

Artículo publicado en el Repositorio Institucional del IMTA

<i>Título</i>	Tránsito de avenidas en cauces mediante redes neuronales artificiales.
<i>Autor / Adscripción</i>	Juan Pablo Molina Aguilar Universidad Nacional Autónoma de México  Javier Aparicio Instituto Mexicano de Tecnología del Agua
<i>Publicación</i>	Ingeniería Hidráulica en México, 21(4): 65-86
<i>Fecha de publicación</i>	2006
<i>Resumen</i>	Los métodos hidrológicos comúnmente utilizados para el tránsito de avenidas presentan limitantes en el análisis de problemas complejos. Las características de las redes neuronales artificiales las convierten en una posibilidad para su aplicación al tránsito de avenidas. La aplicación de redes neuronales artificiales a distintos casos muestra que, contando con información y seleccionando la topología adecuada, es posible obtener resultados con una precisión similar a los métodos hidráulicos e hidrológicos con datos habitualmente disponibles en registros hidrométricos, que son escasos para la aplicación de dichos métodos. Su aplicación se realiza en la región hidrológica 30 y en la confluencia de los ríos Manso y Cajones, que forman el río Tesechoacán.
<i>Identificador</i>	<a href="http://hdl.handle.net/123456789/809">http://hdl.handle.net/123456789/809</a>

# Tránsito de avenidas en cauces mediante redes neuronales artificiales

Juan Pablo Molina-Aguilar

Universidad Nacional Autónoma de México

Javier Aparicio

Instituto Mexicano de Tecnología del Agua

*Los métodos hidrológicos comúnmente utilizados para el tránsito de avenidas presentan limitantes en el análisis de problemas complejos como, por ejemplo, los casos de trenes de avenidas, la presencia de flujos laterales o las uniones de ríos, principalmente en los casos en que no se cuenta con medición hidrométrica de toda la red hidrológica. Las características de las redes neuronales artificiales las convierten en una posibilidad para su aplicación al tránsito de avenidas, pues presentan diversas ventajas respecto a los métodos hidrológicos tradicionales. La aplicación de redes neuronales artificiales a distintos casos muestra que, contando con información y seleccionando la topología adecuada, es posible obtener resultados con una precisión similar a los métodos hidráulicos e hidrológicos con datos habitualmente disponibles en registros hidrométricos, que son escasos para la aplicación de dichos métodos. La aplicación de redes neuronales artificiales de arquitecturas simples en el caso de avenidas reales aisladas y trenes de avenidas en la región hidrológica 30, así como en registros anuales en la confluencia de los ríos Manso y Cajones, que forman el río Tesechoacán, muestra claramente las bondades de su utilización.*

**Palabras clave:** red neuronal, artificial, tránsito, avenida, flujo lateral, Muskingum.

## Introducción

Las redes neuronales artificiales (RNA) se desarrollaron para simular matemáticamente los sistemas naturales usando estructuras análogas a las redes neuronales biológicas (RNB), como las del cerebro humano. Al comienzo de la década de los cuarenta del siglo anterior, la comunidad científica, en estudios sobre el funcionamiento del cerebro humano, desarrolló la inteligencia artificial (IA) y con ello dio origen a las RNA. El primer trabajo sobre RNA surgió en la década de los cuarenta (McCulloch y Pitts, 1943), en el que se establece una analogía entre el proceso de comunicación de las células nerviosas y el correspondiente a la transmisión eléctrica. En 1949, Hebb sugirió que la conectividad del cerebro cambiaba continuamente cuando un organismo aprende a diferenciar diversas funciones, de tal forma que las conexiones neuronales son generadas por ese cambio. En 1956, Rochester *et al.* desarrollaron un modelo de RNA que simulaba la interconexión de centenas de neuronas y un sis-

tema para verificar el comportamiento de la red durante dos estímulos externos. Rosenblat (1958) creó la red *perceptrón*, que se sustenta en el modelo de McCulloch y Pitts, con la cual clasificó patrones de dos categorías gracias a sinapsis con capacidad de ajustarse. Un par de años después, Widrow y Hoff (1960) desarrollaron la primera red capaz de imitar al cerebro humano utilizando procesadores paralelos, con una estructura de red Adaline (*Adaptive Linear Element*), y más tarde estructuraron una nueva red denominada Madaline (Many Adaline). Algunas limitantes de las redes perceptrón fueron eliminadas por Rumelhart *et al.* (1986), con la creación de la regla delta generalizada para redes perceptrón multicapa, además de que las funciones de activación fueron sustituidas por funciones continuas sigmoidales. Aunque las RNA se han desarrollado en muchas esferas científicas y tecnológicas, incluyendo la hidrología, existen diversas áreas en las cuales su aplicación no está completamente estudiada, ya que, por ejemplo, es difícil saber cuánto tiempo necesita una red para aprender

cierta tarea, cuántas neuronas se requiere como mínimo para realizarla, etcétera. Los trabajos utilizando RNA para resolver problemas en ingeniería civil iniciaron después de los años ochenta, a raíz de los trabajos de Hopfield (Hopfield, 1982).

En materia hidrológica, French *et al.* (1992) utilizaron una RNA de tres capas con algoritmo de retropropagación para predecir el campo de intensidades de lluvia en el espacio y el tiempo; una vez entrenada, pronosticó campos de intensidades de lluvia en intervalos de una hora. Posteriormente, Crespo y Mora (1993) calcularon escurrimientos a partir de precipitaciones, entrenando una red multicapa capaz de calcular gastos pequeños, lo que facilitaba la realización de estudios para periodos secos. Flood y Kartman (1994) describieron conceptos y métodos para la utilización de RNA en la predicción de escurrimientos en ríos, aplicando series temporales con resultados aceptables. En el mismo año, Karunanithi *et al.* (1994) utilizaron un algoritmo de correlación en cascada capaz de sintetizar la arquitectura de la red en el proceso de entrenamiento, con el cual validaron la capacidad de predicción de diferentes modelos de RNA antes de seleccionar la arquitectura definitiva, disminuyendo así el tiempo de entrenamiento. Luego, Hsu *et al.* (1995) utilizaron las RNA para simular procesos hidrológicos no lineales en una cuenca, mediante la implementación de un algoritmo híbrido de entrenamiento para la red de tres capas, que consistió en combinar el algoritmo de mínimos cuadrados con el algoritmo "simplex" de optimización, con lo que se logró una buena aproximación al mínimo global, realizando pocas iteraciones. También se han desarrollado varios modelos para la predicción lluvia-escurrimiento de diferentes grados de complejidad. Shamseldin *et al.* (1996) desarrollaron un sistema general para estimar descargas obtenidas de diferentes modelos para ser usado en sistemas operacionales de predicción de escurrimientos en ríos, el cual básicamente consistió en la combinación de resultados de modelos sencillos que captan una característica particular del proceso físico, y utilizarlos como datos para obtener una combinación global de salida usando una RNA, con lo que se logra una mejor estimación del escurrimiento. Sarmiento (1996) usó las RNA en la modelación de variables hidrológicas y los resultados obtenidos fueron comparados con los generados a partir de la calibración del modelo MODHAC (Lanna y Schwarzbach, 1988), tras lo cual concluyó que las RNA deben ser usadas en el modelado de fenómenos naturales en los casos en que no se dispone de formulación matemática explícita. Estudios relativos a la predicción de escurrimientos en tiempo real fueron hechos por Ballini (1996), usando una red *Kohonen* y una red multicapa con algo-

ritmo retropropagación, con lo que mostró que las redes multicapa presentan capacidad apropiada para la predicción lluvia-escurrimiento, después de haber realizado una comparación de resultados de ambas redes con respecto del método de los vecinos más próximos.

En 1997, Valencia llevó a cabo un análisis de aplicaciones de redes neuronales del tipo perceptrón multicapa de simulación de transformación de lluvia en escurrimiento y la predicción de escurrimientos medios mensuales, incorporando la representatividad de cada estación pluviométrica, en lugar de trabajar con lluvias medias. De igual forma, Shamseldin (1997) utilizó una RNA para el modelado de la transformación lluvia-escurrimiento y después evaluó su desempeño respecto de algunos modelos tradicionales. Los resultados con las RNA fueron prometedores y en la mayoría de los casos se obtuvieron correlaciones más altas que con modelos que utilizan información de entrada similar, como es el caso de los modelos lineal, perturbación lineal con estacionalidad y perturbación lineal de vecinos más próximos. Al mismo tiempo, Ballini *et al.* (1997) utilizaron una RNA para estimar escurrimientos mensuales, compararon sus resultados con los de los métodos de Box-Jenkins y resaltaron la importancia de la estandarización de las series temporales, pues ésta contribuyó a un rápido aprendizaje y una adecuada capacidad de generalización y pronóstico de la RNA, con lo que lograron resultados comparables a los obtenidos por los métodos Box-Jenkins, pero en forma más sencilla. La mayor parte del tiempo que se invierte en el montaje de una RNA es en el proceso de identificación de su mejor estructura y entrenamiento. Para determinar una buena estructura y una forma de disminuir ese tiempo, Kadowaki *et al.* (1997) propusieron una metodología para la determinación del número de entradas de las RNA utilizadas para la predicción de escurrimientos medios mensuales y observaron que es más fácil escoger un modelo de red neuronal que construir un algoritmo del tipo Box-Jenkins. Thirumalaiah y Deo (1998) utilizaron una RNA para analizar y prever eventos futuros con base en registros de nivel en ríos. En 2003, Cigizoglu llevó a cabo una investigación en varios ríos, estimando, pronosticando y extrapolando escurrimientos mediante el uso de una red perceptrón multicapa, y concluyó que la solución dada por la RNA provee mejores resultados que los métodos estadísticos y estocásticos convencionales, como el modelo de regresión múltiple y los modelos AR (1), AR (3), AR (6) y AR (6) - log. Campolo *et al.* (2003) llevaron a cabo un estudio en el río Arno, Italia, en el que trataron de predecir la evolución del nivel de agua en tiempo real mediante un modelo basado en una RNA y observaron un buen funcionamiento de la RNA en la predicción de la evolu-

ción del nivel del río, para lo cual tomaron información en tiempo real de lluvia, hidrometría y política de operación de presas de dos años de registro, llevaron a cabo el entrenamiento y posterior validación del rango seleccionado y observaron que el porcentaje máximo de error en la representación de dicho rango era del 7% y que al llevar a cabo la predicción de otros años los resultados eran similares. Las RNA constituyen una forma rápida y flexible para la creación de modelos para la predicción de escurrimientos en ríos y en muchos casos tienen un mejor desempeño que los métodos estadísticos y estocásticos tradicionales mencionados previamente. Por otra parte, los métodos comunes para el tránsito de avenidas en cauces, como el de Muskingum (Aparicio, 2001) y Muskingum-Cunge (US Army Corps of Engineers, 1994) tienen limitaciones en casos relativamente complejos, como cuando se tienen entradas laterales o en confluencias de cauces.

### Neurona artificial

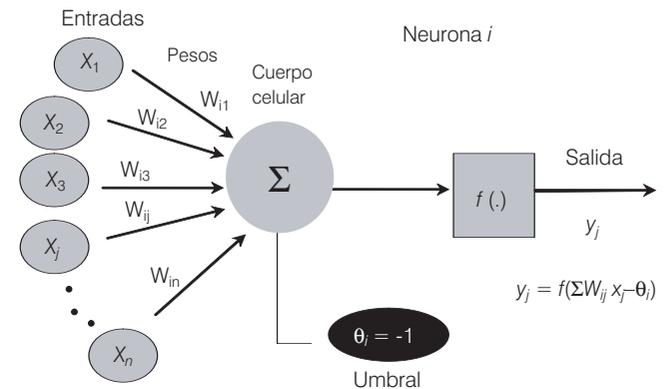
La unidad básica de una RNA es la neurona artificial. La neurona más común es la del tipo McCulloch-Pitts (ilustración 1). Una neurona artificial es un procesador elemental, en el sentido de que procesa un vector  $x$  ( $x_1, x_2, \dots, x_n$ ) de entradas y produce un respuesta o salida única. Los elementos clave de una neurona artificial son las entradas ( $x_i$ ), que reciben los datos de otras neuronas; los pesos sinápticos ( $w_{ij}$ ), que establecen la sinapsis (unión) entre las entradas que vienen de otras neuronas que se modifican junto con el valor del umbral durante el entrenamiento de la red, y la regla de propagación para obtener el valor del potencial postsináptico, que es el valor de salida de la red que pasa a las siguientes neuronas ocultas o de salida de la red, según la capa de que se trate, y que finalmente se toma en cuenta para obtener el valor de respuesta de la función de transferencia.

### Función de transferencia

La neurona recoge las señales por su sinapsis, sumando todas las influencias excitadoras e inhibitoras. Existen cuatro funciones de transferencia típicas que determinan distintos tipos de neuronas: escalón, lineal, sigmoideal y Gaussiana (cuadro 1).

La neurona se activa si la fuerza combinada de la señal de entrada  $\sum X_i W_i$  es superior a un cierto umbral  $\theta_i$ . En una RNA, si la última capa de la red multicapa tiene neuronas con función de transferencia sigmoideal, entonces las salidas de la red estarán limitadas a un intervalo muy pequeño. Si se utilizan neuronas con función de transferencia lineal, las salidas de la red podrán tomar cualquier

**Ilustración 1. Representación de una neurona artificial tipo McCulloch-Pitts. (Fuente: Larrañaga e Iñaki, 1997).**



**Cuadro 1. Funciones de transferencia y rango de aplicación. (Fuente: Larrañaga e Iñaki, 1997).**

Nombre	Función	Rango
Escalón	$y = \text{sign}(x)$ $y = H(x)$	$[-1, +1]$ $[0, +1]$
Lineal	$y = x$	$[-\infty, +\infty]$
Sigmoideal	$y = \frac{1}{1+e^{-x}}$ $y = \text{tgh}(x)$	$[0, +1]$ $[-1, +1]$
Gaussiana	$y = Ae^{-Bx}$	$[0, +1]$

valor, dado el intervalo de la función, por lo que en el caso de avenidas puede representar un intervalo muy amplio de valores. En la retropropagación del error (ajuste de los pesos y umbrales, proporcional al error entre la salida y el objetivo), es importante calcular las derivadas de la función de transferencia utilizada, por lo que comúnmente se usa la función de transferencia sigmoideal, dado que su derivada puede calcularse para cualquier valor del dominio. La función de transferencia sigmoideal ha sido utilizada en RNA con aplicación hidrológica en estudios de escurrimientos directos en cuencas (Zhang y Govindaraju, 2003), transporte de sedimentos (Tayfur, 2002), escurrimientos en ríos (Imrie *et al.*, 2000), entradas y operación en vasos de almacenamiento (Jain *et al.*, 1999), por mencionar sólo algunas.

