

Predicción de caudales medios mensuales del Río Huancané utilizando modelos de redes neuronales

Miguel Sánchez D.¹, Wilber Laqui V.²

Resumen

Esta investigación centró sus esfuerzos en la predicción de caudales medios mensuales del río Huancané, desarrollo de modelos de redes neuronales a partir de datos de caudales, precipitación y evaporación, así como la evaluación de la capacidad de desempeño frente a modelos estocásticos. De esta manera, se desarrollaron 10 modelos de redes neuronales artificiales con distintas arquitecturas, cuyo entrenamiento se realizó con un primer grupo de datos correspondientes al periodo 1964 – 1996, y su validación con un segundo grupo de datos del periodo 1997 – 2002. Los modelos de redes neuronales artificiales mostraron comparativamente mejor desempeño en materia de predicción frente a un modelo autorregresivo periódico de primer orden PAR (1).

Palabras clave: *Predicción de caudales, redes neuronales artificiales, series de tiempo.*

Abstract

This research focused its efforts on the prediction of average monthly discharge of Huancane River, the development of neuron networks from data discharge, precipitation and evaporation, as well as the evaluation of their performance compared to stochastic models. In this manner, 10 artificial neuron networks were developed with distinct architectures, which were trained with the first subset of data corresponding to the 1964 – 1996 period and validated with second subset of data form the 1997 – 2002 period. The artificial neuron network models showed comparatively better performance in the area of prediction compared to a first-order PAR (1) periodic autoregressive model.

Key words: *Prediction of volumes, artificial neuron networks, time series.*

1. 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 importantes 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 los recursos hídricos.

Dada la importancia de la previsión de caudales, durante las últimas décadas se han propuesto gran variedad de modelos, entre ellos, los modelos de alisamiento exponencial, modelos de regresión lineal y no lineal, modelos de series de tiempo estocásticas y muy recientemente se han introducido nuevos modelos de previsión, que manejan de una manera eficiente la aleatoriedad de los caudales y presentan una fácil adaptabilidad a nuevos datos, sin la necesidad de incurrir en laboriosas formulaciones matemáticas. Entre estos modelos se incluyen las técnicas de inteligencia artificial, que comprenden las técnicas de Redes Neuronales, Sistemas Fuzzy y Neuro-Fuzzy (Zúñiga y Jordán, 2005), Teoría Nebulosa, Sistemas Multi-Agentes.

Debido a la gran importancia que tiene la previsión de caudales, es necesaria la adopción de estos nuevos

que permitan conocer con cierta exactitud y facilidad los caudales medios mensuales del río Huancané, los que permitirán su uso racional para diferentes fines.

Los objetivos planteados en esta investigación son:

Objetivo General

Predicir los caudales medios mensuales del río Huancané utilizando modelos de redes neuronales artificiales.

Objetivos Específicos

- Desarrollar modelos predictivos de caudales medios mensuales del río Huancané, basados en redes neuronales artificiales a partir de datos de caudales, precipitación y evaporación.

- Determinar la capacidad de desempeño de los modelos de redes neuronales artificiales como instrumento de modelización y predicción de caudales medios mensuales del río Huancané frente a modelos de series temporales estocásticas.

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 si. Ver Figura 1.

Bravo et al. (2004), dice que las Redes Neuronales Artificiales se presentan como un nuevo modelo empírico (o del tipo "caja negra"). Son modelos matemáticos que intentan imitar el funcionamiento del cerebro humano, a su vez lo define como un sistema de ecuaciones, en general no lineales, interconectados, en los que el resultado de una ecuación es el valor de entrada de otras encadenadas a ella a través de la red. Son considerados como aproximadores universales de funciones y, por lo

