

Análisis de la variación del carbono orgánico del suelo en ecosistemas altoandinos del Ecuador: un estudio con enfoque estadístico multivariante.

Fernando Ricardo Márquez Sañay ¹  Marco Antonio Gavilanes Sagñay*² 

¹ Universidad Politécnica Estatal del Carchi, Riobamba, EC060150, Ecuador

² Escuela Superior Politécnica de Chimborazo, Riobamba, EC060150, Ecuador

✉ Correspondencia: fernando.marquez@esPOCH.edu.ec  + 593 999697304

DOI/URL: <https://doi.org/10.53313/gwj92335>

Resumen: El análisis del Carbono Orgánico del Suelo (COS) en los ecosistemas altoandinos del Ecuador es fundamental debido al importante papel que desempeñan en el proceso físico químico de descomposición de materia, los cambios climáticos, la fertilidad del suelo e interrelaciones con el crecimiento y desarrollo de los seres vivos. Estos ecosistemas, en la actualidad, son considerados importantes reservorios de COS, ya que en estudios previos se ha demostrado su alta capacidad de conservación ambiental, sin embargo, aún quedan vacíos investigativos relacionados a la interacción de las variables en la presencia del carbono en el suelo. En el presente trabajo se propuso analizar la variabilidad del COS de los ecosistemas altoandinos ecuatorianos (región Sierra) mediante el empleo de técnicas estadísticas aplicadas. En primer lugar, se utiliza la base de datos HSDE y la información recopilada de fuentes primarias como trabajos de titulación desarrolladas en la temática, en el área de la investigación de los ecosistemas altoandinos. Luego de la validación de datos, se utilizaron técnicas de estadística descriptiva, análisis exploratorios, análisis de varianza (ANOVA), análisis de regresión lineales múltiples, Análisis aditivos generalizados (GAM) y el análisis de componentes principales (ACP), en el software R. Los resultados analizados en el estudio demuestran que los niveles de COS son muy variables. El ACP explicó el 72.4% de la variabilidad total del sistema, identificando las variables más asociadas con la distribución de COS como la profundidad del muestreo y la densidad aparente. Mediante la comparación de modelos se pudo establecer también que el modelo GAM tuvo un mejor desempeño predictivo ($R^2 = 0.749$) que el modelo lineal múltiple ($R^2 = 0.597$), indicando que existe una relación no lineal entre el COS y las variables analizadas y consideradas en la investigación. En el estudio en general, los resultados comprueban que el uso de estadística multivariante permite identificar y explicar de manera acertada las variables influyentes en la presencia de COS en ecosistemas altoandinos del Ecuador. Esto es



Check for updates

Cita: Márquez Sañay, F. R., & Gavilanes Sagñay, M. A. (2026). Análisis de la variación del carbono orgánico del suelo en ecosistemas altoandinos del Ecuador: un estudio con enfoque estadístico multivariante. *Green World Journal*, 08(02), 335. <https://doi.org/10.53313/gwj92335>

Received: 15/Abril /2026
Accepted: 13/Mayo /2026
Published: 25/Mayo /2026

Prof. Carlos Mestanza-Ramón, PhD.
Editor-in-Chief / CaMeRa Editorial
editor@greenworldjournal.com

Editor's note: CaMeRa remains neutral with respect to legal claims resulting from published content. The responsibility for published information rests entirely with the authors.



© 2026 CaMeRa license, Green World Journal. This article is an open access document distributed under the terms and conditions of the license.

Creative Commons Attribution (CC BY).

<http://creativecommons.org/licenses/by/4.0>

importante para nuevas investigaciones en la temática y para generar estrategias de conservación y el mejor manejo de los suelos.

Palabras clave: Carbono orgánico en suelo , ecosistemas altoandinos, análisis multivariante, modelos estadístico, análisis de componentes principales.

Analysis of soil organic carbon variation in high Andean ecosystems in Ecuador: a study using a multivariate statistical approach.

Abstract: The analysis of soil organic carbon (SOC) in Ecuador's high Andean ecosystems is essential due to the important role it plays in the physicochemical process of matter decomposition, climate change, soil fertility, and its interrelationships with the growth and development of living organisms. These ecosystems are currently considered important reservoirs of SOC, as previous studies have demonstrated their high capacity for environmental conservation; however, there are still gaps in research regarding the interaction of variables affecting the presence of carbon in the soil. This study aimed to analyze the variability of soil organic carbon (SOC) in Ecuadorian high-Andean ecosystems (Sierra region) using applied statistical techniques. First, the HSDE database and information compiled from primary sources, such as degree theses on the subject, were utilized in the research on high-Andean ecosystems. After data validation, descriptive statistical techniques, exploratory analyses, analysis of variance (ANOVA), multiple linear regression, generalized additive models (GAM), and principal component analysis (PCA) were applied using the R software. The results analyzed in the study demonstrate that COS levels are highly variable. The PCA explained 72.4% of the total variability in the system, identifying sampling depth and bulk density as the variables most strongly associated with COS distribution. Through model comparison, it was also established that the GAM model had better predictive performance ($R^2 = 0.749$) than the multiple linear model ($R^2 = 0.597$), indicating that there is a nonlinear relationship between COS and the variables analyzed and considered in the study. Overall, the results of the study confirm that the use of multivariate statistics allows for the accurate identification and explanation of the variables influencing the presence of COS in high-Andean ecosystems. This is important for future research on the topic and for developing conservation strategies and improved soil management practices.

Keywords: Soil Organic Carbon (SOC), high Andean ecosystems, multivariate analysis, statistical models, principal component analysis.

1. Introducción

El cambio climático representa uno de los mayores retos ambientales contemporáneos, impulsado principalmente por el aumento sostenido de las concentraciones de gases de efecto invernadero, en particular el dióxido de carbono (CO_2). En este contexto, los suelos desempeñan un papel fundamental dentro del ciclo global del carbono, ya que constituyen el mayor reservorio terrestre de carbono, almacenando más carbono que la atmósfera y la biomasa vegetal combinadas. [1] [2] La presencia del Carbono orgánico en el suelo (COS) es, por tanto, un componente importante de estudio para comprender los mecanismos de atenuación del cambio climático y la estabilidad de los ecosistemas.

El COS teóricamente se ha determinado que está asociado a las propiedades físicas, químicas y biológicas del suelo, influyendo en procesos clave como la fertilidad del suelo, la estructura edáfica, etc. Su acumulación y estabilidad dependen de complejas interacciones entre el clima, la topografía, la vegetación y las características edáficas, lo que genera patrones espaciales altamente heterogéneos[2]. En ecosistemas de montaña, estas interacciones se intensifican debido a fuertes gradientes altitudinales y climáticos, donde pequeñas variaciones de temperatura o humedad pueden modificar significativamente las tasas de descomposición de la materia orgánica y, en

consecuencia, la distribución del COS [3] Los ecosistemas altoandinos, son por lo tanto fuentes de carbono debido a sus condiciones ambientales particulares, caracterizadas por bajas temperaturas, alta humedad y limitada actividad microbiana. En este grupo destacan ecosistemas como: páramos, humedales, turberas entre otros[4]. Estos ecosistemas no solo contribuyen a la mitigación del cambio climático, sino que también desempeñan funciones clave relacionadas con la regulación de fuentes hidrográficas, la conservación de la biodiversidad y el sostenimiento local[5].

