**PRONÓSTICO DE DEFORESTACIÓN EN COLOMBIA PARA EL PERÍODO COMPRENDIDO ENTRE 2024 A 2032**

Andrea Catalina Burbano Buitrago[[1]](#footnote-1)

Ángel Santiago Castiblanco Falla[[2]](#footnote-2)

Luis Fernando Peña Pineda[[3]](#footnote-3)

Diego Alexander Polo Contreras[[4]](#footnote-4)

Juan David Ruiz Bedoya[[5]](#footnote-5)

Santiago Palacio Soto [[6]](#footnote-6)

Proyecto presentado como requisito parcial para culminar el Bootcamp de Analítica de datos

Docente: Ronald Akerman Ortiz García

Talento Tech

2024

**TABLA DE CONTENIDO**

[**INTRODUCCIÓN** 3](#_Toc181215085)

[**OBJETIVO GENERAL** 4](#_Toc181215086)

[**METODOLOGÍA** 5](#_Toc181215087)

[**Estructura general:** 5](#_Toc181215088)

[**1.** **Descripción de las tablas:** 5](#_Toc181215089)

[**2.** **Relaciones entre tablas:** 7](#_Toc181215090)

[**3.** **Observaciones:** 7](#_Toc181215091)

[**Análisis de datos** 8](#_Toc181215092)

[**RESULTADOS** 9](#_Toc181215093)

[**1.** **Resultados Iniciales:** 9](#_Toc181215094)

[**2.** **Modelo A.R.I.M.A:** 13](#_Toc181215095)

[**3.** **Resultados Finales:** 15](#_Toc181215096)

[**4.** **Metodología de reforestación propuesta:** 16](#_Toc181215097)

[**CONCLUSIONES** 17](#_Toc181215098)

[**RECOMENDACIONES** 18](#_Toc181215099)

[**BIBLIOGRAFÍA** 19](#_Toc181215100)

**ÍNDICE DE ILUSTRACIONES**

[**Ilustración 1. Deforestación por año en Colombia (2001-2023)** 9](#_Toc181215101)

[**Ilustración 2. Cobertura arbórea Colombia por año (2001-2023)** 10](#_Toc181215102)

[**Ilustración 3. Deforestación por departamento (2001-2023)** 10](#_Toc181215103)

[**Ilustración 4. Deforestación por causa (2001-2023)** 11](#_Toc181215104)

[**Ilustración 5. Matriz de correlación de causas de deforestación en Colombia** 12](#_Toc181215105)

[**Ilustración 6. Principales medidas estadísticas obtenidas a partir del modelo A.R.I.M.A** 13](#_Toc181215106)

[**Ilustración 7. Autocorrelación y autocorrelación parcial del modelo predictivo** 14](#_Toc181215107)

[**Ilustración 8. Proyección de la Deforestación en Colombia para el año 2032** 15](#_Toc181215108)

**ÍNDICE DE TABLAS**

[**Tabla 1. Pronóstico de pérdida de cobertura arbórea en el país para los períodos de 2024 a 2032** 8](#_Toc181215109)

[**Tabla 2. Representación de pérdida media de hectáreas forestales cada 5 años** 14](#_Toc181215110)

# **INTRODUCCIÓN**

La deforestación en Colombia ha sido un problema ambiental significativo durante las últimas décadas, causada principalmente por factores como la tala de bosques, la minería ilegal, los cultivos ilícitos, los incendios forestales y la construcción de nueva infraestructura en el país, lo que ha generado problemas para la conservación de la biodiversidad y las comunidades que habitan en estos ecosistemas.

El Ministerio de Ambiente de Colombia estimó que para el año 2023 la deforestación bajó entre un 25% y un 35% a nivel nacional comparado con el año 2022. Sin embargo, Según Susana Muhamad (2024), ministra de ambiente, informó que entre enero y marzo de 2024 Colombia presentó un pico histórico en términos de deforestación, siendo aproximadamente de 109.000 hectáreas de bosque perdido, Presentando un incremento del 40% a comparación del primer trimestre del año 2023, razón por la cual resulta importante pronosticar cual será la pérdida arbórea de los próximos 10 años ya que los bosques desempeñan un papel crucial para la conservación del suelo, la regulación del clima, la provisión de oxígeno la conservación de hábitat de especies naturales entre otras.

