

Enfoque metodológico basado en modelación híbrida para la estimación de la incertidumbre predictiva en la simulación hidrológica de cuencas a escala diaria

¹Javier Eduardo Bonilla Perdomo, ²Miguel Barrios

¹Universidad Surcolombiana. Neiva, Colombia

²Universidad del Tolima. Ibagué, Colombia

Resumen

La modelación hidrológica de cuencas es una herramienta fundamental para comprender los procesos físicos que rigen el movimiento y la distribución del agua, así como para la cuantificación de la disponibilidad hídrica y la planificación de los recursos hídricos. Sin embargo, la complejidad inherente de los procesos hidrológicos, marcada por su heterogeneidad, variabilidad espacio-temporal y comportamiento no lineal, introduce incertidumbres significativas en las simulaciones del ciclo hidrológico. La presente investigación propone un enfoque metodológico para la estimación y reducción de la incertidumbre predictiva en la modelación hidrológica. Para ello, se emplean técnicas de aprendizaje automático junto con el uso de modelos hidrológicos agregados (GR4J), permitiendo evaluar la incertidumbre en la simulación continua de caudales a escala diaria. Como caso de estudio, se analiza una cuenca andina tropical, donde se desarrolla una estrategia de modelos híbridos que combina la conceptualización de procesos físicos con redes neuronales artificiales profundas. Este enfoque busca mejorar la comprensión de los mecanismos que generan incertidumbre en la modelación hidrológica, así como explorar alternativas para reducirla mediante modelos híbridos que integren el conocimiento físico con las ventajas del aprendizaje automático. Los resultados obtenidos tendrán aplicaciones en la planificación de los recursos hídricos, asegurando una distribución eficiente del agua para distintos usos (agrícola, industrial y doméstico), además de contribuir a la evaluación del impacto del cambio climático en la disponibilidad hídrica y en la caracterización de eventos extremos como sequías.

Palabras clave: modelación hidrológica; cuenca; incertidumbre predictiva; incertidumbre en datos; incertidumbre paramétrica; incertidumbre estructural; aprendizaje automático; ensamble de modelos; modelo agregado; caudal; modelo híbrido; redes neuronales artificiales; recurso hídrico

Abstract

Watershed Hydrological modeling is a fundamental tool for understanding the physical processes that govern water movement and distribution, as well as for quantifying water availability and planning water resources. However, the inherent complexity of hydrological processes, characterized by their heterogeneity, spatiotemporal variability, and nonlinear behavior, introduces significant uncertainties into hydrological cycle simulations. This research proposes a methodological approach for estimating and reducing predictive uncertainty in hydrological modeling. To this end, machine learning techniques are used in conjunction with aggregated hydrological models (GR4J), allowing for the assessment of uncertainty in continuous flow simulation at the daily scale. As a case study, a tropical Andean basin is analyzed, where a hybrid modeling strategy is developed that combines the conceptualization of physical processes with deep artificial neural networks. This approach seeks to improve the understanding of the mechanisms that generate uncertainty in hydrological modeling, as well as to explore alternatives to reduce it through hybrid models that integrate physical knowledge with the advantages of machine learning. The results obtained will have applications in water resources planning, ensuring efficient water distribution for different uses (agricultural, industrial, and domestic), in addition to contributing to the assessment of the impact of climate change on water availability and in the characterization of extreme events such as droughts.

Keywords: hydrologic modeling; watershed; predictive uncertainty; data uncertainty; parametric uncertainty; structural uncertainty; machine learning; model ensemble; aggregate model; flow; hybrid model; artificial neural networks; water resources

1. Introducción

La demanda mundial de recursos hídricos está creciendo, impulsada por el aumento de la población y el desarrollo, que requieren más agua para consumo doméstico, industrial y agrícola. Esto ha provocado alteraciones hidrológicas debido al cambio climático y a los cambios en el uso del suelo en las cuencas. Según la FAO (2009), para el año 2050 la agricultura deberá incrementar su producción de alimentos en un 60%, lo que implicará una mayor demanda de agua. Se ha observado que la disponibilidad de agua ha disminuido en algunas partes del mundo y se espera que continúe reduciéndose en las próximas décadas, como ocurrió en Chile en 2013 (Muñoz et al., 2018). En este contexto, Diaz et al. (2009) subraya que el recurso hídrico es fundamental para el desarrollo sostenible e indispensable para los aspectos sociales, económicos, ambientales del ser humano y los demás seres vivos del planeta. Por lo tanto, se presentan desafíos significativos para la planificación y gestión integral del régimen hidrológico.

