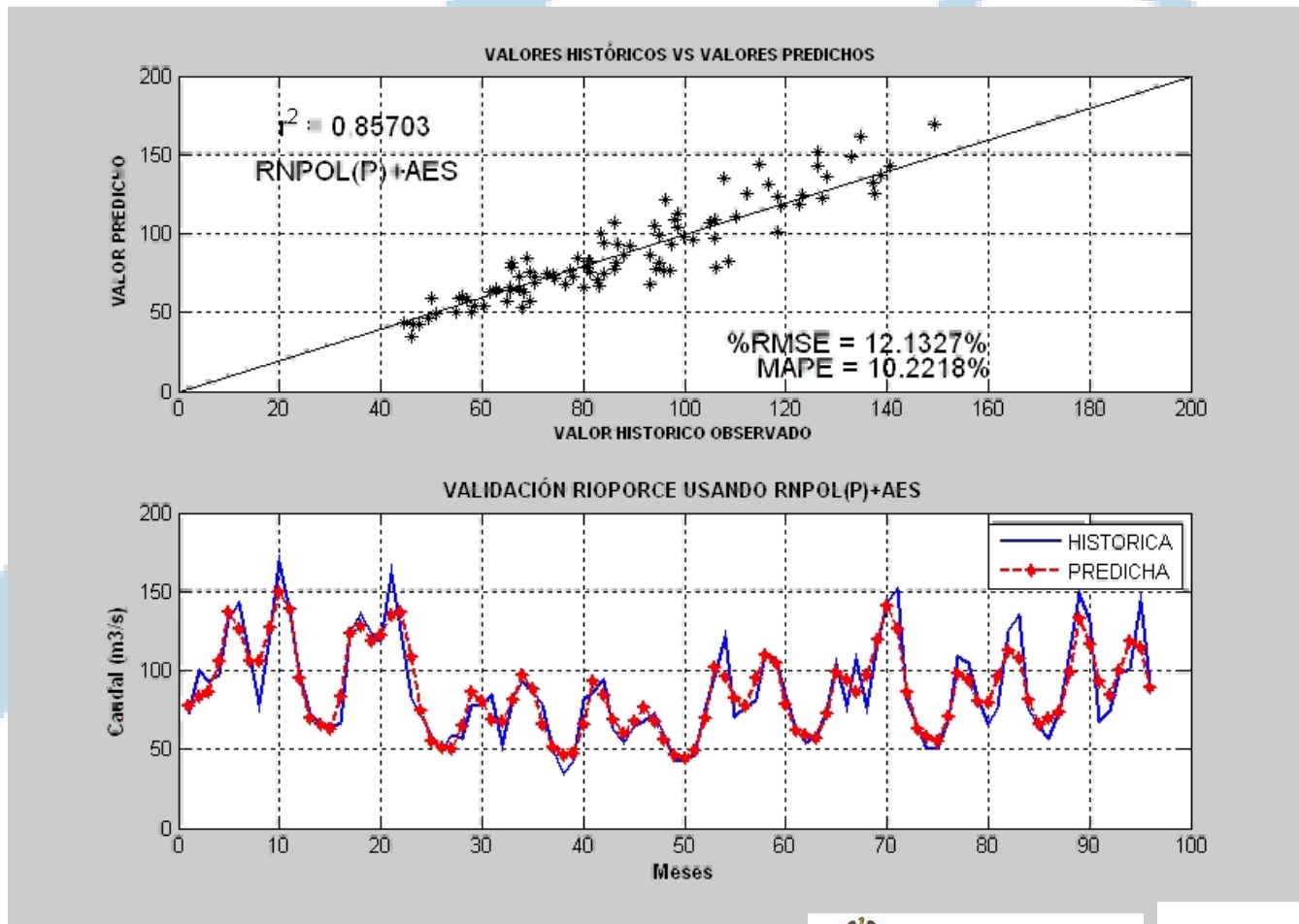
Calibración: 1990-1999
Validación: 2000-2007

Río Guadalupe



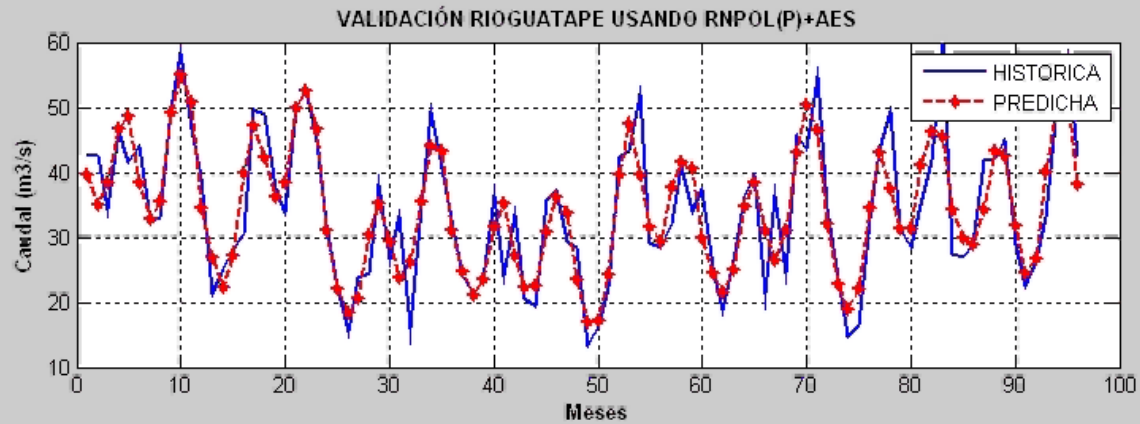
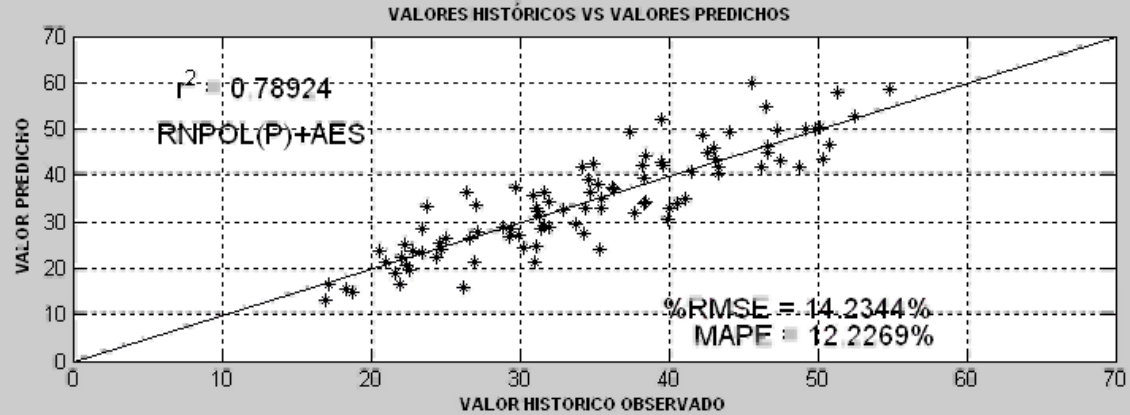
Calibración: 1990-1999
Validación: 2000-2007

Río Porce



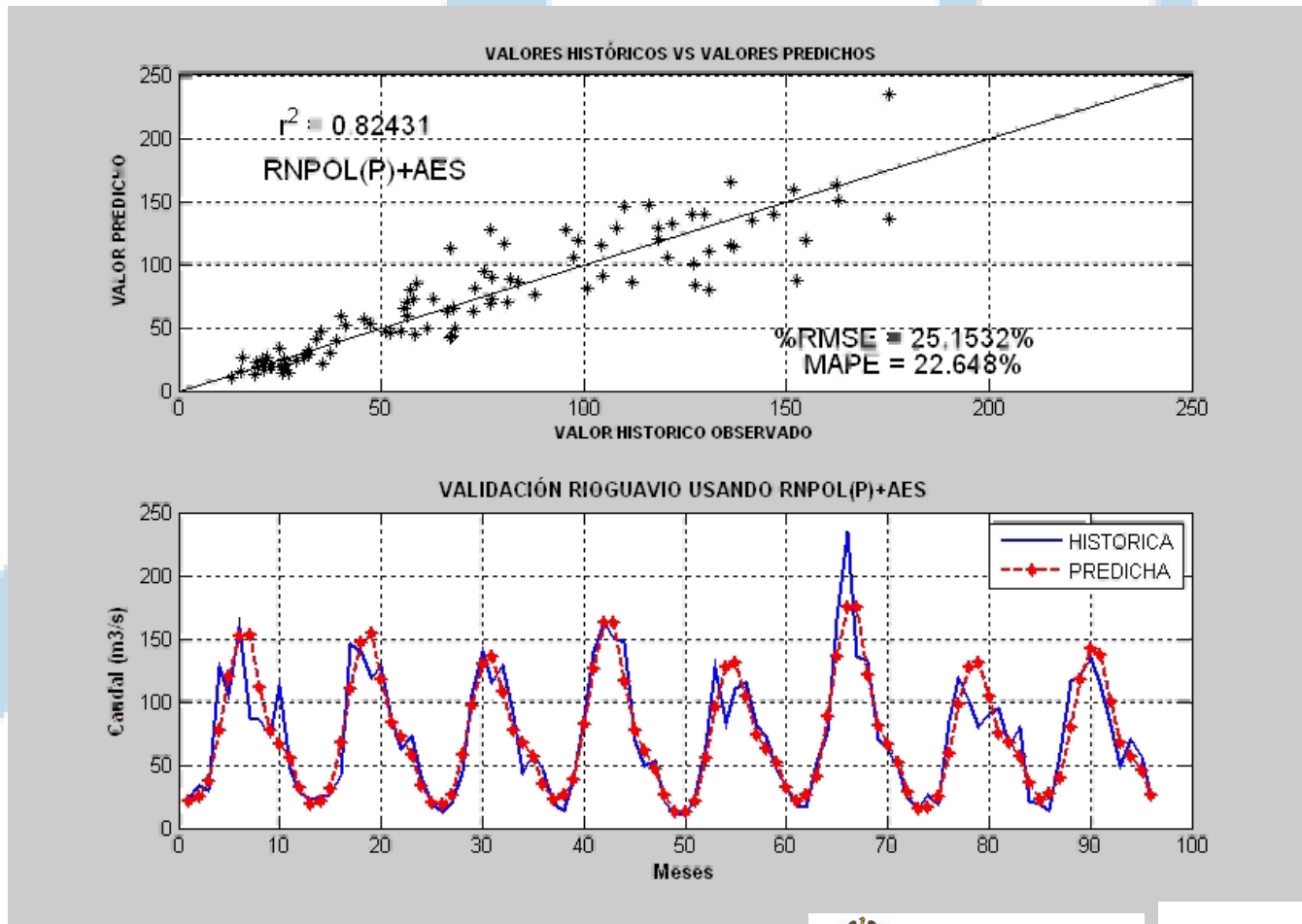
Calibración: 1990-1999
Validación: 2000-2007