En el Ecuador, los ecosistemas altoandinos, generalmente ubicados por encima de los 3 000 m s. n. m., comprenden páramos, bosques altoandinos, matorrales, pastizales y turberas, los cuales presentan una notable capacidad de almacenamiento de carbono orgánico en los horizontes superficiales del suelo. Estudios empíricos han reportado concentraciones y stocks elevados de COS, con valores que superan los 150 Mg C ha⁻¹ en los primeros 30 cm del perfil del suelo, particularmente en áreas poco intervenidas [6] [7]. Luego, en ecosistemas altoandinos, la altitud puede influir en la distribución del COS debido a sus efectos sobre la temperatura, la humedad, la cobertura vegetal, etc.; sin embargo, la intensidad de esta relación pueden variar según el tipo de ecosistema, el uso del suelo y las propiedades características, por lo que requiere ser evaluada empíricamente en cada contexto de estudio. En el apartado de resultados del presente estudio (ecosistemas altoandinos de la sierra ecuatoriana) se evidenciará la influencia de la altitud del ecosistema en la presencia del COS, de acuerdo con los datos obtenidos previamente y análisis estadísticos realizados.

En los últimos años, se han desarrollado investigaciones relevantes a escala local, regional y nacional que han abordado la variabilidad del carbono orgánico del suelo en ecosistemas altoandinos del Ecuador.[6] Por ejemplo, se aplicaron modelos de *Random Forest* para el mapeo predictivo del COS en el páramo de Chimborazo, identificando como variables explicativas clave a la altitud, la precipitación, la unidad geológica y las propiedades edáficas, con coeficientes de determinación superiores a 0,75. De manera complementaria, en estudios previos [7] cuantificaron el almacenamiento de carbono orgánico en humedales altoandinos del páramo Atillo, reportando elevados stocks de carbono tanto en horizontes superficiales como subsuperficiales, lo que evidenció el alto potencial de estos ecosistemas como sumideros de carbono. Otros estudios han mostrado el impacto del uso y cambio de cobertura del suelo sobre el COS. Graham y otros. [8] mostraron que la conversión de páramos a plantaciones forestales disminuye de manera drástica el contenido de carbono orgánico y la capacidad de retención hídrica del suelo. Otros estudios [9] mostraron grandes pérdidas de COS debido a la expansión de la frontera agrícola en los páramos del norte del Ecuador. Recientes estudios [4] han mostrado que las turberas altoandinas, aunque son muy vulnerables a la degradación, contienen grandes cantidades de carbono, y que, por tanto, en procesos de degradación y drenaje, se puede perder, de manera irreversible, una importante cantidad de carbono.

A pesar de estos aportes, los antecedentes revisados muestran que la mayoría de los estudios se han desarrollado a escalas espaciales limitadas y han empleado enfoques metodológicos parciales, centrados en un único tipo de modelo estadístico o predictivo. Además, los estudios previos demuestran que las variables analizadas no están integradas de manera simultánea, por ejemplo: variables edáficas, topográficas, climáticas y de cobertura del suelo dentro de un marco estadístico multivariante, así también evidencia una limitada comparación entre modelos estadísticos explicativos y modelos predictivos [10]. Estas limitaciones evidencian la necesidad de

investigaciones que articulen de manera sistemática los distintos factores ambientales para explicar y predecir la variabilidad del COS en ecosistemas altoandinos. Es así como, el presente estudio plantea como hipótesis principal que la variación del carbono orgánico del suelo en los ecosistemas altoandinos del Ecuador está significativamente asociada a la interacción entre factores edáficos, topográficos y de cobertura del suelo. Luego, un enfoque estadístico multivariante permite explicar y predecir dicha variabilidad de forma más robusta que los análisis univariados. En consecuencia, el objetivo principal de este trabajo es analizar la variación del carbono orgánico del suelo en ecosistemas altoandinos del Ecuador mediante un enfoque estadístico multivariante, integrando variables relevantes de los ecosistemas, y evaluando tanto su capacidad explicativa como predictiva.

Finalmente, para lograr este objetivo, se utilizan bases de datos nacionales y fuentes abiertas validadas, se aplican técnicas de análisis exploratorio, de modelos explicativos (análisis de varianza factorial y modelos lineales mixtos) y modelos predictivos, los cuales son evaluados en función de determinadas métricas de desempeño en estadística. Esta metodología permite lograr una comprensión y una replicabilidad que, no solo será útil para los técnicos en conservación y restauración de los ecosistemas, sino para los científicos en otras disciplinas que se relacionan con el estudio del carbono en los ecosistemas de montaña. A través de ese enfoque, se logra identificar de manera precisa los factores que controlan la distribución del carbono orgánico de la zona, y además, se establece un marco comprensible y reproducible sobre el control del carbono en los ecosistemas de montaña. La aportación científica de este trabajo se encuentra en la integración sistemática de diferentes bloques de variables y enfoques estadísticos, lo que permite plasmar con más precisión los factores que controlan el COS y, además, generar información útil para la formulación de políticas de conservación y de mitigación del cambio climático en los ecosistemas altoandinos del Ecuador.

2. Materiales y métodos

2.1 Área de estudio

El estudio se desarrolló a partir de registros edáficos correspondientes a ecosistemas (suelos) altoandinos del Ecuador, considerando principalmente zonas ubicadas por encima de los 3000 m.s.n.m. Para este análisis se consideraron registros localizados en la región Sierra del Ecuador, con especial representación de las provincias de Chimborazo y Tungurahua. La recopilación de datos fue realizada en fuentes primarias de información como trabajos de titulación relacionados al estudio del COS en la sierra del Ecuador. En específico, se revisaron 50 trabajos de titulación de grado y posgrado de los cuales se extrajeron registros asociados con carbono orgánico del suelo, altitud, profundidad de muestreo, densidad aparente, pH, textura y ubicación geográfica.

El análisis se basó en registros edáficos provenientes de la Harmonized Soil Database of Ecuador [11], [12], una base de datos que integra información de perfiles de suelo recopilados en diferentes regiones del país. Esta base de datos contiene variables asociadas con las propiedades físicas, químicas y morfológicas del suelo. Esto incluye mediciones de carbono orgánico, textura, densidad aparente, profundidad del horizonte, pH, etc.

