

**MODELACION, SIMULACION Y PREDICCION DEL COMPORTAMIENTO DINAMICO DEL RIO  
SAN JUAN USANDO REDES NEURONALES**

**H. Daniel Patiño<sup>1</sup>, Cristian Rodríguez Rivero<sup>2</sup>, Julián Pucheta<sup>3</sup>**

<sup>1</sup>*Instituto de Automática, Fac. de Ing., Universidad Nacional de San Juan  
dpatino@inaut.unsj.edu,*

<sup>2</sup>*Facultad de Ingeniería – Instituto Universitario Aeronáutico, Córdoba  
cristian.rodriguezrivero@gmail.com,*

<sup>3</sup>*LIMAC, Fac. de Ciencia Exactas, Físicas y Naturales, Universidad Nacional de Córdoba  
julian.pucheta@gmail.com*

**Resumen:** La estimación de la predicción del escurrimiento hídrico anual (PEHA) de la cuenca de un río, principalmente en zonas áridas y semiáridas, es información de suma importancia para conocer con antelación la disponibilidad de agua para el consumo residencial, el requerimiento para el riego en la agricultura, el requerido por la industria y la generación de energía hidroeléctrica. Obtener un buen pronóstico es un insumo importante en la planificación, gestión, administración y optimización de todos los recursos del sistema hidráulico disponible, y también, para la planificación, construcción y operación de represas y centrales hidroeléctricas a lo largo de la cuenca, en un contexto de mediano-largo plazo. En general, los modelos de PEHA propuestos en la literatura se basan principalmente en medidas de la acumulación nivea, geología de la cuenca y evolución de la dinámica climática de la misma, variables más significativas que influyen en el proceso del derrame hídrico diario, mensual y anual. Sin embargo, estos modelos en general presentan algunos problemas relativos a su calibración, disponibilidad de mediciones de todas las variables involucradas en el proceso, dependiendo de la infraestructura de medición disponible.

En este trabajo se propone la modelación y estimación del PEHA del río San Juan de Argentina basado en redes neuronales artificiales. El modelo computacional se realiza para un horizonte de predicción de un paso hacia delante basado en la aproximación de un modelo auto-regresivo no lineal (NAR) mediante una red neuronal de Fourier (RNF), ajustado mediante un filtro de Kalman extendido (FKE) con datos de la medición del caudal en m<sup>3</sup>/s. El sistema tiene la ventaja de requerir sólo como datos de entrada la serie temporal histórica del caudal y permite su extensión para un pronóstico de mediano y largo plazo (3-6-12 pasos hacia delante).

El sistema de modelación y predicción propuesto se simula computacionalmente empleando la serie temporal histórica del río, disponible desde 1909 hasta 2012. El desempeño del mismo es evaluado tanto en el proceso de aprendizaje del modelo como en el de test-validación, usando como métricas los valores de sMAPE, RMSE y función de correlación. Los análisis y resultados obtenidos muestran una muy buena performance del sistema en el modelado computacional y una buena estimación de la PEHA en un año hacia delante de la cuenca del río San Juan.

Palabras claves: Inteligencia computacional, redes neuronales de Fourier, filtro de Kalman extendido, predicción de series-temporales, sistemas no lineales, hidroinformática.

## 1. INTRODUCCION

La hidrología en términos generales puede definirse como la ciencia que se dedica al estudio del ciclo hidrológico o ciclo del agua en el planeta, al estudio de su distribución espacial y temporal, y a sus propiedades presente en la atmósfera y corteza terrestre. Esto incluye también a las precipitaciones fluviales, el escurrimiento néveo, humedad del suelo, evapotranspiración y equilibrio de las masas glaciares. En la actualidad la hidrología juega un papel muy importante en la planificación y optimización en el uso y operación de los recursos hidráulicos disponibles, y ha llegado a convertirse en parte fundamental de los proyectos de ingeniería que tienen que ver con el suministro de agua potable, disposición de aguas servidas, drenaje para evitar inundaciones, protección contra la acción de ríos y recreación.

Una de las etapas más importantes en este ciclo, principalmente tenidas en cuenta en regiones áridas y semiáridas, es el proceso de escurrimiento producido por el deshielo y las precipitaciones que contribuyen al caudal de un río. La estimación anual del derrame hídrico de un río, donde la principal fuente de agua proviene de la cantidad de nieve acumulada en alta montaña, es muy importante para la administración y gestión de todo el sistema hidráulico encargado de proveer agua para el consumo de la población, para el riego en la agricultura, para su uso en la industria y para la generación de energía hidroeléctrica, inclusive también para la recreación. Por lo tanto, disponer de un buen pronóstico anual es fundamental en el diseño de estrategias de operación óptimas de todo el sistema hidráulico como así también para la toma de decisiones para la distribución más apropiada para cada uno de los sectores. Esta es una de las razones principales del porqué de la necesidad de obtener un buen pronóstico de los flujos de los ríos en el corto plazo, pero también es importante para la planificación de mediano-largo plazo relativo a requerimientos de obras civiles necesaria para la construcción de represas, la generación de energía hidroeléctrica y el almacenamiento para épocas de sequías, etc. Sin embargo, el problema de realizar un pronóstico, en sí mismo, no es un problema trivial debido a diversos mecanismos y variables físicas-ambientales-geográficas que intervienen en todo el proceso que rigen en el caudal de los ríos; dinámicas que actúan sobre una amplia gama de escalas tanto a nivel temporal como espacial.

