

XIX CONGRESO LATINOAMERICANO DE HIDRÁULICA CORDOBA 2000

PRONÓSTICO DE CAUDALES EN CUENCAS NIVALES USANDO REDES NEURONALES ARTIFICIALES

Oscar Raúl Dölling¹ y Eduardo Varas Castellón²

¹ Dpto. de Hidráulica, Facultad de Ingeniería, Univ. Nacional de San Juan - odolling@unsj.edu.ar

² Dpto. de Ing. Hidráulica y Ambiental, Esc. Ingeniería, P. Univ. Católica de Chile evaras@ing.puc.cl

RESUMEN

Este trabajo expone una metodología destinada a predecir los escurrimientos mensuales en cuencas de montaña, aplicando la técnica de redes neuronales artificiales (ANN). El procedimiento guía la selección de las variables de entrada, la definición de la arquitectura del modelo y los procesos de entrenamiento. Esta metodología logra un modelo capaz de capturar la naturaleza no lineal del proceso precipitación-escorrentía para una escala mensual.

La metodología propuesta fue aplicada para obtener un modelo de predicción de escurrimientos mensuales en la cuenca del río San Juan, Argentina, usando datos climatológicos de la estación Pachón localizada a 1200 mts. de altitud. Los escurrimientos fueron medidos en la estación de aforos Km 47.3 la cual controla una cuenca de 20.000 km².

Los resultados muestran que los escurrimientos mensuales de primavera y verano se representan adecuadamente, mejorando los resultados obtenidos utilizando otros métodos. Esto deberá traer importantes beneficios socio económicos derivados de un uso óptimo de los recursos de agua para riego y producción de energía en la cuenca.

ABSTRACT

This paper presents a methodology for predicting monthly streamflow on mountain watersheds, based on artificial neural networks (ANN). The procedure addresses the selection of input variables, the definition of model architecture and the strategy of the learning process. This type of model captures the non linear nature of the precipitation-runoff process on a monthly scale.

The methodology was applied to obtain a model for streamflow prediction for San Juan River basin, Argentina, using climatological data from Pachón meteorological station located at 1200 m of altitude. Flows were measured at Km 47.3 station which controls a 20.000 square kilometers watershed.

Results show that monthly spring and summer streamflow, can be adequately represented improving the results of calculations obtained using other methods. This should have significant benefits for the optimal use of water resources for irrigation and hydroelectric energy generation.

INTRODUCCIÓN

Distintos investigadores han utilizado diferentes métodos matemáticos y estadísticos para desarrollar modelos de predicción de caudales provenientes de lluvia y de deshielo. Se puede mencionar entre ellos, los modelos de simulación de lluvia–escorrentía y de derretimiento de nieve (Martinec y Rango, 1992), los métodos de regresión lineal múltiple para el pronóstico de volumen total de deshielo de temporada y posterior desagregación mensual de dicho volumen, los modelos de Funciones de Transferencia para pronóstico a un mes plazo construidos a partir de distintas variables explicativas del deshielo, los modelos estocásticos ajustados a la serie de tiempo de los caudales de deshielo y los modelos empíricos como el de potencias ajustado a datos de escurrimientos con una transformación logarítmica. (Karunanithi et al, 1994).

El pronóstico de caudales en cuencas de régimen de precipitación pluvio-nival es esencial para resolver un amplio número de problemas de planificación, diseño y operación de los sistemas de aprovechamiento del recurso agua, los cuales tienen un alto impacto socio económico a nivel regional.

En particular, en las cuencas cordilleranas tanto de Argentina como de Chile existen dificultades para conocer con razonable precisión los caudales mensuales de deshielo. Una de las principales causas de esta dificultad es la influencia que el fenómeno (ENSO) tiene sobre la hidrología de las cuencas del Pacífico sur, hecho referido en numerosas investigaciones en Bolivia, Perú, Argentina y Chile (Salazar et al, 1998; Rossel et al, 1998). Hoy día importantes centros de investigación de la interacción océano–atmósfera–continente tales como el Lamont–Doherty Earth Observatory de la Universidad de Columbia están abocados a estudiar la influencia de los océanos en el clima y su impacto sobre el medio ambiente global (Flamenco y Berri, 1997). Sin embargo, la dificultad en la predicción del clima y las perturbaciones asociadas a las anomalías de temperatura en la superficie del océano Pacífico, reside en su dinámica altamente no-lineal, prototipo de varios fenómenos combinados como la turbulencia, la variable opacidad de la atmósfera a la radiación solar, las transferencias de calor de gran escala en la atmósfera y los complejos procesos termodinámicos, lo que genera incapacidad predictiva (H.Raman, 1995).

