**Logotipo

Descripción generada automáticamente**

**Modelo predictivo de la Actividad Vegetal en el Bosque Tropical Seco del Cañón del Río Cauca**

Eileen Melissa Arévalo Garnica

Pablo Uribe Uribe

Monografía presentada para optar al título de Especialista en Analítica y Ciencia de Datos

Asesor  
Nombres completos, Título académico más alto

Universidad de Antioquia  
Facultad de Ingeniería

Especialización en Analítica y Ciencia de Datos

Medellín, Antioquia, Colombia

2023

| **Cita** | (Arévalo Garnica & Uribe Uribe, 2023) |
| --- | --- |
| **Referencia**  **Estilo APA 7 (2020)** | Arévalo Garnica, E.M, & Uribe Uribe, P. (2023). *Modelo predictivo de la Actividad Vegetal en el Bosque Tropical Seco del Cañón del Río Cauca*. Universidad de Antioquia, Medellín, Colombia. |

**** 

Especialización en Analítica y Ciencia de Datos, CohorteVI.

Centro de Investigación Ambientales y de Ingeniería (CIA).

|  | Diagrama  Descripción generada automáticamente con confianza media |
| --- | --- |

Centro de Documentación Ingeniería (CENDOI)

**Repositorio Institucional:** http://bibliotecadigital.udea.edu.co

Universidad de Antioquia - www.udea.edu.co

Rector: John Jairo Arboleda Céspedes.

Decano: Julio Cesar Saldarriaga Molina

Jefe departamento: Diego José Luis Botia Valderrama

El contenido de esta obra corresponde al derecho de expresión de los autores y no compromete el pensamiento institucional de la Universidad de Antioquia ni desata su responsabilidad frente a terceros. Los autores asumen la responsabilidad por los derechos de autor y conexos.

**Dedicatoria**

Texto de dedicatoria centrado.

**Agradecimientos**

Texto de agradecimientos centrado.

**Tabla de contenido**

[Resumen 8](#_heading=h.3rdcrjn)

[Abstract 9](#_heading=h.26in1rg)

[1. Descripción del problema 10](#_heading=h.35nkun2)

[1.1. Problema de negocio 10](#_heading=h.1ksv4uv)

[1.2. Aproximación desde la analítica de datos 12](#_heading=h.44sinio)

[1.3. Origen de los datos 13](#_heading=h.z337ya)

[1.4. Métricas de desempeño 14](#_heading=h.3j2qqm3)

[2. Objetivos 16](#_heading=h.4i7ojhp)

[2.1. Objetivo general 16](#_heading=h.2xcytpi)

[2.2. Objetivos específicos 16](#_heading=h.3whwml4)

[3. Datos 17](#_heading=h.2bn6wsx)

[3.1. Datos originales 17](#_heading=h.qsh70q)

[3.2. Datsets 18](#_heading=h.3as4poj)

[3.3. Analítica descriptiva 19](#_heading=h.1pxezwc) **22**

[Referencias 23](#_heading=h.1jlao46)

**Lista de tablas**

[**Tabla 1**](#_heading=h.2u6wntf) *Datos original investigación doctoral* [1](#_heading=h.2u6wntf)8

**Lista de figuras**

[**Figura 1**](#_heading=h.2r0uhxc) Distribución de las variables [2](#_heading=h.2r0uhxc)0

[**Figura 2**](#_heading=h.1664s55) Cajas y bigotes de las variables [2](#_heading=h.1664s55)1

[**Figura**](#_heading=h.1664s55) **3** Correlación variable eco-hidroecologicas VS NDVI 22

[**Figura 4**](#_heading=h.1664s55) Media y desviación estándar [2](#_heading=h.1664s55)3