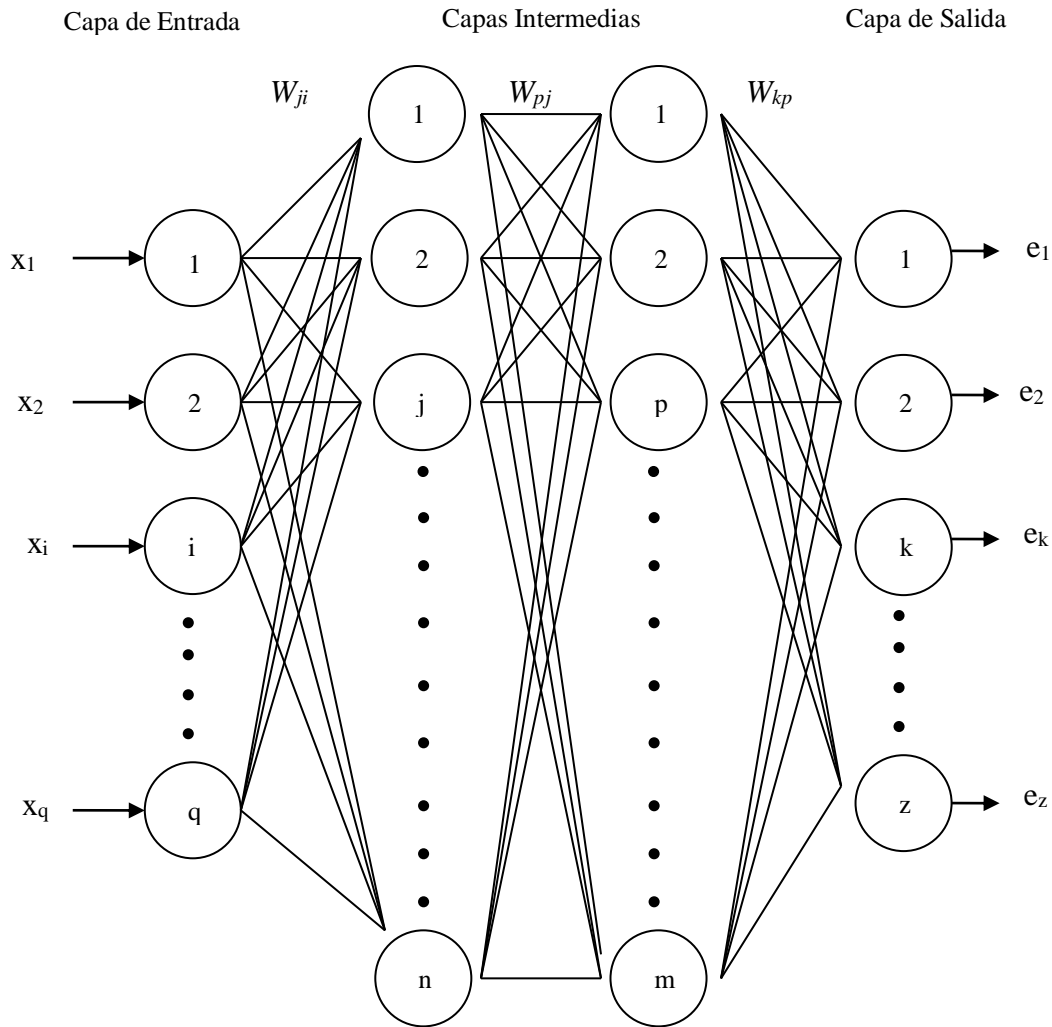
¹ Facultad de Ingeniería Agrícola, Universidad Nacional Agraria La Molina. E-mail: msanchez@lamolina.edu.pe.

² Ingeniero Agrícola, Maestría en Recursos Hídricos, E-mail: wilfer.lv@gmail.com

modelos como instrumento de modelización y predicción, para el desarrollo de modelos predictivos

tanto, tendrían la capacidad de conseguir reproducir el comportamiento de cualquier función matemática.

Figura 1. 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).



Componentes de las redes neuronales

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

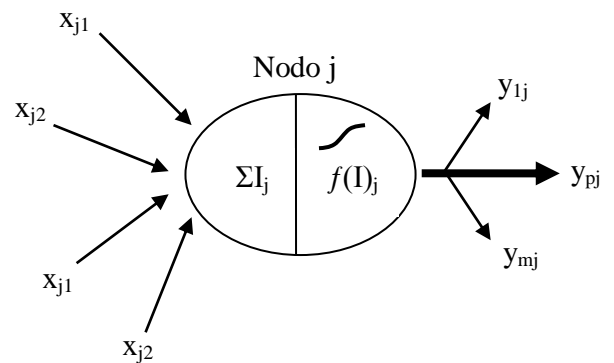
a) Neurona ó unidad procesadora

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. De este modo, dado un conjunto de entradas x_{ji} ($i=1, \dots, q$) en la neurona j , estos valores se multiplican por el conjunto de pesos W_{ji} (conexión del nodo i con el nodo j). Ver Figura 2. (Pulido et al, 2002).

$$I_j = \sum_{i=1}^q x_{ji} W_{ji}$$

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

Figura 2. Elemento de proceso o neurona (Fuente: Pulido et al, 2002).



Funciones de activación

A continuación se describen las funciones de activación más populares:

Funciones lineales: Son funciones que dan una salida lineal:

$f(x) = x; x \in \mathbb{R}$

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:

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

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

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

La función logística:

$$y_{pj} = f(I_j) = \frac{1}{1 + e^{-I_j}}$$

La función tangente hiperbólica: de -1 a 1:

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

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):

1. Aprendizaje supervisado: En este caso, los patrones para el aprendizaje están formados por parejas, $f(a_p; b_p)$ $p = 1; \dots; rg$, 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 y, por lo tanto, el proceso de aprendizaje puede no converger al mínimo global óptimo.

2. 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:

a) Aprendizaje Hebbiano: consiste en modificar los pesos de acuerdo con algún criterio de correlación entre las actividades neuronales.

b) Aprendizaje competitivo: neuronas diferentes se conectan con pesos negativos (inhibitorios) que fuerzan una competición para ganar la actividad neuronal.

c) Representación de características: que concierne a la ordenación geométrica de los vectores peso de las unidades competitivas.

VALIDACIÓN

Alegre (2003), dice que 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), definida como:

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

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

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

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

$$ECM = \frac{1}{n} \sum_{i=1}^n |y_i - \bar{y}_i|^2$$

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.

2. Materiales y métodos

MATERIALES

Dentro de los materiales utilizados en la ejecución de esta investigación se tienen:

- Equipo de Cómputo e impresión.
- Programas de computo (Microsoft Word, Microsoft Excel, Microsoft Power Point, AutoCAD, Surfer 8, Hec 04, Minitab 15, Alyuda NeuroIntelligence 2.2 Versión Trial (evaluación), SAMS, Matlab 6.5, etc.
- Información cartográfica: Cartas Nacionales escala 1:100,000 elaboradas por el Instituto Geográfico Militar – IGM, Hojas 30v, 30x, 30y, 31v, 31x, 31y.
- Información hidrometeorológica: Registros históricos de caudales medios mensuales, series históricas de precipitación total mensual y registro histórico de evaporación mensual.
- Material bibliográfico adquirido, existente y útiles de escritorio.