Los caudales a la salida de una cuenca contienen información de la evolución del sistema físico global que solo puede ser parcialmente captado por un modelo lineal (Salazar et al, 1998; Rossel et al, 1998). Las redes neuronales artificiales, cuya estructura base está formada por una combinación de modelos no lineales elementales, son una técnica adecuada para el desarrollo de un modelo representativo de caudales de deshielo. Los detalles de las propiedades computacionales emergentes de estas redes neuronales artificiales han sido discutidos por Hopfield en 1982.

La utilización de las redes neuronales artificiales ANNs para la predicción de caudales es bastante reciente. Las investigaciones más antiguas se registran a partir del año 1992 con los trabajos de (French et al, 1992; Zurada, 1992). Hoy existen numerosas aplicaciones de ANNs en hidrología, dedicadas a predecir y modelar los fenómenos de lluvia escorrentía en cuencas de distintas características, utilizando datos de entrada tales como precipitación y temperatura.

FORMULACIÓN GENERAL DEL PROBLEMA. NECESIDADES.

La solución de cómo lograr el diseño y validación de un modelo de pronóstico que capture de una manera óptima la naturaleza no lineal de los fenómenos involucrados en la dinámica de los escurrimientos originados por la combinación de lluvia y deshielo en

cuenecas con una alta influencia de fenómenos climáticos locales y global y bajo condiciones de escasez de información, es el problema principal que aborda este trabajo.

Dadas las características del problema , el estudio se centró en el análisis de la aplicación de uno de los paradigmas simbólicos de la inteligencia artificial, las redes neuronales artificiales (ANN).

La aplicación de ANN a la predicción de escurrimientos necesita la definición de cuatro aspectos fundamentales (ver Figura 1): la determinación de las variables explicativas que posean la mayor influencia sobre los escurrimientos, la selección y diseño de la arquitectura óptima de la red, la selección del mejor método de aprendizaje (ajuste de pesos) y la estrategia para validar la respuesta del modelo de predicción.

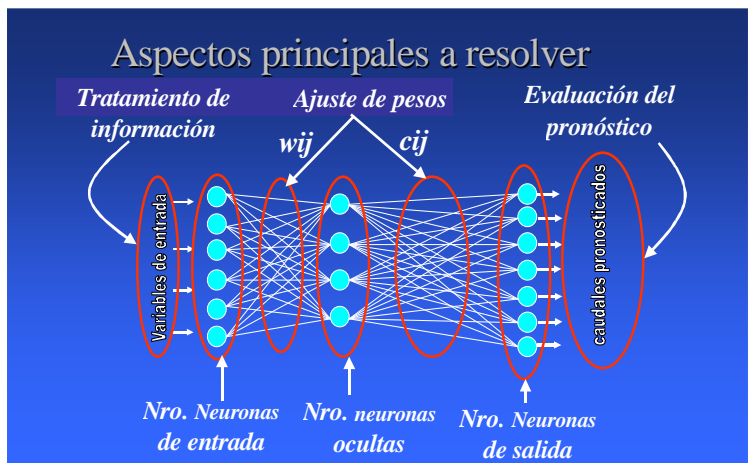


Figura 1: Diseño de una red neuronal artificial multicapa completamente conectada

OBJETIVOS Y ALCANCES DEL TRABAJO

El primer objetivo abordado por este trabajo fue la elaboración de una metodología de diseño y entrenamiento de modelos de redes neuronales artificiales, para lo cual se analizó el trabajo de otros investigadores que utilizaron ANNs a distintos casos de predicción de caudales y experiencias de predicción en otros sistemas de comportamiento no lineal tal como los recientes trabajos de (Zealand et al, 1999; Abrahart et al, 1998; Minns, 1998; Salazar et al, 1998) entre otros.