En general la solución al problema del pronóstico del escurrimiento anual de los ríos involucra un complejo proceso de modelación matemática que se obtiene a partir de las leyes físicas que gobiernan sus subprocesos. Es un problema no lineal de elevada

complejidad por la gran cantidad de variables intervinientes y sus complejas relaciones. Los métodos de modelación y predicción se pueden agrupar en dos grandes grupos principales; uno que considera las leyes físicas involucradas basados en modelos conceptuales, y el otro, basado en modelos de caja negra. Los modelos conceptuales basados en las leyes de la física están diseñados específicamente para simular matemáticamente el comportamiento dinámico (o estático) de los procesos que gobiernan el ciclo hidrológico del río. Estos modelos generalmente incorporan formas simplificadas de las leyes físicas, simplifica o reducen el número de variables intervinientes. Son generalmente modelos no lineales, invariantes en el tiempo, determinísticos, de estructura conocida y parámetros representativos de las características de la cuenca (Hsu, et al.; 1995). En (Kitanidis, and Bras; 1980) se afirma que estos modelos conceptuales del escurrimiento o derrame hídrico de una cuenca son bastante confiables respecto a su modelación matemática, pues involucran la mayoría de las características más importantes de la hidrografía. Sin embargo, su implementación y calibración puede presentar muchas y diversas dificultades, como el requerimiento de herramientas matemáticas sofisticadas, una cantidad significativa de datos de calibración (muchas veces no disponibles), y requerimiento de cierto grado de especialización y experiencia con dicho modelo, (Hsu, et al.; 1995), (Duan, et al.; 1992), (Duan, et al.; 1994), (Sorooshian, et al.; 1993) y (Yapo, et al.; 1996). Se debe considerar también que puede resultar costoso y difícil de medir todas las variables involucradas en el modelo, que como ejemplo se puede citar el caso de la colocación de sensores y estaciones meteorológicas en lugares prácticamente inaccesibles en alta montaña, requiriendo una logística importante en su mantenimiento.

En particular, para las cuencas hidrográficas de montaña, el comportamiento dinámico del caudal de los ríos depende fuertemente del número de glaciares, la acumulación de nieve anual, las lluvias, la topología de su geología, de las dimensiones de la cuenca, la dinámica del clima, etc.; por lo que todos estos factores deben tenerse en cuenta en el modelo conceptual. Los modelos conceptuales además tienen el inconveniente de cómo se incorporan las periodicidades de algunas variables no siempre evidentes, que con frecuencia se enmascaran adicionando al modelo una señal de ruido. Tradicionalmente se han propuesto modelos estocásticos para la modelación y predicción del derrame anual de un río basado en sistemas lineales, como: los modelos Auto Regresivos (AR); de Media Móvil Auto Regresivo con Entrada eXógena (conocido por las siglas ARMAX), (Salas, et al.; 1985),

(Mishra y Abdelhadi; 2004); y regresión no lineal (Chang y Hwang; 1999). La aplicabilidad de estos modelos se limita en general a cuencas con escurrimientos que se han medido durante largos periodos de tiempo y donde no se han producido cambios significativos en las condiciones climáticas y geológicas de las cuencas hidrográficas (Battiti; 1999). La mayoría de los modelos de simulación que surgen de este enfoque resultan de una estructura importante con muchos parámetros desconocidos. En consecuencia, son difíciles, si no prácticamente imposible, su identificación, estimación y validación en términos estadísticos rigurosos, debido principalmente a problemas relacionados con el exceso del número de parámetros y a la falta de datos experimentales.

En ausencia de datos, o de información exacta acerca de los mecanismos físicos subyacentes en todo el proceso hidrológico, o por desconocimiento de las ecuaciones exactas involucradas en la dinámica del derrame hídrico del río, el empleo de un modelo de caja negra tiene ventajas sobre un modelo de base física; ya que este modelo es capaz de representar arbitrariamente el proceso del caudal del río relacionando solamente las entradas con las salidas del sistema sin importar las leyes físicas que las relacionan. Este modelo de caja negra, si bien no conduce necesariamente a una mejor comprensión del comportamiento del proceso dinámico del caudal del río (de una manera física realista), el mismo es fácil de obtener, incluso para diferentes condiciones ya que el proceso de modelado y predicción es casi análogo. Estos modelos fueron desarrollados originalmente basados en la teoría de procesamiento de señales, teoría de sistemas e inteligencia computacional. El uso de redes neuronales (RN) para la modelación y predicción de caudales de ríos y en la estimación de la cantidad de agua acumulada por lluvias ganó mucha popularidad en los últimos años de los noventa – principios del 2000, (Thirumalaiah and Deo; 1998), (Imrie; 2000), (Shamseldin, et al.; 2002), (Sivakumar, and Jayawardena; 2002), (Dolling, and Veras; 2002). El desempeño de esos modelos de tipo NARMAX basados en RN para la predicción del caudal de ríos, es comparable a los modelos conducidos por datos o NAR, probado en (Sivakumar, and Jayawardena; 2002) y (Kisi, 2005). En general, en todos los trabajos se emplean perceptrones multicapa estáticos con dos o tres capa de neuronas para realizar la aproximación de un modelo de caja negra de tipo NARMAX. Sin embargo, existe aún cierta resistencia de su empleo (Young y Chotai; 2001), debido a que presentan los mismos inconvenientes que los modelos de base física, en relación a la necesidad de conocer todas las mediciones de las variables involucradas en las entradas/salidas del sistema.