Los perfiles considerados corresponden a sitios ubicados en gradientes de altitud representativos de ecosistemas altoandinos, lo que permite evaluar la variabilidad espacial del COS en función de variables ambientales y edáficas relevantes

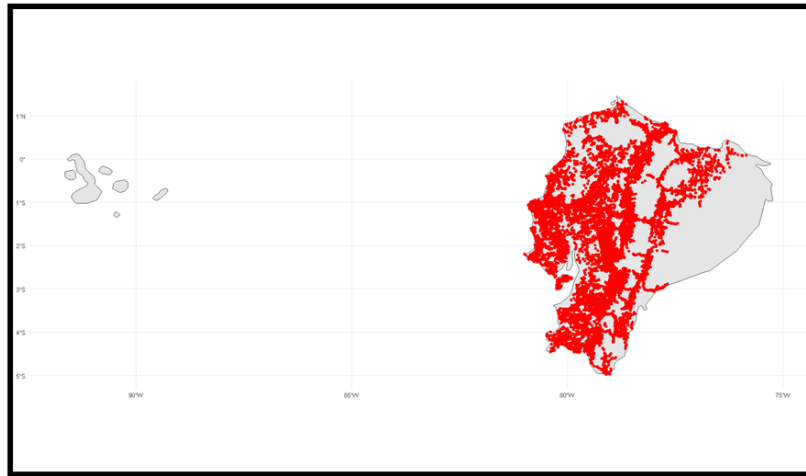


Figura 1 HSDE (puntos muestreados según el Mapa de COS hasta el año 2021)

2.2 Métodos

El estudio adoptó un enfoque cuantitativo con análisis estadístico multivariante, orientado a identificar los factores edáficos y ambientales que explican la variabilidad del carbono orgánico del suelo en ecosistemas altoandinos.

Base de datos y selección de variables

En Ecuador la base HSDE armoniza 13542 perfiles y 92 variables [11]. Para la construcción de la base analítica se aplicaron los siguientes criterios de selección: registros ubicados en ecosistemas altoandinos o zonas de influencia altoandina; observaciones con información disponible de carbono orgánico del suelo; registros con altitud, profundidad de muestreo o variables edáficas asociadas; datos expresados en unidades comparables o susceptibles de transformación; y exclusión de registros duplicados, inconsistentes o sin información mínima para el análisis estadístico. Luego de este proceso, se consolidaron 163 observaciones de carbono orgánico del suelo provenientes de tesis y estudios recopilados, además de los registros disponibles en la HSDE utilizados como referencia y contraste espacial.

Las variables consideradas incluyeron:

- Carbono orgánico del suelo (COS) (%)
- Altitud del sitio (m s.n.m.)
- Profundidad del horizonte (cm)
- Textura del suelo (% de arena, limo y arcilla)
- Densidad aparente (g cm^{-3})
- pH del suelo
- Contenido de materia orgánica (%)

Por otra parte, en algunos estudios no se pudo obtener todos los datos, debido a las características de la investigación y sus objetivos, sin embargo, fueron consideradas para el análisis descriptivo de las zonas estudiadas y posterior análisis comparativo de resultados. Las variables consideradas de los datos recopilados se resumen en la tabla 1.[4], [13]

Tabla 1. Resumen de información obtenida de trabajos de titulación (a nivel nacional)

Categoría	Elemento	Descripción
Fuente de información	Trabajos de titulación compilados	50 tesis de grado relacionadas con estudios de suelo y carbono en ecosistemas altoandinos del Ecuador
	Número total de registros	163 observaciones de carbono orgánico del suelo provenientes de trabajos de titulación; los registros de la HSDE se utilizaron como base complementaria y de referencia espacial, según disponibilidad de variables.
Cobertura espacial	Ubicación de los estudios	Provincias altoandinas del Ecuador (principalmente Chimborazo y Tungurahua)
	Rango altitudinal	Aproximadamente entre 2590 y 4200 msnm, filtrados a altitudes mayores a 3000 msnm
Variables registradas	Número total de variables	24 variables edáficas y ambientales
	Variables de carbono	Carbono orgánico del suelo (%), carbono orgánico del suelo (ton/ha), carbono en biomasa aérea
	Variables edáficas	Densidad aparente, pH del suelo, textura del suelo, conductividad eléctrica
	Variables ambientales	Altitud, ubicación geográfica, ecosistema
Profundidad de muestreo	Número de espesores registrados	Niveles de profundidad
	Espesores de muestreo	Rangos principales: 0–10 cm, 10–20 cm, 15–30 cm, 20–30 cm y >30 cm, según disponibilidad de cada fuente.
Variables derivadas para el análisis	Transformación de carbono	Conversión de carbono orgánico del suelo (%) a Mg por hectárea (ton/ha)
	Variables utilizadas en los modelos	Altitud, profundidad de muestreo, densidad aparente y pH del suelo

Las variables fueron seleccionadas debido a su influencia en los procesos de acumulación, estabilización y distribución del carbono en el suelo, así como los posibles cálculos de Carbono orgánico presentes en cada punto estudiado.

En este estudio se emplea el término “observación” para referirse a cada registro individual de carbono orgánico del suelo asociado con una profundidad, altitud y conjunto de variables edáficas. Debido a que algunas fuentes reportaron varios horizontes o profundidades para un mismo sitio, el número de observaciones no necesariamente equivale al número de puntos geográficos únicos de muestreo. Las profundidades de muestreo reportadas en las fuentes originales fueron heterogéneas; por ello, se organizaron en niveles comparables para el análisis. En total se identificaron siete niveles de profundidad, que fueron agrupados analíticamente en rangos representativos: 0–10 cm, 10–20 cm, 15–30 cm, 20–30 cm y profundidades mayores a 30 cm, según la disponibilidad y estructura de cada fuente. Esta clasificación permitió comparar la

distribución vertical del COS y evaluar el efecto de la profundidad de muestreo sobre su variabilidad.

Procesamiento de datos

El procesamiento de los datos se realizó mediante el software RStudio [14], utilizando paquetes especializados para análisis estadístico y visualización de datos. Inicialmente se efectuó un análisis exploratorio de datos, que incluyó estadísticas descriptivas, evaluación de valores atípicos y análisis de correlación entre variables edáficas.

Posteriormente, se aplicaron técnicas de análisis multivariante, con el objetivo de identificar patrones de variación y relaciones entre el carbono orgánico del suelo y los factores edáficos asociados. Entre los métodos utilizados se incluyeron:

- Análisis de componentes principales (ACP), para reducir la dimensionalidad del conjunto de variables y detectar gradientes ambientales dominantes.[15]
- Análisis de correlación, para evaluar la relación entre el carbono orgánico del suelo y las variables edáficas.
- Modelos estadísticos multivariantes, orientados a cuantificar la influencia de factores como altitud, textura del suelo y profundidad sobre la variabilidad del carbono orgánico.[16]

Finalmente, los resultados se representaron mediante gráficos multivariantes y diagramas de dispersión, que permitieron visualizar las relaciones entre variables y los principales gradientes ambientales asociados a la variabilidad del carbono orgánico del suelo.

Análisis estadístico:

El procesamiento de los datos se realizó mediante el software RStudio, utilizando paquetes especializados para análisis estadístico y visualización de datos. Inicialmente se efectuó un análisis exploratorio de datos, que incluyó estadísticas descriptivas, evaluación de valores atípicos y análisis de correlación entre variables edáficas.