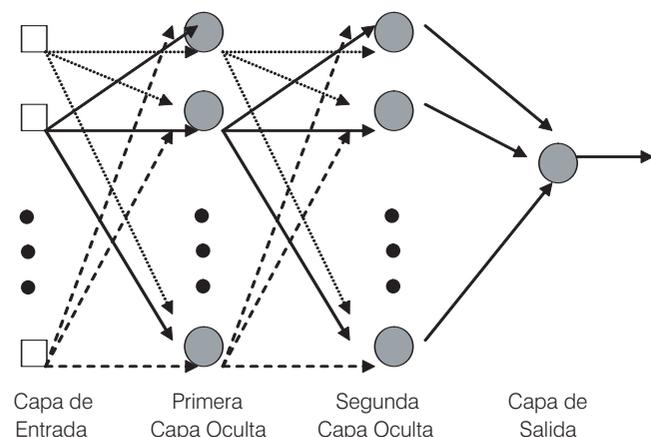
### Redes neuronales

Las RNA están compuestas de un cierto número de elementos procesadores interconectados entre sí (neuronas) trabajando en paralelo para resolver un problema

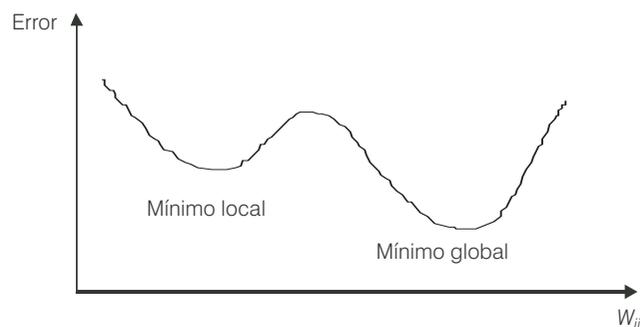
específico. Para tal efecto, las RNA deben entrenarse, es decir, se seleccionan dos vectores: uno de entrada, correspondiente a valores de las variables que generan el fenómeno físico, y otro de salida, correspondiente a la magnitud del mismo, con los cuales la red calculará, mediante un algoritmo, el valor de todas las conexiones y umbrales, de tal forma que la RNA determine la relación existente entre ambos vectores seleccionados. El perceptrón es la red neuronal pionera y la forma más simple de las RNA. El algoritmo de entrenamiento del perceptrón se basa en el Teorema de Convergencia Perceptrón (Rosenblatt, 1958). La red Adaline está limitada a una única neurona de salida y utiliza la denominada regla Delta de Widrow-Hoff (Widrow y Hoff, 1960), basada en la búsqueda del error mínimo de una expresión entre la salida deseada y la salida lineal obtenida. En 1986, Rumelhart *et al.* desarrollaron el algoritmo de entrenamiento llamado "retropropagación", con lo que demostraron que era posible entrenar con eficiencia redes con capas intermedias. Esto dio como resultado el modelo de RNA más utilizado actualmente: las redes perceptrón multicapas (MLP). Una MLP es una extensión del perceptrón con algunas particularidades. Las redes multicapa unidireccionales poseen un mejor desempeño debido al paso unidireccional de la señal por la red, de acuerdo con estudios realizados por la ASCE (Govindaraju, 2000) a diferentes tipos de RNA con aplicaciones hidrológicas.

En la ilustración 2 se puede apreciar una red perceptrón multicapa con capa de entrada (variables que dan origen al fenómeno físico), de salida (resultado del fenómeno físico) y dos capas donde se lleva a cabo la sinapsis de la red, denominadas ocultas, dado que no tienen una conexión directa con el entorno del fenómeno físico. Estas redes pueden ser utilizadas para clasificación de

**Ilustración 2. Arquitectura de un MLP con dos capas ocultas (Fuente: Matlab 6.5, 2002b).**



**Ilustración 3. Representación de un mínimo local y el mínimo global. (Fuente: Larrañaga e Iñaki, 1997).**



patrones no linealmente separables, es decir, patrones cuya relación entre variables dentro del fenómeno físico no puede clasificarse mediante una línea recta en un hiperplano debido a su compleja dependencia ni realizar gran variedad de mapeos complejos. La red posee buena capacidad de generalización: es capaz de clasificar correctamente un patrón complejo, aun cuando éste no pertenezca al conjunto de entrenamiento de la red (Thirumalaiah y Deo, 1998). El algoritmo de retropropagación utilizado para el entrenamiento de esta arquitectura opera en una secuencia de dos pasos:

*Primero.* Un patrón se representa por la capa de entrada de la red. La actividad resultante fluye a través de la red, capa por capa, hasta que la respuesta se produce por la capa de salida.

*Segundo.* La salida obtenida se compara con la salida deseada para ese patrón en particular y se calcula el error. El error es propagado desde la capa de salida hasta la capa de entrada, y los pesos de las conexiones de las unidades de las capas internas se van modificando conforme al error retropropagado.

En el entrenamiento de una RNA compleja, la solución obtenida puede no corresponder al mínimo global, porque una función puede presentar mínimos locales (ilustración 3). Cuando el tipo de superficie de error no es conocido *a priori*, el algoritmo verifica una gran cantidad de posibilidades antes de determinar la mejor solución; por tanto, para garantizar que se alcanza el mínimo global con el algoritmo de entrenamiento seleccionado, se debe utilizar un valor pequeño de la tasa de aprendizaje, pues esto asegura que el algoritmo analiza toda la superficie de error.

La tasa de aprendizaje (incremento del paso en el algoritmo de entrenamiento) es, pues, un elemento importante en este proceso, ya que controla el tiempo de aprendizaje, tomando en cuenta la tasa de convergencia entre la solución actual y el mínimo global.

El entrenamiento de una red neuronal se lleva a cabo mediante el ajuste de los pesos que conectan sus neuronas, lo que se logra mediante el análisis del comportamiento de la red a varios ejemplos de entrenamiento, cada uno de los cuales se compone de una muestra de datos de entrada específicos y la correspondiente respuesta de salida. En los pesos es donde reside el conocimiento de una red. Entre los diferentes tipos de aprendizaje podemos mencionar los siguientes:

*Aprendizaje supervisado.* En este tipo de aprendizaje se le proporciona a la RNA una serie de ejemplos consistentes en unos patrones de entrada, junto con la salida que debería dar la red.

*Aprendizaje no supervisado o autoorganizado.* En este tipo de aprendizaje se presenta a la red una serie de ejemplos, pero no se presenta la respuesta deseada.

*Aprendizaje híbrido.* Unas capas de la red tienen un aprendizaje supervisado y otras capas de la red tienen un aprendizaje de tipo no supervisado.

*Aprendizaje reforzado.* Es un aprendizaje con características del supervisado y con características del autoorganizado, diferenciándose del híbrido en que en éste sólo se proporciona un porcentaje de error que debe cumplirse al no indicarle ninguna salida deseada.

### Función de desempeño

Las funciones de desempeño permiten evaluar el comportamiento de una red. Esto es útil para muchos algoritmos de aprendizaje, como la retropropagación, la cual opera ajustando los pesos y umbrales de la red neuronal para mejorar el desempeño respecto de los datos dados a la red para el entrenamiento. En el cuadro 2 se presentan las funciones de desempeño más comunes. De acuerdo con Karunanithi *et al.* (1994), comparando los resultados de sus estudios, el error cuadrático medio (MSE) proporciona una buena medida de la bondad de ajuste para valores de escurrimientos altos, mientras que los errores relativos medios (RMSE) proporcionan una perspectiva más equilibrada de la bondad de ajuste en valores de escurrimiento moderados. Sin embargo, estas medidas están fuertemente afectadas por las características del cauce y se debe tener cuidado cuando se comparan estudios utilizando tales estadísticas. Por tanto, los pesos se ajustan durante el proceso de entrenamiento mediante el algoritmo de aprendizaje, como se explicará más adelante.

### Normalización de datos

Es conveniente que todos los valores de las variables consideradas en la modelación de un fenómeno hidrológico se normalicen, para asegurar que recibirán igual

**Cuadro 2. Funciones de desempeño (fuente: Dolling *et al.*, 2003).**

Funciones de desempeño		
MAE	Error Medio Absoluto	$\frac{\sum_i  y - \hat{y} }{N}$
MSE	Error Medio Cuadrático	$\frac{\sum_i (y - \hat{y})^2}{N}$
RMSE	Raíz del Error Medio Cuadrático	$\sqrt{\frac{\sum_i (y - \hat{y})^2}{N}}$
SSE	Suma del Error Cuadrático	$\sum_i (y - \hat{y})^2$

$y$  = valor objetivo.

$\hat{y}$  = valor de salida.

$N$  = número de neuronas en la capa de salida.

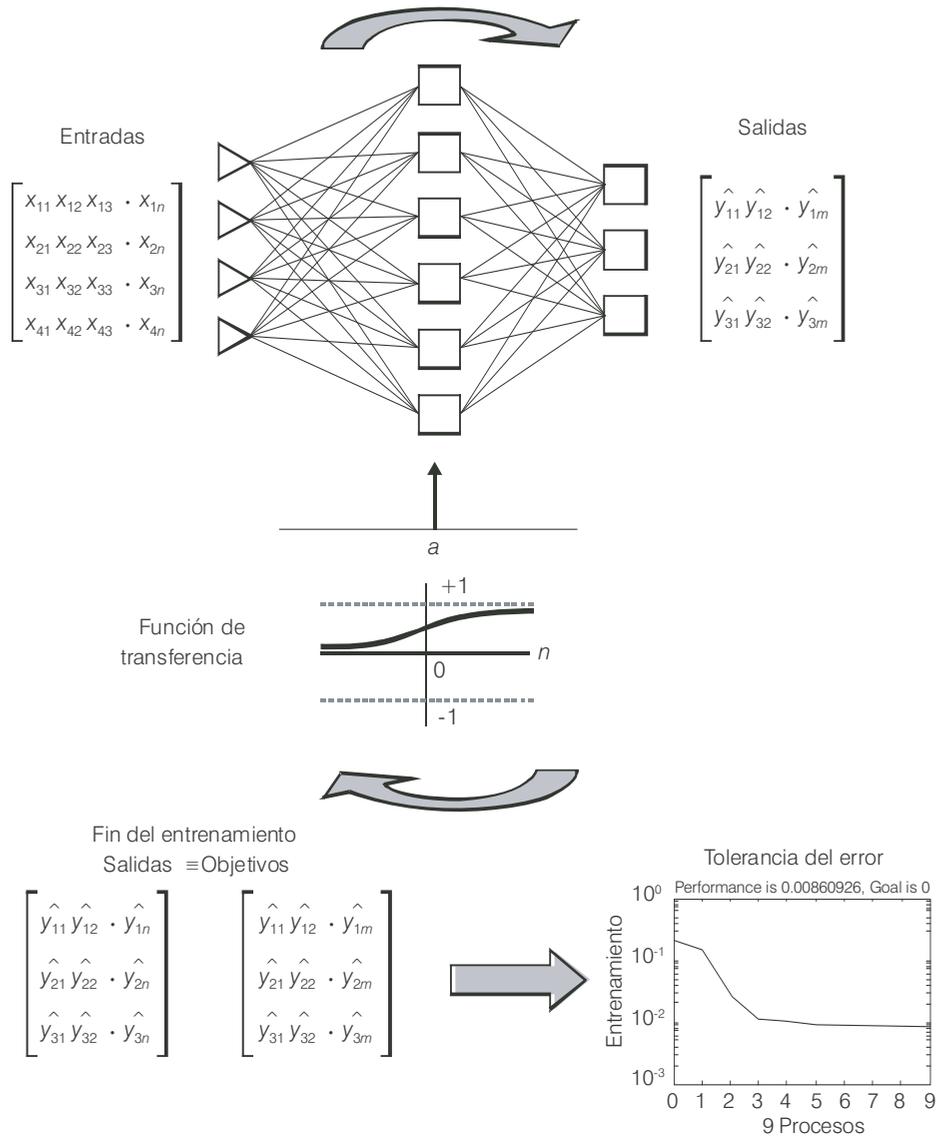
atención durante el proceso de entrenamiento, es decir, para evitar que valores extremos —cercaos a uno o cercaos a cero— dominen el resultado y con ello la red pierda su capacidad de generalizar posteriormente. Esto es particularmente importante en las redes MLP, ya que sin la normalización, los valores de las variables de entrada medidas en diferentes escalas dominarán el entrenamiento en una mayor o menor extensión, porque los pesos iniciales dentro de una red son aleatorios para el mismo intervalo finito. La ventaja de utilizar un intervalo de los valores normalizados entre 0.1 y 0.9 para modelar el escurrimiento es que pueden tomarse en cuenta los eventos para gastos extremos (altos y bajos) que ocurren fuera del intervalo de calibración.

### Desarrollo

La aplicación descrita en este artículo siguió la metodología que se describe a continuación.

En la ilustración 4 se muestran, en primer término, los ejemplos seleccionados para llevar a cabo el entrenamiento (entradas). Posteriormente se hace una primera propuesta de la topología de la red, en la cual se debe especificar claramente la función de transferencia en las capas ocultas, las cuales nos darán un primer resultado (salidas), que deberá compararse con el valor deseado en el entrenamiento (objetivos), con lo que se obtendrá un error. Si dicho valor está fuera de la tolerancia deseada, el algoritmo de entrenamiento procede a retropropagarlo de manera equitativa a través de todos los pesos y umbrales de la topología propuesta hasta llegar de nue-

Ilustración 4. Metodología ilustrativa de RNA.



va cuenta a los pesos y umbrales de la primera capa, para repetir el proceso cuantas veces sean necesarias hasta que las salidas sean idénticamente iguales a los objetivos y así llegar a la tolerancia del error especificada al inicio del entrenamiento.