Características generales del área de estudio

Se identifica como área de estudio a la Cuenca del río Huancané, comprendida desde sus nacientes, hasta la Estación Hidrométrica Puente Huancané (próxima a la desembocadura al Lago Titicaca). La que políticamente se ubica en la región Puno, en las provincias de Huancané, San Antonio de Putina, San Pedro de Moho y Azángaro, se encuentra ubicada en la parte noreste de la región Puno, en el extremo sureste del Perú, (ver Figura 3). Geográficamente la cuenca del río Huancané está comprendida entre las coordenadas UTM; Este: 376,159 – 468,954 y Norte: 8'397,619 - 8'302,196, su cota máxima es 5,162 msnm y una cota mínima de 3,820 msnm, que corresponde al cerro Surupuna y a la desembocadura del río en el lago Titicaca, respectivamente.

La cuenca Hidrográfica del río Huancané, tienen una extensión aproximada de 3,521.98 km², desde la partición de aguas, hasta la desembocadura en el

Lago Titicaca, presenta una forma mas o menos “alargada”, ver Figura 4.

Figura 3. Plano de ubicación de la zona de estudio.

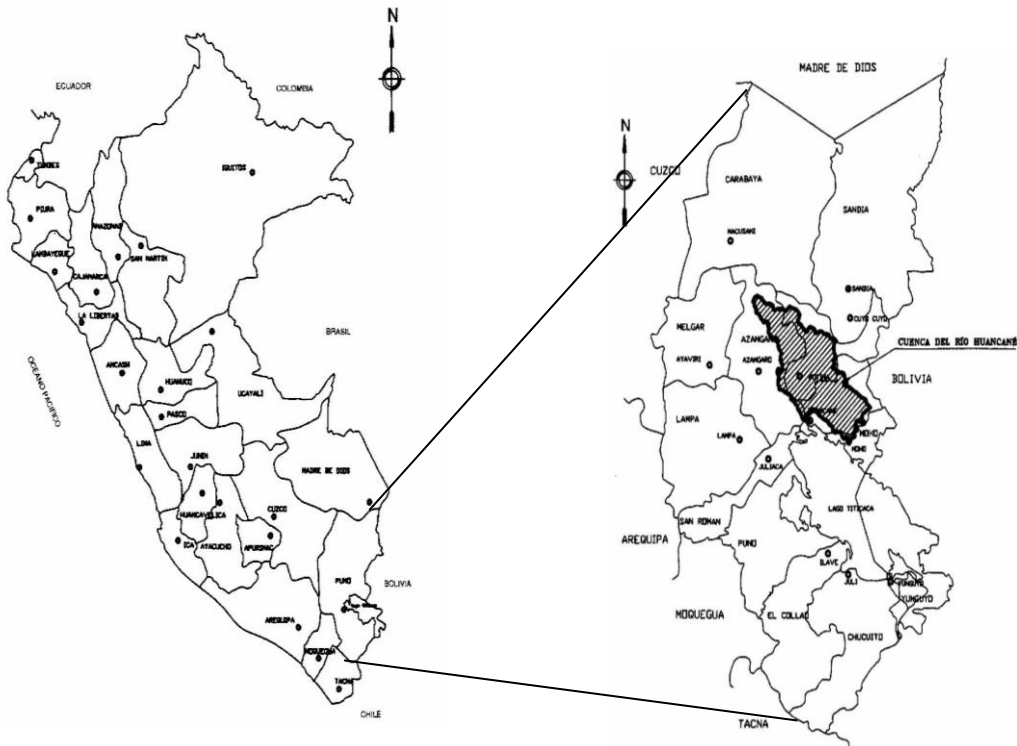
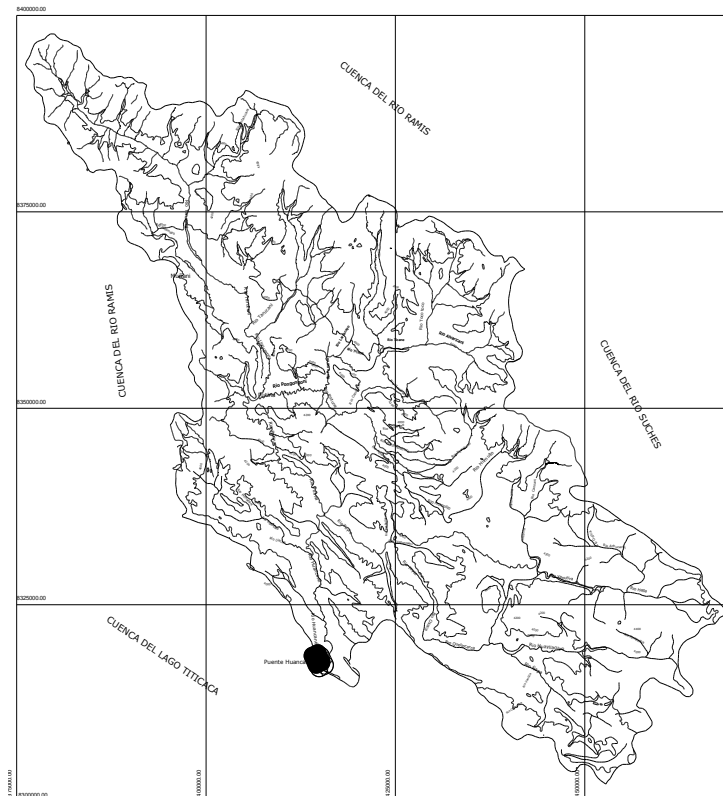


Figura 4. Cuenca hidrográfica del río Huancané



Métodos

La metodología seguida consta de cinco etapas, las que son:

- 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.
- 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.
- Predicción 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.