Una solución alternativa a los problemas antes mencionados es el empleo de un modelo de caja negra Auto Regresivo No Lineal (NAR), también conocido como modelos conducidos por los datos, para el cual es necesario sólo conocer la serie temporal histórica del derrame hídrico del río. En este modelo se supone que en la variable de estado caudal se concentra y refleja toda la historia, efectos y dinámicas de todas las variables involucradas en el proceso, es decir, modelado como un proceso Markoviano. Este modelo de modelación y predicción del comportamiento dinámico del caudal permite además, resolver el problema de predicción de múltiples pasos hacia adelante, problema que no pueden resolver los anteriores modelos.

En este trabajo se propone una estructura y ley de aprendizaje para la modelación y estimación de la predicción del escurrimiento anual de un río basada en una red neuronal de Fourier (RNF), la cual explota las características de estacionalidad del comportamiento típico de las series temporales de los caudales de los ríos. La predicción se realiza con un horizonte de un paso hacia adelante y se emplea un filtro de Kalman extendido (FKE) como algoritmo de ajuste de los parámetros de la red durante el proceso de aprendizaje. El modelo pertenece a la categoría NAR, el cual emplea solamente la serie temporal del caudal diario del río. Este procedimiento de modelado y predicción es aplicado a la previsión anual del derrame hídrico anual de la cuenca del Río San Juan, en San Juan, Argentina. Para ello se emplea su serie temporal histórica disponible desde 1909 al 2012.

## 2. FORMULACION DEL PROBLEMA DE MODELACION Y PREDICION DE SERIE TEMPORAL DEL DERRAME HIDRICO ANUAL DEL RIO SAN JUAN

### 1.1 Descripción y estudio de la Cuenca de Estudio

La cuenca de estudio elegida es la cuenca del río San Juan, que se encuentra en la provincia de San Juan en el sector occidental de la República Argentina, que fluye de oeste a este a lo largo de la Cordillera de los Andes Central. San Juan es una región árida-semiárida, y por lo tanto, la estima del escurrimiento hídrico anual, proveniente principalmente de precipitaciones níveas y glaciares en alta montaña, es de suma importancia para la planificación de los recursos hidráulicos disponibles y el manejo y distribución óptima del recurso para el consumo humano, el riego en la agricultura, el requerido por la industria, recreación, y la generación de energía hidroeléctrica.

El río nace en la localidad de Calingasta, en el sudoeste de la provincia, a partir de la confluencia de tres ríos principales: el río Calingasta, el río Castaño Viejo y el Río los Patos. Estos ríos nacen en

una altitud de aproximadamente 4.000 m sobre el nivel del mar en el oeste y suroeste de la provincia. Tiene una longitud de unos 500 Km, una descarga hídrica o módulo promedio de 56.41m<sup>3</sup>/s, y su cuenca tiene una superficie de unos 20.000 km<sup>2</sup> aproximadamente. Como se dijo, este río tiene su origen en la Cordillera de los Andes, con una elevación media de alrededor de 3.500 m con varios picos que alcanzan más de 6.000 m. Este río sirve a una población de unas 600.000 personas en la provincia de San Juan. Con precipitaciones promedio por año de menos de 150 mm, la gran mayoría de la agricultura intensiva depende de su riego provisto de este afluente. El escurrimiento de nieve es la principal fuente de agua en la cuenca. El derrame de agua óptimo para el consumo anual del Valle de Tulúm (Ciudad de San Juan y gran San Juan) es de aproximadamente de 2080.46 hm<sup>3</sup>, distribuidos en: 91,92% para la agricultura; 4,28% para la generación de energía hidroeléctrica; 2,8% para el consumo residencial; 0,96% para la minería; 0,03% para la recreación; y 0,01% para las industrias (Dirección de Hidráulica, Relevamiento Agrícola en la Provincia de San Juan, Ciclo 2006-2007). Este patrón es consistente con la distribución encontrada en los países en vías de desarrollo (Molden; 2007).

### 1.2 Datos y Métodos

En el presente trabajo, se usa la serie temporal del caudal diario del río San Juan medido en la estación de aforo Km 47,3 por la Dirección de Hidráulica, dependiente del Ministerio de la Producción y Desarrollo Económico de la provincia de San Juan. Los datos de esta serie temporal están disponibles desde 1909 hasta 2012. Los mismos son empleados para calcular el escurrimiento anual acumulado de agua necesario para la obtención del modelo y predicción propuesta. A partir del caudal medio diario  $q$  en [m<sup>3</sup>/s], puede calcularse el derrame hídrico diario, en [hm<sup>3</sup>/día], de acuerdo a la siguiente ecuación,

$$Q(t) = \frac{q \cdot 86400}{10^6} \quad 1)$$

dónde,  $Q(t)$  es el derrame hídrico acumulado promedio en el día. Luego, realizando la sumatoria de todos los días del año, es posible obtener el acumulado hídrico anual de la cuenca, como:  $Q = \sum q$  [hm<sup>3</sup>]. De esta manera es posible obtener una serie temporal de 102 valores correspondiente a los valores acumulados anualmente. En la Fig. 1 se muestra la serie temporal del derrame hídrico acumulado anual del río San Juan para los datos históricos del río desde 1909 al 2012.