Río Guatapé



Calibración: 1990-1999
Validación: 2000-2007

Río Guavio

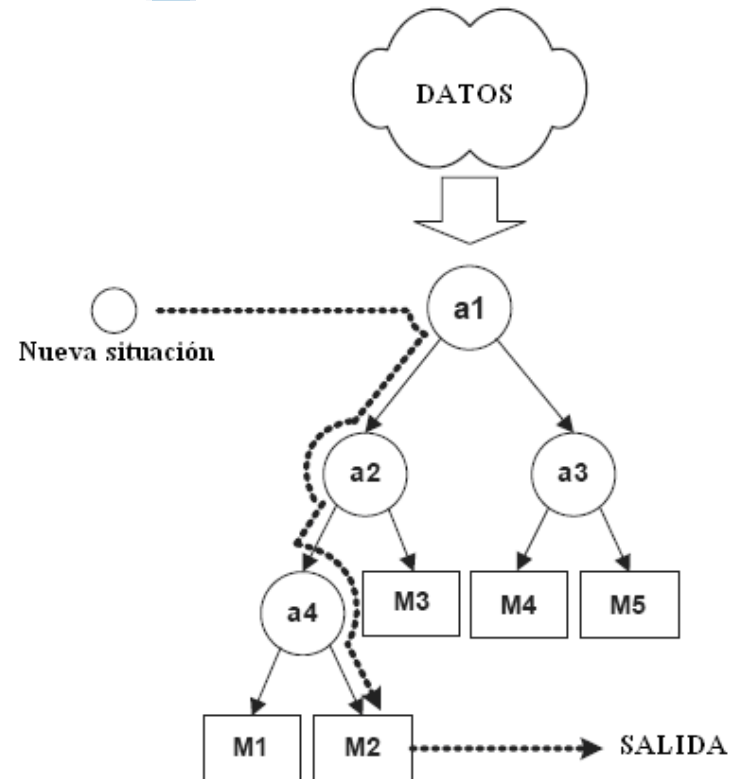
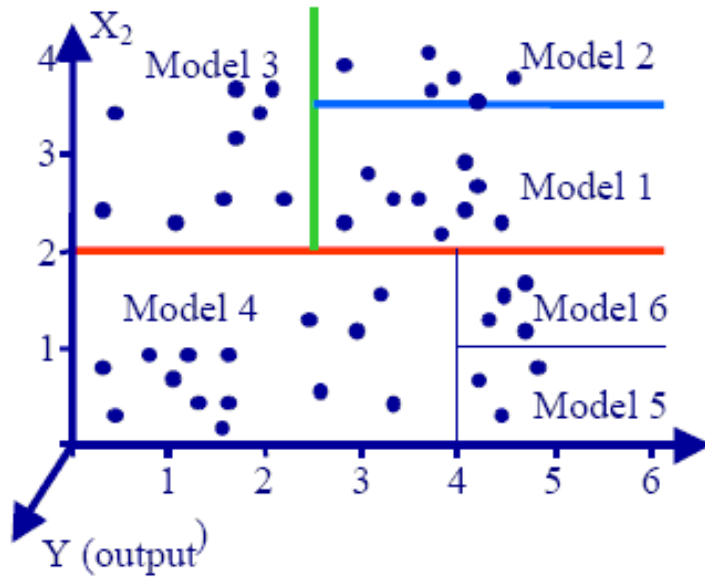


CONTENIDO

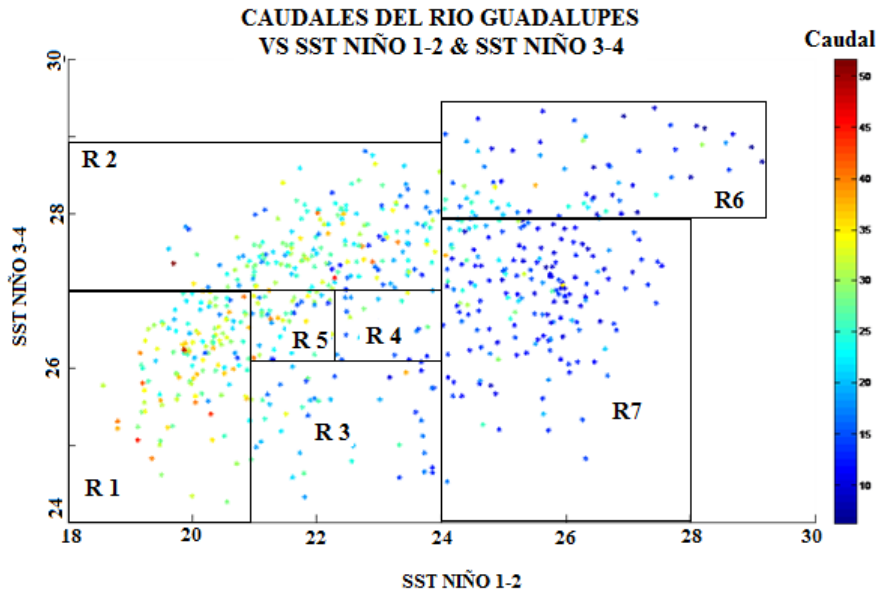
1. PROBLEMÁTICA DE LA PREDICCIÓN DE CAUDALES
2. RELACIONES ENTRE LOS CAUDALES DE COLOMBIA Y LA CLIMATOLOGÍA GLOBAL
3. MÉTODOS ACTUALES PARA EL PRONÓSTICO DE CAUDALES
4. INCORPORACIÓN DE NUEVAS TÉCNICAS DE PRONÓSTICO
5. COMPARACIÓN Y COMBINACIÓN DE PRONÓSTICOS HIDROLÓGICOS
- 6. PREDICCIÓN BASADA EN REGLAS DE JUICIO**
7. PROTOTIPO DE SOFTWARE
8. CONCLUSIONES Y RECOMENDACIONES



6.1 ÁRBOLES DE REGRESIÓN- DECISIÓN



6.2 ÁRBOLES DE REGRESIÓN- DECISIÓN



Regla 1:

Si $SST N1-2 \in [18;21]$

\wedge Si $SST N3-4 \in [24;26.5]$

Entonces $Q = f1(x)$

End

End

Regla 2:

Si $SST N1-2 \in [18;24]$

\wedge Si $SST N3-4 \in [26.5;28.5]$

Entonces $Q = f2(x)$

End

$R1: SST N1-2 \in [18;21] \wedge SST N3-4 \in [24;26.5]$

$R2: SST N1-2 \in [18;24] \wedge SST N3-4 \in [26.5;28.5]$

$R7: SST N1-2 \in [24;28] \wedge SST N3-4 \in [26.5;28.5]$



6.3 ALGORITMO M5

Para agrupar los datos el algoritmo M5 usa método heurístico, conocido como análisis clúster por reducción de varianza, que busca minimizar la variación interna de los valores de cada subconjunto eligiendo aquel atributo que maximice la reducción del error de acuerdo a la siguiente fórmula Quinlan (1992),:

$$SDR = sd(T) - \sum_i \frac{|T_i|}{|T|} sd(T_i)$$

Donde T corresponde a la colección de datos en el nodo a dividir, T_i el conjunto de datos correspondiente al atributo i considerando en la división de T , y el operador $sd()$ estima la desviación típica de los datos.



6.4 VARIABLES DEL ALGORITMO M5

- **Variables cualitativas**

Son aquellas variables cuyos valores son un conjunto de cualidades no numéricas a las que se llama categorías o modalidades.

Ejm: El Niño, La Niña, Normal

Ene, Feb, Dic...

- **Variables Binarias**

Son variables que adoptan valores 0, 1 (si, no; existe, no existe) y se usan para resolver problemas del tipo inclusión-exclusión.

Ejm: llueve-no llueve; verano-invierno.

- **Variables Cuantitativas**

Variable discreta: Es la variable que presenta separaciones o interrupciones en la escala de valores que puede tomar.:

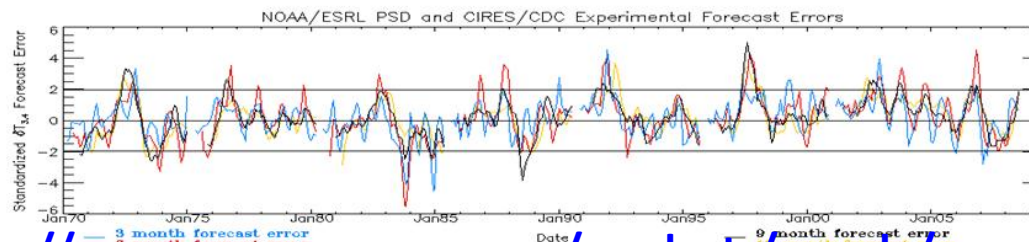
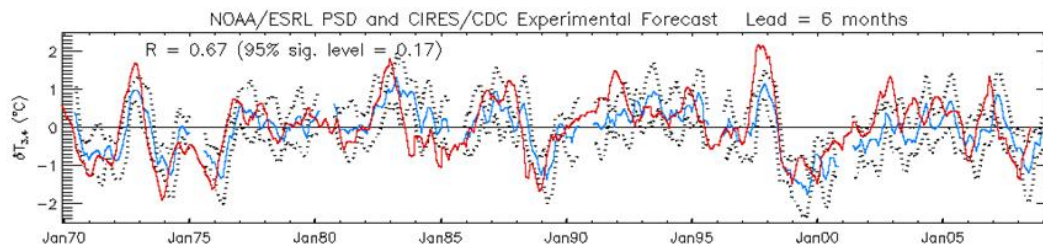
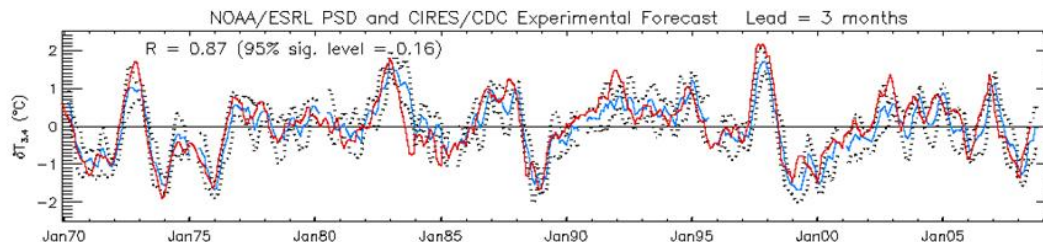
Ejm: Número de huracanes,

Variable continua: Es la variable que puede adquirir cualquier valor dentro de un intervalo especificado de valores.

