

Pronóstico de la salud de la vegetación mediante aprendizaje automático aplicado a imágenes
multiespectrales de satélite

Cardenas Lopez Miller Alexander y Garcia
Arboleda Brandon Stiven

Programa de ingeniería de sistemas, facultad de ingeniería, Universidad CESMAG

Nota del autor

El presente Trabajo de Grado tiene como propósito cumplir el requisito exigido para optar al título de pregrado como Ingenieros de sistemas en la Universidad CESMAG. La correspondencia referente a este trabajo debe dirigirse al Programa de Ingeniería de sistemas de la Universidad CESMAG. Correo electrónico: ingenieriadesistemas@unicesmag.edu.co

Pronóstico de la salud de la vegetación mediante aprendizaje automático aplicado a imágenes multiespectrales de satélite

Cardenas Lopez Miller Alexander y Garcia
Arboleda Brandon Stiven

Ingeniería de sistemas, facultad de ingeniería, Universidad CESMAG

Asesora: Mg. Dalila Mercedes Pachajoa

11 de agosto de 2026

Nota de aceptación

“Va la aceptación definida desde la universidad”.

Nota de exclusión

DEDICATORIA

Dedico este trabajo a mis padres, por su esfuerzo y dedicación que hicieron posible este logro. A mi madre, por su amor y apoyo incondicional en cada momento. A mi padre, quien, aunque ya no está físicamente, sé que desde donde se encuentre me brinda fuerza y motivación. A mi hermana quien ha sido mi apoyo incondicional.

A mi familia y a quienes creyeron en mí, por su acompañamiento y confianza durante este proceso.

Miller Alexander Cardenas López

DEDICATORIA

Este trabajo es dedicado principalmente a mis padres, quienes me apoyaron con la decisión de estudiar y creyeron en mí, principalmente por sus esfuerzos y sacrificios que hicieron para que esto sea posible, estuvieron presentes de la mejor manera posible dentro de sus posibilidades, brindándome lo que estaba a su alcance y más.

Segundo a Dios quien ha estado siempre presente en los momentos en los cuales sentía que no podría y siendo esa compañía cuando mis padres no podían estar.

Tercero a mí, por no darme por vencido pese a las adversidades de la vida, dificultades de entorno y ganas para pensarlo creerlo y ejecutarlo.

Brandon Estiven García Arboleda

AGRADECIMIENTOS

Agradecemos de manera especial a los docentes Dalila Mercedes Pachajoa y Héctor Mora, por su orientación, acompañamiento y aportes durante el desarrollo de este proyecto, los cuales fueron fundamentales para el cumplimiento de los objetivos planteados.

De igual forma, agradezco a la Universidad de Estudios Superiores María Goretti CESMAG por la formación académica brindada y por proporcionar los espacios y recursos necesarios para el desarrollo de esta investigación.

Finalmente, agradezco a mi familia por su apoyo constante durante este proceso.

TABLA DE CONTENIDO

	Pág.
INTRODUCCIÓN	14
I. PROBLEMA DE INVESTIGACIÓN	15
A. OBJETO O TEMA DE ESTUDIO	15
B. LÍNEA DE INVESTIGACIÓN	15
C. PLANTEAMIENTO DEL PROBLEMA.....	15
D. FORMULACIÓN DEL PROBLEMA	16
E. OBJETIVOS.....	16
F. JUSTIFICACIÓN.....	17
G. DELIMITACIÓN.....	18
II. MARCO TEÓRICO	19
A. ANTECEDENTES.....	19
B. SUPUESTOS TEÓRICOS.....	23
C. VARIABLES DEL ESTUDIO.....	28
D. FORMULACIÓN DE HIPOTESIS	34
III. METODOLOGÍA	36
A. PARADIGMA.....	36
B. ENFOQUE	36
C. MÉTODO.....	36
D. TIPO DE INVESTIGACIÓN.....	37
E. DISEÑO DE LA INVESTIGACIÓN	37
F. POBLACIÓN	38
G. MUESTRA.....	38
H. TÉCNICAS DE RECOLECCIÓN DE INFORMACIÓN	38
I. INSTRUMENTO DE RECOLECCIÓN DE DATOS	41
J. METODOLOGÍA DE DESARROLLO.....	41
IV. RESULTADOS DE LA INVESTIGACIÓN	45
Objetivo 1.....	45
Objetivo 2.....	49
Objetivo 3.....	54
V. ANÁLISIS Y DISCUSIÓN DE LOS RESULTADOS.....	63
CONCLUSIONES	66

PRONOSTICO DE LA SALUD DE LA VEGETACIÓN	8
RECOMENDACIONES	67
REFERENCIAS BIBLIOGRÁFICAS	68

LISTA DE FIGURAS

Fig 1. Teledetección.	23
Fig 2. Imágenes multiespectrales	24
Fig 3. NDVI multiespectrales.....	25
Fig 4. Calculo NDVI	46
Fig 5. Mapa de predicción del índice NDVI obtenido mediante la arquitectura ConvLSTM	47
Fig 6. Mapas de NDVI correspondientes a los valores mínimo, promedio y máximo, generados a partir de las bandas espectrales.	48
Fig 7. Dataset final consolidado utilizado para el entrenamiento de los modelos de inteligencia artificial.	49
Fig 8. Visualización inicial del conjunto de datos utilizado para el entrenamiento del modelo de inteligencia artificial.....	51
Fig 9. Estructura de la ventana temporal utilizada para el entrenamiento del modelo de inteligencia artificial	51
Fig 10. Diagrama de la arquitectura del modelo diseñado para el análisis de vegetación.	52
Fig 11. Curvas de aprendizaje (Loss) obtenidas durante el entrenamiento del modelo para la ventana temporal seleccionada.....	53
Fig 12. model.keras almacenado en el repositorio del proyecto	54
Fig 13 Curvas de aprendizaje (Loss) para las cuatro configuraciones de ventana temporal evaluadas. La ventana de 15 días presenta la convergencia más estable	58
Fig 14 Mapas comparativos NDVI real vs. NDVI predicho para la ventana de 3 días. Se observa escasa correspondencia espacial, consistente con el R^2 negativo obtenido.	59
Fig 15 Mapas comparativos NDVI real vs. NDVI predicho para la ventana de 7 días. El modelo comienza a reproducir los patrones generales de distribución de la vegetación.	59
Fig 16 Mapas comparativos NDVI real vs. NDVI predicho para la ventana de 15 días (configuración óptima). Alta correspondencia espacial con los valores reales en la mayoría del departamento	59
Fig 17 Mapas comparativos NDVI real vs. NDVI predicho para la ventana de 30 días. Desempeño ligeramente inferior al de 15 días, especialmente en las zonas costeras del Pacífico	59
Fig 18 Dashboard correspondiente a los valores observados y predicción.....	60
Fig 19 Representación espacial final.....	61

LISTA DE TABLAS

TABLA I. Resultados lees et al. (2022).....	19
TABLA II. Ejemplo de estructura tabular del dataset multiespectral utilizado en el entrenamiento	24
TABLA III. Proceso de preprocesamiento.....	42
TABLA IV. Configuración de modelo	43
TABLA V. Proceso de evaluación.....	44
TABLA VI. Bandas espectrales	45
TABLA VII. Modelos auxiliares entrenados para estimación de bandas espectrales	50
TABLA VIII. Desempeño de modelos auxiliares de preprocesamiento.....	54
TABLA IX. Desempeño según ventana temporal	55
TABLA X. Desempeño comparativo del modelo convLSTM2D.....	56

LISTA DE ECUACIONES

Ecuación 1 Modelo LSTM estándar.....	26
Ecuación 2 Modelo ConvLSTM	27
Ecuación 3. SVM	31
Ecuación 4 clasificación.....	31
Ecuación 5 Regresión.....	31
Ecuación 6 KNN	32
Ecuación 7 Forget Gate	33
Ecuación 8 Input Gate	33
Ecuación 9 Cell State	33
Ecuación 10 Output Gate	33
Ecuación 11 Normalización	34
Ecuación 12 Estandarización	34
Ecuación 13 Imputación.....	34

GLOSARIO

Imágenes satelitales: Representaciones de la superficie terrestre captadas por sensores instalados en satélites, que registran información en diferentes regiones del espectro electromagnético. Estas imágenes permiten analizar características del territorio como la cobertura vegetal, siendo fundamentales en estudios ambientales, agrícolas y de monitoreo de la vegetación.

Multiespectrales: Propiedad de los datos obtenidos en múltiples bandas del espectro electromagnético, lo que facilita el análisis detallado de elementos presentes en la superficie terrestre. En el contexto de este estudio, las imágenes multispectrales permiten identificar y evaluar la vegetación mediante la combinación de bandas visibles e infrarrojas.

NDVI (Índice de Vegetación de Diferencia Normalizada):

Índice utilizado en teledetección para evaluar la presencia y el estado de la vegetación a partir de imágenes satelitales. Se calcula mediante la relación entre la reflectancia del infrarrojo cercano (NIR) y la banda roja (RED), permitiendo identificar áreas con mayor o menor actividad vegetal.

Transformación de datos: Conjunto de métodos que se emplean para transformar datos de un formato a otro con el propósito de optimizar su calidad, simplificar su análisis o adecuarlos a modelos de aprendizaje automático. Incorpora procedimientos como la normalización, la estandarización, el reemplazo de escalas y la eliminación de valores anómalos, los cuales son esenciales para mejorar el desempeño de los algoritmos predictivos.

Predicción de series de tiempo: Técnica que permite estimar el comportamiento futuro de una variable a partir de datos históricos organizados en el tiempo. En este proyecto, se utiliza para predecir la evolución de la vegetación mediante el análisis de secuencias temporales de imágenes satelitales, empleando modelos de inteligencia artificial capaces de capturar patrones espaciales y temporales.

RESUMEN

La presente investigación tiene como objetivo analizar y predecir la salud de la vegetación en el departamento de Nariño mediante el uso de imágenes satelitales multiespectrales provenientes de Landsat y MODIS. En el marco referencial se abordan conceptos relacionados con la teledetección, el índice de vegetación de diferencia normalizada (NDVI) y el uso de modelos de inteligencia artificial para el análisis de datos geoespaciales.