Las redes neuronales usan un enfoque diferente para la solución de problemas que el utilizado por las computadoras convencionales. Las computadoras convencionales usan un enfoque algorítmico, es decir, la computadora sigue un grupo de instrucciones para resolver un problema. Si los pasos específicos que la computadora necesita seguir no son conocidos, no podrá resolver el problema. Esto restringe la capacidad de solucionar sólo los proble-

mas que únicamente nosotros entendemos y sabemos resolver. Las redes neuronales aprenden por medio de ejemplos que deben seleccionarse cuidadosamente; de lo contrario, se perderá tiempo útil o, peor aún, se puede propiciar que la red no funcione correctamente.

### Casos de estudio

#### Experimento numérico

Como primer caso de estudio se aplicó una RNA a un experimento numérico similar al "evento Wilson" discutido por O'Donnell (1985) y estudiado por Aldama (1990).

El experimento considera un canal trapecial con ancho de plantilla  $b=100$  m, talud  $k=2$ , longitud  $L=50,500$  m, pendiente de plantilla  $S_0=0.0001$  y rugosidad de Manning  $n=0.08$ . Para fines de calibración, se supondrá que el evento de avenida discutido por Aldama (1990) (cuadro 3) ocurre en dicho canal. Para poder evaluar la capacidad de la metodología propuesta en este primer caso, en comparación con otros procedimientos existentes, las ordenadas del hidrograma del citado evento se multiplican por factores 2, 5 y 10, y los hidrogramas resultantes se transitan por el canal empleando un método hidráulico, con lo que se obtienen los correspondientes hidrogramas de salida. Estos hidrogramas de salida se toman como base de comparación para los calculados con el método de advección-difusión (AD) (Aldama, 1990), el de Muskingum calibrado con las técnicas propuestas por Gill (1997) y O'Donnell (1985), y con los resultados de las redes neuronales.

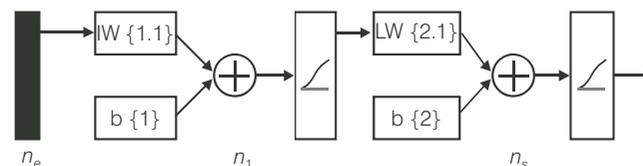
**Cuadro 3. Datos para el experimento numérico (fuente: Aldama, 1990; Aguilar, 1995).**

Número de intervalo	Tiempo (h)	Gasto de entrada I (m³/s)	Gasto de salida O (m³/s)
1	0	22,000	22,000
2	6	23,000	22,000
3	12	35,000	22,001
4	18	71,000	22,005
5	24	103,000	22,078
6	30	111,000	22,624
7	36	109,000	25,541
8	42	100,000	36,093
9	48	86,000	57,113
10	54	71,000	76,565
11	60	59,000	85,625
12	66	47,000	86,555
13	72	39,000	82,866
14	78	32,000	76,915
15	84	28,000	70,106
16	90	24,000	63,251
17	96	22,000	56,781
18	102	22,000	50,894
19	108	22,000	45,651
20	114	22,000	41,066
21	120	22,000	37,137
22	126	22,000	33,834
23	132	22,000	31,114
24	138	22,000	28,910
25	144	22,000	27,158
26	150	22,000	25,789
27	156	22,000	24,746
28	162	22,000	23,963
29	168	22,000	23,384
30	174	22,000	22,964
31	180	22,000	22,662

Una vez hecha la normalización de la forma  $Q_i/Q_{m\acute{a}x}$  del hidrograma de entrada y el de salida del cuadro 3, se procedió a realizar diversas arquitecturas de RNA unidireccionales con retropropagación, con un intervalo de entradas de 0 a 1. El algoritmo de aprendizaje utilizado fue el de Levenberg-Marquardt (*Matlab*, 2002a) con una función de aprendizaje de gradiente descendiente con *momentum* (*Matlab*, 2002b). Para llevar a cabo la comparación del desempeño con respecto a otras estructuras se utilizó el error medio cuadrático (MSE), y la transferencia de la información en las neuronas ocultas se llevó a cabo con la función sigmoideal (*Matlab*, 2002b).

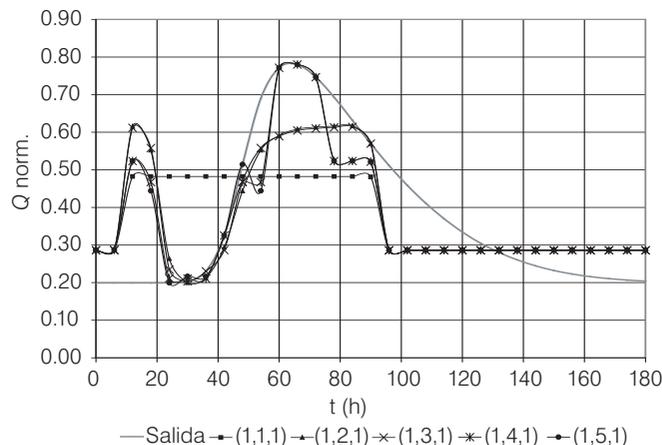
Realizado el entrenamiento de la red con arquitectura  $n_e - n_1 - n_s$ , variando el número de neuronas en la capa oculta ( $n_1$ ) como se observa en la ilustración 5, los resultados para una entrada y salida se pueden apreciar en las ilustraciones 6, 7 y 8, donde se graficaron gastos normalizados ( $Q_{norm}$ ). En este primer desarrollo de aprendizaje de la red no se logra un buen entrenamiento con una sola entrada y salida. Se hicieron otros desarrollos modificando la arquitectura de la red. Por ejemplo, con

**Ilustración 5. RNA de tres capas.**

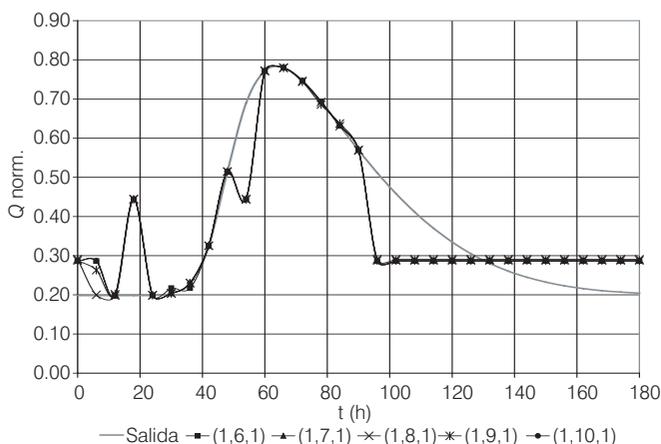


$n_e$  = número de entradas de la RNA.  
 $n_1$  = número de neuronas de la primera capa oculta.  
 $n_s$  = número de neuronas de la capa de salida.

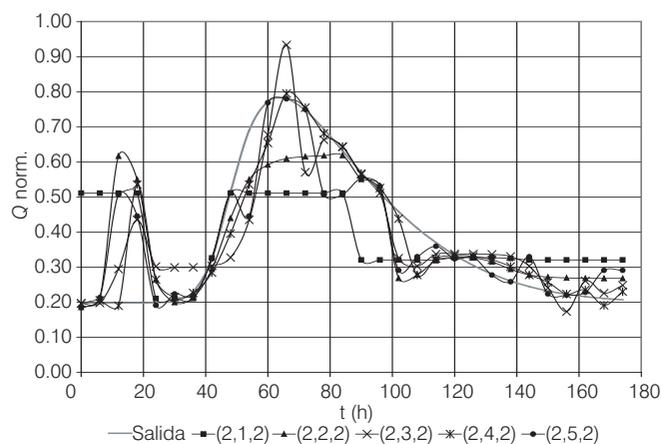
**Ilustración 6. Entrenamiento de la red con estructura 1 - n - 1; (1 < n < 5).**



**Ilustración 7. Entrenamiento de la red con estructura 1 - n - 1; (6 < n < 10).**



**Ilustración 8. Entrenamiento de la red con estructura 2 - n - 2; (1 < n < 5).**



dos neuronas tanto en el hidrograma de entrada como en el de salida se observó la existencia de zonas donde el aprendizaje no es aún satisfactorio, como se muestra en la ilustración 9.

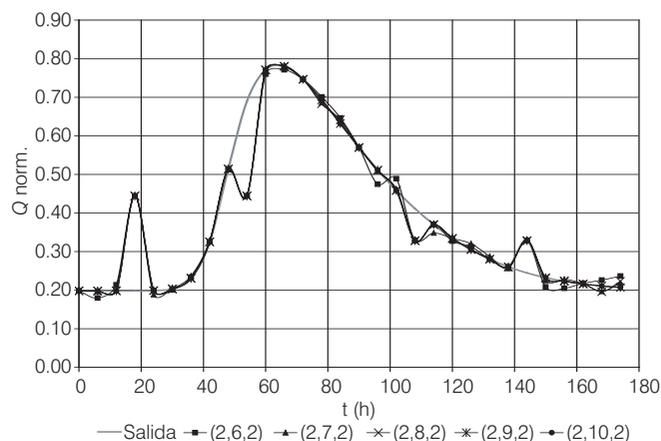
En las ilustraciones 10 y 11 se presenta el caso de tres neuronas para los hidrogramas de entrada y salida. Se aprecia que se tiene un entrenamiento prácticamente perfecto a partir de una RNA 3 - n - 3,  $n \geq 6$ . En la ilustración 12 se aprecian los resultados del tránsito obtenidos por los diferentes métodos. Se observa que la RNA se ajusta perfectamente al hidrograma de salida, puesto que eso se buscó en el entrenamiento. Además se observa que el resto de los métodos tienen problemas al inicio de la rama ascendente, ya que presentan valores menores que el gasto base (hasta en 18 m<sup>3</sup>/s). Fuera de

dicha zona, el método AD es el que más se aproxima al hidrograma de salida, mientras que en los de Gill y O'Donnell el gasto pico es menor en 5.013 y 16.784 m<sup>3</sup>/s, respectivamente

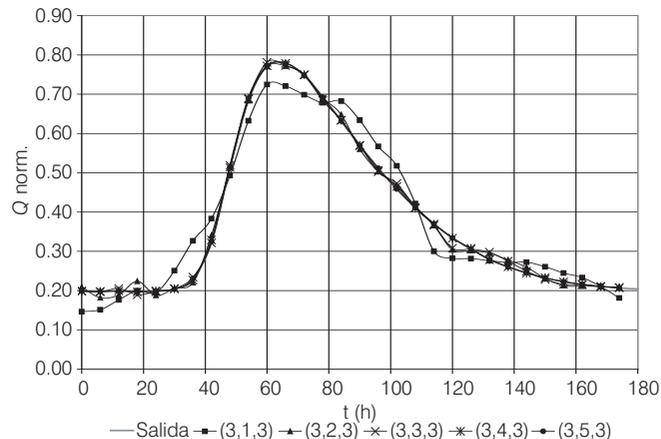
Una vez entrenada la red, las ordenadas del hidrograma de entrada del cuadro 3 se multiplicaron por los factores 2, 5 y 10. En la ilustración 13, correspondiente a un factor de 2, se aprecia cómo el método AD disminuye el problema de ordenadas menores al gasto base que presentaba en la rama ascendente, pero los de Gill y O'Donnell continúan con la misma tendencia. Se aprecia que la RNA se aproxima mejor al hidrograma de salida que los métodos de Gill y O'Donnell.

Al quintuplicar y decuplicar los hidrogramas del experimento numérico se obtienen resultados similares (ilus-

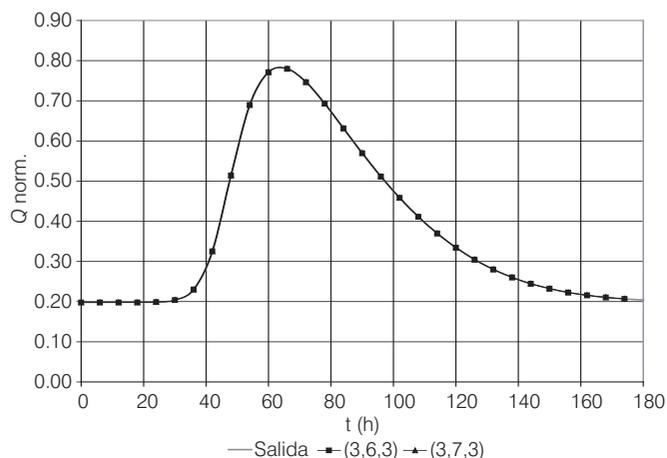
**Ilustración 9. Entrenamiento de la red con estructura 2 - n - 2; (6 < n < 10).**



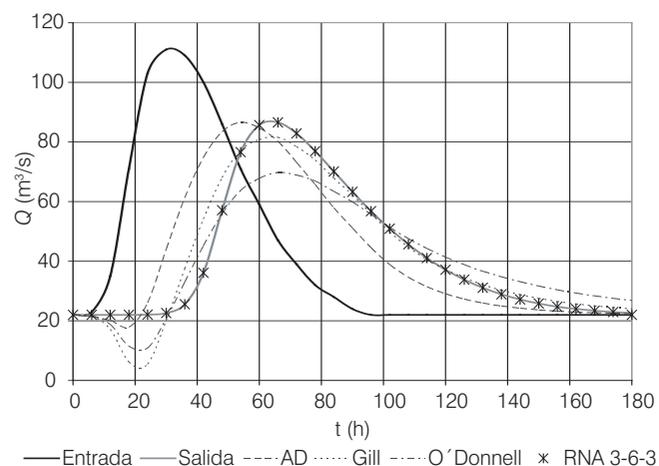
**Ilustración 10. Entrenamiento de la red con estructura 3 - n - 3. (1 < n < 5).**