Mediante el siguiente proyecto se realizará un análisis de datos tomando como referencia el histórico de deforestación entre el año 2001-2023 a nivel país, de forma que se puedan identificar patrones, analizar tendencias y con el fin de proyectar la pérdida de cobertura arbórea de los próximos 10 años mediante un modelo ARIMA. Así mismo se presentará una metodología de reforestación efectiva que pueda ser utilizada en Colombia basados en los resultados obtenidos de forma que ayude a mitigar el problema y así contribuir a salvar el país.

# **OBJETIVO GENERAL**

Aplicar técnicas de analítica de datos para pronosticar la pérdida arbórea en Colombia tomando como referencia datos históricos del 2001 al 2023 para la siguiente década.

**Objetivos Específicos**

1. Analizar los datos y presentar gráficamente el comportamiento de la pérdida de cobertura arbórea durante los años 2001 al 2023
2. Pronosticar mediante la metodología ARIMA la pérdida de cobertura arbórea de los próximos 10 años.
3. Proponer una metodología de reforestación efectiva para ser implementada a nivel país

# **METODOLOGÍA**

El siguiente estudio proyectó la deforestación en Colombia de los próximos 10 años tomando como referencia los datos históricos comprendidos entre los períodos 2001 a 2023, la información fue obtenida del web site Global Forest Watch en su publicación titulada “Colombia Deforestation Rates & Statistics”.

Luego de identificar la fuente de información se procedió a realizar la extracción, transformación y limpieza de los datos para poder ser analizados. Posteriormente, se crearon las tablas para poder realizar la información mediante el uso del lenguaje SQL, las cuales se presentan a continuación:

## **Estructura general:**

La base de datos consta de 6 tablas interrelacionadas: países, departamentos, tipos, eventos departamento, eventos país, razones y df. Estas tienen el prefijo “Deforestación” para indicar el proyecto referencia y a continuación se presenta la descripción de las tablas:

### **Descripción de las tablas:**

* *Tabla países: Almacena información sobre los países*

**Campos:**

* **codigo\_país** (INT, clave primaria, autoincremental): Hace referencia al identificador único de cada país
* **país** (VARCHAR (100), not null): Hace referencia al nombre del país
* *Tabla departamentos: Contiene información sobre los departamentos de cada país*

**Campos:**

* **id\_departamentos** (INT, clave primaria, auto incremental): Hace referencia al identificador único de cada departamento
* **departamento** (VARCHAR (50), no nulo): Hace referencia al nombre del departamento
* **area\_ha** (FLOAT): Representa el número de hectáreas de las que se compone el departamento
* **codigo\_pais** (INT, clave foránea que referencia a países, codigo\_pais): Llave que relaciona las tablas de países y departamentos
* *Tabla tipos:**Almacena los diferentes tipos de eventos o categorías*

**Campos:**

* **id\_tipos** (INT, clave primaria, auto incremental): Hace referencia al identificador único de cada tipo de evento forestal
* **tipos** (VARCHAR (50)): Hace referencia al tipo de evento forestal sucedido (Si corresponde a una ganancia o pérdida arbórea)
* *Tabla eventos\_departamento*: *Registra eventos de deforestación a nivel departamental*

