

# Estudio comparativo de modelos de aprendizaje automático para la predicción de precipitaciones en contextos andinos

*Comparative study of machine learning models for precipitation prediction in andean contexts*

<https://doi.org/10.5281/zenodo.20291626>

## AUTORES:

Daniel S. Paredes <sup>1</sup>

Johanna Dueñas <sup>2</sup>

Joel Fernando Montero Jiménez<sup>3</sup>

José Luis Palacios Vélez<sup>4</sup>

**DIRECCIÓN PARA CORRESPONDENCIA:** [daniel.paredes@ueb.edu.ec](mailto:daniel.paredes@ueb.edu.ec)

**Fecha de recepción:** 01 / 11 / 2025

**Fecha de aceptación:** 17 / 11 / 2025

## RESUMEN

El presente estudio propone una comparación de 4 modelos de aprendizaje automático como: XGBoost, Random Forest, Long Short-Term Memory y Temporal Convolutional Network, se evaluó el desempeño integrando métricas de regresión como MAE, RMSE y R2 para cuantificar la magnitud del error, y métricas de clasificación binaria como accuracy, precision, recall y F1-score para medir la capacidad del modelo de discriminar entre eventos de lluvia y no lluvia. Cabe indicar que estos modelos se entrenaron sobre los datos previamente procesados, utilizando validación cruzada y búsqueda de hiperparámetros para optimizar su desempeño. En cuanto a los resultados evidencian que cada modelo posee fortalezas particulares: Random Forest se destaca por su equilibrio general, XGBoost por su sensibilidad extrema, Long Short-Term Memory por su precisión en predicciones continuas, y Temporal Convolutional Network por su estabilidad y capacidad de interpretar secuencias.

**Palabras Clave:** *Aprendizaje automático, aprendizaje profundo, precipitaciones, predicción.*

<sup>1</sup> <https://orcid.org/0009-0005-3591-7008>, Universidad Estatal de Bolívar, [daniel.paredes@ueb.edu.ec](mailto:daniel.paredes@ueb.edu.ec)

<sup>2</sup> <https://orcid.org/0009-0002-2442-314X>, Universidad Estatal de Bolívar, [johanna.duenas@ueb.edu.ec](mailto:johanna.duenas@ueb.edu.ec)

<sup>3</sup> <https://orcid.org/0009-0003-4174-3863>, Investigador independiente, [Joel\\_fer\\_95@hotmail.com](mailto:Joel_fer_95@hotmail.com)

<sup>4</sup> <https://orcid.org/0009-0009-6583-7403>, Investigador independiente, [Josepalaciosvelez98@gmail.com](mailto:Josepalaciosvelez98@gmail.com)

## ABSTRACT

This study proposes a comparison of four machine learning models: XGBoost, Random Forest, Long Short-Term Memory, and Temporal Convolutional Network. Performance was evaluated by integrating regression metrics such as MAE, RMSE, and R2 to quantify the magnitude of error, and binary classification metrics such as accuracy, precision, recall, and F1-score to measure the model's ability to discriminate between rain and non-rain events. It should be noted that these models were trained on previously processed data, using cross-validation and hyperparameter search to optimize their performance. The results show that each model has particular strengths: Random Forest stands out for its overall balance, XGBoost for its extreme sensitivity, Long Short-Term Memory for its accuracy in continuous predictions, and Temporal Convolutional Network for its stability and ability to interpret sequences.