Otro objetivo importante del trabajo fue probar la metodología desarrollada por medio de su aplicación a un caso real de diseño y validación de un modelo destinado a realizar el pronóstico de volúmenes de escurrimiento mensual en una cuenca montañosa (se eligió la cuenca del río San Juan, Argentina), para un horizonte desde uno a siete meses, pronóstico que debía realizarse previo a la ocurrencia de las nevadas más importantes (mes de Julio) y bajo condiciones de escasez de información.

Se seleccionaron como variables de entrada: la información histórica del indicador del fenómeno ENSO (anomalías de la temperatura superficial del mar en la Zona 3 del Pacífico Ecuatorial, determinada entre los paralelos 5°Norte y 5°Sur y entre los meridianos 90° y 180°Oeste), la temperatura, la precipitación, el equivalente de agua, el viento, la radiación solar, entre otras variables controladas en la estación Pachón ubicada

a 1200 m.s.n.m. en la alta cuenca del río San Juan y los caudales de deshielo observados en la estación Km 47.3 sobre el río San Juan .

El desarrollo del modelo de pronóstico se encaró desde dos puntos de vista distintos. El primer enfoque consistió en desarrollar siete modelos ANNs dedicados cada uno al pronóstico del volumen de escurrimiento de cada mes de julio a enero, es decir siete modelos con una sola unidad de salida, basados en la estructura de entradas identificada como la más explicativa del deshielo. El segundo enfoque consistió en ajustar una sola red neuronal para la misma estructura de entradas anterior, pero con una capa de salida provista con siete unidades, una para cada mes a pronosticar, con lo cual se pretendió investigar la capacidad de las ANNs para aprender a reconocer varios conceptos simultáneamente (Wins, 1997 pp 499), es decir la capacidad de predecir la onda completa de deshielo en un solo paso de pronóstico.

HIPÓTESIS DE TRABAJO. DESCRIPCIÓN DE ELEMENTOS.

Una red neuronal artificial (figura 1), es un modelo matemático del tipo conexionista “grafo dirigido”, formado por “unidades o neuronas” y “conexiones o enlaces” entre ellas, que poseen un “peso” o valor de ponderación asociado al enlace. El procesamiento de la información entre las unidades se modela a través de elementos computacionales básicos (suma, multiplicación, umbral). Una neurona está formada básicamente por una “función de activación” y una “función umbral o de salida”. Para los efectos del procedimiento de aprendizaje dichas funciones deben ser funciones continuas y diferenciables y pueden ser funciones no lineales, tales como la función sigmoide o la función gaussiana. Un conjunto de neuronas forma lo que se llama una capa. A su vez las neuronas de estas capas pueden conectarse a las neuronas de otras capas o recursivamente unirse a sí mismas lo que permite construir estructuras complejas conectadas. Existen muchos tipos distintos de arquitecturas y topologías de redes neuronales. Las características que diferencian los distintos tipos de redes son: la arquitectura, el método de enseñanza y la estrategia de entrenamiento.

En el caso de las redes feedforward, el algoritmo clásico de aprendizaje es el denominado backpropagation error con término de momentum. El objetivo de este algoritmo de aprendizaje es minimizar el gradiente de una función objetivo de desempeño P , que es la suma de errores cuadráticos entre los valores de salida deseados y los calculados por la red para todas las muestras usadas en el entrenamiento, respecto de los pesos de los enlaces w_{ij} . (Wins,1997)

En particular en situaciones de escasez de información y por lo tanto con un número de duplas de entrenamiento bajo, la estrategia sugerida por Mitchell (1997) es la que parece ser actualmente la más adecuada y la que se utilizó en este trabajo. Algunos problemas que pueden derivarse de una mala estrategia de entrenamiento son que las redes multicapas del tipo feedforward entrenadas con el algoritmo backpropagation no garantizan la convergencia a un óptimo global. Conforme se modifiquen los pesos el algoritmo se desliza en un hiperespacio sobre una n -brana de error la cual puede presentar mínimos locales en los cuales el algoritmo puede quedar atrapado. Existen distintos procedimientos para escapar de los mínimos locales que tienen que ver con el adecuado uso de los parámetros de aprendizaje y del término de momentum (Figura 2).

