

## **Acta de Correcciones al Documento de Trabajo de Grado**

**Santiago de Cali, agosto 2 de 2024**

**Autor: Manuel Andrés Anduquia Ramos, Jesús Alberto Garcés Sevillano**

**Título del Trabajo de Grado: “Evaluación del efecto de la pandemia sobre la Deforestación en Colombia por medio de imágenes de satélite y usando los datos de Terra-i”**

**Director: David Arango Londoño**

Como indica el artículo 2.13 de las Directrices para Trabajo de Grado de Maestría, he verificado que el estudiante indicado arriba ha implementado todas las correcciones que los Jurados del Proyecto de Trabajo de Grado definieron que se efectuaran, como consta en el Acta de Evaluación correspondiente.

*David Arango Londoño*

---

David Arango Londoño

Santiago de Cali, 28 de Junio del 2024

Doctor

**Diego Luis Linares.**

Director Maestría en Ciencia de Datos  
Facultad de Ingeniería y Ciencias  
Pontificia Universidad Javeriana de Cali

**Asunto:** Presentación para evaluación del proyecto aplicado

Cordial Saludo,

Con el fin de cumplir con los requisitos exigidos por la Universidad para optar por el título de Magíster en Ciencia de Datos, nos permitimos presentar a su consideración el proyecto denominado “Evaluación del efecto de la pandemia sobre la Deforestación en Colombia por medio de imágenes de satélite y usando los datos de Terra-i”, el cual fue realizado por el (los) estudiante (s) Manuel Andrés Anduquia Ramos y Jesús Alberto Garcés Sevillano con código (s) 8974535 y 8980165 pertenecientes a la Maestría en Ciencia de Datos, bajo la dirección de David Arango Londoño.

El suscrito director del Proyecto Aplicado autoriza para que se proceda a hacer la evaluación de este proyecto, toda vez que ha revisado cuidadosamente el documento y avala que ya se encuentra listo para ser presentado y sustentado oficialmente.

Atentamente,



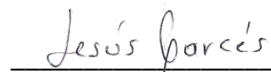
---

Manuel Andrés Anduquia Ramos (Estudiante)  
C.C. 75098249 - Manizales



---

David Arango Londoño (Director)  
C.C. 1130586950 - Cali



---

Jesús Alberto Garcés Sevillano (Estudiante)  
C.C. 1087202648 - Tumaco

**Documentación anexa:**

Resumen del Proyecto Aplicado en formato digital (máximo 1 página).  
Una copia digital (PDF) del documento del proyecto aplicado

## FICHA RESUMEN

### PROYECTO APLICADO – MAESTRÍA EN CIENCIA DE DATOS

**TÍTULO:** Evaluación del efecto de la pandemia sobre la deforestación en Colombia por medio de imágenes de satélite y usando los datos de Terra-i.

1. **ÁREA DE TRABAJO:** Conservación y Biotecnología.
2. **TIPO DE PROYECTO (Aplicado, Innovación, Investigación):** Aplicado
3. **ESTUDIANTE(S):** Jesús Garcés, Manuel Anduquia.
4. **CORREO ELECTRÓNICO:** [jesusgarces101@javerianacali.edu.co](mailto:jesusgarces101@javerianacali.edu.co), [manuelanduquia@javerianacali.edu.co](mailto:manuelanduquia@javerianacali.edu.co)
5. **DIRECCIÓN Y TELEFONO:** Tumaco, Rionegro. (+57 312 5199700, +57 3117807114)
6. **DIRECTOR:** David Arango Londoño
7. **VINCULACIÓN DEL DIRECTOR:** Profesor Facultad Ingeniería y Ciencias - Javeriana Cali
8. **CORREO ELECTRÓNICO DEL DIRECTOR:** [david.arango@javerianacali.edu.co](mailto:david.arango@javerianacali.edu.co)
9. **CO-DIRECTOR (Si aplica):**
10. **GRUPO O EMPRESA QUE LO AVALA (Si aplica):**
11. **OTROS GRUPOS O EMPRESAS:**
12. **PALABRAS CLAVE (al menos 5):** deforestación, colombia, ciencia de datos, machine learning, aprendizaje de máquina supervisado, modelo aprendizaje estadístico, análisis de imágenes, terra-i.
13. **FECHA DE INICIO:** Junio 2023
14. **FECHA FINALIZACIÓN:** Mayo 2024
15. **RESUMEN:**

El objetivo de este proyecto fue evaluar el impacto que la pandemia del COVID-19 tuvo en los niveles de deforestación en Colombia. Se emplearon técnicas de ciencia de datos, que incluyeron modelos estadísticos mixtos y aprendizaje supervisado de máquinas, para analizar grandes volúmenes de datos provenientes de imágenes satelitales de la plataforma Terra-i, la cual monitorea la deforestación. Además, se utilizaron modelos estadísticos para comprender las relaciones y los factores que contribuyeron a la deforestación en áreas específicas identificadas. Estos modelos permitieron identificar patrones y tendencias, incluyendo factores específicos que pudieron haber influido durante la pandemia. Buscando contribuir a la toma de decisiones informadas y a la formulación de estrategias más efectivas para la conservación y el manejo forestal en Colombia.



Pontificia Universidad  
**JAVERIANA**  
Cali

**EVALUACIÓN DEL EFECTO DE LA PANDEMIA SOBRE LA DEFORESTACIÓN EN  
COLOMBIA POR MEDIO DE IMÁGENES DE SATÉLITE Y USANDO LOS DATOS DE  
TERRA-I**

*Manuel Andrés Anduquia Ramos – 8974535*  
*Jesús Alberto Garcés Sevillano - 8980165*

*Proyecto Aplicado para optar al título de  
Magister en Ciencia de Datos*

Director(a)  
David Arango Londoño

FACULTAD DE INGENIERÍA Y CIENCIAS  
MAESTRÍA EN CIENCIA DE DATOS  
SANTIAGO DE CALI, JUNIO 28 DE 2024

## TABLA DE CONTENIDO

<b>INTRODUCCIÓN .....</b>	<b>9</b>
<b>1. DEFINICIÓN DEL PROBLEMA .....</b>	<b>10</b>
<b>1.1. PLANTEAMIENTO DEL PROBLEMA .....</b>	<b>10</b>
<b>1.2. FORMULACIÓN DEL PROBLEMA .....</b>	<b>11</b>
<b>2. OBJETIVOS DEL PROYECTO .....</b>	<b>12</b>
<b>2.1. OBJETIVO GENERAL.....</b>	<b>12</b>
<b>2.2. OBJETIVOS ESPECÍFICOS .....</b>	<b>12</b>
<b>3. MARCO TEÓRICO Y ANTECEDENTES .....</b>	<b>13</b>
<b>3.1. MARCO TEÓRICO .....</b>	<b>13</b>
3.1.1 Clasificación de imágenes.....	13
3.1.2. Análisis de series temporales .....	15
3.1.3. Detección de cambios .....	15
3.1.4. Índices de vegetación .....	16
3.1.5. Modelo de análisis estadístico.....	17
<b>3.2. ANTECEDENTES / TRABAJOS RELACIONADOS .....</b>	<b>18</b>
3.2.1. Relación de los trabajos con el proyecto .....	19
3.2.2. Aportes de los trabajos relacionados .....	20
3.2.3. Diferencias con el proyecto .....	20
<b>4. METODOLOGÍA .....</b>	<b>21</b>
<b>5. ANÁLISIS DESCRIPTIVO .....</b>	<b>23</b>
<b>5.1. PLANEACIÓN .....</b>	<b>23</b>
5.1.1. Entendimiento del negocio .....	23
5.1.2. Entendimiento de los datos.....	24
5.1.3. Documentación de librerías.....	25
<b>5.2. PREPARACIÓN DE LOS DATOS .....</b>	<b>26</b>
5.2.1. Definición fuente datos .....	26
5.2.2. Recolección y exploración variables .....	27
5.2.3. Preparación de los datos .....	28
5.2.4. Transformación de los datos .....	29
<b>5.3. ANÁLISIS DE ZONAS AFECTADAS .....</b>	<b>31</b>
5.3.1. Zonas afectadas por la deforestación en Colombia - 2018.....	31
5.3.2. Zonas afectadas por la deforestación en Colombia – 2019 y 2020 .....	32
5.3.3. Zonas afectadas por la deforestación en Colombia - 2021.....	33

5.3.4. Zonas afectadas por la deforestación en Colombia - 2022.....	34
<b>6. ANÁLISIS EXPLORATORIO DE DATOS.....</b>	<b>35</b>
6.1. ANÁLISIS DE LAS SERIES DE TIEMPO .....	36
6.2. ANÁLISIS ESTADÍSTICO .....	39
6.3. ANÁLISIS DE COMPORTAMIENTO DE LOS DATOS .....	40
<b>7. MODELO ESTADÍSTICO MIXTO .....</b>	<b>42</b>
7.1. MODELO ESTADÍSTICO .....	42
7.2. ANÁLISIS DESCRIPTIVO.....	42
7.3. VALIDACIÓN Y EVALUACIÓN DEL MODELO.....	43
7.3.1. Modelo Estadístico Global.....	43
7.3.2. Modelo Estadístico del Caqueta .....	45
7.3.4. Modelo Estadístico del Guaviare .....	49
7.3.5. Evaluación de Correlación del Modelo .....	51
<b>8. CONCLUSIONES Y TRABAJOS FUTUROS .....</b>	<b>53</b>
8.1. CONCLUSIONES .....	53
8.2. TRABAJOS FUTUROS .....	54
<b>9. REFERENCIAS BIBLIOGRÁFICAS.....</b>	<b>55</b>

## LISTA DE FIGURAS

Figura 1. Ejemplo de índice de vegetación de diferencia normalizado (NVDI) .....	16
Figura 2. Terra-i sistema de monitoreo de la pérdida de hábitat en LATAM.....	23
Figura 3. MODIS - Mapa de índice vegetación .....	24
Figura 4. Análisis de píxeles en Raster .....	25
Figura 5. Interacción de mapas con Rnaturalearth .....	25
Figura 6. Archivos Tiles ASCII – Colombia del año 2018 al 2022 .....	27
Figura 7. Tipos de objetos del RasterLayer unificado (ejemplo Colombia 2022) .....	28
Figura 8. Información de coordenadas de los archivos ASCII .....	29
Figura 9. Transformación de datos para la generación de mapa con reducción de vegetación..	29
Figura 10. Mapa Colombia transformado con reducción de vegetación 2022 .....	30
Figura 11. Mapa: Caquetá - Depto. con mayor pérdida de cobertura forestal en el 2018 .....	31
Figura 12. Mapa: Guaviare - Depto. con mayor pérdida de cobertura forestal en el 2019 .....	32
Figura 13. Mapa: Meta - Depto. con mayor pérdida de cobertura forestal en el 2020 .....	32
Figura 14. Mapa: Caquetá - Depto. con mayor pérdida de cobertura forestal en el 2021 .....	33
Figura 15. Mapa: Caquetá - Depto. con mayor pérdida de cobertura forestal en el 2022 .....	34
Figura 16. Código R: Series de tiempo usando calendario Juliano.....	37
Figura 17. Código R: Gráfica Análisis Series de Tiempo .....	37
Figura 18. Análisis series de tiempo - Deptos. Deforestación Colombia 2018 – 2022 .....	38
Figura 19. Comportamiento de los datos - Deptos. Deforestación Colombia 2018 - 2022 .....	40
Figura 20. Código R: Gráfica de comportamiento de los datos .....	41
Figura 21. Modelo de análisis estadístico Global .....	44
Figura 22. Modelo de análisis estadístico del Caquetá .....	46
Figura 23. Modelo de análisis estadístico del Meta .....	48
Figura 24. Modelo de análisis estadístico del Guaviare .....	50
Figura 25. Evaluación de correlación del modelo estadístico .....	51
Figura 26. Código R: Gráfica de correlación .....	51

## LISTA DE TABLAS

Tabla 1. Eventos de El Niño (rojo) y La Niña (azul) desde el 2010 hasta el 2022 .....	27
Tabla 2. Conversión de Píxeles a Hectáreas .....	30
Tabla 3. Análisis series de tiempo – Deptos. Deforestación Colombia 2018 - 2022 .....	36
Tabla 4. Análisis estadístico – Deptos. Deforestados antes y después COVID-19.....	39

## INTRODUCCIÓN

La deforestación es un problema ambiental grave que afecta a varios países del mundo en especial a los suramericanos, ya que muchos de estos hacen parte de la Amazonia dentro de sus límites territoriales el cual es considerada como "el pulmón del planeta" [1], esto incluye a Colombia la cual cuenta con una rica biodiversidad, pero desafortunadamente, ha sido víctima de la actividad humana especialmente en las últimas décadas donde la tala indiscriminada de árboles, la minería ilegal, entre otros factores ambientales, han llevado a una pérdida significativa de los bosques y selvas [2].

Colombia es uno de los países con mayor cobertura forestal del planeta, por ende, se requiere tener control y acceso del estado a diversas regiones históricamente afectadas por la deforestación y así prever ciertas situaciones de emergencia que puedan conducir a pérdidas contundentes y degradación ambiental. Sin embargo, la pandemia del COVID-19 tuvo un efecto significativo en los niveles de deforestación en Colombia, y no precisamente de manera favorable [3]. La crisis sanitaria, la cual nos puso en cuarentena indefinida y aislamiento obligatorio, fomentó el aumento de conductas ilícitas que afectan a los bosques y selvas colombianas debido a una creciente necesidad de monitoreo y falta de presencia en estas regiones, la deforestación ha aumentado en tiempos de pandemia [4]. Esta situación representa una amenaza para la sostenibilidad ambiental y pone en peligro una de las principales fortalezas del país que es su biodiversidad.

Ante ello este proyecto se propuso evaluar el impacto de la pandemia del COVID-19 en los niveles de deforestación en zonas definidas y previamente afectadas en Colombia. Para esto se utilizaron técnicas y herramientas de ciencia de datos, aprendizaje de máquina y un modelo estadístico mixto, junto con imágenes de satélite y datos de proporcionados por la plataforma Terra-i, que permitieron analizar y comprender cómo la pandemia ha influido en los patrones de deforestación en el país.

