

Artículo

Aplicación de Redes Neuronales Artificiales para la Estimación de Precipitaciones: Caso de Estudio de la Cuenca del Río Pastaza, Ecuador

Alexander Rogel^[1]  Andrés Hidalgo^[1]  Fidel Castro^[1]  Fabián Morales^[1]  Dilon Moya^[1]  Bolívar Paredes^[1] 

[1] Universidad Técnica de Ambato, Facultad de Ingeniería Civil y Mecánica, Carrera de Ingeniería Civil, Ambato 180206, Ecuador



Autor para correspondencia: be.paredes@uta.edu.ec

Resumen

La cuenca del Río Pastaza en Ecuador, crucial por su biodiversidad y gestión hídrica, enfrenta desafíos significativos debido al cambio climático. Este estudio presenta la aplicación de redes neuronales artificiales (RNA) para abordar las deficiencias en los datos pluviométricos de esta cuenca. Implementando un modelo optimizado con 5000 iteraciones, se logró una fiabilidad del 95% en la estimación de datos pluviométricos. Se analizaron datos de múltiples estaciones meteorológicas, ajustando el modelo según las distancias entre estaciones, y se demostró una mejora en precisión y coherencia en comparación con métodos tradicionales. Los resultados destacan la capacidad de las RNA para adaptarse a variaciones significativas en los datos, mejorando la planificación hídrica y mitigando los efectos de eventos climáticos extremos mediante una mejor predicción pluviométrica. La capacidad de las RNA para procesar grandes volúmenes de datos con complejas interacciones es particularmente relevante en el campo hidrometeorológico, donde la variabilidad espacial y temporal de los datos es sustancial. Este avance evidencia la aplicabilidad de las RNA en hidrología y climatología, contribuyendo al entendimiento de la variabilidad climática de la región. La integración de técnicas avanzadas de inteligencia artificial en la estimación y homogeneización de datos hidrológicos proporciona una base sólida para desarrollar estrategias de adaptación y mitigación más efectivas frente al cambio climático. A medida que la tecnología evoluciona, se abren nuevas perspectivas para aplicar técnicas similares en otras cuencas hidrográficas de la región, mejorando la gestión de los recursos hídricos en Ecuador.

Palabras Clave: *redes neuronales artificiales; estimación de precipitaciones; gestión hídrica; cuenca del río Pastaza*

Application of Artificial Neural Networks for Rainfall Estimation: Case Study in the Pastaza River Basin, Ecuador

Abstract

The Pastaza River basin in Ecuador, crucial for its biodiversity and water management, faces significant challenges due to climate change. This study presents the application of artificial neural networks (ANN) to address deficiencies in pluviometric data for this basin. By implementing an optimised model with 5000 iterations, a 95% reliability in precipitation data estimation was achieved. Data from multiple meteorological stations were analysed, adjusting the model based on the distances between stations, demonstrating improved accuracy and coherence compared to traditional methods. The results highlight the ANN's capability to adapt to significant data variations, enhancing water management planning and mitigating the effects of extreme weather events through better precipitation prediction. The ability of ANN to process large volumes of data with complex interactions is particularly relevant in the hydrometeorological field, where spatial and temporal data variability is substantial. This advancement demonstrates the applicability of ANN in hydrology and climatology, contributing to the understanding of regional climate variability. The integration of advanced artificial intelligence techniques in the estimation and homogenization of hydrological data provides a solid foundation for developing more effective adaptation and mitigation strategies in response to climate change. As technology evolves, new perspectives emerge for applying similar techniques in other river basins in the region, improving the management of water resources in Ecuador.

Keywords: *artificial neural networks; rainfall estimation; water management; pastaza river basin*

1. Introducción

En un contexto global cada vez más influenciado por el cambio climático, la precisión en la recopilación de datos meteorológicos es fundamental para una gestión eficaz de los recursos hídricos. Las transformaciones climáticas se manifiestan intensamente en regiones sensibles como la cuenca del Río Pastaza, rica en biodiversidad y vital para el equilibrio ecológico del Ecuador. Este entorno requiere una monitorización detallada para enfrentar los desafíos climáticos y asegurar un manejo sostenible y justo de sus recursos hídricos (Flores Llampazo et al., 2022; López & Sierra, 2010). Además, es crucial entender no solo la dinámica actual de los recursos hídricos, sino también anticipar futuros escenarios. En regiones como la cuenca del Río Pastaza, la variabilidad temporal de los fenómenos meteorológicos podría aumentar, lo que exige estrategias proactivas basadas en datos precisos para fortalecer las tácticas de gestión hídrica frente a la variabilidad climática (Serrano-Vincenti et al., 2022).

La cuenca del Río Pastaza, situada entre la región Interandina y la Amazonía de Ecuador, desempeña un papel crucial en la geografía y ecología del país, siendo uno de los principales sistemas fluviales que sustentan tanto a grandes comunidades humanas como a ecosistemas clave a nivel regional (Cruzatty et al., 2021). La interacción entre dinámicas naturales y antropogénicas ha fortalecido la importancia socioeconómica y ambiental de la cuenca. Su extenso recorrido por diversas regiones geográficas le confiere una biodiversidad única, estableciéndola como un recurso hídrico esencial para actividades agrícolas, industriales y domésticas. La cuenca se caracteriza por una diversidad climática que abarca zonas subtropicales, templadas y altoandinas, lo cual contribuye a una rica biodiversidad de flora y fauna, destacándola como un área de significativa importancia ecológica y económica. Además, es responsable del 45% de la producción nacional de hortalizas, subrayando su papel vital en la economía y sustento alimenticio del país. Su confluencia en el río Marañón, principal afluente del Amazonas, refuerza su importancia como una de las arterias hídricas más críticas a nivel global (López & Sierra, 2010; Siren, 2014). Sin embargo, a pesar de su potencial como una de las mayores reservas de agua dulce del mundo y su capacidad para la generación de energía, la cuenca enfrenta desafíos significativos que comprometen su gestión y conservación (Abril-Saltos et al., 2021; Herrmann, 2002; López & Sierra, 2010). Estos desafíos se deben a factores como la alteración del uso del suelo, la variabilidad meteorológica y el cambio climático, que se manifiestan en alteraciones de los patrones de precipitación y eventos extremos, comprometiendo la capacidad del ecosistema para seguir funcionando como fuente primordial de agua para la región (Carrera-Villacrés et al., 2016; Cruzatty et al., 2021; Serrano-Vincenti et al., 2022).

Las redes neuronales artificiales (RNA), inspiradas en la estructura y función de las neuronas biológicas, se han consolidado como herramientas poderosas en diversos ámbitos, incluyendo la predicción de series temporales y el reconocimiento de patrones (Goos et al., 2012; Lee & Kang, 2015). Estos modelos computacionales aprenden y generalizan a partir de los datos, adaptándose y mejorando su desempeño con cada nueva información. Gracias a su capacidad para modelar relaciones no lineales y procesar grandes volúmenes de datos, las RNA se han convertido en una alternativa prometedora a los métodos tradicionales en las ciencias atmosféricas e hidrológicas, permitiendo capturar complejas interacciones sin necesidad de un conocimiento exhaustivo de los procesos físicos subyacentes (Gómez Guerrero & Aguayo Arias, 2019; Najarchi et al., 2020). En el ámbito hidrometeorológico, las RNA podrían ofrecer soluciones innovadoras para la modelización y predicción de variables clave, desempeñando un rol crucial en un mundo donde el cambio climático presenta desafíos sin precedentes para la gestión de los recursos hídricos. En Ecuador, un desafío notable en la gestión hídrica es la calidad y precisión de los datos meteorológicos recolectados, que a menudo presentan deficiencias debido a factores como mediciones incorrectas, falta de presupuestos adecuados, variabilidad de equipos de medición o cambios inesperados en el entorno. Estas inconsistencias representan un riesgo para la toma de decisiones y la planificación hídrica basadas en esta información. Por ello, es crucial someter cualquier serie de datos pluviométricos a un proceso de homogeneización riguroso, asegurando que los datos se ajusten a estándares que permitan su uso efectivo en Investigación, toma de decisiones y otras aplicaciones prácticas. Mientras que los métodos tradicionales, como la regresión lineal y el cálculo de razones de distancia, han sido la norma, el surgimiento del aprendizaje automático (Machine Learning) ofrece nuevas posibilidades. En particular, las redes neuronales artificiales modulares (RNAM) han mostrado ventajas significativas sobre métodos convencionales en la simulación y corrección de datos hidrometeorológicos, indicando un camino prometedor para superar las limitaciones presentes en diversas cuencas hídricas (Gómez Guerrero & Aguayo Arias, 2019; Kalauzi et al., 2009; Matovelle et al., 2022; Millán et al., 2008, 2010).

Numerosos estudios han abordado los desafíos hídricos en regiones críticas, incluida la cuenca del Río Pastaza. Sin embargo, la acelerada dinámica del cambio climático y su impacto en este sistema fluvial exigen soluciones más avanzadas y precisas. En este contexto, las técnicas contemporáneas, como las RNA, se perfilan como herramientas valiosas, mostrando resultados alentadores en estudios hidrometeorológicos a nivel global. La implementación de estas metodologías en la cuenca del Río Pastaza podría representar un avance significativo en la predicción y gestión de sus recursos hídricos, dada la

complejidad de los factores que definen su dinámica. Este estudio tiene como objetivo evaluar la efectividad de las RNA en la estimación de datos pluviométricos faltantes en la cuenca del Río Pastaza, proporcionando una base más confiable para la gestión de recursos hídricos. La innovación de este trabajo radica en la aplicación específica y detallada de un modelo optimizado, que a través de 5000 iteraciones, ha logrado una fiabilidad del 95% en la estimación de datos pluviométricos. Esta metodología combina la precisión de las RNA con técnicas de validación cruzada, proporcionando una comparación detallada con métodos tradicionales y estableciendo un estándar de referencia para futuras investigaciones. Motivado por la necesidad de proteger y conservar uno de los recursos hídricos más cruciales de Ecuador, este estudio busca contribuir al avance científico y ofrecer soluciones prácticas y efectivas a corto y medio plazo. Con este trabajo, esperamos aportar a la construcción de un futuro más sostenible y resiliente para la cuenca del Río Pastaza y, por extensión, para las comunidades y ecosistemas que dependen de ella.

