



Aplicación de Redes Neuronales Artificiales a la Modelización y Previsión de Caudales Medios Mensuales del Río Huancané

Application of Artificial Neural Networks to Modeling and Forecasting of Mean Monthly Discharge in the Huancane River

WILBER FERMÍN LAQUI VILCA^{1,*}

¹Ingeniero Agrícola, Especialista en Recursos Hídricos.

* wilfer.lv@gmail.com

RESUMEN

En este trabajo se desarrollaron modelos basados en redes neuronales artificiales del tipo retro propagación, para la modelización y previsión de los caudales medios mensuales del río Huancané, a partir de datos medios mensuales de caudales, precipitación y evaporación. El entrenamiento y la validación de los modelos de redes neuronales artificiales se realizaron utilizando 39 años de información meteorológica correspondiente a estaciones localizadas en la cuenca del río Huancané, separados en 33 años como conjunto de datos de entrenamiento y 6 años como conjunto de datos de validación. Los modelos de redes neuronales artificiales para la modelización y previsión de los caudales medios mensuales del río Huancané mostraron un excelente comportamiento en términos de medidas de error, frente a un modelo autorregresivo periódico de primer orden - PAR (1).

Palabras claves: Modelización y previsión de caudales, redes neuronales artificiales, series de tiempo, río Huancané

ABSTRACT

In this work, models based on artificial neuron networks of the back-propagation type, were developed for modeling and forecasting, at monthly step, mean monthly discharge in the Huancane River; from median monthly discharge, precipitation and evaporation data. The training and the validation of the artificial neuron networks models were made using 39 years of meteorological data corresponding to stations localized in the Huancane river watershed. These data were grouped as follow: 33 years corresponding to training data set and 6 years to the validation data set. The artificial neuron network models used for modeling and forecasting, at monthly step, mean monthly discharge of the Huancane River showed a good performance in terms of error measures, compared to a first-order PAR (1) periodic autoregressive model.

keywords: Modeling and forecasting of discharge, artificial neural networks, time series, Huancané river.

INTRODUCCIÓN

La disponibilidad de agua en cantidad y calidad insuficientes para la atención de los diversos usos y usuarios es sin duda uno de los más serios problemas a ser enfrentados en el manejo de los recursos hídricos en una cuenca hidrográfica. Generalmente, la asignación de agua para la atención de las demandas está condicionada por la disponibilidad de agua en el espacio y tiempo, por lo que se hace necesario conocer con aproximación la disponibilidad presente y futura del recurso hídrico, con la finalidad de formular propuestas actuales y futuras para una adecuada gestión de éstos (Laqui, 2008).

Debido a la escasa y limitada información de caudales de los ríos del altiplano peruano, especialmente en la cuenca del río Huancané, en general se ha utilizado modelos de transformación lluvia-caudal, como Témez (Sovero, 1991) y Lutz Scholz (Nina, 2008). La transformación de lluvia en caudal es uno de los procesos más difíciles de ser entendidos, debido al gran número de variables que se modifican en el espacio y tiempo (Weigert, 2005), así como la limitación de datos hidrometeorológicos y la no linealidad del proceso que dificultan la utilización de este tipo de modelos (Santos & Pereira, 2003).

El análisis y previsión de caudales estacionales es de fundamental importancia en el planeamiento y operación de sistemas de recursos hídricos. Una de las dificultades en la previsión de esas series es su naturaleza no estacionaria debido a los periodos de lluvia y estiaje durante el año. Estas previsiones son realizadas, actualmente, por metodologías basadas en modelos de series temporales y se recomienda los modelos autorregresivos periódicos (Ballini et al., 2003).

Dada la importancia de la modelización y previsión de caudales medios mensuales, durante las últimas décadas se han propuesto gran variedad de modelos, y muy recientemente se han introducido nuevos modelos, que manejan de una manera eficiente la aleatoriedad de los caudales y presentan una fácil

adaptabilidad a nuevos datos, sin la necesidad de recurrir a laboriosas formulaciones matemáticas. Entre estos modelos se incluyen las técnicas de inteligencia artificial, que comprenden las técnicas de Redes Neuronales Artificiales, Sistemas Fuzzy y Neuro-Fuzzy (Zúñiga & Jordán, 2005), Teoría Nebulosa, Sistemas Multi-Agentes.

Los modelos de Redes Neuronales Artificiales (RNA) tienen entre sus principales características: tratan incertezas de fácil adaptabilidad a nuevos datos y abstraen el conocimiento físico de los sistemas a ser modelados, ya que el modelo puede ser generado únicamente a partir de datos de entrada y salida (Ballini, 2000). Con el objetivo de superar las deficiencias de los modelos matemáticos se ha incrementado el número de trabajos que utilizan modelos basados en redes neuronales artificiales (Figueiredo et al., 2007). Entre sus principales aplicaciones en la solución de problemas relacionados a recursos hídricos, tenemos la previsión de caudales (Ballini, 2000; Zúñiga & Jordán, 2005; Laqui, 2008; Bravo et al., 2008; Gomes & Lima, 2010; Cruz et al., 2010), modelamiento lluvia-caudal (Baganha, 1999; Tokar & Markus, 2000; Weigert, 2005), operación de reservorios (Kilinç & Cigizoglu, 2003) y otras aplicaciones relacionadas a los recursos hídricos (Magalhães, 2004; Hillman et al., 2004; Ovanto et al., 2005; Jaimes, 2006; Gibilini, 2006), obteniéndose en todos buenos resultados.

Debido a la gran importancia que tiene la determinación y previsión de caudales, es necesaria la adopción de estos nuevos modelos como instrumentos de modelización y posterior previsión, para que permitan conocer con cierta exactitud y facilidad los caudales medios mensuales actuales y futuros del río Huancané, que permitirá su uso racional para diferentes fines.

Esta investigación se enfoca en la aplicación de redes neuronales artificiales a la modelización y previsión de caudales medios mensuales del río Huancané, el desarrollo de modelos basados en redes neuronales artificiales para la modelización y previsión de caudales medios mensuales, a partir de datos medios mensuales de caudales, precipitación y

evaporación; así como la determinación de la capacidad de desempeño de estos modelos como instrumento de previsión de caudales medios mensuales del río Huancané frente a modelos de series temporales estocásticas.

MATERIALES Y MÉTODOS

Zona de estudio

Se identifica como zona de estudio a la cuenca del río Huancané, comprendida desde sus nacientes hasta la Estación Hidrométrica Puente Huancané (próxima a su desembocadura en el Lago Titicaca). Políticamente se ubica en la región Puno, en las provincias de Huancané, San Antonio de Putina, San Pedro de Moho y

Azángaro (noreste de la región Puno), en el extremo sureste del Perú (Ver Fig. 1).

Geográficamente la cuenca del río Huancané está comprendida entre las coordenadas (UTM) Este: 376 159-468 954 y Norte: 8 397619-8 302196, su cota máxima es 5162 msnm y su cota mínima de 3820 msnm, que corresponde al cerro Surupuna y a la desembocadura del río en el Lago Titicaca, respectivamente.

La cuenca hidrográfica del río Huancané tiene una extensión aproximada de 3522 km², desde la partición de aguas hasta la desembocadura en el Lago Titicaca, presenta una forma más o menos alargada. Ver Fig. 2.