3. Resultados y discusión

3.1. Recolección y análisis de la información hidrometeorológica

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

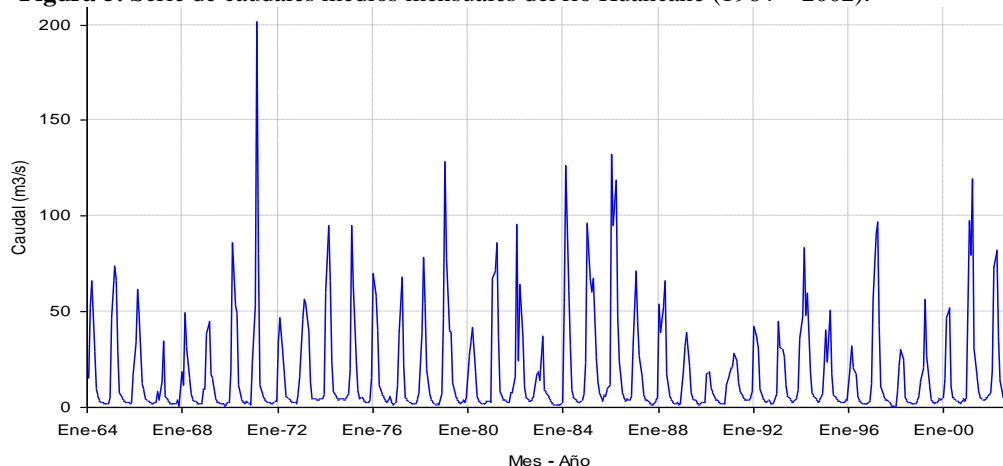
Tabla 1. Información Hidrometeorológica utilizada.

Estación	Longitud	Latitud	Altitud	Parámetro	Periodo	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

En la Figura 5, se presenta la serie temporal de caudales medios mensuales históricas del río

Huancané, con un total de 39 años de registro a partir de 1964 a 2002.

Figura 5. Serie de caudales medios mensuales del río Huancané (1964 – 2002).



La serie de caudales medios mensuales del río Huancané fue corregida en el periodo (Junio 1988 – Diciembre 2002), así como la serie de evaporación mensual de la estación Huaraya – Moho el periodo (Febrero 1984 – Diciembre 2000), sin embargo las series de precipitación total mensual de las estaciones en estudio no requirieron corrección alguna. Se ha efectuado la completación de la series de precipitación total mensual para las estaciones Taraco, Arapa, Huancané, Huaraya, 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. En la Tabla 2, se muestran los coeficientes pluviométricos calculados por el método de Thiessen Modificado, para la determinación de la precipitación media de la cuenca del río Huancané.

Tabla 2. Coeficientes Pluviométricos de Thiessen Modificado.

Estación	Coefficiente Pluviométrico
Crucero	0.0220
Muñani	0.2166
Ananea	0.0906
Putina	0.2667
Huancané	0.1739
Cojata	0.1619
Huaraya Moho	0.0744

3.2. Modelamiento estocástico de la serie mensual de caudales del Río Huancané

De acuerdo a los análisis realizados, se ha optado por un Modelo Autorregresivo Periódico de Primer Orden PAR (1), como el modelo más adecuado para representar la serie de caudales medios mensuales del río Huancané – estación Puente Huancané, los parámetros estimados por el método de Momentos para el modelo PAR (1), se presentan en la Tabla 3.

Tabla 3. Parámetros de modelo PAR (1)

Meses	Phi 1	Alfa 1
Enero	0.3959	0.3959
Febrero	0.4162	0.4162
Marzo	0.5263	0.5263
Abril	0.5540	0.5540
Mayo	0.8976	0.8976
Junio	0.9036	0.9036
Julio	0.8533	0.8533
Agosto	0.7706	0.7706
Septiembre	0.5886	0.5886
Octubre	0.6428	0.6428
Noviembre	0.4143	0.4143
Diciembre	0.5760	0.5760

De los resultados de la prueba de normalidad de asimetría y prueba de independencia de Porte Manteau de los residuos del modelo PAR (1). Puede decirse que los residuos del modelo siguen una distribución normal y son independientes los valores calculados no superan los valores tabulares (0.637 y 14.070, respectivamente), lo que lleva a asegurar, concluyentemente, que el modelo autorregresivo periódico PAR (1) es el más adecuado para la serie de caudales medios mensuales del río Huancané – estación Puente Huancané.

Tabla 5. Comparación estadística de las medias.