2. Material y métodos

2.1. Área de estudio

El área de estudio es la cuenca del Río Pastaza (Figura 1), ubicada entre la región oriental de los Andes y la llanura amazónica en Ecuador, extendiéndose sobre un área total de 32,128 km². La información geográfica detallada, provista

por el Ministerio de Agricultura y Ganadería del Ecuador (MAGAP), revela una diversidad en el uso del suelo dentro de la cuenca: 9.5% de áreas de cultivo, 7.4% de pastizales, 66.8% de tierras forestales, 0.2% de tierras urbanas, 2.0% de zonas acuáticas y 14.1% de otras categorías. Esta variabilidad en el uso del suelo refleja una compleja interacción entre los sistemas naturales y las actividades humanas, ofreciendo una plataforma única para estudiar las dinámicas hídricas y terrestres en un entorno diverso y ecológicamente rico. Los patrones agrícolas, según el III Censo Nacional Agropecuario del año 2000, destacan una diversidad significativa en los tipos de cultivo de la cuenca.

Los cultivos permanentes y transitorios abarcan vastas áreas, reflejando una estructura agrícola diversificada influenciada por la geografía de la cuenca, desde las regiones andinas hasta las llanuras amazónicas (Herrmann, 2002; López & Sierra, 2010). Asimismo, la infraestructura hidráulica de la cuenca, que incluye embalses y zonas tecnificadas de riego, juega un papel crucial en la gestión del agua, con instalaciones hidroeléctricas destacadas como las presas de Agoyán y Pisayambo, que satisfacen gran parte de los requerimientos eléctricos de la región central del Ecuador, y las presas de Mulacorral y Chiquihurco, que proveen servicios invaluable para el abastecimiento de agua potable y riego agrícola, especialmente en la provincia de Tungurahua (Kalauzi et al., 2009; López & Sierra, 2010).

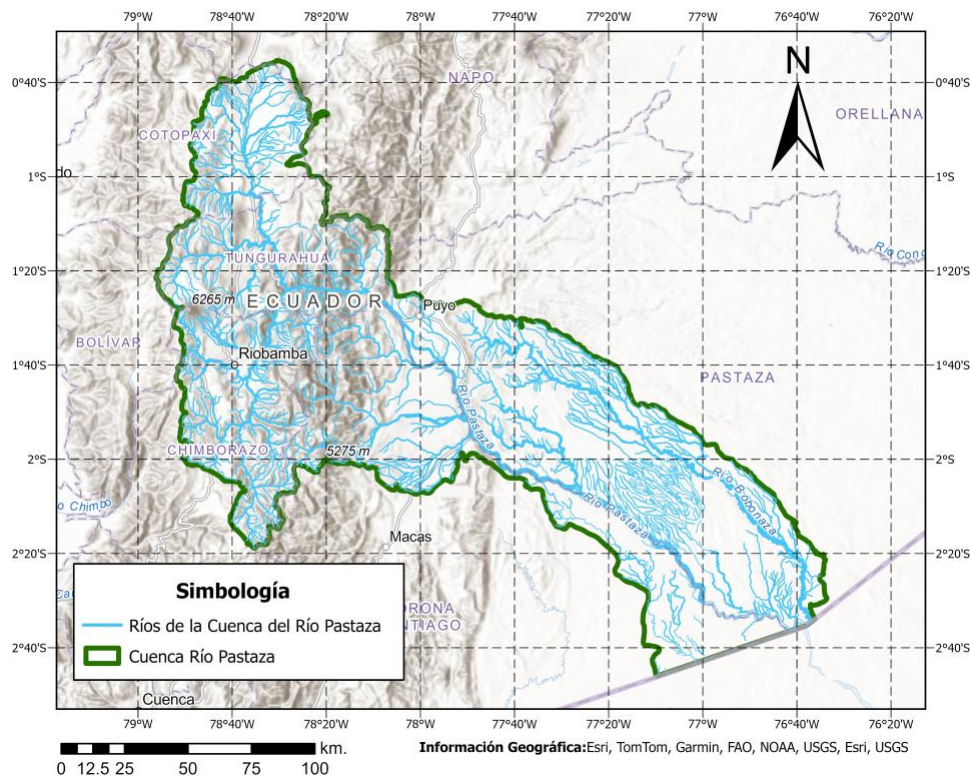


Figura 1: Ubicación de la Cuenca del Río Pastaza, Ecuador.

2.2. Metodología

Este estudio se centró en el empleo de RNA para la estimación y homogeneización de datos pluviométricos faltantes. La metodología implementada comenzó con la recolección meticulosa de datos históricos pluviométricos desde enero de 1995 hasta diciembre de 2013, cubriendo un total de 29 estaciones hidrológicas en la cuenca, obtenidos de anuarios del Instituto Nacional de Meteorología e Hidrología (INHAMI, 2024). Tras la recolección, se llevó a cabo un detallado proceso de preparación de datos que incluyó limpieza, normalización y análisis preliminar para identificar patrones y anomalías. Los datos fueron introducidos en un modelo de RNA configurado para reconocer y aprender de las complejidades de los patrones de precipitación a través de un proceso iterativo y adaptativo. Esta fase de preparación y análisis se apoyó en software avanzado y librerías de Python (Van Rossum, 1995) especializadas en aprendizaje automático como Scikit-learn (Pedregosa et al., 2011), TensorFlow (Abadi et al., 2016) y Keras (Chollet, 2015), garantizando un tratamiento eficaz y riguroso de los datos. La estructura del modelo de RNA se diseñó siguiendo el esquema de un perceptrón multicapa, adecuado para manejar información no lineal característica en datos hidrometeorológicos. El entrenamiento del modelo se realizó mediante múltiples iteraciones, ajustando los parámetros y afinando la precisión hasta alcanzar una convergencia óptima, evidenciada por la significativa mejora en la estimación de los datos pluviométricos. La validación del modelo incluyó comparaciones exhaustivas con datos reales, lo que permitió calibrar y confirmar la fiabilidad y precisión del modelo. Además, se implementó un proceso de retroalimentación y optimización continua para adaptar el modelo a cambios o nueva información.

La metodología completa, desde la recolección de datos hasta la aplicación final del modelo, se resume en la Figura 2. Esta figura ilustra un diagrama de flujo detallado que encapsula el proceso integral de aplicación de las RNA en la estimación y homogeneización de datos pluviométricos. La representación visual destaca la secuencia lógica y las interconexiones entre las diversas fases de nuestra metodología, desde la recopilación inicial de datos hasta la aplicación práctica y evaluación del modelo. Cada etapa está delineada para reflejar el rigor y la precisión con la que los datos fueron tratados, asegurando la robustez y fiabilidad del modelo resultante. Además, proporciona una referencia clara para replicar nuestra metodología y afirma la cohesión y consistencia de nuestro enfoque científico. La recolección de datos inició con una revisión exhaustiva de fuentes bibliográficas centradas en la homogeneización de datos y el uso de redes neuronales para el análisis de datos meteorológicos, con el fin de comprender en detalle los procesos de precipitación de la cuenca de estudio, así como los desafíos asociados con la medición y la heterogeneidad en los registros meteorológicos.

Homogeneización y estimación de datos mediante RNA

Las RNA son modelos computacionales que emulan la estructura neuronal del cerebro humano, compuestas por unidades de procesamiento interconectadas o "neuronas". Su habilidad para adaptarse a patrones complejos las convierte en herramientas ideales para tareas como el reconocimiento de patrones, la predicción y el modelado de sistemas complejos (Goos et al., 2012; Houari et al., 2014; Lee & Kang, 2015; Matovelle et al., 2022). En este estudio, se seleccionó una RNA tipo perceptrón multicapa debido a su eficacia demostrada en la modelización y predicción de datos meteorológicos. Este modelo fue entrenado para predecir datos pluviométricos ausentes, optimizando el manejo de no linealidades en los datos hidrológicos y mejorando la precisión en series históricas. La configuración de la RNA se diseñó meticulosamente, ajustando los parámetros según la necesidad de mejorar los resultados en iteraciones sucesivas. La calidad de las predicciones se verificó utilizando funciones de validación y documentando los resultados en formatos accesibles para análisis y revisión posterior.

La RNA, adecuadamente entrenada, fue responsable de identificar y corregir las inconsistencias y vacíos de información de precipitación mensual, resultando en un conjunto de datos coherente y homogéneo. En total, se seleccionaron 231 estaciones del país, con 29 situadas específicamente en la cuenca del Río Pastaza y las restantes distribuidas en otras cuencas del Ecuador que cumplieran con los criterios específicos para asegurar la confiabilidad del entrenamiento de la RNA. Este enfoque garantizó que las predicciones de la RNA fueran comparables y basadas en datos verificables de las cuencas externas, mientras que la cuenca del Río Pastaza sirvió como estudio de caso independiente para validar los resultados del modelo.

Posteriormente, se realizó un riguroso proceso de calibración y validación, utilizando datos reales como comparación para evaluar la precisión del modelo. Se establecieron parámetros de exclusión para estaciones con datos incompletos y se aplicó un método de razones de distancia en la cuenca del Río Pastaza para corroborar la coherencia geográfica y climatológica de las estimaciones. Este enfoque integral permitió ajustar el modelo de RNA de manera precisa, mejorando continuamente su capacidad para prever y homogeneizar datos pluviométricos en contextos complejos y variados.