La gestión de los recursos hídricos se apoya en el uso de modelos hidrológicos, los cuales son actualmente indispensables para cuantificar las variables del balance hídrico. Dada la alta complejidad de los procesos hidrológicos (variabilidad, heterogeneidad, no linealidad), las representaciones hechas por estos modelos presentan incertidumbres asociadas a la estimación de parámetros de entrada, la estructura del modelo y los datos de entrada, por lo que un tema de interés para mejorar las capacidades de la simulación hidrológica es estimar y reducir estas

incertidumbres. En este proceso, pueden obtenerse múltiples soluciones óptimas, ya que diferentes conjuntos de parámetros pueden realizar predicciones similares y esta característica le agrega un desafío adicional a los procesos de modelación hidrológica (Hernández et al., 2014).

Estas herramientas, especialmente los modelos hidrológicos, han evolucionado considerablemente en las últimas décadas debido al crecimiento computacional. Esto ha permitido analizar una mayor disponibilidad de datos hidrológicos y una comprensión física y dinámica de los modelos cada vez más robusta (Salgado et al., 2023). Sin embargo, Vrugt et al. (2005) señalan que los modelos hidrológicos enfrentan problemas de incertidumbre debido a errores en los procesos de modelación relacionados con los datos de entrada, las condiciones físicas de contorno y la conceptualización de los procesos hidrológicos. Por consiguiente, según Moges et al. (2021) las fuentes de incertidumbre en los modelos hidrológicos provienen de los parámetros, la estructura del modelo y la calibración de los datos de entrada, lo que resulta en la falta de un conjunto único de parámetros óptimos.

Por otro lado, la combinación de modelos hidrológicos con técnicas de aprendizaje automático permite la creación de modelos híbridos que pueden adaptarse dinámicamente a diferentes condiciones hidrológicas y climáticas. Estas técnicas innovadoras no solo buscan reducir la incertidumbre estructural, sino también mejorar la generación de conocimiento y la comprensión de los procesos hidrológicos en cuencas andinas tropicales. Por lo tanto, Neiting y Aftery (2007) indican que las técnicas de aprendizaje automático pueden mejorar considerablemente la calidad predictiva al identificar patrones complejos y ajustar los modelos en función de datos observados. En las últimas décadas, el aprendizaje automático se ha aplicado de manera eficiente en la hidrología, mejorando las predicciones de inundaciones, las estimaciones de precipitaciones, la calidad del agua y el agua subterránea (Mosaffa et al., 2022).

Adicionalmente, las aplicaciones de técnicas de aprendizaje automático en la modelación hidrológica se muestran eficaces en la mejora de la estimación y predicción de caudales. Esto lo afirma Kuok y Bessaih (2007) quienes indican que los comportamientos naturales de los procesos hidrológicos son apropiados para trabajar con redes neuronales, las cuales se utilizan para construir modelos conceptuales de lluvia-escorrentía, generando resultados de la simulación de la escorrentía con alta precisión para las cuencas. Chen et al. (2013) mencionan que los modelos basados en redes neuronales pueden capturar la complejidad de los procesos hidrológicos y mejorar la precisión en la predicción de caudales, además de utilizar técnicas para recuperar datos faltantes. La integración de estos métodos con modelos hidrológicos tradicionales en un enfoque de ensamble puede ofrecer una solución para reducir la incertidumbre estructural y mejorar las predicciones en cuencas. Este trabajo tiene por objetivo analizar el desempeño de realizar la modelación hidrológica híbrida combinando un modelo hidrológico agregado (GR4J) y un modelo de red neuronal profunda con memoria de largo plazo (LSTM) en una cuenca andina tropical de montaña.

2 Materiales y Métodos

2.1 Caso de estudio

Se seleccionó la cuenca hidrográfica del río Recio en el departamento del Tolima (Colombia), la cual presenta un área drenante de 645.51 km² hasta la estación hidrométrica La Nueva (Código 21257100) operada por el Instituto de Hidrología y Estudios Ambientales de Colombia (IDEAM), que presenta un caudal medio de 27.1 m³/s en el periodo comprendido entre el año 2002 y 2012. El régimen pluviométrico en la cuenca es bimodal con dos periodos húmedos en el año hidrológico y una precipitación media anual de 1826 mm y una temperatura media de 22.5 °C. Para el análisis se seleccionó el periodo continuo de datos hidrológicos y meteorológicos correspondiente al 1 de enero de 2002 al 31 de diciembre de 2012, cuya consistencia fue analizada mediante el análisis de curvas de doble masa.