Mes	Media		Prueba Estadística T		Diferencia Significativa
	Historica	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

En la Tabla 4 se presentan las ecuaciones de generación mensual del modelo autorregresivo periódico de primer orden PAR (1).

Tabla 4. Ecuaciones de generación mensual modelo PAR (1).

Mes	Ecuación
Enero	$Z_{v,t} = 0.396 Z_{v,t-1} + E_{v,t}$
Febrero	$Z_{v,t} = 0.416 Z_{v,t-1} + E_{v,t}$
Marzo	$Z_{v,t} = 0.526 Z_{v,t-1} + E_{v,t}$
Abril	$Z_{v,t} = 0.554 Z_{v,t-1} + E_{v,t}$
Mayo	$Z_{v,t} = 0.898 Z_{v,t-1} + E_{v,t}$
Junio	$Z_{v,t} = 0.904 Z_{v,t-1} + E_{v,t}$
Julio	$Z_{v,t} = 0.853 Z_{v,t-1} + E_{v,t}$
Agosto	$Z_{v,t} = 0.771 Z_{v,t-1} + E_{v,t}$
Septiembre	$Z_{v,t} = 0.589 Z_{v,t-1} + E_{v,t}$
Octubre	$Z_{v,t} = 0.643 Z_{v,t-1} + E_{v,t}$
Noviembre	$Z_{v,t} = 0.414 Z_{v,t-1} + E_{v,t}$
Diciembre	$Z_{v,t} = 0.3576 Z_{v,t-1} + E_{v,t}$

Donde: $Z_{v,t}$ es el valor estandarizado de la serie mensual generada, $Z_{v,t-1}$ es el valor inicial estandarizado del modelo y $E_{v,t}$ es el ruido blanco. Para el análisis de bondad del modelo se generaron 10 series de 33 años de registro a partir de las ecuaciones de generación mensual mostradas en la Tabla 4, a las que previamente se les aplicó el proceso inverso de estandarización y transformación.

Tabla 6. Comparación estadística de las desviaciones estándar.

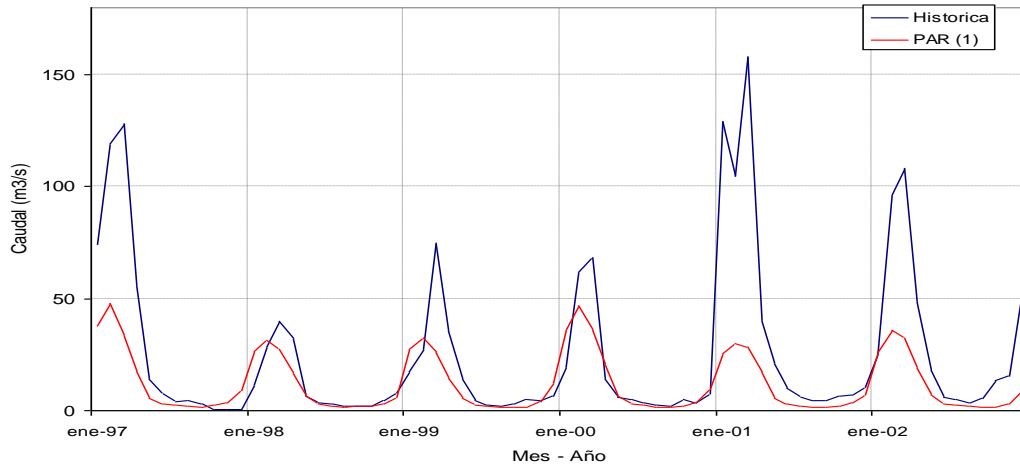
Mes	Desviación Estándar		Prueba Estadística F		Diferencia Significativa
	Historica	Generada	F Calculada	F Tabular	
Enero	30.9071	33.2269	1.0751	2.1888	No
Febrero	36.0375	34.7080	1.0383	2.8543	No
Marzo	22.1623	22.3468	1.0083	2.1888	No
Abril	17.4943	15.2912	1.1441	2.8543	No
Mayo	5.7909	4.8692	1.1893	2.8543	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

De los resultados mostrados en las Tablas 5 y 6, se puede afirmar que no existe diferencia estadísticamente 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.

Una vez comprobada la bondad del modelo PAR (1), se efectuó la predicción de caudales medios mensuales tomando como valor inicial de predicción (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 históricos y generados de la serie de caudales medios mensuales del río Huancané – estación Puente Huancané correspondientes al periodo 1997 – 2002.

Figura 6. Predicción de caudales mensuales usando el Modelo PAR (1).



3.3. Modelamiento con redes neuronales artificiales

Se ha formulado 10 modelos distintos (ver Tabla 7) y el tipo de red usada en el modelamiento, es del tipo Perceptron Multi – Capas (MultiLayer Perceptron – MLP), 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 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	$Q_t = f(Q_{t-1}, Q_{t-2}, Q_{t-3}, Q_{t-11}, Q_{t-12}, Q_{t-13})$
5	$Q_t = f(P_t, E_t)$
6	$Q_t = f(P_t, P_{t-1}, E_t, E_{t-1})$
7	$Q_t = f(P_t, P_{t-1}, P_{t-2}, E_t, E_{t-1}, E_{t-2})$
8	$Q_t = f(P_t, E_t, Q_{t-1})$
9	$Q_t = f(P_t, P_{t-1}, E_t, E_{t-1}, Q_{t-1})$
10	$Q_t = f(P_t, P_{t-1}, P_{t-2}, E_t, E_{t-1}, E_{t-2}, Q_{t-1}, Q_{t-2})$

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 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 10 modelos propuestos capaces de minimizar el error cuadrático varía entre 140 y 46,122 ciclos.