Metodológicamente, el estudio se desarrolló bajo un enfoque cuantitativo aplicado con diseño experimental computacional y alcance predictivo, utilizando datos históricos satelitales que fueron preprocesados, organizados y transformados en un conjunto de datos estructurado. Se aplicaron técnicas de extracción de características espectrales y cálculo del NDVI, así como la construcción de ventanas temporales para capturar la dinámica de la vegetación. Posteriormente, se diseñó e implementó un modelo de aprendizaje profundo basado en redes neuronales convolucionales recurrentes (ConvLSTM), capaz de aprender patrones espaciales y temporales a partir de secuencias de imágenes.

Como principales resultados, se obtuvo un dataset estructurado que integra información espacial, temporal y espectral, además de representaciones visuales del NDVI que permiten identificar la variabilidad de la cobertura vegetal en la región. Los modelos desarrollados fueron evaluados mediante el coeficiente de determinación (R^2), evidenciando un mejor desempeño al utilizar el modelo ConvLSTM2D en la ventana temporal de 15 días, alcanzando un valor de R^2 de 0.699, lo que indica una adecuada capacidad predictiva.

Se concluye que el uso de modelos de aprendizaje profundo aplicados a imágenes satelitales permite analizar y predecir de manera efectiva el comportamiento de la vegetación, constituyéndose en una herramienta útil para el monitoreo ambiental y la toma de decisiones en contextos agrícolas y ecológicos.

Palabras Clave: aprendizaje profundo, imágenes satelitales, NDVI, vegetación

INTRODUCCIÓN

El monitoreo de la salud de la vegetación a través de tecnologías de vanguardia es un campo que combina disciplinas como el aprendizaje automático y la teledetección satelital, fundamentales para entender y cuidar nuestro planeta ante el cambio climático. Estas herramientas permiten anticipar cambios en la salud de la vegetación, un factor clave para la gestión sostenible de los recursos naturales [5].

Por lo tanto, la capacidad de combinar datos satelitales con algoritmos inteligentes permite abordar los problemas medioambientales desde ángulos novedosos y con mayor eficacia. Partiendo de esta idea, el presente proyecto se enfoca en el desarrollo de un modelo de pronóstico para evaluar la salud de la vegetación en el departamento de Nariño, Colombia, a partir de imágenes multiespectrales de satélite y técnicas de aprendizaje automático (machine learning).

Este trabajo busca responder a la necesidad de contar con herramientas predictivas que identifiquen las zonas más vulnerables y ayuden a planificar las estrategias de conservación en un territorio en el que la ausencia de sistemas de este tipo ha frenado las acciones preventivas [40]. El presente trabajo se divide en dos grandes partes: en primer lugar, se enmarcan el problema y su contexto regional; a continuación, se hace un repaso de los antecedentes relacionados con el uso de teledetección y aprendizaje automático en el estudio de la vegetación. Después, se aborda la viabilidad del proyecto y se explica la metodología que contempla la adquisición de datos satelitales, su procesamiento y la construcción de modelos predictivos. Por último, se exponen las fases de evaluación y los resultados esperados, con la puesta a punto de una herramienta útil y fácilmente extrapolable para la gestión ambiental [18].

I. PROBLEMA DE INVESTIGACIÓN

A. OBJETO O TEMA DE ESTUDIO

Pronóstico de la salud de la vegetación en el departamento de Nariño mediante la aplicación de algoritmos de aprendizaje automático con datos de imágenes multiespectrales de los satélites Landsat [27] y MODIS [28], con el fin de generar información para su análisis de calidad

B. LÍNEA DE INVESTIGACIÓN

Tecnología de la información y la Comunicación

Sub línea de investigación

Inteligencia artificial

C. PLANTEAMIENTO DEL PROBLEMA

El departamento de Nariño se caracteriza por su extraordinaria diversidad de ecosistemas, que abarcan desde páramos y bosques andinos hasta zonas áridas y regiones costeras del Pacífico. Esta riqueza natural no solo es un valor ecológico, sino también un recurso vital para las comunidades que dependen de la vegetación para actividades como la agricultura y la conservación de suelos. Sin embargo, esta diversidad enfrenta amenazas crecientes debido a factores como la deforestación, el cambio climático, la expansión agrícola y la degradación del suelo [52].

Contar con herramientas que permitan pronosticar la salud de la vegetación es esencial para una gestión ambiental efectiva. Anticipar cómo evolucionarán los ecosistemas ayudaría a identificar áreas en riesgo y planificar estrategias de conservación. En Nariño, las herramientas actuales para monitorear la vegetación presentan limitaciones importantes. Aunque se dispone de estudios que describen el estado pasado y presente de los ecosistemas mediante observaciones de campo o imágenes satelitales básicas, estos enfoques no permiten realizar predicciones robustas [21].

Los métodos tradicionales no aprovechan el potencial de tecnologías avanzadas para procesar grandes volúmenes de datos ni identificar patrones complejos en la vegetación. Por ejemplo, las imágenes satelitales multiespectrales ofrecen una oportunidad única para evaluar la salud de la

vegetación a gran escala, pero su análisis manual o con técnicas básicas no es suficiente para generar pronósticos precisos [22].

Este proyecto propuso el desarrollo de un modelo basado en aprendizaje automático, una rama de la inteligencia artificial que permitió identificar patrones y hacer predicciones a partir de datos recientes. Al integrar imágenes satelitales multiespectrales con algoritmos de aprendizaje automático, se pudieron analizar variables críticas para pronosticar cómo cambiaría la salud de la vegetación en el futuro [10].

D. FORMULACIÓN DEL PROBLEMA

¿Qué modelo de aprendizaje profundo, aplicado a series temporales de imágenes multiespectrales satelitales, permite pronosticar con mayor precisión la salud de la vegetación en el departamento de Nariño?

E. OBJETIVOS

1) Objetivo general

Desarrollar y evaluar modelos de Machine Learning, utilizando datos de imágenes multiespectrales de satélite, para pronosticar la salud de la vegetación en el departamento de Nariño, con el fin de facilitar su análisis de calidad.

2) Objetivos específicos

- Preprocesar y limpiar imágenes satelitales multiespectrales (Landsat y MODIS) para generar un conjunto de datos estructurado, espacial y temporalmente consistente
- Diseñar y entrenar modelos de inteligencia artificial para analizar la vegetación a partir de imágenes satelitales.
- Evaluar la precisión y la robustez de los modelos para definir el más preciso para el pronóstico de la salud de la vegetación, utilizando métricas apropiadas para su posterior análisis de calidad.

F. JUSTIFICACIÓN

El departamento de Nariño es uno de los más biodiversos del país, sin embargo a su alrededor giran varios problemas que hacen que su diversidad esté en peligro, problemas comunes como el cambio climático, la deforestación y la expansión agrícola [51] han hecho que los ecosistemas del departamento estén en peligro a tal punto que ya se hacen necesarias herramientas tecnológicas que permitan monitorear constantemente estos ecosistemas [41] y conocer por anticipado su estado, problema que se ha estudiado pero con métodos que analizan datos históricos y no se enfocan en pronosticar el estado o la salud de estos ecosistemas.

En este contexto, el uso de imágenes multiespectrales combinadas con algoritmos de aprendizaje profundo representa una oportunidad innovadora para transformar la gestión ambiental del territorio. A diferencia de los métodos estadísticos convencionales, las redes neuronales convolucionales recurrentes (ConvLSTM2D) son capaces de procesar simultáneamente la dimensión espacial (distribución geográfica de la vegetación) y la dimensión temporal (evolución de los índices espectrales a lo largo de semanas), capturando patrones fenológicos que los modelos tradicionales no pueden detectar.

Esta capacidad de análisis espacio-temporal es decisiva para ecosistemas andinos como los de Nariño, donde la variabilidad climática y la topografía compleja generan dinámicas de vegetación heterogéneas que requieren enfoques metodológicos avanzados.

Este proyecto propone el desarrollo de un modelo de inteligencia artificial que analice imágenes satelitales multiespectrales para pronosticar la salud de la vegetación en Nariño. Su valor metodológico radica en la integración de técnicas avanzadas de preprocesamiento, modelado y evaluación, que permitirán detectar patrones complejos en grandes volúmenes de datos [47]. Además, este enfoque tiene potencial de replicabilidad en otras regiones del país, lo que lo convierte en una contribución significativa tanto al ámbito académico como al desarrollo de tecnologías aplicadas al análisis geoespacial [8].

Los beneficios esperados abarcan dimensiones sociales, económicas y académicas. A nivel social, facilitará la toma de decisiones anticipadas para prevenir la degradación ecológica. En lo económico, permitirá optimizar recursos al priorizar intervenciones preventivas sobre acciones correctivas. En el plano académico, fortalecerá la formación en inteligencia artificial aplicada y

aportará conocimiento valioso sobre el uso de tecnologías emergentes en la conservación ambiental. En conjunto, este proyecto ofrece una solución práctica e innovadora a un problema ambiental crítico en la región [25].

G. DELIMITACIÓN

Este trabajo se enfoca en estudiar la vegetación del departamento de Nariño a partir de imágenes tomadas por satélites. El análisis cubre zonas representativas de diferentes ecosistemas del departamento, como los páramos, los bosques andinos y la costa del Pacífico, durante el período comprendido entre 2015 y 2024.

Para el desarrollo del estudio se utilizaron únicamente datos de satélites que capturan múltiples bandas de luz (principalmente Landsat-8 y MODIS), por lo que no se recurrió a otros tipos de tecnología como drones, radares ni mediciones directas en campo. Los resultados se centran en medir la salud de la vegetación mediante índices espectrales como el NDVI o el VCI, sin entrar en temas económicos, sociales ni en el análisis de especies de plantas en particular.

El período de análisis sigue la frecuencia con que los satélites capturan imágenes (aproximadamente cada quince días o cada mes), y se comparan años con alta lluvia frente a años secos para entender cómo responde la vegetación al clima. Por lo tanto, el estudio no pretende analizar escenarios futuros de cambio climático ni evaluar otros indicadores ambientales más allá de lo que se puede observar a través de la cobertura vegetal registrada por los satélites [42].

II. MARCO TEÓRICO

A. ANTECEDENTES

1) Antecedentes internacionales

Diversos estudios internacionales han aplicado técnicas avanzadas de teledetección [30] y aprendizaje automático al análisis de la vegetación. En el contexto del monitoreo climático y agronómico, el uso de índices satelitales como NDVI y VCI es común. Por ejemplo, T. Lees *et al.* [48] desarrollaron un modelo de pronóstico de la salud de la vegetación en territorios semidesérticos de Kenia mediante redes neuronales recurrentes tipo LSTM aplicadas a series temporales de imágenes satelitales multiespectrales.

