



Ciencia Latina Revista Científica Multidisciplinar, Ciudad de México, México.
ISSN 2707-2207 / ISSN 2707-2215 (en línea), julio-agosto 2026,
Volumen 10, Número 4.

https://doi.org/10.37811/cl_rem.v10i4

**MODELO BASADO EN REDES LSTM PARA LA
PREDICCIÓN DE CAUDALES Y EL
FORTALECIMIENTO DE ALERTAS TEMPRANAS
DE INUNDACIONES EN LA CUENCA DEL RÍO
HACHA, AMAZONIA COLOMBIANAS**

**LSTM-BASED MODEL FOR STREAMFLOW PREDICTION
AND FLOOD EARLY WARNING STRENGTHENING IN THE
HACHA RIVER WATERSHED, COLOMBIAN AMAZON**

Olga Lucía Narváez Losada

Universidad Americana de Europa-UNADE, Colombia

Rosa Gabriela Camero Berrones

Tecnológico Nacional de México, México

Ermes Guarnizo Motta

Universidad Nacional Abierta y a Distancia UNAD, Colombia

Nichol Dayane Riveros Gaitan

Universidad Nacional Abierta y a Distancia UNAD, Colombia

Modelo basado en Redes LSTM para la Predicción de Caudales y el Fortalecimiento de Alertas Tempranas de Inundaciones en la cuenca del Río Hacha, Amazonia colombianas

Olga Lucía Narváez Losada¹

olunalo@gmail.com

<https://orcid.org/0000-0003-1091-3543>

Universidad Americana de Europa-UNADE
Colombia

Rosa Gabriela Camero Berrones

gabriela.camero@gmail.com

<https://orcid.org/0000-0003-4438-1645>

Universidad Americana de Europa-UNADE
Tecnológico Nacional de México, campus Cd.
Mader
México

Ermes Guarnizo Motta

ermes.guarnizo@unad.edu.co

<https://orcid.org/0000-0002-0017-4022>

Universidad Nacional Abierta y a Distancia
UNAD
Colombia

Nichol Dayane Riveros Gaitan

nichol.riveros@unad.edu.co

Universidad Nacional Abierta y a Distancia
UNAD
Colombia

RESUMEN

Las inundaciones representan uno de los riesgos naturales con mayor impacto en la Amazonia colombiana, agravado por la variabilidad hidrológica, las consecuencias del cambio climático y la escasa cobertura de redes de monitoreo hidrometeorológico. El propósito de esta investigación fue diseñar y validar un modelo de inteligencia artificial basado en redes neuronales recurrentes Long Short-Term Memory (LSTM) para así estimar los caudales en la cuenca del río Hacha, en Florencia, Caquetá, enfocado al fortalecimiento de los sistemas de alerta temprana y al apoyo de la gestión del riesgo de inundaciones. Se utilizó un enfoque cuantitativo, aplicado y predictivo, con un diseño no experimental, longitudinal y prospectivo. La metodología cubrió el análisis de variables hidrológicas e hidroclimáticas, las cuales se tomaron mediante el sistema de monitoreo IoT RioMonitor durante 6 meses (259.200 registros con frecuencia de 1 dato / min.), se eligieron las variables desde el análisis estadístico y la minería de datos, el entrenamiento y validación del modelo LSTM, y su comparación con modelos de referencia (Persistencia, Regresión lineal múltiple, ARIMA y Random Forest), todo a través de métricas como RMSE, MAE, R^2 , NSE y KGE, y su integración en un sistema de alerta temprana denominado Amazon Smart Hydrology Framework (ASHF). Los hallazgos demostraron que el modelo LSTM tuvo el mejor desempeño predictivo (RMSE = 8,72 m³/s; MAE = 6,25 m³/s; NSE = 0,89; R^2 = 0,91), con diferencias estadísticamente significativas frente a los demás modelos comparados ($p < 0,05$), apelando a las pruebas de Wilcoxon y Friedman. Por otra parte, el análisis SHAP logró identificar el nivel del agua y la variación temporal del nivel como las variables de mayor peso en las predicciones. Todo lo anterior, permite concluir que el modelo es una herramienta sólida para mejorar la capacidad de pronóstico hidrológico y apoyar la generación de alertas tempranas en la Amazonia colombiana.

Palabras clave: inteligencia artificial, redes neuronales LSTM, predicción de caudales, sistemas de alerta temprana, gestión del riesgo de inundaciones

¹ Autor principal

Correspondencia: olunalo@gmail.com

LSTM-Based Model for Streamflow Prediction and Flood Early Warning Strengthening in the Hacha River Watershed, Colombian Amazon

ABSTRACT

Floods rank among the most damaging natural hazards in the Colombian Amazon, intensified by hydrological variability, climate change effects, and sparse hydrometeorological monitoring networks. This study aimed to design and validate an artificial intelligence model based on Long Short-Term Memory (LSTM) recurrent neural networks for streamflow estimation in the Hacha River watershed (Florencia, Caquetá), with the purpose of strengthening early warning systems and supporting flood risk management. A quantitative, applied, and predictive approach was adopted under a non-experimental, longitudinal, and prospective design. The methodology encompassed hydrological and hydroclimatic variable analysis from the IoT-based RioMonitor system over six months (259,200 records at one-minute resolution), variable selection through statistical analysis and data mining, LSTM model training and validation, benchmark comparison against Persistence, Multiple Linear Regression, ARIMA, and Random Forest models using RMSE, MAE, R^2 , NSE, and KGE metrics, and integration into an early warning framework called the Amazon Smart Hydrology Framework (ASHF). Results showed that the LSTM model achieved the best predictive performance (RMSE = 8.72 m³/s; MAE = 6.25 m³/s; NSE = 0.89; R^2 = 0.91), with statistically significant differences compared to all benchmark models ($p < 0.05$) according to Wilcoxon and Friedman tests. SHAP analysis identified water level and temporal level variation as the most influential predictor variables. It is concluded that the proposed model provides a robust tool for improving hydrological forecasting and supporting early warning generation in the Colombian Amazon.

Keywords: artificial intelligence, LSTM neural networks, streamflow prediction, early warning systems, flood risk management

Artículo recibido 15 junio 2026

Aceptado para publicación: 22 julio 2026



INTRODUCCIÓN

La gestión del recurso hídrico afronta presiones crecientes a escala global. Los episodios de inundación se han multiplicado en las últimas décadas. La variabilidad climática natural y los efectos acumulados del calentamiento global los intensifican, y la exposición al riesgo de poblaciones ribereñas crece en paralelo (IPCC, 2023; WMO, 2024). La Organización Meteorológica Mundial afirma que las inundaciones dejan un 44 % las víctimas mortales por causa de desastres a nivel mundial. Ese dato tiende a crecer en regiones tropicales con infraestructura de monitoreo deficiente. La Amazonia colombiana concentra esta problemática de forma aguda, donde las precipitaciones superan 3500 mm anuales, mientras que la geomorfología del piedemonte caqueteño es compleja y la densidad de estaciones hidrometeorológicas es baja. Todo ello convierte las inundaciones en eventos repetitivos y difíciles de anticipar (IDEAM, 2023).

