



**Ciencia Latina**  
Internacional

Ciencia Latina Revista Científica Multidisciplinar, Ciudad de México, México.  
ISSN 2707-2207 / ISSN 2707-2215 (en línea), septiembre-octubre 2024,  
Volumen 8, Número 5.

[https://doi.org/10.37811/cl\\_rcm.v8i5](https://doi.org/10.37811/cl_rcm.v8i5)

# COMPARACIÓN Y EVALUACIÓN DE MÉTODOS PARA LA IMPUTACIÓN DE PRECIPITACIÓN FALTANTE

COMPARISON AND EVALUATION OF METHODS FOR  
IMPUTING MISSING PRECIPITATION

**Silva-Tipantasig, Lenin Gabriel**  
Universidad Central del Ecuador, Ecuador

## Comparación y Evaluación de Métodos para la Imputación de Precipitación Faltante

Silva-Tipantasig, Lenin Gabriel<sup>1</sup>

[lgsilvat@uce.edu.ec](mailto:lgsilvat@uce.edu.ec)

<https://orcid.org/0009-0005-3696-3305>

Carrera de Ingeniería Civil  
Facultad de Ingeniería y Ciencias Aplicadas  
Universidad Central del Ecuador  
Quito, Ecuador

### RESUMEN

En estudios que involucran series temporales, como la precipitación, el análisis de datos depende en gran medida de la calidad y completitud de estos. En este trabajo se exploran diversas técnicas para la imputación de datos faltantes en registros de precipitación mensual, utilizando información proveniente de seis estaciones meteorológicas del Ecuador. Se han empleado conjuntos completos de datos para evaluar la efectividad de distintos métodos de estimación, incluyendo el método de promedio aritmético, la regresión lineal múltiple y árbol de decisión. El método de regresión lineal múltiple proporcionó una estimación aceptable de los datos de precipitación faltantes, lo cual es consistente con estudios previos cuyos hallazgos indican que la regresión lineal múltiple ha demostrado ser un enfoque confiable para la estimación de datos faltantes en diversos campos. Además, se implementó un algoritmo de árbol de decisión, conocido por su capacidad de ofrecer transparencia en la toma de decisiones, una característica útil en el análisis de fenómenos meteorológicos.

**Palabras clave:** datos faltantes, promedio aritmético, regresión lineal múltiple, árbol de decisión

---

<sup>1</sup> Autor principal

Correspondencia: [lgsilvat@uce.edu.ec](mailto:lgsilvat@uce.edu.ec)

# Comparison and Evaluation of Methods for Imputing Missing Precipitation

## ABSTRACT

In studies involving time series, such as precipitation, data analysis largely depends on their quality and completeness. This work explores various techniques for the imputation of missing data in monthly precipitation records, using information from six weather stations in Ecuador. Complete data sets have been used to evaluate the effectiveness of different estimation methods, including the arithmetic mean, multiple linear regression, and decision tree. The multiple linear regression method provided an acceptable estimate of the missing precipitation data, which is consistent with previous studies whose findings indicate that multiple linear regression has proven to be a reliable approach for estimating missing data in various fields. Additionally, a decision tree algorithm was implemented, known for its ability to offer transparency in decision-making, a useful feature in the analysis of meteorological phenomena.

**Keywords:** missing data, arithmetic averaging, multiple linear regression, decision tree

*Artículo recibido 10 septiembre 2024*

*Aceptado para publicación: 12 octubre 2024*



## INTRODUCCION

La ausencia de datos de precipitación es un problema común que surge con frecuencia en investigaciones ambientales y climáticas. Las estaciones meteorológicas a menudo pueden no registrar mediciones precisas de precipitación debido a diversos factores, como fallos en los equipos, cortes de energía u otros desafíos operativos. La estimación precisa de estos valores faltantes es un paso trascendental en la obtención de datos confiables, que permitirán mejorar nuestra comprensión de los patrones de precipitación, desarrollar modelos robustos para estudiar y predecir el cambio climático, y apoyar procesos críticos de toma de decisiones en campos como la gestión de recursos hídricos y la agricultura (Azman et al., 2021) (Zhan et al., 2019).

En Ecuador, el INAMHI (Instituto Nacional de Meteorología e Hidrología) es la organización encargada de monitorear y recopilar datos de la extensa red de estaciones meteorológicas del país. Sin embargo, a pesar de la importancia fundamental de estos datos de precipitación, un número significativo de estas estaciones han sido abandonadas o no funcionan correctamente. Como resultado, los datos que se recopilan a menudo son erróneos e incompletos, lo que presenta desafíos importantes para comprender con precisión los patrones de precipitación en las diferentes cuencas hidrográficas (Matovelle et al., 2023).