*Keywords: Machine learning, deep learning, precipitation, prediction.*

## INTRODUCCIÓN

El estudio de las precipitaciones es un aspecto fundamental debido a que actúan como un factor desencadenante para la materialización de amenazas hidrometeorológicas como son las inundaciones y aluviones; así también amenazas geológicas como son los movimientos en masa. Asimismo, el impacto socio-económico que generan estos eventos puede ser significativo, afectando infraestructuras críticas, viviendas, actividades productivas y poniendo en riesgo la vida de las personas (Paredes et al., 2024). Es importante indicar estos eventos son cada vez más recurrente debido al cambio climático, Hartinger et al (2024) menciona que las políticas que permitan la adaptación al cambio climático aún son débiles en América Latina, es por ello que se deben fortalecer los procesos de gobernanza a nivel de cada país. Por otro lado, Paredes et al. (2025) menciona que en efecto la amenaza de inundación afecta a largo plazo la sostenibilidad de las distintas dinámicas territoriales, por ejemplo: a nivel de seguridad hídrica, infraestructuras, suministro de energía, entre otros elementos esenciales.

Santos et al., (2024) afirma que se proyecta un aumento de las precipitaciones anuales en todo Ecuador, con un énfasis notable en las estaciones húmedas aumentando el riesgo de inundaciones. En ese marco, es de especial relevancia el análisis datos mediante tecnologías emergentes que establezcan modelos predictivos enfocados en la reducción de riesgos de desastres. Es ahí en dónde cobra especial relevancia la disponibilidad de datos geoespaciales a nivel país y en específico que sea de uso gratuito. Sin embargo, en la actualidad a nivel de Ecuador existen limitantes económicas y burocráticas que no permiten el establecimiento de una red o plataforma de datos geoespaciales abierta para el público en

general. Torres-Díaz et al. (2025) menciona que uno de los desafíos principales a nivel de gestión de riesgos en el Ecuador es la dependencia de fuentes de datos obsoletas. Por otro lado, Wani et al. (2024) menciona que a pesar de que los métodos de aprendizaje automático o aprendizaje profundo ofrecen pronósticos precisos, se requiere demandas computacionales mayores y en países con menor inversión tecnológica con es el caso de Ecuador puede resultar un reto mayor.

Tomando en consideración que estos eventos con cada vez más recurrentes, resulta importante el análisis de los datos de precipitaciones que faciliten su predicción. Si bien tradicionalmente se han utilizado para su análisis métodos estadísticos y modelos numéricos de pronóstico de tiempo, en la actualidad existen técnicas de predicción que utilizan grandes volúmenes de datos meteorológicos con la finalidad de detectar patrones complejos. En ese sentido Tricha & Moussaid (2024) afirma que el aprendizaje automático ha mejorado la precisión en la predicción de precipitaciones al captar relaciones complejas entre variables meteorológicas. An et al. (2025) hace alusión a que la predicción de series temporales mediante técnicas asociadas al aprendizaje profundo ha tomado especial relevancia en el campo del pronóstico de las precipitaciones en un corto período de tiempo, y en específico técnicas de *deep learning* que representan un hito significativo a nivel de procesamiento de grandes volúmenes de datos de precipitación.

En virtud de lo expuesto el presente estudio propone una comparación de 4 modelos de aprendizaje automático: XGBoost, Random Forest, Long Short-Term Memory (LSTM) y Temporal Convolutional Network (TCN), está evaluación de los modelos desarrollados para la predicción de precipitaciones trasciende la mera precisión de la predicción, integrando métricas tanto de regresión como: error absoluto medio (MAE), Raíz del error cuadrático medio (RMSE) y Coeficiente de determinación (R<sup>2</sup>) para cuantificar la magnitud del error. Además, de clasificación binaria como: Accuracy, Precision, Recall y F1-score, para evaluar la capacidad de distinguir entre eventos de lluvia y no lluvia. De esta manera se integra un enfoque dual que proporciona una visión más completa del rendimiento del modelo en el contexto específico de la predicción meteorológica. De acuerdo con lo señalado Kagabo et al. (2024) menciona en su estudio que se ha demostrado con éxito la eficacia del aprendizaje automático, en particular el LSTM, a la hora de predecir precipitaciones extremas y a que su ves este tipo de avances puede reducir significativamente las pérdidas sociales, ambientales y económicos derivadas de este tipo de eventos. Por otro lado, El Hafyani et al. (2024) en su estudio desde un enfoque específico asociado a la precipitación analizó eficacia de modelos como: Decision Tree, Random Forest, K-Nearest Neighbors (KNN), AdaBoost, XGBoost, y LSTM, en dónde XGBoost mostró tener un mejor rendimiento en comparación con los otros.

