

Aplicación de Redes Funcionales al pronóstico de niveles en dos cursos de llanura del Gran Rosario

Carlos M. Scuderi¹, Gerardo A. Riccardi^{1,2}, Erik D. Zimmermann^{1,3}

¹Departamento de Hidráulica y CURIHAM. Facultad de Ciencias Exactas, Ingeniería y Agrimensura.

Universidad Nacional de Rosario. Riobamba 245 bis. Rosario (2000). Santa Fe. Argentina. Telefax: +54(341)4808541,

²CIC-CIUNR, ³CONICET

E-mail: cscuderi@fceia.unr.edu.ar

RESUMEN: Se presenta la utilización de redes funcionales (RF) para pronóstico de niveles en cursos de llanura. El objetivo es evaluar el comportamiento de este modelo de caja negra aplicado en diferentes secciones de los arroyos Ludueña y Saladillo (Rosario, sur de Santa Fe, Argentina). Se trabajó con 10 sensores pluviométricos y 16 limnimétricos, con transmisión automática cada 15 minutos, por un período que abarca desde el 01/07/2007 al 30/06/2010. Las cuencas de los arroyos fueron divididas en subcuencas, en correspondencia con los sensores limnimétricos; mientras que la precipitación asociada a cada limnómetro fue estimada por Thiessen. Los eventos registrados en cada subcuenca, fueron divididos en dos grupos: aprendizaje y validación; proponiendo tres familias de funciones para analizar: polinómica, exponencial y de Fourier. Los datos de entrada empleados consistieron en la precipitación y el nivel para un tiempo T0 con el fin de obtener como datos de salida los niveles para distintos horizontes de pronóstico, desde 15 minutos hasta 6 horas. Se realizó un análisis de sensibilidad de las muestras para calibración y validación a través del análisis combinatorio, calculando todas las muestras posibles. Los resultados, en términos de niveles, fueron evaluados a través de diferentes estadísticos (error porcentual en el pico, coeficientes a y b de la recta de ajuste, coeficiente de determinación, error cuadrático medio, entre otros); así como también en forma visual. En general, los mejores resultados se obtuvieron para la función polinómica. Por otro lado, a medida que el horizonte de pronóstico aumenta la calidad de los pronósticos disminuye. Los resultados obtenidos hasta el presente son alentadores evidenciando la capacidad de ‘aprendizaje’ y reproducción de eventos de este tipo de modelo; sin embargo existe cierta incertidumbre respecto del comportamiento del mismo frente a eventos de mayor magnitud a los utilizados en la etapa de entrenamiento.

INTRODUCCIÓN

La concepción actual en el manejo de inundaciones requiere un enfoque multidisciplinario; no sólo científico, sino también relacionado con aspectos sociales, económicos y políticos (Gruntfest, 1995). En el campo de la ingeniería, las medidas destinadas al 'control' de inundaciones pueden dividirse en estructurales y no estructurales (Berga Casafont, 1995). Las primeras intervienen durante la formación y ocurrencia de los fenómenos al modificar las características hidrológicas e hidráulicas de la cuenca (por ej. alteración de la cobertura vegetal o ensanchamiento de cursos). Mientras que las segundas plantean la convivencia entre el hombre y el agua dentro de un determinado espacio geográfico con el fin de minimizar los daños producto de las inundaciones (por ej. regulación del uso del suelo o sistemas de alerta contra inundaciones). En la actualidad varios autores coinciden en que la tendencia generalizada es el empleo de ambos tipos de medidas en forma conjunta (Berga Casafont, 1995; Lekkas et al., 2004).

Los Sistemas de Alerta contra Inundaciones (SAI) tienen como objetivo el seguimiento continuo y pronóstico de las variables hidrológicas (principalmente lluvia y niveles) que pudieran provocar situaciones de riesgo para la población; siendo una importante herramienta de ayuda en los procesos de toma de decisión ante la emisión de alertas y eventuales evacuaciones (Dueñas Molina, 1995; Singh, 2005). En forma general, un SAI se compone de una red de sensores que recolectan datos hidrológicos, un sistema de transmisión de datos a tiempo real y un modelo matemático de pronóstico de niveles. Los requerimientos del modelo de pronóstico para ser utilizado en tiempo real, según Aldana Valverde (2002), son: sencillez de operación, precisión de resultados, tiempo de ejecución breve y capacidad de autocalibración de parámetros.

En cuanto a los modelos de pronóstico de niveles a emplear, éstos pueden ser distribuidos o concentrados (Hsu et al., 2003). Los primeros, físicamente basados, pueden llegar a representar con mayor rigurosidad los procesos físicos desarrollados durante una crecida en todo punto de la cuenca (Ogden et al., 2000); sin embargo demandan gran cantidad de datos de entrada, recurso computacional y tiempo de cálculo. Por otra parte, los modelos concentrados o tipo 'caja negra', obedecen a una concepción diferente: debido a que el pronóstico de niveles (y/o caudales) en tiempo real es una situación práctica cuyo objetivo es disponer de resultados lo más aproximados en el menor tiempo posible, es válido ignorar ciertos aspectos y componentes físicos del sistema estudiado (Hsu et al., 1995). Estos modelos relacionan la precipitación con el nivel (y/o caudal) en los cursos de forma muy simplificada; siendo precisamente esta sencillez lo que hace que ganen popularidad en el pronóstico de crecidas. Entre los modelos más empleados destacan las redes neuronales y como una generalización de las mismas (Castillo, 1998), las redes funcionales (RF). Dicho autor cita algunas de las principales características de las RF: a) la topología es dependiente del problema, pudiendo ser simplificada mediante ecuaciones; b) no es necesario especificar las funciones neuronales ya que éstas 'aprenden' a partir de los datos de entrada; y c) las funciones neuronales pueden ser diferentes, multivariadas y de múltiples argumentos.

Varios son los trabajos publicados relativos a la aplicación de modelos tipo 'caja negra' al pronóstico de caudales para diferentes pasos de cálculo, desde anual hasta horario. Campolo et al. (1999) emplearon redes

neuronales para estimar caudales en una cuenca de 1950 km², tomando como variables de entrada los caudales previos y las precipitaciones horarias; obteniendo valores de $R^2=0.87$ en calibración y de $R^2=0.80$ en validación para 6 hs de antecedencia. Kothyari y Singh (1999) aplicaron en una cuenca de 17157 Km², en India, un modelo concentrado con paso de tiempo diario utilizando 16 pluviómetros y tomando 6 eventos para calibración y 4 para validación; logrando resultados con valores de $R^2=0.79$ y $RMSE=837 \text{ m}^3 \cdot \text{s}^{-1}$ en calibración y de $R^2=0.79$ y $RMSE=314 \text{ m}^3 \cdot \text{s}^{-1}$ en validación. Bruen y Yang (2005) emplearon redes neuronales y funcionales en dos cuencas urbanas y una rural, con pasos de tiempo de 5 min y 1 h respectivamente, demostrando la superioridad de las redes funcionales sobre las neuronales usando como indicador el $R^2 (>0.95)$.

