

Forecast flows in a section of the Bogotá River by Artificial Intelligent Systems

William D. Moscoso

Facultad de Ingeniería
Universidad de La Sabana, Unisabana
Chía, Colombia
william.moscoso@unisabana.edu.co

Luis Mauricio Agudelo-Otálora

Facultad de Ingeniería
Universidad de La Sabana, Unisabana
Chía, Colombia
mauricio.agudelo@unisabana.edu.co

Abstract— This article presents a comparison between two types of intelligent models: Artificial Neural Networks - ANN and Adaptive Neuro-Fuzzy Interference System - ANFIS, for forecasting flows in a section of Bogotá (Colombia) river, looking for the most efficient. The simulation was performed in the Matlab computer software, with data collected by hydrological stations of the Corporación Autónoma Regional of Cundinamarca (CAR), from September 2009 to October 2013. The findings suggest that by using artificial intelligence models you can reach a successful outcome, with Correlation Coefficients above 90 % (CC), Mean Absolute Percentage Error (MAPE) below 12 %, Concordance Correlation Coefficient to 84 %, six other statistical evaluating precision and accuracy, suggesting that forecasts will be labeled as good and could think of the use of these techniques in Colombia.

Key words—Artificial Neural Networks, Neuro-fuzzy Interference System - ANFIS, flood forecasting, flow rate, volume of flood.

I. INTRODUCCIÓN

Las inundaciones son un fenómeno natural que se produce cuando las lluvias se presentan de manera frecuente o son tan fuertes que la capacidad de absorción del suelo es sobrepasada, generando que el agua cambie de curso y se extienda hacia las zonas adyacentes al mismo. En Colombia, el Fenómeno de La Niña es uno de los desastres naturales que representa una gran amenaza a la población. Al saturarse de humedad de los suelos, con frecuencia ocurre deslizamientos y crecientes rápidas en los sistemas hídricos en la regiones Andina, Caribe y Pacífica. En enero de 2011, tras un fuerte temporal de lluvias fue necesario declarar estado de emergencia económica, social y ecológica en todo el territorio colombiano debido a los devastadores efectos generados por este fenómeno [1][2].

Tradicionalmente, la hidrología ha recurrido a métodos de pronóstico de las inundaciones mediante regresiones lineales, las cuales miden la relación entre las variables dependientes e independientes del fenómeno, utilizando dentro de estas, los caudales [3][4]. Como los comportamientos de los caudales no son lineales, las regresiones no son apropiadas para la predicción. Adicionalmente, estos modelos requieren incluir variables de tipo hidrológico y morfológico de la cuenca que no están disponibles para la aplicación del presente estudio. Otro aspecto que influye en la predicción es el cambio climático que se está presentando en la tierra [5], lo que hace

inapropiado el uso de estos modelos [7][8], por lo que es necesario emplear herramientas de optimización robustas para un cálculo apropiado [6].

Debido a los problemas mencionados con estos modelos tradicionales, diferentes investigadores han indagado nuevos métodos de predicción de inundaciones para implementar planes de emergencia en zonas de alta vulnerabilidad que amortigüen el impacto en la sociedad, generando información útil al momento de tomar decisiones para realizar un sistema de alertas tempranas. Actualmente, se han desarrollado estudios de modelos de predicción integrando técnicas de Inteligencia Artificial, las cuales tienen una estructura matemática flexible que es capaz de identificar relaciones complejas no lineales entre las características de los datos de entrada y de salida, donde es difícil de describir el proceso utilizando ecuaciones físicas [9]. Una de las herramientas más usadas en el campo de la Inteligencia Artificial, en cuanto al pronóstico de inundaciones a nivel global, son las técnicas de Soft Computing como la lógica difusa, que permite tratar información imprecisa pero útil para el análisis de variables, las ANN que simulan el funcionamiento del cerebro para la resolución de problemas o los ANFIS que representan una combinación de las herramientas previamente mencionadas, que pueden ser utilizadas para llevar a cabo la elaboración de dichos modelos [7][10][11][12]. Con el presente estudio se pretende observar cual es el mejor desempeño para predicción de inundaciones entre dos técnicas: ANN y ANFIS.

Recientes estudios han desarrollado modelos de inundaciones implementado técnicas de Inteligencia Artificial, como la realizada en Grecia en el Observatorio Nacional de Atenas (NOA), donde se pronosticó la precipitación máxima diaria aplicando modelos de Redes neuronales ANN, utilizando datos durante los años de 1891-2009. Los resultados de la investigación sugieren que dependiendo de la frecuencia del entrenamiento de las ANN puede tener un impacto en la formación óptima de la predicción [13]. Otro caso de la utilización de las ANN en la disciplina de la hidrología es la realizada en la Cuenca Karasu que se encuentra en la Cuenca del río Eufrates en Turquía [14], donde se describe la implementación de una ANN para predecir los flujos de captación de nieve basados en los datos meteorológicos. Para esto el modelo se calibro con datos de escorrentía disponible en un periodo de 12 años. Los resultados del modelo son alentadores y la exactitud se puede aumentar aún más cuando las simulaciones incluyen datos de profundidad de la nieve. En la actualidad, las técnicas de Soft

Computing son referentes para realizar investigaciones de predicción de eventos futuros debido a que emulan el comportamiento no lineal del fenómeno para dar un pronóstico más acertado [15], además de ayudar en la adecuada toma de decisiones frente al uso del agua. Desde este panorama se presenta esta investigación desarrollada en un tramo de la cuenca alta del río Bogotá, cuyo objetivo es el pronóstico de caudales con los cuales se podrían realizar estudios para el pronóstico de inundaciones.