Florencia, capital del departamento del Caquetá, sufre de desbordamientos frecuentes del río Hacha. Dicho cauce pasa por la zona urbana en 64,5 km, desde la Cordillera Oriental hasta su confluencia con el río Orteguaza (Corpoamazonia, 2018). Se documentan episodios graves como la avalancha de agosto de 1962 con 122 víctimas mortales; una inundación sobre el río Hacha y la quebrada La Perdiz, en 1999, lo cual provocó 14 fallecidos y cerca de 14000 damnificados (Hortúa, 2004). A pesar de todo lo anterior, la cuenca carece de un sistema de predicción basado en inteligencia artificial y su monitoreo depende de estaciones convencionales con limitaciones de cobertura y oportunidad.

La predicción de caudales se abordó tradicionalmente con modelos hidrológicos determinísticos y estadísticos fundamentados en ecuaciones físicas o en relaciones empíricas del ciclo del agua (Beven, 2012). Estos enfoques, sin embargo, presentan restricciones para captar la dinámica fuertemente no lineal de las cuencas tropicales. Esa carencia impulsó metodologías basadas en inteligencia artificial, capaces de identificar patrones complejos en series temporales y de mejorar la precisión de las estimaciones hidrológicas (Mosavi et al., 2018; Shen, 2018).

Las redes neuronales Long Short-Term Memory (LSTM), una variante de las redes recurrentes diseñada para resolver el problema del desvanecimiento del gradiente y capturar dependencias temporales de largo plazo (Hochreiter y Schmidhuber, 1997), han demostrado un desempeño superior al de los modelos hidrológicos convencionales en la predicción de caudales. Kratzert et al. (2018) demostraron que un

modelo LSTM entrenado en 241 cuencas del dataset CAMELS sobrepasó al modelo SAC-SMA acoplado con el esquema de nieve Snow-17. Sin embargo, otras investigaciones posteriores confirmaron estos hallazgos y ampliaron su alcance: Kratzert et al. (2019) aplicaron el Entity-Aware LSTM para aprendizaje multi-cuenca, y Nearing et al. (2021) demostraron que los modelos basados en datos pueden superar a los modelos de procesos físicos incluso en cuencas donde la hidrología convencional ha tenido excelentes resultados.

Al mismo tiempo, la integración de sensores de bajo costo con plataformas de transmisión inalámbrica hace posible adquirir datos hidrológicos con alta resolución temporal en cuencas donde la infraestructura de monitoreo convencional es limitada o inexistente. Sistemas basados en sensores ultrasónicos HC-SR04 con comunicación LoRaWAN han sido validados en campo con errores inferiores a 1 cm para la medición del nivel del agua en ríos (Sit et al., 2020). Esta convergencia entre IoT e inteligencia artificial ha favorecido el desarrollo de sistemas inteligentes de alerta temprana que combinan capacidades de adquisición y procesamiento predictivo en una arquitectura única.

A pesar de estos avances, la mayor parte de la evidencia empírica sobre modelos LSTM procede de cuencas densamente instrumentadas de Norteamérica, Europa y Asia. Las cuencas amazónicas tropicales, con regímenes de alta variabilidad, series históricas cortas y respuestas hidrológicas rápidas, permanecen prácticamente inexploradas desde la óptica de la hidrología basada en datos (Shen et al., 2024). Tres aspectos han recibido una atención insuficiente en la literatura: la interpretabilidad de las predicciones, la cuantificación de la incertidumbre asociada y la integración con sistemas de monitoreo IoT en tiempo real.

La interpretabilidad de los modelos de aprendizaje profundo ha cobrado relevancia en la hidrología reciente. Lundberg y Lee (2017) propusieron los valores SHAP (SHapley Additive exPlanations) como mecanismo unificado de atribución de importancia, y estudios posteriores han aplicado esta técnica a modelos LSTM hidrológicos con resultados congruentes con los procesos físicos subyacentes (Huang y Zhang, 2024; Demiray et al., 2024). Por otra parte, la cuantificación de la incertidumbre predictiva mediante remuestreo Bootstrap (Efron y Tibshirani, 1993) y Monte Carlo Dropout se ha consolidado como requisito para que los pronósticos puedan respaldar decisiones de gestión del riesgo (Tiwari y Chatterjee, 2010).

Una búsqueda exhaustiva en Google Scholar, Scopus y SciELO no arrojó estudios que articulen, dentro de un mismo marco metodológico, monitoreo IoT, modelación LSTM, análisis de interpretabilidad SHAP y cuantificación de incertidumbre en cuencas de la Amazonia colombiana. En el caso concreto del río Hacha, la literatura disponible se reduce a estudios geomorfológicos descriptivos (Hortúa, 2004) y al Plan de Ordenación y Manejo de la Cuenca elaborado por Corpoamazonia (2018), sin componente alguno de aprendizaje automático.

Frente a este vacío, la presente investigación plantea como hipótesis que un modelo predictivo basado en redes LSTM, alimentado con datos de un sistema de monitoreo IoT de bajo costo y complementado con análisis SHAP y remuestreo Bootstrap, mejora de forma estadísticamente significativa la predicción de caudales en la cuenca del río Hacha respecto a modelos de referencia, y produce información suficiente para sustentar un sistema de alerta temprana operativo. El objetivo general busca desarrollar y validar este modelo, incluyendo tecnologías IoT y técnicas de inteligencia artificial, de tal manera que fortalezcan los sistemas de alerta temprana de inundaciones y apoyen la toma de decisiones para la gestión del riesgo en el municipio de Florencia (Caquetá).

METODOLOGÍA

Enfoque y diseño de investigación

Esta investigación se desarrolló bajo un enfoque cuantitativo, aplicado y predictivo, orientado al análisis de series temporales hidrológicas y al desarrollo de un modelo de aprendizaje profundo enfocado en la estimación de caudales. El diseño se enmarca en una investigación no experimental, longitudinal y prospectiva, utilizando registros históricos y datos de monitoreo en tiempo real de las condiciones naturales (Creswell y Creswell, 2018). La validación siguió un esquema cronológico que garantizó la no fuga de información entre los subconjuntos de entrenamiento, validación y prueba, lo que permitiría replicar las condiciones que enfrentaría un sistema de alerta en operación real.

Área de estudio

La investigación se desarrolló en la cuenca del río Hacha, municipio de Florencia, departamento del Caquetá, Colombia (coordenadas aproximadas: 1°37'N, 75°37'W). La cuenca forma parte de la región Amazónica colombiana y presenta un régimen hidrológico tropical con precipitaciones superiores a 3 500 mm anuales, alta humedad relativa, temperaturas medias de 25 °C y una marcada variabilidad

estacional condicionada por procesos atmosféricos regionales (IDEAM, 2023). Las crecientes súbitas afectan de manera recurrente a las comunidades urbanas y rurales asentadas en las zonas de influencia de la cuenca. Esa recurrencia justificó su selección como caso de estudio.