El objetivo del presente trabajo es evaluar la aplicación de diferentes modelos de redes funcionales al pronóstico de niveles en ciertas secciones características de los arroyos Ludueña y Saladillo, cursos de llanura situados en la zona del Gran Rosario (sur de la provincia de Santa Fe, Argentina). Para ello serán propuestas y analizadas tres familias de funciones: polinómica, exponencial y de Fourier; siendo utilizados como datos de entrada la precipitación y el nivel limnimétrico en un tiempo T_0 con el fin de obtener como datos de salida los niveles limnimétricos para varios tiempos de pronóstico, que abarcan desde 15 minutos hasta 6 horas de anticipación.

MATERIALES Y MÉTODOS

Cuencas en estudio

La zona del Gran Rosario (sureste de la provincia de Santa Fe, Argentina) ha sido impactada en forma periódica por diferentes eventos pluviométricos intensos, el último de ellos ocurrido en marzo de 2007. Este hecho sumado a la creciente presión sobre el uso de la tierra producto del cultivo de soja (*Glycine max*) y al desarrollo inmobiliario sin planificación, incrementan la vulnerabilidad de la población frente a las inundaciones. Razón por la cual la protección de vidas humanas, generalmente sectores marginales asentados sobre las márgenes de los cursos de agua, y de bienes materiales adquiere cada vez mayor importancia.

Las dos cuencas utilizadas en este trabajo, cuyas denominaciones son Ludueña y Saladillo, se encuentran situadas en el sureste de la provincia de Santa Fe (República Argentina). El suelo predominante es de tipo arcilloso-limoso con permeabilidad moderada a moderadamente lenta (INTA, 1983). La precipitación media anual en la zona de influencia es de 960 mm (Rosario Aero, SMN, serie 1892-2008), registrándose el 68% en el período primavera-verano (octubre a marzo). Ambas cuencas son atravesadas por numerosas vías de comunicación, tanto viales como ferroviarias; mientras que la agricultura constituye el principal uso del suelo, destacándose la producción de soja, maíz y trigo.

La cuenca del A° Ludueña se ubica entre los paralelos 32° 45' y 33° 07' S y los meridianos 60° 41' y 61° 06' O (Figura 1). La red hídrica se compone del A° Ludueña y los canales Ibarlucea y Salvat con sus tributarios, cuya longitud total es de 140 Km; sin embargo, en época de lluvias la gran cantidad de pequeños cursos

intermitentes acrecienta dicha longitud hasta los 370 Km. El área de aporte es de 740 Km² y su elevación varía entre 70 m y 18 m sobre el nivel del mar, con pendiente media del orden de 2.5 m.Km⁻¹. El caudal base del arroyo Ludueña es de 0.5 m³.s⁻¹, alcanzando los 80 m³.s⁻¹ en crecidas ordinarias y caudales superiores a los 400 m³.s⁻¹ en eventos extraordinarios con recurrencia mayor a los 50 años (Riccardi et al., 2002).

La cuenca del A° Saladillo se ubica entre los paralelos 32° 59' y 33° 37' S y los meridianos 60° 36' y 61° 54' O (Figura 2). La red hídrica principal se compone del arroyo Saladillo, el arroyo La Candelaria y el canal Sanford-Arequito, entre otros. La longitud del curso principal (arroyo Saladillo) es de 145 Km, mientras que la red total de cursos (considerando los afluentes de los anteriores) es de 359 Km. El área de aporte es de 3150 Km² y su elevación varía entre 115 m y 18 m sobre el nivel del mar. El caudal base del arroyo Saladillo es de 1 m³.s⁻¹, alcanzando los 1140 m³.s⁻¹ para eventos extraordinarios (Riccardi et al., 2008).

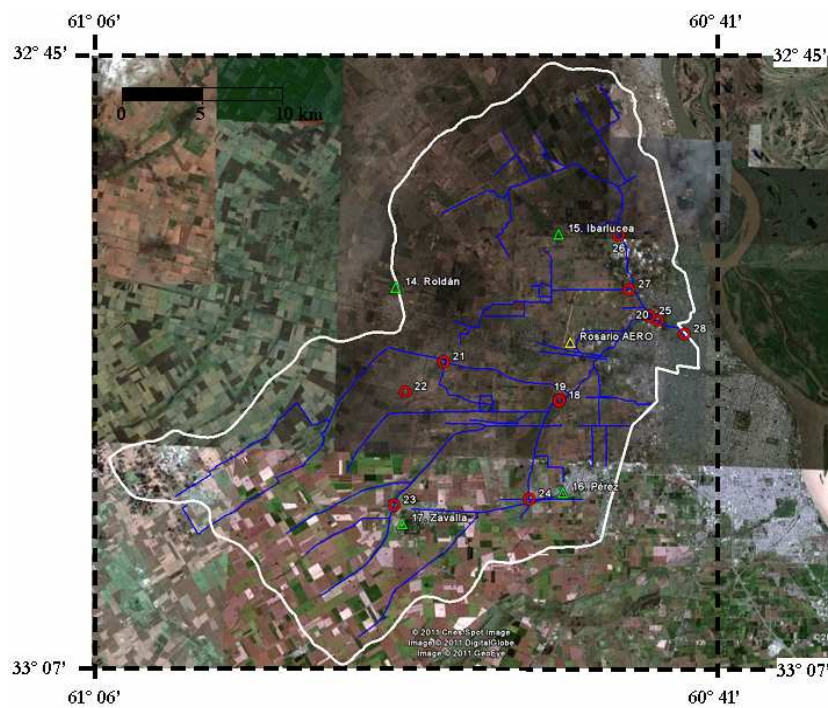


Figura 1.- Cuenca del Arroyo Ludueña.

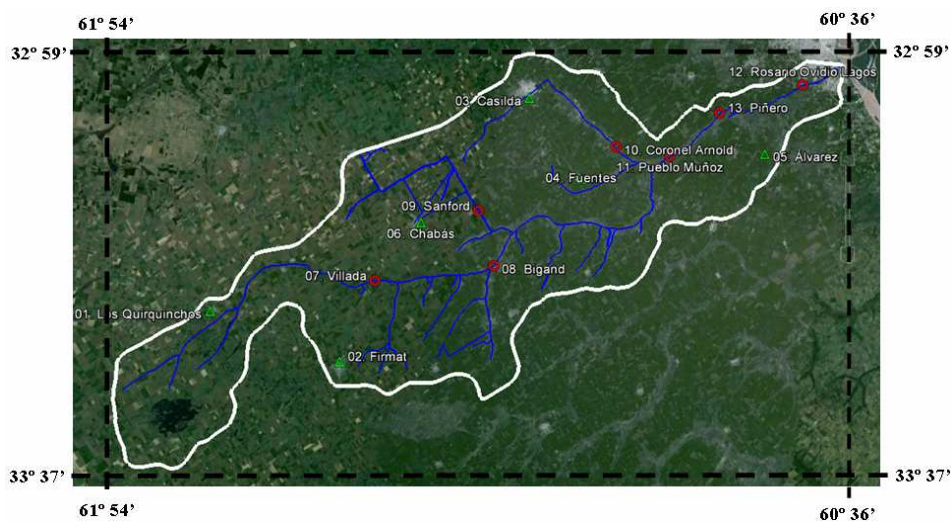


Figura 2.- Cuenca del Arroyo Saladillo.