II. MARCO CONCEPTUAL

A. Redes Neuronales Artificiales (ANN)

Las Redes Neuronales Artificiales (ANN – Artificial Neural Network) son modelos matemáticos inspirados en procesos neurobiológicos en los que el análisis de la información se imita a las acciones desarrolladas por las neuronas en el cerebro. La estructura estándar de una ANN está compuesta por un conjunto de neuronas organizadas en capas (entrada, ocultas y de salida), distribuidas jerárquicamente, constituyendo un sistema funcional autónomo. En este sistema inteligente, como el mostrado en la figura 1, se identifican los siguientes elementos: variables de entrada y salida, pesos sinápticos, que son la intensidad de interacción entre las neuronas, función de propagación, función de activación y función de salida [16]. Además de estos elementos, se deben tener en cuenta el número de capas y neuronas siendo los parámetros más importantes en la modelación de ANN, ya que estos determinan la eficiencia del sistema [17].

Una de las ventajas de las ANN es que pueden ser una herramienta útil para el modelado cuando la relación entre los datos de entrada y salida no es clara o en su defecto es desconocida, razón por la cual este tipo de modelos son llamados de caja negra [12][18], lo que permite que a través de sus composiciones matemáticas sean capaces de modelar sistemas complejos como los sistemas hidrológicos [7]. Otro de los beneficios derivados de las ANN es su capacidad de generar salidas de una combinación específica de entradas y su capacidad de respuesta frente al manejo de datos no lineales, estos sistemas inteligentes pueden llegar a ser mejores que los sistemas de modelos lineales, debido a su flexibilidad al abordar problemas complejos [15].

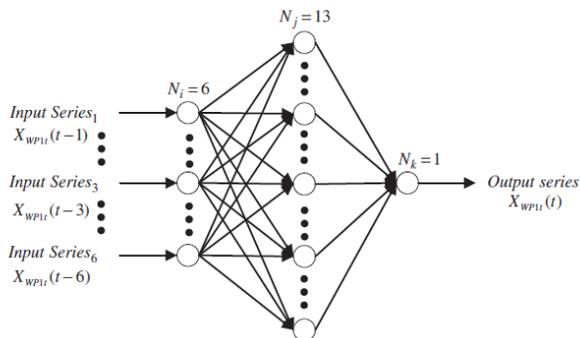


Fig. 1. Redes Neuronal Artificial de tres capas [19]

B. Sistema Neuro-difuso (ANFIS)

Para explicar que es un Sistema Neuro-difuso se debe partir del enfoque de la Lógica Difusa que se basa en expresiones lingüísticas inciertas en lugar de la incertidumbre numérica. Esta técnica se basa en un Sistema de Inferencia Difusa (FIS) como el mostrado en la fig. 2, basado en tres componentes: una base de normas (rulebase) que contiene las reglas difusas si-entonces, una base de datos (database) definida por una función de pertenencia y un sistema de inferencia (interference system), que combina las reglas difusas y produce los resultados del sistema [20].

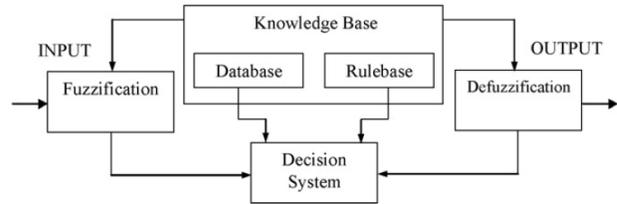


Fig. 2. Estructura de un Sistema de Inferencia Difuso [20]

La estructura de un Sistema Neuro-difuso consiste en la unión de dos técnicas artificiales inteligentes: Redes Neuronales Artificiales y Lógica Difusa, razón por la cual es llamada ANFIS (Adaptative Neuro-fuzzy Interference System). De las ANN utiliza los algoritmos de aprendizaje y su configuración por medio de capas y neuronas, mientras que de los FIS utiliza el razonamiento difuso que permite generar reglas de inferencia a partir de la asignación de variables lingüísticas a la información, caracterizando la combinación de cada una de las entradas a una o varias salidas. Esta combinación permite a la red organizarse a sí misma y generar adaptabilidad del sistema difuso para resolver distintos problemas [18][21]. Una estructura general de un ANFIS se muestra en la fig. 3.