La RNA fue configurada con un número de iteraciones ajustable, de 100 a 5000, para optimizar la precisión del modelo en función de los datos proporcionados y los resultados obtenidos. La iteración óptima se estableció en 5000 tras una evaluación rigurosa, asegurando una alta fidelidad en la correspondencia entre los valores predichos y los reales. La arquitectura del software incluye varias fases

de calibración y validación del modelo, con un sistema adaptativo que ajusta el número de iteraciones basado en la eficacia observada. Este enfoque se complementa con herramientas gráficas integradas que ilustran visualmente la

precisión del modelo, facilitando la interpretación de su desempeño (Figura 2).

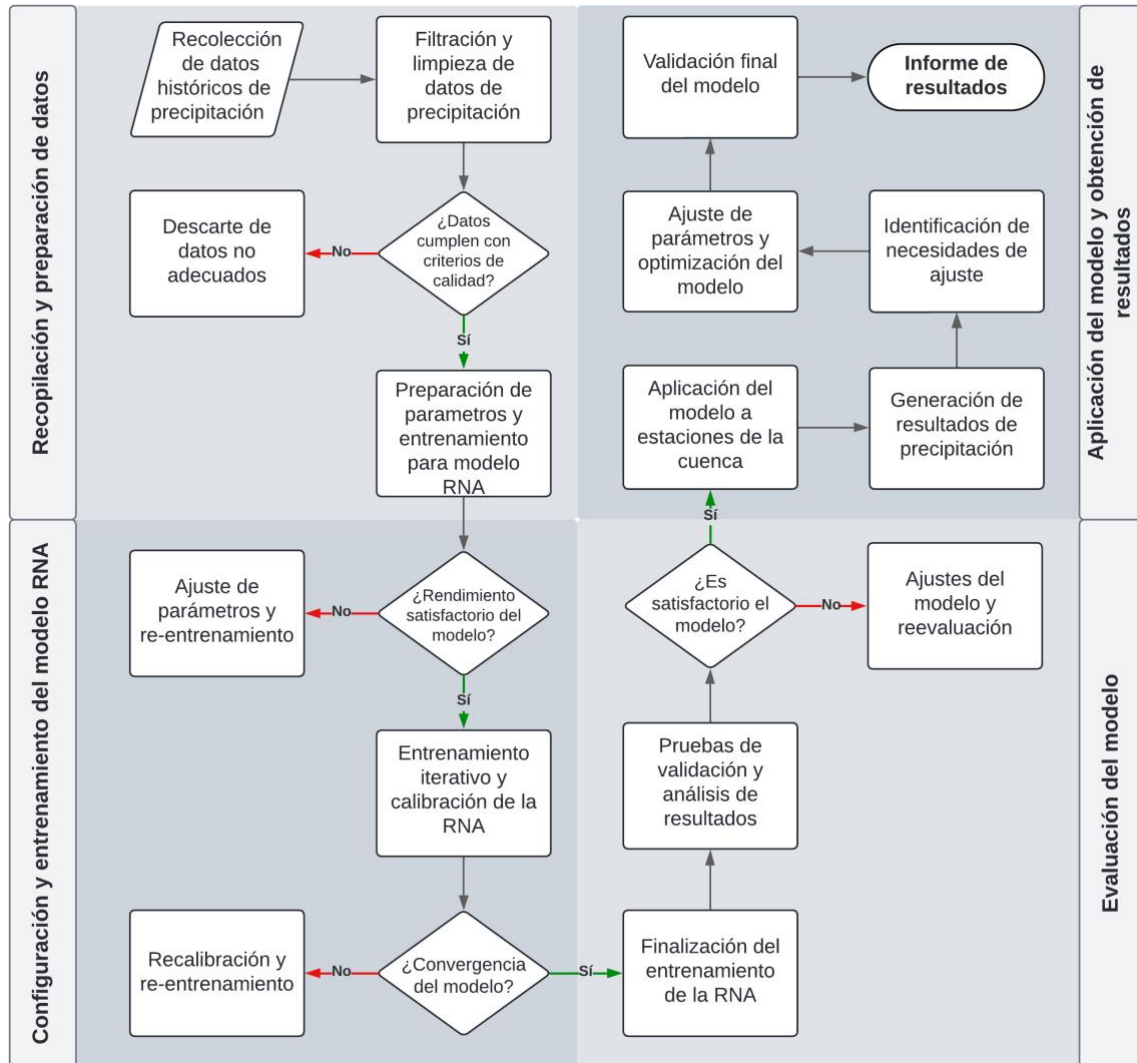


Figura 2: Diagrama de flujo de la metodología utilizada para la aplicación de redes neuronales en la estimación de series de precipitación.

Codificación y Librerías del Proyecto

El desarrollo del software se realizó en Python, aprovechando su flexibilidad y la robustez de librerías como NumPy (Harris et al., 2020) y Pandas (McKinney, 2010) para el manejo eficiente de los datos. TensorFlow y Keras se utilizaron para facilitar la construcción y el entrenamiento de las estructuras neuronales, mientras que Scikit-Learn se usó en la homogeneización y preparación de los datos. Las visualizaciones se realizaron mediante Matplotlib (Hunter, 2007) y Sea-born (Waskom, 2021), permitiendo una clara representación de las tendencias y eficacia del modelo.

Los datos hidrometeorológicos provinieron de estaciones meteorológicas en varias cuencas de Ecuador, con un meticuloso proceso de importación usando Pandas. Este proceso garantizó que los datos fueran óptimos para el análisis y la modelización, excluyendo registros de estaciones con datos incompletos. El entrenamiento de la RNA se llevó a cabo en varias iteraciones, determinando que 5000 iteraciones era el número óptimo para una aproximación precisa a los datos reales. Los resultados se exportaron utilizando Pandas, facilitando la comparación y validación del modelo. La etapa de calibración del software fue crítica, con revisiones exhaustivas de las iteraciones para asegurar la precisión de las predicciones. Se ajustaron los

parámetros del modelo en función del comportamiento de los datos de referencia, mejorando continuamente la eficiencia del modelo. Finalmente, la validación de los resultados se confirmó mediante una detallada comparación con los datos reales, mostrando una alta correlación en la fase de 5000 iteraciones, lo que evidencia la validez y robustez del modelo para predecir y gestionar datos meteorológicos en Ecuador.

Para la predicción de los valores de precipitación faltantes, se implementó un modelo de RNA con un perceptrón multicapa, utilizando la clase MLPRegressor del módulo Scikit-learn. Esta configuración es fundamental para la interpretación y el análisis de las complejas dinámicas hidrometeorológicas de la cuenca del Río Pastaza. Se optó por una estructura de red neuronal que comprende una capa de entrada, varias capas ocultas y una capa de salida, cada una calibrada cuidadosamente para optimizar la detección y corrección de patrones en los datos pluviométricos. El entrenamiento del modelo se efectuó a través de un enfoque iterativo, utilizando un método de retro propagación para ajustar los pesos de la red. El proceso iterativo continuó hasta alcanzar una convergencia satisfactoria, indicada por la disminución del error de predicción y una alineación cercana con los valores de precipitación conocidos. Para ello, el modelo se configuró con el optimizador Adam, que ajusta las tasas de aprendizaje de manera eficiente a través de gradientes estocásticos, favoreciendo una rápida convergencia y precisión en la regresión. La Figura 3 ilustra el esquema del modelo RNA empleado, resaltando el enfoque de descenso de gradiente estocástico utilizado para la optimización. La estructura detallada incluye dos neuronas de entrada correspondientes a las estaciones de referencia 'Est1' y 'Est3', seleccionadas por su cercanía con la estación de estudio, y una neurona de salida para la estación objetivo 'Est2', que se predice durante el entrenamiento. Con 5000 iteraciones, el modelo refina iterativamente los pesos de la red para minimizar la discrepancia entre los valores pronosticados y los reales de la estación de interés. El código para crear y entrenar el modelo es el siguiente:

$$\text{regr} = \text{MLPRegressor}(\text{random_state} = 1, \text{max_iter} = 5000). \text{fit}(x\text{TrainScaled}, y\text{Train}) \quad (1)$$

Este enfoque de optimización individualizada proporciona actualizaciones instantáneas y precisas de los parámetros de la red, lo cual es esencial para abordar la variabilidad inherente a los datos hidrológicos de las estaciones seleccionadas para el análisis. A pesar de los desafíos del enfoque estocástico, como el potencial ruido en los gradientes, su capacidad para evitar mínimos locales y aspirar al óptimo global es una ventaja distintiva para este tipo de análisis (Cartaya et al., 2016; Goos et al., 2012; Lee & Kang, 2015).

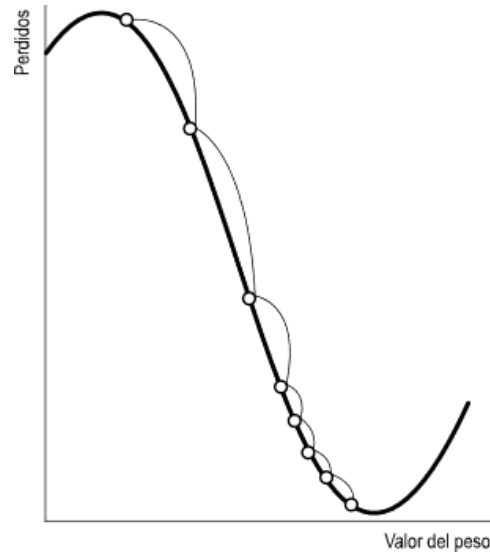


Figura 3: Modelo RNA de algoritmo de descenso de gradiente utilizado para la predicción de datos de precipitación faltantes en la cuenca del Río Pastaza.