**Campos:**

* **id\_eventos** (INT, clave primaria, auto incremental): Hace referencia al identificador único de cada tipo de evento forestal a nivel departamental
* **año** (VARCHAR (5)): Corresponde a un marcador temporal que hace referencia al año en que sucedió el evento
* **id\_departamentos** (INT, clave foránea que referencia a departamentos, id\_departamentos): Llave que relaciona las tablas de departamentos y eventos\_departamento
* **area\_ha** (FLOAT): Hace referencia a la cantidad de hectáreas del departamento que fueron afectadas por el evento
* **id\_tipos** (INT, no nulo, clave foránea que referencia a tipos, id\_tipos): Llave que relaciona las tablas de tipos y eventos\_departamento, indicando si hubo pérdida o ganancia de cobertura arbórea
* *Tabla razones: Almacena las razones o causas de deforestación*
* **id\_razones** (INT, clave primaria, auto incremental): Hace referencia al identificador único de cada causa de pérdida arbórea
* **razón** (VARCHAR (50), no nulo): Explica cada una de las causas de pérdida arbórea almacenadas en la base de datos
* *Tabla eventos país: Registra eventos de deforestación a nivel país*

**Campos:**

* **id\_eventoPais** (INT, clave primaria, auto incremental): Hace referencia al identificador único de cada evento de deforestación a nivel Colombia
* **id\_razones** (INT, no nulo, clave foránea que referencia a razones, id\_razones): Llave que relaciona las tablas de razones y eventos país, de forma que se pueda identificar la razón del evento de deforestación
* **id\_tipos** (INT, no nulo, clave foránea que referencia a tipos, id\_tipos): Llave que relaciona las tablas de tipos y eventos país, de forma que se pueda identificar si el evento corresponde a una ganancia o pérdida arbórea
* **codigo\_pais** (INT, no nulo, clave foránea que referencia a países, codigo\_pais): Llave que relaciona las tablas de países y eventos país, de forma que se pueda identificar a que país corresponde
* **año** (VARCHAR (5), no nulo): Corresponde a un marcador temporal que hace referencia al año en que sucedió el evento
* **area\_ha** (FLOAT, no nulo): Corresponde al número de hectáreas afectada por el evento
* **Total\_año** (FLOAT, no nulo): Corresponde a la suma total de hectáreas que fueron afectadas por cada razón
* **ha\_restantes** (FLOAT, no nulo): Corresponde a la resta entre el número de hectáreas afectadas por cada razón y el número de hectáreas totales del país

### **Relaciones entre tablas:**

* La tabla departamentos está relacionada con países.
* eventos\_departamento está relacionada con departamentos y tipos.
* eventos\_pais está relacionada con razones, tipos y países.

### **Observaciones:**

* La base de datos permite registrar eventos de deforestación tanto a nivel de país como a nivel de departamento.
* Se pueden asociar tipos y razones a los eventos de deforestación.
* El área deforestada se mide en hectáreas (ha).
* Los años se almacenan como cadenas de texto (VARCHAR) en lugar de un tipo de dato de fecha.
* Esta estructura permite realizar análisis detallados sobre la deforestación, comparando datos entre países, departamentos, años, tipos de eventos y razones de deforestación.

## **Análisis de datos**

A partir del proceso anterior, se realizó todo el análisis de la información obtenida utilizando la herramienta de Power BI, esto con el fin de generar diferentes Dashboards para presentar la información de manera más sencilla y poder realizar la exploración de los elementos más importantes. En línea con lo anterior, también se presenta la matriz de correlación.

Posterior a ello se elaboró el modelo ARIMA (Modelo autorregresivo integrado de media móvil) en Python, con el propósito de comprender la interrelación entre las variables planteadas y así proyectar la pérdida arbórea de los próximos diez años partiendo de datos históricos, se validaron los datos estadísticos más relevantes para determinar la eficiencia del modelo y se obtuvo la cantidad de hectáreas que se perderá a nivel país para los próximos años, tal como se muestra en la Tabla 1.

**Tabla 1. Pronóstico de pérdida de cobertura arbórea en el país para los períodos de 2024 a 2032**

Tabla

Descripción generada automáticamente

*Fuente: Elaboración propia, basado en el pronóstico del modelo A.R.I.M.A (2024).*

Finalmente se realizó el análisis de los resultados obtenidos y fueron presentados gráficamente mediante el uso de Dashboards con la herramienta de Power BI.

# **RESULTADOS**

## **Resultados Iniciales:**

A continuación, se presenta el análisis de los resultados de los datos comprendidos entre el período 2001 a 2023.