Por tanto, este estudio es importante en la condición actual del país ya que, proporciona información relevante para fomentar la toma de decisiones acertadas en la conservación y manejo forestal por parte de las entidades competentes que les permita priorizar regiones y analizar patrones, así como contribuir al conocimiento sobre la interacción entre eventos globales y la deforestación.

## **1. DEFINICIÓN DEL PROBLEMA**

### **1.1. PLANTEAMIENTO DEL PROBLEMA**

Es conocido el importante rol que desempeña la cobertura forestal en el planeta, desde servir como proveedores de insumos y medio de vida para muchas poblaciones, hasta ser el sustento de la biodiversidad terrestre del planeta [5]. En Colombia, como en otros países de Latinoamérica existe un gran problema en este aspecto, la conversión de los bosques en tierras agrícolas y ganaderas, la tala y minería ilegal, incendios provocados y la falta de monitoreo, han afectado directamente a la deforestación de miles de hectáreas, siendo la presencia humana el común denominador en estos sectores más afectados [6].

Sin embargo, a finales del 2019 se determinó un estado de emergencia sanitaria debido a la pandemia provocada por el SARS-CoV-2 conocido como COVID-19, lo cual promulgó un aislamiento obligatorio y una cuarentena indefinida que restringía el tránsito libre de las personas, esto promovió en teoría un menor impacto humano en los sectores forestales afectados y por ende una disminución en la deforestación [7].

La necesidad de realizar una evaluación de la deforestación en Colombia en periodos antes, durante y después de la pandemia es importante, ya que con las herramientas basadas en datos se podrían mostrar con claridad los cambios progresivos de la deforestación en estas zonas; y así las entidades competentes pueden contar con la información que les permitan identificar en detalle, las causas principales de esta problemática, esto fomentara la toma acertada de decisiones al momento de implementar planes de choque para disminuir el impacto negativo en las regiones.

## 1.2. FORMULACIÓN DEL PROBLEMA

Se llevó a cabo un análisis de imágenes satelitales suministradas por la plataforma Terra-i que ayudó a evaluar el impacto de la pandemia en la deforestación progresiva en Colombia. Esta evaluación se realizó con el uso de herramientas de ciencia de datos tomando periodos de tiempo claves e identificando las regiones más afectadas en Colombia. Con esto pudimos contestar a la pregunta:

**¿Cuál fue el impacto específico de la pandemia del COVID-19 en los niveles de deforestación en zonas históricamente afectadas en Colombia?**

El modelo estadístico mixto también ayudó a responder las siguientes preguntas:

- ¿Cuál fue el impacto cuantitativo de la deforestación en las diferentes zonas afectadas en Colombia antes, durante y después de la pandemia COVID-19?
- ¿Cómo se puede evaluar de manera efectiva, utilizando herramientas de ciencia de datos, la deforestación en Colombia en tiempos de pandemia?
- ¿Que tanto progresó en la pandemia, los niveles de deforestación en las regiones históricamente más afectadas?

## **2. OBJETIVOS DEL PROYECTO**

### **2.1. OBJETIVO GENERAL**

Evaluar el impacto en la deforestación atribuido al efecto de la pandemia COVID-19 a partir de imágenes de satélite y uso de técnicas de procesamiento de datos y modelos estadísticos para analizar las consecuencias producidas por esta problemática en las zonas que históricamente tienen cifras más altas en Colombia.

### **2.2. OBJETIVOS ESPECÍFICOS**

- Realizar un análisis descriptivo que permita definir indicadores y series de tiempo para cuantificar el impacto de la deforestación en diversas zonas afectadas del país antes, durante y después de la pandemia, contrastando con los resultados obtenidos al analizar las imágenes satelitales proporcionados por Terra-i.
- Identificar y comparar, mediante un análisis exploratorio de datos, el comportamiento progresivo de la deforestación en los periodos de la pandemia analizados, viendo cambios drásticos en las imágenes satelitales, para así definir y priorizar las regiones de Colombia más afectadas.
- Desarrollar un modelo estadístico mixto para evaluar el impacto de la deforestación en las zonas afectadas de Colombia utilizando datos e imágenes satelitales.

### 3. MARCO TEÓRICO Y ANTECEDENTES

#### 3.1. MARCO TEÓRICO

La falta de herramientas tecnológicas para evaluar la deforestación en Colombia representa un desafío significativo en la lucha contra este problema ambiental. La ciencia de datos desempeña un papel crucial en la recopilación, análisis y visualización de información sobre la deforestación, permitiendo identificar patrones, tendencias y áreas prioritarias de intervención.

Existen diversas técnicas de procesamiento de imágenes de satélite [8] que pueden ser de gran utilidad en la detección y monitoreo de la deforestación. Algunas de estas técnicas incluyen:

##### 3.1.1 Clasificación de imágenes

Mediante algoritmos de clasificación, se pudo identificar y distinguir áreas forestales de aquellas que han sido deforestadas. Se utilizaron diferentes enfoques, como clasificación supervisada y no supervisada, para asignar categorías a los píxeles de la imagen y determinar la cobertura forestal.

##### Clasificación supervisada:

- En la clasificación supervisada [9], se utilizaron conjuntos de datos etiquetados previamente para entrenar algoritmos de clasificación. Estos conjuntos de datos consisten en imágenes satelitales junto con información sobre la cobertura del suelo o la vegetación en diferentes ubicaciones.
- Durante el proceso de entrenamiento, el algoritmo aprende a distinguir entre diferentes clases (por ejemplo, bosques, áreas deforestadas, cultivos, agua) mediante la extracción de características de las imágenes.
- Una vez que el modelo está entrenado, se aplica a nuevas imágenes satelitales para asignar categorías a cada píxel en función de las características aprendidas durante el entrenamiento.

### **Clasificación no supervisada:**

- En contraste, la clasificación no supervisada [10] no requiere conjuntos de datos etiquetados previamente. En su lugar, el algoritmo identifica automáticamente patrones y agrupa píxeles en clases basadas en similitudes en sus características espectrales.
- Este enfoque es útil cuando no se dispone de conjuntos de datos etiquetados o cuando se busca descubrir patrones o clases desconocidas en los datos.

### **Técnicas de extracción de características:**

- Para la clasificación de imágenes [11], es crucial extraer características relevantes de los datos, como la reflectancia espectral en diferentes bandas, la textura, la forma y la estructura espacial.
- Se utilizan diversas técnicas para la extracción de características, incluyendo métodos estadísticos, transformaciones matemáticas y filtros espaciales.

### **Algoritmos de clasificación:**

- Existen varios algoritmos de clasificación [12] utilizados en el análisis de imágenes satelitales, incluyendo Máquinas de Vectores de Soporte (SVM), Bosques Aleatorios, Redes Neuronales Convolucionales (CNN), y métodos de clasificación basados en reglas.
- Cada algoritmo tiene sus propias características y es más adecuado para diferentes tipos de datos y problemas de clasificación.

### **Evaluación de la clasificación:**

- Después de aplicar el modelo estadístico mixto, se evalúa la precisión y el rendimiento del modelo utilizando métricas [13] como el error estándar residual,  $r$ -cuadrado múltiple y ajustado, estadístico F, grados de libertad, valor p del modelo, entre otros.
- Esta evaluación proporcionó información sobre la confiabilidad de los resultados de la clasificación y permite ajustar y mejorar el modelo si es necesario.

### **3.1.2. Análisis de series temporales**

Al comparar imágenes de satélite capturadas en diferentes momentos, se pueden detectar cambios en la cobertura forestal a lo largo del tiempo. Esto permitió identificar áreas deforestadas recientemente, así como aquellas que han experimentado procesos de regeneración o degradación.

#### **Concepto de series temporales:**

- Una serie temporal [14] es una secuencia de datos recopilados en intervalos regulares a lo largo del tiempo. En el caso de este proyecto, las imágenes de satélite capturadas en diferentes momentos constituyeron una serie temporal de observaciones de la cobertura del suelo y la vegetación.

#### **Detección de cambios:**

- Al comparar imágenes de satélite tomadas en diferentes momentos, se pudieron detectar cambios en la cobertura forestal, como la deforestación, la regeneración o la degradación.
- Estos cambios se identificaron mediante el análisis de las diferencias en las características espectrales de los píxeles entre diferentes períodos de tiempo.

#### **Métodos de análisis:**

- Existen varios métodos y técnicas para analizar series temporales de imágenes satelitales, como el cálculo de índices de vegetación (por ejemplo, NDVI), el análisis de componentes principales (PCA) y la comparación de firmas espectrales.
- Estos métodos permiten identificar patrones de cambio a lo largo del tiempo y evaluar la magnitud y la dirección de estos cambios.

### **3.1.3. Detección de cambios**

Esta técnica se basa en comparar dos o más imágenes de satélite adquiridas en diferentes momentos para identificar y cuantificar los cambios en la cobertura forestal. Se utilizaron algoritmos de detección de cambios [15] para resaltar las áreas que han sufrido deforestación, degradación o regeneración.

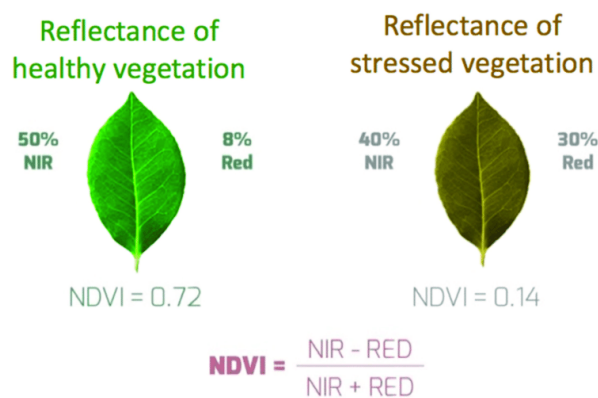
### 3.1.4. Índices de vegetación

Los índices de vegetación, como el Índice de Vegetación de Diferencia Normalizada (NDVI), permitieron evaluar la salud y densidad de la vegetación. Al comparar imágenes de satélite que representan diferentes períodos, se pudo identificar áreas donde ha ocurrido deforestación, ya que presentaron valores de índices de vegetación más bajos.

#### NDVI (Índice de Vegetación de Diferencia Normalizada):

El Índice de Vegetación de Diferencia Normalizada (NDVI) [16] es una herramienta crucial para monitorear la cobertura vegetal. Este índice se calcula utilizando la diferencia entre la reflectancia en el infrarrojo cercano (NIR) y la reflectancia en el rojo (RED), dividida por la suma de ambas (NIR + RED). Los valores del NDVI oscilan entre -1 y 1; valores más altos indican una mayor densidad y salud de la vegetación, mientras que valores más bajos sugieren menor vegetación o áreas deforestadas.

Figura 1. Ejemplo de índice de vegetación de diferencia normalizado (NDVI)



Fuente: Research Gate

#### MODIS (Moderate Resolution Imaging Spectroradiometer):

Sensor satelital a bordo de dos satélites de la NASA: Terra y Aqua. Estos satélites orbitan la Tierra y MODIS [17] recolecta información sobre la Tierra y el clima.

- Tiene la capacidad de observar casi toda la superficie terrestre cada 1 o 2 días
- Captura información en 36 bandas del espectro electromagnético, lo que permite a los científicos estudiar diferentes características de la Tierra, como la vegetación, la temperatura y la composición atmosférica.

### 3.1.5. Modelo de análisis estadístico

Los modelos de análisis estadísticos [18], son herramientas matemáticas diseñadas para explicar la variabilidad en los datos y las conexiones entre variables. Se clasifican en **descriptivos**, **inferenciales** y **predictivos**. Los descriptivos resumen los datos, los inferenciales extraen conclusiones de una muestra y los predictivos anticipan resultados futuros. Este análisis se realizó usando R-Studio, una plataforma potente para el análisis de datos y visualización.

### 3.1.6. Validación de modelos estadísticos

La validación de un modelo estadístico es el proceso de determinar su capacidad para generalizarse a nuevos datos. Existen varias técnicas de validación que se implementaron en R-Studio, entre las cuales destacan:

- **Validación cruzada (Cross-Validation)** [19]: Este método divide el conjunto de datos en múltiples subconjuntos (folds). El modelo se entrena en algunos subconjuntos y se prueba en los restantes, repitiendo el proceso para obtener una medida promedio del rendimiento. En R-Studio, esto se realizó con funciones del paquete caret, como trainControl y train.

### 3.1.7. Evaluación de modelos estadísticos

Evaluar un modelo estadístico [20], implica medir su desempeño en términos de precisión, exactitud y otros criterios relevantes.

- **Métricas de rendimiento:**
  - **Error Cuadrático Medio (Mean Squared Error, MSE):** Mide la media de los cuadrados de los errores o desviaciones, proporcionando una visión del error promedio en las predicciones [21].
  - **Coefficiente de Determinación ( $R^2$ ):** Indica la proporción de la varianza en la variable dependiente que es explicada por el modelo [22].
  - **Precisión, Sensibilidad y Especificidad:** Especialmente importantes en modelos de clasificación, estas métricas evalúan la capacidad del modelo para predecir correctamente las clases de los datos [23].
- **Gráficos de diagnóstico:**
  - **Gráfico de Residuos:** Muestra la diferencia entre los valores observados y los valores predichos por el modelo, ayudando a identificar patrones de error.

### 3.2. ANTECEDENTES / TRABAJOS RELACIONADOS

A continuación, se resaltan algunos estudios de investigación que se relacionan con el tema de análisis propuesto del impacto de la deforestación en Colombia:

**"Monitoreo y Seguimiento al Fenómeno de la Deforestación en Colombia"** [21] (Autores: IDEAM – Instituto de Hidrología, Meteorología y Estudios - 2022): El análisis del IDEAM muestra los mapas de cambio de cobertura de bosque/no bosque (1990-2014) y las alertas tempranas por deforestación AT-D (2022). Este estudio tiene una relación importante con el proyecto que se desarrolló, ya que se utilizó datos de Terra-i para el procesamiento digital de imágenes de satélite de baja resolución, donde se identifican áreas forestales donde existe pérdida de bosque natural.