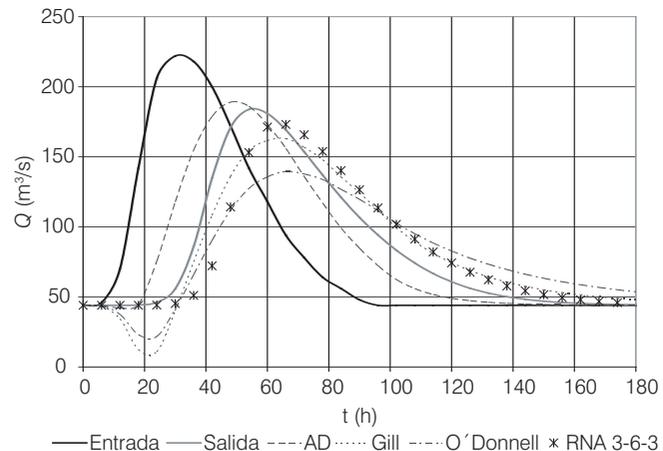
**Ilustración 11. Entrenamiento de la red con estructura 3 – n – 3; (6 < n < 7).**



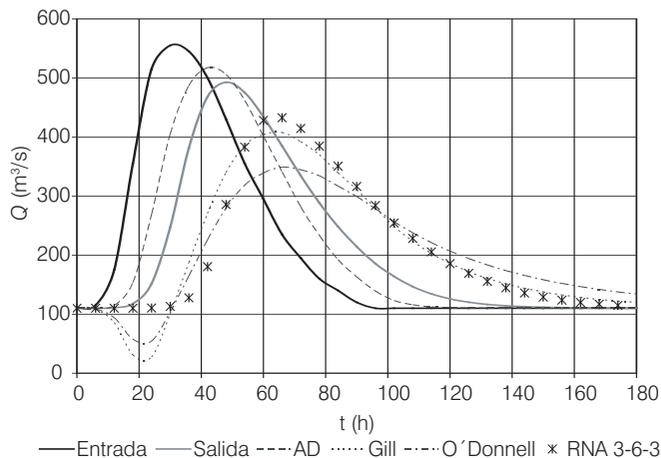
**Ilustración 12. Tránsito de la avenida después de la calibración.**



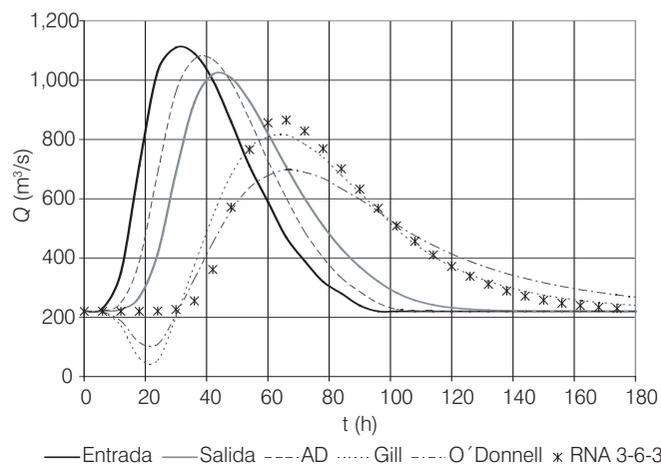
**Ilustración 13. Pronóstico de avenida para dos veces el hidrograma original.**



**Ilustración 14. Tránsito de avenida para cinco veces el hidrograma original.**



**Ilustración 15. Tránsito de avenida para diez veces el hidrograma original.**



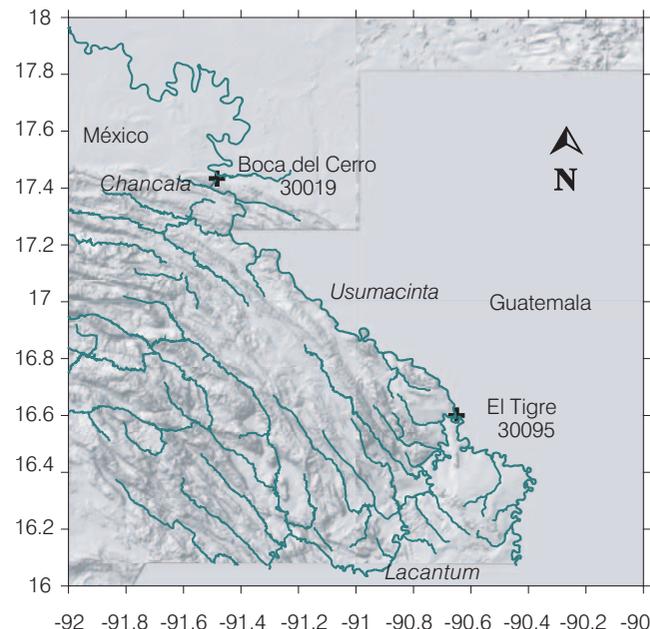
traciones 14 y 15, respectivamente). El método AD presenta gastos mayores a los del hidrograma de salida de referencia en las primeras cincuenta horas aproximadamente y el tiempo pico siempre se adelanta casi seis horas. Se aprecia también que los métodos hidrológicos arrojan gastos pico siempre menores que los del hidrograma de referencia (entre 12 y 32%) y el tiempo pico se retrasa para ambos casos 18 y 24 horas, respectivamente. Los métodos hidrológicos y la RNA presentan el mismo retraso en el tiempo pico con respecto al del hidrograma de salida, pero la RNA tiene mejor aproximación en el gasto pico que los métodos de Gill y O'Donnell, pero no que el AD. En este caso, el método AD es el que presenta mejores resultados.

*Avenidas aisladas con gasto lateral*

El segundo caso estudiado correspondió a datos históricos de avenidas aisladas en la parte baja del río Usumacinta, México. La calibración se realizó con datos de 1964 a 1971 para avenidas históricas registradas en el tramo comprendido entre las estaciones hidrométricas El Tigre (30095) y Boca del Cerro (30019), donde se tienen entradas laterales provenientes de afluentes (ilustración 16). El área drenada aguas arriba de la estación hidrométrica El Tigre es de 41,852 km<sup>2</sup>, y los registros correspondientes permiten conocer el escurrimiento del Usumacinta aguas abajo de la confluencia de los ríos Salinas y Lacantún. La estación hidrométrica Boca del Cerro tiene un área de aportación de 47,697 km<sup>2</sup>. Es en este lugar donde empieza la zona denominada Bajo Usumacinta.

Para fines de entrenamiento se utilizaron cinco avenidas ocurridas entre 1965 y 1970. Estas avenidas se seleccionaron, por un lado, por la disponibilidad de información y por razones de consistencia, porque tienen aproximadamente la misma duración que las avenidas que se usarían para la validación. Se realizará la valida-

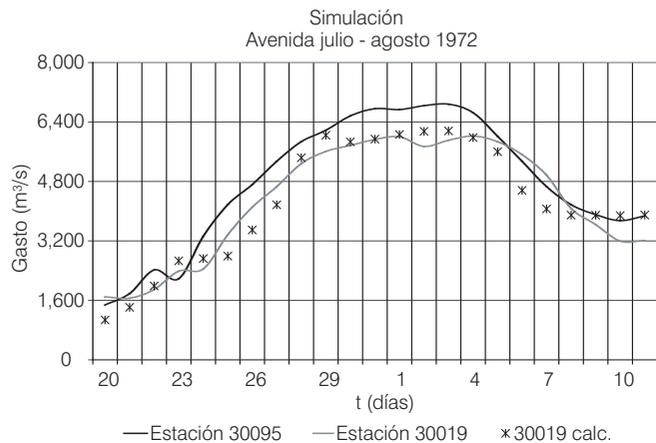
**Ilustración 16. Ubicación de las estaciones de estudio en el río Usumacinta.**



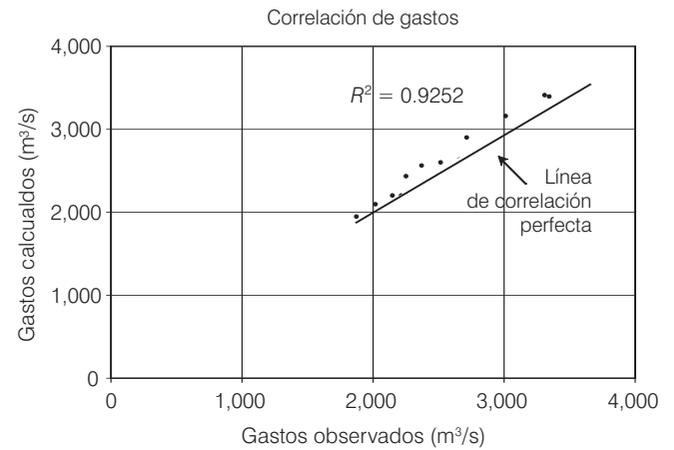
**Cuadro 4. Avenidas utilizadas en el entrenamiento.**

Año Estación	junio-julio 1965		julio 1966		junio-julio 1968		junio-julio 1969		julio-agosto 1970	
	30095	30019	30095	30019	30095	30019	30095	30019	30095	30019
1	532.67	1 731.68	2 337.48	2 613.21	1 515.62	1 934.36	1 763.44	1 978.46	2 422.01	2 831.96
1	823.89	1 726.70	2 677.43	2 566.96	1 622.68	1 815.13	1 997.41	1 933.03	2 465.58	2 766.77
2	115.11	1 721.72	3 049.30	2 867.27	1 831.84	1 849.46	2 907.06	2 098.35	2 509.16	2 701.57
2	957.25	2 304.34	3 032.58	3 160.45	2 034.13	2 046.46	3 757.21	2 776.47	2 824.16	2 758.27
3	218.80	3 018.05	3 493.69	3 225.98	1 903.30	2 252.46	3 570.16	3 538.02	3 139.16	2 814.97
3	071.94	3 331.73	4 309.35	3 640.81	1 846.08	2 187.54	4 269.22	3 663.04	3 102.67	3 162.42
2	889.50	3 361.95	4 791.80	4 270.30	2 043.58	2 261.66	4 815.19	4 160.25	3 114.20	3 464.45
2	704.96	3 235.96	5 165.49	4 758.30	3 026.75	2 452.54	5 208.22	4 725.66	3 391.37	3 475.33
2	432.06	3 062.52	5 405.92	5 247.06	3 860.50	3 215.39	5 462.78	5 049.52	3 561.08	3 676.26
2	250.65	2 834.86	4 974.36	5 554.70	3 934.69	3 906.25	5 415.99	5 270.90	4 019.02	3 896.99
2	302.18	2 635.85	4 204.63	5 233.80	3 271.35	4 010.10	5 201.67	5 260.96	4 135.19	4 255.70
2	602.74	2 602.16	3 649.34	4 566.24	2 659.05	3 600.40	4 913.20	5 017.71	4 236.55	4 420.97
1	820.67	2 390.01	3 815.11	4 036.61	2 430.77	3 127.93	4 386.13	4 699.46	4 051.60	4 472.40
1	621.67	2 137.74	2 997.95	3 620.98	2 172.63	2 834.78	4 231.25	4 382.37	3 628.82	4 334.11
1	470.00	1 940.06	2 981.75	3 380.73	2 058.90	2 570.02	4 112.89	4 211.67	3 191.94	4 083.30
1	598.00	1 739.82	3 210.94	3 316.03	2 284.81	2 433.85	4 217.39	4 116.61	3 126.32	3 747.40
1	685.67	1 790.54	3 116.91	3 412.17	2 149.46	2 554.71	4 185.73	4 137.69	3 189.84	3 570.21
1	634.33	1 875.82	2 852.85	3 311.75	2 072.85	2 475.49	4 243.32	4 142.65	2 915.74	3 550.59
1	523.33	1 837.52	2 583.90	3 076.14	1 873.96	2 380.05	4 005.39	4 140.17	2 792.42	3 347.54
1	532.00	1 750.69	2 395.11	2 837.84	1 942.91	2 200.65	3 606.65	3 942.20	2 560.43	3 143.91
1	525.67	1 718.59	2 270.42	2 618.24	1 752.03	2 195.55	3 153.30	3 596.08	2 451.82	3 015.09
1	434.33	1 701.79	2 136.88	2 461.00	1 619.05	2 043.00	2 733.42	3 198.10	2 343.21	2 886.26
1	408.50	1 665.48	2 167.71	2 326.00	1 529.54	1 879.57	2 491.70	2 852.89	2 276.35	2 762.13
1	382.67	1 629.17	2 138.96	2 304.63	1 634.71	1 758.24	2 879.15	2 653.53	2 209.48	2 637.99

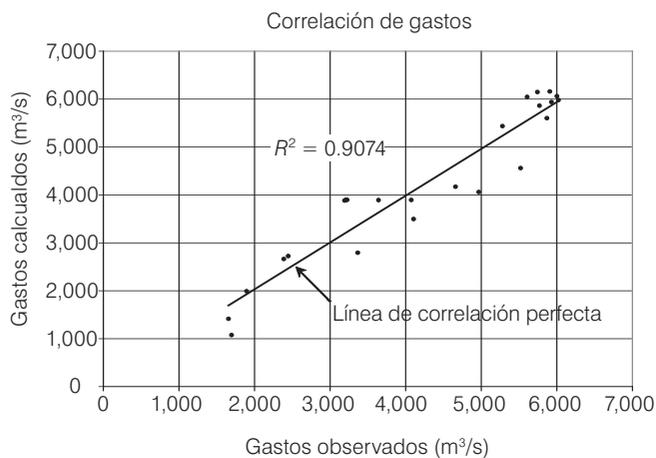
**Ilustración 17. Primer avenida simulada por la RNA entrenada.**



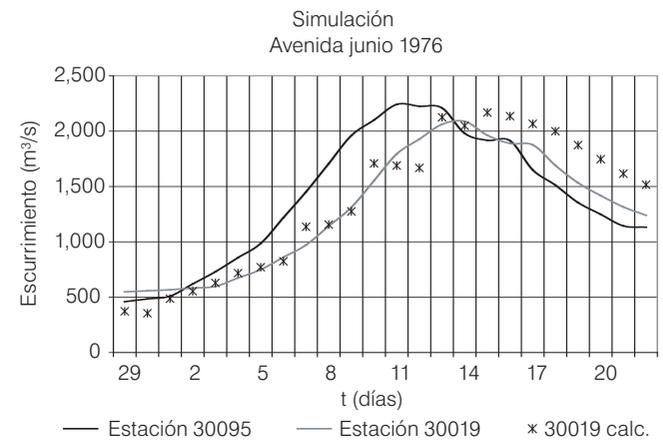
**Ilustración 20. Comparación de gastos calculados contra observados.**