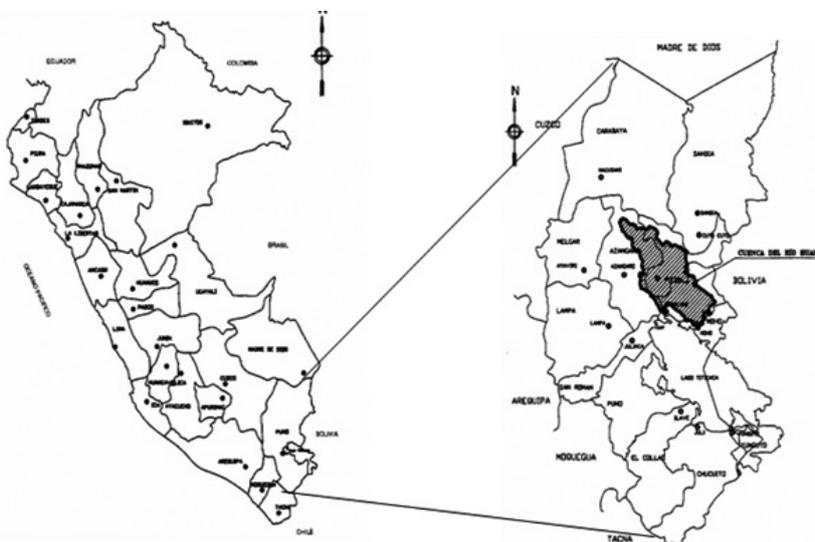


Fig. 1. Plano de ubicación de la zona de estudio.

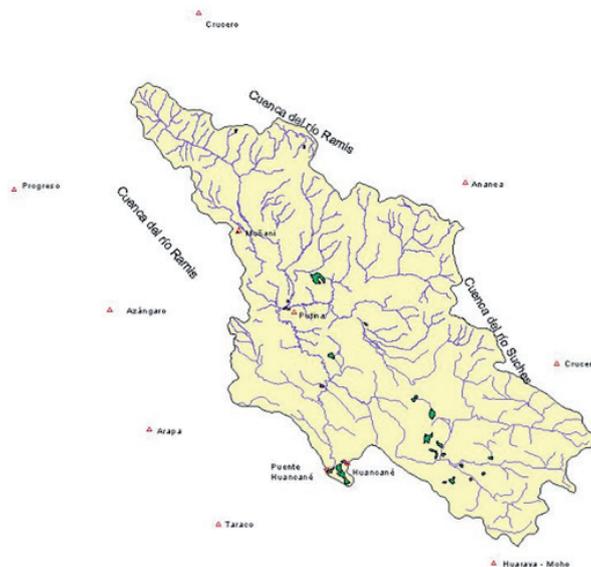


Fig. 2. Cuenca hidrográfica del río Huancané.

Metodología

La metodología seguida consta de cuatro etapas:

- Análisis y estudio de la cartografía existente.
 - Delimitación de la cuenca hidrográfica.
- Recolección y análisis de la información hidrometeorológica.
 - Análisis de consistencia de la información hidrometeorológica.
 - Completación y extensión de la información hidrometeorológica.
 - Determinación de la precipitación media de la cuenca hidrográfica.
 - Análisis estadístico de la serie temporal de caudales medios mensuales.
- Modelización de caudales medios mensuales.
 - Formulación de los posibles modelos de redes neuronales artificiales.
 - Determinación de arquitectura y parámetros de los modelos de redes neuronales artificiales.
 - Fase de entrenamiento de los modelos de redes neuronales artificiales.
 - Fase de validación de los modelos de redes neuronales.
 - Selección del mejor modelo de redes neuronales artificiales.
- Previsión de caudales medios mensuales
 - Modelamiento estocástico de las series mensuales

- Estimación de las características estadísticas de las series mensuales.
- Transformación y estandarización de datos.
- Ajuste a un modelo autorregresivo.
- Pruebas de validación del modelo autorregresivo.
- Generación de series sintéticas y análisis de bondad del modelo.
- Previsión de caudales medios mensuales.
- Modelamiento con redes neuronales artificiales de las series mensuales.
- Formulación de los posibles modelos de redes neuronales artificiales.
- Determinación de arquitectura y parámetros de los modelos de redes neuronales artificiales.
- Fase de entrenamiento de los modelos de redes neuronales artificiales.
- Fase de validación de los modelos de redes neuronales.
- Selección del mejor modelo de redes neuronales artificiales.
- Comparación de modelos
- Análisis de errores de predicción de los modelos.

Información Hidrometeorológica

En la Tabla 1, se muestra la información hidrometeorológica utilizada en esta investigación.

Tabla 1. Información hidrometeorológica.

Estación	Longitud	Latitud	Altitud	Parámetro	Período	Años de registro
Progreso	70°21'55.8"	14°41'21.1"	3 970	Precipitación Total Mensual	1964-83, 87-91, 93-2002	35
Arapa	70°07'05.6"	15°08'10.5"	3 830	Precipitación Total Mensual	1964-2002	39
Muñani	69°57'06.5"	14°46'01"	3 948	Precipitación Total Mensual	1966-74, 77-80, 82-83, 85-2002	33
Putina	70°20'56.9"	14°55'15.5"	3 878	Precipitación Total Mensual	1964, 66, 68, 71-2002	35
Azángaro	70°11'26.7"	14°54'51.7"	3 863	Precipitación Total Mensual	1965-66, 68, 70-81, 93, 95-97, 99-2002	23
Huancané	69°45'12.8"	15°12'05.4"	3 890	Precipitación Total Mensual	1964-73, 75, 77-2002	37
Cojata	69°22'00"	15° 01'00"	4 380	Precipitación Total Mensual	1964, 66-86, 88-94, 97-2002	35
Huaraya Moho	69°29'03.4"	15°23'07"	3 890	Precipitación Total Mensual	1969-2002	34
Taraco	69°59'00"	15°18'00"	3 820	Precipitación Total Mensual	1964, 67, 70-96, 98-2002	34
Ananea	69°32'03.3"	14°40'42.4"	4 660	Precipitación Total Mensual	1965-68, 71-94, 96-2002	35
Crucero	70°01'24.7"	14°21'44.4"	4 130	Precipitación Total Mensual	1964-67, 69, 71-2002	37
Huaraya Moho	69°29'03.4"	15°23'07"	3 890	Evaporación Total Mensual*	1969-82, 84, 86, 91-96, 98-2000	25
Puno	70°01'00"	15°50'00"	3 812	Evaporación Total Mensual*	1968-92, 96, 98-2000	29
Juli	69°27'00"	16°13'00"	3 820	Evaporación Total Mensual*	1964, 68-69, 71-79, 86-87, 90, 92-96, 98-2000	22
Puente Huancané	69°47'31"	15°12'50"	3 814	Caudal Medio Mensual	1964-2002	39
Puente Ramis	69°52'17"	15°15'06"	3 813	Caudal Medio Mensual	1964-2002	39
Puente llave	69°37'47"	15°05'04"	3 825	Caudal Medio Mensual	1964-2000	37