**"Deforestación en la Amazonía y su impacto en el cambio climático"** [22] (Autores: Silva, J., Martínez, R., & López, M. - 2019): Este estudio analiza la deforestación en la región amazónica y su conexión con el cambio climático. Utiliza datos satelitales y modelos de simulación para evaluar las emisiones de gases de efecto invernadero asociadas a la pérdida de cobertura forestal. El estudio destaca la importancia de abordar la deforestación como parte de la estrategia global para mitigar el cambio climático.

**"Análisis de los factores socioeconómicos y políticos que impulsan la deforestación en Colombia"** [23] (Autores: Rodríguez, L., Gómez, S., & Ramírez, E. - 2020): Este trabajo examina los factores socioeconómicos y políticos que contribuyen a la deforestación en Colombia. Utiliza datos de encuestas y entrevistas para identificar las principales causas de la pérdida de bosques en el país. El estudio concluye que la expansión de la agricultura, la minería ilegal y la falta de gobernanza efectiva son factores clave que impulsan la deforestación en Colombia.

**"Evaluación del impacto de la deforestación en la biodiversidad en América Latina"** [24] (Autores: García, A., López, J., & Morales, R. - 2018): Este estudio investiga los efectos de la deforestación en la biodiversidad de América Latina. Utilizando datos de inventarios biológicos y análisis espaciales, el estudio examina los cambios en la diversidad de especies y la pérdida de hábitats debido a la deforestación. Los resultados resaltan la necesidad de conservar los bosques tropicales para proteger la biodiversidad única de la región.

### 3.2.1. Relación de los trabajos con el proyecto

- **Base histórica y metodológica:** El estudio "*Monitoreo y Seguimiento al Fenómeno de la Deforestación en Colombia*" realizado por el IDEAM en 2022 proporcionó una base histórica y metodológica crucial para nuestro proyecto. Este trabajo nos ofreció una línea base metodológica que pudimos utilizar para validar y comparar nuestros resultados, asegurando que nuestras técnicas de procesamiento de imágenes y modelos de aprendizaje automático estuvieran alineadas con prácticas reconocidas y efectivas.
- **Deforestación y cambio climático:** El estudio "*Deforestación en la Amazonía y su Impacto en el Cambio Climático*" de Silva, Martínez y López en 2019 destaca la conexión entre la deforestación y el cambio climático. Este trabajo fue relevante para nuestro proyecto ya que proporcionó una perspectiva global sobre cómo la deforestación influye en el clima. La comprensión de esta relación nos permitió contextualizar mejor el impacto de la pandemia de COVID-19 en la deforestación en Colombia, integrando consideraciones climáticas en nuestro análisis para ofrecer una visión más holística de las consecuencias ambientales.
- **Factores socioeconómicos:** El estudio "*Análisis de los Factores Socioeconómicos y Políticos que Impulsan la Deforestación en Colombia*" de Rodríguez, Gómez y Ramírez en 2020 identifica los factores socioeconómicos y políticos que contribuyen a la deforestación en Colombia. Este trabajo fue esencial para nuestro proyecto porque proporcionó un contexto sobre los factores subyacentes que pueden haber exacerbado la deforestación durante la pandemia. Al incorporar estas perspectivas, pudimos entender mejor las causas detrás de los patrones de deforestación observados en nuestros análisis, lo que nos permitió desarrollar modelos predictivos más precisos y relevantes para la formulación de políticas y estrategias de mitigación.

### 3.2.2. Aportes de los trabajos relacionados

- **Metodología y validación:** Proporcionaron métodos y datos históricos que ayudaron a validar nuestros resultados y mejorar la precisión de nuestros modelos.
- **Perspectiva multidimensional:** Ofrecieron una comprensión más amplia de cómo la deforestación se relaciona con el cambio climático y la biodiversidad.
- **Herramientas y Técnicas:** Nos brindaron técnicas de análisis y procesamiento de imágenes que pudimos adaptar y aplicar en nuestro proyecto.

### 3.2.3. Diferencias con el proyecto

- **Enfoque Específico en la Pandemia:** Nuestro proyecto se centró en el impacto específico de la pandemia de COVID-19 en la deforestación, algo que los otros trabajos no abordan.
- **Uso de Ciencia de Datos y Aprendizaje Automático:** Empleamos técnicas avanzadas de ciencia de datos y modelos de aprendizaje automático para analizar imágenes satelitales, lo cual difiere de los enfoques metodológicos en los otros estudios.
- **Cobertura Temporal y Geográfica:** Nos enfocamos en periodos específicos antes, durante y después de la pandemia en Colombia, mientras que los otros estudios tienen un enfoque temporal y geográfico más amplio o diferente, como es el caso del proyecto *“Evaluación del impacto de la deforestación en la biodiversidad en América Latina”*.
- **Aplicación Práctica:** Nuestro objetivo fue proporcionar información que pueda ser utilizada directamente para la toma de decisiones en la conservación y manejo forestal en Colombia, con un enfoque en el contexto de la pandemia.

## 4. METODOLOGÍA

Este proyecto se llevó a cabo siguiendo la metodología del Proceso de Minería de Datos Cruzado (CRISP-DM) [25], ampliamente reconocida por su efectividad en proyectos de ciencia de datos, incluidos aquellos que involucran análisis de imágenes satelitales. Esta metodología se dividió en seis fases claramente definidas: comprensión del negocio, comprensión de los datos, preparación de los datos, modelado, evaluación y despliegue. Este enfoque metodológico estructurado, garantizó una implementación completa y eficiente del análisis de datos necesario para lograr los objetivos del proyecto.

### **Comprensión del negocio:**

- Identificar los objetivos del proyecto, como la identificación de zonas de deforestación y la comprensión de su evolución en el tiempo.
- Definir los criterios y métricas para determinar qué se considera como deforestación.
- Identificar a los interesados y comprender sus necesidades y expectativas con respecto al proyecto.
- Establecer un plan de proyecto inicial que incluya un cronograma tentativo y un presupuesto estimado.

### **Comprensión de los datos:**

- Recopilar los datos de Terra-i disponibles, incluyendo imágenes satelitales y otros datos relacionados con la cobertura forestal.
- Explorar la estructura de los datos y comprender su formato y calidad.
- Identificar posibles problemas de calidad de datos, como datos faltantes o errores de medición.
- Evaluar la idoneidad de los datos para los objetivos del proyecto y determinar si se necesitan datos adicionales de otras fuentes.

### **Preparación de los datos:**

- Limpiar los datos eliminando duplicados, datos irrelevantes o inconsistencias.
- Transformar los datos en un formato adecuado para su análisis, como convertir imágenes satelitales en conjuntos de datos estructurados que se puedan utilizar en modelos de análisis.
- Dividir los datos en conjuntos de entrenamiento y prueba para su posterior modelado y evaluación.

### **Modelado:**

- Seleccionar y aplicar técnicas de análisis de imágenes satelitales, como clasificación supervisada o no supervisada, para identificar zonas de deforestación.
- Desarrollar modelos estadísticos que puedan detectar cambios en la cobertura forestal a lo largo del tiempo.
- Ajustar los parámetros para evaluar y comparar diferentes modelos para determinar cuál es el más adecuado para los objetivos del proyecto.

### **Evaluación:**

- Evaluar la precisión y eficacia de los modelos desarrollados mediante la comparación de los resultados del modelo con datos de referencia conocidos.
- Realizar análisis de sensibilidad para evaluar la robustez del modelo frente a diferentes escenarios.
- Identificar posibles limitaciones y áreas de mejora para los modelos desarrollados.

### **Despliegue:**

- Implementar los modelos desarrollados en un entorno operativo o en un sistema de información geográfica (SIG).
- Proporcionar los resultados del análisis a los interesados y asegurar su comprensión y aceptación.
- Establecer procesos para el monitoreo continuo de las zonas identificadas como deforestadas y actualizar los modelos según sea necesario.

## 5. ANÁLISIS DESCRIPTIVO

### 5.1. PLANEACIÓN

#### 5.1.1. Entendimiento del negocio

Terra-i [26] es un sistema de monitoreo ambiental que utiliza tecnología de detección remota satelital y aprendizaje automático para rastrear cambios en la cobertura vegetal y la deforestación en América Latina. Fue desarrollada como una herramienta para la vigilancia en gestión de recursos naturales y conservación ambiental en la región, con la finalidad de proporcionar información oportuna y detallada sobre la pérdida de vegetación, lo que es crucial para la detección temprana de la deforestación.

Terra-i monitorea en tiempo casi real la cobertura de la tierra o el cambio de hábitat usando el sensor MODIS de la NASA y las plataformas Terra y Aqua [27]. Emplean algoritmos de aprendizaje automático para generar mapas y análisis del cambio de hábitat cada 16 días desde 2004 hasta la actualidad. Estos datos pueden visualizarse en un geonavegador (mapa web) o descargarse libremente en archivos de Sistemas de Información Geográfica (SIG) para su posterior análisis.

La información proporcionada por Terra-i fue fundamental para el estudio y análisis de este proyecto al contar con datos precisos y actualizados sobre los cambios en la cobertura vegetal, se pudo evaluar con mayor exactitud el impacto de la pandemia en la deforestación. La capacidad de visualizar y analizar estos datos en formatos ASCII/píxeles facilitó la identificación de áreas críticas y la toma de decisiones informadas para la gestión de recursos naturales y la conservación ambiental en Colombia.

*Figura 2. Terra-i sistema de monitoreo de la pérdida de hábitat en LATAM.*

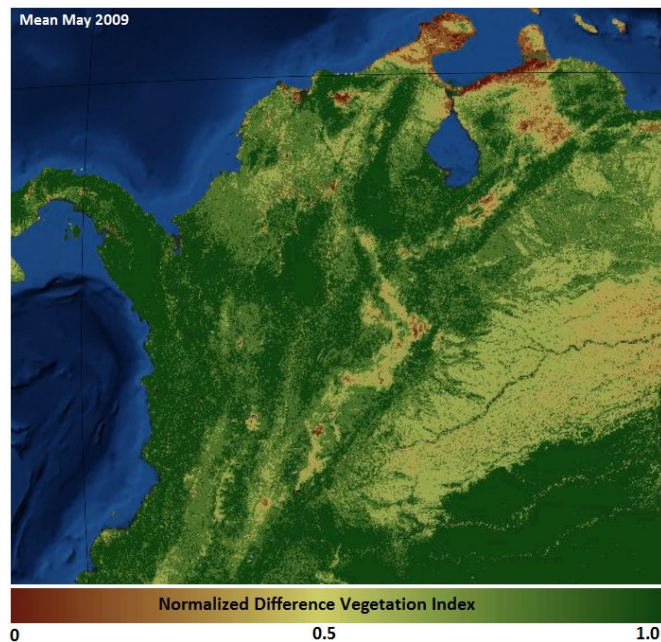


*Fuente: Terra-i*

### 5.1.2. Entendimiento de los datos

- **Datos de entrada (Input Data):** Terra-i utiliza datos del sensor Moderate-Resolution Imaging Spectroradiometer (MODIS), que proporciona imágenes de toda la superficie del globo cada 1 o 2 días con una resolución espacial de 250 m por píxel [28].
- **Datos de salida (Output Data):** Terra-i genera mapas de cambios en el hábitat cada 16 días para cada "tile" de MODIS en LATAM desde el año 2004 hasta la fecha actual.
- **Datos de los índices de vegetación de MODIS:** Los índices de vegetación de MODIS se diseñan para comparar de manera constante los cambios en la vegetación a lo largo del tiempo y el espacio, considerando la radiación fotosintéticamente activa, la clorofila, el área de hojas y las características de las plantas.

*Figura 3. MODIS - Mapa de índice vegetación*



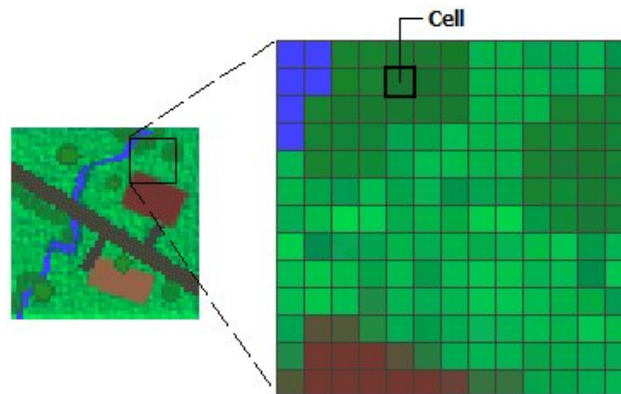
*Fuente: Terra-i*

### 5.1.3. Documentación de librerías

#### Análisis de imágenes en R

- **Raster**: Librería para el procesamiento y análisis de datos raster y geospaciales en R. [29]
- **Rgdal**: Para manejo de datos geospaciales en R. [30]
- **Maps**: Librería para la creación y visualización de mapas temáticos en R.
- **Leaflet**: Librería para la creación de mapas interactivos en R.
- **Rnaturalearth**: Librería para la interacción de vectores con mapas naturales. [31]

Figura 4. Análisis de píxeles en Raster



Fuente: Documentación RasterLayer

Figura 5. Interacción de mapas con Rnaturalearth



Fuente: Documentación Rnaturalearth

## 5.2. PREPARACIÓN DE LOS DATOS

### 5.2.1. Definición fuente datos

Esta etapa implicó identificar y definir las fuentes de datos relevantes para el análisis de la deforestación en Colombia durante la pandemia COVID-19. Se determinaron las fuentes primarias, como las imágenes satelitales proporcionadas por Terra-i, así como otras fuentes secundarias de datos ambientales.

**Datos Terra-i:** Los datos de Terra-i están en formato ARC ASCII [32], con una resolución espacial de 250 m y datum WGS84. Derivan del USGS/NASA MODIS y han sido procesados por CIAT para crear mapas de cambio de hábitat. Los algoritmos de Reymondin et al. (2012) detectan cambios en la cobertura del suelo anualmente desde 2004 hasta el 2022. Un valor de 0 indica que el píxel no cambió, mientras que otros valores representan cambios detectados en períodos de 16 días.