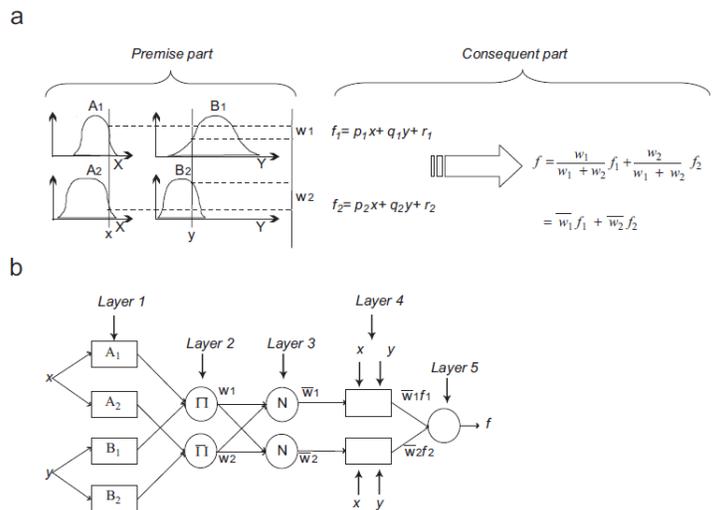


Fig. 3. (a) Sistema de Inferencia difuso. (b) Estructura de un ANFIS [21]

III. METODOLOGÍA

A. Información hidrológica

La cuenca del río Bogotá está ubicada sobre el departamento de Cundinamarca (Colombia), con una extensión de 170km dividida en tres cuencas: alta, media y baja. Su cuidado y monitoreo está a cargo de la Corporación Autónoma Regional – C.A.R, y la cartografía de la misma fue creada por el Instituto Geográfico Agustín Codazzi – IGAC.

Se realizó un estudio de las estaciones hidrológicas ubicadas en la cuenca alta del río, encontrando inconvenientes con la información suministrada, debido a que la mayoría de esta información está en una frecuencia diaria lo que no permite ver los sucesos de inundación en una forma detallada. Por este motivo, se decidió identificar dos estaciones cercanas que tuvieran información horaria, teniendo en cuenta que este tipo de información solo está presente desde septiembre de 2009, fecha en la que instalaron estas estaciones satelitales por parte de la C.A.R.

El tramo elegido para la simulación de los modelos está ubicado entre las estaciones Puente Florencia y Tocancipá que atraviesa los municipios de Gachancipá y Tocancipá. Esta selección se realizó después de un minucioso estudio de las estaciones ubicadas en la cuenca alta del Río Bogotá. Para esto se tuvo en cuenta: estabilidad del tramo, frecuencia de muestreo y subcuencas encontradas de fácil modelación.

En la fig. 4 muestra el tramo elegido y la ubicación correspondiente.

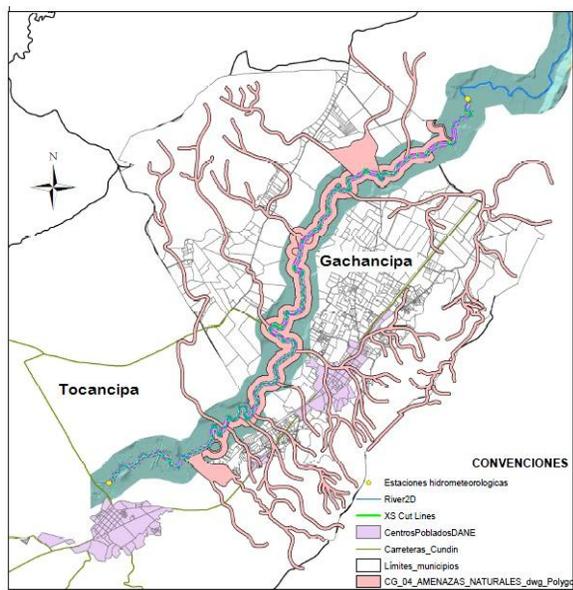


Fig. 4. Tramo elegido (CAR e IGAC)

El periodo tomado para esta investigación es desde el año 2009 hasta el año 2013 con la información de los caudales registrados en las estaciones mencionadas. Los aportes de 12 subcuencas encontradas en este tramo se calcularon por medio del método racional [22], el cual es útil en cuencas pequeñas, donde el límite está dado por las características de las lluvias intensas en la zona y donde la lluvia puntual se pueda

considerar uniformemente distribuida en el área [23]. La ecuación para obtener los caudales es:

$$Q = \frac{C * I * A}{360} \quad (1)$$

Dónde: Q es el caudal en m³/s, C es el coeficiente de escorrentía adimensional, I es intensidad de la lluvia en mm/hora y A es el área en hectáreas.

B. Redes Neuronales Artificiales (ANN)

La simulación de las ANN se realizó en el toolbox del software Matlab 2013 ® llamado nntool (neural network tool) para el cual se hizo una preparación de los datos tanto de entrenamiento como de pronóstico de la red, normalizándolos en un rango de -1 a 1 [24]. La estructura de la red neuronal se compone de 13 conjuntos de datos de entrada que representan los caudales medidos en el tramo seleccionado (1 conjunto de datos del caudal en la estación hidrológica del extremo superior del tramo – estación de Puente Florencia y 12 conjuntos de datos de caudales aportados por subcuencas que alimentan el río) y un conjunto de datos de salida del caudal del extremo inferior del tramo seleccionado – estación de Tocancipá. En cuanto a los parámetros utilizados para la calibración del modelo, se estableció un tipo de entrenamiento backpropagation [25], debido a que la arquitectura implementada fue Multi-capas, con neuronas ocultas, además se utilizó la función de aprendizaje Levenberg-Marquardt (trainlm), ya que realiza mejor la función de ajuste para el reconocimiento de patrones del sistema [24], utilizada en conjunto a la función de medida característica de error cuadrático (MSE).

La función de activación establecida fue del tipo de distribución Triangular, con un máximo de 2000 repeticiones para correr el modelo hasta llegar a la validación total del procesamiento de la información, un gradiente mínimo de 1e-05, y un máximo de 6 revisiones de validación para evaluar la calidad del modelo. Es importante mencionar que el programa computacional divide los datos en dos muestras, el 90% del total de los datos ingresados son para el diseño de la red en el entorno de desarrollo y el 10% para la validación de la misma.