Figura 1. Área de estudio: cuenca del río Hacha, municipio de Florencia, departamento del Caquetá, Colombia



Nota. Elaboración propia.

Población, unidad de análisis y muestra

La población serían todos los registros hidrológicos e hidroclimáticos de la cuenca del río Hacha en el período de estudio. La unidad de análisis correspondería al registro temporal que cubre las variables medidas por el sistema RioMonitor. Todos estos datos estuvieron conformados por 259200 registros obtenidos entre el 1 de marzo de 2025 y el 27 de agosto de 2025, con resolución de 1 dato / min. No se usó muestreo probabilístico pues se trabajó con la totalidad de los registros que cumplieron los criterios de calidad o inclusión (ausencia de errores instrumentales, consistencia temporal, valores dentro de rangos hidrológicamente plausibles), lo que termina en un censo de datos hidrológicos, donde se excluyeron registros con datos faltantes superiores al umbral establecido, errores instrumentales evidentes y valores atípicos incorregibles.

Instrumentos de recolección

Las fuentes de información fueron registros hidrológicos históricos del IDEAM y de Corpoamazonia, y datos en tiempo real del sistema RioMonitor. Este último incluye sensores ultrasónicos HC-SR04 (rango

de medición 2–400 cm, precisión de hasta 3 mm) para la medición del nivel del agua, módulos de comunicación inalámbrica HC-12 (433 MHz) para la transmisión remota de datos, controladores Arduino UNO como nodos de adquisición y Arduino Mega 2560 como estación receptora y una plataforma de almacenamiento y visualización desarrollada con HTML, JavaScript y Web Serial API. La implementación a nivel de sistemas informáticos de los modelos se realizó en Python 3.10 con las bibliotecas TensorFlow 2.12, Keras, Scikit-learn 1.3, Pandas, NumPy, SHAP 0.42 y Matplotlib.

Procedimiento metodológico

El proceso se estructuró en seis fases secuenciales. La primera consistió en la integración de registros históricos con datos provenientes de los sensores IoT del sistema RioMonitor. La segunda abarcó el control de calidad, incluyendo tratamiento de datos faltantes mediante interpolación lineal, eliminación de valores atípicos identificados con el criterio del rango intercuartílico ($IQR \times 1,5$), normalización Min-Max en el intervalo $[0, 1]$ y construcción de secuencias temporales con una ventana de 60 observaciones. La tercera fase comprendió el diseño de la arquitectura LSTM y la optimización de hiperparámetros, cuya configuración final se detalla en la Tabla 1. En la cuarta fase se llevaron a cabo el entrenamiento (70 % de los datos), la validación (15 %) y la prueba (15 %) del modelo mediante división estrictamente cronológica. La quinta fase se centró en la comparación del modelo LSTM con cuatro algoritmos de referencia. La sexta fase aplicó el método SHAP para interpretar la contribución de cada variable de entrada a las predicciones del caudal.

Tabla 1. Arquitectura del modelo LSTM y parámetros de entrenamiento

Componente	Configuración
Tipo de modelo	Red neuronal recurrente LSTM
Variables de entrada	Nivel del río, crecimiento, caudal antecedente, variables derivadas
Ventana temporal (timesteps)	60 observaciones
Capa LSTM 1	128 neuronas, return_sequences = True
Regularización	Dropout = 0,20
Capa LSTM 2	64 neuronas, return_sequences = False
Capa densa (Dense)	32 neuronas, activación ReLU
Capa de salida	1 neurona, activación lineal
Función de pérdida	Mean Squared Error (MSE)
Optimizador	Adam (tasa de aprendizaje = 0,001)
Tamaño de lote (Batch Size)	32 observaciones
Número de épocas	100

Nota. Elaboración propia.

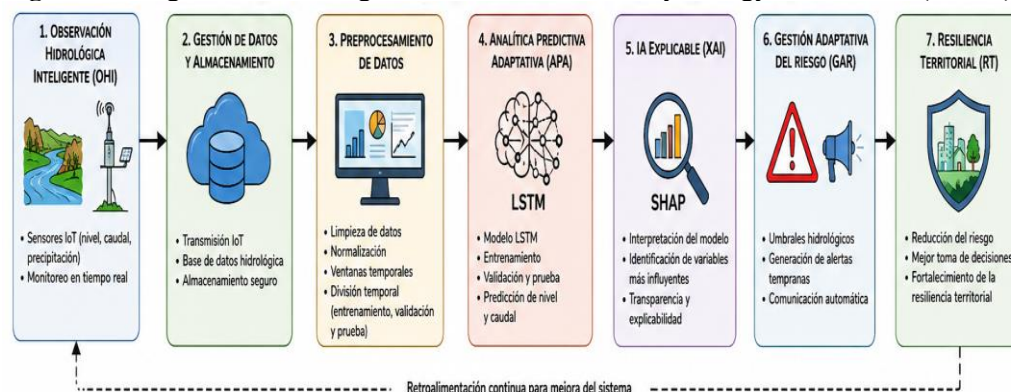
Métricas de evaluación y pruebas estadísticas

El desempeño predictivo se midió a través de cinco indicadores muy bien aceptados en la literatura hidrológica: la Raíz del Error Cuadrático Medio (RMSE), el Error Absoluto Medio (MAE), el coeficiente de determinación (R^2), la eficiencia de Nash-Sutcliffe (NSE) y la eficiencia Kling-Gupta (KGE). El NSE, propuesto por Nash y Sutcliffe (1970), mide qué proporción de la variabilidad observada es explicada por el modelo, donde valores superiores a 0,75 se consideran indicativos de un muy buen desempeño. El KGE, desarrollado por Gupta et al. (2009), toma el error en correlación, la variabilidad y el sesgo, y ofrece una valoración más balanceada que el NSE para contextos hidrológicos específicos. La significancia de las diferencias entre modelos se contrastó con la prueba de rangos con signo de Wilcoxon para comparaciones pareadas y con la prueba de Friedman para la comparación simultánea de los cinco modelos evaluados.

Amazon Smart Hydrology Framework (ASHF)

Para integrar el monitoreo hidrológico en tiempo real con las técnicas de inteligencia artificial en una arquitectura coherente, se desarrolló el Amazon Smart Hydrology Framework (ASHF), un marco metodológico orientado al fortalecimiento de los Sistemas de Alerta Temprana en cuencas de la Amazonia colombiana. El ASHF vincula seis módulos funcionales interrelacionados: adquisición inteligente de datos mediante sensores IoT; control de calidad y preprocesamiento de series temporales; modelado predictivo mediante redes LSTM; interpretabilidad de predicciones con SHAP; evaluación estadística del desempeño y cuantificación de incertidumbre, y; generación automática de alertas hidrológicas con semaforización de cuatro niveles (Kratzert et al., 2019; Lundberg y Lee, 2017; Efron y Tibshirani, 1993).