Asegurar la homogeneidad e identificar cualquier tendencia o anomalía en los datos utilizados para la modelización hidrológica o el análisis de recursos hídricos son pasos esenciales. Sin embargo, la naturaleza heterogénea de la precipitación, con su variabilidad espacial y temporal inherente, presenta un desafío significativo al momento de estimar con precisión los datos faltantes (Chávez De La Lima et al., 2022).

Se ha propuesto una variedad de métodos para estimar o imputar valores de precipitación faltantes, los cuales demuestran diferentes niveles de precisión y complejidad. Estas técnicas para estimar y reconstruir datos faltantes pueden categorizarse de manera general en métodos empíricos, métodos estadísticos y enfoques de funciones de ajuste. (Xia et al., 1999). La mayoría de estos métodos derivan los valores faltantes aprovechando las observaciones de las estaciones cercanas, teniendo en cuenta factores como las relaciones espaciales y temporales de los patrones de precipitación.

Seleccionar el método apropiado para estimar los datos de precipitación faltantes es crucial, ya que puede mejorar significativamente la precisión y confiabilidad de los subsecuentes modelos hidrológicos y sus resultados, lo que lleva a procesos de toma de decisiones mejor informados en campos como la gestión de los recursos hídricos y la adaptación al cambio climático.

Una de las técnicas más sencillas es el promedio aritmético, donde el valor faltante correspondiente a un registro se reemplaza con el promedio de los valores disponibles de las estaciones meteorológicas cercanas que cuenten con valores en el mismo registro. Este método es fácil de implementar, pero puede no funcionar bien en áreas con alta variabilidad espacial en la precipitación (Azman et al., 2021).

Otro método empleado es la regresión lineal múltiple, que utiliza los datos de precipitación disponibles de las estaciones circundantes como variables predictoras para estimar el valor faltante. Este enfoque puede capturar las relaciones espaciales en los datos, pero puede ser más complejo de implementar. (Jafri et al., 2020). Además de estos métodos clásicos reconocidos, existen varios trabajos que han investigado el uso de modelos de árboles de decisión para la imputación de precipitación faltante.

Los modelos de árbol de decisión son métodos no paramétricos que pueden capturar potencialmente relaciones no lineales en los datos y pueden ser más robustos a los valores atípicos en comparación con los enfoques de regresión lineal. En las últimas dos décadas, este tipo de métodos han ganado una popularidad significativa en la investigación hidrológica (Wangwongchai et al., 2023).

Estudios previos han examinado el rendimiento de estas diferentes técnicas en la estimación de datos de precipitación faltantes, algunos de estos concuerdan en que el promedio aritmético es el más preciso (Radi et al., 2015); mientras que otros han reportado que los modelos de regresión lineal múltiple y árboles de decisión son enfoques superiores (Wangwongchai et al., 2023).

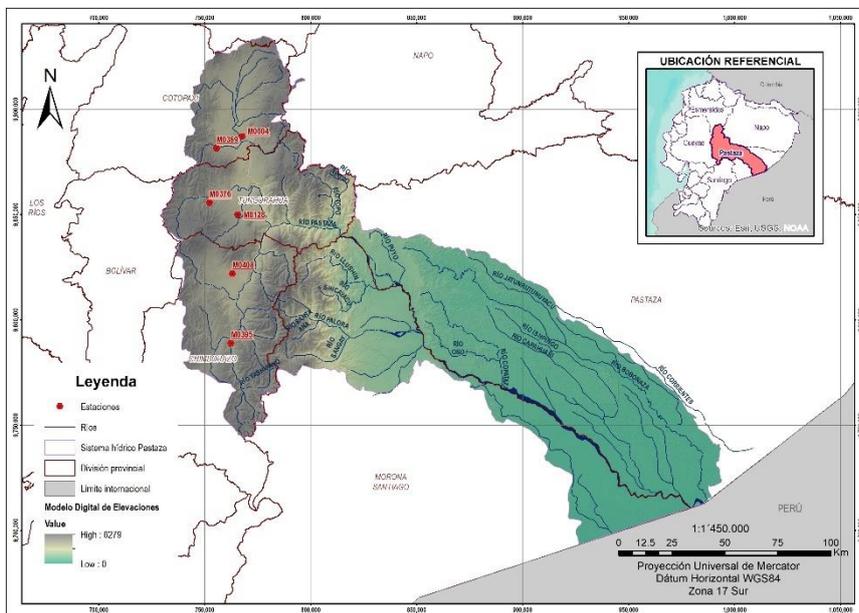
El objetivo de este estudio es basarse en la literatura técnica existente para llevar a cabo una comparación y evaluación exhaustiva del rendimiento de estos tres métodos, a saber, el promedio aritmético, la regresión lineal múltiple y el modelo de árbol de decisión, en el contexto de la imputación de los datos de precipitación ausentes procedentes de seis estaciones meteorológicas situadas en Ecuador.