Por otro lado, para cada escenario entrenado, entendido como el conjunto de criterios predeterminados en el entorno del toolbox de Matlab 2013 ®, se utilizó una estructura Multi-capas [6] modificando el número de capas (entre 2 y 20) y de neuronas (entre 2 y 50). En total se realizaron 168 escenarios, los cuales se dividieron según su función de propagación, esto quiere decir que se entrenaron 85 escenarios con una función de propagación Sigmoidal-Sigmoidal, por lo cual el Sistema Neuronal adquiere un comportamiento de distribución Sigmoidal tanto en sus capas de entrada como en la oculta. Se realizaron 83 escenarios con una función de propagación con distribución Sigmoidal-Lineal para las capas de entrada y en la oculta respectivamente. En cada escenario entrenado se realizó la simulación del pronóstico durante el horizonte de tiempo elegido y con los resultados obtenidos se hallaron cada uno de los índices estadísticos por medio de un programa realizado en Matlab 2013 ®.

C. Sistema Neuro-difuso (ANFIS)

La simulación de los ANFIS se realizó en el toolbox de Matlab, llamado anfisedit con dos entradas: una con los caudales de la estación de Puente Florencia y otra con la suma del caudal total aportado por las 12 subcuencas, ya que los modelos ANFIS presentan limitaciones al simular un modelo con muchas entradas [21] y una salida que corresponde a los caudales de la estación de Tocancipá. A los conjuntos de datos se les hizo una preparación de los datos tanto de entrenamiento como de pronóstico al igual que para los modelos con ANN, normalizándolos en un rango de -1 a 1 [24]. Posterior a este proceso los datos fueron partidos en tres conjuntos: entrenamiento, testeo y chequeo, aclarando que los periodos de datos tomados con los mismos que los empleados en el modelo ANN.

Se simularon en total 30 escenarios, variando la cantidad de épocas entre 2 y 20 y las capas de la red neuronal en un rango de 3 a 20 capas. Respecto a las reglas de inferencia, se usó un sistema tipo Sugeno, usando reglas de inferencia tipo producto (and) entre las entradas, funciones de pertenencia trapezoidal [26], campana [27] y triangular [21], y un método de defusificación tipo 'wtaver', el cual retorna la media ponderada a la salida del sistema difuso del ANFIS.

IV. RESULTADOS

Para comparar la precisión y exactitud de los dos modelos inteligentes trabajados en esta investigación se usaron seis estadísticos, usados en la mayoría de artículos observados y utilizados como métricas de evaluación de modelos simulados [7]. Estos son:

MAE (Mean Absolute Error): el error medio absoluto es una medida de precisión usada para evaluar pronósticos, en el se evalúa el valor absoluto promedio de la diferencia entre el dato real y el dato pronosticado, entre más pequeño sea el error o tienda a cero, será más preciso el pronóstico [28].

MSE (Mean Squared Error): el error medio cuadrático es una medida de precisión usada para evaluar pronósticos, en la que se evalúa el desempeño del valor promedio de la diferencia entre el dato real y el dato pronosticado al cuadrado [29].

MAPE (Mean Absolute Percentage Error): el error porcentual absoluto de la media permite analizar la exactitud del modelo en términos porcentuales, teniendo en cuenta el valor absoluto de la relación entre el dato real y el dato pronosticado. La escala para evaluar la exactitud del modelo usando el MAPE, determina que un pronóstico muy exacto es el que tiene un valor menor o igual al 10%, un buen pronóstico tiene un valor entre un rango entre el 11% al 20%, un pronóstico razonable entre el 21% al 50% y un pronóstico inadecuado mayor al 50% [30].

RMSE (Root Mean Squared Error): la raíz cuadrada del error cuadrático medio es una medida de precisión usada para evaluar pronósticos, en ella se evalúa el valor de la raíz del promedio cuadrático de la diferencia entre el dato real y el dato pronosticado, entre más pequeño sea el error o tienda a cero, será catalogado como el pronóstico más preciso [28]

CC (Correlation Coefficient): el coeficiente de correlación de Pearson determina la relación del dato real y el dato pronosticado, en cuanto a la covarianza con las desviaciones de los dos tipos de datos, a diferencia del CCC, el Coeficiente de correlación ignora componentes de exactitud [31].

CCC (Concordance Correlation Coefficient): el coeficiente de correlación de concordancia indica la relación entre la precisión y exactitud del modelo, este mide el grado en que la covarianza del dato real y el dato pronosticado se acercan a la recta de 45° del modelo, este factor se evalúa con valores entre de 0-1 [32].

A. ANN

Se eligieron 10 de los 85 escenarios con función de propagación Sigmoidal-Sigmoidal, y 10 de los 83 escenarios con función de propagación Sigmoidal-Lineal con mayor CCC. Luego de establecer los 10 Escenarios con mayor precisión y exactitud se compararon cada uno entre sí, ponderando aquellos que cumplieran con la mayor cantidad de resultados favorables en los criterios estadísticos a evaluar con el MAE, MAPE, MSE, RMSE y CC para calcular la precisión y exactitud de los modelos, excluyendo el CCC, debido a que este criterio ya había sido utilizado como criterio evaluador anteriormente. A continuación en la tabla 1 se observan los 3 mejores escenarios de cada uno de las funciones de propagación.