Adquisición y procesamiento de datos

Los datos de precipitaciones y niveles fueron provistos por el Ministerio de Aguas, Servicios Públicos y Medio Ambiente de la provincia de Santa Fe. Los sensores pluviométricos (“Global Water”, RG200) de 6 pulgadas de boca tienen una apreciación de 0,25 mm; mientras que los limnimétricos (“Global Water”, WL400) son transductores de presión sumergibles con rango de 0 a 30 pies y precisión de $\pm 0,1\%$ de la escala completa. El registro de datos se efectúa automáticamente cada 15 minutos, siendo transmitidos vía celular a la estación central; el período de datos disponibles abarca desde el 01/07/2007 hasta el 30/06/2010. Los datos pluviométricos fueron contrastados (dobles acumulaciones) con la estación Rosario Aero (SMN) resultando que las estaciones remotas miden por debajo de la estación de testeo, en un rango entre 19% y 44%. Algunos registros limnimétricos pudieron ser cotejados con reglas en ciertas secciones; sin embargo, los mismos han estado sometidos a situaciones de vandalismo como rotura de cables o hurto de equipos. Un análisis pormenorizado sobre el comportamiento de los sensores puede ser consultado en Scuderi et al. (2009).

El procesamiento de datos consistió, primeramente, en dividir cada cuenca en subcuencas correspondientes al área de aporte a cada sensor limnimétrico (11 en Ludueña y 7 en Saladillo); utilizando curvas de nivel espaciadas cada 1,25 m junto con la traza de los terraplenes viales y ferroviarios. Un esquema de las subcuencas y limnímetros asociados se representa en las Figuras 3a y 3b. Posteriormente se estimó el área de influencia de cada sensor pluviométrico (4 en Ludueña y 6 en Saladillo) mediante polígonos de Thiessen (Figuras 4a y 4b). En función de las áreas obtenidas anteriormente (subcuencas y polígonos de Thiessen), fue calculado el porcentaje de contribución de cada sensor pluviométrico a cada subcuenca a través de la intersección de ambas superficies. Una tercera etapa fue compatibilizar los datos de precipitación y niveles registrados; se definieron como eventos válidos aquellos que cumplieran dos requisitos: (a) la precipitación caída en la subcuenca supera los 10 mm (función del aporte proporcional de cada pluviómetro) y (b) el nivel de agua en la sección estudiada registró un incremento mayor al 15% del máximo de la diferencia entre nivel pico y nivel base de todos los eventos medidos en la sección. Obteniendo de esta forma diferente número de eventos para cada subcuenca, entre 4 y 22 tormentas, con un valor medio de 9 eventos por subcuenca.

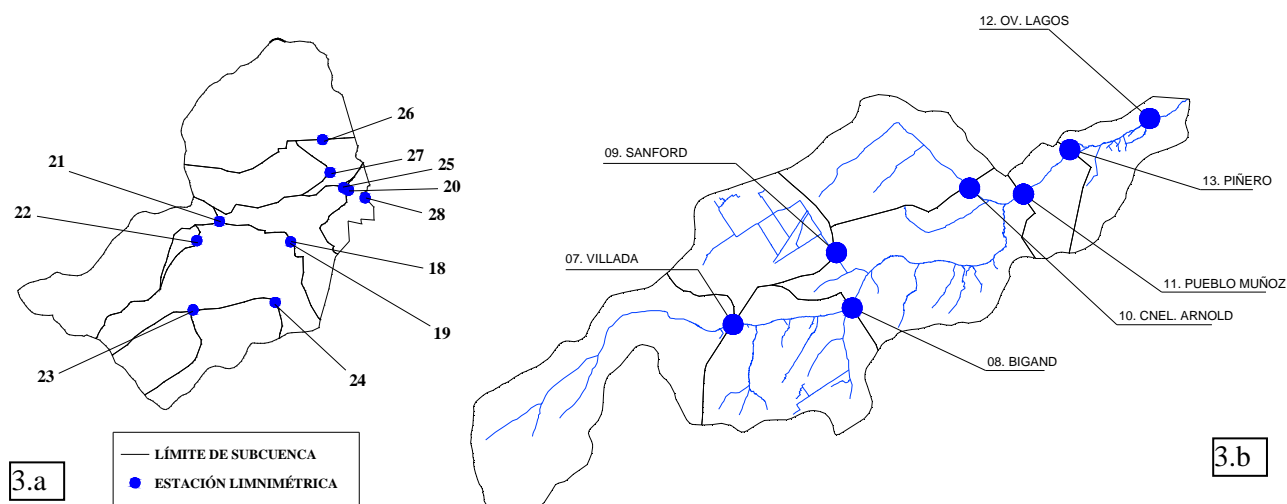


Figura 3.- Subcuencas asociadas a cada sensor limnimétrico: (a) Ludueña y (b) Saladillo.