Figura 2. Diagrama metodológico del Amazon Smart Hydrology Framework (ASHF)



Nota. Elaboración propia.

El módulo de adquisición inteligente captura la distancia entre el sensor y la superficie del agua mediante pulsos ultrasónicos, convierte esa distancia en nivel hidrométrico y estima el caudal a partir de una curva de calibración específica de la sección transversal monitoreada. El módulo de control de calidad utiliza filtros de rango, detección de valores atípicos y sincronización temporal para así poder almacenar los registros en la base de datos. Mientras que el módulo de modelado predictivo recibe las secuencias temporales preprocesadas y arroja estimaciones del caudal a horizontes de 1 hora. El módulo de interpretabilidad calcula valores SHAP para cada predicción, lo que deja identificar qué variables condicionaron el pronóstico en cada instante. El módulo de evaluación estadística aplica las métricas de desempeño y las pruebas de significancia descritas en la sección 2.6, así como el remuestreo Bootstrap con 1 000 réplicas para estimar intervalos de confianza del 95 %. Por último, el módulo de generación de alertas compara el caudal predicho con umbrales hidrológicos predefinidos y asigna automáticamente el nivel de semaforización correspondiente.

El marco teórico que sustenta la integración de estos módulos se apoya en la propuesta teórica de la Hidrología Inteligente Amazónica (HIA), formulada como parte de la investigación doctoral de la cual se deriva este artículo. La HIA identifica cuatro dimensiones interdependientes en la producción de conocimiento hidrológico para cuencas tropicales: observación hidrológica inteligente, analítica predictiva adaptativa, inteligencia artificial explicable y gestión adaptativa del riesgo. Las seis proposiciones derivadas de este marco fueron verificadas empíricamente a través de las hipótesis de investigación contrastadas en la sección de resultados.

Limitaciones del diseño metodológico

El diseño presenta restricciones que deben declararse antes de los resultados. El modelo se entrenó y validó en una sola cuenca, lo que no permite una generalización de los hallazgos como tal vez que quisiera. La serie experimental abarca 6 meses, lo que cubre la variabilidad intraanual pero no interanual. La cobertura espacial del monitoreo IoT se ubicó en un punto de medición sobre el río Hacha, mientras que el sensor HC-SR04, con un costo muy bajo y una muy buena resolución para medición del nivel, es sensible a interferencias de temperatura ambiente y vegetación flotante, lo que es muy común en la Amazonia. Todos estos factores fueron mitigados con filtros de rango y calibración periódica, aunque no fueron eliminados en su totalidad.

Consideraciones éticas

La investigación usó solo información hidrológica y ambiental provenientes de fuentes institucionales y de sensores instalados con fines científicos, sin afectar personas y anonimizándolos. El estudio se clasifica como investigación de riesgo mínimo conforme a la Resolución 8430 de 1993 del Ministerio de Salud de Colombia, y en todo momento se respetaron los principios de integridad científica, transparencia, reproducibilidad y uso responsable de los datos.

RESULTADOS

Sistema de monitoreo RioMonitor

El sistema RioMonitor operó de forma continua durante el período comprendido entre el 1 de marzo y el 27 de agosto de 2025. Acumuló 259 200 observaciones con una frecuencia de un registro por minuto. La disponibilidad operativa superó el 95 %, sin vacíos mayores a 24 horas consecutivas, lo cual es valioso si se consideran las condiciones ambientales del piedemonte amazónico como humedad relativa por encima del 85 %, lluvias frecuentes y presencia de material flotante en el cauce durante las crecientes.

El caudal máximo registrado alcanzó 111,38 m³/s, lo que está relacionado con un evento hidrológico más severo del período de observación y una variación del nivel de 118,5 cm respecto al valor de referencia, donde el caudal promedio fue de 34,51 m³/s y el mínimo en 20,00 m³/s. La amplitud entre los valores extremos (91,38 m³/s) confirma la alta variabilidad del régimen fluvial, característica

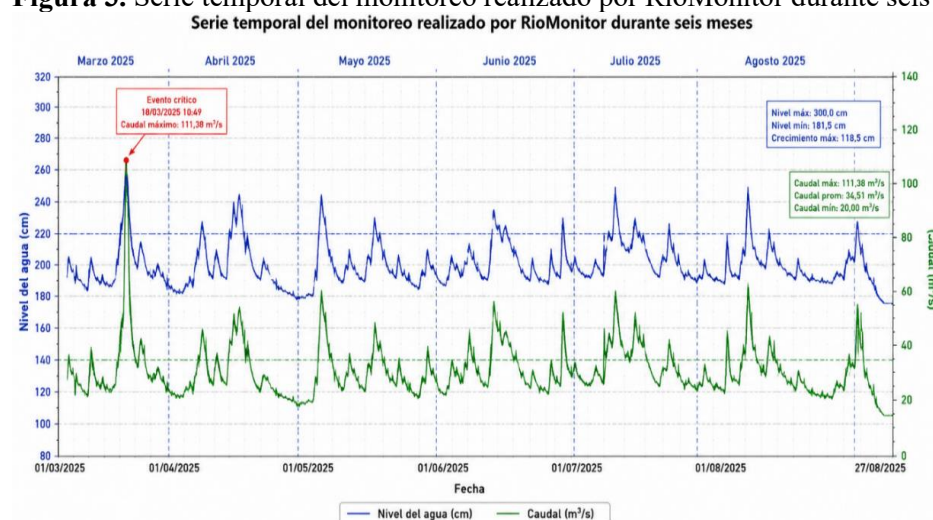
coherente con los patrones hidrológicos documentados por Corpoamazonia (2018) para la cuenca del río Hacha.

Tabla 2. Caracterización general de la base de datos experimental

Variable	Resultado
Período de monitoreo	1 marzo – 27 agosto 2025
Duración	6 meses
Total de registros	259 200
Frecuencia de muestreo	1 registro/min
Caudal máximo	111,38 m ³ /s
Caudal promedio	34,51 m ³ /s
Caudal mínimo	20,00 m ³ /s
Crecimiento máximo	118,5 cm

Nota. Elaboración propia a partir de la base experimental RioMonitor.

Figura 3. Serie temporal del monitoreo realizado por RioMonitor durante seis meses.



Nota. Elaboración propia.

Selección de variables

El análisis de correlación (Pearson y Spearman) y la evaluación de importancia mediante SHAP permitieron identificar las variables con mayor capacidad explicativa sobre el caudal estimado. Los datos históricos del caudal tuvieron un valor medio absoluto más alto (0,41) en el valor SHAP, seguido por el nivel del agua (0,31) y la variación temporal del nivel (0,16), mientras que las variables derivadas, que surgieron a partir de las diferencias temporales y de los promedios móviles, aportaron información complementaria que mejoró la capacidad del modelo para capturar los eventos de creciente, y así la tendencia temporal contribuyó con un SHAP de 0,12, lo que indica que es una variable de apoyo para la captura de patrones estacionales.