TABLA 1. MEJORES ESCENARIOS MODELO ANN

	<i>Criterio</i>	<i>CCC</i>	<i>CC</i>	<i>MAE</i>	<i>MAPE</i>	<i>MSE</i>	<i>RMSE</i>
Sig-Sig	Esc 3	0.863	0.903	2.065	13.425	12.287	3.505
	Esc 4	0.866	0.903	2.060	13.470	12.230	3.497
	Esc 9	0.877	0.921	1.900	11.959	10.178	3.190
Sig-Lin	Esc 2	0.872	0.913	1.946	12.519	10.975	3.312
	Esc 3	0.859	0.910	2.043	12.700	11.651	3.413
	Esc 6	0.873	0.910	1.983	13.009	11.332	3.366

Luego de analizar todos los criterios estadísticos en los mejores escenarios elegidos, el mejor escenario de función de propagación Sigmoidal-Sigmoidal es el número 9, debido a que cumple con la mayor cantidad de resultados favorables en los estadísticos: menor MAE (1,90), menor MAPE (11,9%), menor MSE (10,2), menor RMSE (3,2), mayor CC (0,92) y mayor CCC(0,88). Tomando como referencia el resultado obtenido con el MAPE se puede decir que el pronóstico arrojado es un buen pronóstico ya que se encuentra entre un rango entre el 11% y el 20%, si fuese menos del 10% sería muy exacto [30]. Sin embargo, el CC nos indica que el modelo presenta el 92% de precisión del pronóstico en cuanto a la relación de los datos reales con los datos simulados. En este escenario se implementaron 20 capas y 25 neuronas, lo cual indica que al implementar una ANN con distintos números de neuronas en un modelo de tipo de percepción multicapa, es posible obtener un pronóstico más adecuado [17].

En cuanto a la función de propagación Sigmoidal-Lineal se eligió escenario 2 como el de mejor pronóstico, debido a que

cumple con la mayor cantidad de resultados favorables en los criterios estadísticos evaluados: menor MAE (1,94), el menor MAPE (12,5%), menor MSE (10,97), menor RMSE (3,1) y el mayor CC 0,914. En este escenario se implementaron 2 capas y 50 neuronas, lo cual indica una variación fuerte en el número de capas respecto a la configuración Sigmoidal-Sigmoidal donde el mejor escenario tiene un número de 20 capas. Respecto al número de neuronas las dos configuraciones arrojaron el mejor pronóstico con una cantidad considerable de ellas en cada capa.

A continuación se observa la fig. 5 donde se muestra la comparación entre los datos reales con el mejor escenario simulado bajo la configuración de la función de propagación Sigmoidal-Lineal. En este hidrograma se muestran los caudales en un periodo de 6 meses del año 2011, tiempo donde se dieron varias inundaciones en la región. Se observa que el modelo ANN se acerca bastante a los datos de caudal reales.

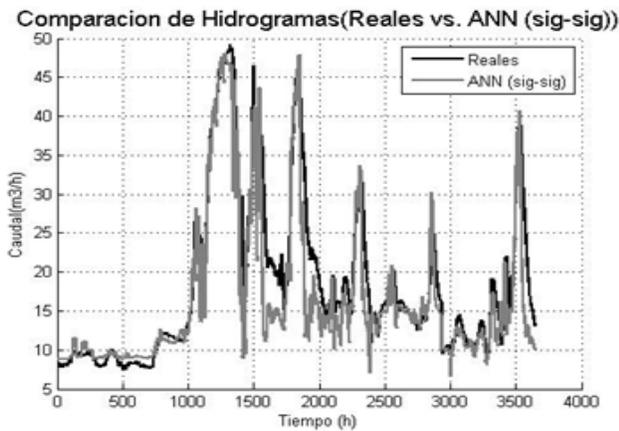


Figura 5. Comparación hidrogramas Reales vs ANN (sig-sig)

B. ANFIS

De igual manera que las ANN, se eligieron los 3 escenarios con el sistema Neuro-difuso.

TABLA 2. DESCRIPCIÓN MEJORES 3 ESCENARIOS ANFIS

Criterio	CC	CC	MAE	MAPE	MSE	RMSE
Esc 9	0.845	0.905	2.005	12.364	11.975	3.516
Esc 5	0.843	0.902	2.028	12.650	12.045	3.556
Esc 8	0.843	0.902	2.030	12.691	12.061	3.562

Luego de observar los anteriores resultados se eligió el escenario 9 como el de mejor pronóstico, ya que cumple con la mayor cantidad resultados favorables de criterios estadísticos evaluados: menor MAE (2,00), el menor MAPE (12,3%), menor MSE (11,97), menor RMSE (3,51), mayor CC (0,905) y el mayor CCC (0,845). Este escenario se simuló con función de pertenencia trapezoidal, con 15 capas y 3 épocas. La comparación con el mejor de los escenarios se muestra en

la fig. 6, donde se muestran los caudales reales registrados en el tiempo en un periodo de tiempo de 6 meses en el año 2011, donde ocurrieron las mayores inundaciones en la zona observada. Estos caudales se comparan con los resultados entregados por el ANFIS, los cuales reproducen de manera muy cercana a los datos reales.