2.2 Implementación de la modelación hidrológica

El modelo GR4J (Génie Rural à 4 paramètres Journalier) es un modelo hidrológico conceptual de tipo lluvia-escorrentía, diseñado para simular el caudal diario a partir de la precipitación y la evapotranspiración de referencia (). Para su implementación en la cuenca hidrográfica del río Recio, se siguieron los siguientes pasos:

- **Selección y preparación de datos**

Se recopilaron y depuraron las series temporales diarias de precipitación, temperatura y caudal observado en la cuenca. La evapotranspiración de referencia (ET₀) se estimó utilizando el método Penman-Monteith (Allen, et. Al., 1998). Dado que el GR4J es un modelo agregado espacialmente, los valores de precipitación y evapotranspiración provenientes de las estaciones seleccionadas fueron promediados arealmente para representar de forma integrada las condiciones meteorológicas de la cuenca.

- **Configuración del modelo GR4J**

Se adoptó la estructura estándar del modelo GR4J, que incluye un almacenamiento de producción de escorrentía, un almacenamiento de enrutamiento y dos funciones de transferencia. El modelo se calibró utilizando sus cuatro parámetros libres:

- X1. Capacidad de almacenamiento de producción de escorrentía
- X2. Coeficiente de intercambio subterráneo
- X3. Capacidad de almacenamiento de enrutamiento
- X4. Tiempo base del enrutamiento unitario

- **Calibración y Validación**

Se dividió la serie de datos en tres periodos: un de calentamiento del modelo comprendido del 01-01-2002 al 31-12-2002 para definir las condiciones iniciales de humedad en la cuenca, un segundo set de datos del 01-01-2003 al 31-12-2009 para la calibración del modelo y un tercer set de datos del 01-01-2010 al 31-12-2012 para la validación del modelo.

La calibración se realizó de forma automática aplicando el algoritmo de optimización global propuesto por Michel (1991). para ajustar los cuatro parámetros del modelo, maximizando la eficiencia entre el caudal simulado y observado, para ello se utilizó el coeficiente de Nash-Sutcliffe (NSE). Para el análisis de los resultados de modelación se implementaron las siguientes métricas:

$$KGE = 1 - \sqrt{(r - 1)^2 + \left(\frac{\sigma_{Y, sim}}{\sigma_{Y, obs}} - 1\right)^2 + \left(\frac{\mu_{Y, sim}}{\mu_{Y, obs}} - 1\right)^2}$$

KGE: eficiencia de Kling-Gupta. se obtienen tres elementos correspondientes a la correlación lineal (coeficiente de Pearson r), la relación entre desviaciones estándar (σ) y la relación entre las medias (μ).

$$NSE = 1 - \frac{\sum_{t=1}^N (Q_{obs,t} - Q_{sim,t})^2}{\sum_{t=1}^N (Q_{obs,t} - \bar{Q}_{obs})^2}$$

NSE: eficiencia de Nash-Sutcliffe. Se define como uno menos la relación entre la varianza del error de la serie de caudales simulado dividida en la varianza de la serie de caudales observada.

2.3 Modelo propuesto para la corrección del sesgo

En este trabajo se asume que la precipitación es responsable del mayor porcentaje de error en la simulación, por lo que se propone representar la incertidumbre de la precipitación a partir de un estimador del sesgo relativo en cada paso de tiempo ($S_{r,t}$) entre el Caudal Observado ($Q_{o,t}$) y el Caudal Simulado ($Q_{s,t}$), este último estimado con la lluvia observada considerada incierta por efecto de errores de medición y variabilidad espacial, donde:

$$S_{r,t} = \frac{Q_{obs,t} - Q_{sim,t}}{Q_{obs,t}}$$

La precipitación corregida por el error de medición y efecto espacial queda definida por:

$$P_c = \frac{P_i}{1 - S_{r,t} + E_{sr}}$$

Donde, P_c es la precipitación corregida, $S_{r,t}$ es el error relativo modelado por una red neuronal del tipo Long Short-Term Memory (LSTM) y E_{sr} es el error aleatorio del estimador, en este trabajo lo presentamos como cero a efectos de concreción.