La metodología consistió en procesar secuencias históricas del índice NDVI derivadas de imágenes MODIS con resolución temporal de 8 días. El modelo LSTM fue entrenado con ventanas temporales deslizantes de 16 días para predecir el valor del NDVI con horizontes de pronóstico de 8, 16 y 24 días.

TABLA I. Resultados lees et al. (2022)

Días	R^2	MAE
8 días	0.60	0.018
16 días	0.75	0.018

El estudio concluyó que las redes LSTM son capaces de capturar la dinámica fenológica de la vegetación en ecosistemas con estacionalidad marcada, lo que resulta crucial para anticipar sequías agrícolas [5].

Este trabajo aporta a nuestra investigación la validación de que las arquitecturas de aprendizaje profundo basadas en memoria temporal son técnicamente viables para el pronóstico de índices de vegetación en regiones con alta variabilidad climática, estableciendo un precedente metodológico para la aplicación de ConvLSTM2D en el departamento de Nariño. Por su parte, la literatura de agricultura de precisión resalta la integración de datos multiespectrales con Wang *et al.* [31], presentaron una revisión sistemática sobre el uso de sensores multiespectrales e hiperespectrales combinados con algoritmos de Machine Learning para aplicaciones de agricultura de precisión. El estudio analizó 127 investigaciones publicadas entre 2015 y 2021, clasificándolas por sensor utilizado (Landsat, Sentinel-2, MODIS, UAV hiperespectral), donde

su tarea objetivo fue la clasificación de cobertura, detección de plagas y pronóstico de rendimiento.

Los algoritmos que usaron fueron SVM, Random Forest, redes neuronales profundas. Los hallazgos principales fueron:

- Los sensores multispectrales convencionales como Landsat y Sentinel siguen siendo ampliamente utilizados debido a su cobertura global gratuita y continuidad temporal, representando el 68% de los estudios analizados.
- Los algoritmos de aprendizaje profundo (CNN, LSTM, Transformer) superan a los métodos tradicionales (SVM, Random Forest) en tareas de clasificación y pronóstico temporal, con mejoras promedio de R^2 de 0.15 a 0.25 puntos.
- La principal limitación identificada fue la escasez de datos de alta resolución temporal para entrenar modelos en regiones con alta nubosidad persistente, lo que genera vacíos en las series temporales [31] [49].

Este antecedente justifica de que Landsat es una fuente de datos válida y ampliamente validada para el análisis de vegetación a escala regional, y que las arquitecturas de aprendizaje profundo son el estado del arte para procesamiento de series temporales satelitales.

M. Bisquert desarrollaron un sistema de predicción del riesgo de incendio forestal en Galicia y Asturias mediante el análisis temporal de índices de vegetación derivados de imágenes del sensor TERRA-MODIS [54].

La metodología consistió en calcular tres índices de vegetación: EVI, GEMI y SAVI, a partir de series temporales diarias de MODIS con resolución espacial de 250 metros, durante el periodo 2000 y 2015.

Los investigadores aplicaron análisis de series temporales para identificar la variación estacional para luego comparar su capacidad para detectar anomalías de sequía previas a eventos de incendio documentados.

Los resultados mostraron que el índice EVI presentó la mejor correlación con la probabilidad de incendio ($R^2 = 0.72$), mientras que GEMI y SAVI tuvieron desempeños menores ($R^2 = 0.58$ y 0.61 respectivamente). Concluyeron que la variación temporal del EVI puede utilizarse como indicador de estrés hídrico acumulado en la vegetación, permitiendo generar alertas de riesgo de incendio con hasta 30 días de anticipación. Con este estudio podemos con validar que los índices derivados del infrarrojo cercano (como NDVI y EVI) capturan adecuadamente el estado

fisiológico de la vegetación, y que las series temporales satelitales de larga duración son esenciales para modelar dinámicas fenológicas en ecosistemas con estacionalidad climática marcada, como ocurre en el departamento de Nariño [15].

2) Antecedentes nacionales

En Colombia se han realizado estudios locales relevantes sobre índices de vegetación. En el ámbito urbano, ejemplo, Dobbs et al [58]. desarrollaron un sistema de monitoreo de la expansión urbana y su impacto sobre ecosistemas en Bogotá mediante el procesamiento de series temporales de imágenes Landsat. La metodología consistió en calcular el índice NDVI a partir de imágenes Landsat-5 TM y Landsat-8 OLI del periodo 1985–2017, con frecuencia quinquenal cada 5 años. Los investigadores clasificaron la cobertura del suelo en vegetación densa, vegetación dispersa, suelo desnudo y urbanizado mediante análisis de umbrales del NDVI y validación en campo con puntos GPS.

Los resultados mostraron que la cobertura vegetal en la Sabana de Bogotá disminuyó un 23% entre 1985 y 2017, con una tasa de pérdida acelerada después del año 2000 (1.8% anual vs. 0.9% en el periodo anterior). Se identificó que las zonas con NDVI menores a 0.4 se correlacionan con áreas de alta densidad urbana, mientras que ecosistemas como humedales mantienen NDVI mayor a 0.6. Este estudio nos permitió validar que las series temporales largas de Landsat son adecuadas para capturar dinámicas de cambio de cobertura vegetal en ecosistemas colombianos, y que el NDVI es un indicador robusto para cuantificar el impacto antrópico sobre la vegetación [16].

En contexto agroforestal, González et al. [59] documentaron pérdidas de áreas sembradas en Cundinamarca durante el periodo 2013 a 2014 mediante el análisis temporal del índice NDVI derivado de imágenes Landsat-8. La metodología consistió en procesar imágenes con frecuencia mensual (12 imágenes por año) sobre los municipios de Facatativá, Zipaquirá y Ubaté, calculando el NDVI para identificar zonas con disminución quebrada asociada a abandono de cultivos o cambio de uso del suelo. Los investigadores aplicaron un umbral de detección de cambio cuando el NDVI descendía más de 0.15 entre meses consecutivos, correlacionando estos eventos con registros catastrales de suelo desnudo.

Los resultados mostraron que aproximadamente 2.400 hectáreas de cultivos fueron abandonadas entre 2013 y 2014, con mayor concentración en zonas de ladera con baja disponibilidad hídrica.

El análisis identificó que el NDVI captura adecuadamente el estrés hídrico acumulado en cultivos de ciclo corto, con valores críticos menores a 0.35 asociados a pérdida de productividad. Este estudio aportó a la demostración de que el NDVI derivado de Landsat-8 es sensible a cambios de corto plazo en la vegetación agrícola colombiana, validando su uso como variable objetivo para modelos de pronóstico temporal [9].

En la región Caribe colombiana se aplicó NDVI al estudio de perturbaciones: Hoyos et al. (2017) desarrollaron un sistema de análisis espacial de incendios forestales entre 2003 y 2015 mediante la integración del índice NDVI con variables climáticas. La metodología incluyó el procesamiento de imágenes MODIS diarias para generar series temporales del NDVI con resolución de 250 metros sobre los departamentos de Atlántico, Bolívar y Magdalena.

Los investigadores aplicaron análisis de correlación cruzada entre la variación del NDVI y la ocurrencia de incendios reportados por el IDEAM, identificando ventanas temporales críticas donde la disminución del NDVI precede a eventos de incendio. Los resultados mostraron que valores de NDVI menores a 0.4 mantenidos durante 21 días consecutivos incrementan la probabilidad de incendio en un 68%, especialmente en ecosistemas de bosque seco tropical.

Se concluyó que el monitoreo del NDVI en la región Caribe permite anticipar zonas de alto riesgo con hasta 30 días de antelación de donde pudimos evidenciar que el NDVI es un indicador predictivo de estrés vegetal en ecosistemas colombianos con estacionalidad climática marcada, y que las series temporales satelitales son viables para generar pronósticos en horizontes de semanas.

Por otro lado, Millán et al. [60] emplearon 48 imágenes Landsat-8 para cartografiar humedales costeros en base a su cobertura vegetal[6] Más recientemente, Torres-Bejarano et al. (2022) encontraron una correlación significativa entre NDVI y parámetros de calidad del agua en el embalse del Guájaro, evidenciando la relación entre baja vegetación y deterioro hídrico [7].

Aunque algunos de estos estudios son descriptivos, demuestran la utilidad del NDVI para monitorear la vegetación colombiana en diferentes contextos ecológicos y escalas espaciales. Sin embargo, son en su mayoría retrospectivos y no emplean modelos de predicción basados en Machine Learning. En particular, en Nariño no se conocen trabajos publicados que combinen teledetección histórica con pronósticos automatizados de salud vegetal mediante aprendizaje profundo. Esto indica la pertinencia de la investigación propuesta, ya que integrará estas

aproximaciones en un contexto regional colombiano con alta biodiversidad y vulnerabilidad ambiental [50].

B. SUPUESTOS TEÓRICOS

Teledetección: la teledetección es una disciplina científica que se encarga de adquirir información sobre la superficie terrestre sin contacto directo, utilizando sensores montados en satélites o aeronaves. Estos sensores capturan datos en diferentes bandas del espectro electromagnético, permitiendo detectar características biofísicas del suelo y la vegetación. En este contexto, la teledetección facilita el monitoreo ambiental a gran escala y en periodos de tiempo prolongados.

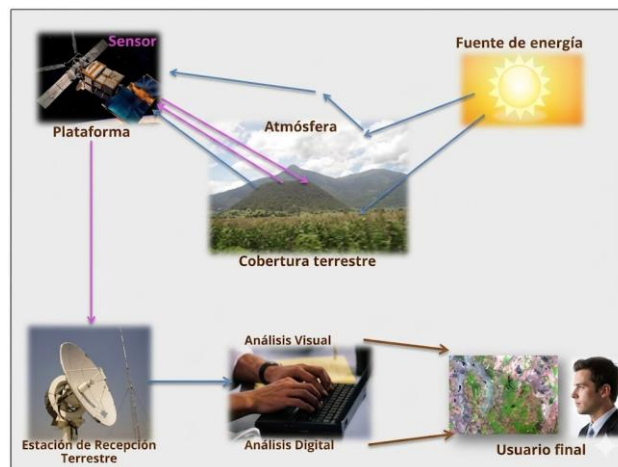


Fig 1. Teledetección.