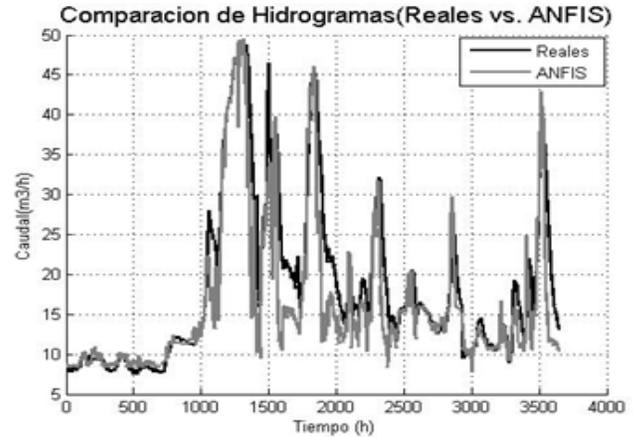


Fig. 6. Comparación hidrogramas Reales vs ANFIS

C. ANN vs. ANFIS

Los resultados de los tres mejores modelos desarrollados en esta investigación (ANN configuración sigmoidal-sigmoidal y ANN configuración sigmoidal-lineal y ANFIS) se encuentran en la tabla 3.

TABLA 3. MEJORES RESULTADOS DE MODELOS ANN Y ANFIS

Modelo	CCC	CC	MAE	MAPE	MSE	RMSE
ANN Sig-Sig	0.877	0.921	1.900	11.95	10.17	3.190
ANN Sig-Lin	0.872	0.913	1.946	12.51	10.97	3.312
ANFIS	0.845	0.905	2.005	12.36	11.97	3.516

Los resultados obtenidos sugieren que los modelos indican resultados de pronósticos muy similares mostrando en orden de efectividad, primero el ANN sigmoidal – sigmoidal, segundo al ANN sigmoidal-lineal, y por último el modelo ANFIS. Para observar mejor el comportamiento de los datos y el por qué de la elección del CCC, es necesario graficar los datos reales vs. datos simulados comparándolos con la recta de la ecuación $y=x$, procedimiento realizado en varias investigaciones de comparación de modelos [7][12][20][33][34].

En la fig. 7 se comparan todos los modelos utilizados en la investigación y la relación de los datos reales con los simulados. En primer lugar con el modelo ANFIS (a), ANN configuración sigmoidal-sigmoidal (b) y ANN configuración sigmoidal-lineal (c). En ésta se observa que existen datos por encima y debajo de la línea de referencia, siendo mayor la cantidad por debajo.

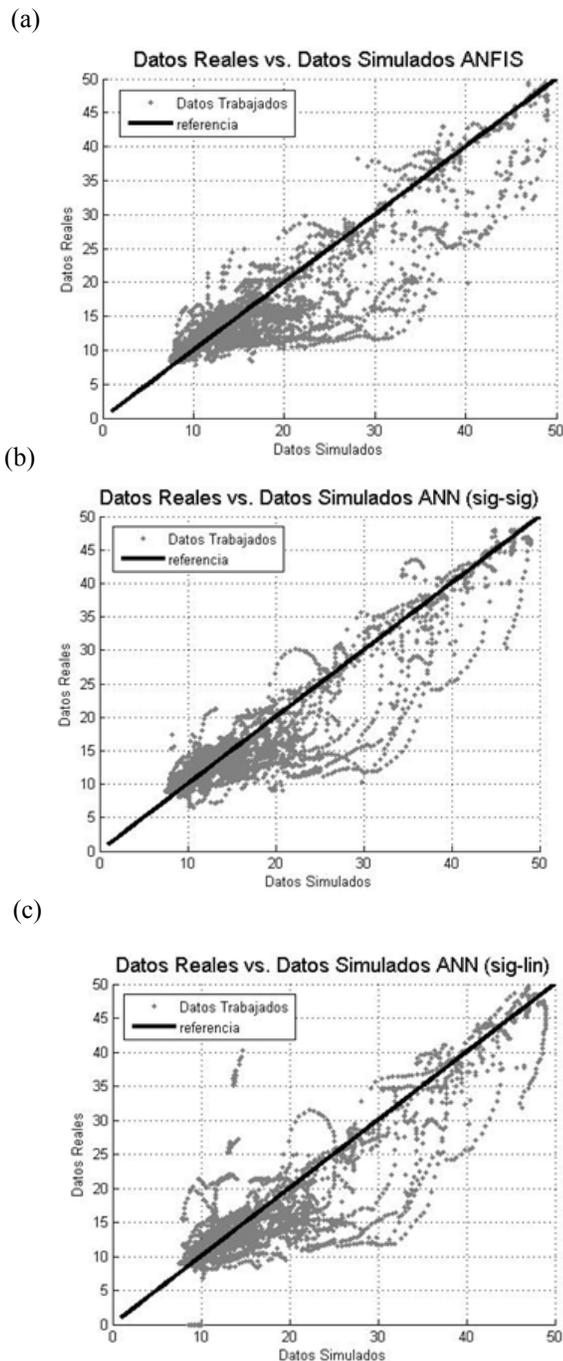


Fig. 7. Comparación datos reales y simulados

Al observar los valores mayores de caudal se puede observar que los modelos inteligentes son capaces de pronosticarlos, y aunque la dispersión de datos es alta los errores dados en los modelos indican que entre el dato pronosticado y el real el error es inferior al 10%.