**Siglas, acrónimos y abreviaturas**

**APA** American Psychological Association

**Cms.** Centímetros

**ERIC** Education Resources Information Center

**Esp.** Especialista

**MP** Magistrado Ponente

**MSc** Magister Scientiae

**Párr.** Párrafo

**PhD** Philosophiae Doctor

**PBQ-SF** Personality Belief Questionnaire Short Form

**PostDoc** PostDoctor

**UdeA** Universidad de Antioquia

# Resumen

Este proyecto se enfoca en entender las relaciones que puedan existir entre la actividad vegetal y las variables eco-hidrológicas para la creación de un modelo predictivo de la actividad vegetal en el bosque seco tropical del Cañón del Río Cauca en Colombia. La creación de una herramienta que ayude a predecir la actividad vegetal podría ser fundamental para hacer seguimiento de la salud del bosque en cuestión, para la gestión de los recursos hídricos y la conservación de estos ecosistemas. Para abordar este desafío, el proyecto utiliza técnicas de machine learning y métricas de evaluación con el fin de desarrollar un modelo preciso y eficaz, utilizando como entrada variables eco-hidrológicas y ambientales, obtenidas de estaciones del IDEAM y de fotos satelitales de Google Engine, logrando comprender y predecir la actividad vegetal (NDVI) en esta región. Aunque el alcance del proyecto se limita a esta cuenca específica y sus datos, su impacto potencial es relevante para las autoridades ambientales, los investigadores y las comunidades locales que buscan tomar decisiones informadas y promover la conservación de los bosques secos tropicales en la región

*Palabras clave*: Machine Learning, Bosque Tropical Seco, Eco-hidrológico, Actividad Vegetal.

https://github.com/emarevalog/Machine\_learning\_project

# Abstract

This project focuses on understanding the relationships that may exist between vegetation activity and eco-hydrological variables in order to create a predictive model of plant activity in the tropical dry forest of the Cauca River Canyon in Colombia. The creation of a tool for predicting vegetation activity could be fundamental for monitoring the health of the forest in question, for managing water resources and for conserving these ecosystems. To address this challenge, the project uses machine learning techniques and evaluation metrics to develop an accurate and efficient model, using as inputs eco-hydrological and environmental variables obtained from IDEAM stations and Google Engine satellite images, to understand and predict vegetation activity (NDVI) in this region. Although the scope of the project is limited to this specific watershed and its data, its potential impact is relevant for environmental authorities, researchers and local communities seeking to make informed decisions and promote the conservation of tropical dry forests in the region.

*Keywords***:** Machine Learning, Tropical Dry Forest, Eco-hydrological, Vegetation Activity.

# Descripción del problema

## Problema de negocio

La necesidad de predecir la actividad vegetal en los bosques secos tropicales de la cuenca del río Cauca, entre los municipios de La Pintada y Valdivia en Colombia. Este problema implica conocer y entender la relación entre la vegetación y los recursos hídricos, y cómo los cambios en estos sistemas pueden afectar la conservación de los bosques secos tropicales y la gestión de los recursos hídricos.

El problema se manifiesta debido a la intensificación de las actividades antropogénicas en la región, como la agricultura, la ganadería, la minería, el desarrollo urbano y el turismo. Estas actividades han llevado a la fragmentación de los bosques secos tropicales, poniendo en riesgo su futuro y los servicios ambientales que proporcionan. El aumento en la frecuencia e intensidad de los incendios forestales, atribuido en gran medida a las disminuciones en las precipitaciones y a los períodos prolongados de sequía relacionados con el cambio climático, agrava aún más el problema. A pesar de estos desafíos, la falta de herramientas efectivas para medir el impacto o riesgo futuro de estas actividades antropogénicas en el ecosistema y en la cuenca del Río Cauca indican la necesidad de una herramienta que sirva a las autoridades y las comunidades para medir el impacto de esta problemática y tener la posibilidad de mejorar la calidad del ecosistema en cuestión.