Desde este enfoque al emplear algoritmos y metodologías estadísticas avanzadas, los sistemas son capaces de discernir patrones, generar predicciones y ejecutar decisiones independientemente de la supervisión humana explícita. En ese sentido el estudio, Kala et al. (2023) reporta que la aplicación de técnicas de aprendizaje automático, en particular la regresión basada en el árbol de decisiones acelerado (BDTR), ha mostrado mejoras significativas en la predicción de las precipitaciones, y superando métodos tradicionales.

Distintos autores mencionan que entre los aspectos fundamentales a tomarse en cuenta para el análisis de datos precipitaciones es pertinente escoger el modelo de aprendizaje automática adecuado, esto hace alusión al que tenga un mejor rendimiento en torno a las métricas de regresión, entre las que tenemos: MAE, RMSE y R2 Score. Además, de las métricas de clasificación: Accuracy, Precision, Recall y F1-score. Es importante indicar que Geng (2024) cita que, el uso de varias métricas al mismo tiempo podría proporcionar una imagen más completa del rendimiento de un modelo en un enfoque, conocido como evaluación multimétrica. Por otro lado, Strauss et al. (2023) afirma que las métricas se eligen en función de los requisitos y restricciones específicos del problema en cuestión, como el equilibrio del conjunto de datos y la importancia relativa de los falsos positivos frente a los falsos negativos.

## **METODOLOGIA**

### **3.1. Carga y preparación de datos**

La metodología se inicia con la carga y preparación de datos multivariados obtenidos de seis estaciones meteorológicas del Distrito Metropolitano de Quito. Cada conjunto incluye, el valor de precipitación con su respectiva unidad de tiempo. Se realiza un proceso de limpieza para manejar valores faltantes y se organiza la información en formato de series temporales, estandarizando las variables mediante técnicas de normalización o escalado, asegurando su compatibilidad con los modelos de aprendizaje automático y profundo. Además, se estructuran ventanas temporales (sliding windows) para capturar las dependencias temporales de los datos.

### **3.2. Función de evaluación de modelos**

Para garantizar una evaluación coherente y homogénea de todos los modelos aplicados, se definió una función que calcula métricas tanto de regresión como de clasificación binaria. Las métricas de regresión incluyen MAE, RMSE y R2, mientras que para la clasificación de lluvia (valor mayor a cero) o no lluvia (cero), se emplean métricas como accuracy, precisión, recall y F1-score. Además, se

generan visualizaciones como la gráfica de valores reales vs predichos y la matriz de confusión para cada modelo.

### **3.3. Modelos de aprendizaje automático**

Se implementaron varios modelos de aprendizaje automático tradicionales para establecer una línea base comparativa. Entre estos modelos se encuentran el algoritmo de Random Forest, el modelo XGBoost, LSTM y TCN. Estos modelos se entrenaron sobre los datos previamente procesados, utilizando validación cruzada y búsqueda de hiperparámetros para optimizar su desempeño. Las predicciones generadas fueron posteriormente evaluadas con la función previamente definida.

### **3.4. Preprocesamiento para redes neuronales**

Dado que las redes neuronales requieren estructuras de entrada específicas, se implementó un proceso de preprocesamiento adicional. Esto incluye la transformación de las series temporales en tensores tridimensionales con la forma (muestras, ventanas temporales, características), asegurando la correcta secuencialidad de los datos. Asimismo, los conjuntos de entrenamiento y prueba se dividieron respetando la temporalidad de las observaciones para evitar fugas de información.