## METODOLOGÍA

La cuenca del río Pastaza es parte de la Demarcación Hidrográfica Pastaza según consideraciones del Ministerio del Ambiente, Agua y Transición Ecológica. La cuenca del Pastaza se encuentra localizada entre las regiones andina y amazónica del Ecuador Continental. Las provincias andinas que forman parte de la hoya son: Cotopaxi, Tungurahua, Chimborazo y en menor proporción Pichincha.

La delimitación del área de estudio cubre la región andina de la cuenca del río Pastaza. Esta región, de aproximadamente 9134 km<sup>2</sup>, experimenta una temperatura media anual de 7 – 20 °C, mientras que la precipitación promedio anual está entre 600 - 2000 mm.

**Figura 1:** Área de estudio y localización de estaciones meteorológicas



Elaborado por: Elaboración propia

Para la investigación se procedió a descargar información oficial del INAMHI, disponible en: [www.datosabiertos.gob.ec/dataset/precipitacion-total-mensual/resource/98c77d18-e863-4e00-8a22-eb47f2981d9c](http://www.datosabiertos.gob.ec/dataset/precipitacion-total-mensual/resource/98c77d18-e863-4e00-8a22-eb47f2981d9c), y con esto conseguir acceso a los registros de precipitación de las diferentes estaciones meteorológicas consideradas. Este estudio utilizó datos de precipitación de seis estaciones meteorológicas en el período entre 1976 y 2018. Estas estaciones, designadas como M0004 (RUMIPAMBA - SALCEDO), M0128 (PEDRO FERMIN CEVALLOS - COLEGIO), M0369 (CUSUBAMBA), M0376 (PILAHUIN), M0395 (CEBADAS) y M0408 (GUANO), se encuentran dentro del área geográfica anteriormente mencionada y se las puede apreciar en la Figura 1.

La Tabla 1 presenta las coordenadas geográficas, las elevaciones y las características de los datos de precipitación mensual de las seis estaciones meteorológicas en el período previamente indicado.

**Tabla 1.** Estadísticas de los datos de precipitación y posición geográfica de las estaciones pluviométricas

Estación Meteorológica	Localización Geográfica			Resumen Estadística Descriptiva				
	Latitud °(N)	Longitud °(E)	Elevación (msnm)	Precipitación mínima (mm)	Precipitación máxima (mm)	Precipitación promedio (mm)	Desviación estándar (mm)	% de valores faltantes
<b>M0004</b>	1,02000	-78,59460	2685	0,0	155,5	45,02	33,47	1,16
<b>M0128</b>	-1,35611	-78,61278	2910	0,0	199,5	43,74	28,57	15,70
<b>M0369</b>	-1,07139	-78,70278	3175	0,0	252,6	45,52	40,07	4,07
<b>M0376</b>	-1,30453	-78,73281	3314	0,0	220,5	58,82	30,94	0,97
<b>M0395</b>	-1,90778	-78,64083	2930	0,0	304,3	47,84	51,82	8,91
<b>M0408</b>	-1,60917	-78,63528	2620	0,0	323,9	38,02	31,29	13,18

Elaborado por: Elaboración propia

Después del análisis estadístico y el control de calidad de los datos de precipitación disponibles, incluyendo pruebas de homogeneidad e identificación de cualquier tendencia; se evaluó y comparó la eficiencia de diferentes métodos estadísticos clásicos, como el promedio aritmético y la regresión lineal múltiple, además se tomó en cuenta en el estudio un modelo de árbol de decisión.