Para modelar el error relativo, se implementó una red neuronal LSTM utilizando la biblioteca TensorFlow (en Python). La serie original fue normalizada mediante MinMaxScaler para asegurar un adecuado comportamiento del modelo. Posteriormente, se estructuraron las entradas en secuencias de 10 pasos diarios para capturar dependencias temporales. La arquitectura de la red constó de dos capas LSTM con 50 unidades cada una (la primera con salida secuencial), seguidas por una capa densa intermedia con 25 neuronas y una capa de salida con una sola unidad. El modelo fue entrenado con el optimizador Adam y la función de pérdida de error cuadrático medio

(MSE), utilizando el 80 % de los datos para entrenamiento y el 20 % restante para validación. Finalmente, se evaluó el desempeño del modelo tanto en el conjunto de entrenamiento como en el de prueba mediante predicciones, graficas de residuos y análisis visual del ajuste entre los valores observados y los estimados.

Finalmente, el modelo hidrológico fue nuevamente calibrado y validado utilizando la serie de precipitación corregida, y comparados sus resultados mediante las métricas descritas (KGE y NSE) con el modelo original.

3 Resultados

Los resultados de la modelación hidrológica con el modelo GR4J muestran un desempeño satisfactorio en ambos periodos (calibración y validación; Figuras 1 y 2), con mejoras aparentes al aplicar una corrección en la precipitación de entrada. En la versión sin corrección, el modelo alcanza un KGE de 0.80 y 0.82 en calibración y validación, respectivamente, y un NSE de 0.62 y 0.67, lo cual indica una buena capacidad para reproducir la dinámica del caudal. No obstante, al introducir la corrección en la precipitación, se observa una ligera disminución del KGE en calibración (0.69), pero se mantiene un desempeño comparable en validación (0.81), mientras que el NSE mejora en ambos periodos.