Fase Inicial: Preparación de datos y análisis descriptivo

- Recopilación de datos de fuentes confiables como HESD, SoilGrids, mapa COS Ecuador, verificando variables, valores y unidades. (Verificación de fuentes, DOI, etc)
- Limpieza de datos y filtración a valores mayores a 3000 msnm
- Tratamiento de datos e ingreso a software RStudio.
- Descripción de datos generales recopilados, mediante estratificación: bandas altitudinales (3000–3400; 3400–3800; >3800 m), bandas de profundidad (0–10; 10–20; 20–30 cm) y clases de cobertura.
- Descriptivos para carbono orgánico medidos en toneladas por hectárea: media aritmética, mediana, rango intercuartilar, rango interpercentil, mín, máx por altitud, por cobertura, etc. Características edáficas: ph, textura, etc.[17]
- Visualizaciones estadísticas: diagramas de caja y bigotes de carbono orgánico (por altitud y profundidad), dispersión carbono vs altitud, perfiles verticales (carbono vs profundidad y cobertura) [18]

- Correlación multivariante y exploratoria: Prueba "Spearman" entre carbono orgánico en el suelo y predictores continuos.

Fase 2 Cuantificación de efectos e interacciones (Desarrollo estadístico aplicado)

- Diagnóstico previo
- ANOVA factorial (explicativo)[19]
- Regresión lineal múltiple (de ser posible)
- Comparación de modelos.

Fase 3: Predicción, selección y validación del modelo estadístico

Modelos estadísticos [20]

Modelo lineal múltiple es una técnica estadística utilizada para explicar la relación entre una variable dependiente y varias variables independientes. Este modelo asume que la relación entre las variables es lineal, permitiendo estimar el efecto individual de cada predictor sobre la variable respuesta. [21]

Modelo GAM: Los modelos aditivos generalizados (GAM) son una extensión flexible de los modelos lineales generalizados que permiten modelar relaciones no lineales entre las variables mediante funciones suavizadas.[22]

- Evaluación del desempeño mediante métricas RMSE, MAE, R^2 (con desviación estándar e intervalos de confianza).
- Diagnóstico espacial de residuos del mejor modelo).
- Intervalo de confianza del 95% de efectos (ANOVA/LMM/GAM).

3. Resultados

3.1. Estadística descriptiva respecto a los datos analizados

3.1.1. Determinación de COS relacionada a la altitud del ecosistema analizado

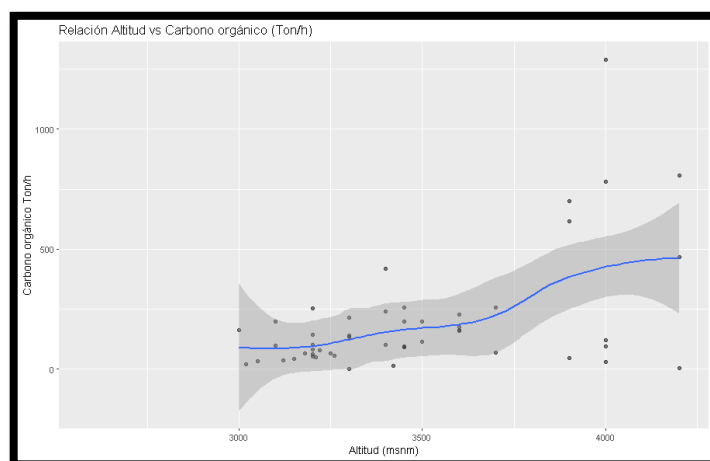


Figura 2 Valores de COS en función de la altitud de los ecosistemas analizados
Fuente: Recopilación de Tesis de grado ESPOCH

Las figuras 2 y 3 muestra la relación entre la altitud (msnm) y el carbono orgánico del suelo (ton/ha) en ecosistemas altoandinos del Ecuador. Cada punto representa una observación individual proveniente de los diferentes estudios recopilados en la base de datos. La línea azul representa una tendencia suavizada (LOESS) que permite observar el comportamiento general de la relación entre ambas variables, mientras que la banda gris corresponde al intervalo de confianza de la tendencia estimada. Se observa una relación positiva entre la altitud y el carbono orgánico del suelo. En términos generales, a medida que la altitud aumenta, también se incrementa el contenido de carbono orgánico almacenado en el suelo. El incremento del carbono orgánico con la altitud puede explicarse por varios factores característicos de los ecosistemas altoandinos: Temperaturas más bajas, que reducen la tasa de descomposición de la materia orgánica, mayor acumulación de residuos vegetales en ecosistemas como páramos o bosques altoandinos, condiciones de humedad y saturación del suelo, que favorecen la conservación del carbono en el perfil edáfico.

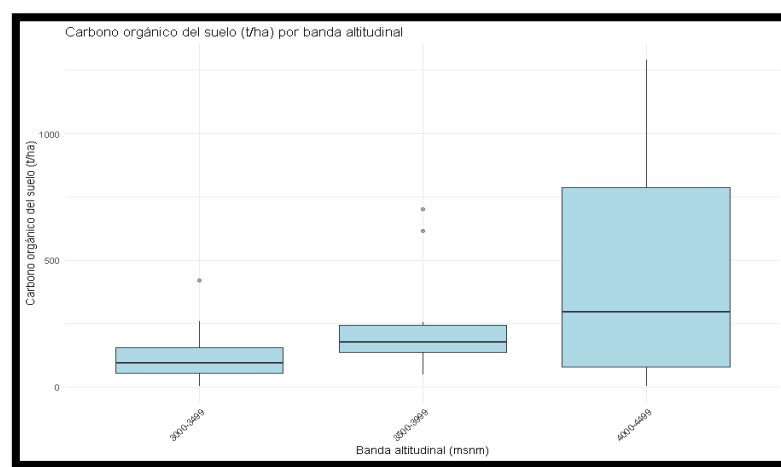


Figura 3. Boxplot relacionada a cantidad de COS respecto a altitud del ecosistema

3.1.2 Determinación de COS relacionada a la altitud del ecosistema analizado

Tabla 2 Resultados de obtención de COS en función de la profundidad de muestreo

Banda altitudinal (msnm)	n	Media COS (ton/ha)	Desv. Estándar	Mediana	Q25	Q75
2500–2999	0	NA	NA	NA	NA	NA
3000–3499	31	116	91.3	93	54.4	153
3500–3999	11	248	213	177	137	242
4000–4499	8	449	471	294	79.5	787
Sin clasificación	34	124	152	59.7	39.9	138

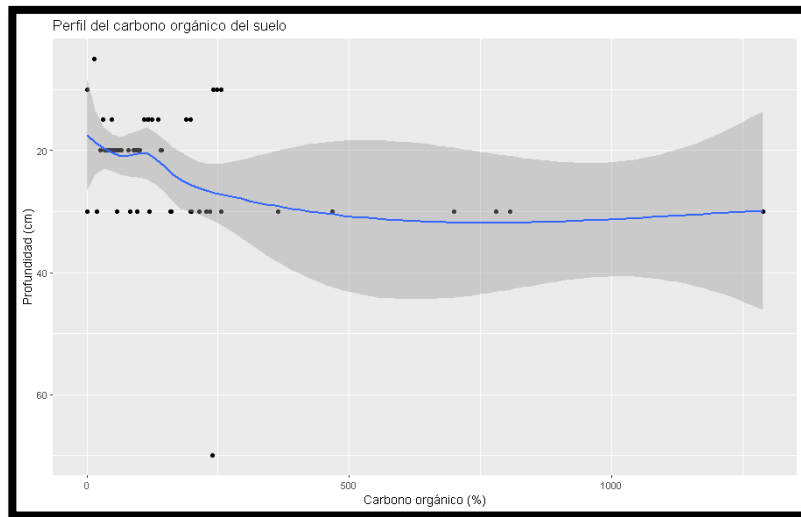


Figura 4 Valores de COS en función de la profundidad del muestreo de los ecosistemas analizados
Fuente: Recopilación de Tesis de grado ESPOCH