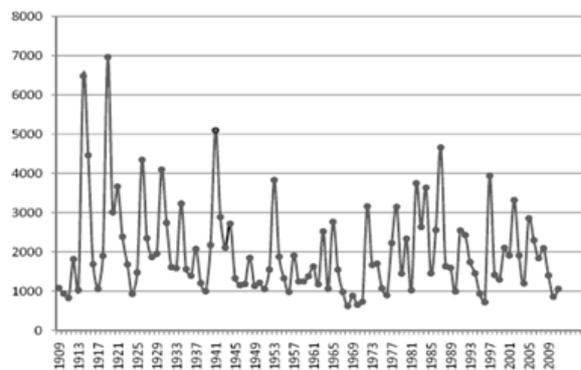


Fig. 1: Serie temporal del derrame hídrico anual en [hm<sup>3</sup>] del río San Juan.

En la Figura 2 se muestra la serie temporal del valor mínimo acumulado en el período antes mencionado y su tendencia lineal, información considerada importante a tener en cuenta para la planificación a largo plazo. Puede verse que su valor se mantiene aproximadamente constante en los 1.300 hm<sup>3</sup>. Puede además observarse que el valor mínimo del derrame hídrico anual fue de 617.9 hm<sup>3</sup>, correspondiente al año 1968, período de aproximadamente de 5 años de sequía.

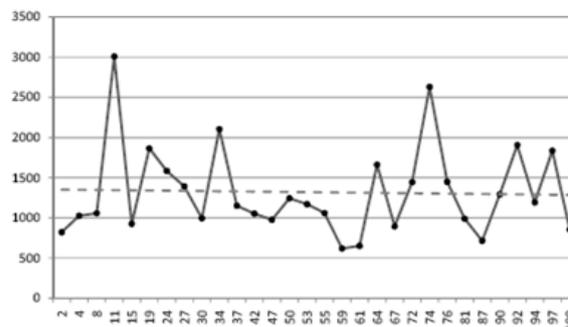


Fig. 2: Serie temporal del derrame hídrico manual en [hm<sup>3</sup>] del río San Juan.

### 3. FORMULACION DEL PROBLEMA DE MODELACION Y PREDICCIÓN

El objetivo de la estimación de la predicción en series temporales puede establecerse de la siguiente manera: teniendo en cuenta una serie del tiempo de la variable de interés, y considerando que sus valores proceden de un sistema dinámico, se desea obtener la predicción de los valores futuros en un paso hacia delante basado en los valores presentes y anteriores de la serie temporal. En otras palabras, se trata de encontrar un modelo computacional, de tipo NAR, que pueda ser capaz de modelar y predecir el comportamiento del sistema en un paso hacia adelante. El valor presente y pasados de la serie temporal se utilizan como entradas del modelo de predicción. Cuando se requiere del pronóstico en un paso hacia delante se conoce este problema, en

general, como la predicción a corto plazo. Mientras que si se requiere conocer una estimación de la predicción con un horizonte de varios pasos hacia delante, el problema de predicción es de mediano-largo plazo (3-6-10 años). En este trabajo se propone un modelo para la predicción con un horizonte de un paso hacia delante del valor del derrame hídrico anual para la cuenca del río San Juan.

La estructura propuesta para aproximar el modelado y la predicción de la serie temporal del derrame hídrico anual del río San Juan en un paso hacia delante se basa en un sistema no lineal de caja negra de tipo NAR, representado mediante la siguiente expresión:

$$\hat{y}_{t+H} = f(y_t, y_{t-1}, y_{t-2}, \dots, y_{t-M+1})$$

dónde, H es el horizonte de predicción, M denota el número de entradas o longitud de la cadena de retardos de la entrada del modelo, y  $f \in \mathbb{R}$  es una función no lineal a ser aproximada. El regresor para este modelo se define como el vector de las entradas  $y_t, y_{t-1}, y_{t-2}, \dots, y_{t-M+1}$ , el cual puede ser completo o incompleto, según estén todas o falten algunas de sus componentes. Es también posible considerar otras variables exógenas como entradas en el regresor, pero que no se han considerado aquí. La ventaja de este enfoque propuesto es que el modelo no depende de otras variables como, acumulación nívica en la cuenca, otras variables dependientes de la geología de cuenca, variables climáticas, número de glaciares, acumulación de agua por precipitaciones, etc., las cuales como se dijo, presentan, en general, serias dificultades de medición y calibración en modelos tipo NARMAX. En el presente trabajo, se propone un modelo de tipo NAR de predicción considerando un horizonte de un paso hacia delante, el cual es aproximado por una RNF de dos capas de neuronas. El proceso de aprendizaje se realiza mediante la serie temporal usando un FKE como algoritmo de aprendizaje.