Los bosques secos tropicales, se ubican en zonas bajas de regiones tropicales caracterizadas por climas cálidos durante todo el año y largas estaciones secas, representan una parte significativa de los ecosistemas naturales, especialmente en América Latina. Estos ecosistemas son cruciales debido a la amplia gama de servicios ambientales que proporcionan y la rica biodiversidad que albergan. Además, cumplen un papel esencial en la protección contra la erosión eólica e hídrica. A pesar de su importancia, los bosques secos tropicales enfrentan varios peligros significativos. La deforestación debido a la extracción de madera y la expansión agrícola representa una amenaza constante para su preservación. Sin embargo, el peligro más relevante al que se enfrentan estos bosques es el aumento de los incendios forestales. En las últimas décadas, se ha observado un aumento en la frecuencia e intensidad de los incendios, lo que se ha atribuido en gran medida a la disminución de las precipitaciones y a los períodos prolongados de sequía relacionados con el cambio climático. (Schröder et al., 2021)

Existe una evidencia contundente de la estrecha relación entre la actividad vegetal y las variables climáticas. Los análisis de correlación han revelado que la temperatura ejerce un impacto negativo en la vegetación en regiones áridas y semiáridas, mientras que tiene un efecto positivo en el crecimiento de la vegetación en zonas húmedas a lo largo de toda la temporada de crecimiento. (Muradyan et al., 2019). Además, se ha observado una relación bidireccional entre la vegetación y las fuentes de agua. La disponibilidad de agua desempeña un papel fundamental en la salud de los ecosistemas y en los procesos esenciales para mantener las condiciones ecológicas y los servicios ecosistémicos. La vegetación utiliza diversas fuentes de agua, incluyendo las superficiales y las aguas subterráneas. El nivel freático puede fluctuar debido a la absorción de agua por parte de las plantas, disminuyendo durante el día e incrementando durante la noche, cuando cesa la actividad de evapotranspiración de las plantas. Aunque el agua subterránea representa aproximadamente el 37% del suministro de agua para la vegetación, su importancia aún es poco comprendida y estudiada. (Wang et al., 2023)

Para estudiar las interacciones entre la actividad vegetal y las variables climáticas se pueden realizar modelado eco-hidrologico. Los modelos basados en productos de teledetección se basan en detectar los cambios en las características aéreas de la vegetación en su mayoría utilizando el Índice de Vegetación de Diferencia Normalizada (NDVI) (Wang et al., 2023). Los índices de vegetación desempeñan un papel importante en la evaluación del crecimiento y salud de la vegetación, lo que es de gran relevancia en diversas aplicaciones, especialmente en el seguimiento de cambios en la cobertura terrestre a partir de imágenes satelitales. Los índices de vegetación son expresiones matemáticas que integran mediciones de reflectancia en las bandas espectrales. Uno de los índices más ampliamente reconocidos y utilizados es el NDVI, que proporciona valores que oscilan entre -1 y 1. En este contexto, los valores positivos indican un aumento en la densidad y vigor de la vegetación, mientras que los valores negativos señalan áreas carentes de cobertura vegetal (Ferchichi et al., 2022)

En Colombia, el conocimiento sobre el bosque seco tropical es limitado y carece de datos sólidos que permitan una gestión integral de este importante ecosistema. La distribución de este tipo de bosque en Colombia está estrechamente relacionada con procesos históricos de deforestación y colonización. En la actualidad, varios factores ejercen una gran presión sobre estos ecosistemas estratégicos, incluyendo la agricultura, la ganadería, la minería, el desarrollo urbano y el turismo. Esta presión ha llevado a la fragmentación de los bosques secos tropicales, lo que representa una amenaza para su conservación y para los servicios ecosistémicos que proporcionan. Es crucial abordar esta problemática a través de una gestión adecuada y la generación de información sólida que respalde la toma de decisiones en la conservación de estos ecosistemas (Pizano & Garcia, 2014).