La figura 4 presenta la relación entre el contenido de carbono orgánico del suelo y la profundidad de muestreo (cm). Cada punto representa una observación individual del conjunto de datos recopilado en diferentes estudios de ecosistemas altoandinos. La línea azul corresponde a una curva suavizada (LOESS) que muestra la tendencia general del comportamiento del carbono en función de la profundidad, mientras que la banda gris representa el intervalo de confianza de la estimación. El gráfico muestra una relación inversa entre la profundidad del suelo y el contenido de carbono orgánico. A menor profundidad (capas superficiales), mayor contenido de carbono orgánico. A mayor profundidad, el contenido de carbono tiende a disminuir.

3.1.3 Estadísticos descriptivos importantes y correlaciones

Tabla 3 Resumen de descriptivos de COS en ecosistemas analizados

Estadístico Descriptivo	Valor (ton/ha)
Mínimo	0.09
Primer cuartil (Q1)	50.36
Mediana	93.95
Media	168.56
Tercer cuartil (Q3)	198.11
Máximo	1289.25
Valores faltantes (NA)	79

Los resultados de la tabla 3 muestran la variabilidad de cada uno de los registros de carbono orgánico del suelo que fueron presentes en la base de datos. La media establecida en 168.56 Ton/h, establece una gran dispersión entre los puntos analizados. Se identificaron 79 valores que fueron borrados o filtrados antes de usar modelos estadísticos

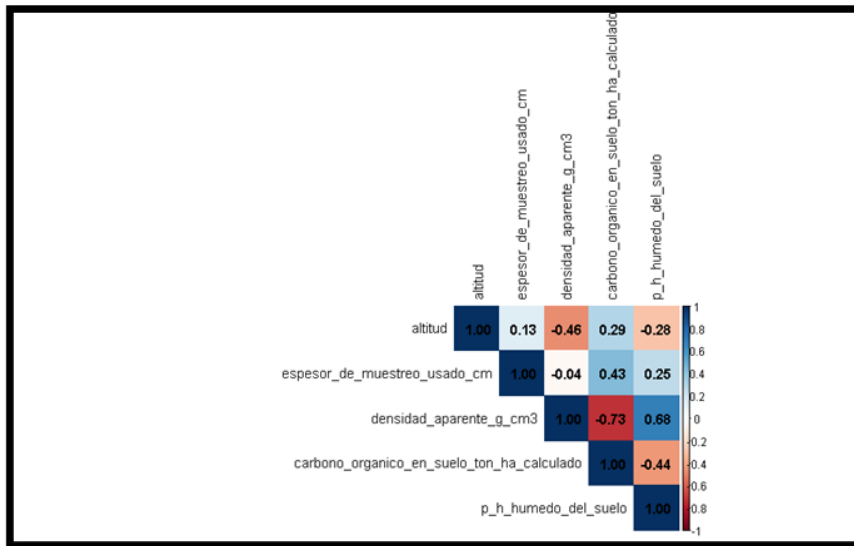


Figura 5 Correlaciones entre variables más importantes en el COS

Fuente: RStudio de variables analizadas

En conjunto, la matriz de correlaciones sugiere que la variabilidad del carbono orgánico del suelo está controlada principalmente por propiedades físicas del suelo, especialmente: densidad aparente, profundidad de muestreo y condiciones edáficas relacionadas con el pH. Por otro lado, la altitud actúa como factor ambiental que modela estas condiciones.

3.2. Modelos multivariados y métricas

Se plantea algunas alternativas de modelos multivariados, que permitan establecer la relación del COS y las variables analizadas:

Para este efecto, deben cumplirse los supuestos en cada uno de los modelos, estos fueron verificados previamente, se tratan de: Normalidad de los datos (residuos), variabilidad constante (homocedasticidad), independencia (prueba Durbin-Watson), ausencia de multicolinealidad

– Modelo Lineal Múltiple:

$$\text{COS} \sim \text{Altitud} + \text{espesor del muestreo} + \text{densidad aparente} + \text{ph del suelo}$$

Tabla 4. Resultados de coeficientes de modelo lineal múltiple

Variable	Coficiente (Estimate)	Error estándar	t value	p-value	Significancia
Intercepto	453.05	153.06	2.96	0.0066	**
Altitud	-0.043	0.0348	-1.24	0.225	ns
Espesor de muestreo (cm)	3.14	0.786	3.99	0.00051	***
Densidad aparente (g/cm ³)	-172.67	45.53	-3.79	0.00084	***
pH del suelo	-15.97	15.07	-1.06	0.299	ns

Tabla 5 Métricas del modelo LM

Estadístico	Valor
Error estándar residual	44.64
R ²	0.667
R ² ajustado	0.614
F estadístico	12.56
p-value del modelo	9.769 × 10 ⁻⁶

El modelo explica aproximadamente 66.7 % de la variabilidad del carbono orgánico del suelo (R² = 0.667), lo que indica una capacidad explicativa moderadamente alta considerando la complejidad de los sistemas edáficos. El valor obtenido en la prueba F es significativo ($p < 0.001$), lo que indica que el modelo es aceptado estadísticamente.

- Modelo log LM

$\log(COS) \sim \text{Altitud} + \text{espesor del muestreo} + \text{densidad aparente} + \text{ph del suelo}$

Tabla 6 Métricas del modelo log LM

Estadístico	Valor
Error estándar residual	0.1613
Grados de libertad	2
Observaciones eliminadas	130
R ² (coeficiente de determinación)	0.9971
R ² ajustado	0.9529
Estadístico F	22.6
Grados de libertad (modelo)	30 y 2
p-value del modelo	0.04322

El modelo presenta un coeficiente de determinación extremadamente alto (R² = 0.9971), lo que indica que aproximadamente el 99.7 % de la variabilidad de la variable dependiente es explicada por el modelo. El R² ajustado (0.9529) también es elevado, lo que sugiere un buen ajuste del modelo considerando el número de variables incluidas. El estadístico F es significativo ($p = 0.04322$), lo que indica que el modelo en conjunto tiene capacidad explicativa estadísticamente significativa.

ANOVA DE LOS MODELOS

- Modelo LM

Tabla 7 ANOVA modelo LM

Variable	Df	Sum Sq	Mean Sq	F value	Pr(>F)	Significancia
Altitud	1	2517	2517	1.263	0.2718	ns
Espesor de muestreo (cm)	1	29862	29862	14.983	0.00069	***
Densidad aparente (g/cm ³)	1	65503	65503	32.865	5.68e-06	***
pH del suelo	1	2236	2236	1.122	0.2996	ns
Residuals	25	49828	1993			

En conjunto, el análisis muestra que las propiedades físicas del suelo, particularmente la densidad aparente y el espesor de muestreo, son los principales factores que explican la variabilidad del carbono orgánico del suelo en el conjunto de datos analizado. Las variables ambientales como la altitud presentan una influencia menor cuando se consideran simultáneamente las propiedades edáficas.