**Ilustración 1. Deforestación por año en Colombia (2001-2023)**

Gráfico, Gráfico de barras, Histograma

Descripción generada automáticamente

*Fuente: Elaboración propia, basado en los datos obtenidos de la entidad Global Forest Watch (2024).*

Al analizar los datos iniciales se observa que, según el gráfico 1, durante los 22 años la pérdida arbórea fue muy representativa, alcanzando un pico de 424.870,18 ha en el año 2017 a nivel país; A su vez, el año 2023 es el año que registra los menores niveles de deforestación, con un total de 122.417,12 ha a nivel país, lo cual representa un 28,81% del total del pico máximo del año 2017, siendo una cifra muy elevada, ya que en promedio se pierden alrededor de 234.383,791 ha cada año.

**Ilustración 2. Cobertura arbórea Colombia por año (2001-2023)**

Gráfico, Gráfico de líneas

Descripción generada automáticamente

*Fuente: Elaboración propia, basado en los datos obtenidos de la entidad Global Forest Watch (2024).*

En el gráfico 2, podemos apreciar el acumulado de la pérdida de cobertura arbórea que ha tenido Colombia en los últimos años, destacando que hay una constancia en la pérdida en los últimos 20 años, representado en más de 20 millones de hectáreas.

**Gráfico, Gráfico de barras

Descripción generada automáticamenteIlustración 3. Deforestación por departamento (2001-2023)**

*Fuente: Elaboración propia, basado en los datos obtenidos de la entidad Global Forest Watch (2024).*

El gráfico número 3 representa los 5 departamentos con mayor pérdida de cobertura arbórea acumulada en los últimos años, mostrando de manera visual la cantidad de hectáreas perdidas por año, para poder analizar a detalle por año de cada departamento, se tiene el dashboard realizado en la herramienta de Power BI en los anexos del proyecto.

**Ilustración 4. Deforestación por causa (2001-2023)**

Gráfico, Histograma

Descripción generada automáticamente

*Fuente: Elaboración propia, basado en los datos obtenidos de la entidad Global Forest Watch (2024).*

El gráfico número 4 representa las principales causas de deforestación en Colombia, siendo las dos más influyentes en los últimos años, la agricultura migratoria, la cual consiste en el proceso por el cual las personas llegan a un ecosistema determinado, lo destruyen y luego se van, por otro lado, la extracción de materia prima también es uno de los principales factores de deforestación, destacando actividades como la minería.

**Ilustración 5. Matriz de correlación de causas de deforestación en Colombia**

Imagen que contiene Gráfico

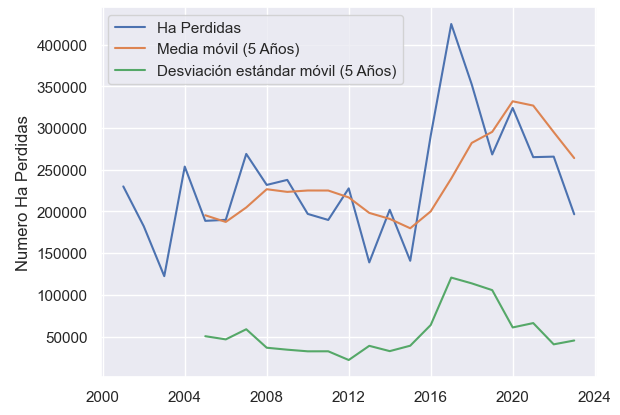
Descripción generada automáticamente

*Fuente: Elaboración propia, basado en los datos obtenidos de la entidad Global Forest Watch (2024).*

La matriz de correlación nos indica que tan relacionadas están unas causas de deforestación las otras, destacando en este caso 2 conexiones principalmente con una correlación del 76% aproximadamente, siendo las de la urbanización con la forestación y la de la agricultura con la de extracción, sin bien existen relaciones muy significativas entre unas y otras, en la vida real se debe de verificar la información ya que cada una es una causa diferente, fuera de esto, la representación de la información se encuentra en millones de hectáreas, por lo cual pueden no estar correlacionadas.