Verificación de resultados

Con el fin de verificar la validez del trabajo de homogeneización de datos pluviométricos de la cuenca del Río Pastaza mediante RNA, se realizó una comparación de datos utilizando otro método de homogeneización más estandarizado. Para ello, se empleó el método de las razones de distancia, un enfoque clásico en la región que utiliza dos estaciones de referencia para calcular los valores faltantes en una estación evaluada. Esta técnica considera la proximidad geográfica de las estaciones para mejorar la precisión en la estimación de los valores faltantes, lo cual es esencial para preservar la coherencia de los datos dentro de una misma cuenca (Goulven, 1988; Guachamin et al., 2019). Este método es útil para completar los datos faltantes en las series meteorológicas en cuencas. Integra de manera efectiva la distancia física entre las estaciones en los cálculos, siendo eficaz al utilizar series completas de dos estaciones de referencia, identificadas como PA y PB, para estimar los datos faltantes en una estación con registros incompletos, PX, asegurando que las estimaciones sean geográficamente lógicas y meteorológicamente válidas. La fórmula utilizada para este método es:

$$PX = PA + a + ab \times (PB - PA) \quad (2)$$

donde 'a' y 'b' representan las distancias desde PX hasta las estaciones PA y PB, respectivamente. Esta fórmula no solo ajusta los datos por proximidad, sino que también mantiene la integridad de las tendencias meteorológicas observadas en las estaciones de referencia.

3. Resultados

Posterior a un exhaustivo análisis bibliográfico y recopilación de datos, se identificaron 29 estaciones hidrológicas dentro de la cuenca del Río Pastaza, cuya distribución está detallada en la Figura 4. La recolección de datos abarcó un período de 19 años a una escala mensual, desde enero de 1995 hasta diciembre de 2013, equivalente a 228 meses de registros pluviométricos. Durante el análisis preliminar, se determinó que 19 de estas estaciones superaban el umbral del 30% de datos faltantes, lo que justificó enfocar el tratamiento y análisis en estas estaciones debido a la significativa cantidad de información incompleta.

El estudio de las precipitaciones en la cuenca se visualiza mediante el diagrama de cajas y bigotes en la Figura 5, que ofrece un resumen estadístico de las variaciones y tendencias de las precipitaciones durante el periodo estudiado. Las estaciones de Puyo (M008) y Sangay (M041) destacan por registrar niveles de precipitación significativamente más altos en comparación con otras estaciones, una diferencia atribuible a sus ubicaciones geográficas y factores climáticos locales. Por el contrario, el resto de las estaciones mostraron

Tabla 1 presenta la información relevante de las estaciones analizadas en este estudio. En ella se detallan el código nacional de cada estación, su nombre, y su ubicación por latitud y longitud en coordenadas geográficas, información necesaria para determinar las distancias entre las estaciones. Asimismo, se desglosan los porcentajes de datos faltantes antes y después de la aplicación de técnicas de recuperación utilizando RNA, junto con el número de datos recuperados. Por ejemplo, la estación M133 (Guaslán) mostró una notable

niveles de precipitación menores y fluctuaciones más moderadas. Los valores atípicos observados en el diagrama indican eventos de lluvias intensas que, aunque esporádicos, son cruciales para entender la dinámica hídrica y las respuestas de la cuenca a condiciones meteorológicas extremas.

Estos eventos de precipitación en la cuenca del Río Pastaza son cruciales para evaluar la vulnerabilidad y capacidad de respuesta de la cuenca a fenómenos climáticos severos. La heterogeneidad en la precipitación documentada enfatiza la necesidad de considerar las variaciones espaciales y temporales en la gestión del agua y en la planificación hidrológica (Millán et al., 2010; Siren, 2014). La consistencia de información entre las estaciones en términos de precipitación promedio y la detección de eventos extremos proporcionan una base sólida para futuras investigaciones, que podrían profundizar en la modelización climática y la evaluación de riesgos hidrológicos en la región.

La

mejora en sus series de precipitación, con datos faltantes reducidos del 28.1% al 16%, recuperando 28 registros. De igual forma, en la estación M377 (Tisaleo), los datos faltantes disminuyeron del 15.8% al 4%, recuperando 26 registros. Estos resultados subrayan la eficacia de los métodos de homogeneización mediante RNA, demostrando su capacidad para reconstituir series de datos pluviales esenciales para estudios hidrológicos detallados en la cuenca del Río Pastaza.

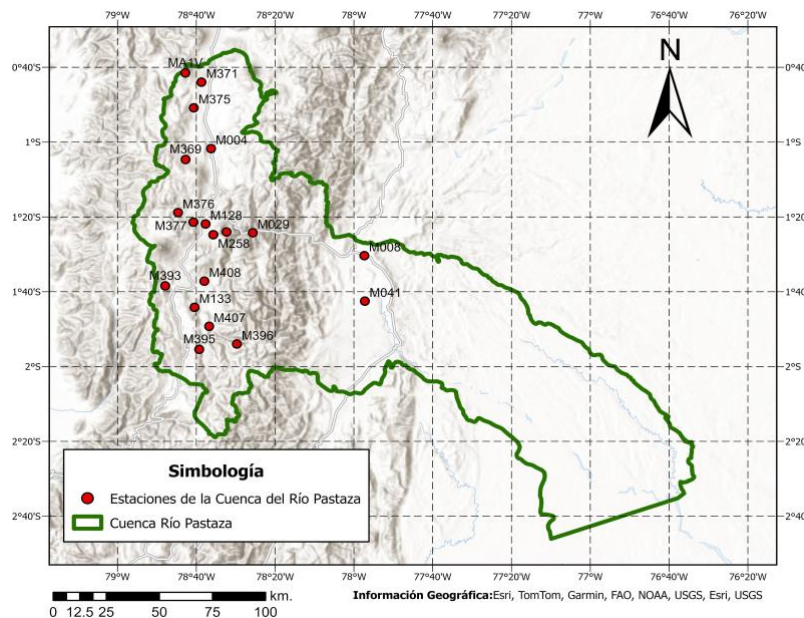


Figura 4: Ubicaciones de las estaciones meteorológicas consideradas para análisis dentro de la cuenca del Río Pastaza.

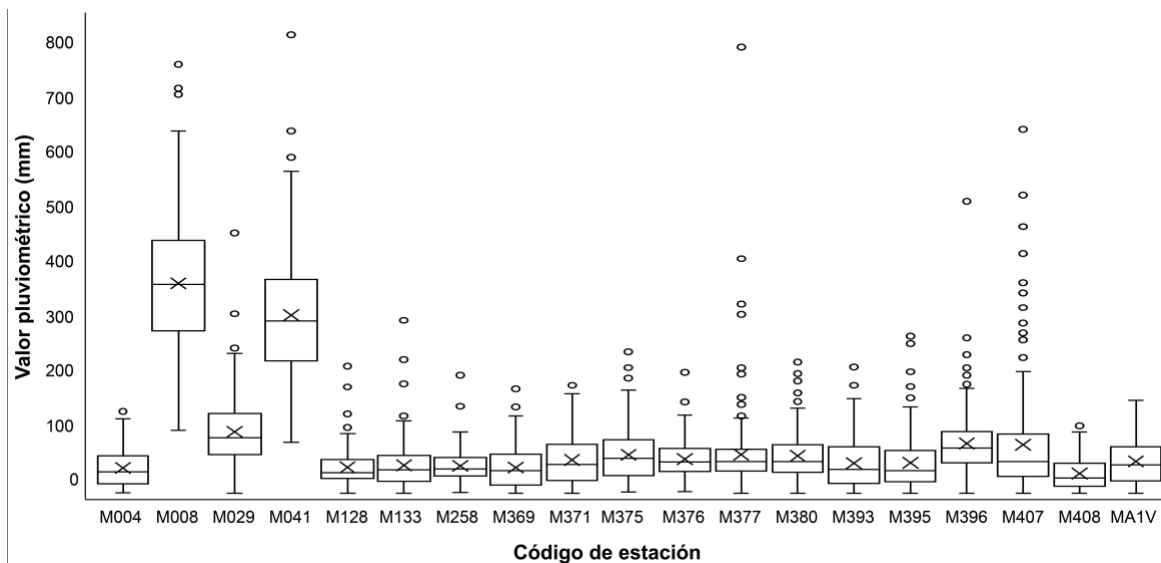


Figura 5: Diagrama de caja y bigotes de las precipitaciones anuales de 1995 a 2013 de las estaciones consideradas para análisis en la cuenca del Río Pastaza.

Tabla 1: Información de las estaciones meteorológicas analizadas en la cuenca del Río Pastaza

Código	Nombre Estación	Latitud	Longitud	Altitud	Datos faltantes	Datos faltantes después	Datos recuperados
M004	Rumipamba	-78.604287	-1.028794	2685	0.0%	0%	0
M008	Puyo	-77.956	-1.5057	960	0.9%	0%	2
M029	Baños	-78.427975	-1.404053	1695	14.0%	0%	32
M041	Sangay	-77.953618	-1.708765	880	5.3%	0%	12
M128	Pedro Fermín	-78.627353	-1.364711	2910	7.5%	0%	17
M133	Guaslán	-78.674136	-1.735808	2850	28.1%	16%	28
M258	Querochaca	-78.594954	-1.412856	2865	4.4%	4%	0
M369	Cusubamba	-78.712371	-1.077123	3175	6.1%	4%	4
M371	Pastocalle	-78.644612	-0.732838	3074	25.9%	23%	7
M375	Saquisilí	-78.677113	-0.847221	2892	10.1%	4%	13
M376	Pilahuín	-78.744245	-1.313889	3314	7.0%	4%	6
M377	Tisaleo	-78.679258	-1.357131	3266	15.8%	4%	26
M380	Huambaló	-78.538765	-1.400113	2880	5.3%	4%	2
M393	San Juan	-78.798557	-1.640619	3220	5.7%	4%	3
M395	Cebadas	-78.654252	-1.923329	2930	7.9%	4%	8
M396	Alao	-78.494818	-1.899688	3200	13.2%	4%	20
M407	Licto	-78.612362	-1.820821	2840	21.9%	15%	16
M408	Guano	-78.632549	-1.619477	2620	7.5%	4%	7
MA1V	Cotopilaló	-78.712922	-0.691507	3250	27.6%	21%	15

La integridad de los conjuntos de datos de precipitación en la cuenca del Río Pastaza fue minuciosamente evaluada utilizando la biblioteca *missingno* de Python, reconocida por su capacidad para visualizar eficazmente los patrones de datos faltantes. La Figura 6(a) muestra detalladamente la

distribución y cantidad de datos disponibles en las estaciones analizadas, destacando los vacíos significativos en la información recolectada. En esta visualización, cada línea blanca indica un dato ausente en los registros de cada estación, donde la fila 1 muestra una secuencia completa de

228 valores, y la fila 6 resalta una carencia notoria con solo 164 valores registrados. De igual forma, la Figura 6(b) evidencia cómo, después de implementar técnicas avanzadas de homogeneización utilizando RNA, se logró una notable

reducción de los valores faltantes. Este resultado demuestra la efectividad del enfoque computacional avanzado en la mejora de la calidad de los datos.