### **3.5. Modelos de redes neuronales**

Se desarrollaron modelos de redes neuronales recurrentes como LSTM, además de modelos basados en TCN y Transformers. Estos modelos fueron diseñados para capturar las dependencias temporales a corto y largo plazo presentes en los datos meteorológicos. La arquitectura, número de capas, neuronas y funciones de activación fueron definidas tras un proceso iterativo de ajuste, y se utilizó early stopping para prevenir el sobreajuste.

### **3.6. Comparación final de métricas**

Finalmente, se consolidaron y compararon todas las métricas obtenidas de los modelos de aprendizaje automático, redes neuronales profundas y métricas. Esta comparación permite identificar qué enfoque proporciona el mejor rendimiento tanto en tareas de regresión como en clasificación binaria de precipitaciones. Los resultados fueron representados en tablas y gráficos comparativos, lo que facilitó la interpretación de fortalezas y debilidades de cada modelo evaluado.

## RESULTADOS

La presente sección detalla los diferentes resultados obtenidos según los métodos aplicados entre los modelos Random Forest, LSTM, TCN y XGBoost. En la tabla 1 se indican las diferentes métricas de evaluación de los cuatro modelos de clasificación para la predicción de precipitaciones, tomando en cuenta cuatro indicadores principales: Puntuación F1, Recall, Exactitud y Precisión.

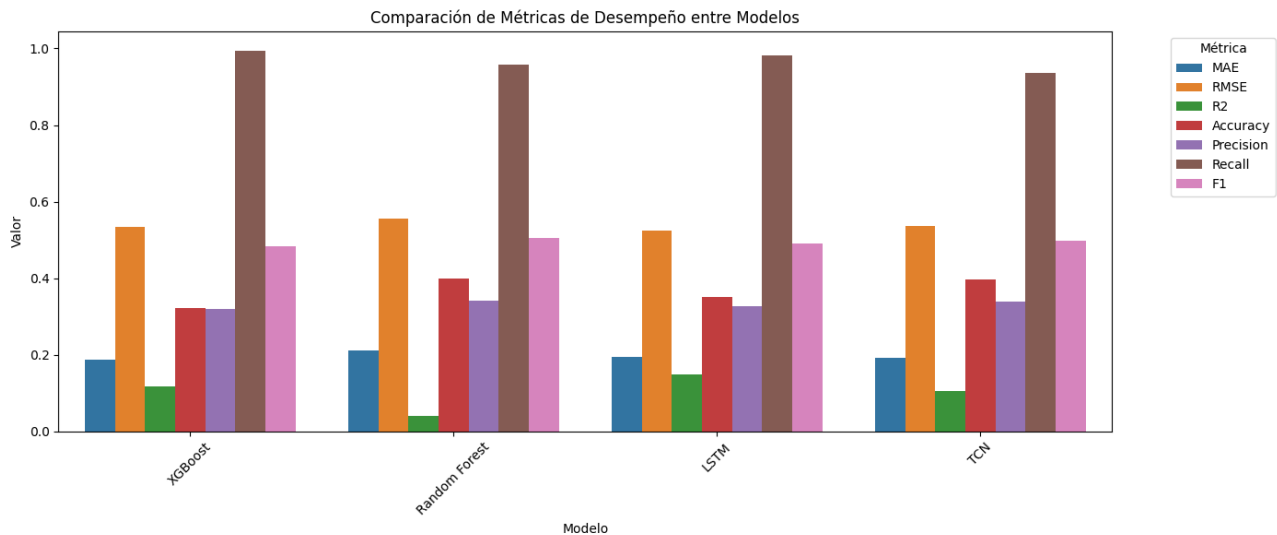
**Tabla 1.** Resultados de las métricas de los cuatro modelos usados.