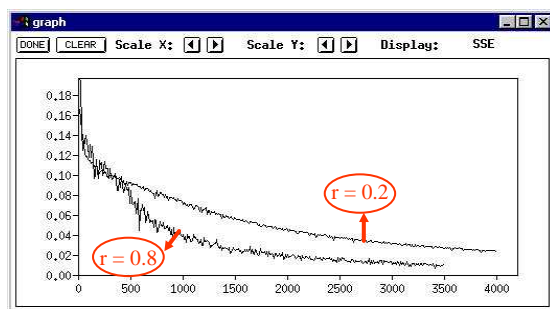


Figura 2.- Curvas de entrenamiento de una red ANN para distintas estrategias (parámetros de rapidez r) por el método backpropagation momentum

Otro problema que suele presentarse durante el entrenamiento de redes feedforward entrenadas con algoritmos backpropagation error es por una mala selección del punto óptimo de finalización del entrenamiento. Un entrenamiento demasiado largo puede conducir al fenómeno indicado como sobreentrenamiento (overfitting), que puede derivar en una disminución considerable de la capacidad predictiva de la red. (Mitchell, 1997)

El procedimiento de validación de las redes neuronales se basó en un conjunto de análisis sobre los residuos (valores calculados como la diferencia entre los valores de escurrimiento observados y los calculados por la red) generados por cada modelo para los distintos conjuntos (entrenamiento, validación y test). También se usó como indicador de comportamiento el número de ciclos necesario para lograr un entrenamiento que asegure una buena generalización.

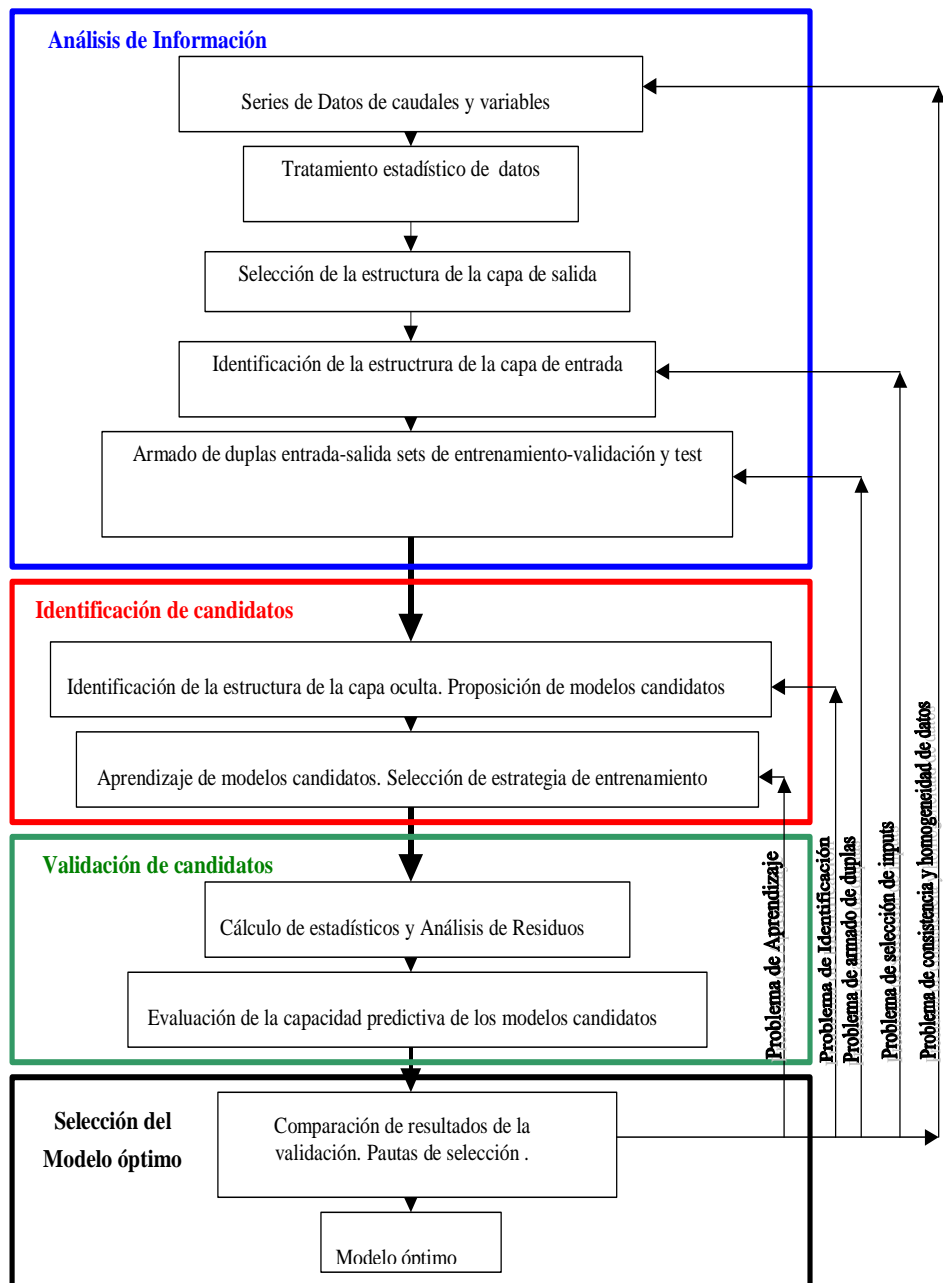
Para comprobar en los modelos el grado de ajuste a los datos de entrenamiento, la capacidad de generalización fuera del conjunto de entrenamiento y detectar problemas de sobreentrenamiento, se usó el estadístico suma de errores cuadráticos, (SSE) y la función de autocorrelación de los residuos. La detección de coeficientes de autocorrelación de la serie de residuos que superen un determinado nivel de confianza, indica problemas en la identificación de las variables de entrada, o bien una arquitectura defectuosa de la red para capturar toda la información predecible.