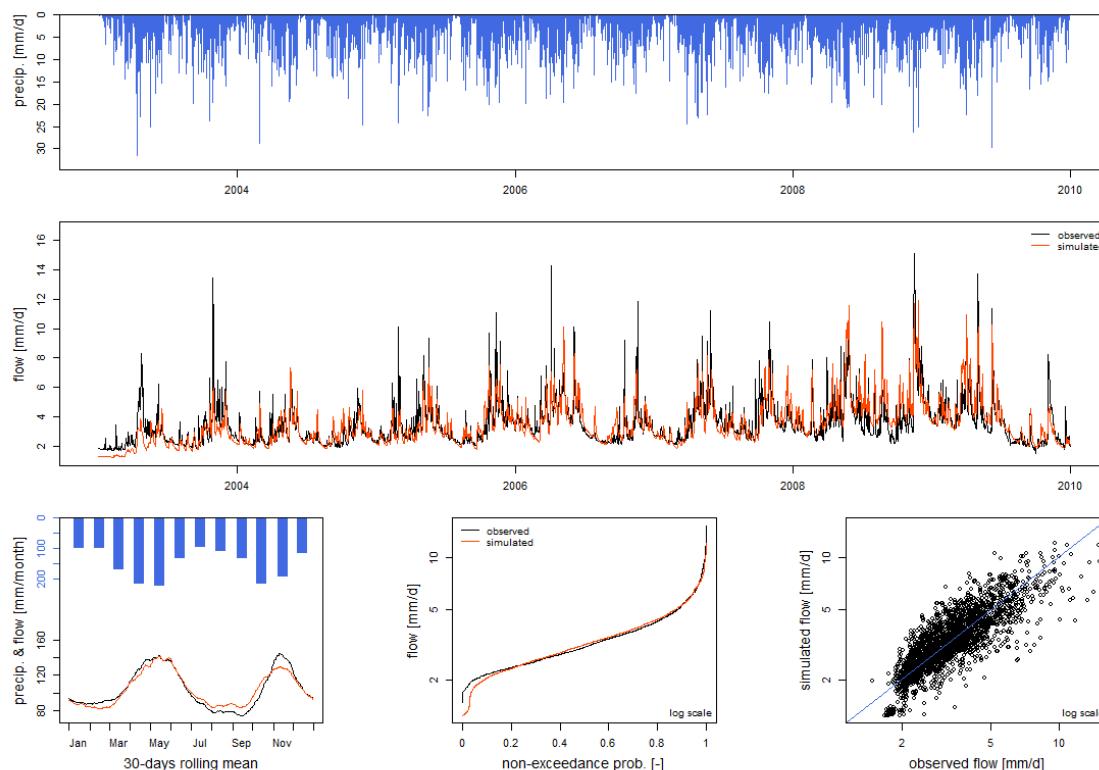


Figura 1. Resultados de calibración del modelo hidrológico. Periodo 2003-2009



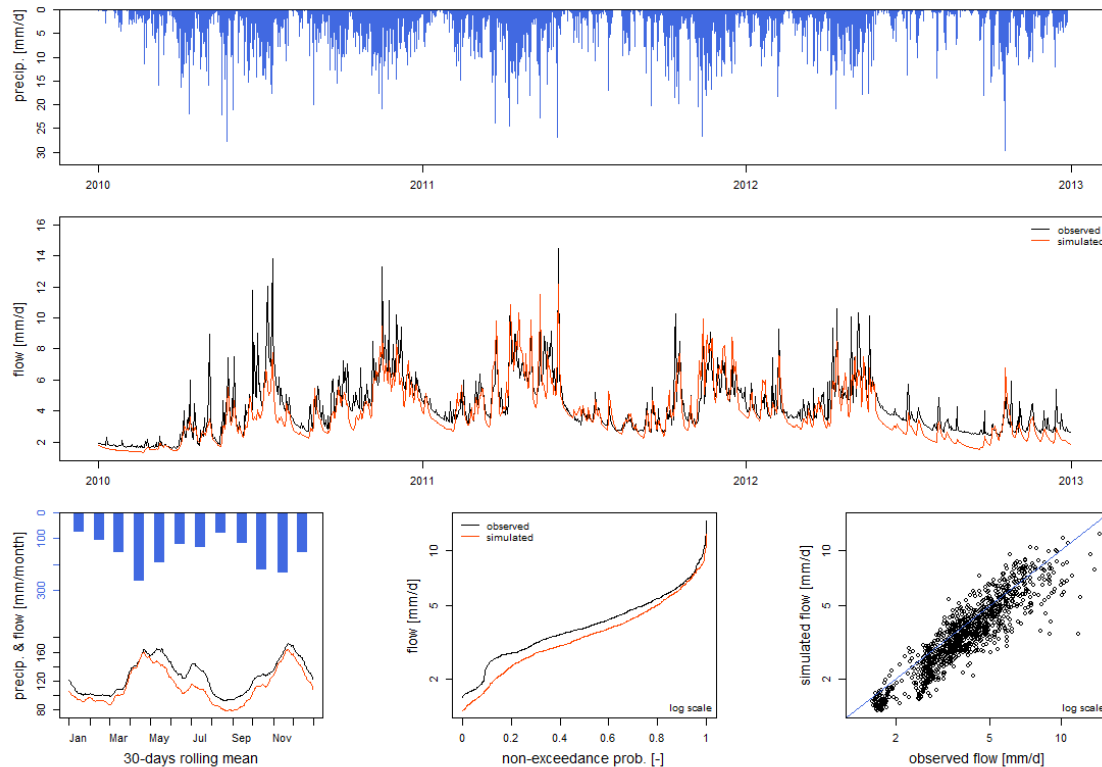


Figura 1. Resultados de validación del modelo hidrológico. Periodo 2010-2012

El coeficiente de correlación (R) se incrementa de 0.80 a 0.81 en calibración y de 0.86 a 0.87 en validación, lo que indica una mejora en la relación lineal entre los caudales simulados y observados al aplicar el modelo de corrección de la lluvia. Por su parte, la relación entre las desviaciones estándar simulado/observado no logra representar una mayor eficiencia en la representación de la variabilidad de los caudales al aplicar el método de corrección, sugiriendo la necesidad de mejoras conceptuales en el modelo de corrección de la precipitación. Finalmente, la razón entre medias se aproxima a la unidad en ambas versiones, lo que indica ausencia de sesgo sistemático.