Nota: (*) Evaporación medida en tanque de evaporación tipo A

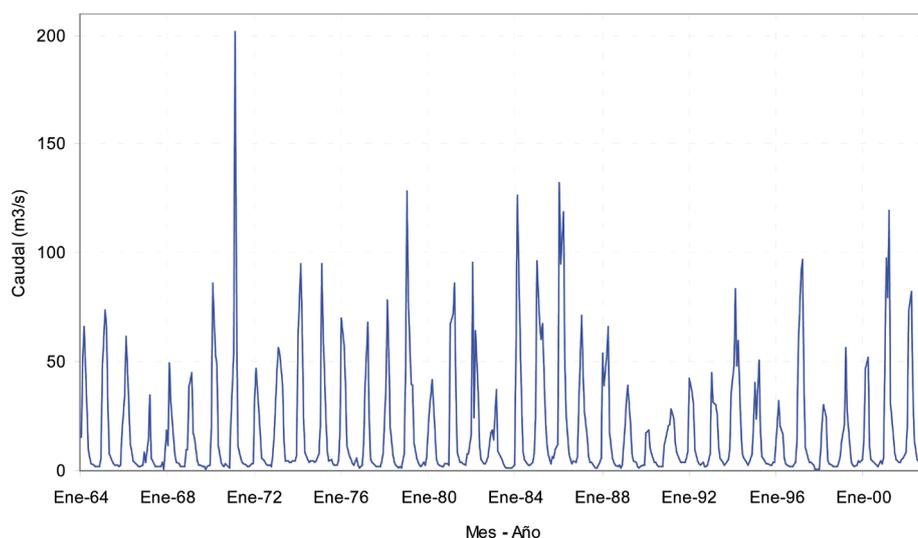


Fig. 3. Serie de caudales medios mensuales del río Huancané (1964-2002).

En la Fig. 3, se presenta la serie temporal de caudales medios mensuales históricos del río Huancané, con un total de 39 años de registro a partir de 1964 al 2002.

Se ha efectuado la completación de las series de precipitación total mensual para las estaciones Taraco, Arapa, Huancané, Huaraya, Moho, Progreso, Muñani, Putina, Azángaro, Ananea, Crucero y Cojata, así como la extensión de la serie de evaporación mensual de la estación Huaraya Moho, a partir del 2000 al 2002, determinándose la precipitación media de la cuenca del río Huancané por el método de Thiessen Modificado (Sovero, 1991).

Redes Neuronales Artificiales

Según Pulido et al. (2002), las Redes Neuronales Artificiales (RNA's) son modelos matemáticos que copian la estructura de los sistemas neuronales biológicos con el fin de alcanzar una funcionalidad similar. En general, una red neuronal consiste en el establecimiento de un conjunto de nodos o neuronas agrupadas en varias capas e interconectadas entre sí. Ver Figura 4.

Bravo et al. (2004), las Redes Neuronales Artificiales se presentan como un nuevo modelo empírico (o del tipo "caja negra"), que son modelos matemáticos que intentan imitar el funcionamiento del cerebro humano; a su

vez definen en general, como un sistema de ecuaciones, las no lineales, interconectadas, en las que el resultado de una ecuación es el valor de entrada de otras encadenadas a ésta a través de la red. Son consideradas como aproximadores universales de funciones y, por lo tanto, tendrían la capacidad de conseguir reproducir el comportamiento de cualquier función matemática.

Componentes de las Redes Neuronales

Las siguientes definiciones describen las principales componentes de una red neuronal artificial (RNA):

a) Neurona o unidad procesadora

El procesador elemental o neurona es un dispositivo simple de cálculo que, a partir de un vector de entrada procedente del exterior o de otras neuronas, proporciona una única respuesta o salida.

Las neuronas de entrada reciben los datos de partida, sin llevar a cabo ninguna operación matemática, solamente transfieren los valores a la siguiente capa.

Cada neurona de las capas intermedias (capas ocultas) y de la capa de salida realiza la suma I_j de los valores ponderados sobre los datos que recibe como entrada (Pulido et al, 2002). De este modo, dado un conjunto de entradas x_{ji} ($i=1... q$) en la neurona j , estos valores

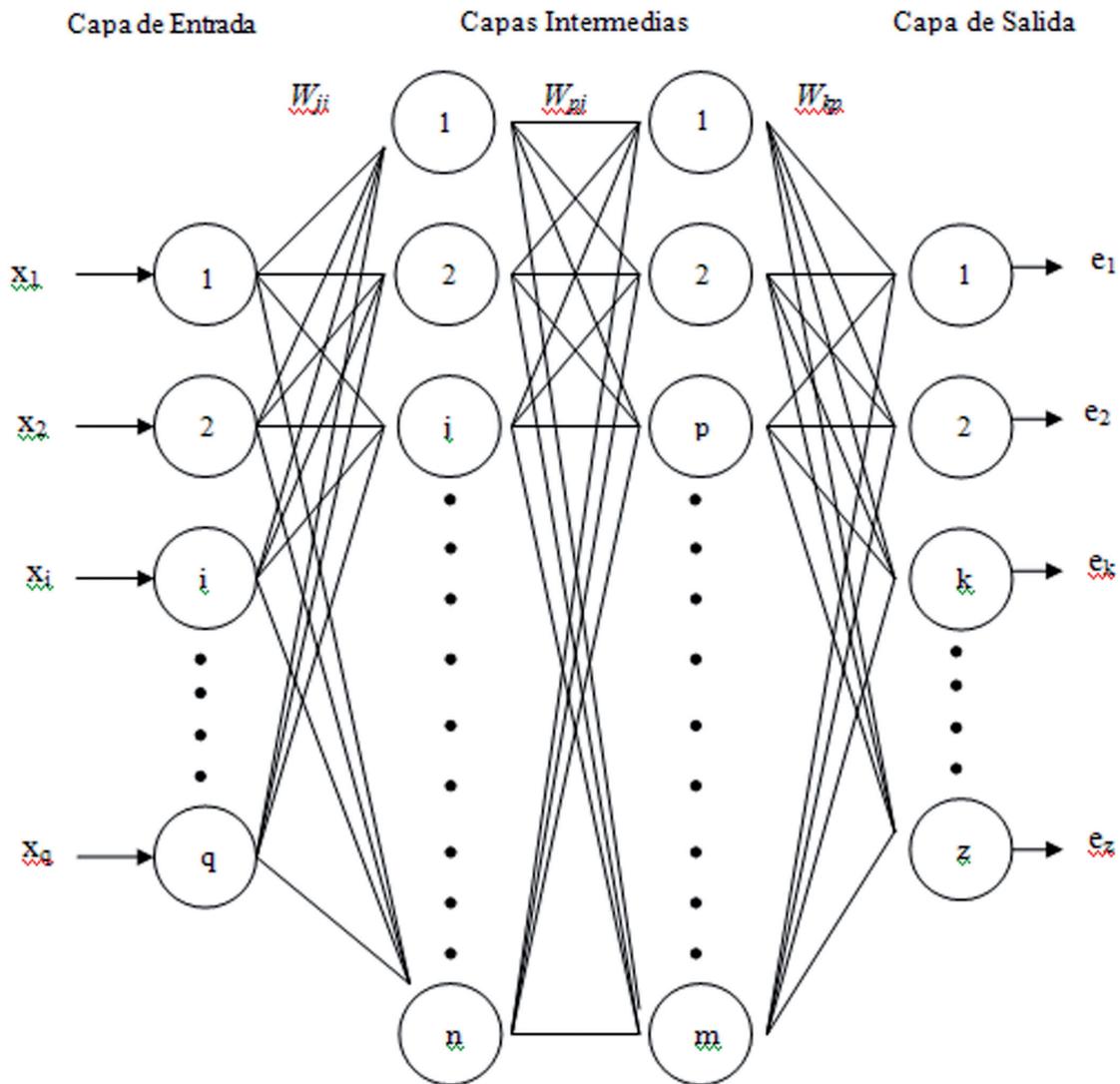


Fig. 4. Red neuronal artificial con una capa de entrada de q neuronas, dos capas intermedias de n y m neuronas, respectivamente, y una capa de salida de z neuronas (Fuente: Adaptado de Pulido et al., 2002).

se multiplican por el conjunto de pesos W_{ji} (conexión del nodo i con el nodo j). Ver Figura 5. (Pulido et al., 2002.)