Ejm: los caudales de un río (210 ,25 , 6 ,52) m³/s o el valor de las temperaturas en el océano (10.3 , 0.5 , 23.8) °C,



6.5 CONSTRUCCIÓN DE ÁRBOLES DE DECISIÓN USANDO EL ALGORITMO M5



VARIABLE
CUANTITATIVA
CONTINUA

http://www.cpc.ncep.noaa.gov/products/people/yxue/SL_forecast_clim71-00_godas.html

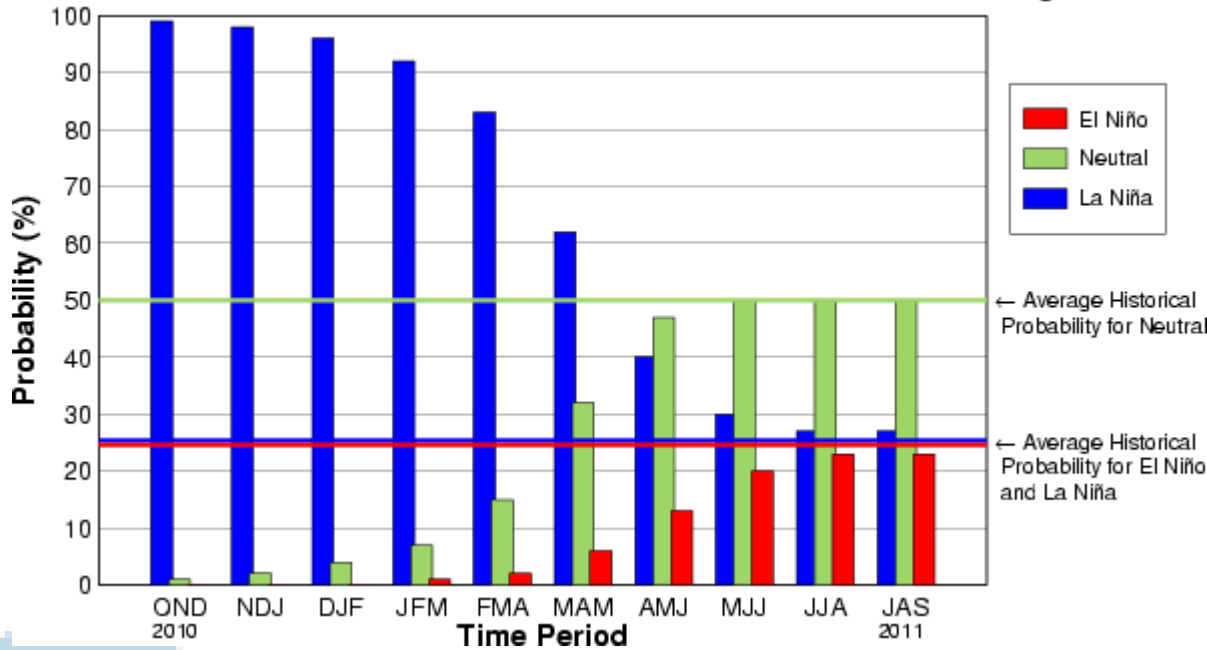


UNIVERSIDAD
NACIONAL
DE COLOMBIA
SEDE MEDELLÍN

POSGRADO EN APROVECHAMIENTO
DE RECURSOS HIDRÁULICOS

6.6 CONSTRUCCIÓN DE ÁRBOLES DE DECISIÓN USANDO EL ALGORITMO M5

IRI Probabilistic ENSO Forecast for NINO3.4 Region



VARIABLE CUALITATIVA

$$ENSO = \begin{cases} NIÑO = 0 \\ NORMAL = 1 \\ NIÑA = 2 \end{cases}$$

<http://portal.iri.columbia.edu/portal/server.pt>

PRONÓSTICOS DEL ENSO (IRI)

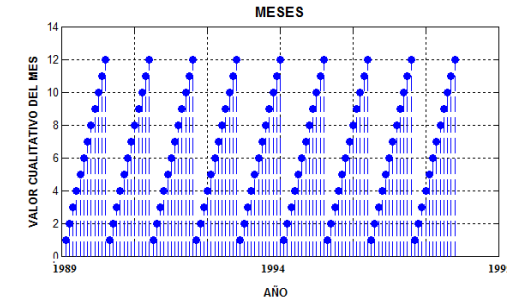
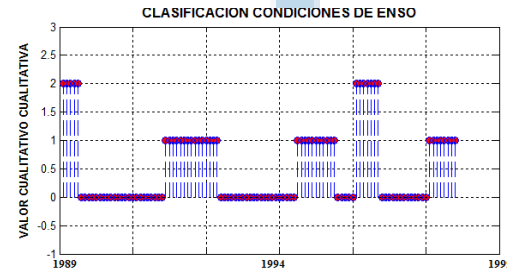
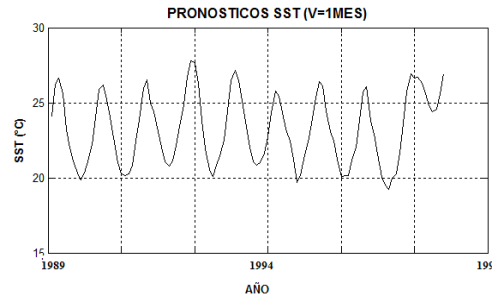
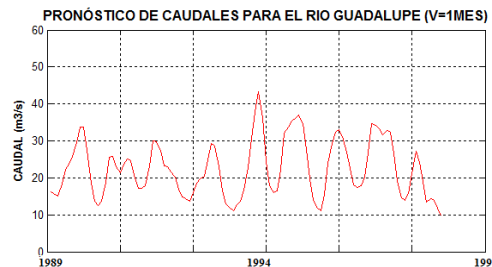
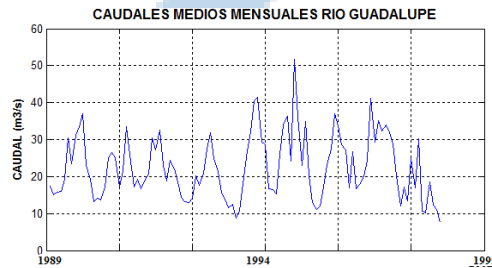


UNIVERSIDAD NACIONAL DE COLOMBIA
SEDE MEDELLÍN



6.7 CONSTRUCCIÓN DE ÁRBOLES DE DECISIÓN USANDO EL ALGORITMO M5

$$Q_{\text{corregido}} = f(Q_{\text{predicho}}, SST, ENSO, \text{mes})$$



$$ENSO = \begin{cases} NIÑO = 0 \\ NIÑA = 1 \\ NORMAL = 2 \end{cases}$$

$$Q_{\text{corregido}} = f(Q_{\text{predicho}}, SST, ENSO, \text{mes})$$



6.7 RESULTADOS ALGORITMO M5

Synthetic variables:

$z1 = x1$

$z2 = 1$ if $x2$ is in {3, 1, 4, 12, 11, 5, 8, 7, 10, 6, 9} else = 0

$z3 = 1$ if $x2$ is in {1, 4, 12, 11, 5, 8, 7, 10, 6, 9} else = 0

$z4 = 1$ if $x2$ is in {4, 12, 11, 5, 8, 7, 10, 6, 9} else = 0

$z5 = 1$ if $x2$ is in {12, 11, 5, 8, 7, 10, 6, 9} else = 0

$z6 = 1$ if $x2$ is in {11, 5, 8, 7, 10, 6, 9} else = 0

$z7 = 1$ if $x2$ is in {5, 8, 7, 10, 6, 9} else = 0

$z8 = 1$ if $x2$ is in {8, 7, 10, 6, 9} else = 0

$z9 = 1$ if $x2$ is in {7, 10, 6, 9} else = 0

$z10 = 1$ if $x2$ is in {10, 6, 9} else = 0

$z11 = 1$ if $x2$ is in {6, 9} else = 0

$z12 = 1$ if $x2$ is in {9} else = 0

if $z1 \leq 15.9826$

$y = 15.2$ (7)

else

$y = 17$ (22)

else

$y = 20.2 - 3.28 * z9$ (29)

else

if $z1 \leq 29.8819$

if $z1 \leq 24.1946$

if $z15 = 0$

if $z13 \leq 24.36$

if $z13 \leq 20.38$

$y = 22.6$ (4)

else

if $z13 \leq 21.04$

$y = 23.9$ (8)

else

$y = 26.6$ (17)

else

$y = 28.9$ (8)

else

if $z1 \leq 32.276$

if $z1 \leq 31.586$

$y = 30.9$ (10)

else

$y = 33.4$ (5)

else

if $z1 \leq 33.8226$

if $z13 \leq 20.19$

$y = 34.2$ (6)

else

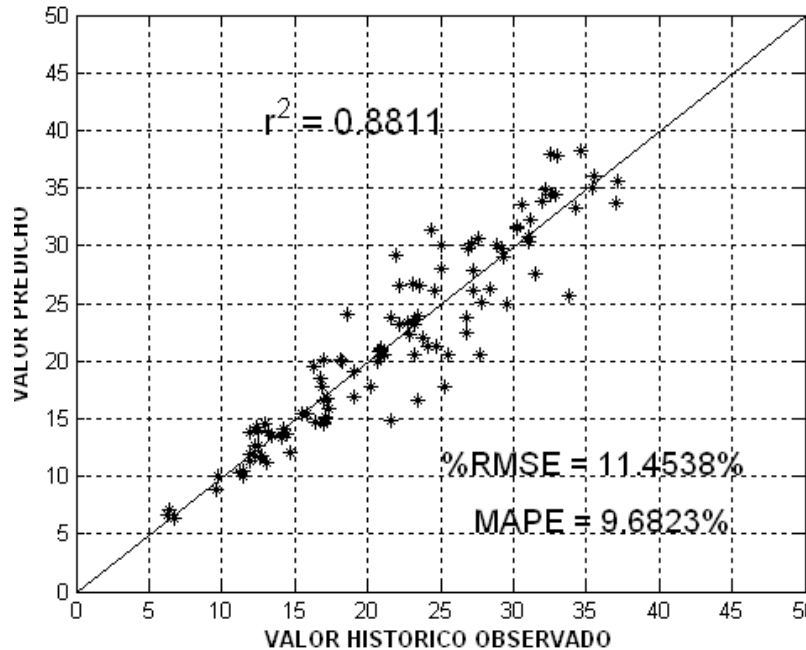
$y = 37.2$ (5)