### Promedio Aritmético PA

El promedio aritmético es el método más sencillo comúnmente utilizado para llenar los datos meteorológicos faltantes en climatología. El valor ausente se calcula con el promedio de los datos disponibles de las estaciones meteorológicas más cercanas, como así se puede evidenciar en Ecuación 1:

$$P_o = \frac{\sum_{i=1}^n P_i}{n} \quad \text{Ecuación 1}$$

donde  $P_o$  es el valor estimado de los datos faltantes en un punto temporal estudiado,  $P_i$  es la precipitación en la  $i$ -ésima estación meteorológica más cercana, en el mismo registro temporal, y  $n$  es el número de estaciones más cercanas. El método PA funciona bien si las estaciones pluviométricas están distribuidas uniformemente en toda el área de estudio y las mediciones de las estaciones individuales no se desvían sustancialmente de la media.

Sin embargo, este enfoque puede no capturar adecuadamente la variabilidad espacial de la precipitación, particularmente en regiones con topografía compleja o distribución desigual de las estaciones meteorológicas (Chow et al., 1988).

## Regresión Lineal Múltiple RLM

La regresión lineal múltiple es un método estadístico empleado para estimar la relación entre una variable dependiente y dos o más variables independientes o predictoras. La RLM identifica la combinación ponderada de variables independientes para predecir la variable dependiente. Se ha demostrado que este método otorga muchas ventajas en la interpolación de datos y la estimación de datos faltantes (Gómez & Aguayo, 2019). El valor faltante ( $P_0$ ) puede ser estimado con Ecuación 2:

$$P_0 = \beta_0 + \beta_1 P_1 + \beta_2 P_2 + \dots + \beta_n P_n \quad \text{Ecuación 2}$$

donde  $\beta_0, \beta_1, \beta_2, \dots, \beta_n$  son los coeficientes de regresión.

### Árbol de decisión AD

Los árboles de decisión son un método de aprendizaje supervisado no paramétrico utilizado tanto para tareas de clasificación como de regresión. A diferencia del enfoque de regresión, que asume una relación lineal entre las variables independientes y dependientes, los árboles de decisión pueden capturar patrones no lineales en los datos. Los modelos de árboles de decisión dividen recursivamente el espacio de entrada en un conjunto de regiones y luego ajustan un modelo simple, como una constante, en cada región hoja (Sattari et al., 2016).

El proceso general de un árbol de decisión para la imputación de precipitación faltante implica los siguientes pasos:

1. Identificar los datos de precipitación disponibles de las estaciones meteorológicas circundantes como variables independientes y los datos de precipitación de la estación objetivo como variable dependiente.
2. Entrenar un modelo de regresión de árbol de decisión utilizando los datos disponibles.
3. Utilizar el modelo entrenado para predecir los valores de precipitación faltantes en la estación objetivo en función de las variables independientes.

La ventaja de los modelos de árboles de decisión es que pueden capturar relaciones complejas y no lineales entre las variables predictoras y la variable objetivo sin hacer suposiciones sobre la distribución subyacente de los datos.

## Métricas de Rendimiento

Para comparar la precisión de los métodos discutidos en la reconstrucción de datos de precipitación mensual faltantes, se utilizaron las siguientes tres métricas de rendimiento: raíz del error cuadrático medio (RMSE), error absoluto medio (MAE) y coeficiente de determinación ( $R^2$ ). El RMSE proporciona una medida de la raíz de la magnitud promedio de los errores cuadráticos, el MAE da la diferencia absoluta promedio entre los valores observados y estimados, y el  $R^2$  indica la proporción de la varianza en la variable dependiente que se puede predecir a partir de las variables independientes. Las expresiones empleadas fueron Ecuación (3) – (5).

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (y_i - x_i)^2}{n}} \quad \text{Ecuación 3}$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |x_i - y_i| \quad \text{Ecuación 4}$$

$$R^2 = \frac{(\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}))^2}{\sum_{i=1}^n (x_i - \bar{x})^2 * \sum_{i=1}^n (y_i - \bar{y})^2} \quad \text{Ecuación 5}$$

donde  $x$  es el valor observado e  $y$  denota el valor calculado.

Para llevar a cabo las pruebas, el análisis de los datos y la gestión de la información, se empleó el programa de código abierto Python, en su versión 3.13.0, aprovechando su amplia gama de bibliotecas especializadas en análisis estadístico y tratamiento de datos. Además, es importante destacar que el presente estudio se desarrolló dentro del marco de una investigación cuantitativa con un enfoque correlacional descriptivo.