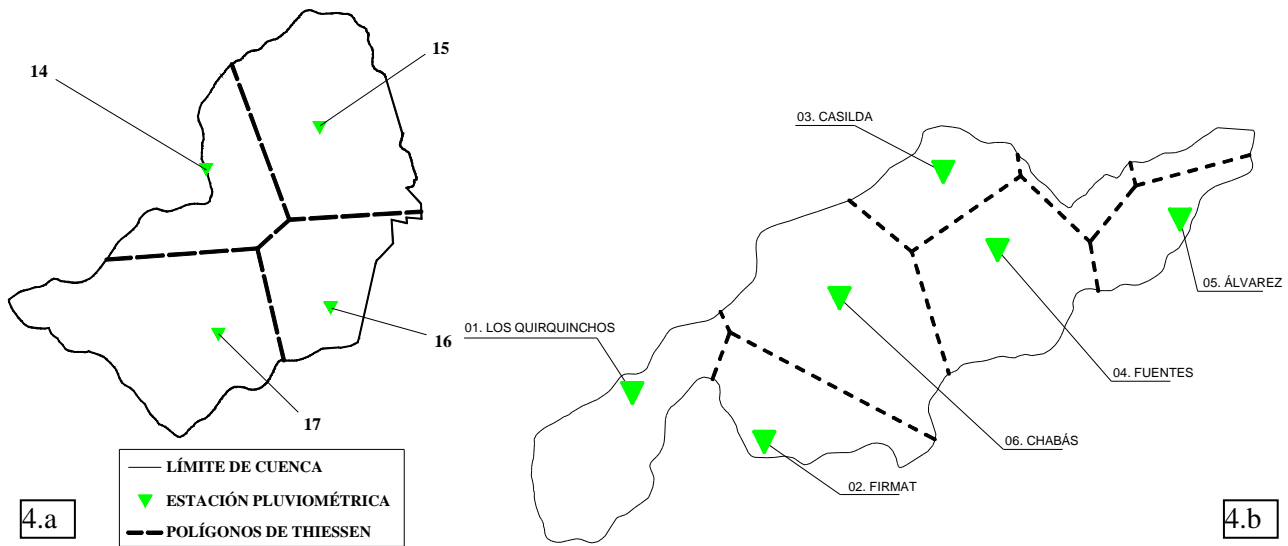


Figura 4.- Polígonos de Thiessen para calcular influencia de cada pluviómetro: (a) Ludueña y (b) Saladillo.

Modelo de redes funcionales

Uno de los atractivos del uso de redes neuronales (y por extensión de las redes funcionales) es la habilidad de ‘aprender’ a partir de datos y su potencial para describir en forma precisa el comportamiento de sistemas no lineales complejos, como el proceso de transformación lluvia – caudal (Bruen y Yang, 2005). Castillo (1998) cita los elementos que componen una red funcional (Figura 5): a) una capa de unidades de entrada, que contiene los datos de entrada; b) una capa de unidades de salida, es la última capa y contiene los datos de salida; c) una o varias capas de neuronas o unidades de cómputo, donde se evalúa un conjunto de valores de entrada provenientes de la capa anterior y produce un conjunto de valores de salida a la próxima capa; d) un conjunto de enlaces dirigidos, que conectan la capa de entrada a la primera capa de neuronas, neuronas de una capa a otras de la próxima capa, y la última capa de neuronas con la capa de salida. La dirección de la información se produce en una única dirección, desde la capa de entrada hacia la capa de salida.

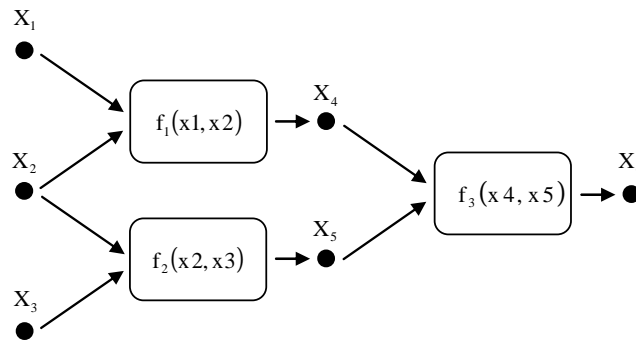


Figura 5.- Esquema típico de una red funcional (RF) (Fuente: Bruen y Yang, 2005).

En este trabajo se propone la utilización de un esquema de red funcional simple separable con dos variables de entrada (precipitación y nivel para un tiempo T_0) y una de salida (nivel para diferentes T_i), como el esquematizado en la Figura 6.

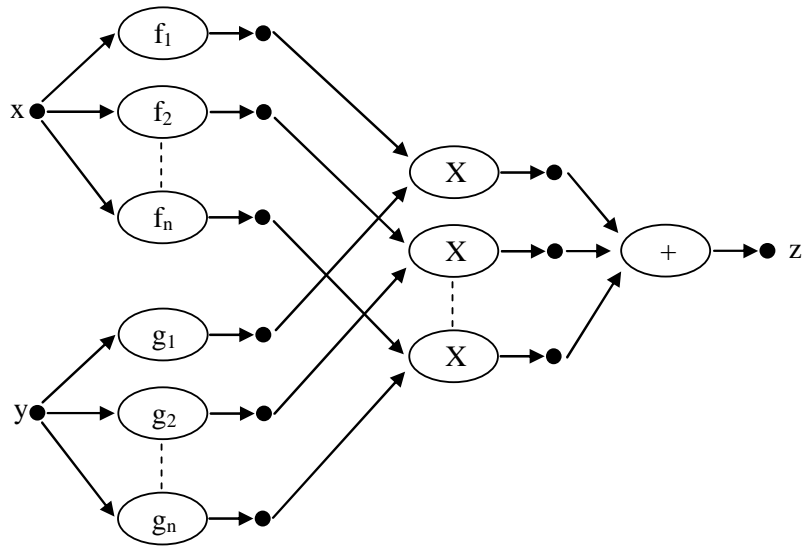


Figura 6.- Arquitectura de una RF separable con dos entradas y una salida (Fuente: Castillo y Gutiérrez, 1998).

La relación entre las variables x , y y z puede ser establecida matemáticamente por la ecuación (1):

$$z = F(x, y) = \sum_{i=1}^n f_i(x)g_i(y) \quad (1)$$

donde x e y son las variables de entrada (en esta aplicación lluvia y nivel para un tiempo T_0), z es la variable de salida (en este caso el nivel para un tiempo T_i , diferente al T_0) y f_i y g_i son las funciones neuronales a ser calculadas. El proceso de entrenamiento consiste en obtener las funciones f y g a partir de los datos de entrada y es, en la práctica, equivalente al proceso de calibración de un modelo hidrológico convencional. Las funciones neuronales f y g pueden ser expresadas según la ecuación 2 (Bruen y Yang, 2005):

$$\hat{f}(x) = \sum_{j=1}^p a_j \varphi_j(x) \quad , \quad \hat{g}(y) = \sum_{j=p+1}^{p+q} a_j \varphi_{j-p}(y) \quad (2)$$

en que los coeficientes a_j son los parámetros de las funciones neuronales φ_j que serán ‘aprendidos’ durante el proceso de entrenamiento de la red, y p y q son los órdenes de cada familia de funciones. El objetivo durante el entrenamiento es la minimización de una función de error que contempla la diferencia entre el valor calculado por el modelo y el valor medido. La función objetivo usada es la suma de los errores cuadráticos. Castillo y Gutiérrez (1998) demostraron que para lograr una representación única de la red es necesario dar un valor inicial a alguna de las funciones neuronales. En este caso pueden utilizarse cualquiera de las dos condiciones iniciales: $f(x_0) = u$ o $g(y_0) = v$. Este término se agrega a la función objetivo, quedando como se presenta en la ecuación 3 (Bruen y Yang, 2005):

$$FO = \sum_{i=1}^k \left[z_i - \sum_{j=1}^p a_j \varphi_j(x_i) - \sum_{j=p+1}^{p+q} a_j \varphi_{j-p}(y_i) \right]^2 + c \cdot \left[\sum_{j=1}^p a_j \varphi_j(x_0) - u \right] \quad (3)$$

siendo k la cantidad de datos de entrenamiento y c una constante. La resolución de la función objetivo FO es equivalente a resolver un conjunto de derivadas de FO respecto de los parámetros a_j y del multiplicador c ,

obteniendo un sistema de $p+q+1$ ecuaciones lineales donde los coeficientes a calcular son los parámetros a_j y la constante c . Para conocer con mayor detalle el análisis matemático consultar Bruen y Yang (2005).