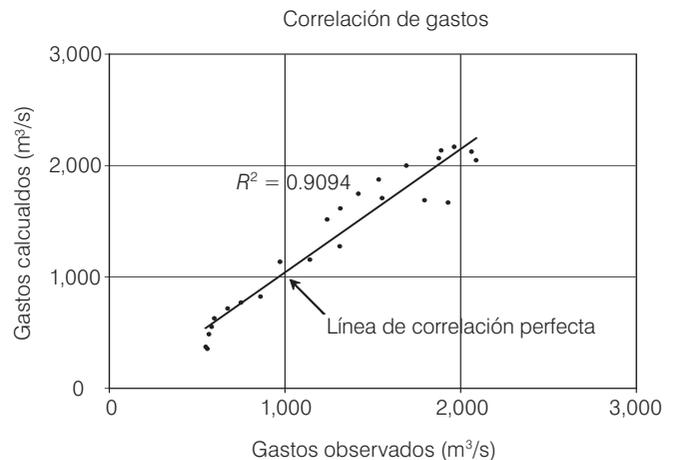
**Ilustración 18. Comparación de gastos calculados contra observados.**



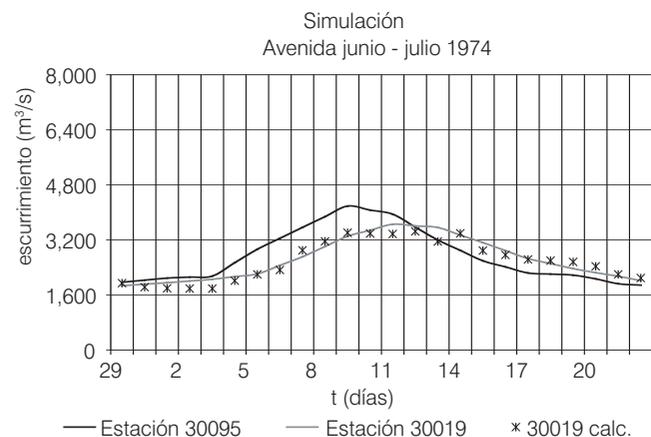
**Ilustración 21. Tercer avenida simulada por la RNA entrenada.**



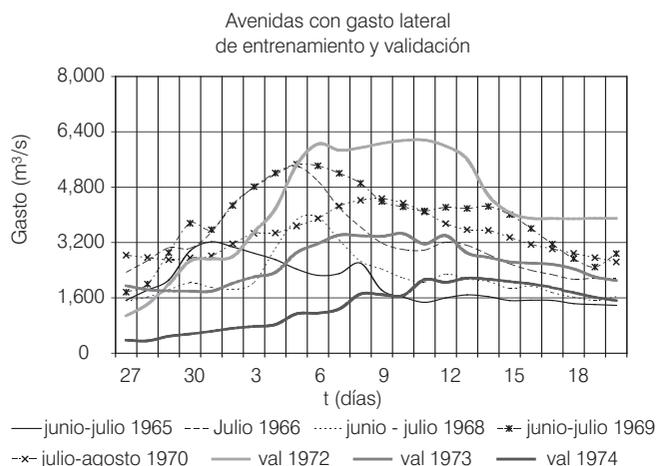
**Ilustración 22. Comparación de gastos calculados contra observados.**



**Ilustración 19. Segunda avenida simulada por la RNA entrenada.**



**Ilustración 23. Avenidas de entrenamiento y validación de la RNA entrenada.**



ción de la RNA con tres avenidas de 1972, 1974 y 1965. Los gastos medios diarios registrados durante junio-julio de 1965, julio de 1966, junio-julio de 1968, junio-julio de 1969 y julio-agosto de 1970 de las avenidas seleccionadas se presentan en el cuadro 4. El intervalo de gastos cubiertos en el entrenamiento de la red va desde los 1,300 hasta los 6,900 m<sup>3</sup>/s, y se seleccionaron avenidas con formas muy diferentes entre sí para evaluar el comportamiento de la metodología en estas condiciones.

Las avenidas utilizadas para fines de validación se presentaron en julio-agosto de 1972 (ilustración 17), junio-julio de 1974 (ilustración 19) y junio de 1976 (ilustración 21). En las ilustraciones 18, 20 y 22 se aprecian las correlaciones existentes entre los gastos observados en la estación hidrométrica Boca del Cerro y los calculados con la red neuronal entrenada, las cuales se consideran aceptables ( $R^2 = 0.9074$ ,  $R^2 = 0.9252$  y  $R^2 = 0.9294$ , respectivamente).

Es importante notar que las avenidas simuladas son notablemente diferentes en forma y desarrollo de las avenidas utilizadas en el entrenamiento, como se muestra en la ilustración 23. De las avenidas de entrenamiento seleccionadas, se determina que el efecto de la cuenca propia entre ambas estaciones repercute en un intervalo del 6.25 al 11.53% del volumen total registrado para la avenida en la estación hidrométrica Boca del Cerro.

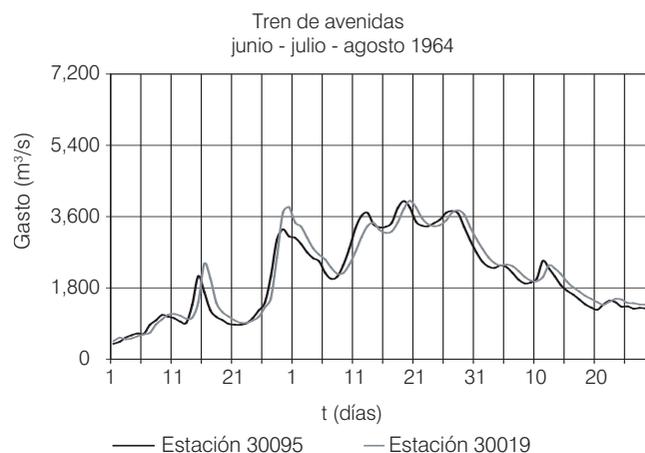
Se llevó a cabo la comparación del volumen de escurrimiento correspondiente a la avenida observada con el resultado de la simulación hecha con la RNA entrenada para observar el comportamiento; en los primeros dos casos se observó que se subestimó el volumen real en un 0.52 y 1.82%, respectivamente, mientras que en el último caso se sobrestimó su valor en 5.58%.

### Trenes de avenidas

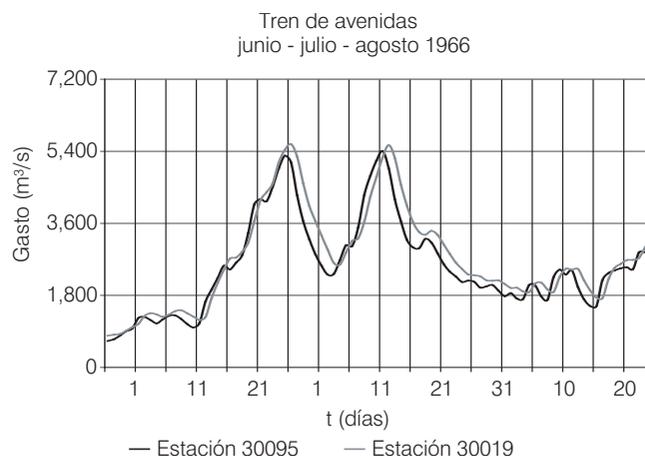
El tercer caso corresponde a la aplicación de una RNA a trenes de avenidas registrados en las mismas estaciones hidrométricas Boca del Cerro (30095) y El Tigre (30019). En este caso se utilizó información de ocho años (de 1964 a 1971) para realizar el entrenamiento de la red con arquitectura 3 - 6 - 3. En las ilustraciones 24 a 26 se muestran ejemplos de tres de los trenes de avenidas usados para el entrenamiento. Por falta de espacio no se muestra el resto de las avenidas, pero éstas se pueden consultar en Molina, 2005. Los años de 1973 a 1976 se usaron para la aplicación de la RNA entrenada (ver ilustraciones 27 a 30).

Como se aprecia en las ilustraciones 27 a 30, la red entrenada sobrestima en algunas zonas el valor del escurrimiento, que corresponde a 2.40, 5.96, 11.19 y

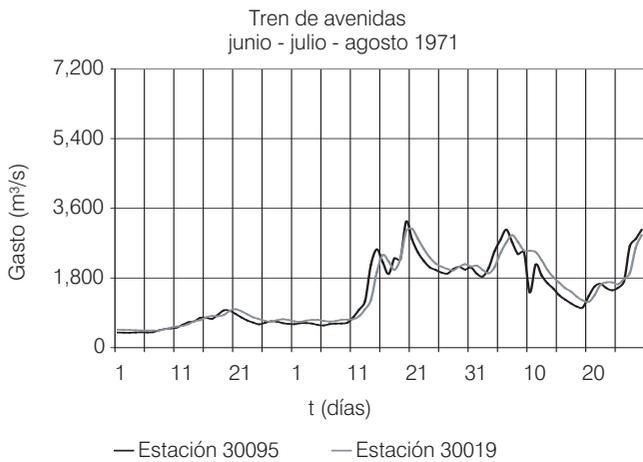
**Ilustración 24. Primer tren de avenidas de entrenamiento.**



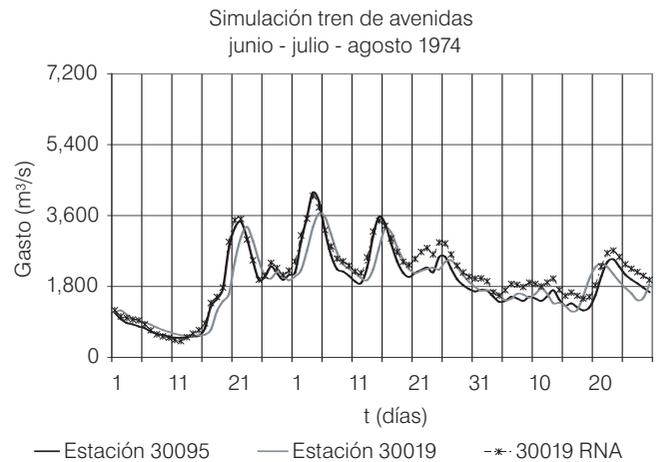
**Ilustración 25. Tercer tren de avenidas de entrenamiento.**



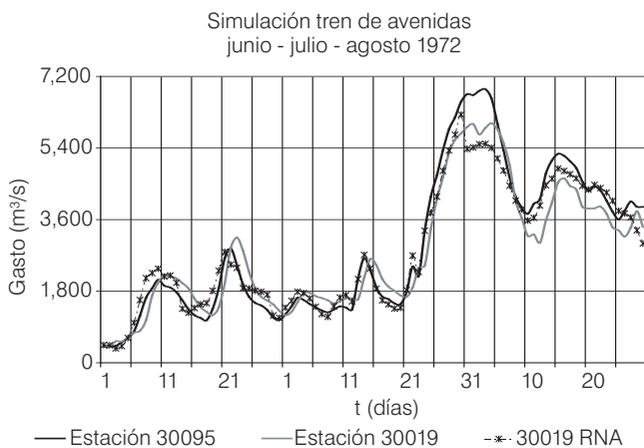
**Ilustración 26. Octavo tren de avenidas de entrenamiento.**



**Ilustración 29. Tercer tren de avenidas simulado.**



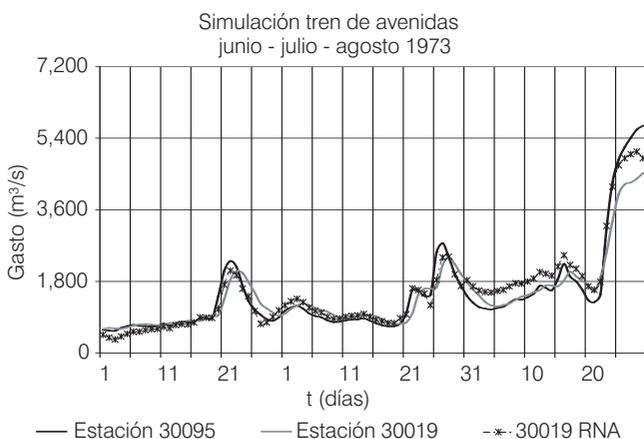
**Ilustración 27. Primer tren de avenidas simulado.**



**Ilustración 30. Cuarto tren de avenidas simulado.**



**Ilustración 28. Segundo tren de avenidas simulado.**

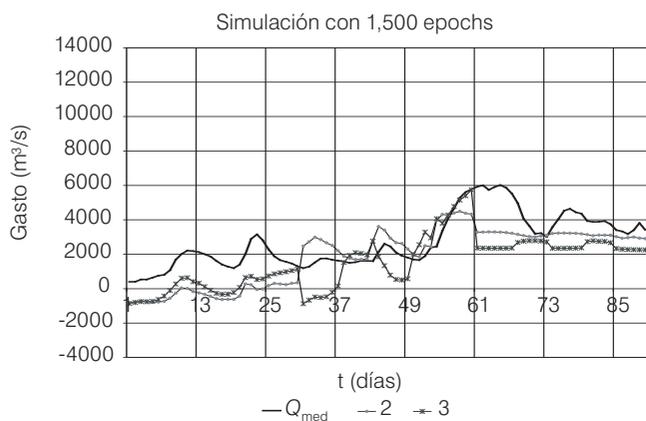


13.05%, respectivamente, del volumen total escurrido para cada tren de avenidas simulado.

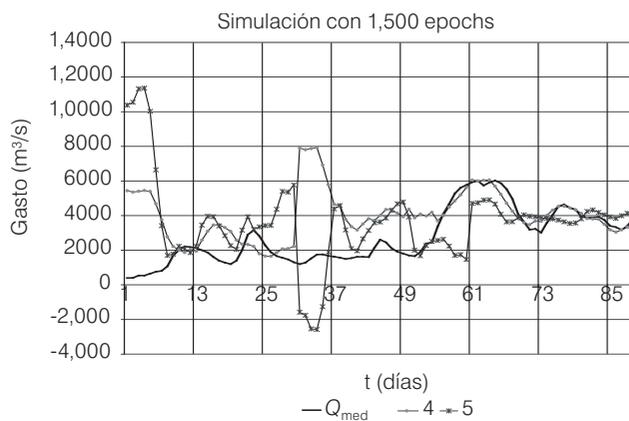
En las ilustraciones 31 (a, b, c) se pueden observar las diferencias existentes en la simulación por parte de la RNA 3 – 6 – 3 para el desarrollo de este caso al variar la cantidad de ejemplos para el aprendizaje. Como es natural, a medida que el aprendizaje de la red se lleva a cabo con más información, mejor es su capacidad de generalización, y el valor de MSE varía según la cantidad de neuronas ocultas para llevar a cabo el aprendizaje, por lo que dicho valor es punto de comparación, siempre y cuando se esté tomando en una red la misma cantidad de información.