Figura 7. Esquema de la configuración de la red para cada mes del Modelo 10.



En la Tabla 8 se observa claramente que el modelo 7 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).

Tabla 8. Estadísticas de desempeño de modelos - Fase de entrenamiento.

Nº	Modelos	Estadísticas					Calificación
		EQM	EAM	EPM	EPmax	R ²	
1	$Q_t = f(Q_{t-1})$	203.40	7.31	50.25	788.83	0.71	10
2	$Q_t = f(Q_{t-1}, Q_{t-2})$	175.08	6.22	44.77	887.24	0.75	9
3	$Q_t = f(Q_{t-1}, Q_{t-2}, Q_{t-3})$	62.02	4.09	32.63	767.46	0.91	6
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	3
5	$Q_t = f(P_t, E_t)$	92.17	4.99	40.15	441.63	0.87	8
6	$Q_t = f(P_t, P_{t-1}, E_t, E_{t-1})$	69.27	3.72	25.13	375.23	0.90	5
7	$Q_t = f(P_t, P_{t-1}, P_{t-2}, E_t, E_{t-1}, E_{t-2})$	16.53	2.29	23.71	228.74	0.98	1
8	$Q_t = f(P_t, E_t, Q_{t-1})$	83.79	4.30	24.78	290.98	0.88	7
9	$Q_t = f(P_t, P_{t-1}, E_t, E_{t-1}, Q_{t-1})$	76.82	3.70	19.88	341.27	0.89	4
10	$Q_t = f(P_t, P_{t-1}, P_{t-2}, E_t, E_{t-1}, E_{t-2}, Q_{t-1}, Q_t)$	27.28	2.71	22.10	284.63	0.96	2

Donde 1 representa el mejor modelo y 10 el menos eficiente.

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 9. Estadísticas de validación

Nº	Modelos	Estadísticas					Calificación
		EQM	EAM	EPM	EPmax	R ²	
1	$Q_t = f(Q_{t-1})$	442.55	10.57	78.58	1390.19	0.74	10
2	$Q_t = f(Q_{t-1}, Q_{t-2})$	346.38	9.37	69.69	1180.64	0.78	7
3	$Q_t = f(Q_{t-1}, Q_{t-2}, Q_{t-3})$	386.79	9.37	60.20	1220.15	0.72	8
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	9
5	$Q_t = f(P_t, E_t)$	263.19	8.59	81.51	1404.34	0.82	5
6	$Q_t = f(P_t, P_{t-1}, E_t, E_{t-1})$	309.28	8.54	77.91	1987.55	0.79	6
7	$Q_t = f(P_t, P_{t-1}, P_{t-2}, E_t, E_{t-1}, E_{t-2})$	237.49	7.98	65.59	1068.29	0.85	4
8	$Q_t = f(P_t, E_t, Q_{t-1})$	256.31	6.62	39.24	664.39	0.83	2
9	$Q_t = f(P_t, P_{t-1}, E_t, E_{t-1}, Q_{t-1})$	245.51	6.80	38.50	524.03	0.83	1
10	$Q_t = f(P_t, P_{t-1}, P_{t-2}, E_t, E_{t-1}, E_{t-2}, Q_{t-1}, Q_t)$	297.12	7.15	54.21	914.08	0.84	3

Donde 1 representa el mejor modelo y 10 el menos eficiente.

En la fase de validación el modelo 9 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 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 8 y 9), los modelos seleccionados son:

- El modelo 3 es el que presenta mejor desempeño dentro de los 4 modelos que consideran como

variables de entrada únicamente caudales medios mensuales. (ver Figura 8).
 - El modelo 7 es el que presenta mejor desempeño dentro de los 3 modelos que consideran como variables de entrada valores de precipitación y evaporación mensual (ver Figura 9).

- El modelo 10 es 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 10)

Figura 8. Predicción de caudales mensuales con el modelo 3.

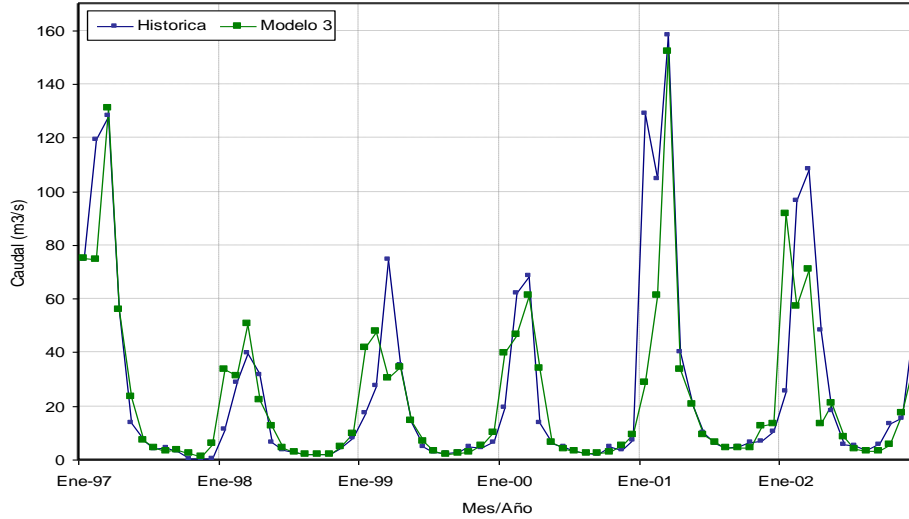


Figura 9. Predicción de caudales mensuales con el modelo 7.