**Datos Ambientales:** El clima en los departamentos de Caquetá, Guaviare y Meta está influenciado por los fenómenos de El Niño y La Niña, que afectan los ciclos de lluvia y sequía, impactando la cobertura forestal. Durante el período estudiado, se observó lo siguiente:

- **Primer Trimestre:** Estos meses suelen ser secos, especialmente con El Niño, como en 2019. Sin embargo, en 2018, incendios forestales y la actividad humana devastaron la región debido a este fenómeno. [33]
- **Segundo y Tercer Trimestres:** La temporada de lluvias se extiende de abril a noviembre, con picos en mayo y octubre. La ganadería extensiva aumenta debido al pasto abundante, aunque la deforestación es menor en este período. [34]
- **Cuarto Trimestre:** La transición hacia la temporada seca comienza en diciembre, generalmente con la llegada de El Niño, iniciando un nuevo ciclo climático.

Cabe resaltar que en los años posteriores a la pandemia, el predominio de La Niña ha resultado en un aumento de la precipitación y una reducción en los incendios forestales.

Tabla 1. Eventos de El Niño (rojo) y La Niña (azul) desde el 2010 hasta el 2022

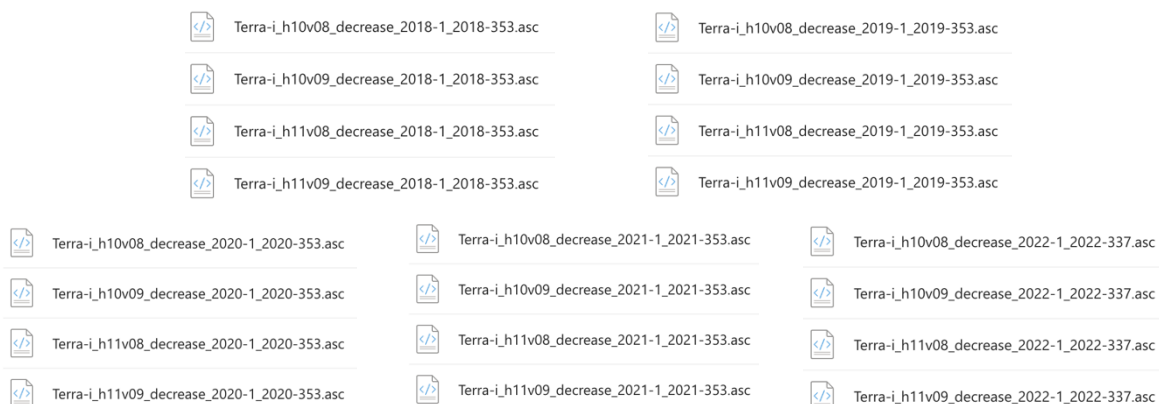
Year	DJF	JFM	FMA	MAM	AMJ	MJJ	JJA	JAS	ASO	SON	OND	NDJ
2010	1.5	1.2	0.8	0.4	-0.2	-0.7	-1.0	-1.3	-1.6	-1.6	-1.6	-1.6
2011	-1.4	-1.2	-0.9	-0.7	-0.6	-0.4	-0.5	-0.6	-0.8	-1.0	-1.1	-1.0
2012	-0.9	-0.7	-0.6	-0.5	-0.3	0.0	0.2	0.4	0.4	0.3	0.1	-0.2
2013	-0.4	-0.4	-0.3	-0.3	-0.4	-0.4	-0.4	-0.3	-0.3	-0.2	-0.2	-0.3
2014	-0.4	-0.5	-0.3	0.0	0.2	0.2	0.0	0.1	0.2	0.5	0.6	0.7
2015	0.5	0.5	0.5	0.7	0.9	1.2	1.5	1.9	2.2	2.4	2.6	2.6
2016	2.5	2.1	1.6	0.9	0.4	-0.1	-0.4	-0.5	-0.6	-0.7	-0.7	-0.6
2017	-0.3	-0.2	0.1	0.2	0.3	0.3	0.1	-0.1	-0.4	-0.7	-0.8	-1.0
2018	-0.9	-0.9	-0.7	-0.5	-0.2	0.0	0.1	0.2	0.5	0.8	0.9	0.8
2019	0.7	0.7	0.7	0.7	0.5	0.5	0.3	0.1	0.2	0.3	0.5	0.5
Year	DJF	JFM	FMA	MAM	AMJ	MJJ	JJA	JAS	ASO	SON	OND	NDJ
2020	0.5	0.5	0.4	0.2	-0.1	-0.3	-0.4	-0.6	-0.9	-1.2	-1.3	-1.2
2021	-1.0	-0.9	-0.8	-0.7	-0.5	-0.4	-0.4	-0.5	-0.7	-0.8	-1.0	-1.0
2022	-1.0	-0.9	-1.0	-1.1	-1.0	-0.9	-0.8	-0.9	-1.0	-1.0	-0.9	-0.8

Fuente: Climate Prediction Center - ONI

## 5.2.2. Recolección y exploración variables

Se llevó a cabo la recolección de datos desde los repositorios de **Terra-i** se descargaron los tiles en formato ASCII de **pérdida de vegetación de Colombia** y las fronteras de Ecuador, Venezuela, Brasil del año **2018 al 2022**. Cada archivo ASCII tiene un tamaño aproximado de 90MB y contiene coordenadas de pixels en formato texto.

Figura 6. Archivos Tiles ASCII – Colombia del año 2018 al 2022



Fuente: Terra-i

### 5.2.3. Preparación de los datos

Los archivos shapes identifican áreas afectadas por la pérdida de vegetación mediante polígonos que marcan sus límites geográficos. Cada archivo contiene atributos como el nombre del país, el área afectada, la fecha de detección y otros metadatos relacionados con la pérdida de cobertura vegetal.

Se realizó una exploración análisis de los tipos de objetos y variables incluidas de los mapas dentro del rasterlayer de cada archivo shape dentro de los tiles. Se depuró e integró la información de los mapas de Colombia arrojando un nuevo archivo shape unificado por cada año (2018 al 2022) con una matriz de 6833x5832x1 pixels.

Figura 7. Tipos de objetos del RasterLayer unificado (ejemplo Colombia 2022)

Name	Type	Value
map_all_col_22	S4 [6833 x 5832 x 1] (raster:)	S4 object of class RasterLayer
file	S4 (raster:::RasterFile)	S4 object of class .RasterFile
data	S4 (raster:::SingleLayerData)	S4 object of class .SingleLayerData
values	double [39850056]	NA NA NA NA NA NA ...
offset	double [1]	0
gain	double [1]	1
inmemory	logical [1]	TRUE
fromdisk	logical [1]	FALSE
isfactor	logical [1]	FALSE
attributes	list [0]	List of length 0
haveninmax	logical [1]	TRUE
min	double [1]	0
max	double [1]	337
band	integer [1]	1
unit	character [1]	"
names	character [1]	'layer'
legend	S4 (raster:::RasterLegend)	S4 object of class .RasterLegend
title	character [0]	
extent	S4 (raster:::Extent)	S4 object of class Extent
rotated	logical [1]	FALSE
rotation	S4 (raster:::Rotation)	S4 object of class .Rotation
ncols	integer [1]	5832
nrows	integer [1]	6833
crs	S4 (sp::CRS)	S4 object of class CRS
srs	character [1]	'GEOGCRS["unknown",\n DATUM["World
history	list [0]	List of length 0
z	list [0]	List of length 0

Fuente: Propia

#### Datos analizados:

- Cantidad de archivos tiles analizados: 20 archivos en formato ASCII (4 por cada país)
- Peso por archivo: 90 MB aproximadamente.
- Peso total de todos los archivos analizados: 2.5 GB aproximadamente
- Analizados más de 30 millones de datos en formato ASCII – píxeles



Figura 10. Mapa Colombia transformado con reducción de vegetación 2022



Fuente: Propia

Tener una tabla de conversión de píxeles a hectáreas fue crucial porque permitió traducir la información digital en unidades comprensibles, facilitando la cuantificación precisa de las áreas afectadas y mejorando la precisión de los análisis de deforestación.

Tabla 2. Conversión de Píxeles a Hectáreas

Píxeles	Hectáreas
1	6,25
50	312,5
100	625
500	3125
1000	6250
2000	12500
3000	18750
4000	25000
5000	31250
6000	37500

Fuente: Resolución espacial 250m por píxel – Terra-i

### 5.3. ANÁLISIS DE ZONAS AFECTADAS

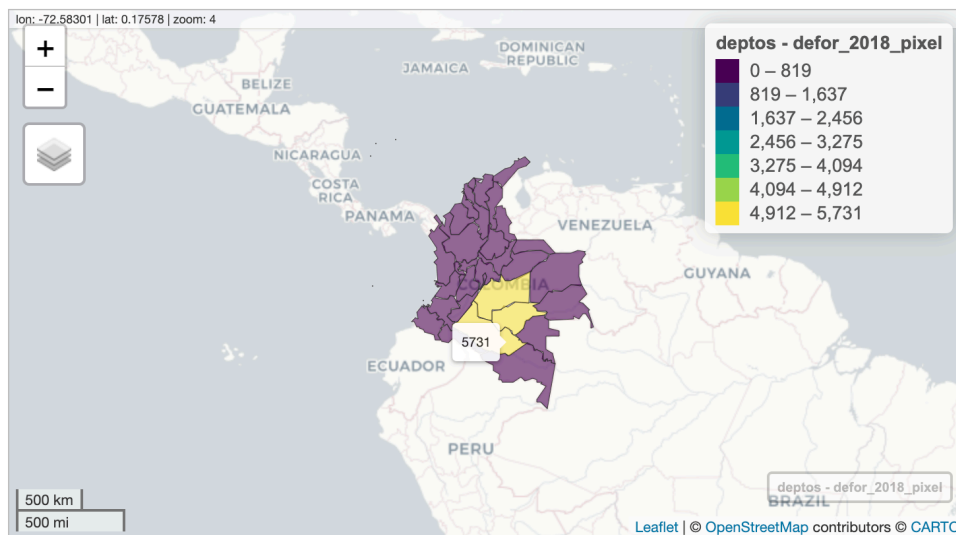
En estos mapas, se pudo analizar el contraste significativo de la deforestación entre los años **2018 y 2022**. Esta evolución se evidenció claramente a través de los patrones de colores que representaron los niveles de deforestación en diferentes regiones de Colombia. Notablemente, los departamentos más afectados durante este periodo fueron **Caquetá, Guaviare y Meta**, donde se observó un incremento alarmante en la pérdida de cobertura forestal.

Los patrones de deforestación mostraron una tendencia preocupante en estos departamentos, lo que los convirtió en áreas prioritarias para este estudio. El análisis de estos mapas fue crucial para entender las dinámicas de la deforestación y para diseñar estrategias efectivas de conservación y recuperación de la cobertura forestal en estas regiones. Los datos se representan mediante una escala de colores, donde los tonos más oscuros indican un mayor índice de deforestación.

#### 5.3.1. Zonas afectadas por la deforestación en Colombia - 2018

El mapa mostró que en el año 2018 el departamento más afectado fue **Caquetá**, que se encontraba en la categoría más alta de deforestación, con un total de **(5731)** píxeles deforestados. Otros departamentos como Meta (5544) y Guaviare (5480) también mostraron niveles preocupantes de pérdida de cobertura forestal.

Figura 11. Mapa: Caquetá - Depto. con mayor pérdida de cobertura forestal en el 2018



Fuente: Propia

### 5.3.2. Zonas afectadas por la deforestación en Colombia – 2019 y 2020

El mapa reveló que en el año 2019 hubo una caída drástica en comparación con el año anterior. El departamento más afectado fue **Guaviare**, situado en la categoría más alta de deforestación, con un total de **(1659)** píxeles deforestados. Otros departamentos como Meta (1626) y Caquetá (1458) también presentaron niveles similares de pérdida de cobertura forestal.

En el año 2020, la tendencia a la baja continuó. El departamento más afectado fue **Meta**, ubicado en la categoría más alta de deforestación, con un total de **(1377)** píxeles deforestados. Los departamentos de Caquetá (1348) y Guaviare (868), también presentaron niveles de pérdida de cobertura forestal.

Figura 12. Mapa: Guaviare - Depto. con mayor pérdida de cobertura forestal en el 2019

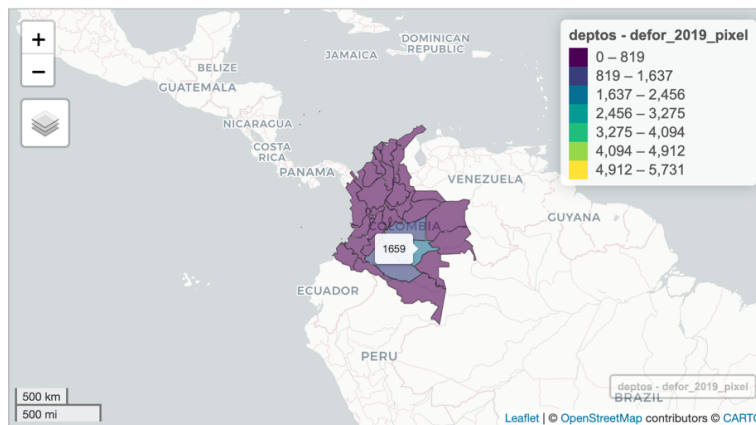
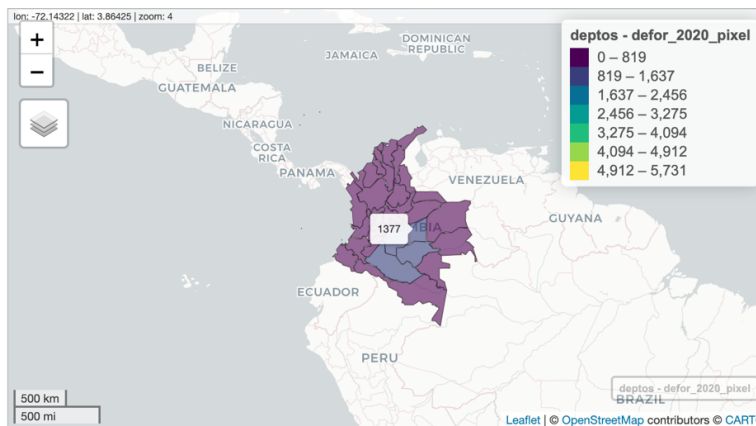


Figura 13. Mapa: Meta - Depto. con mayor pérdida de cobertura forestal en el 2020