En la ilustración 31a, con dos trenes de avenidas para el aprendizaje, la red trabaja con 180 datos en la capa de entrada y 90 en la de salida; con tres trenes lo hace con

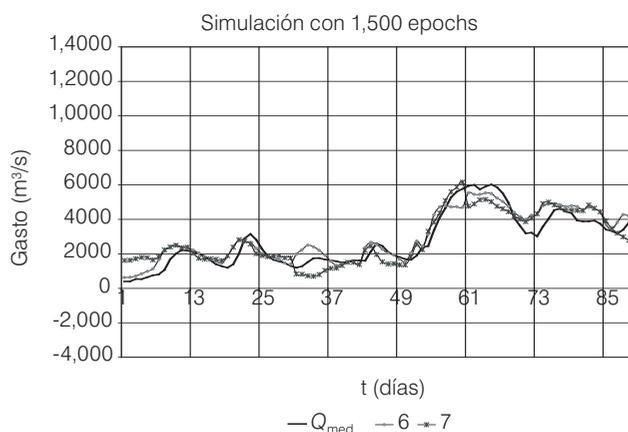
**Ilustración 31. Desempeño de la RNA utilizando diferente número de trenes de avenidas durante el aprendizaje.**



a) Dos y tres trenes de avenidas en el aprendizaje



b) Cuatro y cinco trenes de avenidas en el aprendizaje



c) Seis y siete trenes de avenidas en el aprendizaje

270 en la capa de entrada y 90 en la de salida, y así sucesivamente. En este trabajo se procuró trabajar con el mínimo indispensable de trenes de avenidas en el entrenamiento para conocer hasta qué grado de simplicidad se pueden generar resultados comparables con los métodos hidrológicos, y apreciar la tendencia a obtener un mejor entrenamiento a medida que se aumenta el número de trenes. Se observa una considerable mejoría al utilizar de seis a siete trenes de avenidas en el entrenamiento (ilustración 31c). Se aprecia la ventaja de trabajar con un registro amplio de datos en la fase de entrenamiento.

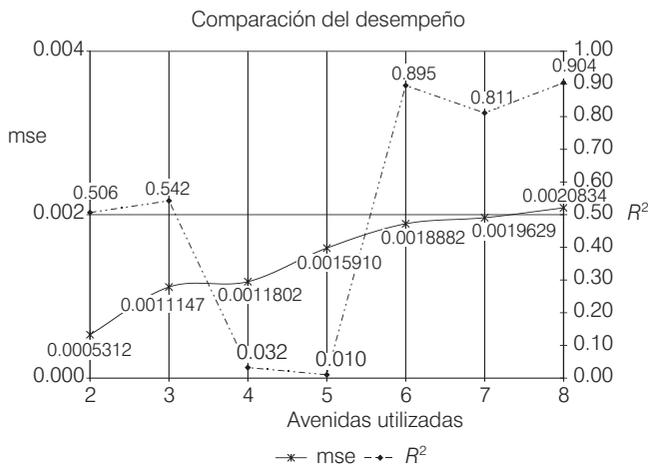
En la ilustración 32 se aprecian los valores de MSE y de  $R^2$  obtenidos en el entrenamiento de la RNA con arquitectura 3 – 6 – 3 con distinto número de trenes de avenidas de entrenamiento. Se puede apreciar que conforme aumentan los ejemplos de entrenamiento (es decir, al aumentar la información), el valor de MSE también se incrementa, de tal manera que no es un parámetro adecuado de comparación para determinar la cantidad

de ejemplos necesarios para llevar a cabo el entrenamiento. Con respecto al valor de  $R^2$ , se observa que la RNA entrenada no presenta una tendencia definida al aumentar la cantidad de ejemplos para el entrenamiento. Por ello se recomienda graficar los resultados y tras una inspección de la gráfica definir el número de ejemplos adecuado para el entrenamiento. Finalmente, las simulaciones de las ilustraciones 27 a 30, realizadas para la RNA entrenada con ocho trenes de avenidas, muestran un comportamiento satisfactorio en términos de los coeficientes de correlación obtenidos. Por ejemplo, el coeficiente de correlación resultó  $R^2 = 0.9125$  para el tren de avenidas de 1972.

### Confluencias

El cuarto caso se enfocó al análisis de gastos históricos (con influencia de flujos laterales) registrados en la estación hidrométrica Azueta, ubicada en el río Tesechoacán,

**Ilustración 32. Valor de MSE y  $R^2$  de la RNA para diferentes números de trenes de avenidas en el entrenamiento.**

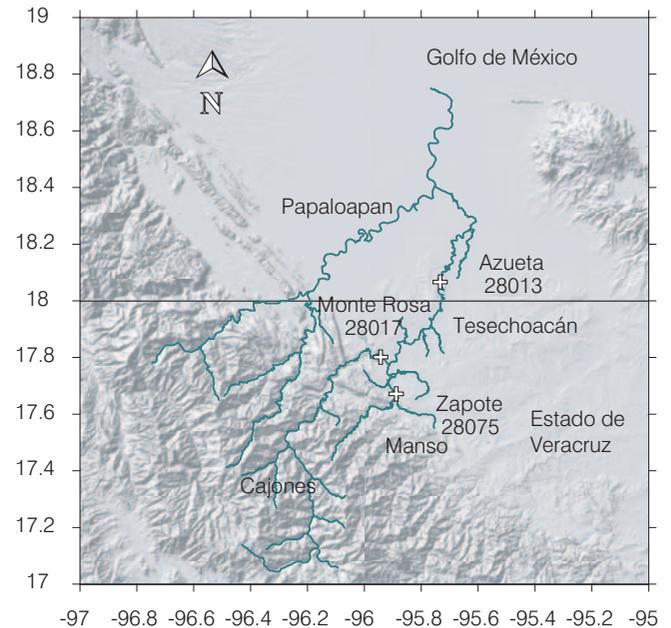


y generados aguas arriba debido al escurrimiento superficial de la confluencia de los ríos Manso y Cajones, donde se encuentran ubicadas las estaciones hidrométricas Zapote y Monte Rosa, respectivamente (ilustración 33). La subcuenca del río Cajones, hasta la estación Monte Rosa, tiene un área de 2,850 km<sup>2</sup> y la subcuenca del río Manso, 805 km<sup>2</sup>.

Éste es un caso que tiene mayor complejidad que los anteriores, pues involucra tres estaciones. Para abordarlo, una vez analizados el comportamiento de los casos anteriores y basado en resultados de estudios con RNA previos (Kavzoglu y Mather, 2003), se propone utilizar una RNA con estructura 6 – 12 – 3, asignando tres neuronas por cauce, que presente la información de manera cuatrimestral para el número de años tomados para llevar a cabo el entrenamiento y la posterior simulación; así, cada neurona manejará 122 datos por cada año seleccionado para el entrenamiento. Considerando la experiencia ganada con los casos anteriores, se usará el doble de neuronas en la capa oculta que en la de salida. En este caso se tomarán los registros históricos correspondientes a los años 1975, 1976 y 1977, y se destinarán los dos primeros al entrenamiento de la red y el último a la simulación. En las ilustraciones 34, 35 y 36 se muestran los registros históricos seleccionados para realizar el entrenamiento, así como la simulación de los registros en las estaciones Monte Rosa, Zapote y Azueta.

En la ilustración 37 se muestra el resultado de la simulación para 1977, una vez que se llevó a cabo el entrenamiento de la red. Se observa que a pesar de la dificultad que representa el tránsito de avenidas en uniones de cauces, con la presencia de flujos laterales y para un amplio registro de datos, el comportamiento presentado

**Ilustración 33. Ubicación de las estaciones de estudio para el caso de confluencias.**



es razonable en cuanto a que se reproducen los picos de los hidrogramas con un error máximo del 21.99% y prácticamente no existe desfase de los hidrogramas simulados con respecto a los registrados, aunque en algunas zonas puede haber diferencias relativamente pequeñas en relación con los valores observados. En la ilustración 38 se graficaron los gastos medidos y simulados para una arquitectura 6 – 12 – 3 y en el cuadro 5 están los coeficientes de correlación resultantes para las distintas propuestas de arquitectura, todas ellas con los mismos parámetros, con excepción del número de neuronas en la capa oculta. Se observa para los casos 6 – 6 – 3 y 6 – 12 – 3 que los valores de estos coeficientes son prácticamente los mismos. Sin embargo, graficando los valores calculados en la simulación de ambas arquitecturas, se observó la diferencia del entrenamiento de cada una de ellas y es de notar que la RNA 6 – 12 – 3 muestra un mejor desempeño no sólo en el valor de dichos coeficientes, sino porque representa más adecuadamente el hidrograma de salida registrado en la estación 28013 en 1977.

En la ilustración 39 se han graficado seis series de valores simulados con una RNA de arquitectura 6 – 6 <math>n\_1</math> <math>16</math> – 3 contra los valores registrados en la estación hidrométrica Azueta. Se aprecia, en los valores calculados, que para gastos de hasta 300 m<sup>3</sup>/s, las redes tienen un desempeño aceptable, pero, a medida que el valor del gasto aumenta, se marca una diferencia entre

Ilustración 34. Registro hidrométrico 1975-1977 de la estación 28017 Monte Rosa.

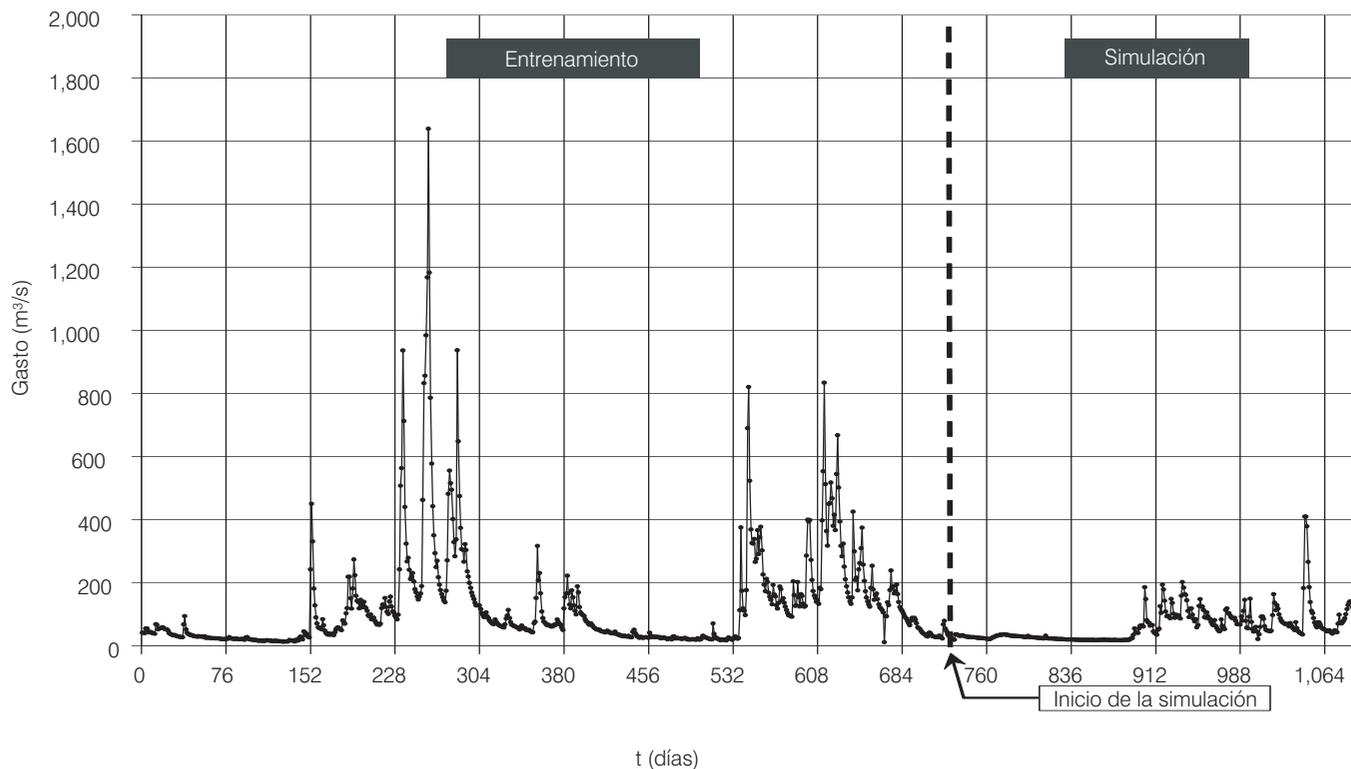


Ilustración 35. Registro hidrométrico 1975-1977 de la estación 28075 Zapote.