El área ribereña del río Cauca en el departamento de Antioquia, Colombia, ha sido objeto de una explotación intensiva de sus recursos naturales a lo largo de la historia, desde la época de la conquista y colonización española hasta la actualidad. Esta explotación ha involucrado actividades como la minería, la tala de árboles y la pesca. Como resultado de estas actividades, la proporción de bosque seco tropical en la región, tanto en el Alto Cauca como en el Medio y Bajo Cauca, ha disminuido significativamente y actualmente representa menos del 20% de la cobertura total. Este proceso de fragmentación ha dado lugar a la presencia de parches de bosque seco tropical dispersos en un paisaje transformado por actividades humanas (Pizano & Garcia, 2014).

## Aproximación desde la analítica de datos

Los modelos predictivos desarrollados en este proyecto buscan anticipar y comprender la actividad vegetal en los bosques secos tropicales del Cañón del Río Cauca. Estos modelos se centran en la relación entre las variables eco-hidrológicas, como la precipitación, humedad del suelo, evapotranspiración, recarga, y la actividad vegetal, medida mediante el Índice de Vegetación de Diferencia Normalizada (NDVI).

Su función principal es ofrecer un sistema de monitoreo continuo de la actividad vegetal en la región, permitiendo la detección temprana de cambios en la salud de los bosques. Emiten alertas tempranas sobre posibles amenazas, como incendios forestales o sequías prolongadas, para facilitar respuestas proactivas. Además, respaldan la toma de decisiones informada para autoridades ambientales, investigadores y comunidades locales al proporcionar insights sobre cómo las variables eco-hidrológicas impactan la vegetación.

Estos modelos contribuyen a la planificación sostenible al ofrecer información valiosa sobre la relación entre las actividades humanas, el cambio climático y la salud de los bosques. Ayudan a desarrollar políticas y prácticas que promuevan la conservación a largo plazo de estos ecosistemas. Además, podrian desempeñar un papel importante en la gestión de recursos hídricos, al proporcionar entendimiento sobre cómo las variables ambientales afectan la vegetación y mantener el equilibrio hídrico en la cuenca del Río Cauca para un uso sostenible del agua.

En resumen, estos modelos no solo ofrecen una visión profunda de las interacciones eco-hidrológicas en los bosques secos tropicales, sino que también brindan herramientas prácticas para abordar desafíos específicos, proteger la biodiversidad y promover la gestión sostenible de estos valiosos ecosistemas..

## Origen de los datos

La fuente de datos abarca diversas fuentes que proporcionan una perspectiva integral para el desarrollo del modelo predictivo. Se emplean imágenes satelitales de alta resolución, accesibles a través de la plataforma Google Earth Engine, provenientes de satélites como Landsat o Sentinel. Estas imágenes, capturadas a lo largo del tiempo, ofrecen una visión detallada de la cuenca de estudio y son fundamentales para rastrear cambios en la actividad vegetal.

El Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM) contribuye significativamente al proyecto mediante datos de sus estaciones de monitoreo distribuidas en la región. Estos datos, que abarcan parámetros hidroclimáticos como precipitación, evapotranspiración y humedad del suelo, proporcionan una comprensión detallada de las condiciones ambientales locales.

La información topográfica y geoespacial se integra para contextualizar el entorno del estudio. Datos topográficos y mapas de cobertura terrestre enriquecen la comprensión de las características del paisaje, complementando así el análisis de las variables eco-hidrológicas.

Los datos recopilados durante la investigación doctoral, que abarcan mediciones de campo y análisis de suelos, aportan una perspectiva única y específica relacionada con la investigación en curso.

Además, la revisión de la bibliografía científica y estudios previos sobre la ecología de los bosques secos tropicales agrega profundidad al conjunto de datos. Esta revisión complementaria asegura que el proyecto se beneficie de los conocimientos acumulados en el campo, fortaleciendo así las bases teóricas y prácticas del modelo predictivo..