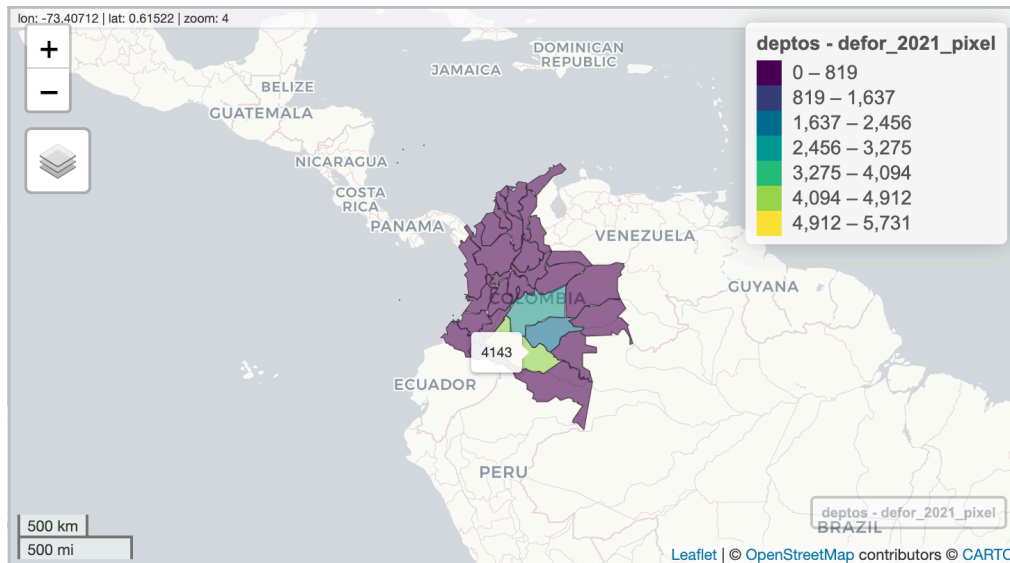
Fuente: Propia

### 5.3.3. Zonas afectadas por la deforestación en Colombia - 2021

El mapa indicó que en el año 2021, el departamento más afectado fue **Caquetá**, con la categoría más alta de deforestación, acumulando un total de **(4143)** píxeles deforestados. Otros departamentos, como Meta (3149) y Guaviare (2456), también exhibieron niveles cambiantes en la pérdida de cobertura forestal.

En el año 2021, se evidenció un crecimiento en la deforestación en todos los departamentos en comparación con los años 2019 y 2020. Esta tendencia al alza podría estar relacionada con las repercusiones de la pandemia del COVID-19, que posiblemente exacerbó las actividades de deforestación debido a la reducción de la vigilancia y el control ambiental durante este período.

Figura 14. Mapa: Caquetá - Depto. con mayor pérdida de cobertura forestal en el 2021



Fuente: Propia

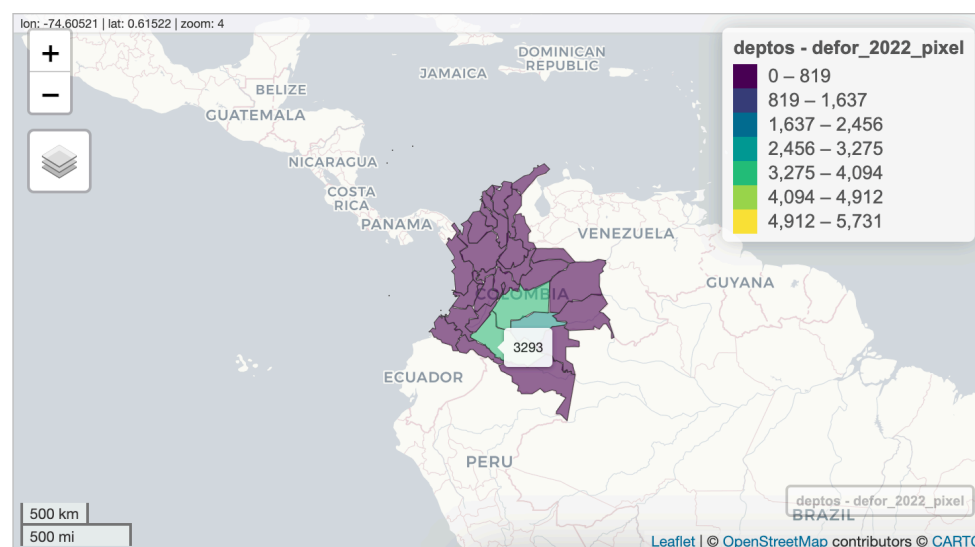
#### 5.3.4. Zonas afectadas por la deforestación en Colombia - 2022

En el año 2022, el análisis reveló que el departamento de **Caquetá** sufrió la mayor deforestación, con **(3293)** píxeles deforestados, manteniéndose en la categoría más alta de pérdida de cobertura forestal. El departamento de Meta le siguió de cerca con 3290 píxeles deforestados, mientras que Guaviare también mostró niveles significativos de deforestación con 2792 píxeles afectados.

Sin embargo, el comportamiento en 2022, el último año de la pandemia, indicó un leve descenso en comparación con el año anterior. Este ligero retroceso sugiere que la salida del estado de emergencia por la pandemia de COVID-19 y la subsiguiente reactivación económica podrían tener un impacto directo en la deforestación de estos tres departamentos.

A medida que las actividades económicas se reanudaron, es probable que las medidas de control y vigilancia ambiental se hayan fortalecido, contribuyendo a esta disminución. No obstante, la relación entre la reactivación económica y la deforestación sigue siendo un tema crucial que requiere un monitoreo constante para mitigar los efectos adversos en el medio ambiente.

Figura 15. Mapa: Caquetá - Depto. con mayor pérdida de cobertura forestal en el 2022



Fuente: Propia

## 6. ANÁLISIS EXPLORATORIO DE DATOS

Para llevar a cabo el análisis exploratorio sobre el comportamiento progresivo de la deforestación, se utilizaron tres momentos clave. En primer lugar, se accedió a un conjunto de datos significativos y relevantes que mostraban la evolución de la deforestación en Colombia. Para este propósito, se empleó el conjunto de datos suministrados por Terra-i. En segundo lugar, se determinó la deforestación antes, durante y después de la pandemia en las diferentes regiones de Colombia objeto de análisis. Y, en tercer lugar, se generaron conclusiones satisfactorias que respondieron al objetivo planteado, con el fin de fomentar la toma de decisiones.

Durante esta fase, se identificaron patrones, tendencias y relaciones dentro de los datos, lo que permitió detectar posibles problemas de calidad o anomalías que podrían haber afectado el modelado posterior. Explorar los datos de manera exhaustiva proporcionó conocimientos significativos que guiaron la selección de características, la elección de modelos y la interpretación de resultados. Además, se emplearon técnicas de visualización de datos para facilitar la comprensión de las variaciones espaciales y temporales en la deforestación. Estas visualizaciones incluyeron gráficos de líneas que destacaron las áreas más afectadas y los periodos críticos de incremento en la tasa de deforestación.

El análisis también involucró la segmentación de los datos en subgrupos basados en variables geográficas y temporales, lo que permitió un estudio más detallado de las causas subyacentes de la deforestación. Esta segmentación reveló diferencias significativas en la dinámica de la deforestación entre las diversas regiones del país. Por ejemplo, se observó que algunas áreas experimentaron una disminución en la tasa de deforestación durante la pandemia, mientras que otras, especialmente aquellas con alta presión por el uso de la tierra, mostraron un aumento. Este análisis detallado fue crucial para hallar conclusiones sobre el comportamiento de la pandemia del COVID-19 en la pérdida de índice de vegetación en los departamentos del estudio.

## 6.1. ANÁLISIS DE LAS SERIES DE TIEMPO

El análisis de las series de tiempo reveló que la deforestación en los departamentos considerados estaba muy elevada, pero mostraba una tendencia a la baja mientras se acercaba el periodo de pandemia. Durante la pandemia, los niveles de deforestación se mantuvieron relativamente bajos, pero comenzaron a aumentar una vez que se declaró el fin del COVID-19 como riesgo sanitario en el país. Para este análisis, se utilizó el calendario juliano para definir las series de tiempo.

De este modo, se pudo concluir que hubo una disminución en la deforestación en Colombia durante los años 2019 y 2020, coincidiendo con los períodos más críticos de la pandemia.

El comportamiento en 2022, como muestra la tabla, indicó un leve descenso en la deforestación en comparación con el año anterior. Esto sugirió que la reactivación económica post-pandemia pudo haber tenido un impacto directo en la reducción de la deforestación en los departamentos afectados. Este fenómeno subrayó la importancia de mantener un equilibrio entre el desarrollo económico y la conservación ambiental para asegurar un desarrollo sostenible.

*Tabla 3. Análisis series de tiempo – Deptos. Deforestación Colombia 2018 - 2022*

<b>Año</b>	<b>Deforestados</b>	<b>Departamento</b>	<b>COVID-19</b>
2018	5731	Caquetá	Antes
2019	1458	Caquetá	Antes
2020	1348	Caquetá	Durante
2021	4143	Caquetá	Durante
2022	3293	Caquetá	Después
2018	5480	Guaviare	Antes
2019	1659	Guaviare	Antes
2020	868	Guaviare	Durante
2021	2456	Guaviare	Durante
2022	2792	Guaviare	Después
2018	5544	Meta	Antes
2019	1626	Meta	Antes
2020	1377	Meta	Durante
2021	3149	Meta	Durante
2022	3290	Meta	Después

*Fuente: Propia*

El uso del calendario juliano en el análisis de las series de tiempo fue importante porque permitió una estandarización precisa de los datos a lo largo del período estudiado. Esta elección facilitó la comparación de eventos y tendencias a través de los años, asegurando que las variaciones temporales observadas fueran atribuibles a cambios reales en los fenómenos analizados, y no a inconsistencias en la medición del tiempo. Además, el calendario juliano, al mantener una secuencia continua de días, evitó las complicaciones derivadas de los ajustes y correcciones que presentan otros calendarios, proporcionando una base temporal uniforme y confiable para el análisis estadístico y la interpretación de resultados.

Figura 16. Código R: Series de tiempo usando calendario Juliano

```
calc_pixeles(inicio = 1,fin = 31) #enero
calc_pixeles(inicio = 32,fin = 59) #febrero
calc_pixeles(inicio = 60,fin = 90) #marzo
calc_pixeles(inicio = 91,fin = 120) #abril
calc_pixeles(inicio = 121,fin = 151) #mayo
calc_pixeles(inicio = 152,fin = 181) #junio
calc_pixeles(inicio = 182,fin = 212) #julio
calc_pixeles(inicio = 213,fin = 243) #agosto
calc_pixeles(inicio = 244,fin = 273) #septiembre
calc_pixeles(inicio = 274,fin = 304) #octubre
calc_pixeles(inicio = 305,fin = 334) #noviembre
calc_pixeles(inicio = 335,fin = 365) #diciembre
```

Fuente: Propia

Figura 17. Código R: Gráfica Análisis Series de Tiempo

```
require(ggplot2)
month=c(sort(rep(1:12,3)),sort(rep(1:12,3)),sort(rep(1:12,3)),sort(rep(1:12,3)),sort(rep(1:12,3)))
year=tabla_series$Año
day=rep(1,180)

fecha=as.Date(paste(year,month,day,sep="-"))
tabla_series$fecha=fecha

g1=ggplot(tabla_series[-c(150:180),],aes(x=fecha,y=Deforestados,colour = Covid))+geom_smooth(method = "lm")+
  facet_grid(~Departamento)+theme_bw()
```

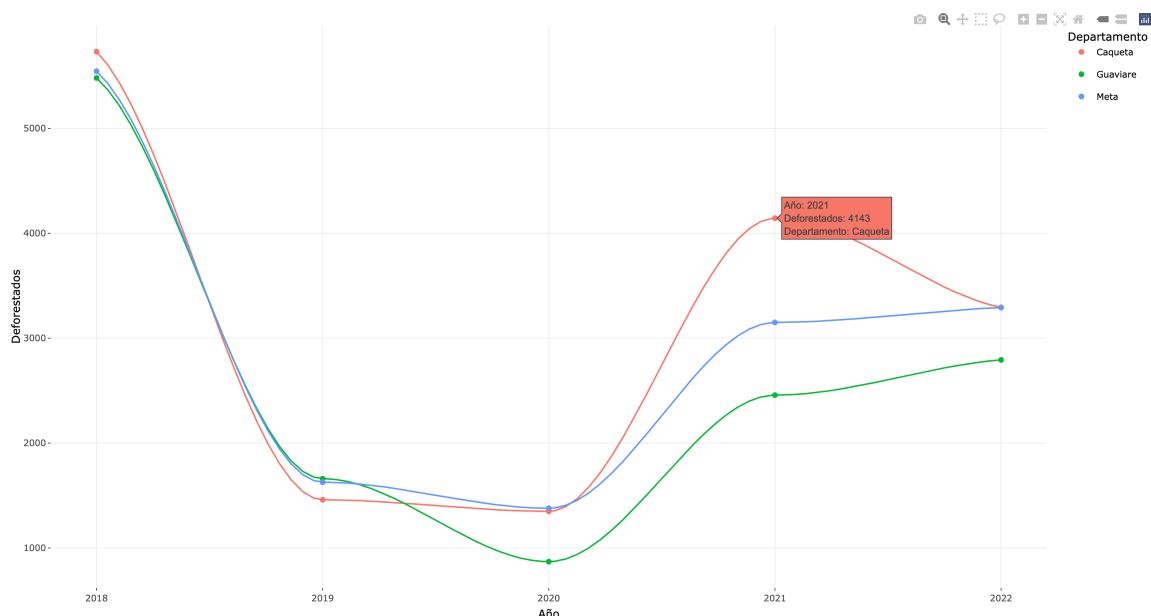
Fuente: Propia

La gráfica muestra una notable disminución en la deforestación en los departamentos de **Caquetá, Meta y Guaviare** desde **2019** hasta **2020**. Durante este periodo, la deforestación alcanzó sus niveles más bajos en 2020, coincidiendo con la pandemia de COVID-19. Esto sugiere que las restricciones y cambios en las actividades humanas durante la pandemia contribuyeron a la reducción de la deforestación.