Desempeño del modelo LSTM

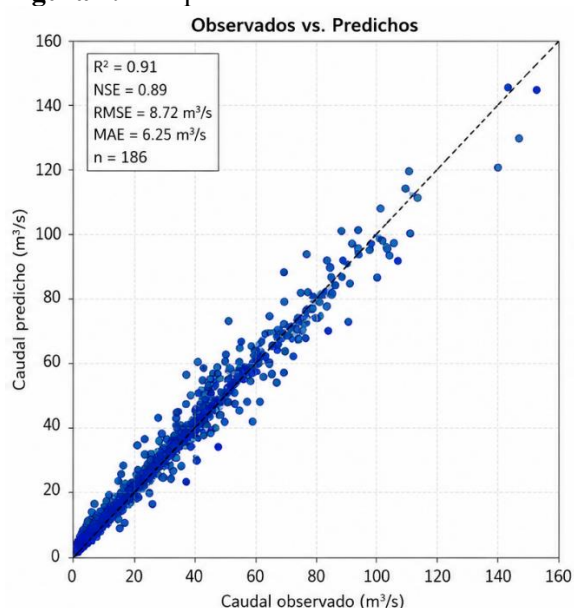
La base de datos se dividió en tres subconjuntos en orden cronológico: entrenamiento (70 %), validación (15 %) y prueba (15 %), por lo que el modelo mantuvo un desempeño consistente durante las tres fases. Sin embargo, el conjunto de entrenamiento obtuvo el error más bajo ($RMSE = 7,94 \text{ m}^3/\text{s}$), en comparación con los mayores índices de desempeño ($NSE = 0,92$; $KGE = 0,93$; $R^2 = 0,93$). De allí que, durante la validación, las métricas experimentaron una variación mínima ($RMSE = 8,38 \text{ m}^3/\text{s}$; $R^2 = 0,92$), que evidencia la ausencia de sobreajuste significativo. En el conjunto de prueba, el modelo alcanzó un $RMSE$ de $8,72 \text{ m}^3/\text{s}$, un MAE de $6,25 \text{ m}^3/\text{s}$, un NSE de $0,89$, un KGE de $0,91$ y un R^2 de $0,91$.

Tabla 3. Resultados de la validación temporal del modelo LSTM

Conjunto	RMSE (m^3/s)	MAE (m^3/s)	NSE	KGE	R^2
Entrenamiento	7,94	5,68	0,92	0,93	0,93
Validación	8,38	5,97	0,90	0,91	0,92
Prueba	8,72	6,25	0,89	0,91	0,91

Nota. Elaboración propia.

Figura 4. Comparación entre caudales observados y predichos por el modelo LSTM.



Nota. Elaboración propia.

Comparación con modelos de referencia

El modelo LSTM presentó el mejor desempeño global entre los cinco algoritmos evaluados. Alcanzó el menor error de predicción y los mayores valores de ajuste en los tres indicadores compuestos (R^2 , NSE y KGE), superando tanto a los modelos estadísticos clásicos (Persistencia, Regresión lineal múltiple y

ARIMA) como al algoritmo de aprendizaje automático Random Forest. La brecha de desempeño fue particularmente marcada frente a los modelos de Persistencia y Regresión lineal, mientras que Random Forest ofreció el desempeño más cercano al LSTM entre los modelos de referencia.

Tabla 4. Comparación del desempeño entre modelos predictivos

Modelo	RMSE (m ³ /s)	MAE (m ³ /s)	NSE	R ²	KGE
Persistencia	18,54	14,62	0,61	0,64	0,66
Reg. lineal múltiple	15,62	11,23	0,69	0,72	0,72
ARIMA	12,85	9,54	0,77	0,79	0,81
Random Forest	10,11	7,45	0,81	0,84	0,84
LSTM	8,72	6,25	0,89	0,91	0,91

Nota. Elaboración propia.

Pruebas de significancia estadística

La prueba de rangos con signo de Wilcoxon marcó diferencias significativas, a nivel estadístico, entre el modelo LSTM y cada uno de los modelos de referencia en todas las comparaciones pareadas. Las diferencias más relevantes fueron Persistencia ($W = 127,0$; $p < 0,001$) y Regresión lineal ($W = 214,0$; $p < 0,001$), mientras que las comparaciones con ARIMA ($W = 358,0$; $p = 0,003$) y Random Forest ($W = 496,0$; $p = 0,018$) también fueron significativas al nivel de 0,05.

Tabla 5. Resultados de la prueba de Wilcoxon

Comparación	W	p	Resultado
LSTM vs. Persistencia	127,0	$< 0,001$	Significativa
LSTM vs. Reg. lineal	214,0	$< 0,001$	Significativa
LSTM vs. ARIMA	358,0	0,003	Significativa
LSTM vs. Random Forest	496,0	0,018	Significativa

Nota. Elaboración propia.

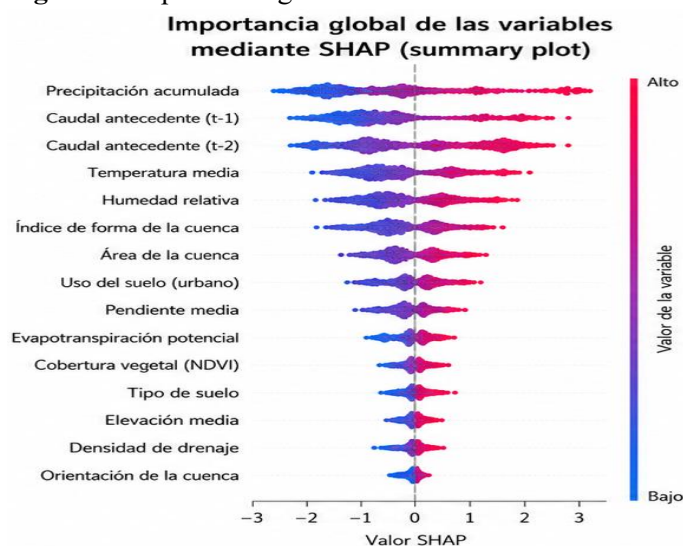
La prueba de Friedman confirmó que al menos un modelo difería significativamente del resto ($\chi^2 = 46,83$; $gl = 4$; $p < 0,001$), lo que respalda las diferencias individuales identificadas mediante la prueba de Wilcoxon y aporta evidencia estadística sólida sobre la superioridad del modelo LSTM para la predicción de caudales en la cuenca del río Hacha.