N	Modelo	Precisión	Exactitud	Recall	Puntuación F1	RMSE	MAE
1	XGBoost	31.95	32.25	99.51	48.37	53.33	18.77
2	Random Forest	34.20	39.92	95.68	50.39	55.59	21.11
3	LSTM	32.89	35.74	97.62	49.20	52.10	18.75
4	TCN	33.90	39.79	93.53	49.76	53.69	19.31

A partir de los resultados obtenidos en los diferentes modelos de predicción de lluvia, se puede observar que el modelo Random Forest en la predicción de lluvia, se destaca su buen desempeño en clasificación, con una precisión del 34.20%, exactitud del 39.92%, recall del 95.68% y una puntuación F1 de 50.39%, lo que indica que el modelo logra detectar adecuadamente la mayoría de los eventos de lluvia, manteniendo al mismo tiempo un mejor control de los falsos positivos en comparación con otros modelos. En cuanto a su desempeño en regresión, el MAE fue de 21.11, lo que significa que, en promedio, las predicciones del modelo se desvían en 21.11 unidades respecto a los valores reales. Por otro lado, el RMSE fue de 55.59, indicando que, si bien la magnitud promedio del error es aceptable, existen algunos errores grandes que incrementan la penalización, ya que el RMSE da mayor peso a desviaciones extremas. Estos resultados sugieren que Random Forest no solo logra un equilibrio eficaz en clasificación, sino que mantiene un nivel de error razonable en sus predicciones cuantitativas, haciéndolo un modelo robusto para este tipo de análisis. Le sigue el modelo TCN, con una precisión de 33.90%, exactitud de 39.79% y una puntuación F1 de 49.76%, además de un MAE de 19.31 y un RMSE de 53.69, que reflejan errores moderados en la magnitud de las predicciones. El modelo LSTM mantiene una alta capacidad de detección (recall de 97.62%) y errores ligeramente más bajos en regresión (MAE de 18.75, RMSE de 52.10), aunque con menor exactitud que TCN. Por último, el modelo XGBoost, si bien alcanza el mayor recall (99.51%), muestra la menor precisión (31.95%) y exactitud (32.25%), lo que indica una fuerte tendencia a generar falsas alarmas, aunque con un MAE competitivo de 18.77. En conjunto, los resultados sugieren que Random Forest es el más equili-brado en términos de clasificación, mientras que LSTM y XGBoost muestran un mejor comportamiento en

magnitud de error, siendo adecuados dependiendo de si se prioriza la detección o la precisión del evento.

**Figura 1.** Comparación de Métricas de Desempeño.



En la figura 1 se aprecia la evaluación comparativa del desempeño de los modelos XGBoost, Random Forest, LSTM y TCN, el modelo Random Forest se posiciona como el de mejor rendimiento general en tareas de clasificación, al obtener los valores más altos en precisión (34.20 %), exactitud (39.92 %) y puntuación F1 (50.39 %). Estas métricas indican una mayor proporción de predicciones correctas y un equilibrio adecuado entre la detección de verdaderos positivos y la minimización de falsos positivos. Aunque XGBoost alcanza el mayor valor de recall (99.51 %), y LSTM reporta los menores errores de predicción en términos de RMSE (52.10) y MAE (18.75), Random Forest logra un desempeño más consistente y robusto en los indicadores clave de clasificación. Por tanto, se considera el modelo más adecuado para contextos donde se requiere alta precisión y equilibrio en la clasificación, sin comprometer significativamente la capacidad predictiva.

Los resultados muestran que los modelos de redes neuronales como LSTM y TCN tienen una mayor capacidad para capturar la dinámica temporal de las precipitaciones en comparación con los modelos de árbol como XGBoost y Random Forest, como se puede ver en las gráficas de "Real vs. Predicho". Pero cuando se trata de métricas de error globales como MAE y RMSE, XGBoost tiene un rendimiento leve-mente mejor. La discrepancia entre el buen rendimiento de XGBoost en métricas de error y el mejor desempeño de LSTM y TCN en la captura de patrones temporales y clasificación de eventos de lluvia sugiere que la elección del modelo óptimo depende de los objetivos específicos del estudio.