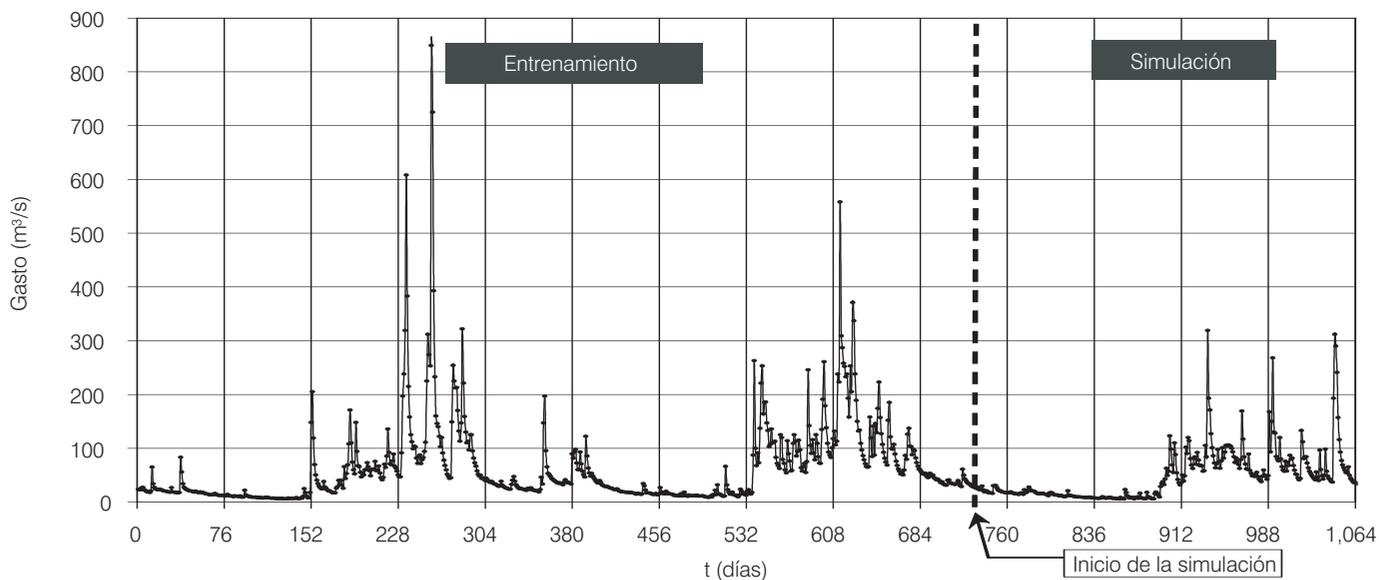


Ilustración 36. Registro hidrométrico 1975-1977 de la estación 28013 Azueta.

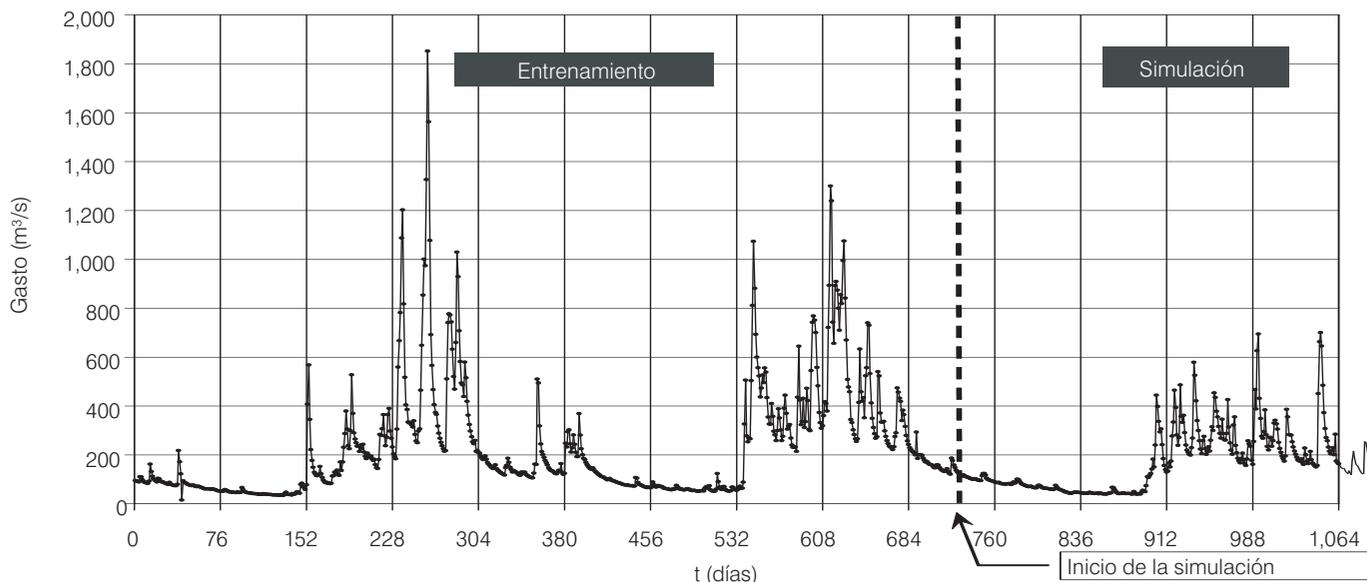
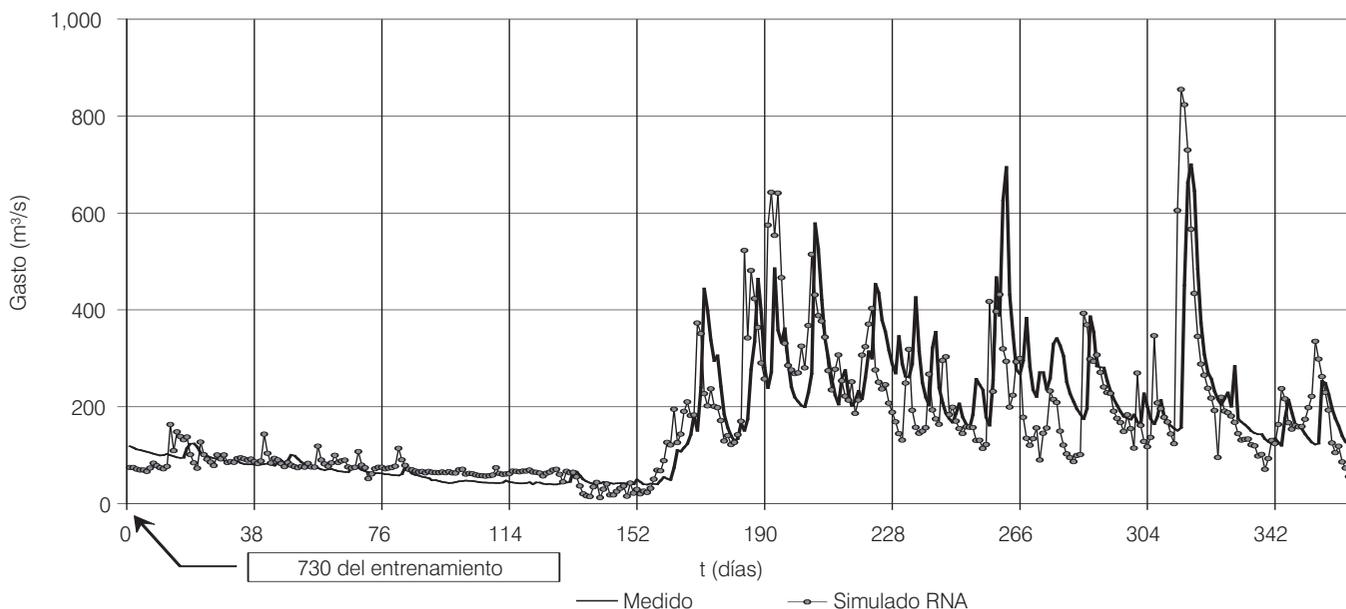


Ilustración 37. Resultado de la simulación del registro hidrométrico de 1977 de la estación 28013 Azueta mediante la RNA entrenada.

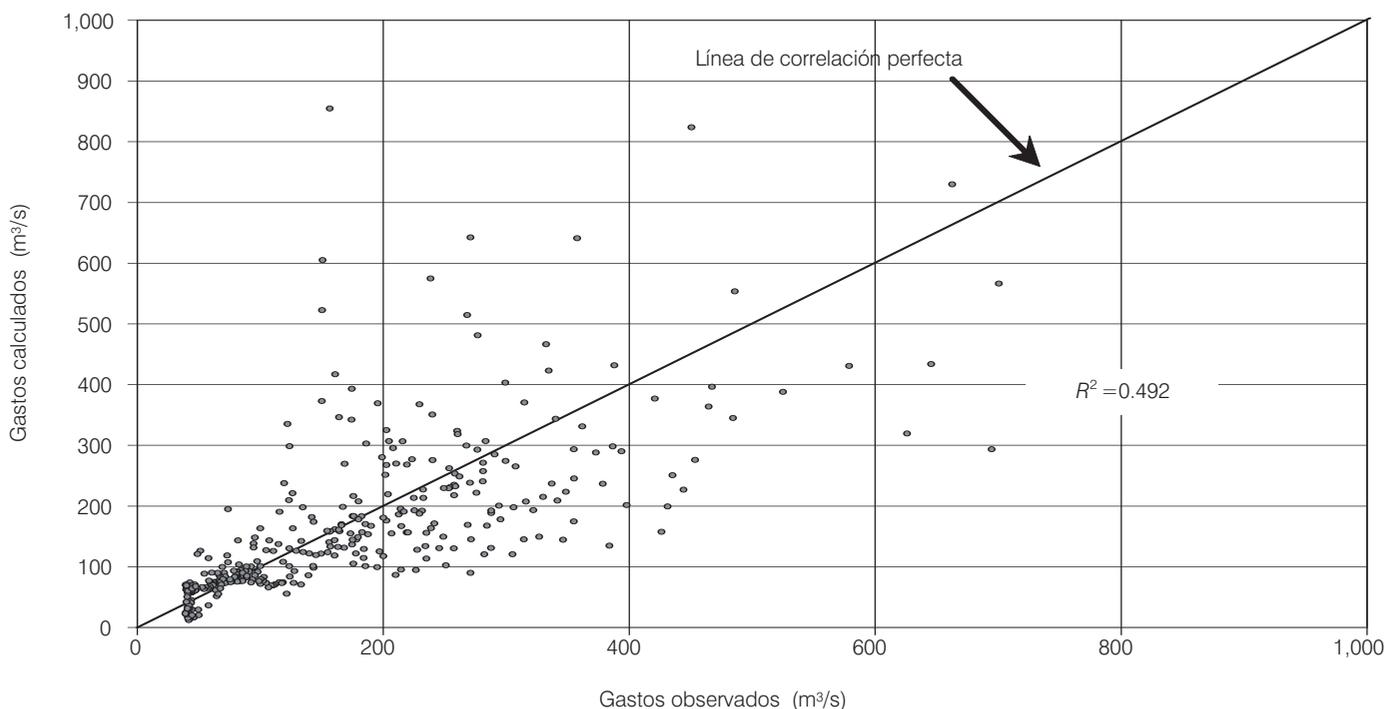


las distintas arquitecturas (número de neuronas en la capa oculta). Para  $n_1 = 6$  existe una serie de gastos en el intervalo de 200 a 400  $m^3/s$  calculados que prácticamente son cero, por lo que la red no tuvo suficientes eventos para aprender los ejemplos presentados en el entrenamiento, dado que es una arquitectura insuficiente (en términos del tamaño de la capa oculta).

En el caso de  $n_1 = 8$ , aparecen gastos negativos debido a que incluso la arquitectura de la red no es capaz de aprender los ejemplos de entrenamiento. De igual forma, los valores calculados en un intervalo de 200 a 400  $m^3/s$  permanecen en su mayoría por debajo de la línea de correlación perfecta debido a la insuficiencia de aprendizaje, por tener pocas unidades procesadoras en

**Ilustración 38. Comparación de gastos calculados contra observados en la simulación de 1977 en Azueta para la RNA 6 – 12 – 3 entrenada.**

Gráfica de correlación



**Cuadro 5. Coeficientes de correlación para distintas arquitecturas de RNA.**

Coeficiente	Arquitectura					
	6 – 6 – 3	6 – 8 – 3	6 – 6 – 3	6 – 8 – 3	6 – 6 – 3	6 – 8 – 3
Correlación	0.699	0.483	0.451	0.702	0.431	0.684

la capa oculta. Para el resto de las arquitecturas se aprecia que para gastos menores a 200 m<sup>3</sup>/s la correlación es buena en general; en el intervalo de 200 a 400 m<sup>3</sup>/s se comienza a marcar una diferencia en el desempeño de los diferentes tamaños de la capa oculta y, finalmente, para valores mayores a 400 m<sup>3</sup>/s, el mejor desempeño en la simulación es el de la arquitectura de  $n_1 = 12$ . Asimismo, puede observarse en el cuadro 5 que el aumentar la cantidad de neuronas en la capa oculta no produjo un mejor resultado.

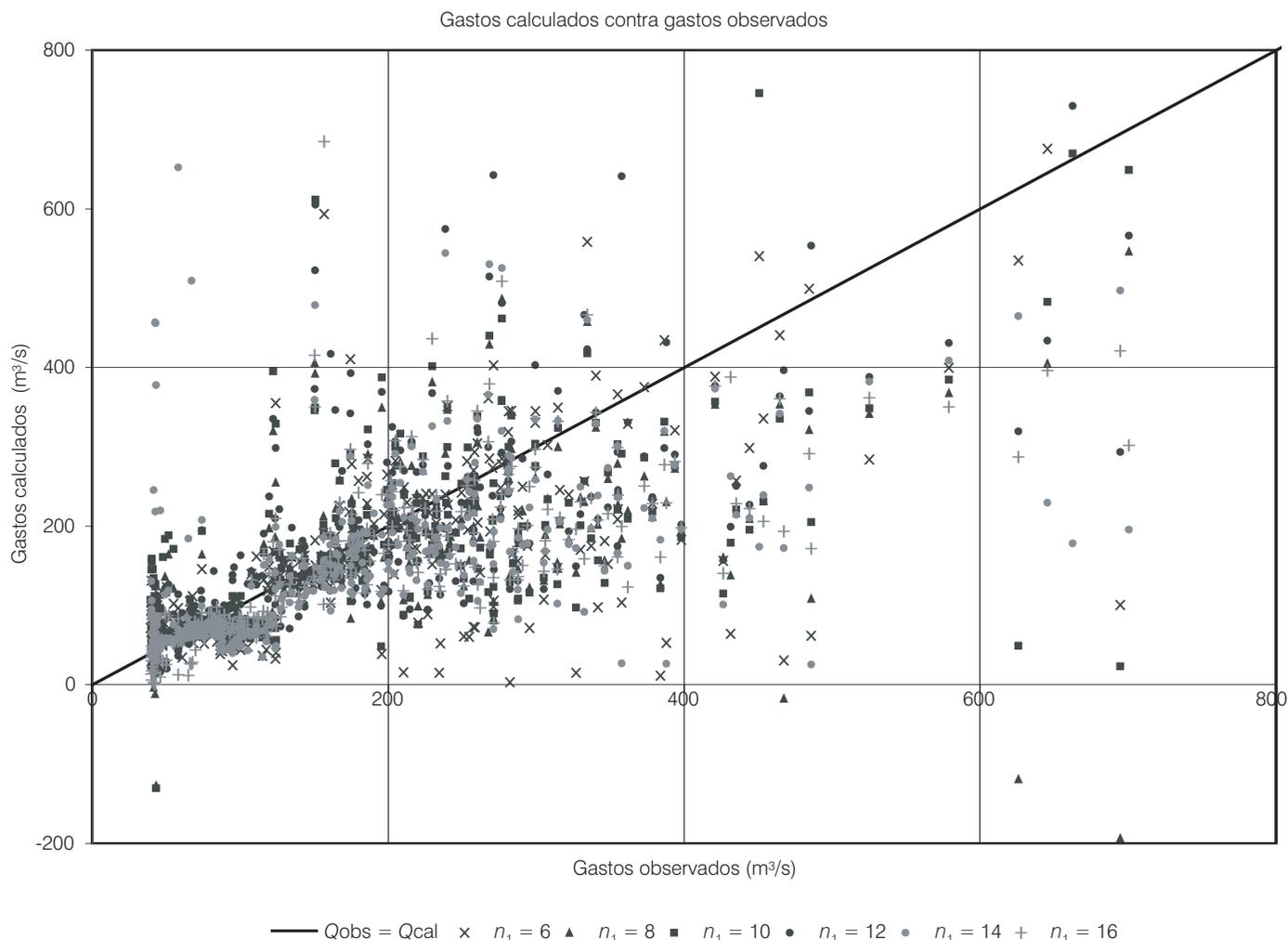
### Conclusiones

Los sistemas hidrológicos contienen procesos complejos influidos por diversos factores, como la topografía, la