Donde: I_j es la salida de una neurona, x_{ji} son las variables de entrada y W_{ji} los pesos (conexión del nodo i con el nodo j).

$$I_j = \sum_{i=1}^q x_j W_j$$

(1)

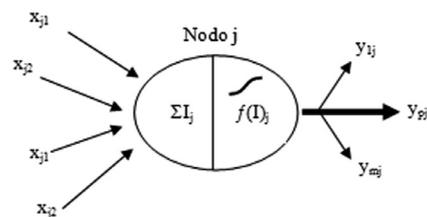


Fig. 5. Elemento de proceso o neurona (Fuente: Pulido et al., 2002)

b) Funciones de activación

A continuación se describen las funciones de activación más comunes:

Funciones lineales: Son funciones que dan una salida lineal:

$$f(x) = x, x \in R \quad (2)$$

Funciones escalón: Son funciones que dan una salida binaria dependiente de si el valor de entrada está por encima o por debajo del valor umbral. Por ejemplo la función signo, $\text{sgn}(x)$, y la escalón estándar, $A(x)$, definidas como sigue:

$$\text{sgn}(x) = \begin{cases} -1; & \text{si } x < 0 \\ 1; & \text{en otro caso} \end{cases}$$

$$A(x) = \begin{cases} 0; & \text{si } x < 0 \\ 1; & \text{en otro caso} \end{cases} \quad (3)$$

Funciones sigmoidales: Son funciones monótonas acotadas que dan una salida gradual no lineal para entradas. Las funciones sigmoidales más populares son:

La función logística:

$$y_{\text{log}} = f(I_j) = \frac{1}{1 + e^{-I_j}} \quad (4)$$

La función tangente hiperbólica: de -1 a 1 :

$$f(I_j) = \tanh(I_j) \quad (5)$$

Aprendizaje

Una de las características de las RNA es su capacidad de aprender a partir de ciertos datos. Una vez que ha sido elegida la arquitectura de red para un determinado problema, los pesos de las conexiones se ajustan para codificar la información contenida en un conjunto de datos de entrenamiento. Los métodos de aprendizaje se clasifican en dos categorías, dependiendo del tipo de información disponible (Alegre, 2003):

a) Aprendizaje supervisado: En este caso, los patrones para el aprendizaje están formados por parejas, $f(ap; bp) p = 1; \dots; rg$, que constan de un vector de variables de entrada, a , junto con las salidas correspondientes b . Por tanto cada valor de salida es la respuesta deseada a las señales de entrada. En este caso, los pesos se obtienen minimizando alguna función de error que mide la diferencia entre los valores de salida deseados y los calculados por la red neuronal. En este tipo de aprendizaje se debe tener en cuenta el problema de la convergencia del error. En general, la función de error resultante puede contener múltiples mínimos locales, por lo que, el proceso de aprendizaje puede no converger al mínimo global óptimo.

b) Aprendizaje no supervisado: En este caso, los datos se presentan a la red sin información externa, y la red tiene que descubrir por sí misma patrones o categorías. Este tipo de aprendizaje se encuadra dentro de las técnicas auto-organizativas, o técnicas automáticas para descubrir la estructura de datos. Algunos métodos de aprendizaje no supervisado son:

1. Aprendizaje Hebbiano: consiste en modificar los pesos de acuerdo con algún criterio de correlación entre las actividades neuronales.
2. Aprendizaje competitivo: neuronas diferentes se conectan con pesos negativos (inhibitorios) que fuerzan una competición para ganar la actividad neuronal.
3. Representación de características: que concierne a la ordenación geométrica de los vectores peso de las unidades competitivas.

Validación

Alegre (2003), una vez que ha terminado el proceso de aprendizaje y los pesos de la red neuronal han sido calculados, es importante comprobar la calidad del modelo resultante. Por ejemplo, en el caso del aprendizaje supervisado, una medida de la calidad puede darse en términos de los errores entre los valores de salida deseados y los obtenidos por la red neuronal. Algunas medidas estándar del error son:

1. El Error Cuadrático Medio (ECM), definido como:

$$ECM = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y}_i)^2 \quad (6)$$

2. El Error Porcentual Medio (EPM), definido como:

$$EPM = \frac{100}{n} \sum_{i=1}^n \frac{(y_i - \bar{y}_i)^2}{y_i} \quad (7)$$

3. El Error Absoluto Medio (EAM), definido como:

$$EPM = \frac{100}{n} \sum_{i=1}^n \frac{(y_i - \bar{y}_i)^2}{y_i} \quad (8)$$

Donde n representa el número de elementos que intervienen en el pronóstico, y_i representa el valor real y \bar{y}_i representa la salida pronosticada.

RESULTADOS Y DISCUSIÓN

Modelización de caudales medios mensuales del río Huancané, período 1964-2002

Debido a que el número de entradas está en función del tamaño del banco de datos disponibles, es decir, que el número de entradas puede ser tan grande como el tamaño del banco de datos y que este tipo de modelos permiten cualquier combinación de las variables de entrada (Tokar & Joshnon, 1999 citado por Weigert, 2005), en esta investigación se han formulado 6 modelos distintos (Ver Tabla 2) y el tipo de red usada en el modelamiento es del tipo Perceptron MultiCapas (MultiLayer Perceptron-MLP)

(Pulido et al., 2002), que comprende tres capas de entrada, intermedias y de salida; donde la capa de entrada tiene un número variable de neuronas en función al modelo propuesto, el número de neuronas de la capa intermedia varía entre 1 a 20 neuronas (según el mes y modelo), y la capa de salida con una única neurona de salida para todos los modelos propuestos.

Tabla 2. Modelos propuestos

Nº	Modelos
1	$Q_t = f(P_t, E_t)$
2	$Q_t = f(P_t, P_{t-1}, E_t, E_{t-1})$
3	$Q_t = f(P_t, P_{t-1}, P_{t-2}, E_t, E_{t-1}, E_{t-2})$
4	$Q_t = f(P_t, E_t, Q_{t-1})$
5	$Q_t = f(P_t, P_{t-1}, E_t, E_{t-1}, Q_{t-1})$
6	$Q_t = f(P_t, P_{t-1}, P_{t-2}, E_t, E_{t-1}, E_{t-2}, Q_{t-1}, Q_{t-2})$

Donde: Q es el caudal medio mensual (m³/s), P es precipitación media mensual (mm/mes) y E es la evaporación mensual (mm/mes).