## Métricas de desempeño

En el transcurso de desarrollo de nuestro modelo de predicción, en la etapa de experimentación donde el objetivo será seleccionar el mejor tipo de modelo y adicional seleccionar los mejores hiperparametros para llevar las predicciones al nivel de acierto mas alto tendremos 2 métricas fundamentales para comparar cada unos de estas combinaciones de modelo y sus hiperparametros, estas métricas serán:

* **Error Cuadrático Medio (RMSE):** Mide la raíz cuadrada de la diferencia entre los valores observados y los valores predichos. Un valor más bajo de RMSE indica un mejor ajuste del modelo a los datos.
* **Coeficiente de Determinación (R²):** Evalúa la proporción de la varianza en la variable de respuesta que es explicada por el modelo. Un valor más alto de R² indica una mejor capacidad del modelo para explicar la variabilidad en los datos.

En nuestro alcance no está llevar el modelo a producción ni a ser usado por las entidades que gestionan la regulación para el cuidado de estos ecosistema, pero como métricas de negocio a usar en un futuro recomendamos, tener mensualmente un informe con:

* **Coeficiente de Determinación (R²) mes vencido**: este indicador nos podrá mostrar la precisión del modelo en ese momento del tiempo y tomar acciones como reentrenar, cambiar o cancelar el modelo.
* **R²/Costo mensual**: Este modelo deberá estar corriendo y consumiendo algunos recursos que pueden ser costeados, una relación entre el R² y este costo podría ser un valor importante para decidir si vale la pena seguir con el modelo corriendo.
* **Efectividad de las acciones tomadas**: Las entidades encargadas deberían de poder evidenciar la efectividad de sus acciones, y una vez ejecutada una intervención o aprobada e impuesta una ley, el modelo debería mostrar cómo sus predicciones se tornan hacia una actividad vegetal positiva.
* **Comparación de Resultados del Modelo con Variaciones del NDVI**: Este indicador nos permite comprobar la consistencia entre las predicciones del modelo y los datos observados de NDVI.

# Objetivos

## Objetivo general

Desarrollar un modelo predictivo preciso mediante técnicas de machine learning que relacione variables eco-hidrológicas y ambientales con la actividad vegetal, medida a través del Índice de Vegetación de Diferencia Normalizada (NDVI), en el Bosque Tropical Seco del Cañón del Río Cauca.

## Objetivos específicos

* Recopilar y preprocesar datos de imágenes satelitales, estaciones de monitoreo del IDEAM, información topográfica, datos de investigación doctoral y bibliografía científica. Realizar un proceso de limpieza, eliminación de valores atípicos y normalización para garantizar la calidad y coherencia de los datos.
* Identificar relaciones entre las variables eco-hidrológicas y el NDVI mediante un análisis exploratorio de datos para analizar posibles correlaciones y patrones que sirvan como base para el desarrollo del modelo predictivo.
* Seleccionar variables relevantes utilizando técnicas de selección para identificar y evaluar la importancia de cada variable y su contribución al modelo en la predicción de la actividad vegetal.
* Implementar modelos de machine learning y comparar diferentes algoritmos de machine learning, con cada uno se deberá entrenar y ajustar los modelos utilizando el conjunto de datos recopilado.
* Evaluar el desempeño de los modelos utilizando métricas de evaluación, como RMSE y R², para medir la precisión y eficacia de los modelos en la predicción del NDVI.
* Seleccionar y optimizar al máximo el modelo con mejores métricas de rendimiento, para su futura entrega y publicación.

# Datos

## Datos originales

Nuestro trabajo es posterior a una investigación doctoral donde se realizaron un gran trabajo de preparación de datos y construcción del dataset que nosotros usaremos para el modelo. Para llegar a nuestro dataset original, en la siguiente tabla se puede observar la descripción de los datos originales usados en la investigación.