## RESULTADOS Y DISCUSIÓN

Para garantizar la precisión y confiabilidad de los datos de precipitación utilizados en este estudio, fue importante aplicar dos pruebas estadísticas clave: la prueba de homogeneidad normal estándar (SNHT) y la prueba de tendencia de Mann-Kendall (MK). La SNHT se utilizó para detectar cambios o discontinuidades en la serie temporal de datos de precipitación mensual (Alexandersson, 1986). Mientras tanto, la prueba MK se empleó para evaluar si había tendencias monotónicas ascendentes o descendentes estadísticamente significativas en la variable de interés durante el período de estudio (Kendall, 1975).

Estas rigurosas medidas de control de calidad de los datos fueron un paso importante en la aprobación de los datos de precipitación para el posterior análisis comparativo de los diferentes métodos de imputación.

La prueba SNHT asumió la hipótesis nula de que los datos de precipitación eran homogéneos y la hipótesis alternativa de que los datos eran heterogéneos. Del mismo modo, la prueba de tendencia de MK planteó la hipótesis nula de que los datos mostraban aleatoriedad y ausencia de tendencias, frente a la hipótesis alternativa de que los datos mostraban no aleatoriedad y presencia de tendencias. Si el p-valor supera el nivel de significancia ( $\alpha$ ), no se puede rechazar la hipótesis nula; de lo contrario, se acepta la hipótesis alternativa. Como refleja la Tabla 2, se determinó que los datos de precipitación mensual en todas las estaciones, a excepción de M0376 y M0395, fueron estadísticamente homogéneos y carecían de tendencias significativas, lo que afirmó su idoneidad para el relleno de datos faltantes y el posterior análisis comparativo de los métodos de imputación.

**Tabla 2.** Resultados de la prueba de homogeneidad y tendencia de las estaciones seleccionadas

Estación	SNHT		MK		$\alpha$
	p-valor	Riesgo de rechazar $H_0$ (%)	p-valor	Riesgo de rechazar $H_0$ (%)	
M0004	0,9192	91,92	0,9096	90,96	0.05
M0128	0,6193	61,93	0,1326	13,26	0.05
M0369	0,8323	83,23	0,3034	30,34	0.05
M0376	0,0075	0,75	0,0051	0,51	0.05
M0395	0,0113	1,13	0,0001	0,01	0.05
M0408	0,0502	5,02	0,0748	7,48	0.05

Elaborado por: Elaboración propia

Analizar la relación entre los niveles de precipitación mensual en diferentes estaciones es relevante y útil para los propósitos de modelado. Por ello, se examinó la correlación lineal existente de la precipitación registrada en las estaciones M0004, M0128, M0369 y M0408. Los coeficientes de correlación calculados se clasificaron como moderados. Los resultados de este análisis se presentan en la

Tabla 3.



**Tabla 3.** Matriz de correlación de las estaciones investigadas

	<b>M0004</b>	<b>M0128</b>	<b>M0369</b>	<b>M0376</b>	<b>M0395</b>	<b>M0408</b>
<b>M0004</b>	1,0000000	0,5242914	0,7582032	0,5821128	0,5707585	0,6462359
<b>M0128</b>	0,5242914	1,0000000	0,4951450	0,7936810	0,5231268	0,5582562
<b>M0369</b>	0,7582032	0,4951450	1,0000000	0,5551818	0,5752333	0,6123940
<b>M0376</b>	0,5821128	0,7936810	0,5551818	1,0000000	0,5983971	0,5344677
<b>M0395</b>	0,5707585	0,5231268	0,5752333	0,5983971	1,0000000	0,5955841
<b>M0408</b>	0,6462359	0,5582562	0,6123940	0,5344677	0,5955841	1,0000000

Elaborado por: Elaboración propia

El conjunto de datos inicial presentaba diferentes proporciones de valores faltantes en las estaciones meteorológicas, como se puede visualizar en la Tabla 1. Para obtener un conjunto de datos completo, los registros con valores faltantes en cualquier estación, fue removido. Posteriormente, resultó necesario crear una serie de tiempo sintética, para la cual una parte de los datos se designó aleatoriamente como la fracción faltante. Estos datos faltantes se emplearon como un conjunto de prueba, mientras que los datos restantes se emplearon para entrenar los tres modelos de imputación.

El número de estaciones vecinas utilizadas en los métodos de imputación fue un aspecto clave a considerar. Los enfoques de promedio aritmético y regresión lineal múltiple utilizaron valores de todas las estaciones vecinas disponibles en el conjunto de datos. En contraste, el método de árbol de decisión implicó probar diversas combinaciones de parámetros de entrada en la totalidad de las estaciones meteorológicas para determinar la configuración que produjera el mejor rendimiento.