Si el enfoque principal es minimizar el error global en la predicción de la cantidad de precipitación, XGBoost podría ser preferible. Sin embargo, si la detección precisa de los eventos de lluvia es primordial, LSTM o TCN pueden ser más adecuados. El bajo rendimiento promedio de Random Forest indica que este modelo potencialmente no es la mejor opción para predecir precipitaciones en este conjunto de datos. Su incapacidad para absorber la complejidad temporal y su menor precisión en la clasificación de eventos de lluvia lo hacen menos competitivo en comparación con los otros modelos mantenidos dentro de la comparación.

En general, la comparación de los cuatro modelos indica que cada uno de ellos tiene puntos fuertes y débiles. LSTM y TCN son más aptos para capturar la naturaleza secuencial de los datos de precipitación, mientras que XGBoost es mejor en minimizar el error global. En relación a lo dicho, Waqas & Humphries (2024) afirma que los modelos LSTM se han convertido en puntos focales debido a la efectividad que logran en el modelado de sistemas hidrológicos no lineales y de variación de tiempo. A su vez Qin et al. (2024) plantea que los modelos TCN modelan eficazmente características temporales cruzadas, y su entrenamiento es mucho más rápido que un modelo LSTM. Por otro lado, en la investigación de Wani et al. (2024a) se menciona que los enfoques de aprendizaje profundo y en particular LSTM superan sistemáticamente a modelos de aprendizaje automático como RF.

El aprendizaje automático y profundo se encuentra en un desarrollo vertiginoso y su aplicación a nivel de procesamiento de datos climáticos como es la precipitación es indudable, y un aspecto a mejorar es el realizar estudios comparativos de los diversos modelos en dónde se incluya un mayor número de estaciones meteorológicas combinando con datos satelitales de alta resolución con alta resolución espacial (He et al., 2023; Kossieris et al., 2024).

## **DISCUSION**

En el estudio se analizó el desempeño de cuatro modelos de aprendizaje automático para la predicción de lluvias, observándose que no existe una solución única que domine todas las métricas, sino que cada algoritmo presenta ventajas específicas según el objetivo de la predicción. Al evaluar la clasificación de eventos, se determinó que el modelo Random Forest es el más equilibrado, ya que alcanzó los niveles más altos de precisión (34.20%), exactitud (39.92%) y puntuación F1 (50.39%), lo que refleja un control superior sobre los falsos positivos. Por el contrario, aunque con XGBoost se obtuvo la mayor sensibilidad o recall (99.51%), se notó que su baja precisión conlleva una tendencia significativa a generar "falsas alarmas" en la detección de lluvia.

En lo que respecta a la predicción cuantitativa o regresión, los hallazgos sugieren que las redes neuronales son más eficaces para capturar la naturaleza secuencial y temporal de los datos meteorológicos. Específicamente, el modelo LSTM reportó el mejor desempeño en esta categoría, con los errores más bajos del estudio, registrando un MAE de 18.75 y un RMSE de 52.10. Asimismo, se comprobó que el modelo TCN es una alternativa competitiva y estable, que se destaca por su capacidad para modelar características temporales complejas con una mayor rapidez de entrenamiento.

## CONCLUSIONES

El estudio comparativo de modelos de predicción de lluvia manifiesta que Random Forest es la mejor opción en términos de desempeño general, con un F1 Score de 50.39%, debido a su adecuado balance entre precisión (34.20%) y recall (95.68%), lo que le permite detectar eventos de lluvia con alta sensibilidad y reducir falsas alarmas. Aunque sus métricas de regresión (MAE: 21.11; RMSE: 55.59) no fueron las más bajas, se consideran aceptables, destacando su robustez en aplicaciones operativas como sistemas de alerta temprana. Por su parte, XGBoost mostró un recall sobresaliente de 99.51%, útil para la detección exhaustiva de eventos, pero su baja precisión (31.95%) genera muchas falsas alarmas. En contraste, el modelo LSTM tuvo el mejor desempeño en regresión (MAE: 18.75; RMSE: 52.10), con una buena capacidad clasificatoria (F1 Score: 49.20%), siendo útil para estimaciones numéricas precisas en problemas de series temporales.