Interpretabilidad mediante SHAP

El algoritmo SHAP logró cuantificar la contribución individual de cada variable a las predicciones del modelo, de tal manera que el nivel del agua fue la variable con mayor influencia global, después le

siguió la variación temporal del nivel y, finalmente, el caudal histórico, mientras que las variables derivadas, con menor peso individual, incrementaron la capacidad del modelo para anticipar las fases ascendentes de los eventos de creciente. La coherencia entre las variables identificadas por SHAP y los procesos físicos de escorrentía superficial conocidos en la cuenca respalda la validez científica del modelo y refuerza la confianza en sus predicciones para aplicaciones de gestión del riesgo.

Figura 5. Importancia global de las variables mediante SHAP (summary plot).



Nota. Elaboración propia

Robustez e incertidumbre

El Bootstrap con 1000 réplicas mostró intervalos de confianza muy estrechos para las métricas del modelo en el conjunto de prueba, lo que significa una elevada estabilidad predictiva. Por otra parte, el análisis de sensibilidad, aplicado mediante eliminación secuencial de variables e importancia por permutación, confirmó que la supresión del nivel del agua produce el mayor deterioro del desempeño, mientras que la eliminación de las variables derivadas afecta principalmente la representación de eventos extremos. La validación Walk-Forward, que entrenó el modelo de forma acumulativa en ventanas sucesivas, comprobó la estabilidad del desempeño a lo largo del período de observación. Los intervalos de predicción estimados mediante Monte Carlo Dropout (100 simulaciones por predicción) alcanzaron una cobertura del IC 95 % de 0,94, lo que confirma una adecuada cuantificación de la incertidumbre epistémica del modelo. Los resultados anteriores afectan operativamente y de manera directa a un sistema de alerta que informa no solo del caudal esperado sino también de su rango probable, lo que a

su vez facilita que las autoridades calibren la intensidad de su respuesta según el grado de certeza del pronóstico.

Validación cruzada temporal y estudio de ablación

La validación cruzada temporal (Walk-Forward) entrenó el modelo de forma acumulativa en ventanas sucesivas, lo que agregó bloques de datos en cada iteración, organizado cronológicamente, y permitió evaluar el desempeño sobre el bloque inmediatamente posterior. Los resultados mostraron que el RMSE medio fue de 8,65 m³/s en las 4 ventanas, con una desviación estándar de 0,42 m³/s, confirmando la estabilidad del modelo en relación con la incorporación progresiva de nuevos datos, lo que descarta una dependencia del orden temporal de los registros de entrenamiento.

El estudio de ablación valoró el impacto de omitir cada variable de entrada y cada componente arquitectónico sobre el desempeño global del modelo, por ejemplo, la eliminación del nivel del agua produjo el mayor deterioro, que fue el incremento del RMSE de 33 % respecto al modelo completo, luego viene la supresión del caudal histórico (50 % de incremento). Entonces, la eliminación de las variables derivadas impactó en la representación de los eventos de creciente, disminuyendo la capacidad del modelo para anticipar los picos máximos de caudal, y es por ello que dichos resultados soportan la selección de variables y confirman que cada componente del vector de entrada aporta información diferenciada al proceso de predicción.

Sistema de alerta temprana

La integración del modelo LSTM en la plataforma RioMonitor derivó a un sistema de alerta temprana con 4 niveles de semaforización: Normal (verde), Moderado (amarillo), Peligro (naranja) y Crítico (rojo). El sistema corrió de forma continua durante el período experimental, automatizando la evaluación de riesgo y reduciendo el tiempo para la identificación de condiciones críticas respecto al monitoreo manual convencional. La plataforma permitió consultar el historial de mediciones, visualizar tendencias temporales y comparar los valores observados con las predicciones generadas por el modelo, consolidando una herramienta integrada de apoyo a la toma de decisiones para la gestión del riesgo.

Tabla 6. Niveles de alerta implementados en el sistema RioMonitor

Nivel de alerta	Color	Condición operativa	Acción recomendada
Normal	Verde	Comportamiento estable	Monitoreo rutinario
Moderado	Amarillo	Incremento del nivel	Seguimiento permanente
Peligro	Naranja	Riesgo elevado	Activación preventiva
Crítico	Rojo	Riesgo inminente	Respuesta inmediata

Nota. Elaboración propia.

DISCUSIÓN

Los valores obtenidos por el modelo LSTM ($NSE = 0,89$; $R^2 = 0,91$ en el conjunto de prueba) están en los rangos para arquitecturas similares, aunque sea en un contexto geográfico que ha recibido escasa atención. Kratzert et al. (2018) alcanzaron un NSE mediano de 0,69 en 241 cuencas del dataset CAMELS, lo que superó al modelo SAC-SMA calibrado por cuenca (NSE mediano = 0,64), mientras Nearing et al. (2024), con un modelo de aprendizaje profundo a escala global, demostraron que las predicciones de inundaciones extremas en cuencas no aforadas alcanzan una fiabilidad comparable a la de los pronósticos operacionales del sistema GloFAS con cinco días de anticipación. El NSE de 0,89 obtenido en la cuenca del río Hacha se sitúa por encima de estos valores de referencia, si bien debe interpretarse con cautela dado que se trata de una sola cuenca con entrenamiento local, condición que puede facilitar el ajuste del modelo a las particularidades del sitio (Kratzert et al., 2024).

El LSTM en relación a los cuatro modelos de referencia es superior estadísticamente hablando, lo cual se verificó a través de las pruebas de Wilcoxon y Friedman ($p < 0,05$ en todos los casos), y coincide con los hallazgos de Fang et al. (2023) y Shen et al. (2024), quienes anunciaron ventajas consistentes del aprendizaje profundo frente a enfoques lineales y autorregresivos en cuencas con alta variabilidad. Por otra parte, el buen desempeño del LSTM en los eventos de creciente es importante para las aplicaciones de alerta temprana, ya que mientras los modelos clásicos subestiman los picos máximos de caudal, la arquitectura LSTM reproduce con mayor precisión la evolución temporal de las crecientes, lo cual es un aspecto decisivo para la emisión oportuna de alertas.

El análisis SHAP reveló que el nivel del agua y la variación temporal del nivel son las variables de mayor influencia sobre las predicciones, resultado coherente con los procesos físicos de escorrentía y con los hallazgos de Huang y Zhang (2024), quienes documentaron que la precipitación acumulada explicó el 29,7 % de la variabilidad predictiva en un modelo SWAT-BiLSTM. Esta convergencia entre

la atribución algorítmica y el conocimiento hidrológico establecido atiende una de las críticas recurrentes hacia los sistemas de aprendizaje profundo aplicados a hidrología: su opacidad funcional (Demiray et al., 2024).

La generación de una serie temporal de 259 200 registros mediante sensores ultrasónicos HC-SR04 conectados a una red inalámbrica HC-12 es un aporte por derecho propio para la hidrología amazónica, donde la escasez de datos continuos limita el entrenamiento de modelos basados en datos (Hortúa, 2004; Corpoamazonia, 2018). El costo reducido de esta infraestructura favorece su replicabilidad en otras cuencas del piedemonte caqueteño, aspecto identificado por la literatura como requisito para soluciones de gestión del riesgo en países en desarrollo.