Los resultados obtenidos en el promedio aritmético, la regresión lineal múltiple, y el modelo de árbol de decisión, considerando la estación M0004, se presentan en la

Tabla 4. El método de promedio aritmético proporcionó el peor desempeño, con los valores más altos de RMSE y MAE, y el  $R^2$  más bajo. El enfoque de regresión lineal múltiple exhibió mejores resultados, con RMSE y MAE más bajos, y un  $R^2$  más alto.

El mejor desempeño del enfoque de regresión lineal múltiple puede atribuirse a que las estaciones meteorológicas involucradas se encuentran en entornos de elevación comparable y el área geográfica demuestran un régimen de precipitación relativamente constante.

**Tabla 4.** Valores de los criterios de rendimiento para distintos métodos de estimación de los datos pluviométricos

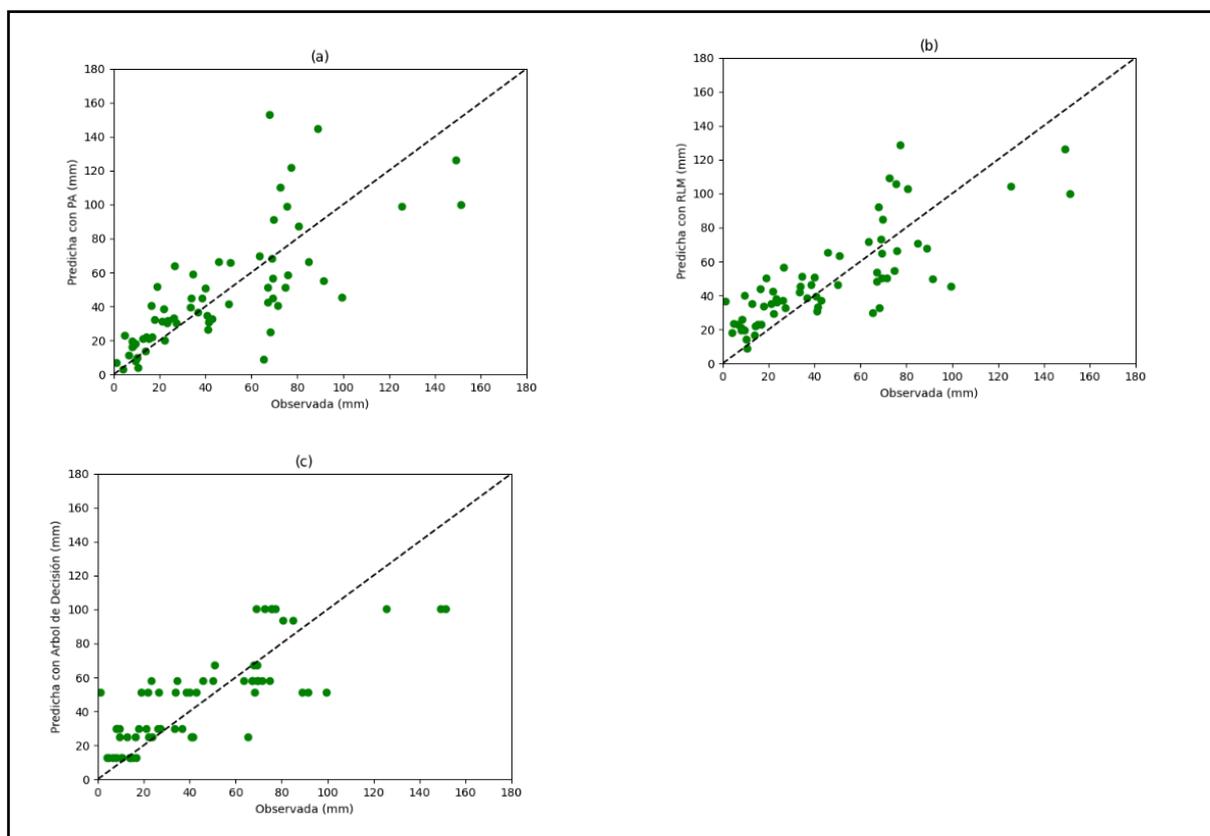
Método		R <sup>2</sup>	RMSE (mm)	MAE (mm)	Promedio (mm)	Desviación estándar (mm)
	Fase de prueba	-	-	-	45,65	34,68
Métodos estadísticos clásicos	AA	0,49	24,64	18,04	47,35	34,29
	RLM	0,61	21,50	17,63	49,64	27,84
Método de minería de datos	AD	0,62	21,15	16,45	47,95	27,84