El modelo TCN también mostró un rendimiento competitivo, con métricas cercanas a Random Forest (F1 Score: 49.76%) y un balance entre precisión y recall, siendo eficaz en entornos con patrones temporales persistentes. En términos generales, los resultados evidencian que cada modelo posee fortalezas particulares: Random Forest se destaca por su equilibrio general, XGBoost por su sensibilidad extrema, LSTM por su precisión en predicciones continuas, y TCN por su estabilidad y capacidad de interpretar secuencias. Se propone, para trabajos futuros, optimizar hiperparámetros y explorar técnicas de ensemble learning para integrar las fortalezas de estos algoritmos, especialmente combinando la baja tasa de error de XGBoost con la capacidad de detección de LSTM, para construir un modelo híbrido más robusto.

Del mismo modo que la mejora del modelado, se propone extender el conjunto de datos con variables exógenas que sean pertinentes para la predicción de lluvias. Se podrían tomar parámetros tales como la humedad relativa, presión atmosférica, temperatura, velocidad y dirección del viento, y topografía local, podrían aportar información útil que refuerce la capacidad predictiva de los modelos. La adición de datos espaciales, tales como imágenes de satélite o datos de radar meteorológico, podría también

adicionar riqueza al análisis y hacer posible la gene-ración de predicciones de precipitación distribuidas en el espacio. Por último, se propone evaluar estos modelos en distintas regiones geográficas y escalas temporales para probar su generalización y flexibilidad para adaptarse a diferentes escenarios climáticos.

## REFERENCIAS BIBLIOGRÁFICAS

- An, S., Oh, T.-J., Sohn, E., & Kim, D. (2025). Deep learning for precipitation nowcasting: A survey from the perspective of time series forecasting. *Expert Systems with Applications*, 268, 126301. <https://doi.org/10.1016/j.eswa.2024.126301>
- El Hafyani, M., El Himdi, K., & El Adlouni, S.-E. (2024). Improving monthly precipitation prediction accuracy using machine learning models: A multi-view stacking learning technique. *Frontiers in Water*, 6. <https://doi.org/10.3389/frwa.2024.1378598>
- Geng, S. (2024). Analysis of the Different Statistical Metrics in Machine Learning. *Highlights in Science, Engineering and Technology*, 88, 350-356. <https://doi.org/10.54097/jhq3tv19>
- Hartinger, S. M., Palmeiro-Silva, Y. K., Llerena-Cayo, C., Blanco-Villafuerte, L., Escobar, L. E., Diaz, A., Sarmiento, J. H., Lescano, A. G., Melo, O., Rojas-Rueda, D., Takahashi, B., Callaghan, M., Chesini, F., Dasgupta, S., Posse, C. G., Gouveia, N., Martins de Carvalho, A., Miranda-Chacón, Z., Mohajeri, N., ... Romanello, M. (2024). The 2023 Latin America report of the Lancet Countdown on health and climate change: The imperative for health-centred climate-resilient development. *Lancet Regional Health - Americas*, 33, 100746. <https://doi.org/10.1016/j.lana.2024.100746>
- He, K., Zhao, W., Brocca, L., & Quintana-Seguí, P. (2023). SMPD: A soil moisture-based precipitation downscaling method for high-resolution daily satellite precipitation estimation. *Hydrology and Earth System Sciences*, 27(1), 169-190. <https://doi.org/10.5194/hess-27-169-2023>
- Kagabo, J., Kattel, G. R., Kazora, J., Shangwe, C. N., & Habiyakare, F. (2024). Application of Machine Learning Algorithms in Predicting Extreme Rainfall Events in Rwanda. *Atmosphere*, 15(6), Article 6. <https://doi.org/10.3390/atmos15060691>
- Kala, K., Vijayakumar, N., Poovizhi, P., & Selvapandian, D. (2023). An Efficient Rainfall Forecasting System using Machine Learning Methods. *International Conference on Electrical Engineering and Information Communication Technology*, 1(6). <https://doi.org/10.1109/ICEEICT56924.2023.10157395>