En este trabajo se propuso utilizar tres familias funcionales: polinómica, exponencial y de Fourier. Además se plantearon 4 modificaciones de las mismas (3 exponenciales y 1 de Fourier) por lo que, finalmente fueron obtenidas 7 familias funcionales. Cada una de éstas se encuentra especificada en la Tabla 1. El grado de cada familia se considera incorporando al término anterior, es decir, por ejemplo la familia polinómica de grado 1 está constituida por $1 + x$, mientras que la de grado 2 se compone por los términos $1 + x + x^2$.

Tabla 1.- Familias funcionales propuestas.

Nombre	Términos
1_Polinómica	$1, x, x^2, \dots, x^9$
2_Exponencial_A	$1, e^x, e^{-x}, e^{2x}, e^{-2x}, \dots, e^{5x}$
3_Exponencial_B	$1, e^x e^{-x}, e^{2x} e^{-2x}, \dots, e^{9x} e^{-9x}$
4_Exponencial_C	$1, e^x, e^{2x}, \dots, e^{9x}$
5_Exponencial_D	$1, e^{-x}, e^{-2x}, \dots, e^{-9x}$
6_Fourier_A	$1, \text{sen} x, \text{cos} x, \text{sen} 2x, \text{cos} 2x, \dots, \text{sen} 5x$
7_Fourier_B	$1, \text{sen} x \text{cos} x, \text{sen} 2x \text{cos} 2x, \dots, \text{sen} 9x \text{cos} 9x$

Software desarrollado

El software fue desarrollado en lenguaje Fortran 90 y su implementación se realizó en una PC de 2 núcleos con 1.86 GHz de velocidad de procesador y 2 GB de RAM. Los datos de entrada están constituidos por la precipitación asociada a la subcuenca en donde se encuentra la estación limnimétrica y el nivel limnimétrico registrado en la misma; ambos con pasos de tiempo de 15 minutos. A partir de este paso de tiempo fueron considerados diferentes grados de agregación de la información, partiendo desde los 15 minutos, pasando por 30, 45 minutos, 1, 2, 3, 4, 5 y 6 horas. La arquitectura del software está conformada por diferentes tareas que responden a la siguiente descripción:

a) Generador de muestras: esta subrutina construye las combinaciones que surgen de tomar, para cada limnómetro, diferentes submuestras del total de eventos registrados para ‘entrenar’ el modelo de pronóstico y los eventos no considerados en la etapa de entrenamiento para validar el mismo. Básicamente calcula el número de combinaciones posibles asociadas al número total de eventos para cada sensor con la siguiente limitante; el número de eventos utilizado para validación no debe ser menor que el 50% del número de eventos usados para el entrenamiento. Es decir en caso de tener 8 eventos en una estación limnimétrica, las combinaciones posibles se definen como todas las submuestras que se pueden obtener a partir de 4 eventos para entrenamiento y todas las submuestras que se pueden obtener a partir de los 4 eventos no considerados en el entrenamiento, que serán utilizadas para validación y viceversa. También son válidas las submuestras para entrenamiento obtenidas a partir de 5 eventos y las submuestras ‘complementarias’ para validación

calculadas a partir de los 3 eventos no considerados en el entrenamiento y viceversa; no serían válidas las submuestras halladas a partir de 6 eventos para entrenamiento y 2 para validación, ya que 2 es menor que el 50% de 6. En el caso de 8 eventos se llega a un valor de 182 muestras posibles de ser analizadas.

b) Tiempo de pronóstico y armado de archivo para entrenamiento: el programa pregunta al usuario el tiempo de pronóstico deseado y crea un único archivo con todos los eventos que fueron seleccionados para el entrenamiento (en función del generador de muestras).

c) Preparación de matrices y vectores y resolución del sistema de ecuaciones: se calculan los términos de las matrices y los vectores según la familia funcional elegida (polinómica, exponencial o de Fourier); las ecuaciones se resuelven mediante una subrutina que utiliza la técnica de la eliminación Gaussiana y el máximo pivote desarrollada por Chapman.

d) Generación de archivos de salida durante el entrenamiento: se generan dos archivos de salida, un archivo donde se especifican los términos de la matriz de resolución, el vector de términos independientes y el vector de coeficientes resultado; y un segundo archivo con los niveles en la estación limnimétrica en estudio según el tiempo de pronóstico seleccionado (desde 15 minutos hasta 6 horas).

e) Cálculo de niveles en etapa de validación del modelo: utilizando el vector de coeficientes resultado obtenido en el entrenamiento, son calculados los niveles limnimétricos empleando las diferentes submuestras de los eventos que fueron considerados para validación.

Filtrado, ordenamiento y selección de muestras

Un primer filtrado que se realizó sobre los niveles calculados, tanto en entrenamiento como en validación, fue implementar el uso de dos tests de significancia para analizar los siguientes parámetros estadísticos: el coeficiente de correlación r , la pendiente a y la ordenada al origen b de la recta de ajuste entre valores medidos y calculados, y el desvío medio. Para el primer estadístico se evalúa si el mismo es diferente de 0 a través de la aplicación del test t de Student ($\alpha=0,05$). Para los parámetros de la recta de ajuste se cuantifica si la pendiente a y la ordenada al origen b , en forma simultánea, no son diferentes de 1 y 0 respectivamente en función del test de Dent-Blackie para $\alpha=0,05$ (Tedeschi, 2004). Finalmente, se verifica si el desvío medio no es diferente de 0 mediante la aplicación del test t de Student para $\alpha=0,05$ (Jacovides y Kontoyiannis, 1995). De esta forma, las muestras que a través de los procesos de entrenamiento y de validación arrojan resultados que 'superan' estos tests de significancia son consideradas como muestras viables o válidas.