Elaborado por: Elaboración propia

Por otra parte, el modelo de árbol de decisión mostró un desempeño ligeramente mejorado en las tres métricas de evaluación. La superioridad del modelo de árbol de decisión se puede atribuir a su capacidad para capturar relaciones no lineales e interacciones complejas entre las variables predictoras, que no son fácilmente capturadas por los métodos tradicionales como el de regresión lineal (Kim, 2008).

Los diferentes métodos de imputación se ilustran mediante diagramas de dispersión. La Figura 2 muestra la correlación entre los valores de precipitación observados y pronosticados para cada método.

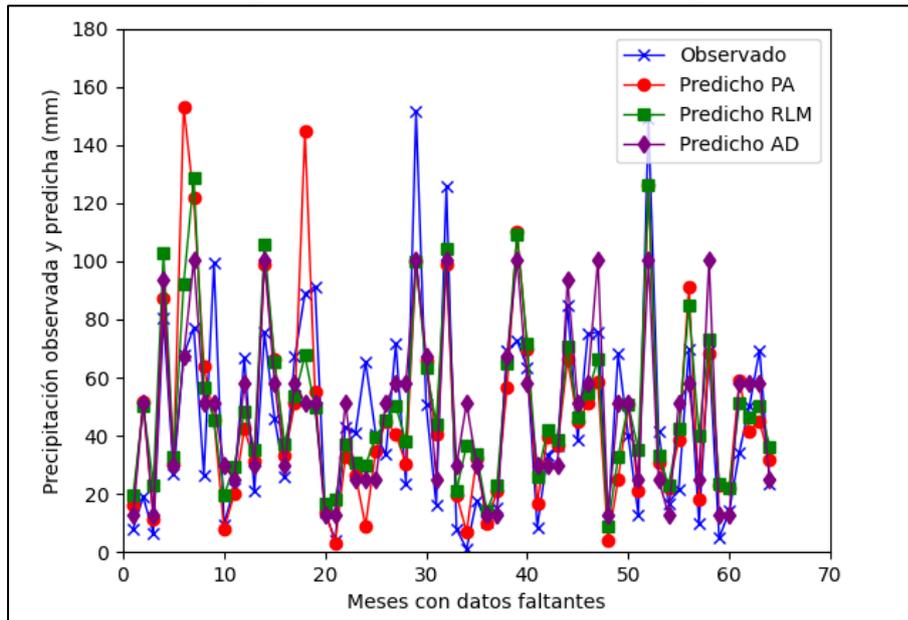
**Figura 2:** Diagrama de dispersión de los valores de precipitación predichos y observados generados por (a) PA, (b) RLM, (c) AD



Elaborado por: Elaboración propia

La Figura 3 muestra los patrones temporales de los datos de precipitación real e imputados por los modelos de promedio aritmético, regresión lineal múltiple y árbol de decisión. Se puede apreciar que el modelo de AD y la RLM proporcionan la aproximación más cercana a los valores de precipitación observados reales, en comparación con el modelo PA que en ciertos puntos incluso expone picos de precipitación.

**Figura 3:** Series temporales de valor observado y valores predichos



Elaborado por: Elaboración propia

## CONCLUSIONES

El análisis presentado en este documento utilizó los datos de precipitación mensual, desde 1976 hasta 2008, de seis estaciones meteorológicas ubicadas en las regiones andinas de la Demarcación Hidrográfica Pastaza en Ecuador como base inicial de la investigación.

Los datos inicialmente recopilados, excluyendo las observaciones de las estaciones M0376 y M0395, fueron estadísticamente homogéneos y carecieron de tendencias significativas, lo que indica su idoneidad para el análisis comparativo de los métodos de imputación. Sin embargo, el conjunto de datos presentó diferentes proporciones de valores faltantes en las estaciones meteorológicas. El análisis mostró que, de los dos métodos estadísticos tradicionalmente utilizados para imputar los datos faltantes, la regresión lineal múltiple exhibió el mejor rendimiento en comparación con el enfoque de promedio aritmético.