6.8 RESULTADOS ALGORITMO M5

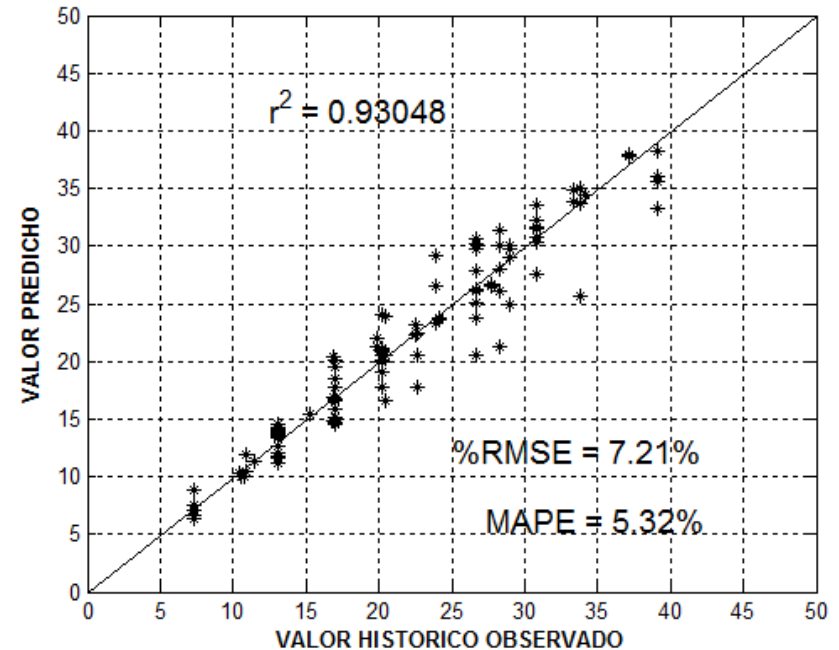
RNPOL (P)+(AES)

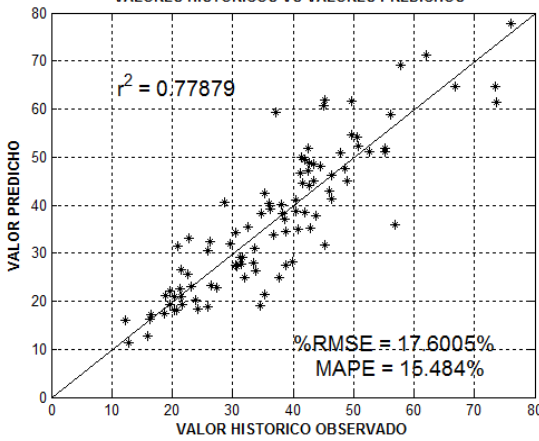
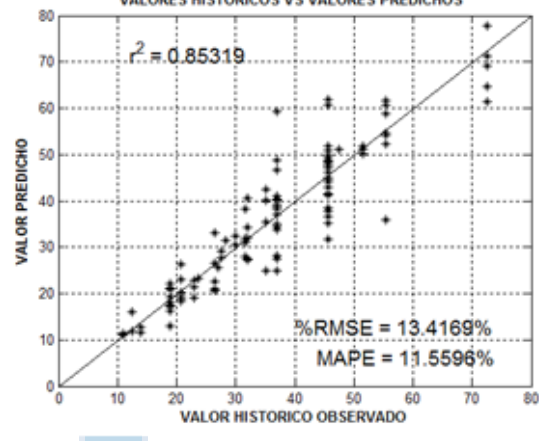
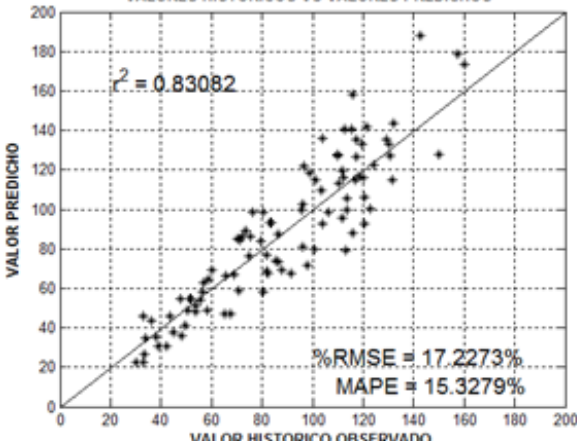
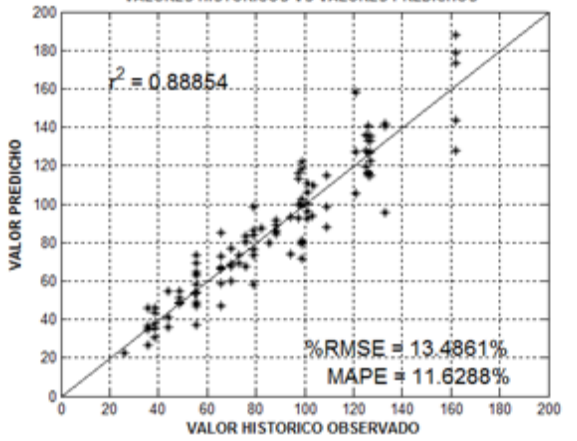
VALORES HISTÓRICOS VS VALORES PREDICHOS



RNPOL (P) + AES CORREGIDO

VALORES HISTÓRICOS VS VALORES PREDICHOS

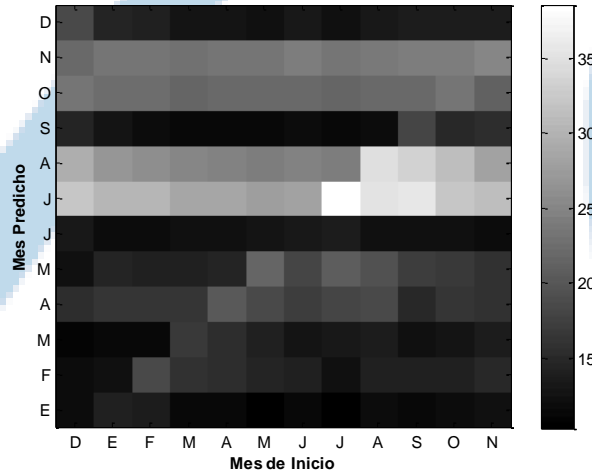


NOMBRE	RNPOL (P) +AES	CORRECCIÓN USANDO M5
RIO TENCHE	<p>VALORES HISTÓRICOS VS VALORES PREDICHOS</p>  <p>$r^2 = 0.77879$</p> <p>%RMSE = 17.6005% MAPE = 15.484%</p>	<p>VALORES HISTÓRICOS VS VALORES PREDICHOS</p>  <p>$r^2 = 0.85319$</p> <p>%RMSE = 13.4169% MAPE = 11.5596%</p>
RIO GUATAPE	<p>VALORES HISTÓRICOS VS VALORES PREDICHOS</p>  <p>$r^2 = 0.83082$</p> <p>%RMSE = 17.2273% MAPE = 15.3279%</p>	<p>VALORES HISTÓRICOS VS VALORES PREDICHOS</p>  <p>$r^2 = 0.88854$</p> <p>%RMSE = 13.4861% MAPE = 11.6288%</p>

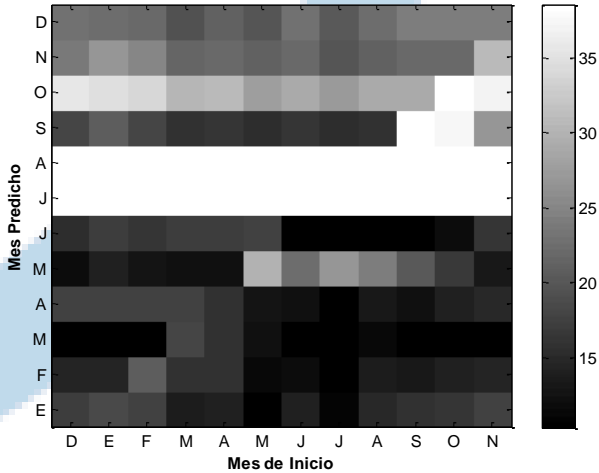


6.9 ÁRBOLES DE DECISIÓN DERIVADOS EMPÍRICAMENTE (I)