## **Modelo A.R.I.M.A:**

**Ilustración 6. Principales medidas estadísticas obtenidas a partir del modelo A.R.I.M.A**



*Fuente: Elaboración propia, basado en los datos obtenidos de la entidad Global Forest Watch (2024).*

Mediante el gráfico 6, se ilustra la media de hectáreas perdidas cada 5 años, junto con su desviación estándar, como podemos observar la desviación es baja, lo cual sugiere que las pérdidas de hectáreas son similares cada año, al estar cerca de la media. Fuera de esto, mediante la tabla 2, podemos observar esta media de perdida cada 5 años, teniendo un promedio general de pérdida de 235.035,33 ha.

**Tabla 2. Representación de pérdida media de hectáreas forestales cada 5 años**

Tabla

Descripción generada automáticamente

*Fuente: Elaboración propia (2024).*

**Ilustración 7. Autocorrelación y autocorrelación parcial del modelo predictivo**

Gráfico, Gráfico de cajas y bigotes

Descripción generada automáticamente

*Fuente: Elaboración propia (2024).*

Por medio del gráfico 7, analizamos las funciones de autocorrelación (ACF) y autocorrelación parcial (PACF) de los datos. Esto nos permitió determinar los parámetros p, d y q, que representan el orden auto regresivo, el grado de diferenciación y el orden de la media móvil respectivamente.

Destacando para el gráfico ACF (Autocorrelation Function) cómo los valores actuales de la serie temporal están correlacionados con los valores en diferentes retardos (lags), visualmente vemos una caída rápida a cero después de unos pocos retardos, indicando que la serie puede ser estacionaria.

## **Resultados Finales:**

**Ilustración 8. Proyección de la Deforestación en Colombia para el año 2032**

Gráfico, Gráfico de líneas

Descripción generada automáticamente *Fuente: Elaboración propia (2024).*

Al analizar los datos obtenidos a partir del pronóstico, según el gráfico 8, durante los 10 años siguientes, se espera una tendencia creciente de la deforestación en Colombia, sin llegar a tener picos tan altos como en los años anteriores, pero de igual forma teniendo un alza constante, fuera de esto se destaca que el año 2024 será el año de menor nivel de pérdida y el año 2030 aquel en donde se perderá la mayor cantidad de cobertura arbórea en el país.

## **Metodología de reforestación propuesta:**

**A partir de los datos obtenidos y según la revisión de las principales técnicas de reforestación usadas a nivel mundial, el mejor método para la reforestación en Colombia dependerá de la región y condiciones ambientales que presente la locación donde será llevada a cabo. Sin embargo, debido a que la mayor cantidad de hectáreas que se han perdido corresponde a zonas principalmente rurales y al no poder realizar grandes inversiones debido a los problemas de corrupción y lucha contra los grupos al margen de la ley, los sistemas silvopastoriles son probablemente la opción más adecuada. Este método según el artículo titulado “Sistemas silvopastoriles – Aspectos teóricos y prácticos” (2020), es muy adecuado para regiones ganaderas como los Llanos Orientales, la región Caribe y el Eje Cafetero, donde la integración de árboles con la producción ganadera no solo ayudaría a restaurar la biodiversidad, sino que también mejoraría la productividad del ganado y la sostenibilidad de la agricultura, además de que mejora la biodiversidad local, factor crucial en un país como Colombia para la preservación de su flora y fauna.**

**Para el país, la industria ganadera es una de las más importantes y los sistemas silvopastoriles pueden ayudar a generar un incremento en su productividad y beneficio para ser implementada, esto al reducir la degradación del suelo y ayudar a mitigar el cambio climático mediante la captura de carbono. Los departamentos de Córdoba o Casanare son un claro ejemplo de los beneficios de esta metodología al reducir la presión sobre los bosques nativos y permitir la recuperación del suelo y la vegetación.**