**Tabla 1**  *Datos original investigación doctoral “Dinámica ecohidrológica entre bosque seco tropical y agua subterránea en el cañón del río Cauca en la región comprendida entre los municipios de Caramanta y Valdivia, Antioquia-Colombia”*

| **INFORMACIÓN** | **DESCRIPCIÓN** |
| --- | --- |
| Datos hidrometeorológicos de la red de monitoreo del IDEAM | Datos de lluvia total, temperatura máxima, temperatura mínima, temperatura media, evaporación, caudal medio, humedad relativa y brillo solar |
| Datos hidrometeorológicos de la red de monitoreo de PIRAGUA- CORANTIOQUIA | Datos de lluvia total, temperatura máxima, temperatura mínima, temperatura media, evaporación, caudal medio, humedad relativa y brillo solar |
| Unidades cartográficas de suelos – escala 1:100.000 | Shape de las unidades cartográficas de suelos del Estudio General de Suelos y Zonificación de Tierras del departamento de Antioquia (IGAC & Gobernación de Antioquia, 2007) |
| Información textural y de propiedades hidráulicas de suelos – escala 1:100.000 | Información de los perfiles de suelos del Estudio General de Suelos y Zonificación de Tierras de Antioquia (IGAC & Gobernación de Antioquia, 2007) |
| Información de suelos de los POMCAS en jurisdicción de CORANTIOQUIA – escala 1:25.0000 | Descripción de suelos a escala 1:25.000 de los POMCAS del río Amagá – quebrada Sinifaná y Río Aurrá |
| Información textural y de propiedades hidráulicas de suelos – escala 1:25.000 | Información de los perfiles de suelos de los POMCAS del río Amagá – quebrada Sinifaná y Río Aurrá |
| Mapa de coberturas de la tierra a escala 1:100.000 | Mapa de coberturas de la tierra para la zona de estudio con su respectiva leyenda – escala 1:100.000. Formato shape. |
| Mapa de coberturas de la tierra a escala 1:25.000 de los POMCAS en jurisdicción de CORANTIOQUIA | Mapa de coberturas de la tierra de los POMCAS del río Amagá – quebrada Sinifaná y Río Aurrá, a escala 1:25.000. Formato shape. |
| Modelo Digital de Elevaciones (MDE) para Antioquia | Modelo Digital de Elevaciones (MDE) para Antioquia, tamaño de pixel 12,5 m, obtenido de la base de datos ALOS – PALSAR y corregido por la Gobernación de Antioquia. Formato raster. |
| Modelo Digital de Elevaciones (MDE) de Colombia | Modelo Digital de Elevaciones (MDE) para Antioquia, tamaño de pixel 30 m, obtenido de información satelital ASTER y corregido por IGAC. Formato raster. |
| Red de drenaje superficial | Red de drenaje superficial (drenaje doble y drenaje sencillo) en escala 1:100.000 y 1:25.000. Formato shape. |

Luego de un tratamiento de datos y consumo de estas imágenes satelitales al final de la investigación doctoral se llegó a un dataset preparado, el cual es el mismo que nosotros usaremos como entrada de nuestro modelo

## Datsets

El dataset de entrada al modelo es un conjunto de datos que contiene información climática mensual promedio donde la variable NDVI fue obtenida por procesamiento de imágenes del satélite MODIS y las demás provienen del modelo Soil-Water Balance (SWB), las variables de nuestro dataset son las siguientes :