AJEDREZ RIO
GUADALUPE USANDO COMBINACIÓN-ANFIS

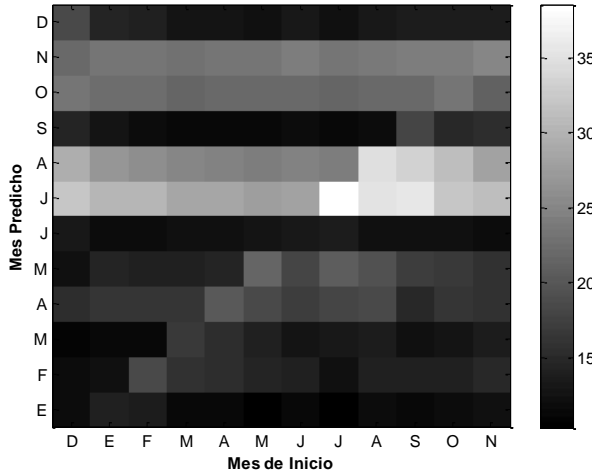


AJEDREZ RIO
GUADALUPE USANDO COMBINACIÓN-ANFIS



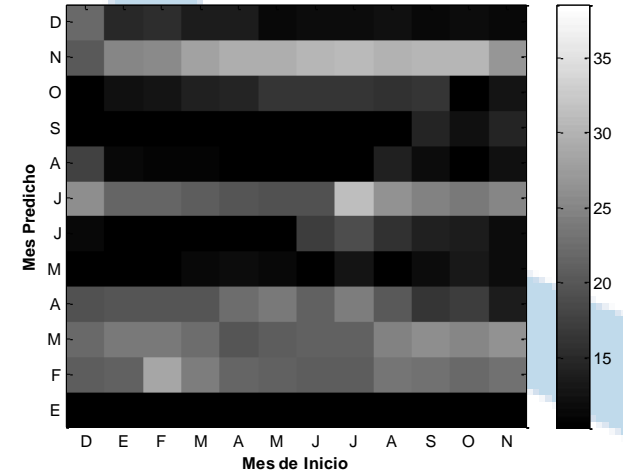
Niño

AJEDREZ RIO
GUADALUPE USANDO COMBINACIÓN-ANFIS



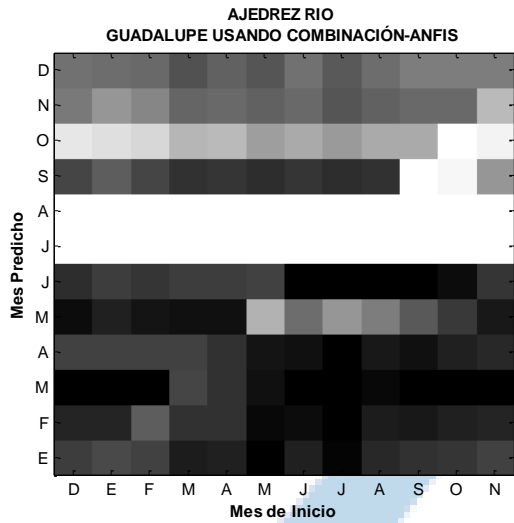
Normal

AJEDREZ RIO
GUADALUPE USANDO COMBINACIÓN-ANFIS

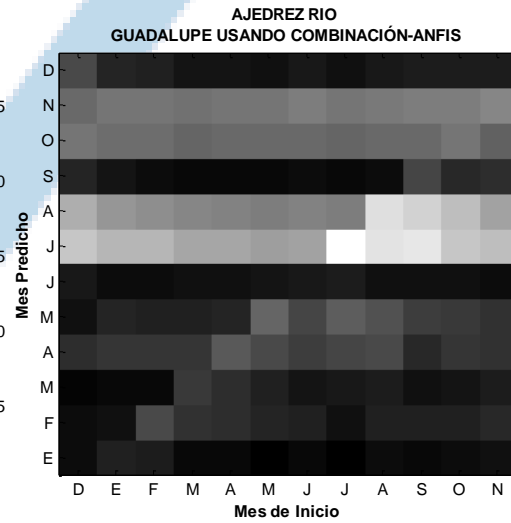


Niña

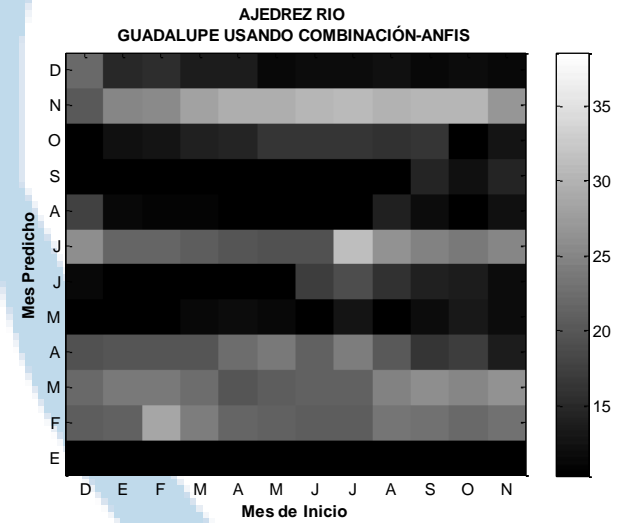
6.9 ÁRBOLES DE DECISIÓN DERIVADOS EMPÍRICAMENTE (I)



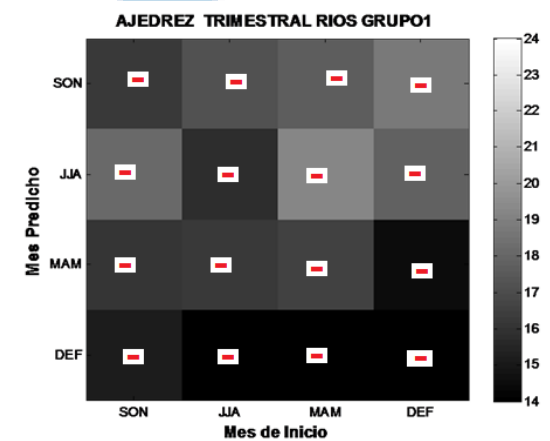
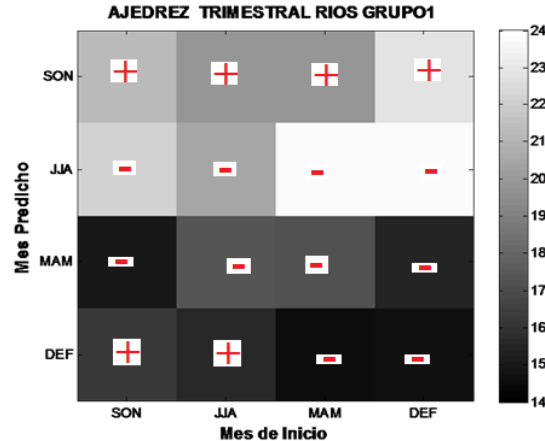
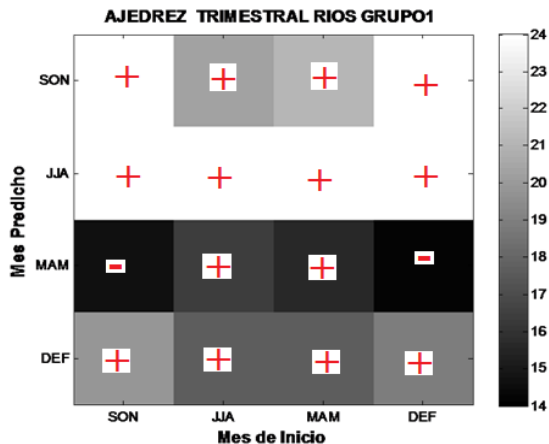
Niño



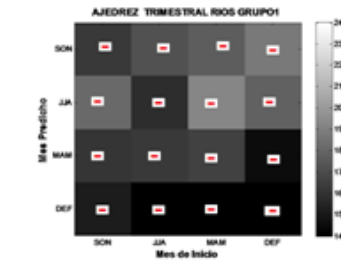
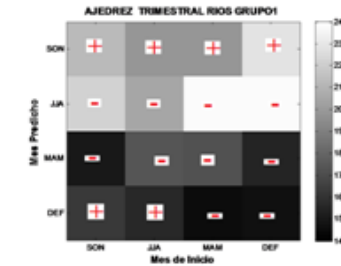
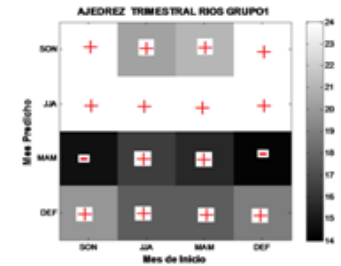
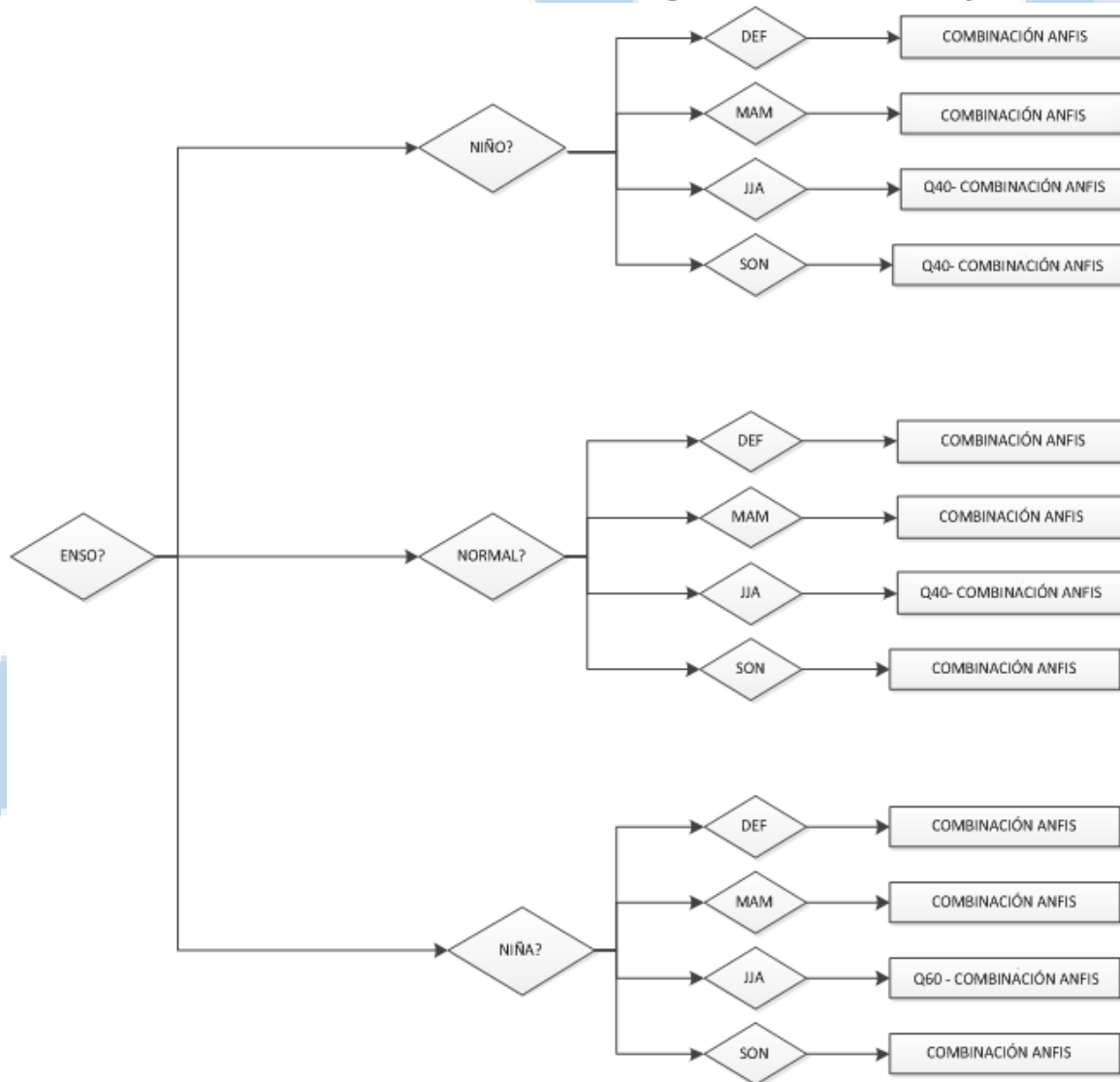
Normal