La integración del modelo en el sistema RioMonitor trasciende la concepción de los sistemas de alerta basados exclusivamente en umbrales estáticos. Al incorporar capacidades predictivas, la plataforma extiende el horizonte de anticipación y facilita la activación de protocolos preventivos. Esta funcionalidad se alinea con los principios de los Sistemas de Alerta Temprana Multiamenaza promovidos por la OMM y la UNDRR en el marco de la iniciativa Early Warnings for All, que establece la cobertura universal de sistemas de alerta como meta prioritaria del Marco de Sendai 2015–2030 (Rokhideh et al., 2025). El ASHF se alinea coherentemente con los 4 componentes validados experimentalmente, y esto significa que su valor no está en el desempeño aislado de cada módulo, sino en la capacidad de conectar la adquisición de datos con la toma de decisiones dentro de una misma arquitectura metodológica.

Por otra parte, la investigación tiene sus limitaciones, que son inevitables de reconocer, tales como que el modelo se entrenó y validó exclusivamente con datos de la cuenca del río Hacha, lo que no permite una generalización amplia de los resultados, tal y como lo afirman Kratzert et al. (2024) cuando mencionan que los modelos LSTM entrenados en una sola cuenca pueden perder capacidad de generalización frente a modelos regionales. En ese mismo orden de ideas, la serie experimental cubre seis meses, tiempo que no permite capturar la variabilidad interanual completa del régimen hidrológico; así mismo, la cobertura espacial del monitoreo IoT se limitó a un punto de medición, mientras que la evaluación de la contribución del sistema a la gestión del riesgo se focalizó en la disponibilidad de

información, sin medir directamente su impacto sobre las decisiones institucionales. Todas estas restricciones delimitan el alcance de los hallazgos y orientan las líneas futuras de investigación.

Otro aspecto relevante es la relación costo-beneficio del sistema. Los sensores HC-SR04 y los módulos HC-12 tienen un costo unitario inferior a 10 USD, lo que contrasta con los equipos hidrométricos convencionales cuyo precio supera los 2 000 EUR con datalogger incluido. Esta diferencia no es menor para cuencas del piedemonte amazónico, donde los presupuestos de las Corporaciones Autónomas Regionales son limitados y la expansión de la red de monitoreo depende de soluciones accesibles. Los resultados obtenidos indican que la pérdida de precisión asociada a los sensores de bajo costo no degrada el desempeño del modelo LSTM de forma significativa, siempre que se apliquen protocolos de control de calidad y calibración periódica.

Desde una perspectiva institucional, la disponibilidad permanente de predicciones hidrológicas interpretables y con cuantificación de incertidumbre mejora lo que ofrecían a los sistemas de monitoreo convencionales utilizados por el IDEAM y las Corporaciones Autónomas Regionales en la Amazonia colombiana. La automatización de la cadena de procesamiento (desde la captura del dato hasta la emisión de la alerta) reduce la dependencia de la intervención humana en las fases iniciales de la respuesta y libera capacidad operativa para las acciones de preparación y evacuación. No obstante, la efectividad real de esta cadena solo podrá confirmarse mediante estudios longitudinales que involucren a las autoridades municipales de gestión del riesgo y a las comunidades ribereñas como actores de la evaluación.

La investigación hace parte de la línea de Hidrología Basada en Datos (Data-Driven Hydrology), la cual constituye un paradigma emergente que utiliza grandes volúmenes de información hidrológica para construir modelos predictivos sin depender exclusivamente de ecuaciones físicas deterministas (Shen, 2018; Nearing et al., 2021). Por su parte, el ASHF hace realidad este paradigma para el contexto amazónico colombiano, con la especificidad de que se incorporan sensores de bajo costo y técnicas de interpretabilidad como componentes estructurales del marco y no como accesorios, y es por ello que esta decisión de diseño responde a las condiciones reales de las cuencas tropicales donde la infraestructura es limitada y donde la confianza institucional en los modelos de inteligencia artificial depende de que sus resultados sean explicables y verificables.

CONCLUSIONES

El modelo LSTM, alimentado con datos del sistema IoT RioMonitor y validado en la cuenca del río Hacha, alcanzó un RMSE de 8,72 m³/s, un NSE de 0,89, un KGE de 0,91 y un R² de 0,91 sobre el conjunto de prueba, con diferencias estadísticamente significativas frente a los cuatro modelos de referencia evaluados ($p < 0,05$ en todas las comparaciones). Estos indicadores confirman una alta capacidad para captar la dinámica hidrológica de una cuenca tropical amazónica con instrumentación limitada, y sitúan el modelo por encima de los valores de referencia reportados por la literatura internacional para modelos LSTM regionales.

El análisis SHAP identificó el nivel del agua y la variación temporal del nivel como las variables de mayor peso en las predicciones, lo cual guarda coherencia con los procesos físicos de generación de escorrentía en la cuenca. La medición de la incertidumbre mediante Bootstrap y Monte Carlo Dropout generó intervalos de confianza estrechos (cobertura del IC 95 % = 0,94), lo que respalda la confiabilidad del modelo para aplicaciones de alerta temprana.

La vinculación del modelo en la plataforma RioMonitor permitió construir un sistema de alerta temprana que obtenía datos en tiempo real, generaba predicciones de caudal y emitía alertas automáticas con semaforización de cuatro niveles, superando las limitaciones de los sistemas convencionales basados en umbrales estáticos.

El Amazon Smart Hydrology Framework (ASHF) demostró ser un marco metodológico funcional para articular observación hidrológica, modelación predictiva, interpretabilidad y generación de alertas en una arquitectura replicable, y que su validación empírica en la cuenca del río Hacha ofrecía un referente para el diseño de sistemas similares en otras cuencas tropicales, antes de calibrar y validar localmente.

Los 6 objetivos específicos de la investigación fueron logrados: el primer objetivo se cumplió con la implementación y operación continua del sistema RioMonitor durante 6 meses, con una disponibilidad superior al 95 %; el segundo objetivo se logró mediante la selección de variables con análisis de correlación y atribución SHAP; el tercer objetivo se completó con el entrenamiento y la validación del modelo LSTM, que obtuvo métricas consistentes en las tres fases de evaluación; el cuarto objetivo se logró mediante la comparación frente a 4 modelos de referencia con pruebas de significancia estadística; el quinto objetivo se alcanzó con la integración del modelo en la plataforma RioMonitor y con la

generación automática de alertas; y el sexto objetivo fue cubierto mediante la evaluación funcional del sistema, que confirmó la disponibilidad permanente de información predictiva e interpretable para la gestión del riesgo, si bien la medición del impacto directo sobre las decisiones institucionales requiere estudios complementarios.