Fuente: Tomada de INEGI,2015 [53]

Imágenes multiespectrales: las imágenes multiespectrales son registros digitales capturados por sensores remotos que recogen información en varias longitudes de onda, tanto del espectro visible como del infrarrojo cercano. Estas imágenes permiten diferenciar tipos de cobertura terrestre y analizar el comportamiento de la vegetación, ya que las plantas reflejan la luz de manera distinta dependiendo de su estado fisiológico.[32]

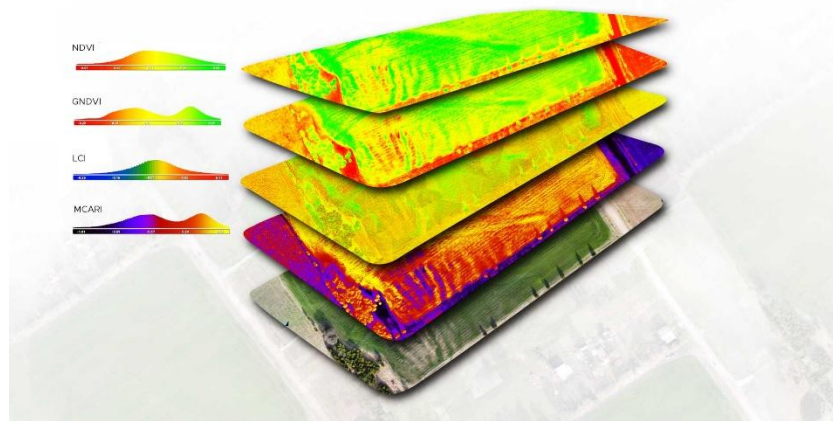


Fig 2. Imágenes multiespectrales

Fuente: tomada de AV3 AEROVISUAL, 2021 [52]

Bandas espectrales utilizadas en el estudio:

Sensor Landsat: Landsat es un programa de observación terrestre operado por la NASA y el USGS desde 1972, proporcionando imágenes multiespectrales con resolución espacial de 30 metros en las bandas del visible e infrarrojo. El sensor captura información en siete bandas espectrales del espectro electromagnético, cada una sensible a diferentes propiedades de la superficie terrestre.

En esta investigación, las bandas espectrales fueron organizadas en una estructura tabular junto con las coordenadas geográficas de cada observación, permitiendo construir el dataset utilizado para el entrenamiento del modelo de inteligencia artificial.

TABLA II. Ejemplo de estructura tabular del dataset multiespectral utilizado en el entrenamiento

Latitud	Longitud	band 1	band 2	band 3	band 4
-8789850	184950	0.1051125905	0.08271564244	0.05275554907	0.3155197183
-8784900	174600	0.1048404213	0.08573163966	0.05352727115	0.2714632767
band 5	band 6	band 7			
0.1326408839	296.1786069	0.04458619665			
0.09024746761	295.9697098	0.02729183302			

Índices de vegetación (NDVI, VCI): los índices de vegetación son fórmulas matemáticas que combinan las bandas espectrales de las imágenes satelitales para evaluar la presencia y condición de la vegetación. El NDVI (Índice de Vegetación de Diferencia Normalizada) es uno de los más utilizados, ya que permite estimar el vigor o verdor de la vegetación a partir de la reflectancia del rojo y el infrarrojo cercano. Por su parte, el VCI (Índice de Condición de la Vegetación) mide la variabilidad del NDVI a lo largo del tiempo para detectar anomalías y estrés vegetal.

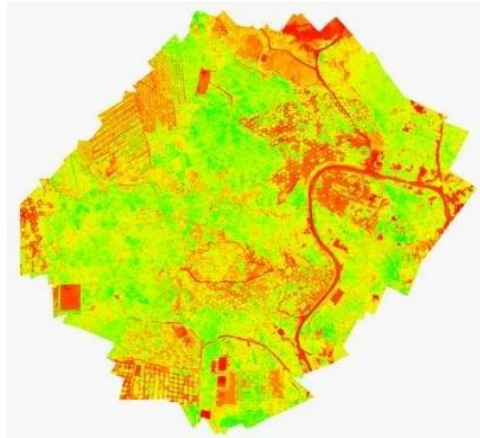


Fig 3. NDVI multiespectrales

Fuente: tomada de AV3 AEROVISUAL, 2021) [52]

Machine Learning: el machine learning es un campo de la inteligencia artificial que permite a los sistemas informáticos aprender a partir de datos, identificar patrones y hacer predicciones sin intervención humana directa. En investigaciones como esta, se utiliza para construir modelos que, basados en datos históricos, puedan pronosticar comportamientos futuros, en este caso, de la salud de la vegetación.

Modelado predictivo: el modelado predictivo consiste en el uso de algoritmos matemáticos y estadísticos para predecir eventos o comportamientos futuros a partir de datos existentes. Su fundamento teórico radica en la identificación de relaciones lógicas y consistentes entre variables observadas. En este estudio, se aplican modelos predictivos a series temporales de datos satelitales para anticipar el estado de la vegetación en distintos escenarios.

Limitaciones de las redes recurrentes clásicas

Problema del gradiente desvaneciente: las redes neuronales recurrentes (RNN) simples presentan dificultades para aprender dependencias a largo plazo debido al problema del gradiente

desvaneciente (vanishing gradient). Durante el entrenamiento mediante retropropagación a través del tiempo (BPTT), el gradiente del error se multiplica repetidamente por los pesos de la red. Si dichos pesos son menores que 1, el gradiente decae exponencialmente con la profundidad temporal, impidiendo que la red actualice sus parámetros para secuencias largas.

Redes Neuronales Convolucionales Recurrentes (ConvLSTM): Las redes ConvLSTM (Convolutional Long Short-Term Memory) constituyen una arquitectura de aprendizaje profundo que integra operaciones de convolución espacial dentro de las ecuaciones de una celda LSTM, permitiendo capturar simultáneamente dependencias espaciales y temporales en secuencias de datos estructurados en cuadrícula, como las imágenes satelitales multitemporales.

La diferencia fundamental frente a una LSTM estándar reside en la naturaleza del operador aplicado sobre las entradas y los estados ocultos. En una LSTM convencional, todas las transformaciones se realizan mediante multiplicaciones matriciales, lo que implica que la entrada X_t es un vector unidimensional y la información espacial de una imagen debe ser previamente aplanada (*flattened*), destruyendo así la estructura de vecindad entre píxeles. Las ecuaciones de una LSTM estándar son:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$

Ecuación 1 Modelo LSTM estándar

donde el operador $\cdot$ representa una multiplicación matriz-vector que no preserva relaciones espaciales.

En contraste, la ConvLSTM reemplaza ese producto matricial por una convolución espacial ($*$), de modo que tanto la entrada X_t como el estado oculto H_{t-1} son tensores tridimensionales (alto $\times$ ancho $\times$ canales). Las ecuaciones reformuladas son:

$$f_t = \sigma(W_f * [H_{t-1}, X_t] + b_f)$$

$$i_t = \sigma(W_i * [H_{t-1}, X_t] + b_i)$$

$$C_t = \tanh(W_C * [H_{t-1}, X_t] + b_C)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot C_t$$

$$o_t = \sigma(W_o * [H_{t-1}, X_t] + b_o)$$

$$H_t = o_t \odot \tanh(C_t)$$

Ecuación 2 Modelo ConvLSTM

donde $*$ es la convolución espacial y $\odot$ es el producto de Hadamard (elemento a elemento). El kernel de convolución W actúa localmente sobre una vecindad espacial definida, lo que significa que cada neurona del estado oculto H_t solo depende de una región local del tensor de entrada, preservando la topología geográfica original de los datos.

Esta propiedad es decisiva para el análisis de imágenes satelitales por tres razones concretas. Primero, la **localidad espacial**: la reflectancia espectral de un píxel en Nariño está fuertemente correlacionada con la de sus vecinos inmediatos debido a la continuidad de los ecosistemas (un páramo o un bosque andino ocupa zonas contiguas, no píxeles aislados). Segundo, la **invarianza traslacional**: los patrones de vegetación pueden aparecer en diferentes posiciones geográficas entre fechas, y la convolución los detecta independientemente de su ubicación. Tercero, la **eficiencia paramétrica**: los pesos del kernel se comparten espacialmente, lo que reduce drásticamente el número de parámetros respecto a una LSTM que procesara cada píxel por separado, haciendo el modelo manejable computacionalmente para grillas geográficas de alta resolución.

Salud de la vegetación: la salud de la vegetación se refiere al estado fisiológico de las plantas, en términos de su crecimiento, vigor, densidad y resistencia al estrés. Una vegetación saludable presenta altos niveles de clorofila, estructura foliar densa y respuesta positiva a factores ambientales. Este concepto se operacionaliza a través de índices espectrales como el NDVI, que reflejan su condición a partir de parámetros cuantificables.

Gestión ambiental basada en datos: este concepto hace referencia al uso de tecnologías de información y análisis cuantitativo para la toma de decisiones ambientales. Parte del principio de que la incorporación de datos objetivos y modelos científicos mejora la capacidad de intervención oportuna frente a problemas ecológicos, como la degradación vegetal. La integración de inteligencia artificial y datos satelitales permite diseñar estrategias de monitoreo, conservación y uso sostenible del territorio.

C. VARIABLES DEL ESTUDIO

Variable dependiente:

Salud de la vegetación (representada principalmente mediante valores de NDVI futuros y su evolución temporal) [33], R^2 (Coeficiente de determinación), Tiempo de predicción [33].

Variable independiente:

Algoritmos Machine Learning Tradicionales

- SVM (Support Vector Machine)
- Random forest
- KNN (K-nearest neighbors)

Deep Learning

- LSTM (Long Short-Term Memory)
- Técnicas de preprocesamiento

Variables de control:

- Tipo de cobertura vegetal.
- Estación o época del año.
- Localización geográfica (región dentro del departamento de Nariño).
- Condiciones climáticas históricas.

Definición nominal de variables

Variable dependiente

- **Salud vegetal NDVI:** se refiere al estado funcional y fisiológico de las plantas, evaluado a través de indicadores biofísicos como el contenido de clorofila, la cobertura foliar, el vigor de crecimiento y la capacidad de respuesta al estrés ambiental (Singh & Srivastava, 2021) . Desde la perspectiva de la teledetección, la salud vegetal se mide mediante índices espectrales como el NDVI (Normalized Difference Vegetation Index) o el VCI (Vegetation Condition Index), que permiten cuantificar la reflectancia del dosel vegetal en bandas específicas del espectro

electromagnético, particularmente el rojo y el infrarrojo cercano (EOS, 2023; Auravant, 2023) [2].