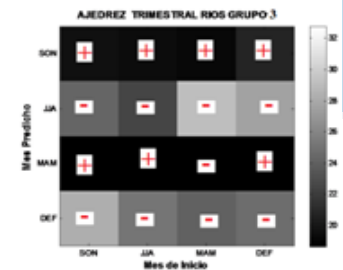
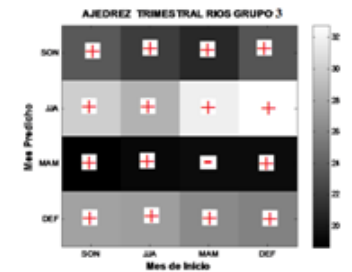
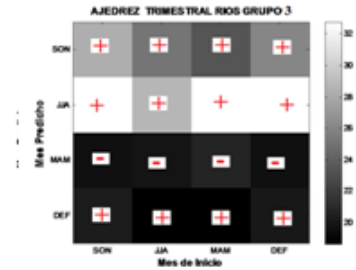
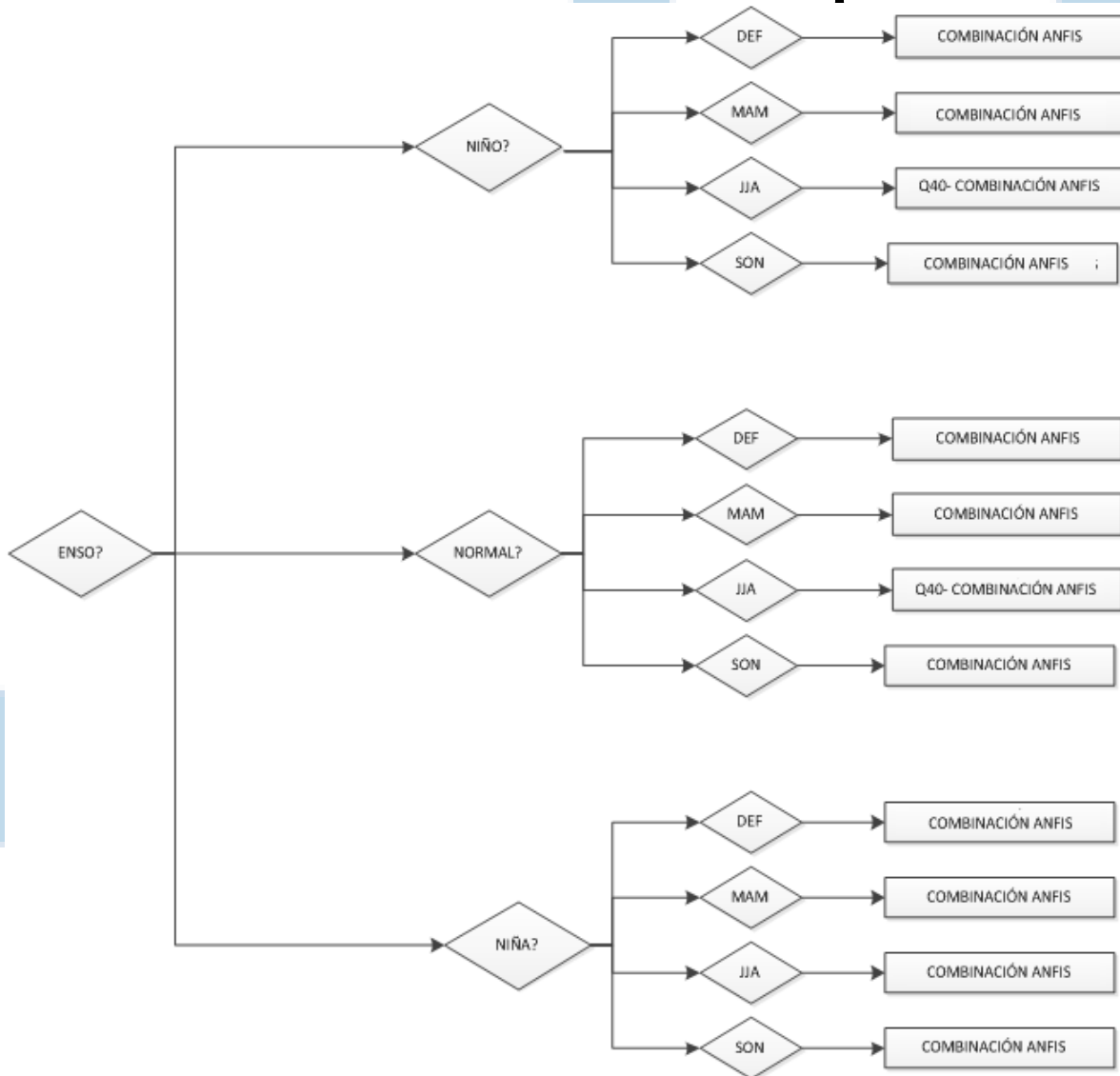
Niña



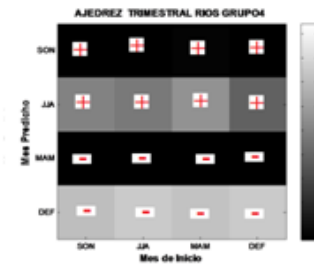
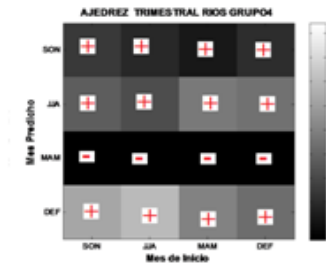
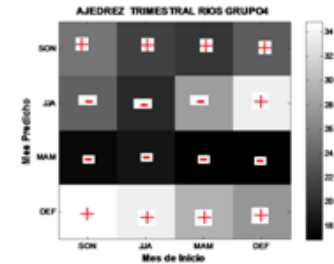
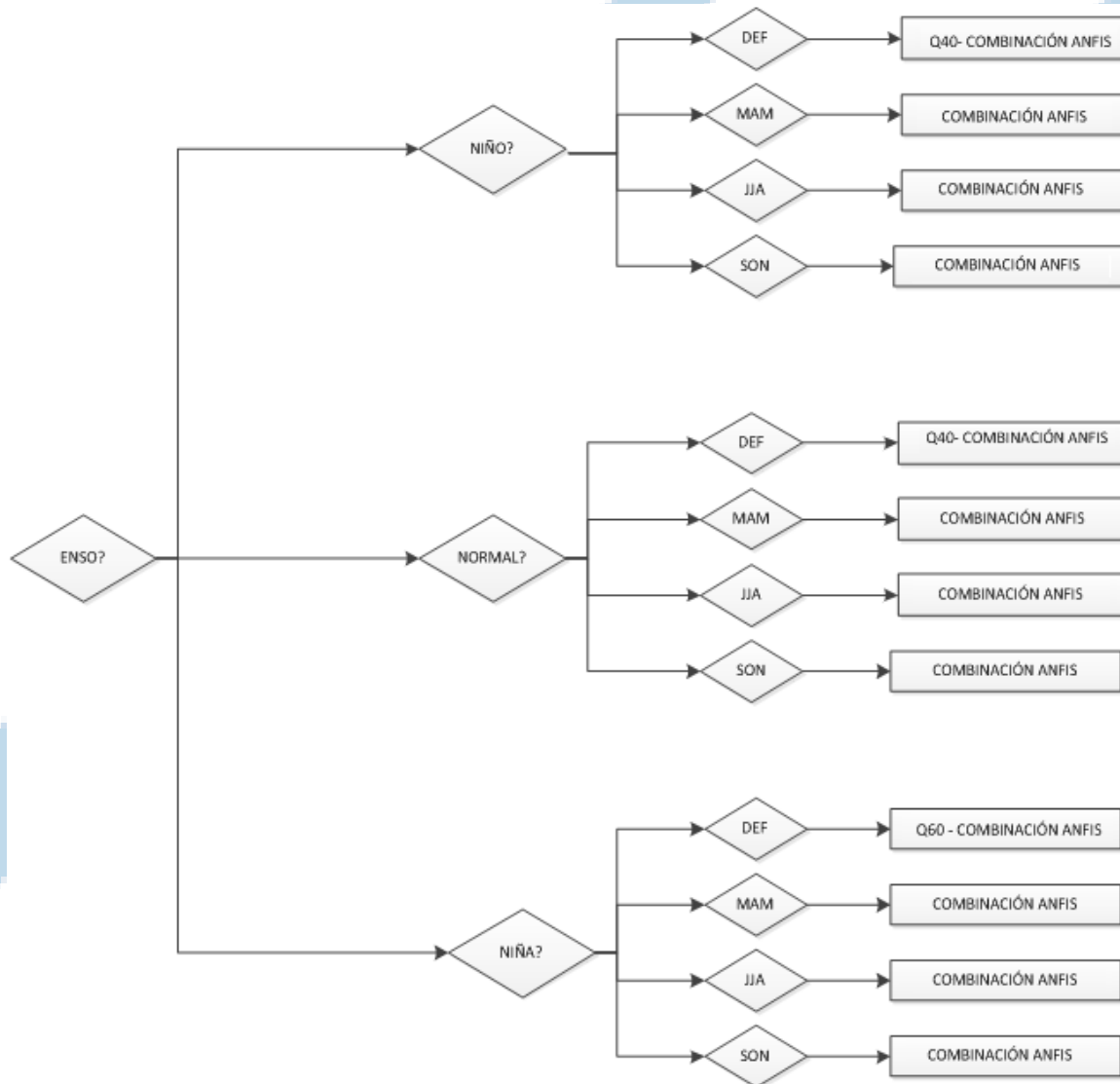
Arboles de decisión ríos Guadalupe, Tenche, Riogrande y Porce.