- Modelo multiplicativo (altitud por espesor)

- **$COS \sim Altitud * espesor\ del\ muestreo$**

Tabla 8 ANOVA modelo multiplicativo

Variable	Df	Sum Sq	Mean Sq	F value	Pr(>F)	Significancia
Altitud	19	1 107 008	58 264	0.595	0.846	ns
Espesor de muestreo (cm)	4	235 404	58 851	0.601	0.670	ns
Altitud × Espesor	6	252 518	42 086	0.430	0.844	ns
Residuales	11	1 077 216	97 929			

No significativo

- Modelo interacción

$COS \sim Altitud * espesor\ del\ muestreo + densidad\ aparente + ph\ del\ suelo$

Tabla 9 ANOVA modelo interacción

Variable	Coficiente (Estimate)	Error estándar	t value	p-value	Significancia
Intercepto	754.20	399.20	1.889	0.071	.
Altitud	-1.337	1.160	-1.152	0.260	ns
Espesor de muestreo (cm)	-11.53	17.96	-0.642	0.527	ns
Densidad aparente (g/cm ³)	-165.8	46.6	-3.557	0.0016	**
pH del suelo	-15.67	15.18	-1.033	0.312	ns
Altitud × Espesor	4.284	5.240	0.818	0.421	ns

No significativo

- Modelo GAM

$COS \sim s(Altitud) * espesor\ del\ muestreo + s(densidad\ aparente) + s(ph\ del\ suelo)$

Tabla 10 Métricas del modelo GAM

Indicador	Valor
R ² ajustado	0.655
Deviance explicada	74.6 %
REML	140.52

Error estándar de escala	1786.2
Tamaño de muestra (n)	30

El modelo explica aproximadamente 74.6 % de la desviación total del carbono orgánico del suelo, lo que indica una buena capacidad explicativa del modelo. El R^2 ajustado de 0.655 sugiere que el modelo describe adecuadamente la variabilidad del carbono en el conjunto de datos analizado.

3.3 Métricas

Al analizar los modelos planteados anteriormente, se descartan: modelo logarítmico debido a la reducción extrema del tamaño muestral al eliminar variables faltantes, modelo LM, por baja significancia de algunas variables, por solo depender del espesor del muestreo, por lo tanto, se decide comparar estadísticamente los modelos LINEAL Y GAM como posibles modelos a aplicar en el estudio de COS.

- Comparación de los predictivos de los modelos lineal y GAM

Tabla 11 Resumen de métricas modelo LM y GAM

Modelo	R^2	RMSE	MAE	RPIQ
LM	0.597	44.87	33.64	2.39
GAM	0.749	35.63	26.32	3.01

Tabla 12 Comparación AIC y BIC de los modelos elegidos

Modelo	Parámetros (df)	AIC	BIC
LM	6	325.36	333.77
GAM	11.16	321.84	337.48

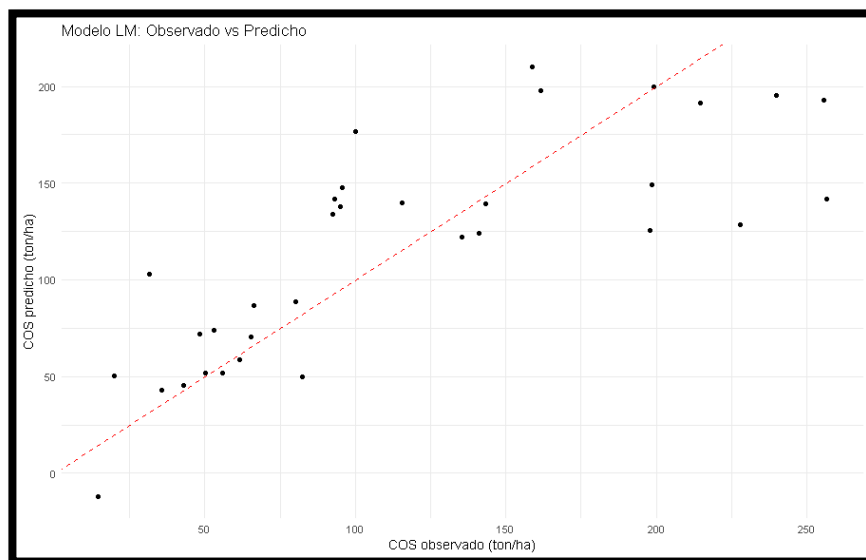


Figura 6 Valores calculados vs reales mediante modelo lineal

Fuente: RStudio (análisis de predicción mediante modelo estadístico lineal)

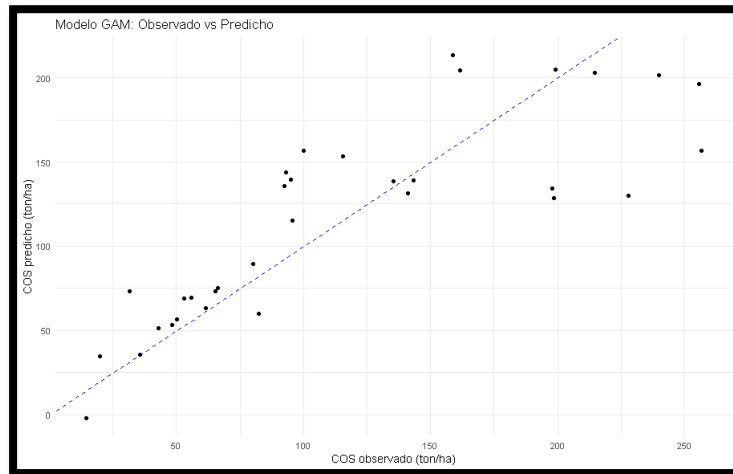


Figura 7 Valores calculados vs reales mediante modelo lineal

Fuente: RStudio (análisis de predicción mediante modelo estadístico lineal)

Los resultados muestran que el modelo aLm generalizado presenta un mejor desempeño predictivo que el modelo lineal múltiple. El GAM logró explicar una mayor proporción de la variabilidad del carbono orgánico del suelo y presentó menores errores de predicción. Esto sugiere que las relaciones entre el carbono del suelo y las variables ambientales consideradas pueden presentar comportamientos no lineales que son capturados de manera más adecuada por el modelo GAM.