#### 4. RECONSTRUCCIÓN DEL ESPACIO DE ESTADO DEL MODELO PREDICTOR NAR

La principal pregunta que se puede hacer respecto a la propuesta de emplear un modelo de tipo NAR como modelo del sistema es la siguiente: "¿Cómo es posible reconstruir el modelo dinámico completo del sistema, Figura 3, midiendo solamente la variable de interés de la serie temporal ( $y_t, y_{t-1}, y_{t-2}, \dots$ )?".

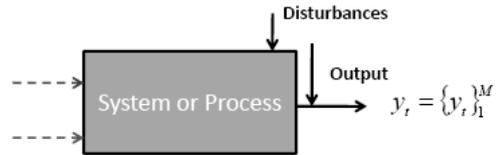


Fig. 3: Representación del modelo NAR del predictor.

Hay dos teorías principales que dan respuesta a esta pregunta. Una teoría proviene de la teoría de series temporales generadas por sistemas dinámicos lineales, propuesta por J.C. Willems, (Willems; 1986-1987), y las representaciones en ecuaciones de estado de la serie temporal (Markovsky, et al.; 2005). La segunda respuesta proviene de la teoría de los sistemas no lineales caóticos propuesta por el Teorema de Taken y Mañé, (Taken; 1980) y (Mañe; 1980), basados en el denominado teorema embebido o conocido en Inglés como Embedding Theorem. En este trabajo, se emplea una RNF para aproximar la función  $f(\cdot)$  de (2), que como entradas es la línea de retardo de  $y$  de la serie temporal. Es bien sabido que una RN puede aproximar cualquier mapeo no lineal estático con un error arbitrario mediante una RN de tipo feedforward de dos capas, con un número suficientemente grande de neuronas y sobre un dominio compacto de un espacio normado de dimensión finita (Funahashi; 1989), (Hornik, et al.; 1989), y (Hunt, and Sbarbaro; 1992). Con esta propuesta, además, el empleo de RN como aproximador explota la característica de aprendizaje a través de datos obtenidos experimentalmente. En la Figura 4 se muestra un modelo de auto-regresión dinámico no lineal basado en la red neuronal.

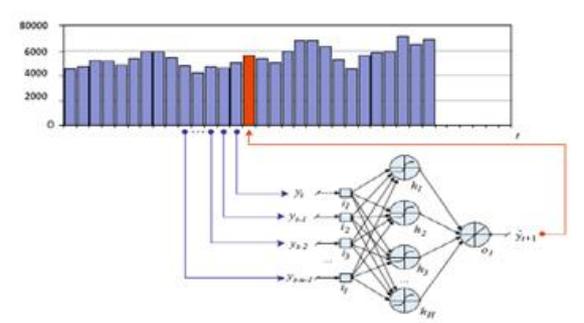


Fig. 4: Esquema del predictor en un paso hacia delante basado en RN.

#### 5. MODELO DE PREDICCIÓN BASADO EN RNF

Del análisis de series de Fourier es bien conocido que cualquier función periódica (o señal periódica) puede descomponerse como una suma infinita de términos de funciones trascendentales de tipo seno y coseno (serie de Fourier),

$$f(x) = \sum_{n=0}^{\infty} [a_n \cos(nx) + b_n \sin(nx)] \quad 2)$$

la cual puede ser reescribirse como la suma de términos de tipo seno de la forma,

$$f(x) = a_0 + \sum_{n=1}^{\infty} c_n \sin(nx + \theta_n) \quad 3)$$

dónde,  $c_n = \sqrt{a_n^2 + b_n^2}$ , y  $\theta_n = \arctan(b_n/a_n)$ .

La ecuación (3) puede ser vista como una red neuronal estática de tipo feedforward estática, con una capa de entrada de nodos de distribución para  $x$ , una neurona simple de salida, para  $f(x)$ , y una capa oculta de neuronas con funciones de activación de tipo  $\sin(\bullet)$ . El parámetro  $a_0$  correspondería a la señal de desviación o polarización de la neurona de salida; los coeficientes  $c_n$  corresponderían a los pesos de las sinapsis de la neurona de salida; el parámetro de fase  $\theta_n$  correspondería al término de desviación o polarización de las neuronas correspondientes a la capa oculta; y  $n$  a los pesos o sinapsis entre los nodos de entrada y los de la capa oculta. La diferencia básica entre el enfoque de Fourier y el enfoque de la RN como aproximadores, es que en el enfoque de Fourier  $n$  son números enteros y los coeficientes deben obtenerse analíticamente; mientras que en la RNF estos "pesos" entre los nodos de la capa de entrada y el de las unidades ocultas pueden tomar cualquier valor fijados por la ley de aprendizaje.

Debido al patrón o característica de estacionalidad que presentan las series temporales de caudales de los ríos, es que puede explotarse y usarse a la serie de Fourier como aproximador natural de este tipo de series, y por ende, emplearse como predictor en el modelo propuesto. En 1976, este planteo fue aplicado por primera vez para el análisis de series temporales en hidrología por (Bloomfield; 1976). Afshar y Fahmi en (Afshar y Fahmi; 2001) proporcionan un modelo para predecir la cantidad de agua aportado por las lluvias mediante la combinación de modelos de Fourier con modelos ARIMA. En (Zuo, and Cai; 2005), (Zuo, et al.; 2009) y (Cordova y Yu; 2009) se han propuesto sistemas de identificación y control de procesos basados en RNF. En (Huang, and Babri; 1998) se prueba que una RN estática puede aproximar cualquier mapeo no lineal, con cualquier precisión deseada, mediante una topología de dos capas, con un número suficientemente grande de neuronas y con funciones de activación de base radial, sinusoidal, coseno, exponencial, y demás funciones simétricas similares.