Posteriormente, al subconjunto de niveles calculados (entrenamiento y validación) derivado de las muestras filtradas por la metodología anterior, se les calculó una serie de parámetros estadísticos con el fin de seleccionar para cada una de las siete familias funcionales y a su vez para diferentes grados de aproximación (desde 1 hasta 9), cuál era la muestra que presentaba mejor 'comportamiento' tanto en entrenamiento como en validación en forma conjunta. Los parámetros calculados fueron: 1) diferencia en exceso entre niveles medidos y pronosticados $dmáx$ (pronosticado hace referencia tanto a la etapa de entrenamiento del modelo

como a la de validación), 2) diferencia en déficit entre niveles medidos y pronosticados $dmín$, 3) diferencia en el pico entre niveles medidos y pronosticados $dpico$, 4) pendiente de la recta de regresión entre valores medidos y pronosticados a , 5) ordenada al origen de la recta de regresión entre valores medidos y pronosticados b , 6) coeficiente de correlación concordancia rc (propuesto por Lin y citado en Tedeschi, 2004), 7) coeficiente de determinación del modelo CD (Loague y Green, 1991), 8) coeficiente de confianza CC , 9) eficiencia del modelo E_{NS} , propuesto por Nash y Sutcliffe, 10) raíz del error cuadrático medio $RMSE$, 11) desvío medio MBE y 12) coeficiente del error igual a la razón entre el valor medio de las diferencias absolutas respecto al valor medio de los valores medidos C , propuesto por Klepper y Rouse (Medina-Peralta et al., 2010).

Una vez calculados los estadísticos, las muestras válidas para cada una de las familias funcionales y para cada grado de aproximación son ordenadas en forma jerárquica otorgando a cada estadístico el mismo peso. El procedimiento de ordenamiento es el siguiente: se calcula el primer estadístico para cada muestra y se ordenan las mismas, en este caso desde la de menor diferencia hasta la mayor, otorgando a la primera un índice de valor 1 y a la última un índice de valor igual a la cantidad de muestras analizadas. Así se procede sucesivamente con cada estadístico de manera tal que a cada muestra le quedan asignados 12 índices; estos índices son sumados y cada muestra obtiene así un único número índice. La muestra que posee el menor número índice es la que ‘mejor respuesta’ presenta en función de los doce estadísticos elegidos. Cabe aclarar que este procedimiento se ejecuta para cada tiempo de pronóstico o delta T (un total de 9 delta T) y para cada grado de aproximación de la familia funcional (un total de 9 juegos), por lo que para cada una de las familias funcionales se tiene un total de 81 muestras. Luego de efectuados los ordenamientos para una estación limnimétrica se contabiliza cuál es la muestra que más veces se repite en las 567 ocasiones (81 muestras por 7 familias funcionales), y dicha muestra es a partir de la cual son comparadas las diferentes familias funcionales para diversos grados de aproximación de las mismas. Para esta muestra ‘testigo’ es que se efectúan las gráficas de los diferentes estadísticos y de los niveles medidos y calculados, tanto en la etapa de entrenamiento como en la de validación de los diferentes modelos.

RESULTADOS

Se detallan algunos resultados obtenidos en las estaciones limnimétricas RTU 07 Villada y RTU 08 Bigand. Por cuestiones de espacio los tiempos de pronóstico presentados son solamente 15 minutos, 1 hora y 6 horas; mientras que el grado de aproximación también fue limitado a seis (de los nueve analizados). La Figura 7 corresponde al análisis efectuado sobre la muestra número 67 formada según la siguiente combinación 1 2 3 0 0 0 7, la cual indica que de los siete eventos registrados cronológicamente en dicha estación, los eventos 1, 2, 3 y 7 se utilizan en la etapa de entrenamiento y los eventos 4, 5 y 6 se emplean en la validación. En la Tabla 2, se indican nueve estadísticos (de los doce calculados), en función de los diferentes tiempos de pronóstico; y también se indica el modelo (familia funcional y grado) que mejor desempeño tiene para cada tiempo de pronóstico. En forma similar, la Figura 8 y la Tabla 3 corresponden a la estación RTU 08 Bigand (muestra analizada número 68).

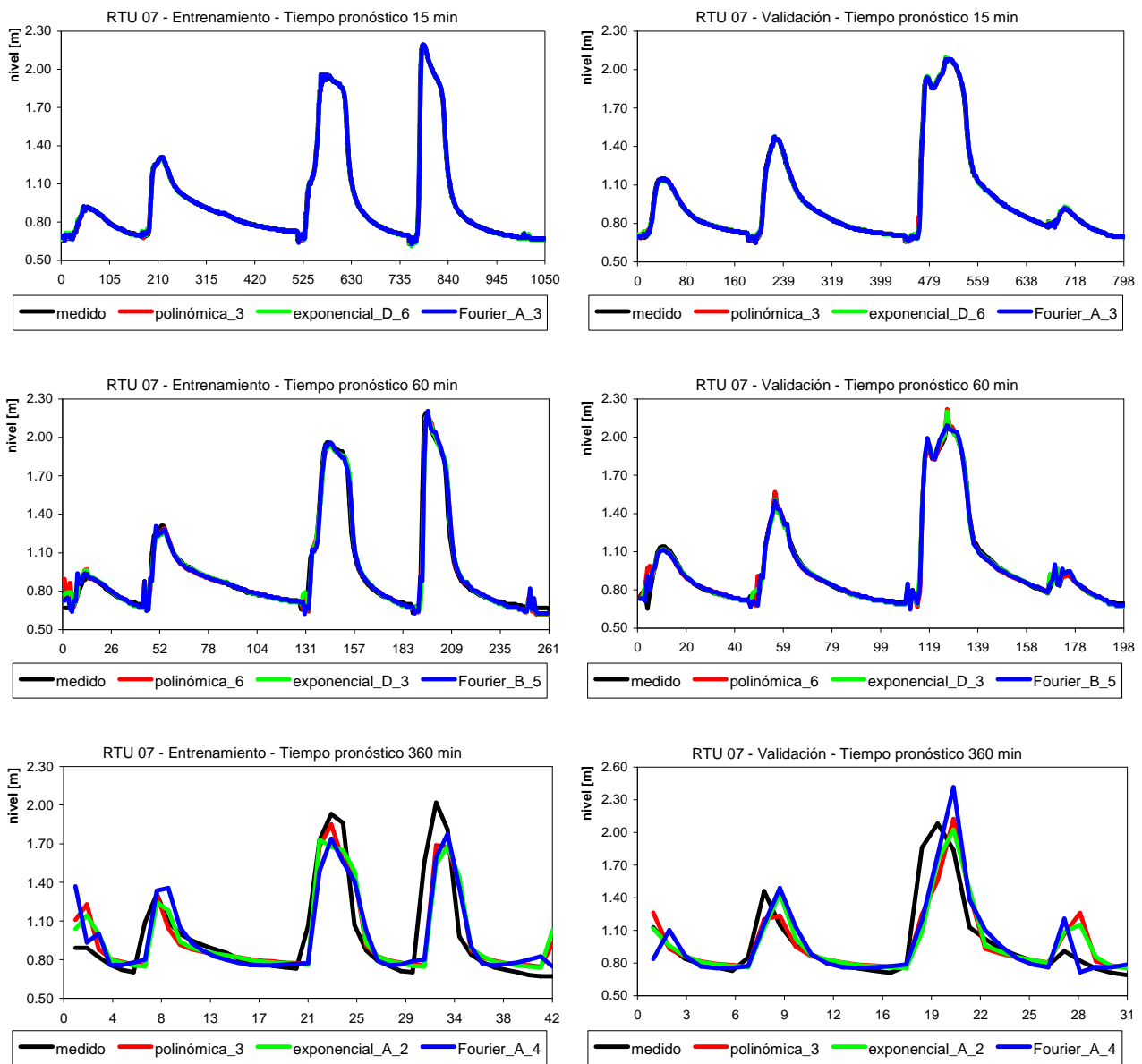


Figura 7.- RTU 07. Villada, muestra 67. Niveles medidos y pronosticados para 15 min, 1 hora y 6 horas.