3.4 Estudio multivariante ACP:

La figura 8 corresponde a un biplot del Análisis de Componentes Principales (ACP) aplicado a las variables edáficas y ambientales del estudio.[23]

Los dos primeros componentes principales explican la mayor parte de la variabilidad del sistema: Dim1 = 46.8 %, Dim2 = 25.6 %. En conjunto explican aproximadamente 72.4 % de la variabilidad total de los datos, lo que indica que el modelo captura adecuadamente la estructura multivariante del sistema COS

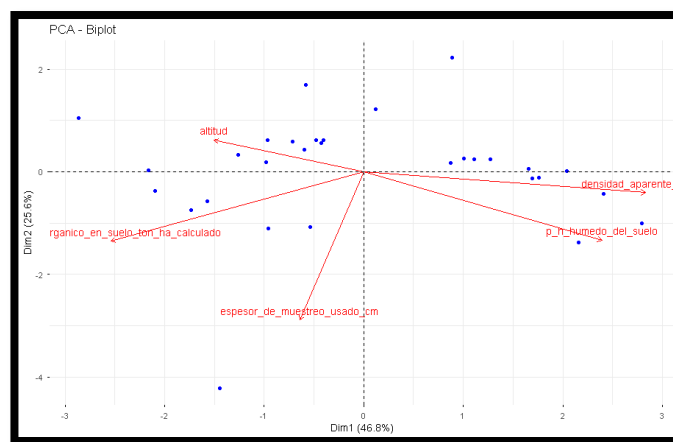


Figura 8 PCA Biplot desarrollada sobre las variables relacionadas al COS

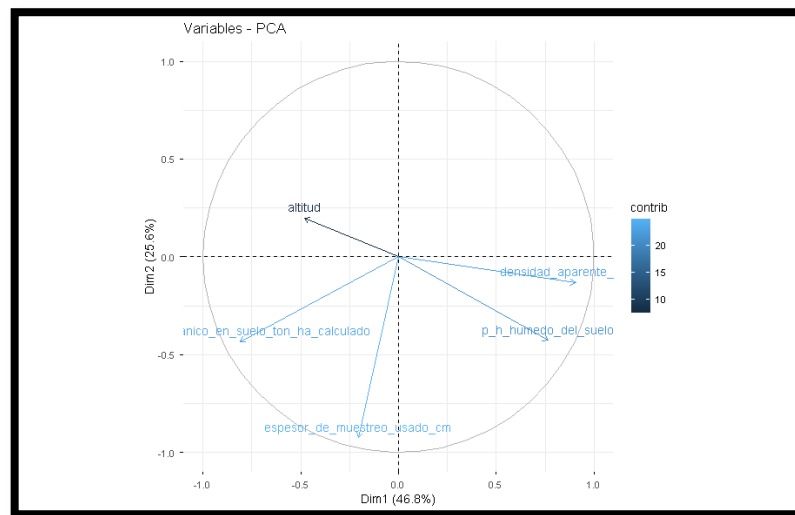


Figura 9 Variables PCA sobre las variables relacionadas al COS

En la figura 9 se observa la representación de las variables dentro del círculo de correlaciones, donde la dirección de las flechas indica la relación entre variables y su longitud representa el grado de contribución al modelo. El estudio de PCA no implica causalidad.

El análisis de componentes principales explica el 72.4 % de la variabilidad del sistema suelo carbono, evidenciando que la densidad aparente y el espesor de muestreo son los principales factores que controlan la distribución del carbono orgánico del suelo.

4. Discusión

Los resultados obtenidos en el presente estudio identifican la alta variabilidad del COS en los ecosistemas estudiados, luego, los valores están influenciados básicamente por propiedades físicas del suelo, en particular la densidad aparente y la profundidad de muestreo realizado. Estos resultados se muestran similares a investigaciones previas que han identificado que las propiedades estructurales del suelo desempeñan un papel determinante en la acumulación y estabilización del carbono orgánico [2], [4]

En los resultados del análisis estadístico descriptivo se demuestra la alta variabilidad de COS dentro de los datos analizados, con valores que van entre 0.09 y 1289 Ton/ha. La dispersión de los resultados demuestra alta heterogeneidad de la presencia de COS en suelos altoandinos ecuatorianos, influenciados por distintos factores y que generan muy diferentes distribuciones de COS. Mediante los análisis realizados, también se puede establecer que existe una tendencia en forma general: se evidencia el aumento o incremento de COS a medida que se incrementa la elevación del muestreo realizado, esto puede explicarse debido a la disminución de temperatura del ecosistema y la poca descomposición de la materia orgánica en esos entornos.

El análisis de correlaciones y el análisis de componentes principales (ACP) permitieron identificar que las variables edáficas presentan mayor influencia en la variabilidad del COS que los factores ambientales considerados. El ACP explicó aproximadamente el 72.4 % de la variabilidad total del sistema suelo-carbono, dentro de ellos, la densidad aparente y el espesor de muestreo son variables con mayor contribución en los componentes principales. Estos resultados están alineados

con lo mencionado por la literatura revisada previamente, donde señalan que suelos con mayor densidad aparente poseen mayor contenido de COS, esto atribuido a suelo más porosos y por lo tanto mayor capacidad de acumulación de material orgánico.

Por su parte, en el análisis de los modelos estadísticos propuestos, el modelo de regresión lineal múltiple explicó aproximadamente el 60% de la variabilidad de COS ($R^2 = 0.597$), mientras que el modelo aditivo generalizado GAM alcanzó un $R^2=0.749$, lo que determina una mayor capacidad explicativa. En el análisis de métricas y errores, el modelo GAM mostró valores más bajos de RMSE y MAE, lo cual confirma una mejor capacidad predictiva de este modelo en comparación al lineal tradicional. Por consiguiente, se demuestra que la relación existente entre el COS y las variables del suelo analizado pueden presentar un comportamiento que puede ser explicado de manera más eficiente en forma no lineal. Luego, en la comparación AIC, el modelo GAM presento valores más bajos, lo que significa una mejora entre el modelo y el ajuste realizado. Contrastando con investigaciones previas, también existe concordancia en donde se aplican modelos no lineales para explicar la dinámica del COS y las variables de los ecosistemas.

En forma general, la distribución de COS en ecosistemas altoandinos está explicada fundamentalmente por variables edáficas y la profundidad del muestreo realizado, y esta variabilidad es mejor explicada por el modelo GAM de forma no lineal. Tanto la dinámica del COS como la comprensión de los procesos subyacentes y el uso de métodos multivariantes para integrar una amplia gama de variables se ven favorecidos por el uso de la multivariada. El conocimiento de los factores que rigen la acumulación de carbono en el suelo alimenta la mejora de los modelos de estimación de carbono y fortalece las políticas de mitigación del carbono en relación con el uso de la tierra

5. Conclusión

1. La variabilidad del carbono orgánico del suelo en ecosistemas altoandinos del Ecuador está asociada principalmente a propiedades físicas del suelo, particularmente a la densidad aparente y al espesor de muestreo. Los resultados obtenidos mediante análisis descriptivo, correlaciones y análisis de varianza evidencian que estas variables presentan una influencia estadísticamente significativa sobre la distribución del carbono orgánico del suelo, lo que confirma su importancia en los procesos de acumulación y estabilización del carbono en sistemas edáficos de ecosistemas altoandinos.
2. El análisis multivariante realizado es adecuado para el estudio del COS en suelos altoandinos, mediante estas técnicas se puede explicar el 72,4% de la variabilidad de los datos analizados, explicado mediante el ACP. Es decir, el hallazgo valida la utilidad de las técnicas multivariantes para utilizar en bases de datos complejas y permite entender las interacciones entre las variables analizadas de los ecosistemas altoandinos.
3. Luego del planteamiento de varios modelos explicativos, el modelo GAM no lineal mostró un mejor rendimiento predictivo para las variaciones de COS, esto debido al valor de $R^2 = 0.749$ que demuestra una mejor explicación y menores errores de predicción en comparación con el modelo lineal múltiple. Esto demuestra que la interacción de las variables analizadas no sigue patrones lineales.
4. Para el desarrollo del análisis estadístico del COS se utilizó análisis descriptivos y exploratorios, luego técnicas multivariantes, modelos explicativos y predictivos con la

finalidad de establecer los factores determinantes en la explicación del COS, luego contribuir con la línea de investigación sobre captura de carbono, modelación ambiental y manejo adecuado de los ecosistemas altoandinos en Ecuador.