Eliminado: que

METODOLOGÍA Y PLAN DE TRABAJO

No existe a la fecha, una metodología única para lograr el diseño y la validación de ANNs. Del análisis bibliográfico realizado y la consulta a especialistas se ha logrado reunir el conocimiento necesario para definir una metodología de trabajo que permite encontrar el diseño y la validación de un modelo ANN de predicción, para un nivel determinado de confianza referido a la capacidad de generalización. Las fases a realizar para lograr la elaboración de un modelo ANNs con dichas características, se presentan en la Figura 3 en forma de Diagrama de bloques. Como se puede observar el proceso de elaboración presenta distintas instancias de decisión y prueba que requieren la intervención de un analista y una etapa de validación final que puede conducir a la decisión de volver atrás iterativamente para resolver problemas identificados en etapas anteriores.

Figura 3 : fases de elaboración de un ANNs para predicción



RESULTADOS ALCANZADOS

A continuación se presentan los resultados obtenidos en el diseño y validación de un modelo ANN dedicado a pronosticar la onda de deshielo completa.

La metodología retroalimentativa de elaboración de ANNs desarrollada, permitió encontrar un modelo capaz de predecir, con suficiente precisión, los caudales de deshielo de la cuenca del río San Juan. El software y algoritmos de apoyo al diseño y entrenamiento de ANNs orientados a investigación aplicada SNNS aportó una base de recursos importante en el desarrollo computacional de las herramientas necesarias para llevar a buen término las distintas etapas planteadas en la metodología utilizada.

Se comprobó el buen comportamiento predictivo de un modelo ANNs feedforward de arquitectura 30-20-7 (30 neuronas de entrada, 20 neuronas ocultas y 7 neuronas de salida) entrenado mediante el método de aprendizaje backpropagation momentum y utilizando la estrategia de aprendizaje sugerida por Mitchell (1997) para ejemplos escasos. Se inicializó la red con valores aleatorios entre -2 y 2 . Se estudió el proceso de convergencia del método backpropagation y se probaron distintas tasas de aprendizaje y distintos coeficientes de momentum, finalmente se adoptaron los valores $0,9$ y $0,7$ respectivamente.