- **Coefficiente de determinación:** (R^2) es una medida estadística que cuantifica la proporción de la varianza de la variable dependiente explicada por las variables independientes en un modelo de regresión.
- **Tiempo de predicción:** predice resultados futuros analizando patrones en datos históricos con al menos dos ciclos temporales (periodos divididos, como meses en años).

Variable independiente:

Se refieren a los datos espectrales obtenidos de sensores satelitales multispectrales (MODIS, Landsat-8), expresados a través de las bandas de reflectancia y el índice NDVI histórico calculado a partir de las bandas roja e infrarroja cercana

- **SVR (Máquinas de Vectores de Soporte):** Es un algoritmo de aprendizaje supervisado que emplea para clasificar los diferentes estados de la vegetación. Su funcionamiento consiste en encontrar la frontera matemática que mejor separa las categorías de cobertura vegetal (por ejemplo, zonas sanas vs. zonas con estrés hídrico) en el espacio de datos espectrales.
- **Random Forest:** Es un algoritmo ampliamente utilizado que combina los resultados de múltiples árboles de decisión para producir una única predicción más confiable
- **KNN (K-Vecinos más Cercanos):** Es un algoritmo que clasifica un dato nuevo comparándolo con los datos más similares que ya conoce.

Deep learning

- **LSTM:** Es un tipo de red neuronal diseñada para recordar información a lo largo del tiempo, lo que la hace ideal para trabajar con datos en secuencia. En este proyecto es especialmente útil para analizar cómo evoluciona el NDVI mes a mes o quincena a quincena en los ecosistemas de Nariño, capturando patrones de cambio vegetal a lo largo del período 2015–2024.
- **ConvLSTM2D:** Arquitectura de aprendizaje profundo que integra operaciones de convolución espacial dentro de las ecuaciones de una celda LSTM, permitiendo

procesar simultáneamente información espacial y temporal de las imágenes satelitales. Este es el modelo principal utilizado en la investigación para pronosticar el NDVI futuro.

- **Técnicas de Preprocesamiento:** Antes de aplicar cualquier modelo, los datos del NDVI obtenidos de Landsat-8 y MODIS deben ser preparados adecuadamente. Esto incluye corregir valores faltantes por nubosidad, eliminar datos anómalos producto de fallas del sensor, normalizar los índices espectrales para que sean comparables entre períodos, y dividir las series temporales en conjuntos de entrenamiento y prueba. Estos pasos son fundamentales para garantizar que los modelos aprendan patrones reales de la vegetación y no errores del dato.

Métricas de evaluación del modelo

- **Coefficiente de determinación: (R^2)** es una medida estadística que cuantifica la proporción de la varianza de la variable dependiente explicada por las variables independientes en un modelo de regresión.
- **Tiempo de predicción:** predice resultados futuros analizando patrones en datos históricos con al menos dos ciclos temporales.

Definición operativa de variables

Variables Dependientes

- **La salud vegetal** se define como el estado fisiológico de la vegetación, evaluado mediante índices espectrales como el NDVI. Valores altos (mayores a 0.6) indican vegetación saludable, mientras que valores bajos reflejan estrés o degradación.
- **Coefficiente de determinación:** el R^2 siempre está entre 0 y 100%.

0% indica que el modelo no explica ninguna porción de la variabilidad de los datos de respuesta en torno a su media.

100% indica que el modelo explica toda la variabilidad de los datos de respuesta en torno a su media.

- **Tiempo de predicción:** el periodo requerido por un modelo para efectuar una predicción puede manifestarse en diferentes unidades de tiempo, tales como segundos, milisegundos, microsegundos o, en ciertas situaciones, minutos, en función de la complejidad del procedimiento.

Variable independiente

Algoritmos Machine Learning

SVM (Maquina de Vectores de Soporte): SVM fue utilizado en su variante de regresión (SVR) para estimar la Banda 1 de las imágenes Landsat a partir de coordenadas geográficas (latitud, longitud) y variables temporales codificadas, su objetivo es maximizar el margen entre las clases o minimizar el error de predicción manteniendo un margen de tolerancia controlado, lo que lo hace robusto ante datos con ruido o valores atípicos.

Ecuación 3. SVM

$$wx + b = 0$$

w es el vector de peso, x es el vector de entrada y b es el término de sesgo.

$$(wx_j + b)y_j \geq a$$

y luego se maximiza el margen: $\max \gamma = \|w\| a$, donde a es el margen proyectado sobre w .

RANDOM FOREST (Bosque Aleatorio): Random Forest es un algoritmo de conjunto (ensemble learning) que combina múltiples árboles de decisión entrenados de forma independiente sobre subconjuntos aleatorios de los datos de entrenamiento. Cada árbol realiza una predicción individual, y la predicción final del modelo se obtiene mediante votación mayoritaria (en clasificación) o promedio (en regresión).

Ecuación 4 clasificación

$$f(x) = \text{modo}\{hb(x): b = 1, 2, \dots, B\}$$

$f(x)$: predicción final.

$hb(x)$: predicción del árbol número b .

B : número total de árboles.

Regresión:

Ecuación 5 Regresión

$$f(x) = \sum_{b=1}^{B_1} B_b b(x)$$

$f(x)$: predicción final (promedio).

$hb(x)$: predicción del árbol número b .

$B1$: número total de árboles.

KNN (Vecinos más Cercanos): KNN es un algoritmo de aprendizaje basado en instancias que no construye un modelo explícito durante el entrenamiento, sino que almacena todos los datos de entrenamiento y realiza predicciones comparando el nuevo punto con sus vecinos más cercanos en el espacio de características. KNN fue utilizado como modelo auxiliar en la cadena de preprocesamiento para estimar las bandas espectrales 2, 3, 4, 5 y 7 de forma secuencial, tomando como entrada las bandas previamente estimadas. El algoritmo funciona calculando la distancia entre el punto a predecir y todos los puntos de entrenamiento, seleccionando los k puntos más cercanos, y asignando al nuevo punto el promedio de los valores de esos k vecinos (en regresión) o la clase mayoritaria (en clasificación). El parámetro k determina cuántos vecinos se consideran: valores pequeños ($k=4$) capturan patrones locales detallados, pero son sensibles al ruido, mientras que valores mayores ($k=51$) generan predicciones más suavizadas, pero pueden perder detalles finos.

Ecuación 6 KNN

$$f(x) = \frac{1}{k} \sum_{i=1}^k y_i$$

$f(x)$: predicción final para el punto x .

x : punto nuevo que quieres clasificar o predecir.

i : cada uno de los puntos de entrenamiento (vecinos).

y_i : valor real del punto. (Clase si es clasificación, número si es regresión.)

k : número de vecinos más cercanos que se consideran.

Deep Learning

ConvLSTM (Convolutional Long Short-Term Memory 2D): ConvLSTM2D es una arquitectura de aprendizaje profundo que integra operaciones de convolución espacial dentro de las ecuaciones de una celda LSTM (Long Short-Term Memory). A diferencia de una LSTM estándar, que procesa secuencias unidimensionales mediante multiplicaciones matriciales, la ConvLSTM2D reemplaza estas operaciones por convoluciones espaciales, lo que permite procesar simultáneamente información espacial (mapas bidimensionales de imágenes satelitales) e información temporal (evolución de esos mapas a lo largo del tiempo). Esta arquitectura es especialmente adecuada para datos geoespaciales como las imágenes satelitales multiespectrales, donde cada píxel tiene correlación espacial con sus vecinos y evolución temporal en secuencias de fechas consecutivas [55].

El funcionamiento de ConvLSTM2D se basa en cuatro componentes principales, implementados mediante convoluciones en lugar de multiplicaciones matriciales:

Forget Gate: decide qué parte de la memoria previa olvidamos:

Ecuación 7 Forget Gate

$$f_t = \sigma(w_f \cdot [h_{t-1}, x_t] + b_f)$$

Input Gate: decide qué parte de la nueva información almacenamos:

Ecuación 8 Input Gate

$$i_t = \sigma(w_i \cdot [h_{t-1}, x_t] + b_i)$$

Cell State: actualiza con nueva información:

Ecuación 9 Cell State

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c)$$

Puerta de salida (Output Gate): decide qué parte de la memoria mostramos como salida:

Ecuación 10 Output Gate

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t$$

En nuestra investigación, el modelo ConvLSTM2D procesa secuencias de imágenes satelitales del departamento de Nariño (ventanas de 3, 7, 15 o 30 días) con las siete bandas espectrales como canales de entrada, y predice el mapa espacial del índice NDVI futuro.

- **TÉCNICAS DE PREPROCESAMIENTO:**

Normalización min – máx.: escala todos los valores de una característica entre 0 y 1.

Ecuación 11 Normalización

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}}$$

x : Valor original.

x_{min} , x_{max} : mínimo y máximo de la característica.

Estandarización (Z-score normalization): convierte los datos para que tengan media 0 y desviación estándar 1.

Ecuación 12 Estandarización

$$x' = \frac{x - \mu}{\sigma}$$

μ : media de los datos.

σ : desviación estándar.

Imputación simple de datos faltantes (Media): cuando falta un valor x_i , se puede imputar usando la media:

Ecuación 13 Imputación

$$\bar{x} = \frac{1}{n} \sum_{j=1}^n x_j$$

Promediando todos los demás valores existentes.

D. FORMULACIÓN DE HIPOTESIS**1) Hipótesis de investigación**

Hi: El modelo de aprendizaje automático entrenado con datos de imágenes multiespectrales (Landsat-8 y MODIS) permite pronosticar con precisión la salud futura de la vegetación en Nariño. Es decir, existe al menos un modelo ML capaz de estimar valores de NDVI (Índice Normalizado de Vegetación) con un error aceptable.

2) Hipótesis nula

Ho: Los modelos de aprendizaje automático aplicados a las imágenes multiespectrales no logran pronosticar significativamente mejor la salud de la vegetación que un modelo no predictivo o aleatorio. En consecuencia, no existiría relación significativa entre las características espectrales pasadas y la salud vegetal futura que permita una predicción útil.