Contribución de autores: Conceptualización, F.R.M.S.; metodología, F.R.M.S. y M.A.G.S; software, F.R.M.S.; validación, F.R.M.S. y M.A.G.S.; análisis formal, F.R.M.S.; investigación, F.R.M.S.; curaduría de datos, F.R.M.S.; redacción del borrador original, F.R.M.S.; redacción–revisión y edición, M.A.G.S.; visualización, F.R.M.S.; supervisión, M.A.G.S.; administración del proyecto, F.R.M.S.; adquisición de fondos, F.R.M.S. y M.A.G.S Todos los autores han leído y aprobado la versión final del manuscrito.

Financiamiento: Los autores financiaron a integridad el estudio.

Conflictos de interés: Los autores declaran no tener ningún conflicto de intereses.

Referencias

- [1] M. P. Salazar *et al.*, "Soil organic carbon," *Revista de la Facultad de Agronomía*, vol. 119, no. 2, p. 053, Dec. 2020, doi: 10.24215/16699513e053.
- [2] K. Georgiou *et al.*, "Global stocks and capacity of mineral-associated soil organic carbon," *Nat. Commun.*, vol. 13, no. 1, Dec. 2022, doi: 10.1038/s41467-022-31540-9.
- [3] C. Li, C. Xiao, M. Li, L. Xu, and N. He, "A global synthesis of patterns in soil organic matter and temperature sensitivity along the altitudinal gradient," *Front. Environ. Sci.*, vol. 10, Sep. 2022, doi: 10.3389/fenvs.2022.959292.
- [4] G. Alavi-Murillo, J. Diels, J. Gilles, and P. Willems, "Soil organic carbon in Andean high-mountain ecosystems: importance, challenges, and opportunities for carbon sequestration," *Regional Environmental Change* 22:4, vol. 22, no. 4, pp. 1–16, Oct. 2022, doi: 10.1007/S10113-022-01980-6.
- [5] W. Buytaert, R. Celleri, P. Willems, B. De Bièvre, and G. Wyseure, "Spatial and temporal rainfall variability in mountainous areas: A case study from the south Ecuadorian Andes," *J. Hydrol. (Amst)*, vol. 329, no. 3–4, pp. 413–421, Oct. 2006, doi: 10.1016/J.JHYDROL.2006.02.031.
- [6] J. Ayala *et al.*, "ESTIMACIÓN DEL CARBONO ORGÁNICO DEL SUELO USANDO DATOS DE TELEDETECCIÓN Y SIG EN DOS ECOSISTEMAS ALTO ANDINOS DEL ECUADOR," Universidad Autónoma de Nuevo León, Ed., México, Oct. 2022.
- [7] A. A. Beltrán-Dávalos *et al.*, "Evaluation of Soil Organic Carbon Storage of Attilo in the Ecuadorian Andean Wetlands," *Soil Syst.*, vol. 6, no. 4, Dec. 2022, doi: 10.3390/soilsystems6040092.
- [8] K. A. Farley, E. F. Kelly, and R. G. M. Hofstede, "Soil organic carbon and water retention after conversion of grasslands to pine plantations in the Ecuadorian Andes," *Ecosystems*, vol. 7, no. 7, pp. 729–739, Nov. 2004, doi: 10.1007/S10021-004-0047-5/METRICS.
- [9] L. G. Thompson *et al.*, "The impacts of warming on rapidly retreating high-altitude, low-latitude glaciers and ice core-derived climate records," *Glob. Planet. Change*, vol. 203, Aug. 2021, doi: 10.1016/J.GLOPLACHA.2021.103538;
- [10] A. A. Beltrán-Dávalos *et al.*, "Driving variables to explain soil organic carbon dynamics: páramo highlands of the Ecuadorian Real mountain range," *J. Soils Sediments*, vol. 25, no. 5, pp. 1578–1597, May 2025, doi: 10.1007/s11368-025-04017-7.

- [11] V. Loayza *et al.*, "Digital mapping of organic carbon in Ecuador soils," *Ecosistemas*, vol. 29, no. 2, 2020, doi: 10.7818/ECOS.1852.
- [12] D. Armas *et al.*, "Harmonized Soil Database of Ecuador (HESD): data from 2009 to 2015," *Earth Syst. Sci. Data*, vol. 15, no. 1, pp. 431–445, Jan. 2023, doi: 10.5194/essd-15-431-2023.
- [13] C. Chunchu Morocho and G. Chunchu, "Páramos del Ecuador, importancia y afectaciones: Una revisión," *Bosques Latitud Cero*, vol. 9, no. 2, Nov. 2019.
- [14] A. Alejandro, G. Correa, S. Ecuatoriana, E. Defensa, and F. Equilibrio, "TRATAMIENTO COMPUTACIONAL DE DATOS."
- [15] W. Álvarez, "Universidad Regional Amazónica Ikiam INTRODUCCIÓN AL ANÁLISIS DE COMPONENTES PRINCIPALES Y AL ANÁLISIS BIPLLOT."
- [16] E. Galindo, "Curso Básico de Regresión Lineal Simple," 2025.
- [17] "Using R for Introductory Statistics Second Edition."
- [18] R. Gentleman and G. Parmigiani, "Series Editors." [Online]. Available: <http://www.springer.com/series/6991>
- [19] J. Cohen, "Statistical Power Analysis for the Behavioral Sciences Second Edition," 1988.
- [20] L. F. Restrepo, "Tipos de suma de cuadrados en el análisis de la varianza Types of sum of squares for analysis of variance."
- [21] D. Montgomery, E. Peck, and Geoffrey Vining, "INTRODUCTION TO LINEAR REGRESSION ANALYSIS," 2012.
- [22] Gareth. James, Daniela. Witten, Trevor. Hastie, and Robert. Tibshirani, *An introduction to statistical learning: with applications in R*. Springer : Springer Science+Business Media, 2017.
- [23] S. Bojinski, M. Verstraete, T. C. Peterson, C. Richter, A. Simmons, and M. Zemp, "The concept of essential climate variables in support of climate research, applications, and policy," *Bull. Am. Meteorol. Soc.*, vol. 95, no. 9, pp. 1431–1443, Sep. 2014, doi: 10.1175/BAMS-D-13-00047.1.



© 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license <http://creativecommons.org/licenses/by/4.0/>