Tabla 2.- RTU 07. Villada. Estadísticos de los modelos de mejor desempeño en función del tiempo de pronóstico.

muestra 67	modelo	dm _{máx} (%)	dpico (%)	a	b	rc	CC	E_NS	RMSE	MBE
15 min.	pol_3	11,8	0,2	1,001	-0,001	0,999	0,998	0,997	0,021	0,000
	expo_D_6	11,1	0,3	1,000	0,000	0,999	0,998	0,997	0,019	0,000
	Fourier_A_3	11,6	0,2	1,000	0,000	0,999	0,998	0,997	0,020	0,000
1 hora	pol_6	25,9	0,8	0,997	0,001	0,982	0,974	0,965	0,068	-0,002
	expo_D_3	25,2	1,0	0,999	-0,001	0,984	0,976	0,968	0,066	-0,002
	Fourier_B_5	29,6	1,8	0,995	0,005	0,983	0,975	0,967	0,067	0,000
6 horas	pol_3	34,6	20,7	0,989	0,007	0,827	0,762	0,703	0,203	-0,005
	expo_A_2	38,8	20,0	0,979	0,016	0,807	0,736	0,672	0,214	-0,004
	Fourier_A_4	34,9	17,5	0,899	0,094	0,819	0,747	0,661	0,217	-0,006

Discusión y análisis de los resultados

Considerando los resultados presentados para la estación RTU 07 Villada, se aprecia en la Figura 7 que para un tiempo de pronóstico de 15 minutos, la diferencia entre familias funcionales es prácticamente nula, esto se corrobora cuando analizamos la Tabla 2, en que para este tiempo de pronóstico los estadísticos permanecen casi invariantes. A partir de ello surge la idea de que es indistinto utilizar cualquiera de las familias funcionales, sin embargo si se prioriza la sencillez de cálculo, la función polinómica de grado 3 es la que presenta mayor simpleza y por lo tanto requiere un menor tiempo de cálculo. Si ahora nos centramos en el tiempo de pronóstico de 1 hora, a partir de la Figura 7 vemos que comienzan a aparecer algunas diferencias sobre todo en los picos de los limnigramas, zona donde se demanda mayor precisión en los resultados. Cuando buscamos en la Tabla 2 el sector para tiempo de pronóstico de 1 hora, se observa que el modelo de Fourier presenta una leve desventaja respecto a los restantes, aunque en términos prácticos no es relevante. Si consideramos el factor sencillez, nuevamente la función polinómica (esta vez de grado 6) aparece como la mejor opción. Analizando ahora para un tiempo de pronóstico de 6 horas, se observa en la Figura 7 una mayor diferencia entre valores medidos y calculados; en entrenamiento las funciones parecen no poder alcanzar los picos de los limnigramas, mientras que en la etapa de validación parecería que contrariamente los picos calculados superan a los medidos. Si nos referimos a la Tabla 2, puede apreciarse que para tiempos de pronóstico de 6 horas, por ejemplo, las diferencias en el pico del limnograma alcanzan valores del orden del 20%, cantidad que contrasta fuertemente con los valores para los tiempos de pronóstico anteriores (hasta 2%). En función de los estadísticos presentados se advierte una pequeña superioridad de la función polinómica de grado 3 por sobre las demás.

Efectuando un análisis similar, para la estación RTU 08 Bigand, se advierte que para un tiempo de pronóstico de 15 minutos los niveles medidos y calculados no difieren visualmente (Figura 8). Si se recurre a la Tabla 3, se aprecia que los estadísticos permanecen casi invariantes. Cuando se realiza el análisis para un tiempo de pronóstico de 1 hora, se localizan nuevamente algunas diferencias en el pico de los limnigramas. Si ahora se considera un tiempo de pronóstico de 6 horas, una diferencia notoria es que tanto la función polinómica como la de Fourier (ambas hasta grado 6) no consiguen generar resultados satisfactorios para pasar los tests de significancia; razón por la cual no aparecen en los resultados. Respecto de la función exponencial existe un desfase temporal entre valores medidos y calculados y conjuntamente con esto una baja sustancial en los estadísticos si los comparamos con la estación RTU 07 Villada.

Sintetizando, en ambas estaciones analizadas, se observa que en general la función polinómica es la que mejor se adapta a los valores medidos por sobre las funciones exponencial y de Fourier, sobre todo si se considera el grado de sencillez en su construcción respecto a las restantes. Como es de esperar se verifica una disminución en la calidad de los estadísticos a medida que se incrementa el tiempo de pronóstico. Respecto al grado de las funciones, no se observa que exista preponderancia de alguno sobre otro; sin embargo parecería que para las funciones exponencial y de Fourier, grados 'altos' generan mejores resultados.

CONCLUSIONES Y RECOMENDACIONES

Se destaca el empleo de medidas estructurales en conjunto con medidas no estructurales en la lucha contra las inundaciones, dentro de las segundas se encuentran los sistemas de alerta contra inundaciones que permiten pronosticar el estado del nivel de agua en ciertas secciones características de los cursos de agua; siendo una herramienta de ayuda en la toma de decisiones por parte de las autoridades.

La confiabilidad de los sensores es de vital importancia para el correcto funcionamiento del sistema, demandando un seguimiento constante tanto en la estación central como en el lugar donde se encuentran instalados. Es recomendable ejecutar el mantenimiento preventivo de la red en forma periódica como el mantenimiento correctivo inmediatamente sean detectados problemas en la recepción de los datos. Es importante lograr una serie continua de datos, tanto de precipitaciones como de niveles, para evitar periodos sin datos y aumentar la robustez de los pronósticos.

Los resultados en términos de niveles limnimétricos fueron evaluados para diferentes familias funcionales y varios tiempos de pronóstico a través de diferentes estadísticos. Los mejores resultados, se obtuvieron en general para la función polinómica; aunque no existe una superioridad notable de la misma por sobre la función exponencial o de Fourier. Por otra parte se observa que no existe un grado predominante para las funciones consideradas. A medida que el tiempo de pronóstico aumenta la calidad de los pronósticos disminuye, lo que tiene sentido por ser más dificultoso pronosticar con mayor tiempo de anticipación.