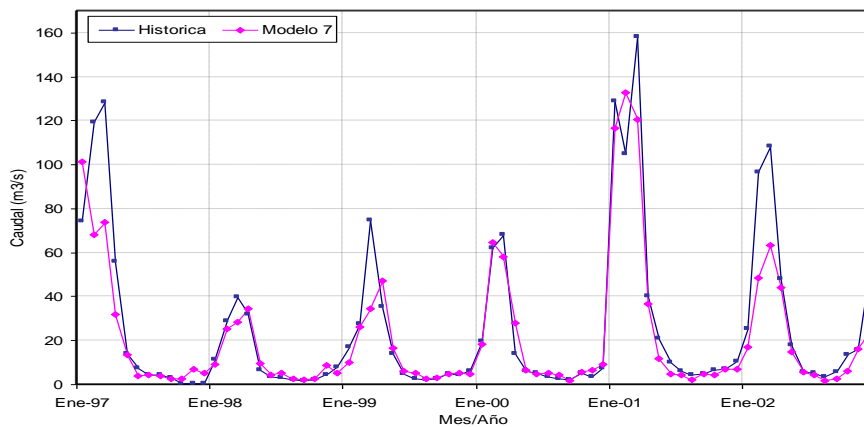
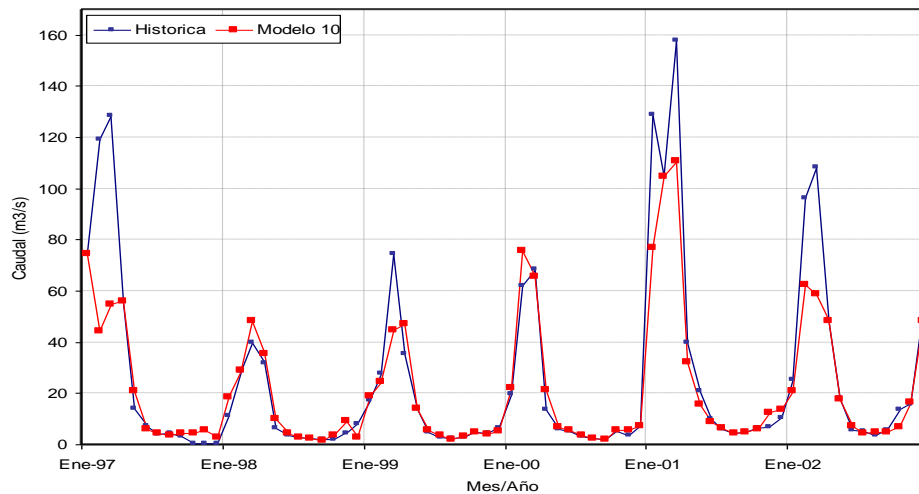


Figura 10. Predicción de caudales mensuales con el modelo 10



3.4. Comparación de modelos

En la comparación de los modelos de predicción, se evaluaron los resultados en términos estadísticos y análisis gráfico de los 3 mejores modelos de redes

neuronales artificiales con los resultados del modelo de serie temporal estocástica PAR (1). (ver Tabla 10 y Figura 11).

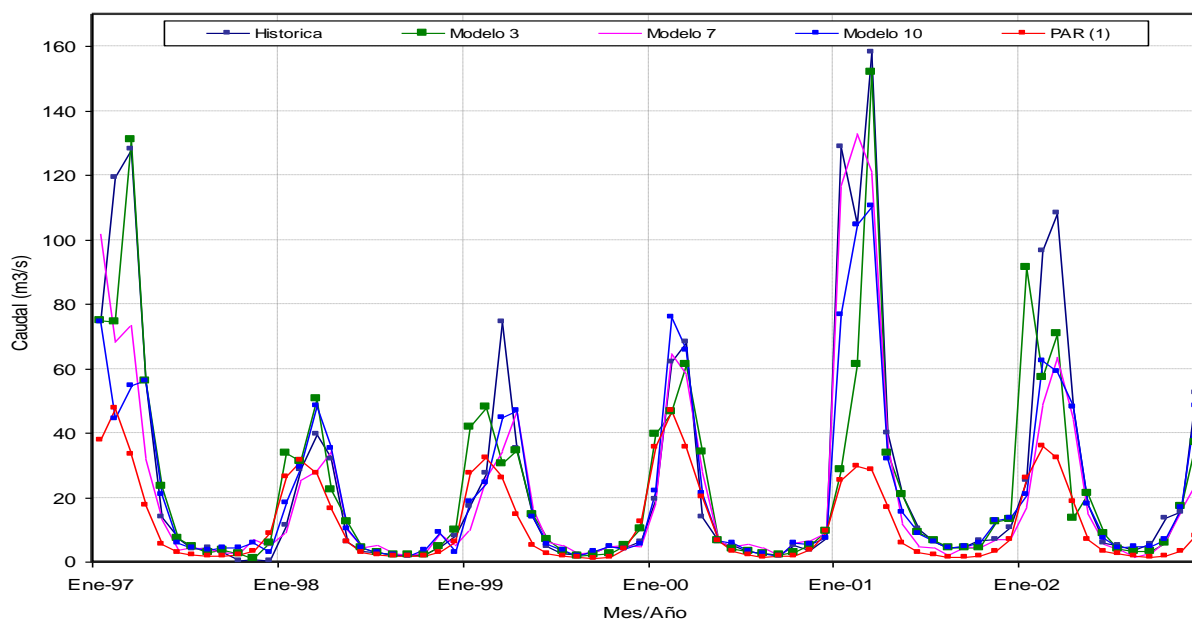
Tabla 10. Resumen de resultados estadísticos.