3) *Hipótesis alterna*

Ha: Solo en ciertos escenarios o con algoritmos específicos se logrará un pronóstico satisfactorio de la salud vegetal. Es decir, el modelo de ML solo pronostica adecuadamente la vegetación bajo condiciones particulares (por ejemplo, ciertos tipos de cobertura o rangos estacionales) y falla en otros casos. Esto implica que la hipótesis general no se cumple universalmente, sino solo de manera parcial según variables de contexto, por ejemplo: El modelo solo tiene éxito en predecir NDVI en ecosistemas como páramos, pero no en zonas costeras

III. METODOLOGÍA

A. PARADIGMA

El proyecto se enmarcó en el paradigma positivista. Se asumió la existencia de relaciones objetivas y cuantificables entre los datos espectrales y la salud de la vegetación, y se buscó medir esas relaciones empíricamente [1]. A través de la recopilación y análisis de datos numéricos (imágenes satelitales históricas), se procurará formular leyes predictivas o modelos matemáticos que expliquen los fenómenos observados.

B. ENFOQUE

El enfoque fue cuantitativo. Se emplearon mediciones numéricas (valores de reflectancia, índices de vegetación, métricas de desempeño) y procedimientos estadísticos para evaluar la precisión de los modelos. Todas las etapas (preprocesamiento, entrenamiento del modelo, validación) generaron datos cuantitativos que fueron analizados objetivamente. No habrá enfoque cualitativo (entrevistas o percepciones subjetivas), ya que el estudio se basa en datos espaciales medibles.

C. MÉTODO

La investigación siguió un método científico analítico-explicativo. Primero se recolectaron y procesaron datos, luego se formuló el modelo predictivo (hipótesis de correspondencia entre entrada y salida), se entrena y valida ese modelo con datos conocidos (experimentación), y al final se evalúa su rendimiento (análisis de resultados). el proceso fue iterativo (ciclo CRISP-DM o KDD), ajustando el modelo con base en la retroalimentación obtenida. Dado que el fin último es evaluar la capacidad predictiva de los algoritmos, el método global puede considerarse como una experimentación con datos reales[23].

D. TIPO DE INVESTIGACIÓN

El tipo de investigación fue aplicada y explicativa. Fue aplicada porque buscó resolver un problema concreto relacionado con la gestión ambiental en Nariño mediante herramientas tecnológicas. Fue explicativa porque pretendió identificar y modelar las relaciones entre variables espectrales e indicadores de salud vegetal, explicando cómo estas influyen en los patrones temporales de la vegetación [11].

E. DISEÑO DE LA INVESTIGACIÓN

El diseño de la investigación es cuantitativo, aplicado y tipo experimental con enfoque predictivo, ya que se trabaja con registros históricos recopilados entre los años 2015 y 2024, sin manipular variables ni intervenir directamente en los ecosistemas [46]. El estudio se basa en el análisis de datos satelitales preexistentes, utilizando herramientas computacionales para extraer patrones y generar modelos predictivos.

El enfoque metodológico se fundamenta en el uso de técnicas de inteligencia artificial para el pronóstico de la vegetación mediante aprendizaje automático aplicado a imágenes satelitales. Para ello, se estructuraron los datos en ventanas temporales de observación, permitiendo que el modelo analice secuencias históricas de imágenes y prediga el comportamiento futuro de la vegetación.

Las redes ConvLSTM combinan la capacidad de las redes convolucionales para extraer características espaciales con el análisis temporal de secuencias de datos, lo que las hace adecuadas para el procesamiento de información satelital multitemporal. Este enfoque permite predecir mapas de vegetación a partir de ventanas temporales de imágenes históricas, evaluando diferentes periodos de observación (3, 7, 15 y 30 días) para determinar el mejor rendimiento del modelo.

Se implementó el modelo de inteligencia artificial mediante técnicas de aprendizaje profundo basadas en redes neuronales convolucionales recurrentes, las cuales permiten analizar la evolución temporal y espacial de los índices de vegetación. Este tipo de modelos es adecuado para el análisis de imágenes satelitales, ya que combina la capacidad de las redes convolucionales para extraer características espaciales con el análisis temporal de secuencias de datos.

El modelo procesa las bandas espectrales e índices de vegetación como entradas, generando mapas predictivos que permiten estimar la distribución futura de la vegetación en la región andina de Nariño.

F. POBLACIÓN

La población objeto de estudio está compuesta por todos los registros satelitales multiespectrales relacionados con la cobertura vegetal en el Departamento de Nariño, capturados por los sensores de los satélites Landsat y MODIS durante los periodos comprendidos entre 2003-2015 (Landsat) y 2005-2015 (MODIS). Esta población comprende más de 20 millones de observaciones espectrales (4.071.185 para Landsat y 18.809.970 para MODIS), que incluyen datos de bandas espectrales, índices de vegetación, y coordenadas geográficas.

El conjunto de datos final se organizó en archivos en formato CSV, para luego ser manipulado con el modelo de inteligencia artificial mediante técnicas de Deep Learning basadas en redes neuronales convolucionales recurrentes (ConvLSTM), las cuales permiten analizar secuencias temporales de imágenes satelitales y generar predicciones espaciales de la vegetación.

G. MUESTRA

La muestra fue seleccionada a partir de la base total de datos mediante técnicas de muestreo aleatorio estratificado y suavizado por promedios móviles, buscando garantizar una representación equilibrada de las diferentes coberturas vegetales, épocas del año y regiones geográficas del departamento de Nariño. El tamaño muestral se definió considerando la capacidad computacional disponible y la necesidad de mantener un balance entre precisión del modelo y eficiencia de procesamiento. De este modo, se extrajeron subconjuntos de datos que contienen secuencias temporales completas y valores válidos de índices como NDVI o VCI, permitiendo entrenar y validar los modelos de predicción sin sesgos significativos.

H. TÉCNICAS DE RECOLECCIÓN DE INFORMACIÓN

En lugar de una recolección primaria de información, este proyecto se fundamentó en la utilización de bases de datos preexistentes, validadas por la comunidad científica para la evaluación de

métodos de Forecasting. Estos datasets, derivados de plataformas como Google Earth Engine, USGS Earth Explorer y Copernicus, proporcionaron series temporales de imágenes satelitales georreferenciadas de la región andina del departamento de Nariño.

Validez de la técnica

La validez de las técnicas empleadas en la recolección de datos se sustenta en el uso de fuentes satelitales reconocidas internacionalmente, como Landsat y MODIS, las cuales han sido ampliamente validadas en estudios ambientales y geoespaciales [34]. Estas plataformas, operadas por la NASA y el USGS, proporcionan datos científicamente validados mediante procesos rigurosos de calibración radiométrica y corrección atmosférica [29].

En el contexto colombiano diversos estudios han demostrado la validez de estas fuentes para el análisis de vegetación, Etter et al. (2006) utilizaron imágenes Landsat para mapear la deforestación en la región Andina durante el periodo 1985–2000, obteniendo una precisión global de clasificación del 87.3% validada mediante puntos de verificación en campo [56].

Armenteras et al. (2013) emplearon datos MODIS para monitorear incendios forestales en la Amazonia colombiana, reportando que el producto de detección de puntos de calor (hot spots) presenta una precisión del 91% comparado con reportes oficiales del IDEAM. Estos antecedentes validan técnicamente el uso de Landsat y MODIS para el análisis de vegetación en ecosistemas colombianos [57].

Landsat ofrece resolución espacial de 30 metros en las bandas multiespectrales, lo que permite observación detallada de la vegetación en paisajes heterogéneos como el departamento de Nariño. MODIS complementa este análisis con frecuencia de adquisición diaria y cobertura global, permitiendo monitoreo continuo de variaciones temporales estacionales. Además, los datos provenientes de ambas fuentes son de acceso público y gratuito a través de plataformas como Google Earth Engine, USGS Earth Explorer y NASA EARTHDATA, lo que garantiza transparencia y reproducibilidad en los procesos investigativos. Esta disponibilidad fomenta la validación cruzada entre diferentes estudios y permite que la comunidad científica replique y verifique los resultados obtenidos, fortaleciendo su aplicabilidad científica[20]. La posibilidad

de integrar estas imágenes con otras fuentes de datos geoespaciales mediante técnicas de análisis espacial y temporal incrementa la precisión de los modelos predictivos desarrollados. En conjunto, estos factores confirman que las técnicas de recolección utilizadas poseen validez científica sólida para el estudio del comportamiento espectral de la vegetación y el pronóstico de su salud mediante aprendizaje profundo.

Confiabilidad de la técnica

La confiabilidad de los instrumentos empleados en esta investigación se fundamenta en la consistencia y calidad comprobada de los datos satelitales utilizados. Tanto Landsat como MODIS han sido objeto de múltiples investigaciones en distintos campos, donde sus datos han demostrado ser precisos, estables en el tiempo y adecuados para el monitoreo ambiental. La estructura sistemática de recolección de datos, junto con el respaldo institucional de agencias como NASA y USGS, asegura que las mediciones sigan estándares rigurosos de calibración y control de calidad.

La confiabilidad se refleja también en la posibilidad de replicar los resultados obtenidos. Los datos históricos disponibles que abarcan más de una década y permiten entrenar modelos de aprendizaje automático con series temporales completas, lo que a su vez incrementa la solidez del análisis estadístico y reduce la posibilidad de errores aleatorios. Esta continuidad temporal reduce la incertidumbre asociada a errores aleatorios y permite capturar la variabilidad interanual de la vegetación en ecosistemas con alta estacionalidad climática como los del departamento de Nariño [43].

Finalmente, la versatilidad de estas bases de datos permite que se aplique en múltiples disciplinas, desde la agricultura hasta el análisis del cambio climático, lo cual respalda su capacidad para generar información coherente, confiable y útil para fines científicos. Por lo anterior, se concluye que las técnicas usadas de recolección garantizan una alta confiabilidad en el desarrollo del presente estudio [44][45].

I. INSTRUMENTO DE RECOLECCIÓN DE DATOS

Para la recolección de información se emplearon bases de datos satelitales públicas, principalmente las proporcionadas por los programas Landsat y MODIS, accesibles a través de plataformas como Google Earth Engine [13]. Estas herramientas permiten extraer, filtrar y procesar imágenes multiespectrales con resolución temporal adecuada para el análisis longitudinal.