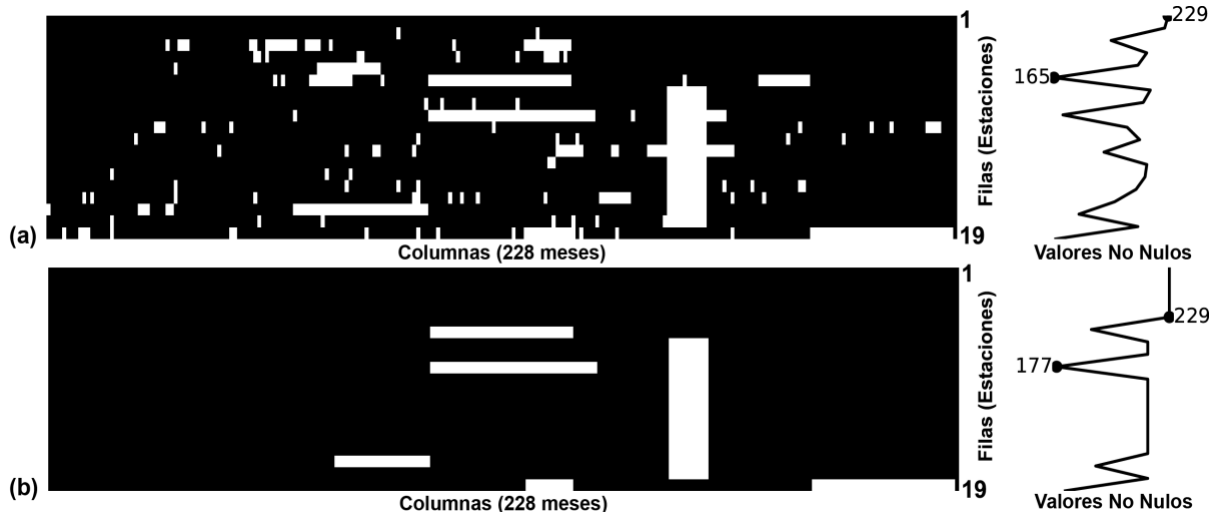


Figura 6: Densidad de datos de precipitación de las estaciones de la cuenca del Río Pastaza, el gráfico a) indica los datos de precipitación de las estaciones meteorológicas consideradas dentro de la cuenca, los espacios en blanco indican los meses sin información de precipitación, el gráfico b) indica los datos de precipitación luego del procesamiento de información, observándose un relleno de información de precipitación.

La eficacia de este proceso se refleja en la Tabla 1, que detalla cuantitativamente los datos recuperados, resaltando la consistencia y la integridad mejoradas del conjunto de datos tras el tratamiento. Estas visualizaciones no solo son esenciales para el preprocesamiento de los datos, sino que también establecen una base sólida para las etapas subsiguientes del análisis en la investigación. Al clarificar la disponibilidad y calidad de los datos recopilados, se facilita una mejor comprensión de las dinámicas de precipitación de la cuenca, permitiendo análisis más precisos y fundamentados.

Para la estimación de datos de precipitación mediante RNA, se utilizó un enfoque de entrenamiento con datos de estaciones meteorológicas externas a la cuenca del Río Pastaza. Este enfoque permitió validar la efectividad del modelo en un contexto más amplio, empleando registros de diversas cuencas a lo largo de Ecuador y excluyendo específicamente la cuenca del Río Pastaza para preservar la imparcialidad en la evaluación de resultados. La selección de estaciones siguió un estricto criterio de filtrado, descartando aquellas con registros incompletos para asegurar la integridad y fiabilidad del entrenamiento. La calibración del modelo alcanzó 5000 iteraciones, demostrando una notable convergencia con los valores reales.

La Figura 7 compara la precisión de las predicciones del modelo en la cuenca del Río Pastaza con datos reales de otras estaciones con registros completos. La línea naranja de la

predicción sigue de cerca la línea azul de los valores reales a lo largo de 350 muestras de precipitación. Las fluctuaciones en las predicciones reflejan la capacidad del modelo de capturar con precisión la variabilidad en los datos pluviométricos, incluso durante eventos de lluvias intensas, lo que es crucial para el análisis hidrológico detallado y la gestión de recursos hídricos.

Esta estrecha correlación subraya la superioridad del modelo de RNA sobre los métodos manuales tradicionales, demostrando no solo una mayor precisión sino también una robustez significativa en la homogeneización de los datos. Al comparar este modelo avanzado con métodos manuales, se observa una mejora sustancial en la precisión y en la gestión de la variabilidad de los datos. La aplicación de un análisis no paramétrico mediante la prueba de rachas también confirmó la validez del modelo, con un índice de confiabilidad del 95%. Este nivel de fiabilidad corrobora que las predicciones no solo son precisas sino consistentes, reafirmando la capacidad de la RNA para ser aplicada en la homogeneización de datos en otras cuencas del país.

En la comparación de los resultados, se observó que la predicción mediante RNA y métodos convencionales difiere en su enfoque y capacidad para manejar datos complejos. Las RNA, al emplear gradientes estocásticos y considerar un conjunto de elementos más amplio, ofrecen una estimación más rica y detallada, lo cual se refleja en un coeficiente de determinación R^2 de 0.71. Este valor sugiere que las

predicciones del modelo capturan una proporción sustancial de la variabilidad en los datos, indicando un buen ajuste y ofreciendo una mejora notable en la precisión en comparación con los métodos convencionales que dependen de relaciones matemáticas simples entre estaciones cercanas.

La Figura 8, Figura 9 y Figura 10 describen visualmente los resultados obtenidos después del proceso de homogeneización de los datos de precipitación para cada una de las estaciones analizadas en este estudio en la cuenca del Río Pastaza utilizando RNA.

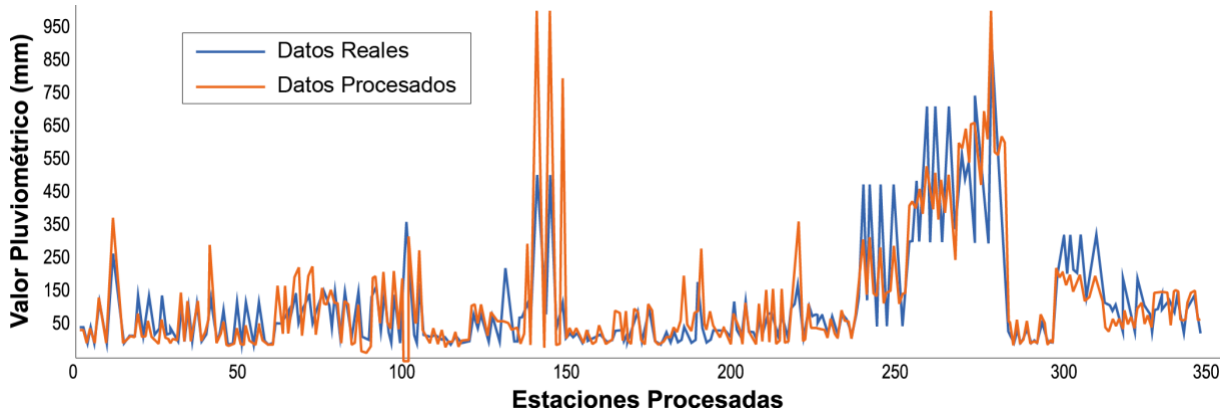


Figura 7: Convergencia con 5000 iteraciones del modelo predictivo para las estaciones de la cuenca del río Pastaza y que no consideradas en el entrenamiento.

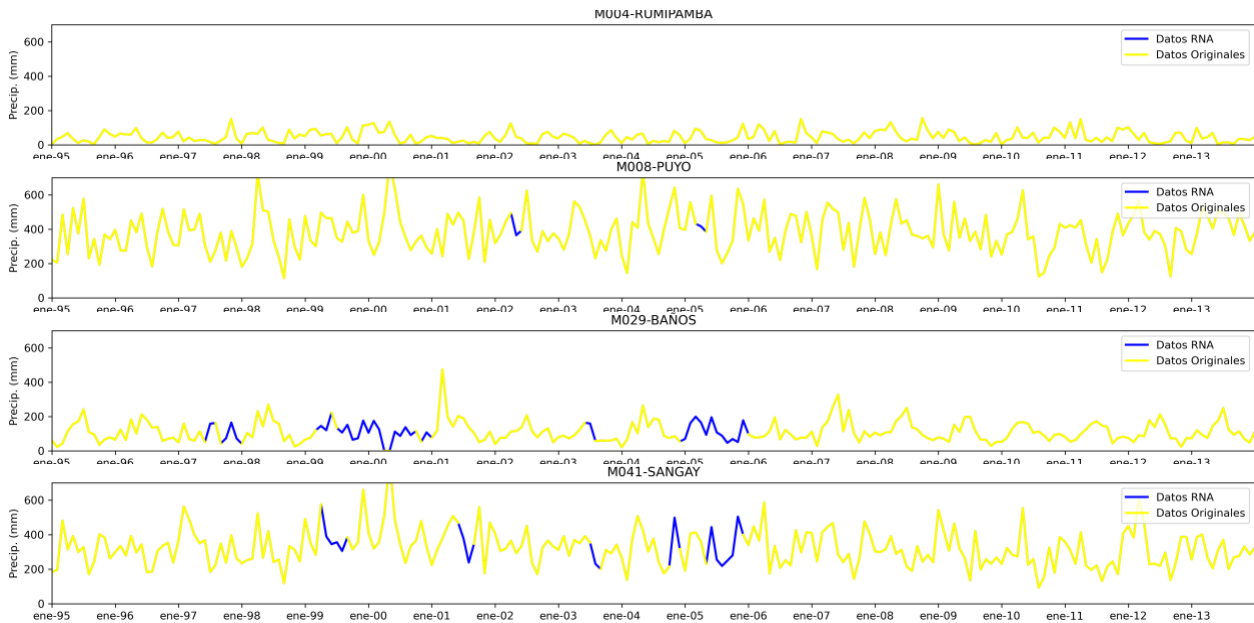


Figura 8: Resultados obtenidos de la homogeneización de datos de precipitación mensual en las estaciones del Río Pastaza (Est: M004, M008, M029 y M041).