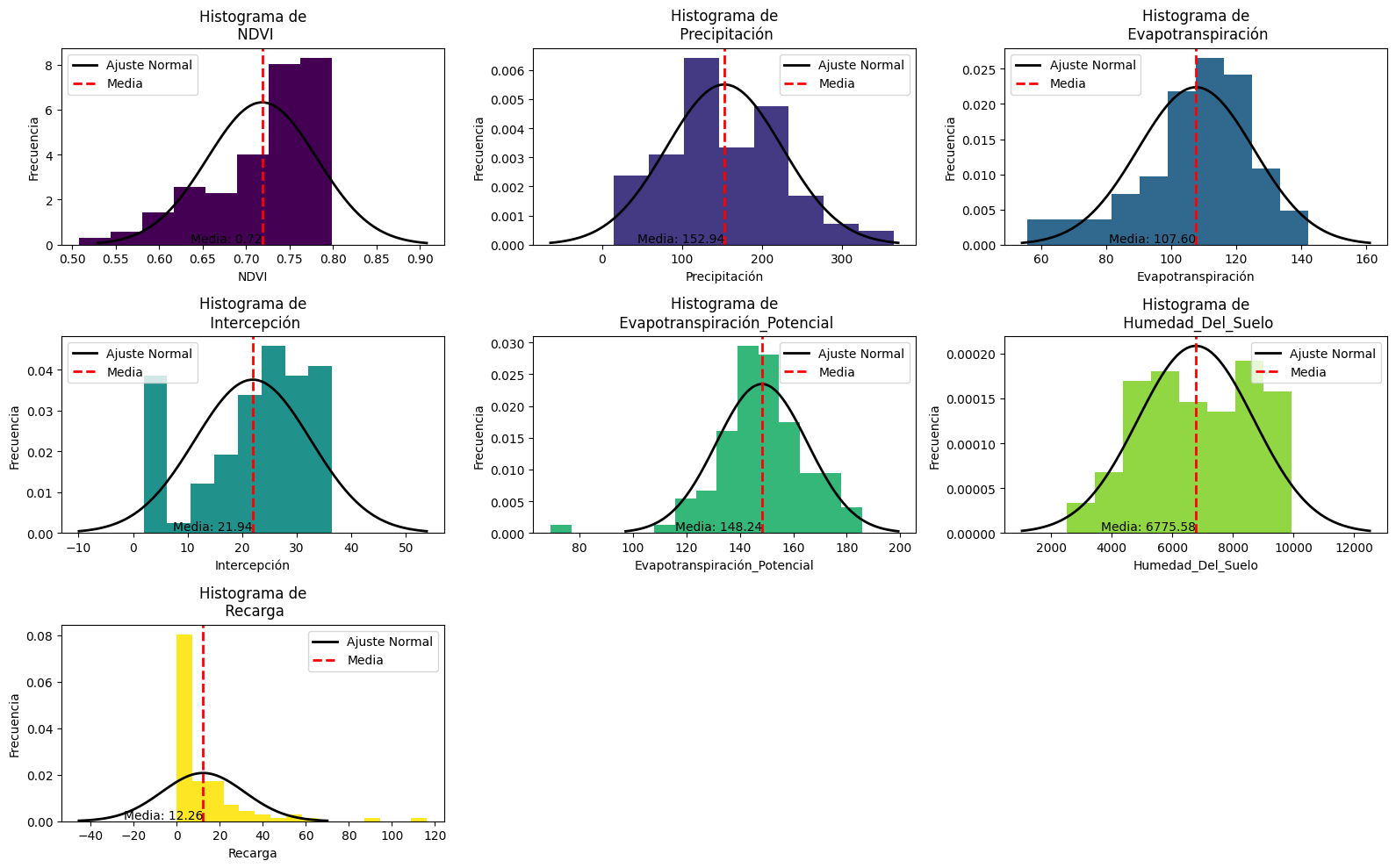
* Mes (1 al 12)
* NDVI
* Precipitación(ml)
* Evapotranspiración (ml)
* Intercepción (ml)
* Evapotranspiración (ml)
* Humedad Suelo (ml)
* Recarga (ml)

## Analítica descriptiva

Con el objetivo de conocer a fondo el comportamiento y composición de nuestras variables hicimos un ejercicio de graficar diferentes escenarios para tener evidencia gráfica y realizar el análisis.