En este trabajo, se propone una red neuronal de Fourier (RNF) para la modelación y predicción de la serie temporal del derrame hídrico anual del río San Juan, a partir de su serie temporal histórica de 1909 al 2012. La predicción se realiza para un horizonte de predicción de un solo paso hacia adelante, considerando la aproximación de un modelo de tipo NAR con la RNF. En comparación con otros modelos tradicionales de aproximación de la modelación y predicción de series temporales basados en RN, una característica sobresaliente de esta estructura (RNF) es que podría extenderse la metodología al análisis y diseño en el dominio frecuencial.

La RNF propuesta tiene una topología estructurada descrita por la siguiente ecuación,

$$y_{t-1} = b_0 + \sum_{i=1}^N b_i \sin(W_i \cdot X) \quad 4)$$

dónde,  $b_0$  son los pesos de la neurona de salida,  $N$  es el número de neuronas de la capa oculta,  $W_i$  es un vector de pesos de las entradas en la capa de entrada,  $[w_{i0}, w_{i1}, w_{i1}, \dots, w_{iN}]$ , y  $X$  es el vector de entrada o regresor de la función  $f$ ,  $[1, y_t, y_{t-1}, \dots, y_{t-M+1}]$ . En el presente trabajo los parámetros de la RNF,  $b_i$  y  $w_{ij}$ , se ajustan mediante una ley de aprendizaje basada en el FKE. Para más detalles ver (Haykin; 1994) y (Patiño; 1998).

## 6. RESULTADOS PRINCIPALES

Para demostrar la viabilidad característica de este modelo que presentan la estructura de la RNF y su análisis de desempeño durante el proceso de aprendizaje y test se basó en la propuesta presentada en (Patiño; 1995; 1998). La topología de la RNF consistió de: una línea de retardo de la señal  $y(t)$  de 5 (regresor de 5 componentes), 50 neuronas en la capa oculta con funciones de activación de tipo seno, y una neurona de salida con función de activación de tipo sigmoide. Se utilizaron 98 de datos de la serie temporal para

Para demostrar la viabilidad característica de este modelo que presentan la estructura de la RNF y su análisis de desempeño durante el proceso de aprendizaje y test se basó en la propuesta presentada en (Patiño; 1995; 1998).

La topología de la RNF consistió de: una línea de retardo de la señal  $y(t)$  de 5 (regresor de 5 componentes), 50 neuronas en la capa oculta con funciones de activación de tipo seno, y una neurona de salida con función de activación de tipo sigmoide. Se utilizaron 98 de datos de la serie temporal para

ajustar las sinápsis de la red neuronal usando un FKE como ley de aprendizaje. Los tres últimos valores de la serie temporal fueron usados para la prueba de test y validación del modelo. En la Figura 5 se muestra un diagrama de dispersión para comparar los valores estimados por la red y los verdaderos de la serie temporal, obteniéndose un coeficiente de correlación de  $R^2=0,99$ . En la Figura 6 se presenta el observado y calculado por la red, y el error de predicción. Para medir el desempeño del sistema de identificación y predicción se empleó las métricas del error absoluto medio porcentual simétrico (SMAPE) y RMSE, obteniéndose:  $SMAPE=1,42$  y  $RMSE=35,22$ . La Tabla 1 muestra el porcentaje de error absoluto de los tres últimos datos de la serie usados en la prueba de test-validación.

Tabla 1: Porcentaje del Error Absoluto Relativo Porcentual.

	Observada [Hm <sup>3</sup> ]	Predicción [Hm <sup>3</sup> ]	Error Absoluto Porcentual %
2009	1397,28	1405,4	0,6
2010	853,05	854,00	0,1
2011	1055,41	995,29	5,7

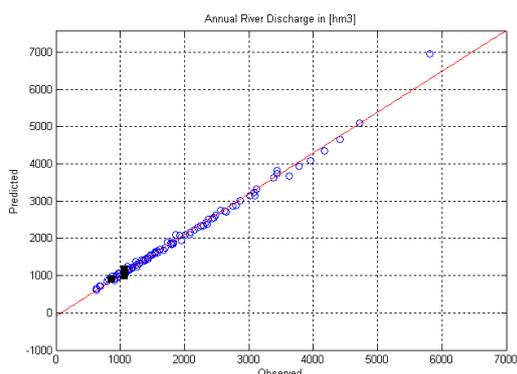


Fig. 5: Diagrama de dispersión entre los valores pronosticados por la red y los verdaderos de la serie temporal.