El proceso de entrenamiento de los modelos de redes neuronales artificiales se efectuó con el primer subconjunto de la serie de caudales medios mensuales del río Huancané, que comprende un total de 33 años de registro (1964-1996). El algoritmo de aprendizaje utilizado en el entrenamiento de todos los modelos de redes neuronales artificiales es el algoritmo de retro-propagación (back-propagation), la función de activación tipo sigmoidea y la tasa de aprendizaje y momento igual a 0.6 y 0.8 respectivamente; donde la cantidad de ciclos de cálculo (épocas) para cada mes de los 6 modelos propuestos capaces de minimizar el error cuadrático varía entre 376 y 46 122 ciclos.

Tabla 3. Estadísticas de desempeño de modelos-fase de entrenamiento periodo 1964-1996

Nº	Modelos	Estadísticas					Calificación
		ECM	EAM	EPM	EPmax	R ²	
1	$Q_t = f(P_t, E_t)$	92.17	4.99	40.15	441.63	0.87	6
2	$Q_t = f(P_t, P_{t-1}, E_t, E_{t-1})$	69.27	3.72	25.13	375.23	0.90	4
3	$Q_t = f(P_t, P_{t-1}, P_{t-2}, E_t, E_{t-1}, E_{t-2})$	16.53	2.29	23.71	228.74	0.98	1
4	$Q_t = f(P_t, E_t, Q_{t-1})$	83.79	4.30	24.78	290.98	0.88	5
5	$Q_t = f(P_t, P_{t-1}, E_t, E_{t-1}, Q_{t-1})$	76.82	3.70	19.88	341.27	0.89	3
6	$Q_t = f(P_t, P_{t-1}, P_{t-2}, E_t, E_{t-1}, E_{t-2}, Q_{t-1}, Q_{t-2})$	27.28	2.71	22.10	284.63	0.96	2

Donde :1 representa el mejor modelo y 6 el menos eficiente.

Tabla 4. Estadísticas de desempeño de modelos-fase de validación periodo 1997-2002

N°	Modelos	Estadísticas					Calificación
		ECM	EAM	EPM	EPmax	R ²	
1	$Q_t = f(P_t, E_t)$	263.19	8.59	81.51	1404.34	0.82	5
2	$Q_t = f(P_t, P_{t-1}, E_t, E_{t-1})$	309.28	8.54	77.91	1987.55	0.79	6
3	$Q_t = f(P_t, P_{t-1}, P_{t-2}, E_t, E_{t-1}, E_{t-2})$	237.49	7.98	65.59	1068.29	0.85	4
4	$Q_t = f(P_t, E_t, Q_{t-1})$	256.31	6.62	39.24	664.39	0.83	2
5	$Q_t = f(P_t, P_{t-1}, E_t, E_{t-1}, Q_{t-1})$	245.51	6.80	38.50	524.03	0.83	1
6	$Q_t = f(P_t, P_{t-1}, P_{t-2}, E_t, E_{t-1}, E_{t-2}, Q_{t-1}, Q_{t-2})$	297.12	7.15	54.21	914.08	0.84	3

Donde: 1 representa el mejor modelo y 6 el menos eficiente.

En la Tabla 3, se observa claramente que el modelo 3 muestra mejor desempeño en el proceso de entrenamiento de los modelos de redes neuronales, ya que presenta errores con valores menores en comparación al resto de modelos propuestos y coeficiente de ajuste del orden de 98% (Coeficiente de determinación).

El proceso de prueba y validación de los modelos se efectuó con el segundo subconjunto de datos de la serie de caudales medios mensuales del río Huancané, que comprende un total de seis años de registro (1997-2002). Los resultados de las estadísticas de validación de los modelos propuestos se muestran en la Tabla 4.

En la fase de validación el modelo 5 es el que presenta un mejor desempeño en comparación al resto de modelos propuestos, debido a los bajos errores y coeficiente de ajuste alto que presenta.

Se efectuó la selección de los mejores modelos teniendo como principal criterio

de evaluación, la capacidad de desempeño y generalización a través de las estadísticas de validación de los modelos, tanto para la fase de entrenamiento como para la de validación. (Ver Tablas 3 y 4). Los modelos seleccionados fueron: el modelo 3, que presenta mejor desempeño dentro de los 3 modelos que consideran como variables de entrada valores de precipitación y evaporación mensual (Ver Figura 5a) y el modelo 5 como el que presenta mejor desempeño dentro de los 3 modelos que consideran como variables de entrada valores de precipitación, evaporación, y caudales medios mensuales (Ver Figura 5b)

Previsión de caudales medios mensuales con horizonte mensual

a) Modelamiento Estocástico de la serie mensual de caudales del río Huancané
De acuerdo a los análisis realizados, se optó por el Modelo Autorregresivo Periódico de Primer Orden PAR (1), como el modelo

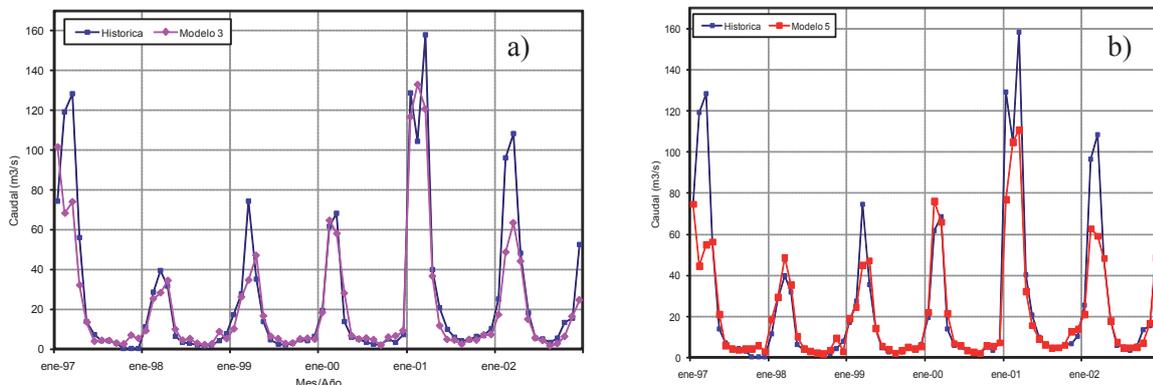


Fig. 5. Caudales medios mensuales observados y estimados con los modelos 3. Fig. 5a) y el modelo 5. Fig. 5b)

más adecuado para representar la serie de caudales medios mensuales del río Huancané, estación Puente Huancané; los parámetros para el modelo PAR (1) fueron estimados por el método de Momentos.

Para el análisis de bondad del modelo se generaron 10 series de 33 años de registro (Tablas 5 y 6), del cual se puede afirmar que no existe diferencia estadística significativa

entre las medias y desviaciones estándar de las series generadas y la serie histórica a un 95% de probabilidad, ya que los valores de T y F calculados son menores a los valores tabulares, concluyéndose que el modelo de generación PAR (1) es adecuado ya que reproduce satisfactoriamente la media y desviación estándar de la serie histórica.