En **2021**, se observó un aumento significativo en la deforestación en todos los departamentos, especialmente en Caquetá, que registró 4143 píxeles deforestados. Este repunte puede estar relacionado con el levantamiento de las restricciones pandémicas y la reactivación de actividades económicas y sociales, lo que incrementó la presión sobre los bosques.

En **2022**, hubo un leve descenso en la deforestación en comparación con el año anterior. Aunque los niveles se mantuvieron altos, la disminución sugiere que las medidas de control y vigilancia ambiental se reforzaron tras la pandemia. Este patrón destaca la importancia de equilibrar la reactivación económica con la conservación ambiental para asegurar la sostenibilidad de los recursos forestales en Caquetá, Meta y Guaviare.

Figura 18. Análisis series de tiempo - Deptos. Deforestación Colombia 2018 – 2022



Fuente: Propia

## 6.2. ANÁLISIS ESTADÍSTICO

En esta tabla se representaron las áreas deforestadas antes, durante y después de la pandemia en los departamentos de Caquetá, Guaviare y Meta. Los datos evidenciaron que, durante la emergencia sanitaria, los índices de deforestación bajaron significativamente en comparación con los otros periodos considerados. Sin embargo, se observó un incremento notable en la deforestación una vez finalizada la emergencia sanitaria y las cuarentenas.

Antes de la pandemia, los promedios de áreas deforestadas eran bastante elevados en los tres departamentos, con Caquetá registrando un promedio de 305 píxeles deforestados, Guaviare 299, y Meta 304. Durante la pandemia, estos promedios disminuyeron drásticamente: Caquetá a 75, Guaviare a 95.6 y Meta a 104 píxeles deforestados. Esta disminución reflejó el impacto de las restricciones y la reducción de actividades humanas durante ese periodo.

Después de la pandemia, se evidenció un repunte en los índices de deforestación. En Caquetá, el promedio aumentó a 849 píxeles deforestados, en Guaviare a 463, y en Meta a 655. Este aumento podría estar relacionado con la reactivación económica y el retorno a las actividades normales, lo que incrementó la presión sobre los bosques. Estos datos subrayaron la importancia de implementar políticas de conservación más efectivas para prevenir un aumento sostenido de la deforestación en el futuro.

En cuanto a la mediana de las áreas deforestadas, antes de la pandemia, Caquetá tenía una mediana de 147 píxeles, Guaviare 112 y Meta 118. Durante la pandemia, estas cifras disminuyeron significativamente: Caquetá a 71.5, Guaviare a 31 y Meta a 81. Sin embargo, después de la pandemia, la mediana de la deforestación aumentó considerablemente: Caquetá a 207, Guaviare a 227 y Meta a 238. Este comportamiento de la mediana confirmó la tendencia observada en los promedios y resaltó la variabilidad de la deforestación en los diferentes periodos. La mediana, al ser menos sensible a valores extremos, proporcionó una visión adicional de cómo la deforestación afectó de manera consistente a las áreas forestales durante y después de la pandemia.

*Tabla 4. Análisis estadístico – Deptos. Deforestados antes y después COVID-19*

	Caquetá			Guaviare			Meta		
	antes (N=25)	durante (N=18)	despues (N=7)	antes (N=25)	durante (N=18)	despues (N=7)	antes (N=25)	durante (N=18)	despues (N=6)
<b>Deforestados</b>									
Mean (SD)	305 (524)	75.0 (71.7)	849 (1020)	299 (510)	95.6 (161)	463 (612)	304 (510)	104 (105)	655 (762)
Median [Min, Max]	147 [13.0, 2590]	71.5 [0, 316]	207 [85.0, 2520]	112 [1.00, 2240]	31.0 [0, 669]	227 [7.00, 1640]	118 [17.0, 2380]	81.0 [14.0, 418]	238 [92.0, 1750]

*Fuente: Propia*

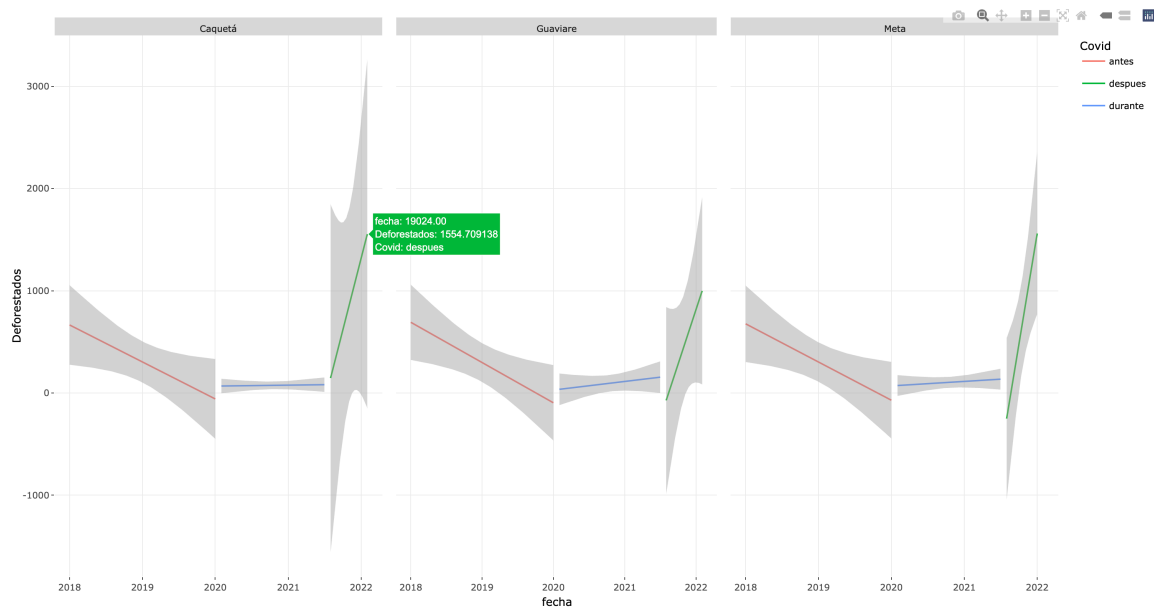
### 6.3. ANÁLISIS DE COMPORTAMIENTO DE LOS DATOS

Se analizó el comportamiento de la deforestación en los periodos antes, durante y después de la pandemia en los departamentos de Caquetá, Guaviare y Meta. Este análisis reveló un incremento significativo en la deforestación post pandemia. Antes de la pandemia, los niveles de deforestación se mantuvieron relativamente bajos y estables en los tres departamentos.

Durante la pandemia, los datos mostraron una ligera disminución o estabilización en la deforestación, reflejando posiblemente las restricciones y la reducción de actividades humanas durante ese tiempo. La gráfica evidencia que, durante este periodo, las líneas correspondientes a cada departamento permanecieron más cercanas al eje horizontal, indicando menores variaciones en los niveles de deforestación.

Después de la pandemia, se observó un notable aumento en la deforestación en todos los departamentos. En 2022, por ejemplo, Caquetá registró 1554 píxeles deforestados en una sola fecha destacada en la gráfica, lo que indica un pico significativo. Guaviare y Meta también mostraron incrementos importantes en las fechas posteriores a la pandemia. Este comportamiento subrayó el impacto de la reactivación económica y el retorno a las actividades normales, incrementando la presión sobre los recursos forestales.

Figura 19. Comportamiento de los datos - Depts. Deforestación Colombia 2018 - 2022



Fuente: Propia

El código en R que se utilizó para analizar la deforestación en los departamentos de Caquetá, Guaviare y Meta, diferenciando los periodos antes, durante y después de la pandemia. Primero, se definió el factor COVID con tres niveles (“antes”, “durante” y “después”). Luego, se ajustó un modelo lineal para cada departamento (mod1, mod2, mod3), evaluando la influencia de los diferentes periodos de la pandemia en la deforestación. Se utilizó la prueba LSD para realizar comparaciones post-hoc y se visualizaron los resultados mediante gráficos de barras para cada departamento, limitando los ejes y mostrando las variaciones en la deforestación a lo largo del tiempo.

Figura 20. Código R: Gráfica de comportamiento de los datos

```
require(table1)
require(plotly)

ggplotly(g1)

tabla_series$Covid=factor(tabla_series$Covid,levels = c("antes", "durante", "despues"))
table1(~Deforestados|Departamento+Covid,data=tabla_series[-c(150:180)],overall = F)

##Modelo por departamento
tabla_series2=tabla_series[-c(150:180),]

##Para Caqueta
pos=tabla_series2$Departamento=="Caquetá"
mod1=lm(Deforestados~Covid,data=tabla_series2[pos,])
summary(mod1)
require(agricolae)
compara1=LSD.test(mod1,"Covid")
compara1

##Para Meta
pos=tabla_series2$Departamento=="Meta"
mod2=lm(Deforestados~Covid,data=tabla_series2[pos,])
summary(mod2)
require(agricolae)
compara2=LSD.test(mod2,"Covid")
compara2
bar.group(compara2$groups,ylim=c(0,800))

##Para Guaviare
pos=tabla_series2$Departamento=="Guaviare"
mod3=lm(Deforestados~Covid,data=tabla_series2[pos,])
summary(mod3)
require(agricolae)
compara3=LSD.test(mod3,"Covid")
compara3
bar.group(compara3$groups,ylim=c(0,600))
```

Fuente: Propia

## **7. MODELO ESTADÍSTICO MIXTO**

El modelo estadístico mixto permitió evaluar el impacto de la deforestación en las zonas afectadas de Colombia utilizando datos históricos de imágenes satelitales de Terra-i desde el 2018-2022. El objetivo del modelo fue integrar múltiples fuentes de información y técnicas estadísticas para comprender mejor la dinámica de la deforestación y sus posibles variaciones durante y después de la pandemia COVID-19.

### **7.1. MODELO ESTADÍSTICO**

Se desarrolló un modelo estadístico mixto en R-Studio para evaluar el impacto de la deforestación en las zonas afectadas de Colombia, utilizando datos históricos de imágenes satelitales de Terra-i desde el 2018-2022. Este modelo permitió integrar múltiples fuentes de información y técnicas estadísticas para comprender mejor la dinámica de la deforestación y sus posibles variaciones durante y después de la pandemia COVID-19. Además, se consideraron los efectos aleatorios y fijos para capturar la heterogeneidad espacial y temporal en los datos, proporcionando una estimación más precisa y robusta de los factores que influyen en la deforestación.

Para validar la efectividad del modelo estadístico mixto, se realizaron múltiples iteraciones de experimentación y ajuste en R-Studio, asegurando que los resultados fueran coherentes y reflejaran adecuadamente la realidad observada en las zonas estudiadas.

### **7.2. ANÁLISIS DESCRIPTIVO**

El análisis descriptivo se centró en proporcionar una comprensión detallada y visual de los datos históricos de deforestación en Colombia durante el período 2018-2022. Este paso es fundamental para identificar patrones, tendencias y anomalías en los datos antes de proceder con el ajuste del modelo estadístico mixto.

Se calcularon estadísticas descriptivas para las variables clave, incluyendo la tasa de deforestación, la extensión de las áreas deforestadas y las características ambientales de las zonas afectadas. Estas estadísticas incluyeron medidas de tendencia central (media y mediana) y de dispersión (desviación estándar y rango intercuartil).

### 7.3. VALIDACIÓN Y EVALUACIÓN DEL MODELO

El modelo desarrollado fue un modelo estadístico explicativo diseñado para examinar la relación entre la deforestación y los distintos periodos antes, durante y después de la pandemia. Este análisis tuvo como objetivo principal determinar si la emergencia sanitaria había tenido un impacto significativo en las tasas de deforestación.

Este modelo se basó en la asignación de niveles de correlación específicos para cada uno de los periodos mencionados. Esto permitía identificar posibles cambios drásticos en los patrones de deforestación que podrían estar relacionados con las medidas de confinamiento, la disminución de actividades industriales y agrícolas, o cualquier otro factor asociado a la pandemia.

#### 7.3.1. Modelo de análisis estadístico de deforestación global

##### **Análisis de Resultados:**

1. Intercepto (Intercept):
  - Estimación: 302.59
  - Error estándar: 53.83
  - Valor t: 5.621
  - Valor p: 9.34e-08 (muy significativo)

Esto indica que, en ausencia de efectos de la pandemia, se espera un valor de deforestación de aproximadamente 302.59 unidades.

2. Codivdurante:
  - Estimación: -210.96
  - Error estándar: 83.20
  - Valor t: -2.535
  - Valor p: 0.01228 (significativo)

Durante la pandemia, se observa una disminución de la deforestación de aproximadamente 210.96 unidades, lo cual es estadísticamente significativo.

3. CovidDespues:
  - Estimación: 353.21
  - Error estándar: 117.32
  - Valor t: 3.011
  - Valor p: 0.00307 (muy significativo)

Después de la pandemia, se observa un incremento en la deforestación de aproximadamente 353.21 unidades, lo cual es estadísticamente significativo.

### Residuos:

- Rango de residuos: Min = -648.80, 1Q = -232.59, Mediana = -74.63, 3Q = 33.41, Max = 2283.41.

Los residuos muestran la diferencia entre los valores observados y los valores predichos por el modelo. Un rango amplio como este sugiere que hay variabilidad considerable en los datos que no está siendo capturada completamente por el modelo.

### Métricas del Modelo:

- Error estándar residual: 466.2 (esto indica la variabilidad de los residuos).
- R-cuadrado múltiple: 0.1307 (indica que el 13.07% de la variabilidad en la deforestación es explicada por las variables independientes).
- R-cuadrado ajustado: 0.1188 (corrige el R-cuadrado múltiple por el número de predictores en el modelo).
- Estadístico F: 10.98
- Grados de libertad: 2 y 146
- Valor p del modelo: 3.612e-05 (muy significativo, indicando que al menos una de las variables independientes tiene un efecto significativo sobre la deforestación).

### Conclusión:

El modelo sugiere que la pandemia tuvo un efecto significativo sobre la deforestación. Durante la pandemia, la deforestación disminuyó significativamente, mientras que después de la pandemia, la deforestación aumentó significativamente. Aunque el modelo es significativo, el R-cuadrado relativamente bajo indica que hay otros factores importantes que no están siendo capturados por este modelo.