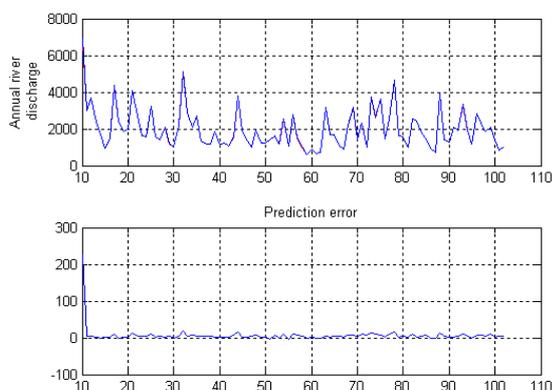


Fig. 6: (a) Pronóstico y observados de la serie temporal. (b) Errores de predicción.

## 7. CONCLUSIONES

En este trabajo se ha presentado un sistema para la modelación y predicción de la serie temporal del derrame anual acumulado de ríos basado en la aproximación de un modelo ARN mediante una RNF. La ley de aprendizaje empleada se basó en un FKE. En particular, se aplicó a un caso de estudio, el escurrimiento anual de agua del río San Juan en Argentina. El sistema emplea sólo como datos de entrada la serie temporal del derrame hídrico anual histórico disponible desde 1909 al 2012. El pronóstico se realiza para un horizonte de predicción de un paso hacia delante, pudiéndose extender a un horizonte de mediano-largo plazo. Este enfoque tiene la ventaja sobre otros modelos basados en NARMA, NARMAX, de que no depende de otras mediciones como, la acumulación de nieve, características de la cuenca geológica, variables climáticas, número de glaciares, precipitaciones, etc. Por lo tanto, resulta un sistema de pronóstico de bajo costo. Además, no necesita un procedimiento de calibración como los otros. Los resultados de los estudios de simulación indican un muy buen desempeño en la modelación de la serie temporal y una buena performance respecto a la estimación del pronóstico. Se ha observado que la salida del sistema de pronóstico es sensible a la cantidad de componentes seleccionadas en la línea de retardo, en el diseño regresor, debido principalmente a que se cuenta con una serie temporal de muy pocos valores, 102 datos. Como trabajos futuros se propone lograr una mayor robustez estadística de la predicción y definición de los límites de confianza para una serie temporal de pocos datos como la estudiada.

## AGRADECIMIENTOS

Los autores desean agradecer el apoyo financiero recibido por la Universidad Nacional de San Juan y la Universidad Nacional de Córdoba para la realización del presente trabajo. Además, a la Dirección de Hidráulica de la provincia de San Juan por permitir al acceso de las mediciones del caudal diario del río San Juan desde 1909 al presente.

## REFERENCIAS

- Hsu, K; Gupta, H. V. (1995) "Artificial neural network modeling of the rainfall-runoff process". *Water Resour. Res.*, 31, (10), pp. 2517-2530.
- Kitanidis, P. K.; Bras, R.L. (1980) "Adaptive filtering through detection of isolated transient errors in rainfall-runoff models". *Water Resour. Res.*, 16 (4), pp. 740-748.