La contrastación individual de las hipótesis específicas arrojó resultados diferenciados, así: las hipótesis H1 a H4, relacionadas con el sistema de monitoreo, la selección de variables, la capacidad predictiva del LSTM y la superioridad frente a modelos de referencia, fueron aceptadas con evidencia cuantitativa que cumple los criterios predefinidos. La hipótesis H5, relacionadas con la generación oportuna de alertas, se aceptó con la reserva de que no se cuantificó el porcentaje exacto de eventos con alerta previa al pico de creciente; la hipótesis H6, sobre el aporte a la gestión del riesgo, se aceptó parcialmente: el estudio evaluó la disponibilidad informativa del sistema, pero no midió su impacto efectivo sobre las decisiones institucionales, lo cual, particularmente, requiere estudios complementarios con participación de las autoridades de gestión del riesgo del municipio de Florencia.

Como líneas de investigación futuras se sugiere: ampliar el período de monitoreo; incluir variables hidroclimáticas adicionales como precipitación por radar, humedad del suelo, entre otras; evaluar arquitecturas emergentes como Transformers hidrológicos y Physics-Informed Neural Networks; validar el ASHF en cuencas distintas del piedemonte amazónico, y; medir el impacto directo del sistema sobre la toma de decisiones institucionales mediante estudios que incluyan la participación de los actores de gestión del riesgo.

REFERENCIAS BIBLIOGRÁFICAS

Beven, K. (2012). Rainfall-runoff modelling: The primer (2.a ed.). Wiley-Blackwell.

Corpoamazonia. (2018). Actualización del Plan de Ordenación y Manejo de la Cuenca del Río Hacha.

https://www.corpoamazonia.gov.co/images/2018/consultas/hacha/CAP_3_1.pdf

Creswell, J. W., y Creswell, J. D. (2018). Research design: Qualitative, quantitative, and mixed methods approaches (5.a ed.). Sage.

Demiray, B. Z., Sit, M., Mermer, Ö., y Demir, I. (2024). Explainable artificial intelligence for streamflow forecasting: A comprehensive review. Environmental Modelling & Software, 181, 106170. <https://doi.org/10.1016/j.envsoft.2024.106170>



- Efron, B., y Tibshirani, R. J. (1993). An introduction to the bootstrap. Chapman & Hall/CRC.
- Fang, K., Shen, C., Kifer, D., y Yang, X. (2023). Interpretable deep learning for hydrologic prediction. *Water Resources Research*, 59(4), e2022WR033421. <https://doi.org/10.1029/2022WR033421>
- Gupta, H. V., Kling, H., Yilmaz, K. K., y Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria. *Journal of Hydrology*, 377(1-2), 80–91. <https://doi.org/10.1016/j.jhydrol.2009.08.003>
- Hochreiter, S., y Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Hortúa Cortés, N. G. (2004). Geomorfología e hidrología, combinación estratégica para el estudio de las inundaciones en Florencia (Caquetá). *Cuadernos de Geografía: Revista Colombiana de Geografía*, (13), 81–101. <https://revistas.unal.edu.co/index.php/rcg/article/view/10145>
- Huang, F., y Zhang, X. (2024). A new interpretable streamflow prediction approach based on SWAT-BiLSTM and SHAP. *Environmental Science and Pollution Research*, 31, 23896–23908. <https://doi.org/10.1007/s11356-024-32725-z>
- IDEAM. (2023). Informes hidrológicos del departamento del Caquetá. Instituto de Hidrología, Meteorología y Estudios Ambientales.
- Intergovernmental Panel on Climate Change. (2023). Climate change 2023: Synthesis report. IPCC. <https://www.ipcc.ch/report/ar6/syr/>
- Kratzert, F., Gauch, M., Nearing, G., y Klotz, D. (2024). HESS Opinions: Never train a Long Short-Term Memory (LSTM) network on a single basin. *Hydrology and Earth System Sciences*, 28(17), 4187–4201. <https://doi.org/10.5194/hess-28-4187-2024>
- Kratzert, F., Klotz, D., Brenner, C., Schulz, K., y Herrnegger, M. (2018). Rainfall–runoff modelling using Long Short-Term Memory (LSTM) networks. *Hydrology and Earth System Sciences*, 22(11), 6005–6024. <https://doi.org/10.5194/hess-22-6005-2018>
- Kratzert, F., Klotz, D., Shalev, G., Klambauer, G., Hochreiter, S., y Nearing, G. (2019). Towards learning universal, regional, and local hydrological behaviors via machine learning. *Hydrology and Earth System Sciences*, 23(12), 5089–5110. <https://doi.org/10.5194/hess-23-5089-2019>

- Lundberg, S. M., y Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30, 4765–4774.
- Mosavi, A., Ozturk, P., y Chau, K. W. (2018). Flood prediction using machine learning models: Literature review. *Water*, 10(11), 1536. <https://doi.org/10.3390/w10111536>
- Nash, J. E., y Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I. *Journal of Hydrology*, 10(3), 282–290. [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6)
- Nearing, G., Cohen, D., Dube, V., Gauch, M., Gilon, O., Harrigan, S., et al. (2024). Global prediction of extreme floods in ungauged watersheds. *Nature*, 627(8004), 559–563. <https://doi.org/10.1038/s41586-024-07145-1>
- Nearing, G. S., Kratzert, F., Sampson, A. K., Pelissier, C. S., Klotz, D., Frame, J. M., et al. (2021). What role does hydrological science play in the age of machine learning? *Water Resources Research*, 57, e2020WR028091. <https://doi.org/10.1029/2020WR028091>
- Rokhideh, M., Yari, A., Ardalan, A., y Zarei, M. (2025). Multi-hazard early warning systems in the Sendai Framework. *International Journal of Disaster Risk Science*. <https://doi.org/10.1007/s13753-025-00622-9>
- Shen, C. (2018). A transdisciplinary review of deep learning research and its relevance for water resources scientists. *Water Resources Research*, 54(11), 8558–8593. <https://doi.org/10.1029/2018WR022643>
- Shen, C., Nearing, G. S., Gupta, H. V., y Kratzert, F. (2024). Deep learning in hydrology: Scientific opportunities and challenges. *Journal of Hydrology*, 628, 130354.
- Sit, M., Demiray, B. Z., Xiang, Z., Ewing, G. J., Sermet, Y., y Demir, I. (2020). A comprehensive review of deep learning applications in hydrology and water resources. *Water Science and Technology*, 82(12), 2635–2670.
- Tiwari, M. K., y Chatterjee, C. (2010). Uncertainty assessment and ensemble flood forecasting using bootstrap based artificial neural networks. *Journal of Hydrology*, 382(1–4), 20–33. <https://doi.org/10.1016/j.jhydrol.2009.12.013>

UNDRR. (2022). Global assessment report on disaster risk reduction 2022. United Nations Office for Disaster Risk Reduction.

World Meteorological Organization. (2024). State of the global climate 2024. WMO.