V. CONCLUSIONES

La implementación de modelos basados en técnicas de Inteligencia Artificial con los cuales se emuló el comportamiento de un tramo del río Bogotá basado en

históricos de caudal, arrojaron resultados estadísticos que demuestran pronósticos muy cercanos a los datos reales y la efectividad de las técnicas (MAPE entre 12 y 12.5, que indica un buen pronóstico; CC entre 0.90 y 0.92, lo que indica una buena relación entre los datos reales y los simulados; CCC entre 0.84 y 0.87 que demuestra precisión y exactitud en el pronóstico.

Se logró observar que sistemas de Inteligencia Artificial (o de caja negra) como las ANN pueden ser iguales o mejores que los modelos hidrológicos comunes en ríos con comportamientos muy variantes, dados por el relieve de un país como Colombia, y que la aplicación de este tipo de soluciones puede ser una buena alternativa para la generación de sistemas de alertas tempranas, a través de la integración de Sistemas de Información Geográfica que permitan visualizar la información pronosticada en estos sistemas.

Se recomienda continuar la investigación, para la construcción de alertas tempranas en los municipios con ayuda de Sistemas de Información Geográfica (SIG) y buscar otras zonas del país donde se pueda adaptar este proyecto piloto para la preparación de la población ante riesgos de inundación.

Además, se recomienda la implementación del uso de técnicas artificiales inteligentes enfocadas al pronóstico de inundaciones en países de relieves tan diversos como los del continente Suramericano, donde no existen muchas investigaciones con el uso de elementos presentados en esta investigación; para ello es necesario continuar validando estos métodos con la integración de más variables no solo hidrometeorológicas, sino también sociales, logísticas, económicas y culturales, lo que podría generar mayores desarrollos tecnológicos para no solo la creación de alertas tempranas, sino para la prevención de desastres naturales de distinto tipo, empleando la programación por código, ya que las cajas de herramientas de Matlab son un poco limitadas.

Por otro lado, surge la posibilidad de estudiar este problema con otros tipos de técnicas pertenecientes al mundo del Soft Computing como los algoritmos genéticos y la Lógica difusa Intuitiva y su comparación con modelos matemáticos y/o hidráulicos.

AGRADECIMIENTOS

Se brindan agradecimientos especiales a las entidades Corporación Autónoma Regional de Cundinamarca y el Instituto Geográfico Agustín Codazzi, quienes entregaron la información y su colaboración para el desarrollo de esta investigación. De igual manera, al Semillero de Sistemas de Información Geográfica (principalmente a Gina Corredor, Laura Vargas y Catalina Mesa) del grupo de Investigación Human Centered Design de la Facultad de Ingeniería de la Universidad de La Sabana, quienes apoyaron con los estudios hidrográficos.

REFERENCIAS

- [1] González Valencia, J. E. (2006). Propuesta metodológica basada en un análisis multicriterio para la identificación de zonas de amenaza por deslizamientos e inundaciones, 5(8), 59–70.