- Kitanidis, K.; Bras, R. L. (1980) "Real-Time Forecasting With A Conceptual Hydrological Model". *Water Resour. Res.*, 16 (4), Pp. 740–748.
- Duan, Q.; Sorooshian, S.; Gupta, V. K. (1992) "Effective and efficient global optimization for conceptual rainfall–runoff models". *Water Resour. Res.*, 28 (4), pp. 1015–1031.
- Duan, Q.; Sorooshian, S.; Gupta, V.K. (1994) "Optimal use of SCEUA global optimization method for calibrating watershed models". *Journal of Hydrology*. 158, pp. 265–284.
- Sorooshian, S.; Daun, Q.; Gupta, V.K. (1993) "Calibration of rainfall–runoff models: application of global optimization to the sacramento soil moisture accounting model". *Water Resour. Res.*, 29 (4), pp. 1185–1194.
- Yapo, P.; Gupta, V. K.; Sorooshian, S. (1996) "Calibration of conceptual rainfall–runoff models: Sensitivity to calibration data". *Journal of Hydrol.* 181, pp. 23–48.
- Salas, J. D.; Delleur, J. W.; Yevjevich, V.; Lane, W. L. (1985) "Applied Modeling of Hydrologic Time Series". *Water*.
- Mishra, A.; Abdelhadi, A. (2004) "Models for recession flows in the upper Blue Nile River". *Hydrological Processes* 18, pp. 2773–2786.
- Chang, F.J.; Hwang, Y.Y. (1999) "A self-organization algorithm for real-time flood forecast". *Hydrological Processes* 13(2), pp. 123–138.
- Battit, R. (1999) "First and second order methods of learning: between the steepest descent and Newton's Method". *Neural Computation* 4(2), pp. 141–166.
- Young, P.; Chotai, A. (2001) "Data-Based Mechanistic Modeling, Forecasting, and Control". *IEEE Control System Magazine*. Oct.
- Willems, J.C. (1986) "From Time Series to Linear System – Part I. Finite Dimensional Linear Time Invariant". *Automatica*, Vol. 22, N° 5, pp. 561–580.
- Willems, J.C. (1986), "From Time Series to Linear System – Part II. Exact Modelling". *Automatica*, Vol. 22, No. 6, pp. 675–694.
- Willems, J.C. (1987) "From Time Series to Linear System – Part III. Approximate Modelling". *Automatica*, Vol. 23, No. 1, pp. 87–115.
- Markovskiy, I.; Willems, J.C.; De Moor, D. (2005) "State Representation From Finite Time Series". *Proceedings of the 44th IEEE Conference on Decision and Control, and the European Control Conference 2005*. Seville, Spain, Dec.
- Taken, F. (1981) "Detecting strange attractors in turbulence". In D.A. Rand and B.-S. Young, editors, *Dynamical Systems and Turbulence*, Vol. 898 of Lecture Notes in Mathematics, pages 366–381. Springer-Verlag, Berlin.
- Mañé, R. (1981) "On the dimension of the compact invariant sets of certain non-linear maps". In D.A. Rand and B.S. Young, editors, *Dynamical Systems and Turbulence*, Vol. 898 of Lecture Notes in Mathematics, pages 230–242. Springer-Verlag, Berlin.
- Funahashi, K. (1989) "On the approximate realization of continuous mapping by neural networks," *Neural Networks*, vol. 2, pp. 183–192.
- Hornik, K.; Stinchcombe, M.; White, H. (1989) "Multilayered feedforward network are universal approximators," *Neural Networks*, vol. 2, pp. 359–366.
- Hunt, K.J.; Sbarbaro, D.; Zbikowski, R.; Gawthrop, P.J. (1992) "Neural networks for control systems—A survey," *Automatica*, vol. 28, pp. 1083–1112.
- Bloomfield, P. (1976) *Fourier analysis of time series: An introduction*. New York: Wiley.
- Saremi, A.; Karimi Pashaki M.; Sedghi H., Rouzbahani, A.; Saremi, A. (2011) "Simulation of River Flow Using Fourier Series Models." *2011 International Conference on Environmental and Computer Science*. Vol. 19, pp. 133–138.
- Zuo, W.; Y. Zhu, Y.; Cai, L. (2009) "Fourier-Neural-Network-Based Learning Control for a Class of Nonlinear Systems With Flexible Component." *IEEE TRANSACTIONS ON NEURAL NETWORKS*. Vol. 20, NO. 1, pp. 139–151. JANUARY.
- Cordova, J.J.; W. Yu. W. (2009) "Stable Fourier Neural Networks with Application to Modeling Lettuce Growth," *Proceedings of International Joint Conference on Neural Networks*, pp. 591–596, Atlanta, Georgia, USA, June.
- Huang, G-B; Babri, H.A. (1998) "Upper Bounds on the Number of Hidden Neurons in Feedforward Networks with Arbitrary Bounded Nonlinear Activation Functions." *IEEE Trans. on Neural Networks*, Vol. 9, NO. 1, JANUARY.
- Haykin, S. (1994) *Neural Networks: A comprehensive foundation*. Prentice Hall; 1th, and 2ed.
- Patiño, H. D. (1998) "Letter of Course EE/CS 676: Neural Information Processing Systems", Electrical and Computer Engineering Department, Stevens Institute of Technology, Hoboken, New Jersey, U.S.A.
- Patiño, H. D. (1995) "Control Dinámico de manipuladores Robóticos Usando Redes Neuronales". *Temas de Automática*. Tesis doctoral en Ingeniería de Sistemas de Control, Fac. de Ing., INAUT, UNSJ. Ed. Fundación Universidad Nacional de San Juan, páginas 277.
- Pucheta, J.; Patiño, H.D.; Kuchen, B. (2007) "Neural Networks-Based Time Series Prediction Using Long and Short Term Dependence in the Learning Process". *International Symposium on Forecasting ISF'07, NN3 Forecasting Competition*, New York, USA.

- Dawson, C.W.; Harpham, C.; Chen Y. (2002) "Evaluation of artificial neural network techniques for flow forecasting in River Yangtze, China." *Hydrol. Earth Syst. Sci.*, **Vol. 6.**, pp. 619-626.
- Dolling, O.R.; Veras, E.A. (2002) "Artificial neural networks for streamflow prediction." *J. Hydraul. Res.*, **Vol. 40**, pp. 547-554.
- Imrie, C.E.; Durucan, S.; Korre, A. (2000) "River flow prediction using artificial neural networks: generalization beyond calibration range." *J. Hydrol.*, **Vol. 233**, pp. 138-153.
- Thirumalaiah, K.; Deo, M.C. (1998) "River stage forecasting using artificial neural networks," *J. Hydrolo. Eng.-ASCE*, **Vol 3**, pp. 26-31.
- Shamseldin, A.Y.; Nasr, A.E.; O'Connor, K.M. (2002) "Comparison of different forms of multi-layer feed-forward neural network method used for river flow forecasting," *Hydrol. Earth Syst. Sci.*, **Vol. 6**, pp. 671-684.
- Sivakumar, B.; Jayawardena, A.W.; Fernando, T.M.K.G. (2002) "River flow forecasting: use of phase-space reconstruction and artificial neural networks approaches," *Journal of Hydrology*, 265, pp. 225-245.
- Kisi, O. (2005) "Daily River Flow Forecasting Using Artificial Neural Networks and Auto-Regressive Models," *Turkish J. Eng. Env. Sci.*, 29, pp. 9-20.
- Molden, D. (2007) Summary water for food, water for life: A Comprehensive Assessment of Water Management in Agriculture. *International Water Management Institute*.