Aunque la cantidad de eventos registrados durante el lapso del estudio no es muy elevada, en media 9 eventos por subcuenca, los resultados obtenidos por esta metodología son alentadores; evidenciando de esta manera la capacidad de ‘aprendizaje’ y reproducción de eventos de este tipo de modelo.

Dado que el modelo empleado es de tipo ‘caja negra’, se observa que para su puesta en funcionamiento no demanda gran cantidad de datos de la cuenca a estudiar; como las características físicas, la topografía, el detalle de las obras de arte presentes, el uso del suelo o la condición de humedad para cada evento. Por lo que su aplicación, resulta relativamente sencilla de llevar a cabo en otras cuencas una vez programada la red funcional. El tiempo de cálculo de este tipo de modelos es ‘corto’ debido a que están basados en la resolución de un sistema de ecuaciones; lo que le otorga cierta ventaja frente a los modelos distribuidos.

En caso de que pueda incrementarse el número de eventos en función de próximas tormentas, su incorporación al modelo de pronósticos es sumamente sencilla dada la arquitectura planteada del software. De esta forma, la mayor cantidad de datos de entrada al modelo permitirá mejorar la robustez del mismo. Un último comentario con relación a los parámetros empleados para evaluar el grado de bondad de los modelos, es que no existe en la actualidad un criterio unánime por parte de la comunidad científica en cuanto a la superioridad de un parámetro sobre otro; siendo deseable profundizar en el estudio de estos índices estadísticos para mejorar el criterio de selección de los modelos.

REFERENCIAS

- Aldana Valverde, A., 2002. Planificación, Diseño y Utilización de Herramientas de Ayuda a la Toma de Decisiones en Tiempo Real: Jornadas sobre sistemas de ayuda a la decisión ante problemas hidráulicos e hidrológicos en tiempo real, CEDEX, Marzo de 2002, pp. 7-16.
- Berga Casafont, L., 1995. *Aplicación de los sistemas de previsión de avenidas. Experiencias y problemática: Jornadas sobre el SAIH en la Gestión Hídrica*. Ministerio de Obras Públicas, Transportes y Medio Ambiente. Dirección General de Obras Hidráulicas. Confederación Hidrográfica del Ebro, Marzo de 1995.
- Bruen, M. y Yang, J., 2005. Functional networks in real-time flood forecasting –a novel application. *Advances in Water Resources*. 18(9), pp. 899-909.
- Campolo, M., Andreussi, P. y Soldati, A., 1999. River flood forecasting with a neural network model. *Water Resources Research*, 35(4), pp. 1191-1197.
- Castillo, E., 1998. Functional Networks. *Neural Processing Letters*, 7, pp. 151-159.
- Castillo, E. y Gutiérrez, J. M., 1998. Nonlinear time series modeling and prediction using functional networks. Extracting information masked by chaos. *Physics Letters A*, 244, pp. 71-84.
- Dueñas Molina, C., 1995. *El SAIH y su aplicación a la Protección Civil: Jornadas sobre el SAIH en la Gestión Hídrica*. Ministerio de Obras Públicas, Transportes y Medio Ambiente. Dirección General de Obras Hidráulicas. Confederación Hidrográfica del Ebro, Marzo de 1995.
- Gruntfest, E., 1995. Long term social and economic impacts of extreme floods: US-Italy Research Workshop on the Hydrometeorology, Impacts and Management of Extreme Floods. Perugia (Italy). November 1995.
- Hsu, K. L., Gupta, V. y Sorooshian, S., 1995. Artificial neural network modeling of the rainfall-runoff process. *Water Resources Research*, 31, pp. 2517-2530.
- Hsu, M., Fu, J. y Liu, W., 2003. Flood routing with real-time stage correction method for flash flood forecasting in the Tanshui River, Taiwan. *Journal of Hydrology*, 283, pp. 267-280.
- INTA, 1983. Instituto Nacional de Tecnología Agropecuaria. Cartas de Suelos de la Provincia de Santa Fe.
- Jacovides, C. P. y Kontoyiannis, H., 1995. Statistical procedures for the evaluation of evapotranspiration computing models. *Agricultural Water Management*, 27, pp. 365-371.
- Kothyari, U. y Singh, V., 1999. A multiple-input single-output model for flow forecasting. *Journal of Hydrology*, 220, pp. 12-26.
- Lekkas, D., Maxey, R. y Wheeler, H., 2004. Intercomparison of forecasting methods for flood warning in the river catchment. *Global Nest: the International Journal*, 5(2), pp. 89-97.
- Loague, K. y Green, R. E., 1991. Statistical and graphical methods for evaluating solute transport models: Overview and application. *Journal of Contaminant Hydrology*, 7, pp. 51-73.
- Medina-Peralta, S., Vargas-Villamil, L., Navarro-Alberto, J., Canul-Pech, C. y Peraza-Romero, S., 2010. Comparación de medidas de desviación para validar modelos sin sesgo, sesgo constante o proporcional. *Universidad y Ciencia. Trópico Húmedo*, 26(3), pp. 255-263.
- Ogden, F., Sharif, H., Senarath, S., Smith, J., Baek, M. y Richardson, J., 2000. Hydrologic analysis of the Fort Collins, Colorado, flash flood of 1997. *Journal of Hydrology*, 228, pp. 82-100.
- Riccardi, G. A., Scuderi, C. M. y Stenta, H. R., 2002. Modelado de escurrimiento superficial con alto nivel de detalle en la cuenca del arroyo Ludueña, Santa Fe, Argentina. *XIX Congreso Nacional del Agua*, T33, V.C. Paz, Córdoba.
- Riccardi, G. A., Zimmermann, E. D., Basile, P. A., Mancinelli, A. M., Scuderi, C. M., Rentería J. P. y García P., 2008. Rehidrología y Modelo de Simulación a Tiempo Real en Sistema de Alerta Hidrológico en las Cuencas de los A° Ludueña y Saladillo. Informe Técnico 1. *Convenio entre el Ministerio del Aguas, Servicios Públicos y Medio Ambiente de Santa Fe y la Facultad de Ciencias Exactas, Ingeniería y Agrimensura de Rosario*. Agosto de 2008.
- Scuderi, C. M., Riccardi, G. A. y Zimmermann, E. D., 2009. Sistema de alerta a tiempo real en la zona del Gran Rosario. Resultados Preliminares. *XXII Congreso Nacional del Agua*, Trelew, Chubut, 11 al 14 de nov., 11 p.
- Singh, R., 2005. *Real time flood forecasting – Indian experiences: International G-WADI Modelling Workshop*, National Institute of Hydrology, Roorkee, India, 28/02 al 05/03 de 2005, 34 pp.
- SMN, 2008. Servicio Meteorológico Nacional. Datos de lluvia correspondientes a la serie 1892-2008.
- Tedeschi, L. O., 2004. Assessment of the Adequacy of Mathematical Models. Department of Animal Science, Texas A&M University, College Station, TX 77843, 29 pp.