4 Conclusión

La implementación del modelo GR4J en la cuenca del río Recio permitió simular de manera satisfactoria el comportamiento hidrológico diario, tanto en el periodo de calibración como en el de validación, con indicadores de eficiencia que superan los umbrales comúnmente aceptados en modelación hidrológica. La propuesta de corregir la precipitación mediante un estimador de sesgo relativo modelado con una red neuronal LSTM contribuyó a mejorar ciertos aspectos del desempeño del modelo, especialmente en la correlación y el ajuste de la media de los caudales simulados. Sin embargo, la reducción en la eficiencia (KGE) durante la calibración y la persistente dificultad para representar adecuadamente la variabilidad del caudal (desviaciones estándar) indican que, si bien el enfoque es prometedor, requiere ajustes adicionales para capturar con mayor precisión la incertidumbre asociada a los insumos meteorológicos. En conjunto, los resultados evidencian la

importancia de considerar estrategias de corrección de entrada en la modelación hidrológica, pero también refuerzan la necesidad de robustecer los aspectos de representación conceptual de los procesos hidrometeorológicos complejos en cuencas tropicales de montaña.

5 Referencias

- Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). Crop evapotranspiration: Guidelines for computing crop water requirements (FAO Irrigation and Drainage Paper No. 56). FAO.
- Chen, S. M., Wang, Y. M., & Tsou, I. (2013). Using artificial neural network approach for modelling rainfall-runoff due to typhoon. *Journal of Earth System Science*, 122(2), 399–405. <https://doi.org/10.1007/s12040-013-0289-8>
- Diaz, A., Chingate, N., Muñoz, D. P., Olaya, W. R., Perilla, C., Sanchez, F., & Sanchez, K. (2009). Desarrollo sostenible y el agua como derecho en Colombia Sustainable Development and the Right to Water in Colombia. *Historia*, 11(1), 84–116.
- FAO. (2009). La agricultura mundial en la perspectiva del año 2050. Fao, 4. https://www.fao.org/fileadmin/templates/wsfs/docs/Issues_papers/Issues_papers_SP/La_agricultura_mundial.pdf
- Fernández, M., Vela, A., & Castaño, S. (1997). La utilización de modelos en hidrología. In *Ensayos: Revista de la Facultad de Educación de Albacete* (Issue 12, pp. 305–318).
- Gattke, C., & Schumann, A. (2007). Comparison of different approaches to quantify the reliability of hydrological simulations. *Advances in Geosciences*, 11, 15–20. <https://doi.org/10.5194/adgeo-11-15-2007>
- Gupta, A., & Govindaraju, R. S. (2023). Uncertainty quantification in watershed hydrology: Which method to use? *Journal of Hydrology*, 616(October 2022), 128749. <https://doi.org/10.1016/j.jhydrol.2022.128749>
- Hassan, Z., Jefferson, A. J., Avellaneda, P. M., & Bhaskar, A. S. (2024). Assessment of hydrological parameter uncertainty versus climate projection spread on urban streamflow and floods. *Journal of Hydrology*, 638(June), 131546. <https://doi.org/10.1016/j.jhydrol.2024.131546>
- Hernández, A., Solera Solera, A., Nowak, W., & Andreu, J. (2014). Reducción de incertidumbre en la parametrización de modelos hidrológicos usados para la gestión conjunta de recursos hídricos. *January 2016*, 11.
- Kuok, K. K., & Bessaih, N. (2007). Artificial Neural Networks (ANNs) for Daily Rainfall Runoff Modelling. *The Journal of the Institution of Engineers*, 68(3), 31–42. https://www.researchgate.net/publication/309546851_Artificial_Neural_Networks_ANNs_for_Daily_Rainfall_Runoff_Modelling
- Martínez, C. M. G., López, E. J. G., & Chamorro, F. R. M. (2018). Estado del arte de modelos hidrológicos. *Boletín Informativo CEI*, 5(2), 116–117.
- Méndez, J. (2017). Análisis comparativo de modelos agregados de balance hídrico en la España Peninsular y su incertidumbre climática.
- Michel, C. (1991), *Hydrologie appliquée aux petits bassins ruraux*. Hydrology handbook (in French), Cemagref, Antony, France <https://dialnet.unirioja.es/servlet/tesis?codigo=153504&info=resumen&idioma=SPA>
- Moges, E., Demissie, Y., Larsen, L., & Yassin, F. (2021). Review: Sources of hydrological model uncertainties and advances in their analysis. *Water (Switzerland)*, 13(1), 1–23. <https://doi.org/10.3390/w13010028>
- Mosaffa, H., Sadeghi, M., Mallakpour, I., Naghdzadegan Jahromi, M., & Pourghasemi, H. R. (2022). Application of machine learning algorithms in hydrology. *Computers in Earth and Environmental Sciences: Artificial Intelligence and Advanced Technologies in Hazards and Risk*