Tabla 5. Comparación estadística de las medias

Mes	Media		Prueba Estadística T		Diferencia Significativa
	Histórica	Generada	T Calculada	T Tabular	
Enero	48.0103	50.6377	-1.2986	2.0195	No
Febrero	61.7209	61.9409	-0.1019	2.0195	No
Marzo	50.5645	49.4992	-0.6263	2.0195	No
Abril	28.0836	26.9610	-0.7540	2.0195	No
Mayo	10.0091	9.7133	-0.3466	2.0195	No
Junio	5.0882	5.0152	-0.1355	2.0195	No
Julio	3.7306	3.6894	-0.1009	2.0195	No
Agosto	2.7112	2.7080	-0.0096	2.0195	No
Septiembre	2.6848	2.4915	-0.4797	2.0195	No
Octubre	3.0227	2.8929	-0.2860	2.0195	No
Noviembre	5.3885	5.5605	-0.2365	2.0195	No
Diciembre	13.8873	14.1852	-0.2394	2.0195	No

Tabla 6. Comparación estadística de las desviaciones estándar

Mes	Desviación Estándar		Prueba Estadística F		Diferencia Significativa
	Histórica	Generada	F Calculada	F Tabular	
Enero	30.9071	33.2269	1.0751	2.1888	No
Febrero	36.0375	34.7080	1.0383	2.8543	No
Marzo	22.1623	22.3468	1.0083	2.1888	No
Abril	17.4943	15.2912	1.1441	2.8543	No
Mayo	5.7909	4.8692	1.1893	2.8543	No
Junio	2.3397	1.8271	1.2806	2.8543	No
Julio	1.3258	1.1171	1.1868	2.8543	No
Agosto	0.8779	0.7743	1.1338	2.8543	No
Septiembre	1.3088	1.0243	1.2778	2.8543	No
Octubre	1.6331	1.3961	1.1698	2.8543	No
Noviembre	3.9503	4.4417	1.1244	2.1888	No
Diciembre	11.4869	13.3059	1.1584	2.1888	No

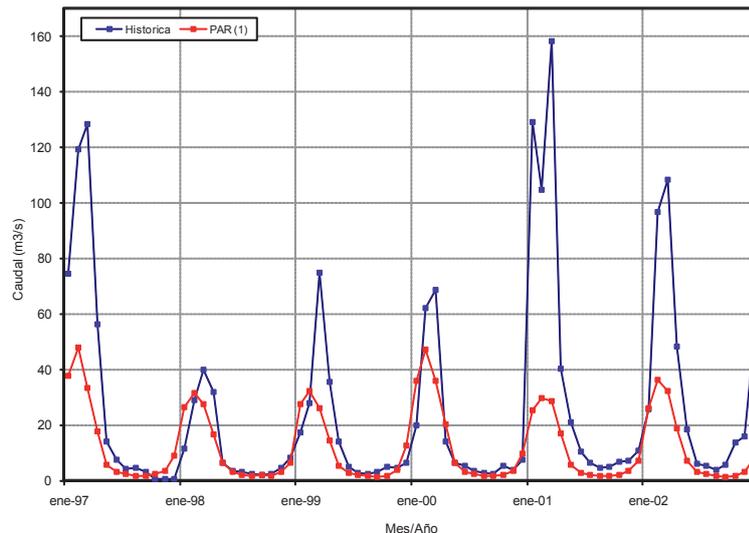


Fig. 6. Caudales medios mensuales observados y estimados con el modelo PAR (1)

Una vez comprobada la bondad del modelo PAR (1), se efectuó la previsión con horizonte mensual de los caudales medios mensuales del río Huancané, tomando como valor inicial de previsión (semilla) el mes de diciembre del año 1996 (12.57 m³/s). En la Figura 6, se muestra la representación gráfica de la comparación de los valores observados y estimados (previstos) de la serie de caudales medios mensuales del río Huancané, estación Puente Huancané correspondientes al periodo 1997-2002.

b) Modelamiento con Redes Neuronales Artificiales

Para esta investigación se ha utilizado 4 modelos distintos (Ver Tabla 7) y el tipo de red usada en el modelamiento es Perceptron MultiCapas (MultiLayer Perceptron-MLP) (Pulido et al, 2002), que comprende tres capas de entrada, intermedias y de salida, donde la capa de entrada tiene un número variable de neuronas en función al modelo propuesto, el número de neuronas de la capa intermedia varía entre 1 a 10 neuronas (según el mes y modelo), y la capa de salida con una única neurona de salida para todos los modelos propuestos.

Tabla 7. Modelos propuestos

Nº	Modelos
1	$Q_t = f(Q_{t-1})$
2	$Q_t = f(Q_{t-1}, Q_{t-2})$
3	$Q_t = f(Q_{t-1}, Q_{t-2}, Q_{t-3})$
4	$Q_t = f(Q_{t-1}, Q_{t-2}, Q_{t-3}, Q_{t-11}, Q_{t-12}, Q_{t-13})$

Donde: Q es el caudal medio mensual (m³/s).

El proceso de entrenamiento de los modelos de redes neuronales artificiales se efectuó con el primer subconjunto de la serie de caudales medios mensuales del río Huancané, que comprende un total de 33 años de registro (1964-1996). El algoritmo de aprendizaje utilizado en el entrenamiento de todos los modelos de redes neuronales artificiales es el de retropropagación (back-propagation), la función de activación tipo sigmoidea y tasa de aprendizaje y momento igual a 0.6 y 0.8 respectivamente; donde la cantidad de ciclos de cálculo (épocas) para cada mes de los 4 modelos propuestos capaces de minimizar el error cuadrático varía entre 140 y 17 526 ciclos.

En la Tabla 8, se observa que el modelo 4 muestra mejor desempeño en el proceso de entrenamiento de los modelos de redes neuronales, ya que presenta errores con valores menores en comparación al resto de modelos propuestos y coeficiente de ajuste del orden de 93% (Coeficiente de determinación).

El proceso de prueba y validación de los modelos se efectuó con el segundo subconjunto de datos de la serie de caudales medios mensuales del río Huancané que comprende un total de 6 años de registro (1997-2002). Los resultados de las estadísticas de validación de los modelos propuestas se muestran en la Tabla 9.

Tabla 8. Estadísticas de desempeño de modelos-fase de entrenamiento periodo 1964-1996

N°	Modelos	Estadísticas					Calificación
		ECM	EAM	EPM	EPmáx	R ²	
1	$Q_t = f(Q_{t-1})$	203.40	7.31	50.25	788.83	0.71	4
2	$Q_t = f(Q_{t-1}, Q_{t-2})$	175.08	6.22	44.77	887.24	0.75	3
3	$Q_t = f(Q_{t-1}, Q_{t-2}, Q_{t-3})$	62.02	4.09	32.63	767.46	0.91	2
4	$Q_t = f(Q_{t-1}, Q_{t-2}, Q_{t-3}, Q_{t-11}, Q_{t-12}, Q_{t-13})$	54.69	3.30	23.77	294.47	0.93	1

Donde: 1 representa el mejor modelo y 4 el menos eficiente.

En la fase de validación el modelo 2 es el que presenta mejor desempeño en comparación al resto de modelos propuestos, debido a los bajos errores y coeficiente de ajuste alto que presenta.