- Kossieris, P., Tsoukalas, I., Brocca, L., Mosaffa, H., Makropoulos, C., & Angehelea, A. (2024). Precipitation data merging via machine learning: Revisiting conceptual and technical aspects. *Journal of Hydrology*. <https://doi.org/10.1016/j.jhydrol.2024.131424>
- Paredes, D. S., Jiménez, J. F. M., & Vélez, J. L. P. (2024). Análisis de la vulnerabilidad de infraestructuras críticas como represas, puentes y carreteras ante eventos asociados al fenómeno de: El Niño – Oscilación Sur en la Costa ecuatoriana. *AlfaPublicaciones*, 6(1.1), 45-61. <https://doi.org/10.33262/ap.v6i1.1.454>
- Paredes, D. S., Rivera, E. F., Baldeón-Egas, P., & Toasa, R. M. (2025). The Multiscale Assessment of Infrastructure Vulnerability to River Floods in Andean Areas: A Case Study of the Chibunga River in the Parish of San Luis, Ecuador. *Sustainability*, 17(17), 7915. <https://doi.org/10.3390/su17177915>
- Qin, Z., Zhang, Y., Qin, H., Mo, L., Ren, P., & Zhu, S. (2024). A Temporal Convolutional Neural Network Fusion Attention Mechanism Runoff Prediction Model Based on Dynamic Decomposition Reconstruction Integration Processing. *Water*, 16(23), 3515. <https://doi.org/10.3390/w16233515>
- Santos, F., Jara, J., Acosta, N., Galeas, R., & De Bièvre, B. (2024). Assessing Annual and Monthly Precipitation Anomalies in Ecuador Bioregions Using WorldClim CMIP6 GCM Ensemble Projections and Dynamic Time Warping. *International Journal of Climatology*, 45, n/a-n/a. <https://doi.org/10.1002/joc.8685>
- Strauss, E., Villas Boas, M., & Ferreira, W. L. (2023). The importance of using appropriate performance evaluation metrics in predictive machine learning models. *Projectus*, 7(2), 52-62. <https://doi.org/10.15202/25254146.2022v7n2p52>
- Torres-Díaz, V., del Río-Rama, M. de la C., Álvarez-García, J., & Venegas-Martínez, F. (2025). Spatial Inequalities and the Sensitivity of Social Vulnerability in Ecuador. *Land*, 14(11), 2110. <https://doi.org/10.3390/land14112110>
- Tricha, A., & Moussaid, L. (2024). Evaluating machine learning models for precipitation prediction in Casablanca City. *Indonesian Journal of Electrical Engineering and Computer Science*. <https://doi.org/10.11591/ijeecs.v35.i2.pp1325-1332>
- Wani, O. A., Mahdi, S. S., Yeasin, M., Kumar, S. S., Gagnon, A. S., Danish, F., Al-Ansari, N., El-Hendawy, S., & Mattar, M. A. (2024). Predicting rainfall using machine learning, deep

learning, and time series models across an altitudinal gradient in the North-Western Himalayas.

*Scientific Reports*, 14(1), 27876. <https://doi.org/10.1038/s41598-024-77687-x>

Waqas, M., & Humphries, U. W. (2024). A critical review of RNN and LSTM variants in hydrological time series predictions. *MethodsX*, 13, 102946. <https://doi.org/10.1016/j.mex.2024.102946>