cobertura vegetal, el tipo de suelo, las características del cauce y otros. Los proyectos de ingeniería y los análisis de impacto ambiental frecuentemente requieren estimaciones del escurrimiento en un amplio ámbito de problemas relacionados con el diseño y la gestión de los recursos hídricos o sus propiedades características, como el gasto pico y tiempo al pico en lugares donde no se tienen datos de gastos. Se ha desarrollado una gran variedad de métodos para este propósito, incluidos los modelos paramétricos o los modelos empíricos. Las redes neuronales unidireccionales con retropropagación se están utilizando ampliamente en este tipo de aplicaciones.

Las RNA constituyen una técnica con estructura matemática flexible, capaz de reproducir relaciones no lineales complejas entre los datos de entrada y salida sin

Ilustración 39. Tendencia del desempeño respecto al número de neuronas utilizadas en la capa oculta de la RNA con arquitectura 6 –  $n_1$  – 3.



consideraciones detalladas de la estructura interna de los procesos físicos. Las RNA proveen una solución nueva e interesante al problema de relacionar las variables de entrada y salida en sistemas complejos. Si hay pocas neuronas en las capas ocultas, la red puede ser incapaz de describir la función implícita a causa de insuficiencia de parámetros (grados de libertad) para poder representar todos los puntos en los datos de entrenamiento. Por el contrario, si hay demasiadas neuronas, la red tiene un exceso de parámetros libres que pueden provocar el llamado sobreentrenamiento, y con ello perder la habilidad de generalizar. En el caso de tránsito de avenidas, se recomienda utilizar el doble de neuronas en la capa oculta con respecto a las utilizadas en la capa de entrada ( $n_i = 2n_e$ ). Por otra parte, un excesivo número de neuronas ocultas puede retardar el proceso de entrenamiento a tal grado que la red necesite una cantidad considera-

ble de tiempo para aprender. En este artículo se mostró que el uso de las RNA puede tener ventajas sobre otras metodologías para representar el tránsito de avenidas tanto simples como complejas (trenes de avenidas y uniones en cauces), que incluyen entradas laterales significativas, con grandes registros de información, gran cantidad de formas de los hidrogramas, amplios intervalos del valor del escurrimiento e influencia de las temporadas de estiaje y lluvia registrados, entre otros, con resultados aceptables.

A lo largo del desarrollo de este trabajo se observa que entrenar una red neuronal con un amplio registro de información garantiza que la red entrenada tendrá suficiente capacidad de generalización, por lo que será capaz de predecir con buena exactitud avenidas futuras o reproducir las variaciones que sufren las avenidas. Se observa claramente entonces que es importante tener

información suficiente para realizar el aprendizaje de la red de acuerdo con la complejidad del problema, aunque este aspecto está aún poco estudiado en el campo de las redes neuronales.

Se recomienda trabajar con las arquitecturas más sencillas, cuidando de garantizar un aprendizaje adecuado. Es importante hacer notar que la arquitectura de la red por utilizar cambiará de acuerdo con la naturaleza del caso de estudio, por lo que no hay nada definitivo en cuanto a la topología de las redes. De acuerdo con los resultados de este trabajo y con la literatura consultada para el mismo (Kavzoglu y Mather, 2003), se recomienda, como primera propuesta de arquitectura, utilizar en las capas ocultas el doble de las utilizadas en la capa de entrada, para con esto dar libertad a la red de aprender más rápidamente y con mayor efectividad.

Recibido: 19/10/2005  
Aprobado: 10/02/2006

## Referencias

- AGUILAR, E. *Método hidrológico con base física para el tránsito de avenidas en cauces*. Tesis de Maestría. México, D.F.: Facultad de Ingeniería, UNAM, 1995.
- ALDAMA, A. Least-squares parameter estimation for Muskingum flood routing. *J. of Hydr. Engineering*. Vol. 116, núm. 4, 1990, pp. 580-586.
- APARICIO, J. *Fundamentos de hidrología de superficie*. México, D.F.: Limusa Noriega, 2001.
- BALLINI, R. *Redes neuronales para la previsión lluvia escurrimiento*. Sao Paulo, Brasil: Instituto de Ciencias Matemáticas, Universidad de Sao Paulo, 1996.
- BALLINI, R., FRANCA, E., KADOWAKI, M., SOARES, S. y ANDRADE, M. *Modelos de redes neuronales y Box-Jenkins para la previsión de escurrimientos medios mensuales*. Simposium Brasileño de Recursos Hídricos, CD-ROM, 1997.
- CAMPOLO, M., SOLDATI, A. y ANDREUSSI, P. Artificial neural network approach to flood forecasting in the River Arno. *Hydrol. Sci. J.* Vol. 48, núm. 3, 2003, pp. 381-398.
- CIGIZOGLU, H. K. Estimation, forecasting and extrapolation of river flows by artificial neural networks. *Hydrol. Sci. J.* Vol. 48, núm. 3, 2003, pp. 349-361.
- CRESPO, J.L. y MORA, E. Drought estimation with neural networks. *Adv. in Engineering Software*. Vol. 18, 1993, pp. 167-170.
- FLOOD, I. y KARTMAN, N. Neural network in civil engineering. I: principles and understanding. *J. of Comp. in Civil Eng.* Vol. 8, núm. 2, 1994, pp. 131-148.
- FLOOD, I. y KARTMAN, N. Neural network in civil engineering. II: systems and application. *J. of Comp. in Civil Eng.* Vol. 8, núm. 2, 1994, pp. 149-162.
- FRENCH, M. N., KRAJEWSKI, W. F. y CUYKENDALL, R. R. Rainfall forecasting in space and time using a neural network. *J. of Hydrol.* Núm. 137, 1992, pp. 1-31.
- GOVINDARAJU, R. Artificial neural networks in Hydrology I: Preliminar concepts. *J. of Hydrol. Eng.* Vol. 5, núm. 2, 2000, pp. 115-123.
- HEBB, D. *The organization of behavior: a neuropsychological theory*. New York: Wiley, 1949.
- HOPFIELD, J.J. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences of the USA*. Vol. 79, 1982, pp. 2554-2558.
- HSU, K., GUPTA, H. V. y SOROOSHIAN, S. Artificial neural network modeling of the rainfall – runoff process. *Water Res. Res.* Vol. 31, núm. 10, 1995, pp. 2517-2530.
- IMRIE, C. E., DURUNCAN, S. y KORRE, A. River flow prediction using artificial neural networks: generalization beyond the calibration range. *J. of Hydrol.* Vol. 233, 2000, pp. 138-153.
- JAIN S. K., DAS, A. y SRIVASTAVA, D. K. Application of ANN for reservoir inflow prediction and operation. *J. of water res., planning and manag.* Vol. 125, núm. 5, 1999, pp. 263-271.
- KADOWAKI, M., SOARES, S. y ANDRADE, M. G. *Previsão de vazões mensais utilizando redes neurais multicamadas com algoritmo backpropagation*. Goiânia, Brasil: 4º Simpósio Brasileiro de Redes Neurais, 1997.
- KARUNANITHI, N., GRENNEY, J.W., WHITLEY, D. y BOVEE, K. Neural networks for river flow prediction. *J. of Comp. in Civil Eng.* Vol. 8, núm. 2, 1994, pp. 201-220.
- KAVZOGU, T. y MATHER, P. M. The use of back propagating artificial neural networks in land cover classification. *Int. J. of Remote Sensing*. Vol. 24, núm. 23, 2003, pp. 4907-4938.
- LARRAÑAGA, P. e IÑAKI, I. *Tema 8. Redes Neuronales*. Euskadi, País Vasco, España: Departamento de Ciencias de la Computación e Inteligencia Artificial, Universidad del País Vasco – Euskal Erico Unibertsitate, 1997.
- LANNA, A.E. y SCHWARZBACH, M. *MODHAC-modelo hidrológico auto-calibrável*. Florianópolis, Brasil: Instituto de Pesquisas Hidráulicas de UFRGS, 1988.
- MATLAB 6.5. The mathworks, Inc. Matlab, 2002a.
- MATLAB 6.5. Neural network toolbox. User's Guide Version 4.0, 2002b.
- McCULLOCH, W.S. y PITTS, W. A logical calculus of the ideas immanent in nervous activity. *Bull. of Math. Biophysics*. Vol. 5, 1943, pp. 115-133.
- MOLINA, A.J.P. *Tránsito de avenidas en cauces mediante redes neuronales artificiales*. Tesis de maestría. México, D.F.: Facultad de Ingeniería, UNAM, 2005.
- ROCHESTER, N., HOLLAND, J.N., HAIBT, L.H. y DUDA, W.L. Test on cell assembly theory of action of the brain, using a large digital computer. *IRE transactions of information theory IT – 2*, 1956, pp. 80-93.

- ROSENBLATT, F. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review*. Núm. 65, 1958, pp. 386-408.
- RUMELHART, D.E., HINTON, G. E. y WILLIAMS, R.J. Learning representations by back – propagating errors. *Nature*. Vol. 323, 1986, pp. 533-536.
- SARMENTO, F.J. Modelagem de séries hidrológicas a través de redes de neuônios. *Revista Brasileira de Recursos Hídricos – RBRH*. Vol. 1, núm. 2, 1996, pp. 19-31.
- SHAMSELDIN, A.Y., O´CONNOR, K.M. y LIANG, G.C. Methods for combining the outputs of different rainfall–runoff models. *J. of Hydrol.* Vol. 197, 1996, pp. 203-229.
- SHAMSELDIN, A.Y. Application of a neural network technique to rainfall – runoff modeling. *J. of Hydrol.* Vol. 199, 1997, pp. 272-294.
- TAYFUR, G. Artificial neural networks for sheet sediment transport. *Hydrol. Sci. J.* Vol. 47, núm. 6, 2002, pp. 879-892.
- THIRUMALAI, K. y DEO, M.C. River stage forecasting using artificial neural networks. *J. of Hydrol. Engineering*. Vol. 3, núm. 1, 1998, pp. 27-32.
- US ARMY CORPS OF ENGINEERS. Flood Runoff Análisis. Chapter 9. 1999, pp. 24.
- VALENCIA, M.J.S. *Aplicação de redes neurais na área de recursos hídricos*. Simpósio Brasileiro de Recursos Hídricos, 120, CD-ROM, 1997.
- WIDROW, B. y HOFF, M. E. *Adaptive switching circuits*. Institute of Radio Engineers, Western Electronic Show and Convention, 1960, pp. 96-104.
- ZHANG, B. y GOVINDARAJU R.S. Geomorphology-based artificial neural networks (GANNs) for estimation of direct runoff over watersheds. *J. of Hydrol.* Núm. 273, 2003, pp. 18-34.

**Abstract**

MOLINA-AGUILAR, J.P. & APARICIO, J. Flood routing in rivers by artificial neural networks. *Hydraulic engineering in Mexico (in Spanish)*. Vol. XXI, no. 4, October-December, 2006, pp. 65-86.

Commonly used hydrological methods for flood routing in rivers have restrictions in the analysis of complex problems, as for example in the case of sequential flows, lateral flows or river junctions, mainly in cases without hydrometrical information in the whole hydrological network. The characteristics of artificial neural networks make them a possibility for their application to stream flood routing, because they have several advantages with respect to the traditional hydrological methods. Application of artificial neural network to different sample cases, with enough information and selecting appropriate topology, shows that it is possible to obtain results with a similar precision to the hydraulic and hydrological methods, with usually available data in hydrometrical records that are scarce for the application of such methods. Application of artificial neural networks of simple architecture in isolated stream flood routing cases and sequential flows in the hydrological region 30 in Mexico, as well as an annual hydrometrical record in the junction of the Manso and Cajones streams flowing into Tesechoacán river, shows clearly their advantages.

**Keywords:** artificial neural network, stream flood routing, lateral flow, Muskingum.

**Dirección institucional de los autores:**

*Juan Pablo Molina-Aguilar*

Universidad Latina de América,  
Licenciatura en Ingeniería Civil,  
Manantial Cointzio Norte 355,  
Fracc. Manantiales, Morelia, Michoacán, México,  
C.P. 58170,  
teléfono: + (52) (443) 322 1500, extensiones 1014 y 1019,  
jpmolina@unla.edu.mx

*Javier Aparicio*

Instituto Mexicano de Tecnología del Agua,  
Paseo Cuauhnáhuac 8532, Progreso, Jiutepec, Morelos, México,  
C.P. 62550,  
teléfono: + (52) (777) 329 3600, extensión 551,  
japaricio@tlaloc.imta.mx