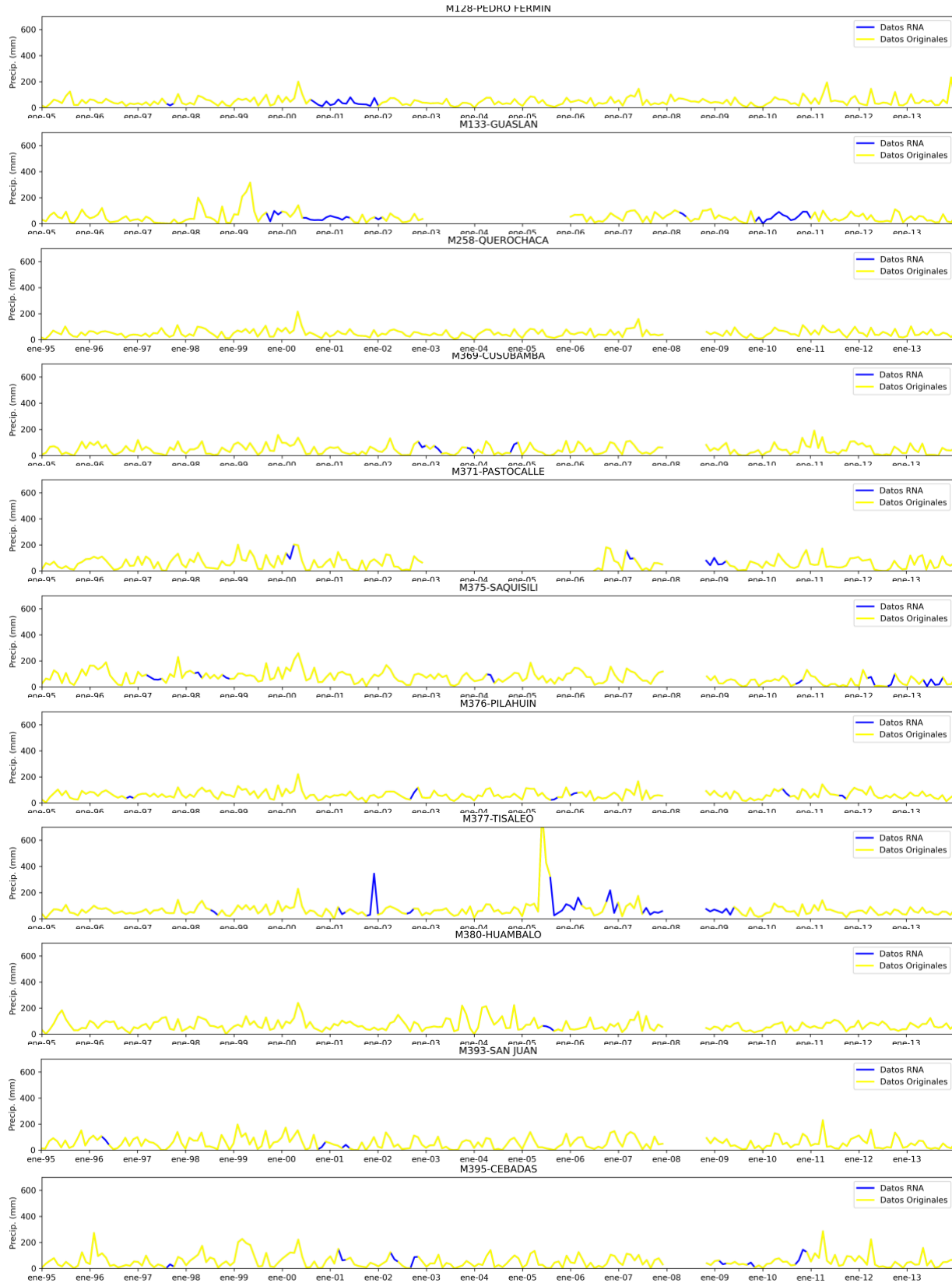


Figura 9: Resultados obtenidos de la homogeneización de datos de precipitación mensual en las estaciones del Río Pastaza (Est: M128, M133, M258, M369, M371, M375, M376, M377, M380, M393 y M395. , M396, M407, M408 y MAIV).

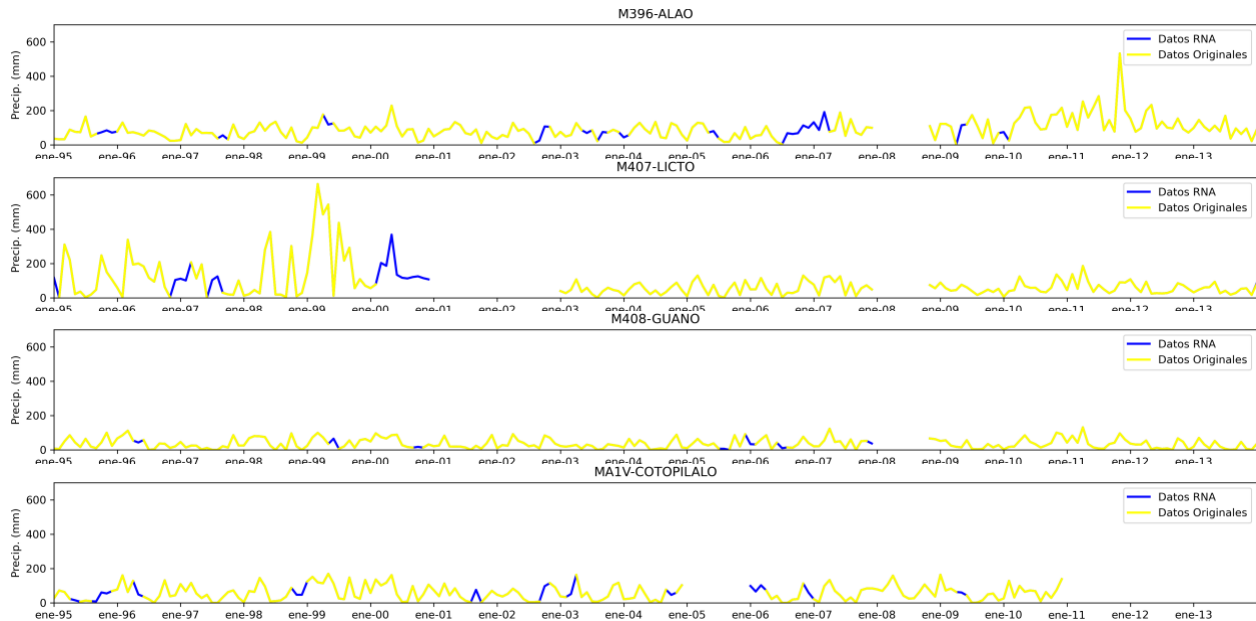


Figura 10: Resultados obtenidos de la homogeneización de datos de precipitación mensual en las estaciones del Río Pastaza (Est: M396, M407, M408 y MA1V).

Al analizar los resultados, se observa que el proceso de homogeneización utilizando RNA mejora la precisión de los datos previstos en comparación con los datos originales, verificando la efectividad del método. La Figura 11 muestra la información de la estación de Alao correspondiente al período entre 2003 y 2006, sugiriendo una incorporación coherente de los datos tratados y reflejando una alta fidelidad en los patrones de precipitación.

4. Discusión

Los resultados alcanzados mediante el empleo de RNA en la cuenca del Río Pastaza abren nuevas perspectivas en el estudio de la hidrología y climatología regional. Este análisis refuerza la importancia de estimaciones pluviométricas precisas para el modelado de patrones climáticos y la gestión de recursos hídricos, demostrando la superioridad de las RNA sobre métodos convencionales en la captura y análisis de la variabilidad climática. Al integrar múltiples variables y trascender la simple relación de proximidad entre estaciones, las RNA proporcionan predicciones que reflejan de manera más fiel la complejidad de los sistemas hidrológicos de la región. La versatilidad del modelo sugiere su aplicabilidad en diversas cuencas del Ecuador y su potencial adopción en otros entornos geográficos con desafíos hidrológicos y climáticos similares.

La relevancia de las RNA en la predicción de información hidrológica es evidente, ya que pueden proporcionar datos críticos para la planificación estratégica y la prevención frente a fenómenos meteorológicos extremos. El enfoque

sofisticado de las RNA para la estimación pluviométrica puede ser crucial para la optimización de infraestructuras y la resiliencia hídrica de las comunidades. Estos resultados actúan como un catalizador para la sostenibilidad, demostrando cómo la tecnología avanzada puede y debe integrarse con la gestión ambiental y climática (Carrera-Villacrés et al., 2016; Cartaya et al., 2016; Goos et al., 2012; Millán et al., 2010).