Arboles de decisión ríos Nare y Guatapé



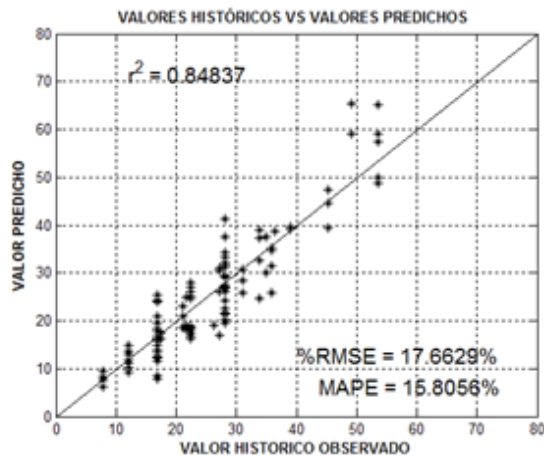
Arboles de decisión ríos Guavio y Bata



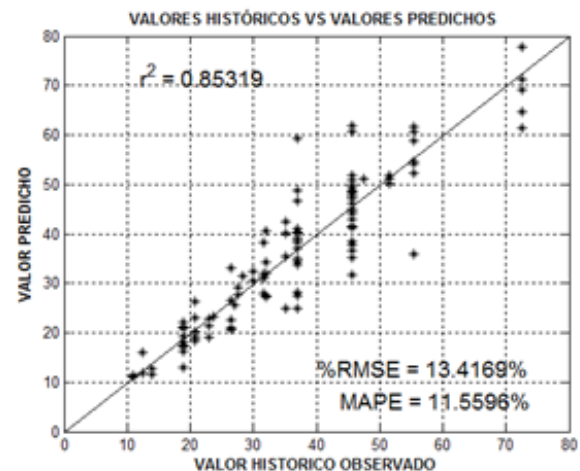
NOMBRE

RIO SAN
LORENZO

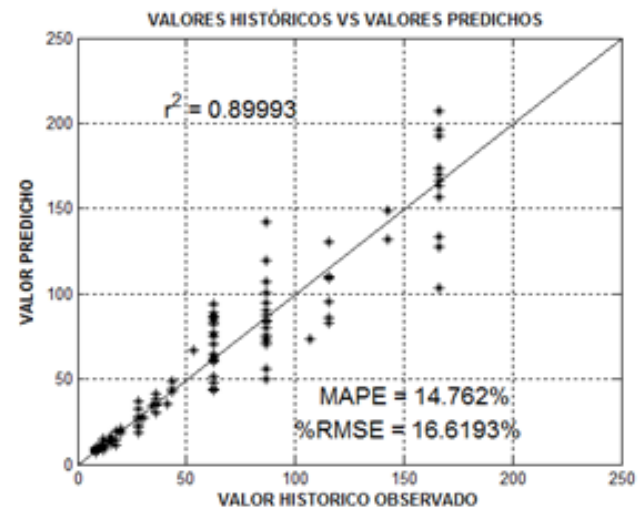
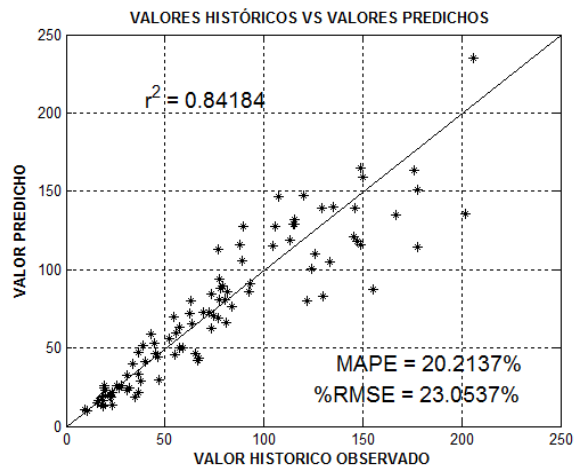
RNPOL (P) +AES



CORRECCIÓN EMPIRICA



RIO BATA

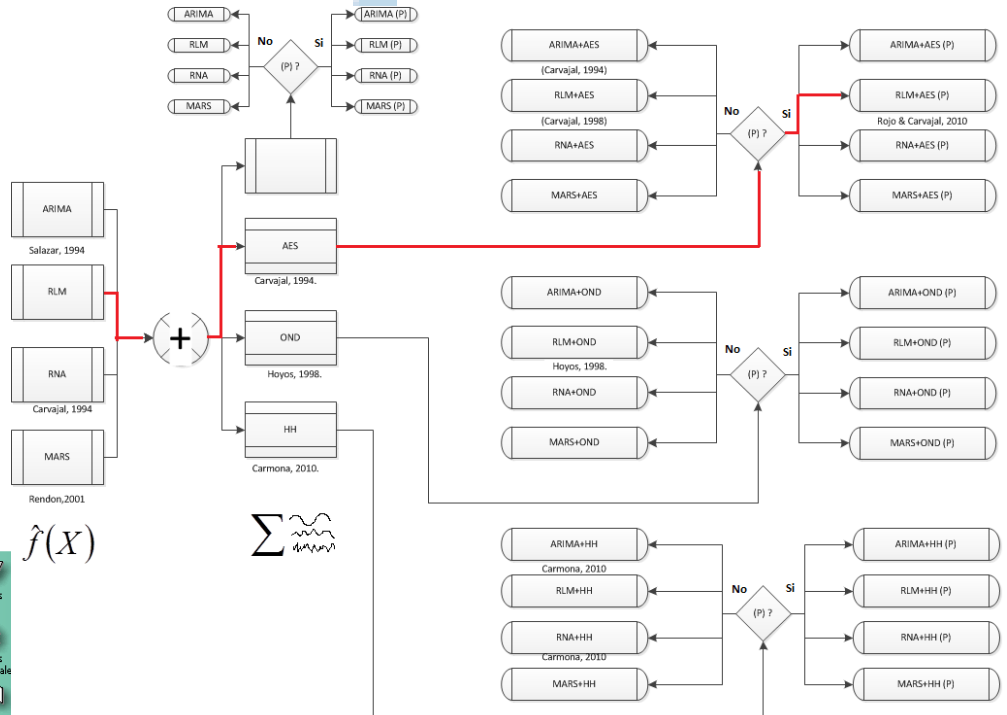
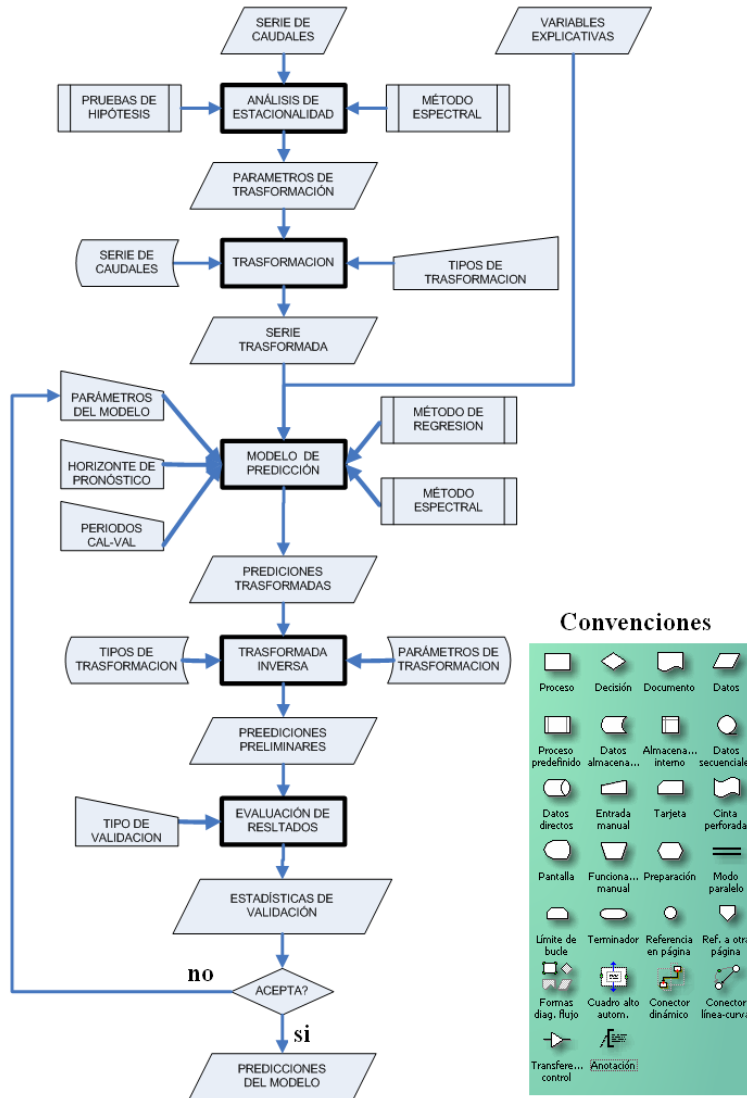


CONTENIDO

1. PROBLEMÁTICA DE LA PREDICCIÓN DE CAUDALES
2. RELACIONES ENTRE LOS CAUDALES DE COLOMBIA Y LA CLIMATOLOGÍA GLOBAL
3. MÉTODOS ACTUALES PARA EL PRONÓSTICO DE CAUDALES
4. INCORPORACIÓN DE NUEVAS TÉCNICAS DE PRONÓSTICO
5. COMPARACIÓN Y COMBINACIÓN DE PRONÓSTICOS HIDROLÓGICOS
6. PREDICCIÓN BASADA EN REGLAS DE JUICIO
- 7. PROTOTIPO DE SOFTWARE**
8. CONCLUSIONES Y RECOMENDACIONES