Modelo	Estadísticas				
	EQM	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
RNA Modelo 7	237.49	7.98	65.59	1,068.29	0.85
RNA Modelo 10	297.12	7.15	54.21	914.08	0.84

Independientemente del conjunto de entradas utilizado en el entrenamiento y validación de los modelos de redes neuronales artificiales (RNA), estadísticamente los modelos de redes neuronales

artificiales presentan los mejores resultados. Los coeficientes de determinación fueron mayores y los errores menores, lo que hace que los valores calculados estén más próximos a los reales.

Figura 11. Comparación de modelos.



De todo esto puede decirse que se confirma de cierta forma la necesidad de utilizar modelos de redes neuronales artificiales en el modelamiento y predicción de caudales medios mensuales, ya que se consiguió una mayor efectividad en los resultados de predicción con porcentajes muy cercanos a los reales, lo que confirma la robustez y grandes ventajas de la modelación inteligente, en comparación a un modelo de serie temporal estocástica PAR (1).

4. Conclusión

✓ Como es observable los modelos de redes neuronales artificiales desde la predicción de caudales mensuales, muestra un excelente comportamiento en términos de las estadísticas de los errores de predicción; más aún cuando de manera comparativa con un modelo autorregresivo periódico de primer orden – PAR (1), muestra grados superiores de eficiencia en las predicciones, ya que presentan porcentajes de ajuste superiores al 72%, en comparación al 59% que presenta el modelo de serie de tiempo estocástica. Estos resultados son altamente

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

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

✓ Independientemente del conjunto de entradas utilizado, todos los modelos de redes neuronales mostraron un mejor desempeño en comparación al modelo PAR (1), sin embargo el modelo 3 ($Q_t = f(Q_{t-2}, Q_{t-3})$), es el que presenta mejor resultado dentro de los modelos que consideran como variables de entrada únicamente caudales medios mensuales. El modelo 7 ($Q_t = f(P_t, P_{t-1}, P_{t-2}, E_t, E_{t-1}, E_{t-2})$), presenta mejor desempeño dentro de los modelos que consideran como variables de entrada valores de precipitación y evaporación mensual. El modelo 10

($Q_t = f(P_t, P_{t-1}, P_{t-2}, E_t, E_{t-1}, E_{t-2}, Q_{t-1}, Q_{t-2})$), es el que responde mejor dentro de los 3 modelos que consideran como variables de entrada valores de precipitación, evaporación, y caudales medios mensuales.

5. Referencias bibliográficas

- Alegre, A. 2003. Simulación de redes neuronales artificiales: Una aplicación didáctica. Tesis Lic. Sistemas. Universidad Nacional del Nordeste. Corrientes, Argentina. 188 p.
- Arias, P. 2006. Modelamiento estocástico de las descargas medias con fines de riego en la cuenca del río Huancané. Tesis Ing. Agrícola. Universidad Nacional del Altiplano. Puno, Perú. 122 p.
- Baganha, A. 1999. Modelagem chuva – vazão em bacias hidrográficas com suporte em redes neurais artificiais. Tesis Dr. Universidade Estadual de Campinas. Brasil. 279p.
- 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.
- 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.
- Ehlers, R. 2007. Análise de séries temporais. Curitiba, Laboratório de Estatística e Geoinformação. Universidade Federal do Paraná. Brasil. 107 p.
- Gibilini, D. 2006. Aplicação de técnicas de inteligencia Artificial na alocação dinâmica de canais em redes sem fio, Tesis Mg. Sc. Universidade de São Paulo. Brasil. 166 p.
- 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. Bariloche, Argentina.
- Kiliç, I; Cigizoglu, K. 2003. Reservoir management using artificial neural networks. *State Hydraulic Works*. 9 p. Istanbul, Turquía.
- Lavado, W. 2002. Generación de descargas medias mensuales mediante el uso de modelos autorregresivos: Aplicación al río Cunas. Tesis Ing. Agrícola. Universidad Nacional Agraria La Molina. Lima, Perú. 165 p.
- Magalhães, M. 2004. Redes neurais, metodologias de agrupamento y coninação de previsores aplicados à previsão de vazões naturais. Tesis Mg. Sc. Universidade Estadual de Campinas. Brasil. 123 p.
- Ovando, G; Bocco, M; Sayazo, S. 2005. Redes neuronales para modelar predicción de heladas. *Agricultura Técnica*. 65(1):65-73. Córdoba, Argentina.
- 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. México.
- 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. Guayaquil, Ecuador.