El uso de técnicas avanzadas en la investigación hidrológica contribuye significativamente a la ciencia del agua y el clima, enriqueciendo el conocimiento científico con métodos innovadores y eficaces. Este desarrollo amplía la comprensión de las dinámicas interrelacionadas entre sistemas naturales y antropogénicos, facilitando descubrimientos que mejoren nuestra capacidad de respuesta frente a los desafíos medioambientales del futuro. La aplicación de las RNA en la estimación de datos pluviométricos de la cuenca del Río Pastaza ofrece un enfoque innovador respecto a estudios anteriores. Mientras que investigaciones previas han dependido mayormente de métodos estadísticos convencionales, este estudio se distingue por integrar y procesar las complejidades inherentes a los datos climáticos de manera dinámica y holística (Kalauzi et al., 2009; Millán et al., 2008, 2010).

La comparación directa con metodologías tradicionales subraya no solo la viabilidad, sino también la superioridad de las RNA en la gestión de grandes volúmenes de datos y la precisión de los resultados obtenidos.

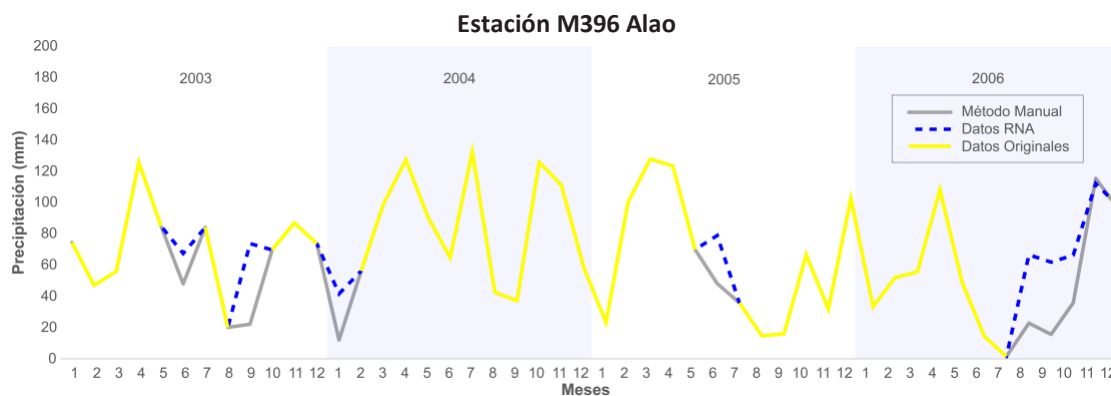


Figura 11: Visualización del proceso de homogeneización de datos de precipitación mediante el uso de RNA en las estaciones de Baños y Alao de la cuenca del Río Pastaza.

El uso de técnicas de aprendizaje automático, como las RNA, en la investigación climática e hidrográfica, está en línea con la tendencia contemporánea de explorar patrones de datos y su interpretación a mayor profundidad. Al comparar los resultados de este estudio con investigaciones similares, se observa un incremento en la capacidad de análisis y precisión predictiva, apoyando la propuesta de que las RNA podrían ser una herramienta viable en futuros análisis hidrológicos y climáticos, especialmente en regiones como Ecuador, donde la información hidrológica y climatológica de calidad es limitada. Este estudio presenta una metodología adaptable y aplicable a diferentes contextos y conjuntos de datos, a diferencia de métodos más rígidos. Esta flexibilidad y la capacidad de las RNA para calibrarse a diversas condiciones refuerzan su potencial como herramienta en la homogeneización de datos.

En el contexto del cambio climático, la creciente variabilidad climática plantea desafíos significativos a la gestión hídrica, donde la predicción precisa de los patrones de precipitación es crucial para la adaptación y mitigación (Gómez Guerrero & Aguayo Arias, 2019; Kalauzi et al., 2009; Matovelle et al., 2022; Millán et al., 2008). Los datos estimados por RNA ofrecen una comprensión avanzada y detallada de la variabilidad pluviométrica, esencial para predecir y adaptarse a las condiciones climáticas cambiantes. Las RNA destacan como herramientas predictivas de relevancia ante el cambio climático, capaces de predecir eventos climáticos extremos con mayor precisión. Esta capacidad es vital para preparar a las comunidades frente a desastres naturales, facilitando la implementación de estrategias de respuesta temprana y reduciendo el impacto negativo tanto en las poblaciones humanas como en los ecosistemas de la región. Además, los resultados obtenidos mediante RNA pueden mejorar significativamente la calidad de los modelos climáticos y de proyecciones futuras, permitiendo entender cómo la alteración del clima influirá en los ciclos

hidrológicos de la cuenca del Río Pastaza. Esto es crucial para proyectar y anticipar cambios en la disponibilidad de recursos hídricos y desarrollar estrategias de gestión robustas y adaptables a futuro.

Aunque nuestros resultados muestran la eficacia del uso de RNA en la homogeneización de datos pluviométricos, existen limitaciones. Una restricción notable es la dependencia de registros históricos, que en esta región pueden presentar imprecisiones debido a métodos de recolección antiguos o errores en la instrumentación. En algunas estaciones analizadas en la cuenca del Río Pastaza, los datos faltantes eran tan extensos que su imputación podía introducir sesgos en el modelo. Esto resalta la importancia de la calidad y la integridad de los datos en la modelización de RNA.

De cara al futuro, es necesario continuar la investigación enfocándose en la mejora y verificación de los modelos de RNA aplicándolos a una variedad más amplia de cuencas hidrográficas, para examinar su aplicabilidad y robustez en diferentes contextos geográficos y climáticos. La integración de tecnologías emergentes y modelos hidrológicos más sofisticados será de gran utilidad en las cuencas de la región. La incorporación de datos de sensores remotos y plataformas de modelización climática podría enriquecer los conjuntos de datos y proporcionar una mayor precisión en las estimaciones. Además, es crucial explorar otros métodos de inteligencia artificial, como las redes neuronales convolucionales y los sistemas de aprendizaje profundo, que podrían ofrecer información adicional valiosa en la compleja tarea de homogeneizar y analizar datos hidrológicos. El desarrollo de enfoques híbridos, que combinen múltiples metodologías de IA, podría superar algunas de las limitaciones identificadas y abrir nuevos horizontes en la investigación hidrológica y climática de la región.

5. Conclusiones

El presente estudio se ha enfocado en la estimación de datos pluviométricos mediante el uso de RNA. De forma destacada, las RNA han demostrado ser una herramienta de gran utilidad en términos de precisión y capacidad de generalización en comparación con los métodos tradicionales. Los hallazgos indican que las RNA pueden predecir eficazmente los vacíos en los registros pluviométricos, incluso en cuencas hidrográficas tan complejas como la del río Pastaza, lo que tiene profundas implicaciones para la gestión hídrica y la planificación climática. La robustez del modelo de RNA fue evidenciada a través de su aplicación en diferentes cuencas en Ecuador, lo que sugiere su potencial de aplicación en una variedad más amplia de contextos hidrológicos. La aplicación de RNA en la homogeneización de datos no solo ha mejorado la calidad de la información disponible para la toma de decisiones, sino que también establece un precedente en la precisión de las predicciones hidrológicas y climáticas de las cuencas de la región.

Esencialmente, estos resultados realzan la validez de las RNA como una metodología confiable y avanzada para la estimación de datos faltantes en estudios hidrológicos, proporcionando un sólido fundamento para la implementación de prácticas de manejo de recursos hídricos más informadas y la formulación de políticas climáticas más efectivas. El uso de RNA en la estimación de datos hidrometeorológicos ha demostrado ser no solo aplicable sino también altamente relevante en la era actual, marcada por la incertidumbre climática y la necesidad de una gestión eficiente de recursos naturales. La capacidad de las RNA para procesar y analizar grandes volúmenes de datos con complejas interacciones y patrones variables es particularmente pertinente en el campo hidrometeorológico, donde la variabilidad espacial y temporal de los datos puede ser sustancial. La precisión mejorada y la capacidad predictiva de las RNA habilitan predicciones más confiables, lo que es vital para la planificación y respuesta ante fenómenos extremos, como inundaciones y sequías. En la práctica, esto se traduce en la mejora de los sistemas de alerta temprana y en la planificación estratégica a largo plazo de las infraestructuras hídricas y agrícolas, contribuyendo significativamente a la resiliencia y sostenibilidad de las comunidades afectadas.

De esta forma, este análisis sugiere una mejoría de las RNA en la predicción de valores faltantes de precipitación, lo que las posiciona como una herramienta poderosa y precisa para la homogeneización de datos en el campo de la hidrología, con implicaciones directas en la gestión de recursos hídricos en cuencas hídricas en países como Ecuador, donde la falta de información confiable es evidente. En definitiva, los resultados del estudio subrayan el potencial transformador de las RNA en la hidrología y la meteorología, así como para

la gestión del agua, reforzando la importancia de la integración de métodos analíticos avanzados en el manejo y análisis de datos hidrometeorológicos para la toma de decisiones estratégicas y fundamentadas. La implementación de RNA para la estimación y homogeneización de datos pluviales en la cuenca del Río Pastaza tiene implicaciones prácticas profundas para la gestión de sus recursos hídricos. El análisis refinado de la precipitación permite a los administradores de recursos hídricos desarrollar modelos hidrológicos más precisos y, por ende, tomar decisiones mejor informadas respecto a la conservación del agua, la agricultura, y la planificación de infraestructuras hídricas. A medida que la tecnología evoluciona y se vuelve más accesible, se abren perspectivas para aplicar técnicas similares en otras cuencas hidrográficas de la región, con la intención de mejorar nuestra comprensión y gestión de los recursos hídricos de Ecuador.