- [2] Corporación Autónoma Regional de Cundinamarca – C.A.R. (2011). Plan de acción para la atención de la emergencia y la mitigación de sus efectos.
- [3] Pandey, G. R., & Nguyen, V. (1999). A comparative study of regression based methods in regional flood frequency analysis. *Journal of Hydrology*, 225(1–2), 92-101.
- [4] Weisberg, Sanford, (1947). Applied linear regression Third Edition. Wiley Series in Probability and Statistics
- [5] Huffman, W. S., (2001). Geographic information systems, expert systems and neural networks: disaster planning, mitigation and recovery. in: River Basin Management. International Conference. pp. 311-324.
- [6] Kia, M. B., Pirasteh, S., Pradhan, B., Mahmud, A. R., Sulaiman, W. N. A., & Moradi, A. (2011). An artificial neural network model for flood simulation using GIS: Johor River Basin, Malaysia. *Environmental Earth Sciences*, 67(1), 251–264.
- [7] Dawson, C. W., Abrahart, R. J., Shamseldin, a. Y., & Wilby, R. L. (2006). Flood estimation at ungauged sites using artificial neural networks. *Journal of Hydrology*, 319(1-4)
- [8] Aqil, Muhammad, Ichiro Kita, Akira Yano, Soichi Nishiyama. (2007). Analysis and prediction of flow from local source in a river basin using a Neuro-fuzzy modeling tool. *Journal of Environmental Management*, Volume 85, Issue 1, Pages 215-223.
- [9] Seckin, N., Cobaner, M., Yurtal, R., & Haktanir, T. (2013). Comparison of artificial neural network methods with L-moments for estimating flood flow at ungauged sites: The case of east mediterranean river basin, turkey. *Water Resources Management*, 27(7), 2103-2124
- [10] Kalteh, A. M. (2013). Monthly river flow forecasting using artificial neural network and support vector regression models coupled with wavelet transform. *Computers & Geosciences*, 54(0), 1-8.
- [11] Goyal, M. K., Bharti, B., Quilty, J., Adamowski, J., & Pandey, A. (2014). Modeling of daily pan evaporation in sub tropical climates using ANN, LS-SVR, fuzzy logic, and ANFIS. *Expert Systems with Applications*, 41(11), 5267-5276.
- [12] Wang, Y., Wang, H., Lei, X., Jiang, Y., & Song, X. (2011). Flood simulation using parallel genetic algorithm integrated wavelet neural networks. *Neurocomputing*, 74(17), 2734–2744. doi:10.1016/j.neucom.2011.03.018
- [13] Nastos, P. T., Paliatso, A. G., Koukouletsos, K. V., Larissi, I. K., & Moustris, K. P. (2014). Artificial neural networks modeling for forecasting the maximum daily total precipitation at athens, greece. *Atmospheric Research*, 144(0), 141-150
- [14] Yilmaz, A. G., Imteaz, M. A., & Jenkins, G. (2011). Catchment flow estimation using artificial neural networks in the mountainous euphrates basin. *Journal of Hydrology*, 410(1–2), 134-140.
- [15] Zou, H. F., Xia, G. P., Yang, F. T., & Wang, H. Y. (2007). An investigation and comparison of artificial neural network and time series models for Chinese food grain price forecasting. *Neurocomputing*, 70(16-18), 2913–2923.
- [16] Komatsu, Misako, Jun Namikawa, Zenas C. Chao, Yasuo Nagasaka, Naotaka Fujii, Kiyohiko Nakamura & Jun Tani. (2014). An artificial neural network model for estimating the network structure underlying partially observed neuronal signals. *Neuroscience Research*, In Press, Corrected Proof.
- [17] Hamzaçebi, C., Akay, D., & Kutay, F. (2009). Comparison of direct and iterative artificial neural network forecast approaches in multi-periodic time series forecasting. *Expert Systems with Applications*, 36(2), 3839–3844. doi:10.1016/j.eswa.2008.02.042
- [18] Chau, K.W., Wu, C.L. Li, Y.S. (2005). Comparison of several flood forecasting models in Yangtze River. *Journal of Hydrologic Engineering*, ASCE, Vol. 10, No. 6, pp. 485-491
- [19] Hui Liu, Hong-qi Tian, Di-fu Pan, Yan-fei Li, (2012) Forecasting models for wind speed using wavelet, wavelet packet, time series and Artificial Neural Networks, *Applied Energy*, Volume 107, July 2013, Pages 191-208
- [20] Firat Mahmut & Güngör Mahmud. (2007). River flow estimation using adaptive neuro fuzzy inference system, *Mathematics and Computers in Simulation*, Volume 75, Issues 3–4, Pages 87-96
- [21] Talei, A., Chua, L. H. C., Quek, C., & Jansson, P.-E. (2013). Runoff forecasting using a Takagi–Sugeno neuro-fuzzy model with online learning. *Journal of Hydrology*, 488, 17–32.
- [22] Mulvaney, T. (1850) “On the use of self registering rain and flood gauges in making observations of the relation of rainfall and flood discharges in a given catchments”. *Trans. Instn civ. Engrs Ire* 4(2) 18.
- [23] Fattorelli, Sergio y Fernandez, Pedro. (2011). *Diseño Hidrológico*, Segunda edición. Ed. Biblioteca virtual Water Assessment & Advisoy Global Network. ISBN:978-987-05-2738-2
- [24] Matworks. (2013). *Neural Network Toolbox, User’s guide MATLAB*, version B.
- [25] Kourentzes, Nikolaos, Barrow, D evon k. & Crone Sven F. (2014). *Neural network ensemble operators for time series forecasting*. *Expert Systems with Applications*, Volume 41, Issue 9, Pages 4235-4244.
- [26] Karimi, S., Kisi, O., Shiri, J., & Makarynsky, O. (2013). Neuro-fuzzy and neural network techniques for forecasting sea level in Darwin Harbor, Australia. *Computers & Geosciences*, 52, 50–59.
- [27] Ghose, D. K., Panda, S. S., & Swain, P. C. (2013). Prediction and optimization of runoff via ANFIS and GA. *Alexandria Engineering Journal*, 52(2), 209–220.
- [28] Singhal, D., & Swarup, K. S. (2011). Electricity price forecasting using artificial neural networks. *International Journal of Electrical Power & Energy Systems*, 33(3), 550-555.
- [29] Gecynalda, S., da S. Gomes, Teresa B. Ludermir. (2013). Optimization of the weights and asymmetric activation function family of neural network for time series forecasting. *Expert Systems with Applications*, Volume 40, Issue 16, 15, Pages 6438-6446.
- [30] Lewis, C. (1982). *Industrial and business forecasting methods*. London: Butterworths.
- [31] Li, J., Feng, P., & Wei, Z. (2013). Incorporating the data of different watersheds to estimate the effects of land use change on flood peak and volume using multi-linear regression. *Mitigation and Adaptation Strategies for Global Change*, 18(8), 1183-1196. doi:http://dx.doi.org/10.1007/s11027-012-9416-
- [32] Lin, L. (2002). *Statistical Methods in Assessing Agreement: Models, Issues, and Tools*. *Journal of the American Statistical Association*.
- [33] Mahabir, C., Hicks, F., & Robinson, A. (2006). Neuro-fuzzy river ice breakup forecasting system, 46, 100–112.
- [34] Fantin-Cruz, I., Pedrollo, O., Castro, N. M. R., Girard, P., Zeilhofer, P., & Hamilton, S. K. (2011). Historical reconstruction of floodplain inundation in the Pantanal (Brazil) using neural networks. *Journal of Hydrology*, 399(3-4), 376–384.