Pudieron elaborarse 18 duplas entrada-salida (30-7), esto es cada dupla compuesta por 30 variables de entrada y 7 variables de salida. De ellas 17 fueron usadas para el entrenamiento y la última dupla se utilizó para Test de la red. Las 30 variables de entrada incluidas en el modelo corresponden a los valores observados para los meses de abril, mayo y junio de: Número del mes, IOS (Índice de Oscilación del Sur), precipitación mensual, temperatura promedio, Humedad relativa promedio, Heliofanía Efectiva Promedio, Viento promedio mensual, altura máxima de nieve, Días totalmente nublados (datos climáticos de la estación meteorológica Pachón) y el Volumen de escurrimiento (estación de aforo km 47.3) ambas estaciones dependientes del Dpto. Hidráulica de la provincia de San Juan. Las 7 variables de salida corresponden a los valores del volumen escurrido mensual (aforo en Km 47.3) para los meses de julio a enero del próximo año.

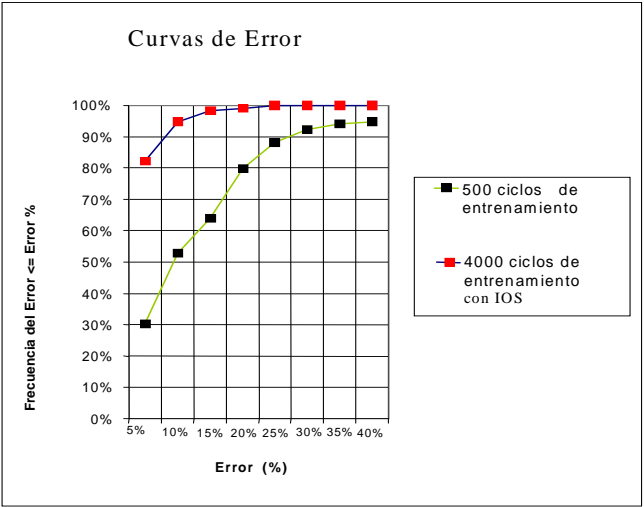
Todos los valores incluidos en las duplas fueron escalados de $-0,7$ a $0,7$ con el fin de homogeneizar las magnitudes de las distintas variables, utilizar la función Sigmoide como función de activación a la salida de la red y para poder extrapolar valores de salida superiores a los usados en el entrenamiento de la red.

El análisis de los valores de SSE (Suma de errores cuadráticos) obtenidos en la predicción del patrón de Test (julio 1997 a enero 1998) indica que la red entrenada presenta una mejor capacidad de generalización para los 4000 ciclos de entrenamiento.

Se realizó un análisis detallado de la frecuencia con que cada modelo entrenado cometía errores de predicción por bajo determinados umbrales. Esto permitió obtener lo que se denominó **curvas de error** para cada modelo. La introducción de indicadores de las anomalías de temperatura superficial del Pacífico Ecuatorial, como el IOS, fue un elemento explicativo de importancia en la cuenca de estudio. En particular, su inclusión en el modelo mejoró su repuesta en cuanto su capacidad de generalización lo cual se puede observar utilizando las curvas de error que se presentan en la (figura 4). Las curvas de error de los modelos de ANN entrenados para distintos ciclos de entrenamiento, ilustran la frecuencia con que un cierto umbral de error no es superado por el pronóstico. Este valor de frecuencia corresponde al promedio de los 7 valores de frecuencia correspondientes a los 7 meses pronosticados. Como un ejemplo de interpretación de esta gráfica se puede decir que para la red ANN entrenada con 4000 ciclos y que posee el indicador IOS como variable de entrada, en el 82% de los pronósticos realizados, el valor absoluto del error de