Se efectuó la selección del mejor modelo de RNA, teniendo como principal criterio de evaluación la capacidad de desempeño y generalización a través de las estadísticas de validación de los modelos, tanto para la fase de entrenamiento, como para la de validación. (Ver Tablas 8 y 9), eligiéndose al modelo 3 como el que presenta mejor desem-

peño dentro de los 4 modelos de previsión, que consideran como variables de entrada valores de caudales medios mensuales. (Ver Fig. 7)

c) Comparación de Modelos

Uno de los objetivos de esta investigación fue evaluar la capacidad de desempeño de los modelos de RNA como instrumento de previsión de caudales medios mensuales del río Huancané, frente a los modelos de serie temporal estocástica PAR (1), por lo que es

Tabla 9. Estadísticas de desempeño de modelos-fase de validación periodo 1997-2002

N°	Modelos	Estadísticas					Calificación
		ECM	EAM	EPM	EPmax	R ²	
1	$Q_t = f(Q_{t-1})$	442.55	10.57	78.58	1390.19	0.74	4
2	$Q_t = f(Q_{t-1}, Q_{t-2})$	346.38	9.37	69.69	1180.64	0.78	1
3	$Q_t = f(Q_{t-1}, Q_{t-2}, Q_{t-3})$	386.79	9.37	60.20	1220.15	0.72	2
4	$Q_t = f(Q_{t-1}, Q_{t-2}, Q_{t-3}, Q_{t-11}, Q_{t-12}, Q_{t-13})$	587.37	10.99	58.80	714.46	0.60	3

Donde: 1 representa el mejor modelo y 4 el menos eficiente.

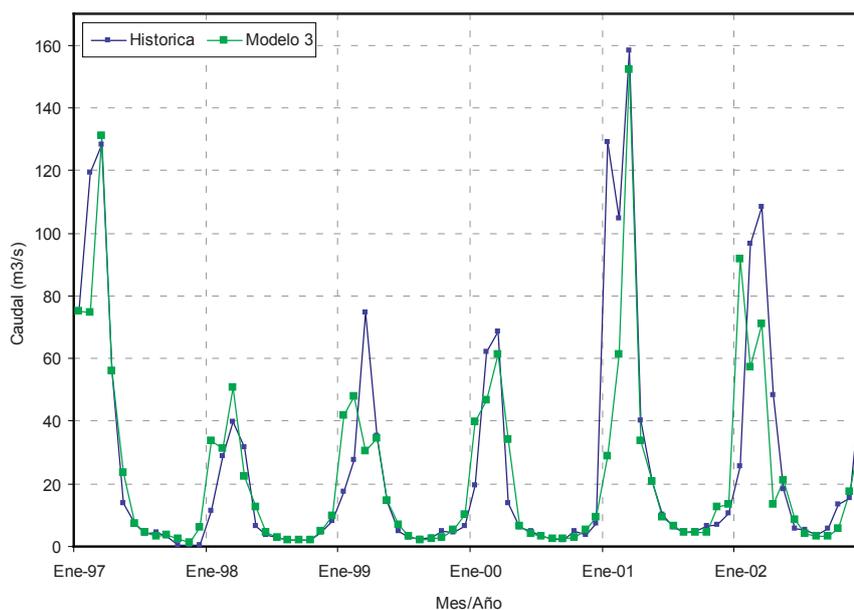


Fig. 7. Caudales medios mensuales observados y estimados con el modelo 3.

coherente pensar que los datos utilizados para calibrar uno y otro modelo deben ser semejantes, sin embargo, independientemente del conjunto de datos de entrada utilizados en la calibración y validación de ambos modelos, se ha efectuado la comparación del mejor modelo de RNA y el modelo de serie temporal estocástica PAR (1), para ello se tiene como antecedentes comparaciones efectuadas en otros trabajos. (Ballini, 2000; Pulido et al., 2002; Zuñiga & Jordán, 2005; Ferreira & Forain, 2007; Figueiredo et al., 2007; Gomes & Lima, 2010).

En la comparación de los modelos de previsión, se evaluó el desempeño de los mismos en términos de medida de errores, coeficiente de ajuste (R²) y análisis gráfico del mejor modelo de redes neuronales artificiales con los resultados del modelo de serie temporal estocástica PAR (1). En el Tabla 10 se muestra un resumen de los resultados estadísticos.

Tabla 10. Resumen de resultados estadísticos

Modelos	Estadísticas				
	ECM	EAM	EPM	EPmax	R ²
PAR (1)	961.03	15.90	86.14	1,858.49	0.59
RNA Modelo 3	386.79	9.37	60.20	1,220.15	0.72

De la Tabla 10, puede decirse que se

confirma de cierta forma la necesidad de utilizar modelos de redes neuronales artificiales (RNA) en la previsión mensual de caudales medios mensuales del río Huancané, ya que se consiguió una mayor efectividad en los resultados de previsión con valores muy cercanos a los reales (Ver Figura 8), obteniéndose medidas de error inferiores en comparación al modelo PAR (1), y un porcentaje de ajuste de 72% en comparación al 59% que presenta el modelo PAR (1), lo que confirma la robustez y gran ventaja de la modelación inteligente, en comparación a un modelo de serie temporal estocástica PAR (1).

CONCLUSIONES

Los modelos de redes neuronales artificiales desde la previsión de caudales medios mensuales muestran un excelente comportamiento en términos de las estadísticas de los errores de previsión; más aún cuando de manera comparativa con un modelo autorregresivo periódico de primer orden-PAR (1) muestra grados superiores de eficiencia en las previsiones, ya que presenta un porcentaje de ajuste de 72%, en comparación al 59% que presenta el modelo de serie de tiempo estocástica.

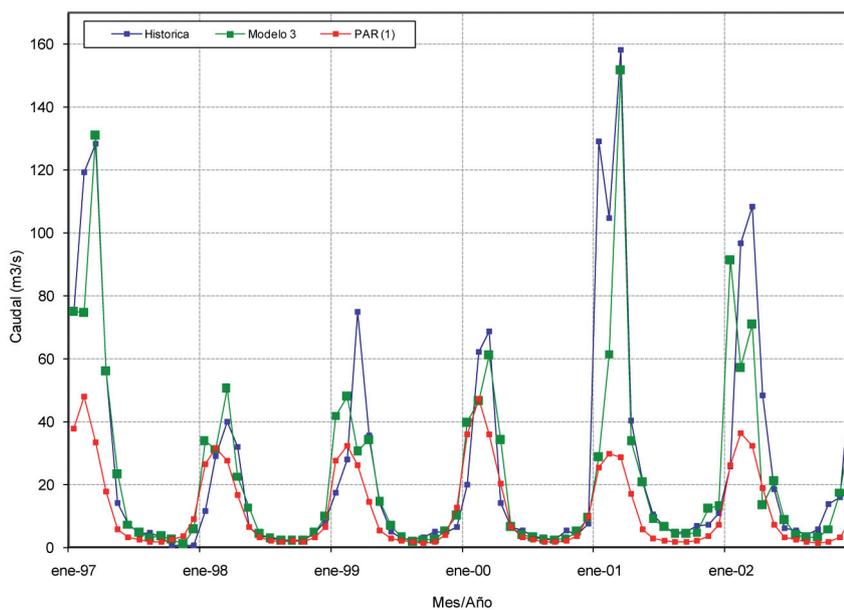


Fig. 8. Caudales medios mensuales observados y estimados.

Estos resultados son altamente promisorios y sugieren que el enfoque propuesto tiene mejores rendimientos que los modelos de series temporales estocásticas en términos de error de predicción, eficiencia y estimación de picos.