Referencias

- Abadi, M., Barham, P., Sanjay Ghemawat, Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D. G., Steiner, B., Tuck-er, P., Vasudevan, V., Warden, P., Wicke, M., Yu, Y., Zheng, X., Google Brain, Chen, J., ... Devin, M. (2016). TensorFlow: A System for Large-Scale Machine Learning. 12th USENIX Symposium on Operating Systems Design and Implementation, 265–283.
- Abril-Saltos, R. V., Armas-Chugcho, P. A., Chamorro, W. P., Toscano-Salazar, V. E., Sucoshañay-Villalva, D. J., & Ríos-Rodríguez, A. F. (2021). Calidad de agua del río Puyo y afluentes, Pastaza, Ecuador. *Tecnología y Ciencias Del Agua*, 12(3), 379–417. <https://doi.org/10.24850/j-tyca-2021-03-10>
- Carrera-Villacrés, D. V., Guevara-García, P. V., Tamayo-Bacacela, L. C., Balarezo-Aguilar, A. L., Narváez-Rivera, C. A., & Morocho-López, D. R. (2016). Relleno de series anuales de datos meteorológicos mediante métodos estadísticos en la zona costera e interandina del Ecuador, y cálculo de la precipitación media. *Idesia (Arica)*, 34(3), 81–90. <https://doi.org/10.4067/S0718-34292016000300010>
- Cartaya, S., Zurita, S., & Montalvo Párraga, V. (2016). Métodos de ajuste y homogenización de datos climáticos para determinar índice de humedad de Lang en la provincia de Manabí, Ecuador. *Revista La Técnica*, 94(106), 94–106.
- Chollet, F. (2015). <https://github.com/fchollet/keras>. GitHub Repository. <https://github.com/keras-team/keras>
- Cruzatty, C., Jimenez, D., Valencia, E., Zambrano, I., Mora, C., Luo, X., & Cando, E. (2021). A Case Study: Sediment Erosion in Francis Turbines Operated at the San Francisco Hydropower Plant in Ecuador. *Energies*, 15(1), 8. <https://doi.org/10.3390/en15010008>

- Flores Llampazo, G., Honorio Coronado, E. N., del Aguila-Pasquel, J., Cordova Oroche, C. J., Díaz Narvaez, A., Reyna Huaymacari, J., Grandez Ríos, J., Lawson, I. T., Hastie, A., Baird, A. J., & Baker, T. R. (2022). The presence of peat and variation in tree species composition are under different hydrological controls in Amazonian wetland forests. *Hydrological Processes*, 36(9). <https://doi.org/10.1002/hyp.14690>
- Gómez Guerrero, J. S., & Aguayo Arias, M. I. (2019). Evaluación de desempeño de métodos de relleno de datos pluviométricos en dos zonas morfoestructurales del Centro Sur de Chile. *Investigaciones Geográficas*, 99. <https://doi.org/10.14350/rig.59837>
- Goos, G., Hartmanis, J., Van, J., Board, L. E., Hutchison, D., Kanade, T., Kittler, J., Kleinberg, J. M., Kobsa, A., Mattern, F., Zurich, E., Mitchell, J. C., Naor, M., Nierstrasz, O., Steffen, B., Sudan, M., Terzopoulos, D., Tygar, D., & Weikum, G. (2012). *Neural Information Processing* (T. Huang, Z. Zeng, C. Li, & C. S. Leung, Eds.; Vol. 7666). Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-34478-7>
- Goulven, P. (1988). Homogeneización de los datos pluviométricos. https://horizon.documentation.ird.fr/exl-doc/pleins_textes/divers11-03/010042844.pdf
- Guachamin, W., Garcia, F., Arteaga, M., & Cadena, J. (2019). Determinación de Ecuaciones para el Cálculo de Intensidades Máximas de Precipitación (2nd ed., Vol. 2). INAMHI. https://www.inamhi.gob.ec/Publicaciones/Hidrologia/ESTUDIO_DE_INTENSIDADES_V_FINAL.pdf
- Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., ... Oliphant, T. E. (2020). Ar-ray programming with NumPy. *Nature*, 585(7825), 357–362. <https://doi.org/10.1038/s41586-020-2649-2>
- Herrmann, P. (2002). Management Conflicts in the Ambato River Watershed, Tungurahua Province, Ecuador. *Mountain Research and Development*, 22(4), 338–340. [https://doi.org/10.1659/0276-4741\(2002\)022\[0338:MCITAR\]2.0.CO;2](https://doi.org/10.1659/0276-4741(2002)022[0338:MCITAR]2.0.CO;2)
- Houari, R., Bounceur, A., Tari, A. K., & Kecha, M. T. (2014). Handling missing data problems with sampling methods. *Proceedings - 2014 International Conference on Advanced Networking Distributed Systems and Applications, INDS 2014*, 99–104. <https://doi.org/10.1109/INDS.2014.25>
- Hunter, J. D. (2007). Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering*, 9(3), 90–95. <https://doi.org/10.1109/MCSE.2007.55>
- INHAMI. (2024). Anuarios Meteorológicos. <https://inamhi.website/anuarios-metereologicos/>
- Kalauzi, A., Cukic, M., Millán, H., Bonafoni, S., & Biondi, R. (2009). Comparison of fractal dimension oscillations and trends of rainfall data from Pastaza Province, Ecuador and Veneto, Italy. *Atmospheric Research*, 93(4), 673–679. <https://doi.org/10.1016/j.atmosres.2009.02.007>
- Lee, H., & Kang, K. (2015). Interpolation of Missing Precipitation Data Using Kernel Estimations for Hydrologic Modeling. *Advances in Meteorology*, 2015, 1–12. <https://doi.org/10.1155/2015/935868>
- López, S., & Sierra, R. (2010). Agricultural change in the Pastaza River Basin: A spatially explicit model of native Amazonian cultivation. *Applied Geography*, 30(3), 355–369. <https://doi.org/10.1016/j.apgeog.2009.10.004>
- Matovelle, C., Heras, D., & Solano-Peláez, J. (2022). Eficiencia de la Imputación de Datos Faltantes de Precipitaciones Utilizando Herramientas Computacionales en la Cuenca Hidrográfica, Jubones - Ecuador. *Revista Politécnica*, 50(2), 23–30. <https://doi.org/10.33333/rp.vol50n2.03>
- McKinney, W. (2010). Data Structures for Statistical Computing in Python. 56–61. <https://doi.org/10.25080/Majora-92bf1922-00a>
- Millán, H., Kalauzi, A., Cukic, M., & Biondi, R. (2010). Nonlinear dynamics of meteorological variables: multifractality and chaotic invariants in daily records from Pastaza, Ecuador. *Theoretical and Applied Climatology*, 102(1–2), 75–85. <https://doi.org/10.1007/s00704-009-0242-6>
- Millán, H., Kalauzi, A., Llerena, G., Sucoshañay, J., & Piedra, D. (2008). Climatic trends in the Amazonian area of Ecuador: Classical and multifractal analyses. *Atmospheric Research*, 88(3–4), 355–366. <https://doi.org/10.1016/j.atmosres.2007.11.030>
- Najarchi, M., Mahdavi, A., Hazaveh, E., Hazaveh, S. M. M. H., & Najafzadeh, S. M. M. (2020). Determination of missing precipitation data by various methodologies. In *Journal of Critical Reviews* (Vol. 7, Issue 6, pp. 925–935). Innovare Academics Sciences Pvt. Ltd. <https://doi.org/10.31838/jcr.07.06.161>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830. <https://doi.org/10.48550/arXiv.1201.0490>
- Serrano-Vincenti, S., Condom, T., Campozano, L., Escobar, L. A., Walpersdorf, A., Carchipulla-Morales, D., & Villacís, M. (2022). Harmonic Analysis of the Relationship between GNSS Precipitable Water Vapor and Heavy Rainfall over the Northwest Equatorial Coast, Andes, and Amazon Regions. *Atmosphere*, 13(11), 1809. <https://doi.org/10.3390/atmos13111809>

Siren, A. (2014). History of natural resource use and environmental impacts in an interfluvial upland forest area in western Amazonia. *Fennia - International Journal of Geography*, 192(1), 36–53. <https://doi.org/10.11143/8825>

Van Rossum, G. (1995). Python tutorial.

Waskom, M. (2021). seaborn: statistical data visualization. *Journal of Open-Source Software*, 6(60), 3021. <https://doi.org/10.21105/joss.03021>

Proyecto de investigación relacionado

Los autores desean expresar su agradecimiento a la Dirección de Investigación y Desarrollo (DIDE) de la Universidad Técnica de Ambato por financiar el proyecto “*CARACTERIZACIÓN DE FUENTES HÍDRICAS SUPERFICIALES EN LA ZONA ANDINA DEL RÍO PASTAZA, ECUADOR*”, el cual fue aprobado mediante la resolución No. UTA-CONIN-2022-0030-R. Además, expresan su gratitud al grupo de investigación “*GESTIÓN DE RECURSOS NATURA-LES E INFRAESTRUCTURA SUSTENTABLE (GeReNIS)*” de la Universidad Técnica de Ambato.

Contribución de los autores (CRediT)

Conceptualización: A.R., A.H., B.P.; Curación de datos y contenidos, Software: A.R., A.H., D.M., B.P.; Análisis formal de datos, Investigación, Validación, Redacción-borrador original: A.R., A.H., F.C., F.M., D.M., B.P.; Adquisición de fondos, Administración de proyecto, Supervisión: D.M., B.P.; Metodología: A.R., A.H., F.C., B.P.; Recursos materiales: A.H., D.M., B.P.; Visualización: A.R., A.H., F.M., D.M., B.P.; Redacción-revisión y edición: F.C., D.M., B.P. Todos los autores han leído y aceptado la versión publicada del manuscrito.

Conflicto de intereses

Los autores han declarado que no existe conflicto de intereses en esta obra.



Derechos de autor 2024. Revista Científica FINIBUS - ISSN: 2737-6451.

Esta obra está bajo una licencia: Internacional Creative Commons Atribución-NoComercial-CompartirIgual 4.0