*Figura 21. Modelo de análisis estadístico Global*

```
> summary(mod_all)

Call:
lm(formula = Deforestados ~ Covid, data = tabla_series2)

Residuals:
    Min       1Q   Median       3Q      Max
-648.80 -232.59  -74.63   33.41 2283.41

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   302.59     53.83   5.621 9.34e-08 ***
Covid durante -210.96     83.20  -2.535  0.01228 *
Covid despues  353.21    117.32   3.011  0.00307 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 466.2 on 146 degrees of freedom
Multiple R-squared:  0.1307,    Adjusted R-squared:  0.1188
F-statistic: 10.98 on 2 and 146 DF,  p-value: 3.612e-05
```

*Fuente: Propia*

### 7.3.2. Modelo de análisis estadístico de deforestación en el Caqueta

#### Análisis de Resultados:

1. Intercepto (Intercept):
  - Estimación: 304.7
  - Error estándar: 104.9
  - Valor t: 2.905
  - Valor p: 0.00559 (muy significativo)

Esto indica que, en ausencia de efectos de la pandemia, se esperaba un valor de deforestación de aproximadamente 304.7 unidades.

2. Codivdurante:
  - Estimación: -229.7
  - Error estándar: 162.2
  - Valor t: -1.417
  - Valor p: 0.16317 (no significativo)

Durante la pandemia, se observa una disminución de la deforestación de aproximadamente 229.7 unidades, pero este resultado no es estadísticamente significativo.

3. CovidDespues:
  - Estimación: 544.7
  - Error estándar: 224.3
  - Valor t: 2.428
  - Valor p: 0.01905 (significativo)

Después de la pandemia, se observa un incremento en la deforestación de aproximadamente 544.7 unidades, lo cual es estadísticamente significativo.

#### Residuos:

- Rango de residuos: Min = -764.43, 1Q = -240.47, Mediana = -58.36, 3Q = 26.00, Max = 2281.28.

Los residuos muestran la diferencia entre los valores observados y los valores predichos por el modelo. Un rango amplio sugiere que hay variabilidad considerable en los datos que no está siendo capturada completamente por el modelo.

### Métricas del Modelo:

- Error estándar residual: 524.6 (esto indica la variabilidad de los residuos).
- R-cuadrado múltiple: 0.1896 (indica que el 18.96% de la variabilidad en la deforestación es explicada por las variables independientes).
- R-cuadrado ajustado: 0.1551 (corrige el R-cuadrado múltiple por el número de predictores en el modelo).
- Estadístico F: 5.496
- Grados de libertad: 2 y 47
- Valor p del modelo: 0.007162 (muy significativo, indicando que al menos una de las variables independientes tiene un efecto)

### Conclusión:

El modelo sugiere que la pandemia tuvo diferentes efectos sobre la deforestación en Caquetá. Durante la pandemia, aunque se observa una disminución en la deforestación, este efecto no es estadísticamente significativo. Sin embargo, después de la pandemia, la deforestación aumenta significativamente. Aunque el modelo es significativo, el R-cuadrado relativamente bajo indica que hay otros factores importantes que no están siendo capturados por este modelo.

Figura 22. Modelo de análisis estadístico del Caquetá

```
> summary(mod1)

Call:
lm(formula = Deforestados ~ Covid, data = tabla_series2[pos,
  ])

Residuals:
    Min       1Q   Median       3Q      Max
-764.43 -240.47  -58.36   26.00 2281.28

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    304.7      104.9    2.905  0.00559 **
Coviddurante  -229.7      162.2   -1.417  0.16317
Coviddespues   544.7      224.3    2.428  0.01905 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 524.6 on 47 degrees of freedom
Multiple R-squared:  0.1896,    Adjusted R-squared:  0.1551
F-statistic: 5.496 on 2 and 47 DF,  p-value: 0.007162
```

Fuente: Propia

### 7.3.3. Modelo de análisis estadístico de deforestación en el Meta

#### **Análisis de Resultados:**

1. Intercepto (Intercept):
  - Estimación: 304.04
  - Error estándar: 90.12
  - Valor t: 3.374
  - Valor p: 0.00151 (muy significativo)

Esto indica que, en ausencia de efectos de la pandemia, se esperaba un valor de deforestación de aproximadamente 304.04 unidades.

2. Codivdurante:
  - Estimación: -199.71
  - Error estándar: 139.29
  - Valor t: -1.434
  - Valor p: 0.15840 (no significativo)

Durante la pandemia, se observa una disminución de la deforestación de aproximadamente 199.71 unidades, pero este resultado no es estadísticamente significativo.

3. CovidDespues:
  - Estimación: 351.13
  - Error estándar: 204.84
  - Valor t: 1.714
  - Valor p: 0.09324 (marginamente significativo)

Después de la pandemia, se observa un incremento en la deforestación de aproximadamente 351.13 unidades, lo cual es marginalmente significativo.

#### **Residuos:**

- Rango de residuos: Min = -563.17, 1Q = -233.04, Mediana = -88.33, 3Q = 31.67, Max = 2071.96.

Los residuos muestran la diferencia entre los valores observados y los valores predichos por el modelo. Un rango amplio sugiere que hay variabilidad considerable en los datos que no está siendo capturada completamente por el modelo

### Métricas del Modelo:

- Error estándar residual: 450.6 (esto indica la variabilidad de los residuos).
- R-cuadrado múltiple: 0.1314 (indica que el 13.14% de la variabilidad en la deforestación es explicada por las variables independientes).
- R-cuadrado ajustado: 0.0936 (corrige el R-cuadrado múltiple por el número de predictores en el modelo).
- Estadístico F: 3.478
- Grados de libertad: 2 y 46
- Valor p del modelo: 0.0392 (significativo, indicando que al menos una de las variables independientes tiene un efecto significativo sobre la deforestación).

### Conclusión:

El modelo sugiere que la pandemia tuvo diferentes efectos sobre la deforestación en Meta. Durante la pandemia, aunque se observa una disminución en la deforestación, este efecto no es estadísticamente significativo. Después de la pandemia, la deforestación aumenta y este efecto es marginalmente significativo. Aunque el modelo es significativo, el R-cuadrado relativamente bajo indica que hay otros factores importantes que no están siendo capturados por este modelo.

Figura 23. Modelo de análisis estadístico del Meta

```
> summary(mod2)

Call:
lm(formula = Deforestados ~ Covid, data = tabla_series2[pos,
])

Residuals:
    Min       1Q   Median       3Q      Max
-563.17 -233.04  -88.33   31.67 2071.96

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   304.04     90.12   3.374  0.00151 **
Covid durante -199.71    139.29  -1.434  0.15840
Covid despues  351.13    204.84   1.714  0.09324 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 450.6 on 46 degrees of freedom
Multiple R-squared:  0.1314,    Adjusted R-squared:  0.0936
F-statistic: 3.478 on 2 and 46 DF,  p-value: 0.0392
```

Fuente: Propia

### 7.3.4. Modelo de análisis estadístico del Guaviare

#### Análisis de Resultados:

1. Intercepto (Intercept):
  - Estimación: 299.00
  - Error estándar: 87.16
  - Valor t: 3.430
  - Valor p: 0.00127 (muy significativo)

Esto indica que, en ausencia de efectos de la pandemia, se esperaba un valor de deforestación de aproximadamente 299.00 unidades.

2. Codivdurante:
  - Estimación: -203.44
  - Error estándar: 134.72
  - Valor t: -1.510
  - Valor p: 0.13770 (no significativo)

Durante la pandemia, se observa una disminución de la deforestación de aproximadamente 203.44 unidades, pero este resultado no es estadísticamente significativo.

3. CovidDespues:
  - Estimación: 163.71
  - Error estándar: 186.36
  - Valor t: 0.878
  - Valor p: 0.38415 (no significativo)

Después de la pandemia, se observa un incremento en la deforestación de aproximadamente 163.71 unidades, pero este resultado no es estadísticamente significativo.

#### Residuos:

- Rango de residuos: Min = -455.71, 1Q = -226.50, Mediana = -74.56, 3Q = 22.36, Max = 1938.00.

Los residuos muestran la diferencia entre los valores observados y los valores predichos por el modelo. Un rango amplio sugiere que hay variabilidad considerable en los datos que no está siendo capturada completamente por el modelo.

### Métricas del Modelo:

- Error estándar residual: 435.8 (esto indica la variabilidad de los residuos).
- R-cuadrado múltiple: 0.08282 (indica que el 8.282% de la variabilidad en la deforestación es explicada por las variables independientes).
- R-cuadrado ajustado: 0.04379 (corrige el R-cuadrado múltiple por el número de predictores en el modelo).
- Estadístico F: 2.122
- Grados de libertad: 2 y 47
- Valor p del modelo: 0.1311 (no significativo, indicando que el modelo en su conjunto no tiene un efecto significativo sobre la deforestación).

### Conclusión:

El modelo sugiere que la pandemia no tuvo efectos estadísticamente significativos sobre la deforestación en Guaviare. Tanto durante como después de la pandemia, los cambios en la deforestación no son significativos. Además, el bajo R-cuadrado indica que hay muchos otros factores que no están siendo capturados por este modelo y que podrían estar influyendo en la deforestación. El modelo en su conjunto no es significativo, lo que sugiere que las variables incluidas no explican bien las variaciones en la deforestación en esta región.

*Figura 24. Modelo de análisis estadístico del Guaviare*

```
> summary(mod3)

Call:
lm(formula = Deforestados ~ Covid, data = tabla_series2[pos,
  ])

Residuals:
    Min       1Q   Median       3Q      Max
-455.71 -226.50  -74.56   22.36 1938.00

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   299.00     87.16   3.430  0.00127 **
Covid durante -203.44    134.72  -1.510  0.13770
Covid despues  163.71    186.36   0.878  0.38415
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 435.8 on 47 degrees of freedom
Multiple R-squared:  0.08282, Adjusted R-squared:  0.04379
F-statistic: 2.122 on 2 and 47 DF, p-value: 0.1311
```

*Fuente: Propia*

### 7.3.5. Evaluación de Correlación del Modelo

A continuación, se presenta un diagrama para evaluar la fortaleza del modelo. Este análisis se basa en la correlación de los niveles: si los niveles de correlación son similares, se concluye que el impacto del COVID-19 en la deforestación no es significativo. Sin embargo, si los niveles de correlación son diferentes, el impacto se juzga en función de la asignación jerárquica **a**, **b** y **c**.

Figura 25. Evaluación de correlación del modelo estadístico

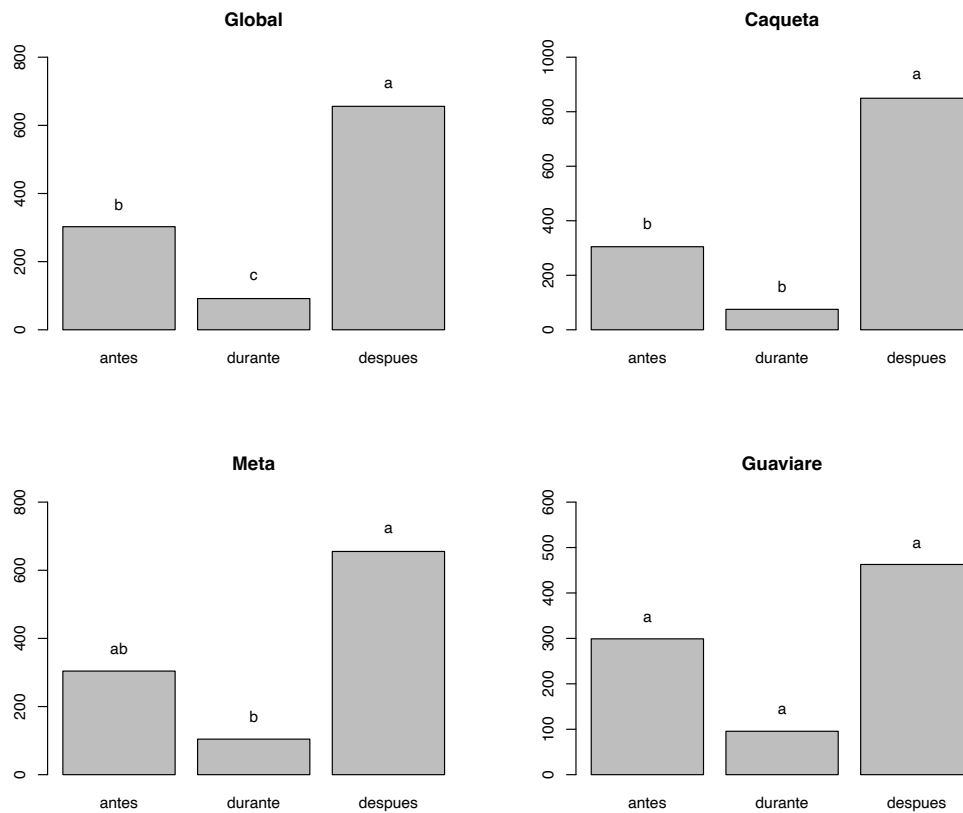


Figura 26. Código R: Gráfica de correlación

```
par(mfrow=c(2,2))
bar.group(compara_all$groups[c(2,3,1)],ylim=c(0,800),main="Global")
bar.group(compara1$groups[c(2,3,1)],ylim=c(0,1000),main="Caqueta")
bar.group(compara2$groups[c(2,3,1)],ylim=c(0,800),main="Meta")
bar.group(compara3$groups[c(2,3,1)],ylim=c(0,600),main="Guaviare")
```

Fuente: Propia

### **Análisis de Correlación del Modelo:**

Al observar los diagramas, podemos llegar a las siguientes conclusiones sobre el impacto de la pandemia de COVID-19 en la deforestación en las distintas regiones:

#### **Global:**

La pandemia tuvo un efecto significativo en la deforestación a nivel global. Durante el periodo de la pandemia, se registró el número más bajo de áreas deforestadas. Sin embargo, después de la emergencia sanitaria, la deforestación aumentó drásticamente, alcanzando los niveles más altos de todo el periodo analizado.

#### **Caquetá:**

En Caquetá, la pandemia también tuvo un impacto considerable en la deforestación. El mínimo número de áreas deforestadas se registró durante la pandemia. Sin embargo, al igual que a nivel global, la deforestación aumentó de manera significativa una vez que la emergencia sanitaria terminó, alcanzando niveles muy altos.