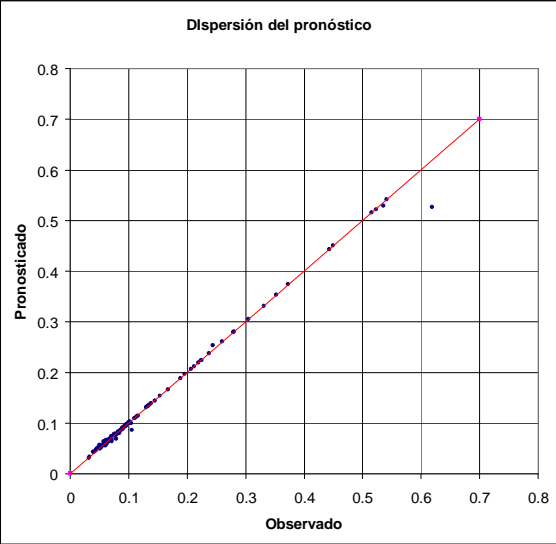
predicción promedio de los siete meses no superó el 5%, mientras que en el 95% de los pronósticos realizados el valor absoluto del error de predicción promedio de los siete meses es inferior al 10%.

Figura 4: Curvas de Garantía de error



También es importante indicar la utilidad del diagrama de dispersión de valores pronosticados sobre los valores observados el cual permite visualizar el comportamiento del modelo de pronóstico en todo el rango de datos analizado (figura 5) y las tendencias hacia la sobre y la sub estimación de cada modelo.

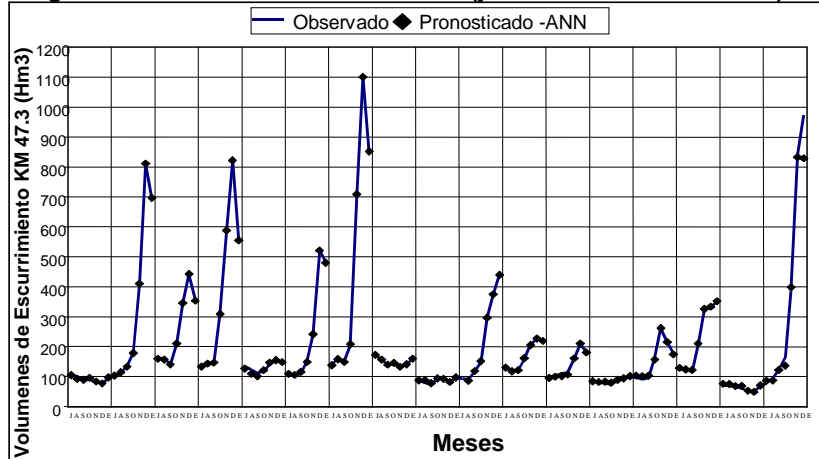
Figura 5: Diagrama de dispersión de valores de volúmenes de escurrimiento observados y pronosticados, escalados (0 a 0,7) para la red ANN – 4000 ciclos -IOS



CONCLUSIONES

Se implementó una metodología retroalimentativa de elaboración de ANNs que permitió encontrar un modelo capaz de predecir con precisión, los caudales de deshielo de la cuenca del río San Juan antes de registrar las nevadas mas importantes en la cuenca (Figura 6).

Figura 6: Volúmenes de Esguerrimiento (julio año T - enero año T+1)



La metodología desarrollada, puede generalizarse a cuencas de similares características hidrológicas con buen grado de éxito, fundamentando esta apreciación, en los excelentes resultados obtenidos por el mismo autor en otras aplicaciones destinadas a identificar modelos ANNs capaces de predecir los esguerrimientos derivados de fenómenos igualmente complejos como el de lluvia-esguerrimiento y lluvia – deshielo- esguerrimiento en cuencas de respuesta notablemente no lineal, como la cuenca del río Maule en Chile.

La introducción de indicadores de las anomalías de temperatura superficial del Pacífico Ecuatorial, es un elemento explicativo de importancia cuya inclusión en el modelo mejoró el comportamiento predictivo de la red ANN ajustada por medio del método de aprendizaje backpropagation momentum y entrenada con la estrategia de aprendizaje sugerida por Mitchell para ejemplos escasos, calibrado para pronosticar los esguerrimientos para la cuenca del río San Juan.