PROTOTIPO DE SOFTWARE (I)- PROCESOS

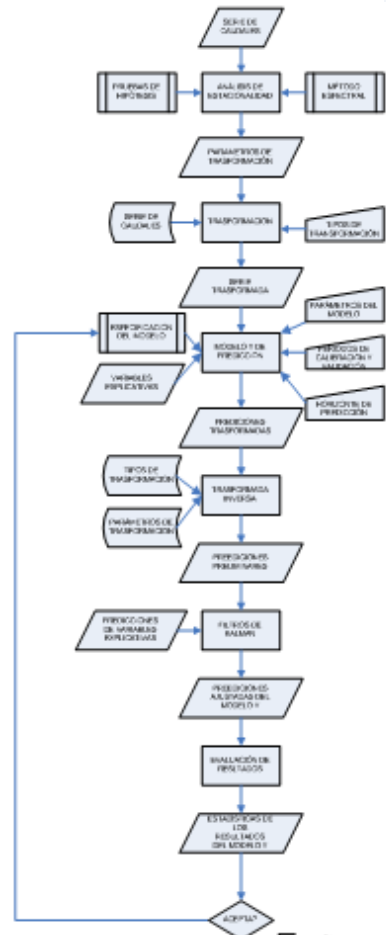


Convenciones

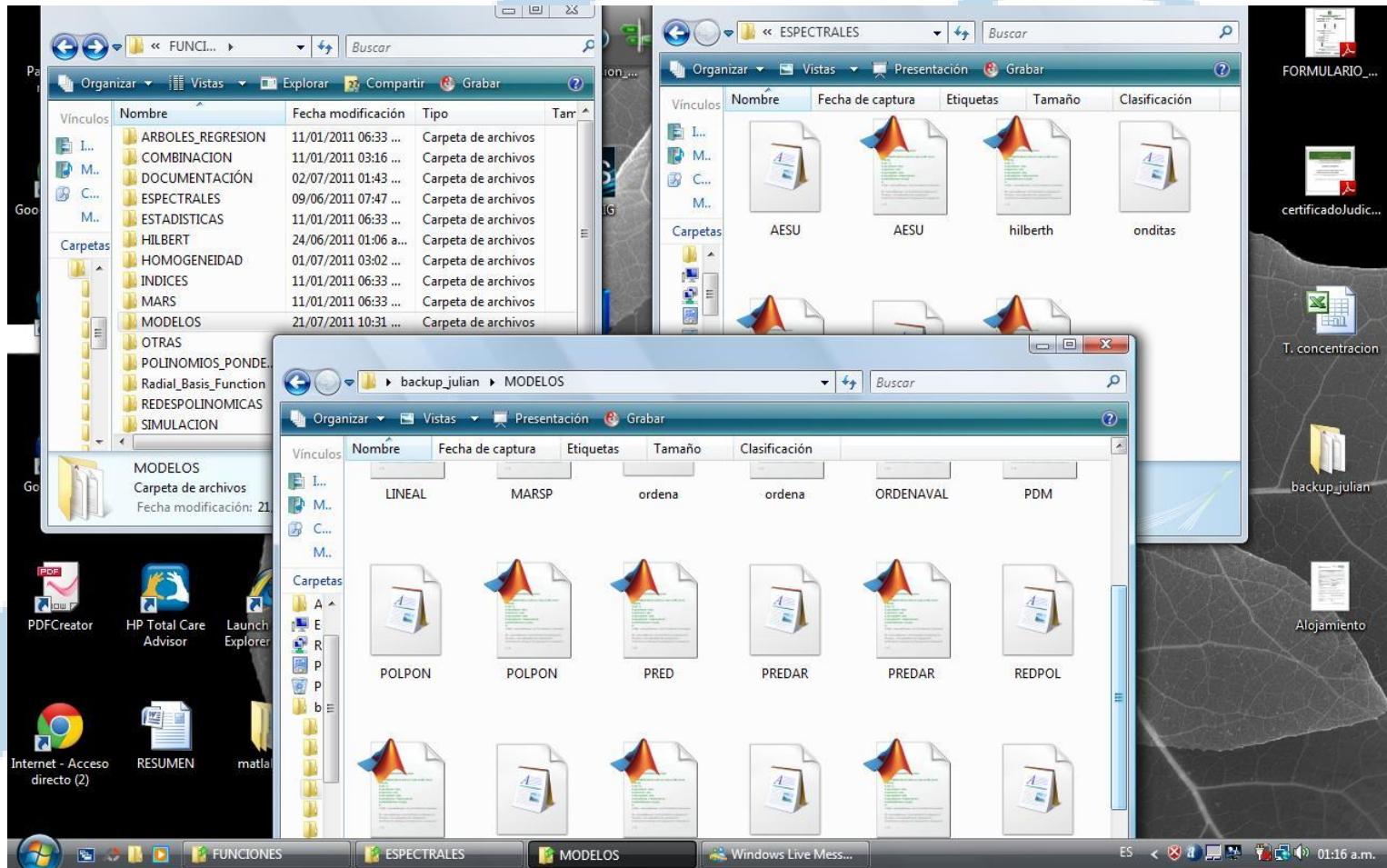
Proceso	Decisión	Documento	Datos
Proceso predefinido	Datos almacenados...	Almacena... interno	Datos secuenciales
Datos directos	Entrada manual	Tarjeta	Cinta perforada
Pantalla	Funciona... manual	Preparación	Modo paralelo
Límite de bucle	Terminador	Referencia en página	Ref. a otra página
Formas diag. flujo	Cuadro alto autom.	Conector dinámico	Conector línea-curva
Transfere... control	Anotación		

$$\hat{f}(X)$$





PROTOTIPO DE SOFTWARE (II)- FUNCIONES





PDFCreator



TeXnicCenter



Facemoods.....



MANUAL DE USUARIO.doc



REDESPL

Debut Video Capture Software © NCH Software

Cancel



R 2.10.1



RESUMEN HOJA DE VI...



2242229.pdf



TeamViewer 5



Hoja de vida extendida.doc



Ranking de desempe...



Formato_le...



ArcMap



Articulo 02b (1).pdf



Untitled 2.rar



guadalupe.rar



CLIPSWin.exe



TALLER FUNCIONE...



Beatriz.doc



certificadoJ...



Google Chrome



ubicacion.kmz



carta1.doc



Centro de soluciones HP



JCreator Pro



rectas.bmp



Babylon



Google Earth



MapWindow GIS



AESMP.rar



FoxTab Video Converter



carta2.doc



HEC-RAS 4.0 Beta



Mozilla Firefox



ANALISIS ESTADISTIC...



aa-ocw-reg...



HP Photosmart Essential



PsStudio.exe



CERCHA_T...



WIENTOSuv...



experiencia...



Mendeley Desktop



Texmaker



Marilia2.JPG



movconvert...



versión de impresión.pdf



redes.bmp

CONTENIDO

1. PROBLEMÁTICA DE LA PREDICCIÓN DE CAUDALES
2. RELACIONES ENTRE LOS CAUDALES DE COLOMBIA Y LA CLIMATOLOGÍA GLOBAL
3. MÉTODOS ACTUALES PARA EL PRONÓSTICO DE CAUDALES
4. INCORPORACIÓN DE NUEVAS TÉCNICAS DE PRONÓSTICO
5. COMPARACIÓN Y COMBINACIÓN DE PRONÓSTICOS HIDROLÓGICOS
6. PREDICCIÓN BASADA EN REGLAS DE JUICIO
7. PROTOTIPO DE SOFTWARE
- 8. CONCLUSIONES Y RECOMENDACIONES**



Seguimiento a los objetivos

- Establecer una base sobre el estado actual de predicción de caudales medios mensuales en el País.
- Fueron calibrados y validados modelos basados en los métodos de regresión y los métodos espectrales conocidos hasta el momento en el país.
- Se propuso el uso conjunto de esquemas de regresión periódica y métodos espectrales a fin de mejorar la eficiencia de modelos típicamente usados en el país para el pronóstico de caudales.
- Los resultados indican que las regresiones periódicas son más eficientes que las regresiones globales.
- Los resultados han demostrado las bondades de utilizar híbridos espectrales en el pronóstico de caudales.
- El mejor de los modelos aplicados en el primera instancia fue un híbrido periódico entre el modelo MARS y la transformada en Onditas. Sus resultados representan la línea de base ante las mejoras planteadas en el presente trabajo.

Seguimiento a los objetivos

Incorporar en la construcción del modelo el uso de variables explicativas que reflejen el comportamiento macro-climático

-
- Se elaboró un diagnóstico climático que permitió establecer las relaciones entre diferentes variables macro-climáticas y los caudales de los ríos del presente estudio.
- Se aplicó el Test de Debold y Mariano para estudiar la sensibilidad de los modelos ante la inclusión de variables macro-climáticas. Dicha herramienta permite dotar de un nuevo instrumento de base estadística para comprar diferentes usos de las variables macro-climáticas.
- La técnica de los polinomios ponderados por funciones de influencia radial representa una nueva forma de incorporar variables macro-climáticas usando el concepto de interpolación.

El método de las redes poli nómicas permite manejar una gran cantidad de información relacionada con las variables explicabas y su funcionamiento permite ponderar dicha información y seleccionar aquella que sea mas significativa.

Seguimiento a los objetivos

Revisar los últimos avances en temas relacionados con la predicción en hidrología e incorporar aquellos que sean relevantes para caso Colombiano.

-Aplicación de dos nuevos métodos al pronóstico de caudales medios mensuales en Colombia, los polinomios ponderados y las redes poli-nómicas. Los resultados indican que ambas metodologías son superiores en eficiencia en muchos de los casos analizados en el presente trabajo.

-Se introducen nuevas herramientas estadísticas para compara pronósticos y estimar cuales modelos son redundantes.

- Se introduce una metodología para mejorar los pronósticos de caudales usando la combinación de diversos métodos, los resultados indican que es mejor hacer una combinación de diferentes métodos de predicción en lugar de adoptar solo aquel modelo que sea mas eficiente en la validación

El mejor modelo de combinación fue el ANFIS, sin embargo se logran mejorar significativas tomando promedios ponderados de diferentes modelos usando la regresión lineal múltiple.

Seguimiento a los objetivos

Ajustar los pronósticos mediante reglas de juicio.

Fueron propuestos los árboles de regresión de decisión como una herramienta para elaborar de forma automática reglas de juicio que permitan corregir los pronósticos de caudales en función de pronósticos de variables explicativas que incluso pueden ser cualitativas.

Los resultados indican que los árboles de regresión de decisión basados en el algoritmo M5 permite corregir de forma significativa los pronósticos usando información de diversas agencias internacionales.

Los resultados de validación de los diversos modelos de pronóstico pueden convertirse en información útil para comprender el desempeño de los modelos para diferentes comportamientos de variables macro-climáticas.

Usando un procedimiento relativamente simple se pudo comprender que durante el fenómeno de El Niño, (fase cálida del ENSO) los modelos suelen sobre-estimar los caudales para los meses que históricamente poseen un invierno más intenso, generando así reglas de juicio para corregir a la baja dichos pronósticos.

Seguimiento a los objetivos

Desarrollar un Prototipo de Software para la predicción de Caudales Medios Mensuales Basado en modelos matemáticos y reglas de juicio.

De forma general, todas las metodologías establecidos en el presente trabajo para el pronóstico de caudales medios mensuales fueron compiladas en varias librerías de funciones que facilitan la aplicación de los diferentes procedimientos de pronóstico explicados en el presente trabajo.

AGRADECIMIENTOS

Al profesor Luis Fernando Carvajal, por su dedicación, apoyo y cercanía.

Al profesor Juan David Velásquez por sus recomendaciones.

A los profesores de la escuela de Geociencias y Medio Ambiente por sus enseñanzas.

A mis compañeros del posgrado por la amistad brindada.

A mi familia, por su apoyo incondicional.

A Leidy Andrea por su compañía.



UNIVERSIDAD
NACIONAL
DE COLOMBIA
SEDE MEDELLÍN



AGRADECIMIENTOS



Departamento Administrativo de
Ciencia, Tecnología e Innovación
Colciencias
República de Colombia



UNIVERSIDAD
NACIONAL
DE COLOMBIA
SEDE MEDELLÍN



POSGRADO EN APROVECHAMIENTO
DE RECURSOS HIDRÁULICOS

A la memoria de mi abuela Alicia...



UNIVERSIDAD
NACIONAL
DE COLOMBIA
SEDE MEDELLÍN



POSGRADO EN APROVECHAMIENTO
DE RECURSOS HIDRÁULICOS



MUCHAS GRACIAS



UNIVERSIDAD
NACIONAL
DE COLOMBIA
SEDE MEDELLÍN

POSGRADO EN APROVECHAMIENTO
DE RECURSOS HIDRÁULICOS