**Sin embargo, se deben contemplar los retos de su implementación en el país, el primero de ellos respecto al ámbito cultural, ya que se debe realizar un cambio de mentalidad y de las prácticas utilizadas por parte de los ganaderos, lo que puede implicar resistencia inicial, además de que el éxito de esta metodología depende de una gestión adecuada del pastoreo y del conocimiento técnico para integrar correctamente árboles, pasto y ganado.**

# **CONCLUSIONES**

* Mediante los resultados obtenidos se puede apreciar una tendencia creciente a la deforestación en Colombia para los próximos años, si bien los datos presentan una tendencia decreciente en el período comprendido entre 2018 a 2023, estos están influenciados por factores como la pandemia, por lo cual, la reactivación de los sectores económicos genera un aumento significativo en las actividades forestales y puede afectar el pronóstico realizado.
* La revisión realizada para cada uno de los departamentos permite identificar las principales tendencias de cada una de las zonas del país, de forma que se puede segmentar la información y se pueden utilizar estrategias como el principio de Pareto, para abordar las problemáticas más graves por tipo de causa y zona en el país y poder solucionar paralelamente los demás problemas en los departamentos de menores niveles de deforestación.
* Los datos obtenidos por medio del Modelo ARIMA son una referencia para establecer un punto de partida en la mitigación de la deforestación en Colombia, sin embargo, la efectividad del modelo puede verse afectado por factores políticos, económicos y sociales los cuales no fueron contemplados al momento de hacer la estimación, por lo cual se debe ajustar el pronóstico realizado para aumentar su efectividad.

# **RECOMENDACIONES**

* Una parte vital del uso de estas herramientas de analítica de datos es poder contemplar diferentes técnicas y conocerlas para determinar su aplicabilidad, en el caso de la predicción de datos, sin bien pueden existir diferentes modelos para obtener el mismo resultado, es importante primero validar la efectividad de la técnica utilizada ya que puede que existan modelos que no generen datos que se ajusten de manera precisa.
* La integración de diferentes herramientas para la creación, análisis y visualización de los datos permite tener un conocimiento holístico en el manejo de cada uno de los programas y permite sacar el máximo provecho a la exploración de la información, ya que se utilizan métodos y procesos acordes a las necesidades que requerimos en cada una de las fases del proyecto, en lugar de utilizar una sola herramienta que puede no ser tan efectiva para llevar cabo todas las etapas.
* La implementación de la metodología de reforestación para Colombia, si bien está adaptada para las condiciones forestales del país, de igual forma se deben analizar otras variables tales como el tiempo, inversión y las condiciones específicas de cada departamento, de forma que se pueda tener un panorama claro y ajustarlo para poder sobreponerse a las cifras obtenidas de deforestación que habrá en el país en los próximos años según las necesidades que presenten cada una de las zonas del país.

# **BIBLIOGRAFÍA**

[1] Global Forest Watch (2023). Colombia Deforestation Rates & Statistics. Available at: https://acortar.link/L4BZmS

[2] Puentes, P. (2024) Colombia: Disminuyó La Deforestación en el 2023, Pero va en aumento en el 2024, Noticias ambientales (MONGABAY). Available at: https://es.mongabay.com/2024/04/colombia-disminuyo-deforestacion-2023-aumento-en-2024/

[3] Rojas Sarmiento, L. (2024) Deforestación en Colombia: Retos y oportunidades, Señal Colombia. Available at: https://www.senalcolombia.tv/general/deforestacion-en-colombia-retos

[4] Zapata Cadavid A. & Silva Tapasco B.E. 2020. Sistemas silvopastoriles: aspectos teóricos y prácticos CARDER, CIPAV. Editorial CIPAV. Segunda edición. Cali, Colombia. 242p.

1. Catalinagh4@yahoo.es [↑](#footnote-ref-1)
2. 2 lpenapin@gmail.com

   3 jdruizbe@unal.edu.co

   4 angel.1.santigor@gmail.com

   5 Santiagopalaciosoto@gmail.com [↑](#footnote-ref-2)
3. [↑](#footnote-ref-3)
4. [↑](#footnote-ref-4)
5. [↑](#footnote-ref-5)
6. [↑](#footnote-ref-6)