- Management, 585–591. <https://doi.org/10.1016/B978-0-323-89861-4.00027-0>
- Muñoz, E., Gutiérrez-Vejar, J. C., & Zapata, P. I. T. (2018). Incertidumbre en los caudales de salida de un modelo hidrológico semidistribuido. *Tecnología y Ciencias Del Agua*, 9(2), 150–172. <https://doi.org/10.24850/j-tyca-2018-02-06>
 - Neiting, T. G., & Aftery, A. E. R. (2007). Strictly Proper Scoring Rules , Prediction , and Estimation. 102(477). <https://doi.org/10.1198/016214506000001437>
 - Renard, B., Kavetski, D., Kuczera, G., Thyer, M., & Franks, S. W. (2010). Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors. *Water Resources Research*, 46(5), 1–22. <https://doi.org/10.1029/2009WR008328>
 - Salgado, F., Barrios, M., & Velez Upegui, J. (2023). Skew-normal distribution model for rainfall uncertainty estimation in a distributed hydrological model. *Hydrological Sciences Journal*, 00(00), 1–10. <https://doi.org/10.1080/02626667.2023.2185149>
 - Scheller, M., van Meerveld, I., Sauquet, E., Vis, M., & Seibert, J. (2024). Are temporary stream observations useful for calibrating a lumped hydrological model? *Journal of Hydrology*, 632(December 2023), 130686. <https://doi.org/10.1016/j.jhydrol.2024.130686>
 - Teweldebrhan, A. T., Burkhart, J. F., & Schuler, T. V. (2018). Parameter uncertainty analysis for an operational hydrological model using residual-based and limits of acceptability approaches. *Hydrology and Earth System Sciences*, 22(9), 5021–5039. <https://doi.org/10.5194/hess-22-5021-2018>
 - Tian, Y., Xu, Y. P., Yang, Z., Wang, G., & Zhu, Q. (2018). Integration of a parsimonious hydrological model with recurrent neural networks for improved streamflow forecasting. *Water (Switzerland)*, 10(11). <https://doi.org/10.3390/w10111655>
 - Tu, T., Li, Y., Duan, K., & Zhao, T. (2024). Enhancing physically-based hydrological modeling with an ensemble of machine-learning reservoir operation modules under heavy human regulation using easily accessible data. *Journal of Environmental Management*, 359(October 2023), 121044. <https://doi.org/10.1016/j.jenvman.2024.121044>
 - Van Beusekom, A. E., Hay, L. E., Bennett, A. R., Choi, Y. D., Clark, M. P., Goodall, J. L., Li, Z., Maghami, I., Nijssen, B., & Wood, A. W. (2022). Hydrologic Model Sensitivity to Temporal Aggregation of Meteorological Forcing Data: A Case Study for the Contiguous United States. *Journal of Hydrometeorology*, 23(2), 167–183. <https://doi.org/10.1175/JHM-D-21-0111.1>
 - Vrugt, J. A., Diks, C. G. H., Gupta, H. V., Bouten, W., & Verstraten, J. M. (2005). Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation. *Water Resources Research*, 41(1), 1–17. <https://doi.org/10.1029/2004WR003059>
 - Zhang, J., Li, Y., Huang, G., Chen, X., & Bao, A. (2016). Assessment of parameter uncertainty in hydrological model using a Markov-Chain-Monte-Carlo-based multilevel-factorial-analysis method. *Journal of Hydrology*, 538, 471–486. <https://doi.org/10.1016/j.jhydrol.2016.04.044>

Sobre los autores

- **Javier Eduardo Bonilla Perdomo:** Ingeniero Agrícola, Máster en Ingeniería-Recursos Hidráulicos, Doctor en formación en Planificación y Manejo Ambiental de Cuencas Hidrográficas de Universidad del Tolima. Profesor Asistente de la Universidad Surcolombiana. Javier.bonilla@usco.edu.co
- **Miguel Barrios:** Ingeniero Forestal, Máster Ingeniería Hidráulica y Medio Ambiente, Doctor en Ingeniería del Agua y Medioambiental de Universidad Politécnica de Valencia. Profesor Asociado Universidad del Tolima. mibarrios@ut.edu.co



Los puntos de vista expresados en este artículo no reflejan necesariamente la opinión de la Asociación Colombiana de Facultades de Ingeniería.

Copyright © 2025 Asociación Colombiana de Facultades de Ingeniería (ACOFI)