Existe numeroso software y algoritmos de apoyo al diseño y entrenamiento de ANNs elaborados en distintos lenguajes de computación, como también paquetes de software orientados a investigación aplicada como el SNNS que aportan una base de recursos muy importante en el desarrollo computacional de las herramientas necesarias para llevar a buen término las distintas etapas planteadas en la metodología de identificación y validación de modelos desarrollada.

Finalmente, este trabajo pone de manifiesto, con resultados concretos, que una utilización metódica y ordenada de las técnicas de Inteligencia Artificial, como lo es la de redes neuronales artificiales, al pronóstico de variables hidrológicas puede dar lugar a una nueva etapa en la gestión de los recursos hídricos, la cual estará apoyada sobre predicciones tanto o mas precisas y confiables que las realizadas hasta ahora con los métodos de simulación o procedimientos estadísticos tradicionales, especialmente cuando se trata de predecir fenómenos de características netamente no lineal.

REFERENCIAS

- (1) ABRAHART R.J., SEE L., KNEALE P.E. (1998) "New Tools for Neurohydrologists : using network pruning and model breeding algorithms to discover optimum inputs and architectures." <http://www.geog.port.ac.uk/geocomp/geo98/20/gc-20.htm>
- (2) FLAMENCO E., BERRI G.J.(1997) "Modelo de pronóstico de la disponibilidad de agua en el río Diamante de la provincia de Mendoza durante el período octubre-marzo, basado en el pronóstico del fenómeno "El Niño". Informe EVARSA y Lamont-Doherty Earth Observatory of Columbia University. <http://www.customw.com/ecoweb/notas/970603.htm>
- (3) FRENCH, M.N., KRAJEWSKI, W.F. AND CUYKENDALL, R.R. (1992). "Rainfall forecasting in space and time using a neural network," Journal of Hydrology, 137, pp. 1-31.
- (4) KARUNANITHI N., GRENNY W.J., WHITLEY D., BOVEE K. (1994) "Neural Networks for River Flow Prediction". ASCE. Journal of Computing in Civil Engineering Vol. 8 N°2.
- (5) MARTINEC J., RANGO A., ROBERTS R. (1992) "USER'S MANUAL FOR THE SNOWMELT RUNOFF MODELO (SRM)". Hydrology Laboratory Technical Report HL-47, update edition, ver. 3.2, Bestville Maryland 10705.
- (6) MINNS A.W. (1998) *Artificial Neural Networks as subsymbolic Process Descriptors*. Thesis PhD Delft University of Technology (Holanda) Director: Abbot M.B.
- (7) MITCHELL T. M. (1997) *Machine Learning. Cap.4 Artificial Neural Networks*. Carnegie Mellon University Mc Graw Hill (pp. 81-127) . <http://www.mhcollege.com>.
- (8) RAMAN, H. AND SUNILKUMAR, N. (1995). "Multivariate modelling of water resources time series using artificial neural networks", Hydrological Sciences Journal, 40, pp.145-163.
- (9) ROSSEL F. et al (1998) "Homogeneización y Regionalización de la influencia de "El Niño" sobre las precipitaciones mensuales y anuales y construcción de modelos explicativos de previsión." XVIII Congreso Latinoam. De Hidráulica, Oaxaca, México. pp. 505-514.
- (10) SALAZAR J.E., CARAVAJAL L.F., MESA O.J., SMITH R. (1998) "Modelos de Predicción de caudales mensuales considerando anomalías climáticas." Ponencia XVIII Congreso Latinoam. De Hidráulica Oaxaca, México. pp. 525-533.
- (11) WINS P.H. (1997) *Inteligencia Artificial. Cap.22 : Aprendizaje mediante entrenamiento de redes neuronales*. Addison Wesley iberoamericana 3° edición, pp 477-505.
- (12) ZEALAND C.M., BURN D.H. y SIMONOVIC S.P. (1999) "Short Term Streamflow forecasting using artificial neural networks". Journal of Hydrology.214 pp. 32-48.
- (13) ZURADA, J.M. (1992). Introduction to Artificial Neural Systems. West Publishing Co., N.Y. U.S.A.

Eliminado: Application of a neural network technique to rainfall-runoff modelling
Shamseldin, A.Y. page 272-294
Journal of Hydrology, Volume 187 Issue 3-4 ¶