Los resultados sugieren que el método de regresión lineal múltiple es adecuado para estimar los datos de precipitación faltantes. Este hallazgo es consistente con estudios previos (Xia et al., 1999) (Wangwongchai et al., 2023), que también han reportado el enfoque de imputación RLM como una técnica confiable para ser empleado en las observaciones de precipitación faltantes.

Finalmente, los hallazgos de este estudio indican que el modelo de regresión de árbol de decisión exhibió un desempeño predictivo comparable con el modelo RLM, evidenciado por un coeficiente de correlación de 0.62, un error cuadrático medio de 21.15 mm y un error absoluto medio de 16.45 mm. Dada la precisión del enfoque de árboles de decisión, se recomienda como un método adecuado para imputar los valores de precipitación faltantes en las regiones andinas de Ecuador Continental.

## REFERENCIAS BIBLIOGRÁFICAS

- Alexandersson, H. (1986). A homogeneity test applied to precipitation data. *Journal of Climatology*, 6(6), 661-675. <https://doi.org/10.1002/joc.3370060607>
- Azman, A. H., Tukimat, N. N. A., & Malek, M. A. (2021). Comparison of Missing Rainfall Data Treatment Analysis at Kenyir Lake. *IOP Conference Series: Materials Science and Engineering*, 1144(1), 012046. <https://doi.org/10.1088/1757-899X/1144/1/012046>
- Chávez De La Lima, J. D., Olgún López, J. L., Guevara Gutiérrez, R. D., Mancilla Villa, Ó. R., Palomera García, C., & Cruz-Sandoval, G. (2022). Consistencia, homogeneidad y distribución de la precipitación pluvial y temperatura, Región Sierra de Amula, Jalisco, México. *Revista Geográfica de América Central*, 68, 353-378. <https://doi.org/10.15359/rgac.68-1.13>
- Chow, V. T., Maidment, D. R., & Mays, L. W. (1988). *Applied Hydrology*. McGraw-Hill.
- Gómez, J. S., & Aguayo, 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/ig.59837>
- Jafri, I. A. M., Noor, N. M., Ul-Saufie, A. Z., & Suwardi, A. (2020). Prediction of Missing Data in Rainfall Dataset by using Simple Statistical Method. *IOP Conference Series: Earth and Environmental Science*, 616(1), 012005. <https://doi.org/10.1088/1755-1315/616/1/012005>
- Kendall, M. G. (1975). *Rank correlation methods* (4th ed., 2d impression). Griffin.

- Kim, Y. S. (2008). Comparison of the decision tree, artificial neural network, and linear regression methods based on the number and types of independent variables and sample size. *Expert Systems with Applications*, 34(2), 1227-1234. <https://doi.org/10.1016/j.eswa.2006.12.017>
- Matovelle, C., Heras, D., & Solano Peláez, J. (2023). Eficiencia de la Imputación de Datos Faltantes de Precipitaciones Utilizando Herramientas Computacionales en la Cuenca Hidrográfica, Jubones-Ecuador. *Revista Politécnica*, 50(2), 23-30.
- Radi, N. F. A., Zakaria, R., & Azman, M. A. (2015). Estimation of missing rainfall data using spatial interpolation and imputation methods. *AIP Conference Proceedings*, 1643(1), 42-48. <https://doi.org/10.1063/1.4907423>
- Sattari, M.-T., Rezazadeh-Joudi, A., & Kusiak, A. (2016). Assessment of different methods for estimation of missing data in precipitation studies. *Hydrology Research*, 48(4), 1032-1044. <https://doi.org/10.2166/nh.2016.364>
- Wangwongchai, A., Waqas, M., Dechpichai, P., Hlaing, P. T., Ahmad, S., & Humphries, U. W. (2023). Imputation of missing daily rainfall data; A comparison between artificial intelligence and statistical techniques. *MethodsX*, 11, 102459. <https://doi.org/10.1016/j.mex.2023.102459>
- Xia, Y., Fabian, P., Stohl, A., & Winterhalter, M. (1999). Forest climatology: Estimation of missing values for Bavaria, Germany. *Agricultural and Forest Meteorology*, 96(1), 131-144. [https://doi.org/10.1016/S0168-1923\(99\)00056-8](https://doi.org/10.1016/S0168-1923(99)00056-8)
- Zhan, C., Wu, F., Wu, Z., & Tse, C. K. (2019). Daily Rainfall Data Construction and Application to Weather Prediction. *2019 IEEE International Symposium on Circuits and Systems (ISCAS)*, 1-5. <https://doi.org/10.1109/ISCAS.2019.8702124>