La visualización de la distribución de las variables, como el NDVI, Precipitación, Evapotranspiración, Intercepción, Humedad del Suelo y Recarga, desempeña un papel crucial en la fase de análisis descriptivo previo a la construcción del modelo de machine learning. Al examinar la forma en que estas variables están distribuidas, se busca comprender la naturaleza de los datos y detectar posibles patrones o anomalías.

**Figura 1**Distribución de las variables

**

# Con el fin de conocer un poco mejor la distribución de los valores de nuestros datos y adicional ver gráficamente cuantos y de qué magnitud son nuestros valores atípicos y decidir si debemos conservarlos, transformarlos o eliminarlos decidimos realizar una gráficas de cajas y bigotes.

**Figura 2**Cajas y bigotes de las variables

# 

La evaluación de la dispersión y correlación de las variables versus el NDVI, es especialmente importante y nos proporciona información sobre la variabilidad de estos parámetros y su relacion entre ellos. Esto es esencial para comprender la consistencia y la estabilidad de los datos, especialmente en el contexto de la gestión de recursos hídricos

**Figura 3**Correlación variable eco-hidroecologicas VS NDVI

# 

# 

# Con el ciclo anual de cada parámetro podemos observar la media y la desviación estándar de las mediciones de los 8 años, desde el 2013 hasta el 2021, gracias a esta gráfica se puede comprender el rango de variabilidad de cada uno de los parámetros,

# Figura 4Media y desviación estándar

# 

# Referencias

Ferchichi, A., Abbes, A. B., Barra, V., & Farah, I. R. (2022). Forecasting vegetation indices from spatio-temporal remotely sensed data using deep learning-based approaches: A systematic literature review. *Ecological Informatics*, *68*.

Muradyan, V., Tepanosyan, G., Asmaryan, S., Saghatelyan, A., & Dell'Acqua, F. (2019). Relationships between NDVI and climatic factors in mountain ecosystems: A case study of Armenia. *Remote Sensing Applications: Society and Environment*, *Volume 14*, 158-169. https://doi.org/10.1016/j.rsase.2019.03.004

Pizano, C., & Garcia, H. (Eds.). (2014). *Bosque seco tropical en Colombia*. Instituto de Investigación de Recursos Biológicos Alexander von Humboldt. http://repository.humboldt.org.co/handle/20.500.11761/9333

Schröder, J. M., Ávila Rodríguez, L. P., & Günter, S. (2021). Research trends: Tropical dry forests: The neglected research agenda? *Forest Policy and Economics*, *122*. https://doi.org/10.1016/j.forpol.2020.102333

Torres-Bejarano, F., Padilla Coba, J., Rodríguez Cuevas, c., Ramírez León, H., & Cantero Rodelo, R. (2016). La modelación hidrodinámica para la gestión hídrica del embalse del Guájaro, Colombia. *Revista Internacional de Métodos Numéricos para Cálculo y Diseño en Ingeniería*, *32*. https://doi.org/10.1016/j.rimni.2015.04.001

Wang, T., Wu, Z., Wang, P., Wu, T., Zhang, Y., Yin, J., Yu, J., Wang, H., Guan, X., Xu, H., Yan, D., & Yan, D. (2023). Plant-groundwater interactions in drylands: A review of current research and future perspectives. *Agricultural and Forest Meteorology*, *341*. https://doi.org/10.1016/j.agrformet.2023.109636

Hussein, E. A., Thron, C., Ghaziasgar, M., Bagula, A., & Vaccari, M. (2020, November 17). Groundwater Prediction Using Machine-Learning Tools. *MDPI*, 16. https://doi.org/10.3390/a13110300

Brassington, G. (2017). Mean Absolute Error and Root Mean Square Error: Which Is the Better Metric for Assessing. *Geophysical Research Abstracts*, *19*, 2. https://meetingorganizer.copernicus.org/EGU2017/EGU2017-3574.pdf