Se desarrollaron scripts en Python para automatizar la descarga y el preprocesamiento de datos. Es importante precisar que las variables topográficas y climáticas se integran al modelo de manera intrínseca a través de la ubicación geográfica (Latitud/Longitud) y la respuesta espectral de las bandas. Dado que la topografía de Nariño influye directamente en la reflectancia capturada por los sensores Landsat/MODIS, el modelo ConvLSTM aprende a reconocer patrones de vegetación asociados a relieves específicos mediante la relación entre las coordenadas espaciales y las firmas espectrales de cada píxel [12][36][37].

J. METODOLOGÍA DE DESARROLLO

a) Preprocesar y limpiar imágenes satelitales multiespectrales (Landsat y MODIS) para generar un conjunto de datos estructurado, espacial y temporalmente consistente

En esta fase se realizó la obtención, organización y limpieza de imágenes satelitales provenientes de sensores multiespectrales, con el fin de construir un conjunto de datos consistente a nivel espacial y temporal.

Inicialmente, se recolectaron imágenes de los satélites Landsat y MODIS, las cuales contienen información en diferentes bandas espectrales. Posteriormente, se llevó a cabo un proceso de organización de los datos en estructuras tipo carpeta, clasificadas según fecha y tipo de sensor. A partir de estas imágenes, se extrajeron los valores de reflectancia correspondientes a cada banda espectral, junto con la información geográfica (latitud y longitud) y temporal (fecha de captura). Estos datos fueron transformados a un formato tabular, permitiendo su manipulación mediante herramientas computacionales.

Durante el preprocesamiento se realizaron tareas de limpieza de datos, incluyendo el manejo de valores faltantes y la prevención de errores en cálculos posteriores, como divisiones por cero en el índice NDVI. Finalmente, se generó un conjunto de datos estructurado en formato CSV, el cual

integra información espectral, espacial y temporal, sirviendo como base para el entrenamiento del modelo.

TABLA III. Proceso de preprocesamiento

Etapa	Actividad	Herramienta	Resultado
Recolección	Descarga de imágenes satelitales	Google Earth Engine / USGS	Imágenes Landsat y MODIS
Organización	Clasificación por carpetas	Sistema de archivos	Estructura ordenada
Extracción	Obtención de bandas espectrales	Python	Datos numéricos
Limpieza	Manejo de valores nulos	NumPy / Pandas	Datos consistentes
Estructuración	Generación de dataset	CSV	Dataset final

b) Diseñar y entrenar un modelo de inteligencia artificial para analizar la vegetación a partir de imágenes satelitales.

En esta etapa se diseñó e implementó un modelo de inteligencia artificial orientado al análisis de la vegetación a partir de datos satelitales. Inicialmente, se cargó el conjunto de datos previamente estructurado en un entorno de desarrollo en Python. Posteriormente, se realizó la preparación de los datos, incluyendo la transformación de la variable temporal en componentes como año, mes y día.

Para el modelado, se implementó una arquitectura basada en redes neuronales convolucionales recurrentes (ConvLSTM), la cual permite capturar patrones tanto espaciales como temporales presentes en los datos. Además, se aplicó una técnica de ventanas deslizantes sobre la serie temporal, permitiendo estructurar los datos como un problema supervisado, donde se utilizan valores históricos para predecir estados futuros.

El modelo fue entrenado utilizando subconjuntos de datos de entrenamiento, permitiendo ajustar sus parámetros y aprender patrones presentes en la información.

TABLA IV. Configuración de modelo

Parámetro	Descripción
Tipo de modelo	ConvLSTM
Entrada	Datos espaciales y temporales
Técnica	Ventanas deslizantes
Lenguaje	Python
Librerías	TensorFlow / Keras

c) Evaluar la precisión y la robustez del modelo de pronóstico de la salud de la vegetación, utilizando métricas apropiadas para su posterior análisis de calidad.

Para evaluar el desempeño del modelo, se emplearon diferentes configuraciones de ventanas temporales (3, 7, 15 y 30 días), con el fin de analizar su impacto en la capacidad predictiva.

Se dividió el conjunto de datos en subconjuntos de entrenamiento y prueba, permitiendo evaluar la capacidad de generalización del modelo.

Las predicciones generadas por el modelo fueron comparadas con los valores reales, utilizando el coeficiente de determinación (R^2) como métrica principal de evaluación.

Adicionalmente, se analizaron visualmente los resultados mediante la comparación entre mapas reales y predichos, lo que permitió observar el comportamiento espacial del modelo.

TABLA V. Proceso de evaluación

Etapas	Actividad	Resultado
División de datos	Train/Test	Validación de modelo
Predicción	Model.predict()	Valores estimados
Evaluación	R^2 score	Métrica de desempeño
Comparación	Mapas	Validación visual

IV. RESULTADOS DE LA INVESTIGACIÓN

Objetivo 1.

Preprocesar y limpiar imágenes satelitales multiespectrales (Landsat y MODIS) para generar un conjunto de datos estructurado, espacial y temporalmente consistente

En este apartado se presentan los resultados obtenidos durante el proceso de preprocesamiento de las imágenes satelitales Landsat y MODIS, los datos satelitales fueron proporcionados por el grupo de investigación GIEEE de la Universidad de Nariño, cuyo propósito fue generar un conjunto de datos estructurado que permitió el posterior entrenamiento de modelos de aprendizaje automático. A partir del procesamiento de las bandas espectrales se construyó un dataset que integra información geográfica y espectral relevante para el análisis de la vegetación.

Resultado 1 – Consolidación del Dataset Estructurado

Se integraron múltiples archivos ráster organizados cronológicamente en un único repositorio de datos. El resultado principal es el archivo `landsat_model.csv`, el cual vincula coordenadas geográficas (Latitude y Longitude) con firmas espectrales, eliminando inconsistencias y valores nulos mediante técnicas de limpieza de datos.

Resultado 2 – Integración de bandas espectrales

Como resultado del procesamiento de las imágenes satelitales se obtuvo un conjunto de variables correspondientes a las diferentes bandas espectrales capturadas por el sensor Landsat. Estas bandas permiten representar diferentes características de la superficie terrestre, lo que resulta fundamental para el análisis de la vegetación.

Las bandas espectrales integradas en el dataset fueron:

TABLA VI. Bandas espectrales

banda	Descripción	Aplicación en proyecto
Banda 1	Banda azul	Detección de atmósfera y cuerpos de agua.
Banda 2	Banda verde	Evaluación del vigor de la vegetación.

Banda 3	Banda roja	Absorción de clorofila (crítico para NDVI).
Banda 4	Infrarojo cercano	Reflectancia de la estructura celular de las hojas.
Banda 5	SWIR 1	Contenido de humedad en la vegetación.
Banda 6	Banda térmica	Estrés hídrico y temperatura superficial.
Banda 7	Infrarojo medio	Geología y humedad del suelo.

Resultado 3 – Generación del índice NDVI

A partir de las bandas 3 (Rojo) y 4 (NIR), se generó el Índice de Vegetación de Diferencia Normalizada (NDVI) como variable objetivo. Este índice permitió estandarizar la medición de la biomasa en la región de Nariño, sirviendo como la variable objetivo para los modelos de aprendizaje profundo.

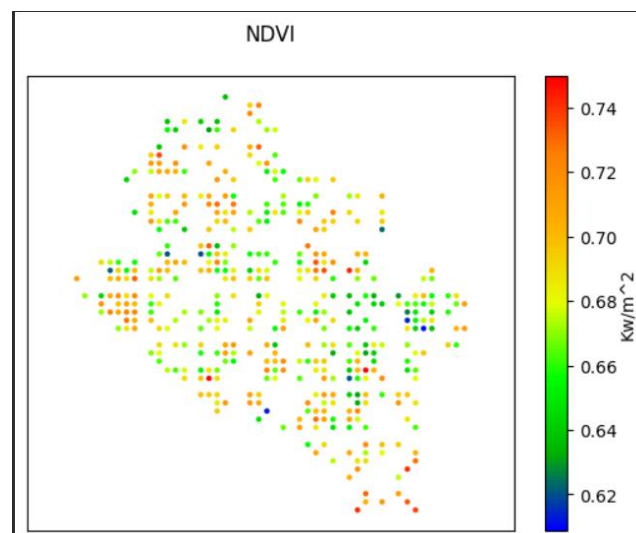


Fig 4. Calculo NDVI

A partir de las bandas espectrales procesadas, se generó la capa de NDVI. Este resultado representa la condición base de la vegetación antes de ser procesada por el modelo predictivo.

3.2. Normalización y Ajuste de Valores

Posteriormente, se realizó un ajuste de los valores obtenidos para asegurar la compatibilidad con los tensores de entrada de la red neuronal ConvLSTM. Esta visualización permite identificar la intensidad de la respuesta espectral en una escala normalizada.

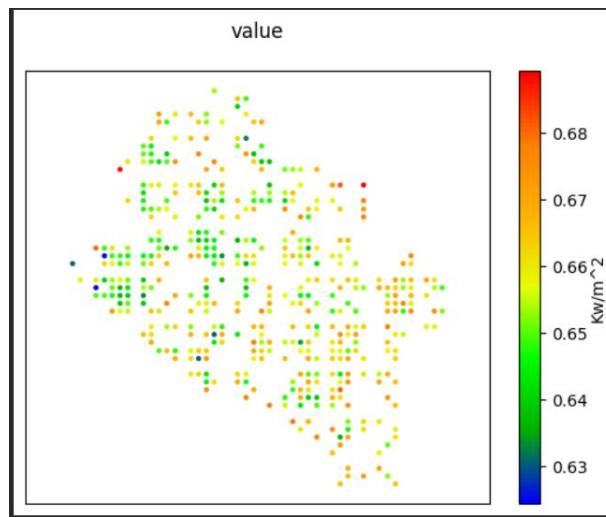


Fig 5. Mapa de predicción del índice NDVI obtenido mediante la arquitectura ConvLSTM

Resultado 4 – Visualización espacial del NDVI

Como resultado del procesamiento de los datos, se generaron diversas visualizaciones espaciales del índice de vegetación de diferencia normalizada (NDVI). Estas representaciones gráficas facilitan la identificación de áreas con distintos niveles de vegetación. Adicionalmente, se generaron mapas correspondientes a los valores mínimo, promedio y máximo del NDVI.