#### **Meta:**

En Meta, aunque hubo un cambio en la deforestación influenciado por la pandemia, los niveles de correlación entre antes y durante la pandemia no son tan diferentes. Sin embargo, en el periodo posterior a la pandemia, la deforestación aumentó significativamente, indicando un fuerte impacto en esta fase.

#### **Guaviare:**

En Guaviare, el efecto de la pandemia en la deforestación fue menos significativo. Aunque se observó una disminución en las áreas deforestadas durante la pandemia, este cambio no fue tan marcado en comparación con las otras regiones. Después de la pandemia, la deforestación aumentó, pero no alcanzó niveles tan altos como en Caquetá o Meta.

### **Conclusión General:**

En general, los datos sugieren que la pandemia de COVID-19 tuvo un efecto significativo en la deforestación, con una notable disminución durante la pandemia y un aumento posterior. Las variaciones regionales indican que el impacto no fue uniforme, siendo más pronunciado en Caquetá y Meta que en Guaviare. Este análisis subraya la importancia de considerar factores específicos de cada región al evaluar el impacto de eventos globales como la pandemia en la deforestación.

## 8. CONCLUSIONES Y TRABAJOS FUTUROS

### 8.1. CONCLUSIONES

El análisis realizado en este proyecto sobre la deforestación en Colombia antes, durante y después de la pandemia de COVID-19 ha proporcionado varias conclusiones significativas. En primer lugar, se observó que los índices de deforestación disminuyeron notablemente durante los períodos más críticos de la pandemia, coincidiendo con las restricciones de movilidad y la reducción de actividades económicas en áreas rurales. Este descenso sugiere que las limitaciones impuestas a las actividades humanas tuvieron un efecto positivo en la disminución de la deforestación a corto plazo.

Sin embargo, una vez que las restricciones fueron levantadas, hubo un repunte significativo en la deforestación, especialmente en los departamentos de Caquetá, Meta y Guaviare. Este incremento pudo estar asociado con la reactivación económica y el retorno a las actividades socioeconómicas, lo que incrementó la presión sobre los bosques. Los resultados también indican que la pandemia no tuvo un efecto uniforme en todas las regiones, siendo más pronunciado en algunas áreas que en otras.

El departamento de Caquetá fue identificado como el más afectado por la deforestación en el rango de tiempo analizado, especialmente después de la pandemia. Este departamento mostró un aumento significativo en la pérdida de cobertura forestal, lo que subraya la necesidad de enfoques específicos de conservación y gestión en esta región.

Además, los fenómenos climáticos de El Niño y La Niña también pudieron haber impactado la deforestación en estos departamentos durante el período analizado. Estos fenómenos afectan los ciclos de lluvia y sequía. Por ejemplo, El Niño tiende a causar sequías que pueden aumentar el riesgo de incendios forestales, mientras que La Niña puede incrementar la precipitación, afectando la accesibilidad y las actividades de tala. Durante el período analizado el fenómeno más común fue La Niña que generó mayor precipitación climática que posiblemente fomentó el incremento de cobertura vegetal.

Estos hallazgos subrayan la importancia de implementar políticas de conservación y gestión ambiental que puedan sostenerse incluso en tiempos de crisis. La variabilidad en la deforestación durante diferentes períodos también resalta la necesidad de un monitoreo continuo y detallado para responder rápidamente a los cambios en las actividades humanas y los factores climáticos que impactan el medio ambiente.

## 8.2. TRABAJOS FUTUROS

El proyecto “Evaluación del Efecto de la Pandemia sobre la Deforestación en Colombia por Medio de Imágenes de Satélite y Usando los Datos de Terra-I” abre varias vías para futuros trabajos de investigación en los campos ambiental y socioeconómico. A continuación, se presentan algunas áreas de estudio que podrían beneficiarse de los hallazgos y metodologías desarrolladas en este proyecto:

- **Relación entre Deforestación y Factores Socioeconómicos:** Futuras investigaciones podrían explorar cómo la pobreza, la migración y el acceso a recursos económicos influyen en la deforestación en Caquetá, Guaviare y Meta, especialmente durante y después de la pandemia.
- **Impacto de Políticas de Reforestación y Conservación:** Evaluar la efectividad de las políticas de reforestación y conservación post-pandemia será crucial para mejorar las estrategias de gestión forestal a largo plazo.
- **Desarrollo de Modelos Predictivos Mejorados:** Futuros trabajos podrían desarrollar modelos predictivos avanzados, incorporando más variables ambientales y socioeconómicas para prever áreas en riesgo de deforestación y diseñar intervenciones efectivas.
- **Evaluación del Impacto Ambiental a Largo Plazo:** Investigar los impactos ambientales a largo plazo en la región, incluyendo la resiliencia y capacidad de recuperación de los ecosistemas forestales ante futuras crisis.
- **Integración con Otros Estudios Regionales y Globales:** Comparar este estudio con investigaciones similares en otras regiones para entender cómo diferentes contextos afectan la deforestación, facilitando la formulación de políticas globales de conservación y desarrollo sostenible.

## 9. REFERENCIAS BIBLIOGRÁFICAS

- [1] Humboldt, «Bosques para las personas,» 11 Junio 2023. [En línea]. Available: <http://repository.humboldt.org.co/bitstream/handle/20.500.11761/31369/230.pdf?sequence=1&isAllowed=y#:~:text=Los%20bosques%20contienen%20el%2080.>
- [2] Cepal.org, «La pérdida de los bosques de América Latina y el Caribe 1990-2020,» [En línea]. Available: [https://repositorio.cepal.org/bitstream/handle/11362/47151/1/S2100266\\_es.pdf](https://repositorio.cepal.org/bitstream/handle/11362/47151/1/S2100266_es.pdf).
- [3] S. T. Amazon, «Las alarmantes cifras de deforestación de la pandemia en la Amazonía,» 15 Julio 2021. [En línea]. Available: <https://savingtheamazon.org/blogs/news/las-alarman-tes-cifras-de-deforestacion-de-la-pandemia-en-la-amazonia>.
- [4] C. C. Petersen, «Universidad Nacional Autónoma de México Facultad de Ciencias Políticas y Sociales la Amazonía Brasileña: La Degradación Ambiental y los Desafíos que enfrenta la Región,» 2021. [En línea]. Available: [https://ru.dgb.unam.mx/bitstream/DGB\\_UNAM/TES01000811052/3/0811052.pdf](https://ru.dgb.unam.mx/bitstream/DGB_UNAM/TES01000811052/3/0811052.pdf).
- [5] H. Romero, «Deforestación en Colombia: Retos y perspectivas,» [En línea]. Available: [https://www.repository.fedesarrollo.org.co/bitstream/handle/11445/337/KAS%20SOPLA\\_Deforestacion%20en%20Colombia%20retos%20y%20perspectivas.pdf?sequence=2&isAllowed=y](https://www.repository.fedesarrollo.org.co/bitstream/handle/11445/337/KAS%20SOPLA_Deforestacion%20en%20Colombia%20retos%20y%20perspectivas.pdf?sequence=2&isAllowed=y).
- [6] M. d. A. y. D. Sostenible, «Se reduce y se contiene la deforestación en Colombia durante los últimos cuatro años,» 15 Julio 2022. [En línea]. Available: <https://www.minambiente.gov.co/bosques-biodiversidad-y-servicios-ecosistemicos/se-reduce-y-se-contiene-la-deforestacion-en-colombia-durante-los-ultimos-cuatro-anos/>.
- [7] Wwf.org.co, «Aumento de la deforestación en Colombia en 2022 exige acciones urgentes,» [En línea]. Available: <https://www.wwf.org.co/?378752/Aumento-de-la-deforestacion-en-Colombia-en-2022-exige-acciones-urgentes>.
- [8] G. R. a. D. C. C. Delrieux, «Visualización y procesamiento de imágenes satelitales,» 2001. [En línea]. Available: <http://sedici.unlp.edu.ar/handle/10915/21753>.
- [9] NASA, «Clasificación supervisada de imágenes,» 2017. [En línea]. Available: <https://appliedsciences.nasa.gov/sites/default/files/Exercise%201%20->

- %20Creating%20a%20Supervised%20Land%20Cover%20Classification%20%28Spanish%209.pdf.
- [10] A. UTPL, «Clasificación supervisada y no supervisada,» 14 Abril 2008. [En línea]. Available: <https://advancedtech.wordpress.com/2008/04/14/clasificacion-supervisada-y-no-supervisada/>.
- [11] J. D. V. R. y. F. M. Romon, «Clasificación de imágenes espaciales mediante visión artificial,» 2022. [En línea]. Available: [https://oa.upm.es/71155/1/TFG\\_ISABEL\\_ROMON\\_JACOB.pdf](https://oa.upm.es/71155/1/TFG_ISABEL_ROMON_JACOB.pdf).
- [12] F. camp, «Cómo funcionan los clasificadores Naive Bayes,» 28 Abril 2021. [En línea]. Available: <https://www.freecodecamp.org/espanol/news/como-funcionan-los-clasificadores-naive-bayes-con-ejemplos-de-codigo-de-python/>.
- [13] R. Díaz, «Métricas de Clasificación,» [En línea]. Available: <https://www.themachinelearners.com/metricas-de-clasificacion/>.
- [14] M. P. y. G. Casimiro, «Análisis de series temporales,» [En línea]. Available: <https://addi.ehu.es/bitstream/handle/10810/12492/04-09gon.pdf>.
- [15] IBM, «¿Qué son los datos geoespaciales?,» [En línea]. Available: <https://www.ibm.com/mx-es/topics/geospatial-data>.
- [16] Eos.com, «NDVI (Índice Diferencial Normalizado De Vegetación),» 27 Septiembre 2021. [En línea]. Available: <https://eos.com/es/make-an-analysis/ndvi/>.
- [17] NASA, «MODIS (Moderate Resolution Imaging Spectroradiometer),» [En línea]. Available: <https://modis.gsfc.nasa.gov/data/>.
- [18] R. H. F. C. y. B. Sampieri, «Metodología de la investigación,» 2014. [En línea]. Available: <https://highered.mheducation.com/sites/dl/free/1456260960/1106855/cap08sampieri.pdf>.
- [19] H. University., «Cross validation,» [En línea]. Available: <https://rafalab.dfci.harvard.edu/dsbook/cross-validation.html>.
- [20] N. S. Chauhan, «Métricas de evaluación de Modelos en el aprendizaje automático,» 25 Septiembre 2023. [En línea]. Available: <https://www.datasource.ai/es/data-science-articles/metricas-de-evaluacion-de-modelos-en-el-aprendizaje-automatico>.

- [21] J. J. B. Bejarano, «Evaluación de modelos: Métricas y técnicas para medir el rendimiento,» [En línea]. Available: <https://universo-ia.com/evaluacion-de-modelos-metricas-y-tecnicas-para-medir-el-rendimiento/>.
- [22] R. Casal, «Construcción y evaluación de los modelos,» [En línea]. Available: [https://rubenfcasal.github.io/aprendizaje\\_estadistico/const-eval.html](https://rubenfcasal.github.io/aprendizaje_estadistico/const-eval.html).
- [23] P. Uno, «Consideraciones al elegir las métricas de evaluación,» [En línea]. Available: <https://prompt.uno/aprendizaje-automatizado/metricas-de-evaluacion-de-modelos/>.
- [24] M. y. E. A. DEAM - Instituto de Hidrología, «Monitoreo y Seguimiento al Fenómeno de la Deforestación en Colombia,» [En línea]. Available: <http://www.ideam.gov.co/web/bosques/deforestacion-colombia>.
- [25] J. M. Silva, «Cambio climático en la región amazónica,» 2019. [En línea]. Available: <http://otca.org/wp-content/uploads/2021/02/El-Cambio-Climatico-en-la-Region-Amazonica-Acciones-de-la-OTCA-1.pdf>.
- [26] L. G. S. & R. Rodríguez, «Análisis político y económico de la deforestación en regiones afectadas por el conflicto en Colombia,» 2020. [En línea]. Available: <https://foronacionalambiental.org.co/wp-content/uploads/2020/05/AnalisisDeforestacionKPMG.pdf>.
- [27] A. L. J. & M. R. García, «El cambio climático y sus efectos en la biodiversidad en América Latina,» 2018. [En línea]. Available: [https://repositorio.cepal.org/bitstream/handle/11362/39855/S1501295\\_en.pdf?sequence=1](https://repositorio.cepal.org/bitstream/handle/11362/39855/S1501295_en.pdf?sequence=1).
- [28] IBM, «Conceptos básicos de ayuda de CRISP-DM,» [En línea]. Available: <https://www.ibm.com/docs/es/spss-modeler/SaaS?topic=dm-crisp-help-overview>.
- [29] Terra-i, «Terra-i,» [En línea]. Available: <http://www.terra-i.org/terra-i.html>.
- [30] Terra-i, «Metodología,» [En línea]. Available: <http://www.terra-i.org/es/terra-i/data/description.html>.
- [31] Terra-i, «Datos,» [En línea]. Available: <http://www.terra-i.org/es/terra-i/data.html>.

- [32] R. Spatial, «Raster package,» [En línea]. Available: <https://rspatial.org/raster/pkg/index.html>.
- [33] RDocumentation, «Rgdal package,» [En línea]. Available: <https://www.rdocumentation.org/packages/rgdal/versions/1.6-7>.
- [34] C.-R. Project, «Rnaturalearth package,» [En línea]. Available: <https://cran.r-project.org/web/packages/rnaturalearth/vignettes/rnaturalearth.html>.
- [35] Terra-i, «Dataset,» [En línea]. Available: <http://www.terra-i.org/terra-i/data.html>.
- [36] Minambiente, «Incendio forestal en Guaviare está liquidado y no se registran incendios en Caquetá,» [En línea]. Available: <https://archivo.minambiente.gov.co/index.php/noticias-minambiente/3636-incendio-forestal-en-guaviare-esta-liquidado-y-no-se-registran-incendios-en-caqueta#:~:text=Marzo%203%20de%202018>.
- [37] C. P. Center, «Cold & Warm Episodes by Season,» [En línea]. Available: [https://origin.cpc.ncep.noaa.gov/products/analysis\\_monitoring/ensostuff/ONI\\_v5.php](https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php).