En esta investigación se logró desarrollar modelos de redes neuronales de gran efectividad, a partir de datos de caudales, precipitación y evaporación, lo que demuestra la capacidad de los modelos de redes neuronales artificiales en la modelización de caudales medios mensuales a partir de este tipo de información.

Independientemente del conjunto de entradas utilizado, los modelos de redes neuronales artificiales mostraron gran desempeño en la modelización de caudales medios mensuales, siendo el modelo 3 ($Q_t = f(P_t, P_{t-1}, P_{t-2}, E_t, E_{t-1}, E_{t-2})$) el que presenta mejor desempeño dentro de los modelos que consideran como variables de entrada valores de precipitación y evaporación mensual y el modelo 5 ($Q_t = f(P_t, P_{t-1}, P_{t-2}, E_t, E_{t-1}, E_{t-2}, Q_{t-1}, Q_{t-2})$) presenta mejor desempeño dentro de los 3 modelos que consideran como variables de entrada valores de precipitación, evaporación, y caudales medios mensuales.

Las redes neuronales artificiales para la concepción de modelos aplicados a la modelización lluvia-caudal y posterior previsión de caudales medios mensuales, muestran resultados satisfactorios. El hecho de que los modelos de redes neuronales artificiales exhiban un desarrollo comparable e incluso mejor que los modelos estocásticos sugiere que esta investigación puede proveer una herramienta útil en la solución de tipos similares de problemas en la gestión y estudio de los recursos hídricos.

Agradecimientos: A mi Alma Mater la Universidad Nacional del Altiplano y la Universidad Nacional Agraria La Molina – Maestría en Recursos Hídricos, así como mi agradecimiento Comité editor de la Revista Peruana Geo-Atmosférica del SENAMHI por la invitación y publicación de este trabajo de investigación.

BIBLIOGRAFÍA

Alegre, A. (2003). Simulación de redes neuronales artificiales: Una aplicación didáctica. Tesis Lic. Sistemas, Universidad Nacional del Nordeste, Corrientes, Argentina. 188 p.

Baganha, A. (1999). Modelagem chuva – vazão em bacias hidrográficas com suporte em redes neurais artificiais. Tesis Dr., Universidade Estadual de Campinas, Brasil. 279 p.

Ballini, R. (2000). Análise e previsão de vazões utilizando modelos de séries temporais, redes neurais e redes neurais nebulosas. Tesis Dr., Universidade Estadual de Campinas, Brasil. 183 p.

Ballini, R., Soares, S., & Andrade, M. (2003). Previsão de vazões usando redes neurais nebulosas. Revista Controle & Automação 14(3) 286-297.

Bravo, J., Pilar, J., & Depettris, C. (2004). Utilización de un modelo de redes neuronales en la obtención de la curva de descarga de la obra control del río Negro en Laguna Blanca. Universidad Nacional del Nordeste, Corrientes, Argentina. 4 p.

Bravo, J., Correa, O., Collischonn, W., Rolim, A., & Bertacchi, C. (2008).

Previsões de curto prazo de vazão afluente ao reservatório de furnas utilizando redes neurais artificiais. Revista Brasileira de Recursos Hídricos RBRH 13(2) 77-88.

Cruz, M., Rodrigues, L., & Versiani, B. (2010). Previsão de vazões com a metodologia DPFT e com redes neurais artificiais. Revista Brasileira de Recursos Hídricos RBRH 15(1) 121-132.

Ferreira, L., & Forain, V. (2007). Comparação de métodos de previsão de vazões naturais afluentes a aproveitamentos hidroelétricos. Revista Brasileira de Recursos Hídricos RBRH 12(3) 13-20.

Figueiredo, K., Barbosa, C., Da Cruz, A., Vellasco, M., Pacheco, M., Contreras, R., Barros, M., Castro, R., Marques, V., & Duarte, U. (2007). Modelo de previsão de vazão com informação de precipitação utilizando redes neurais. Revista Brasileira de Recursos Hídricos RBRH 12(3) 69-82.

Gibilini, D. (2006). Aplicação de técnicas de inteligência Artificial na alocação dinâmica

- de canais em redes sem fio. Tesis Mg. Sc., Universidade de São Paulo, Brasil. 166 p.
- Gomes, L., & Lima, S. (2010).** Modelo baseado na técnica de redes neurais para previsão de vazões na bacia do rio São Francisco. *Revista Brasileira de Recursos Hídricos RBRH* 15(1) 5-15.
- Hillman, G., Dölling, O., Pagot, M., Pozzi, C., & Plencovich, G. (2004).** Aplicación de redes neuronales para el pronóstico de evolución de niveles de agua para la laguna de Mar Chiquita. *Mecánica Computacional*. 23, 1245-1259.
- Jaimes, W. (2006).** Modelado inteligente para soportar decisiones del sector de navegación en el río Magdalena, Trabajo de Grado Ing. Sistemas, Universidad de Pamplona, Colombia. 109 p.
- Kilinç, I., & Ciğizoglu, K. (2003).** Reservoir management using artificial neural networks. *State Hydraulic Works*. Istanbul, Turquía. 9 p.
- Laqui, W. (2008).** Predicción de caudales medios mensuales del río Huancané utilizando modelos de redes neuronales artificiales. Tesis Mg. Sc., UNALM, Lima, Perú. 149 p.
- Magalhães, M. (2004).** Redes neurais, metodologías de agrupamento y conveção de previsores aplicados à previsão de vazões naturais. Tesis Mg. Sc., Universidade Estadual de Campinas, Brasil. 123 p.
- Nina, C. (2008).** Modelamiento hidrológico de la cuenca del río Huancané. Tesis Ing. Agrícola. Universidad Nacional del Altiplano. Puno, Perú.
- Ovando, G., Bocco, M., & Sayazo, S. (2005).** Redes neuronales para modelar predicción de heladas. *Agricultura Técnica*. 65(1), 65-73.
- Pulido, I., Roldán, R., López, R., & Gutiérrez, J. (2002).** Técnicas de predicción a corto plazo de la demanda de agua: Aplicación al uso agrícola. *Ingeniería del Agua*. 9(3), 319-331.
- Santos, C., & Pereira, A. (2002).** Modelagem hidrológica urbana por meio utilizando redes neurais artificiais: Uma aplicação para a bacia do rio Tamandateí em Sao Paulo, SP. *Revista Brasileira de Meteorología*. 18(2), 149-159.
- Sovero, H. (1991).** Modelo matemático de transformación precipitación – descarga con aplicación a la cuenca del río Huancané. Tesis Ing. Agrícola, UNALM, Lima, Perú. 125 p.
- Tokar, A., & Markus, M. (2000).** Precipitation-runoff modeling using artificial neural networks and conceptual models. *Journal of Hydrologic Engineering*. 4(4), 156-161.
- Weigert, F. (2005).** Modelagem chuva-vazão mensal utilizando redes neurais artificiais. Tesis Mg. Sc., Universidade Federal do Paraná, Curitiba, Brasil. 174 p.
- Zuñiga, A., & Jordán, C. (2005).** Pronóstico de caudales medios mensuales empleando sistemas neuro-fuzzy. *Revista Tecnológica de la Escuela Superior Politécnica del Litoral*. 18(1), 17-23.