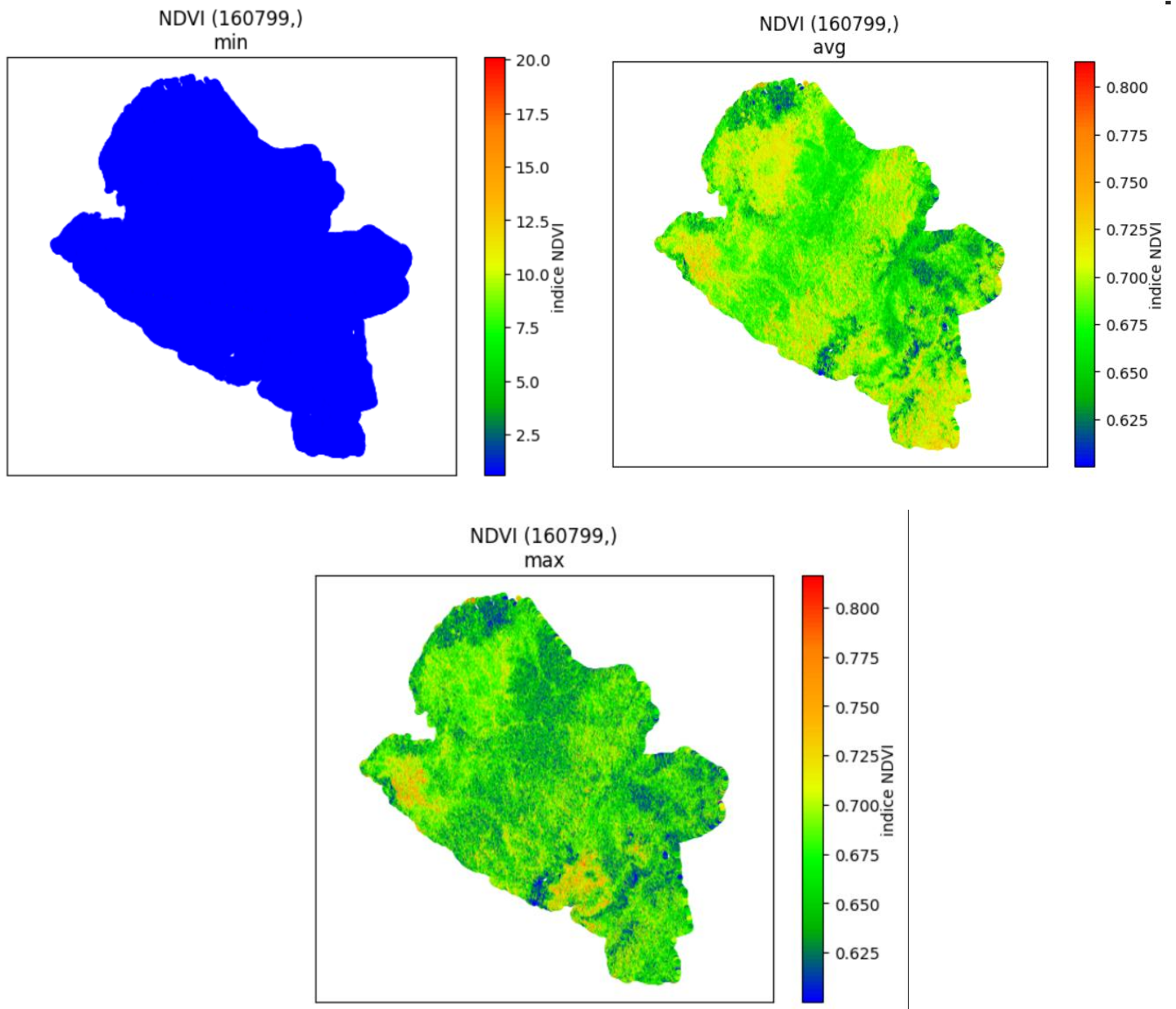
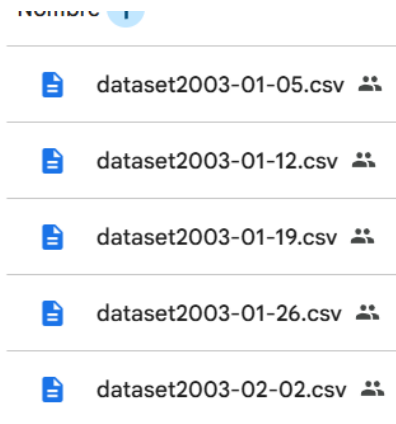


Fig 6. Mapas de NDVI correspondientes a los valores mínimo, promedio y máximo, generados a partir de las bandas espectrales.

Resultado 5 – Consolidación del dataset final

Como resultado final del proceso de preprocesamiento y limpieza de datos se obtuvo un dataset consolidado que integra información geográfica, bandas espectrales y valores asociados al índice de vegetación. Este conjunto de datos constituye la base para el entrenamiento de los modelos.



dataset2003-01-05.csv
dataset2003-01-12.csv
dataset2003-01-19.csv
dataset2003-01-26.csv
dataset2003-02-02.csv

Fig 7. Dataset final consolidado utilizado para el entrenamiento de los modelos de inteligencia artificial.

Objetivo 2

Diseñar y entrenar modelos de inteligencia artificial para analizar la vegetación a partir de imágenes satelitales

Resultado 1 Preparación del conjunto de datos para el entrenamiento

En este apartado se presentan los resultados del diseño, entrenamiento y evaluación de los modelos de Machine Learning implementados en la investigación.

Se entrenaron dos categorías de modelos:

1. Modelos auxiliares de preprocesamiento basados en algoritmos tradicionales de Machine Learning (SVR, KNN) para la estimación de bandas espectrales.
2. Modelo principal de pronóstico basado en aprendizaje profundo (ConvLSTM2D) para la predicción del índice NDVI futuro.

Resultado 1 Preparación del conjunto de datos base

Como resultado del proceso de preparación de los datos descrito en OE1, se obtuvo un conjunto de datos estructurado en formato CSV denominado landsat_model.csv, el cual fue utilizado como base para el entrenamiento del modelo de inteligencia artificial donde se almacena información geográfica, temporal y espectral asociada a cada observación, Este dataset sirvió como base tanto para el entrenamiento de los modelos auxiliares de estimación de bandas como para la construcción posterior del dataset de secuencias temporales requerido por el modelo ConvLSTM2D.

Resultado 2 Entrenamiento de modelos auxiliares SVR y KNN

Dado que las imágenes presentan vacíos temporales y espaciales debido a nubosidad persistente en la región de Nariño, se entrenaron modelos auxiliares con el propósito de estimar valores de bandas espectrales en puntos y fechas sin observación directa. Estos modelos constituyen una cadena secuencial donde cada banda se predice utilizando las bandas anteriormente estimadas como variables de entrada.

TABLA VII. Modelos auxiliares entrenados para estimación de bandas espectrales

MODELO (.PKL)	ALGORITMO	VARIABLES DE ENTRADA	R ²
latlondateb1.pkl	SVR	lat, lon, mes, día, año	-0.011
knnb1b2.pkl	KNN	Banda 1	0.765
knnb1b2b3.pkl	KNN	Bandas 1–2	0.873
knnb1b2b3b4.pkl	KNN	Bandas 1–3	0.685
knnb1b2b3b4b5.pkl	KNN	Bandas 1–4	0.804
knnb1b2b3b4b5b7.pkl	KNN	Bandas 1–5	0.947
svrb1b2b3b4b5b7b6.pkl	SVR	Bandas 1–5, 7	0.507

Estos modelos fueron serializados en formato.pkl mediante la librería joblib y utilizados posteriormente en el proceso de generación del dataset.

Resultado 3 Preparación de tensores y ventanas temporales del modelo principal

La estructura tabular del conjunto de datos facilitó la manipulación de la información y la selección de variables relevantes para el proceso de modelado, permitiendo organizar cada observación como un registro independiente con múltiples atributos asociados.

latitude	longitude	band1	band2	band3	band4	band5	band6	band7	value
-878950	184950	0.1051125905	0.08271564244	0.05275554907	0.3155197183	0.1326408839	296.1786069	0.04458619665	203.4
-8784900	174600	0.1048404213	0.08573163966	0.05352727115	0.2714632767	0.09024746761	295.9697098	0.02729183302	198.3
-8775000	164700	0.1034676758	0.0791469091	0.04893191661	0.2486185447	0.07859331321	296.8426895	0.02209766349	203.1
-8775000	169650	0.108848644	0.0906645707	0.05571492217	0.338889248	0.1282274576	296.1239273	0.03979753696	199.1
-8775000	174600	0.09652626522	0.07102843495	0.04224698498	0.1947884317	0.05446750733	294.4074912	0.01771206647	196.3
-8775000	194850	0.1147502456	0.09446627902	0.06208231261	0.3389840959	0.1331359444	296.8987684	0.04275377112	201.7
-8769600	164700	0.1060903788	0.08960354563	0.05748188957	0.3608179474	0.156995398	296.7202432	0.05395090965	199.4
-8769600	174600	0.109988866	0.09101283304	0.05940039255	0.3429740757	0.1552977377	296.4705482	0.05494264934	196.3
-8769600	184950	0.112143067	0.09338760201	0.05916869817	0.3338078995	0.1551032405	296.4747906	0.05534423227	196.4
-8769600	189900	0.107985264	0.08625670443	0.06454178054	0.2773909273	0.1370125929	297.818269	0.05217630288	196.7
-8769600	194850	0.1036768505	0.08979321575	0.05372435145	0.2759818265	0.09993877713	296.8413022	0.0302681538	198.5
-8764650	154800	0.1066772756	0.08476108172	0.05419595755	0.2744306307	0.1305969463	296.4884722	0.04570082924	198.2
-8764650	164700	0.1119534554	0.08625550102	0.06220420392	0.3628123198	0.1828757688	296.2969027	0.06422799926	198.3
-8764650	169650	0.112206715	0.09389568633	0.06226469248	0.3504757678	0.1518425228	295.3547427	0.05157909327	197.7
-8764650	174600	0.1153344945	0.09622097141	0.06280524034	0.3620510367	0.1867983931	297.0160779	0.0684296131	196.5
-8764650	184950	0.1117671188	0.09599087355	0.06139941867	0.3532152207	0.1693304707	296.8444691	0.05893423771	195.8
-8764650	189900	0.1136025276	0.0942984161	0.06085919825	0.3345190856	0.1409652715	296.4780601	0.04636379914	195
-8764650	194850	0.112423489	0.09276387166	0.05774485659	0.2819382625	0.08938216201	297.0004738	0.02645975353	197
-8759700	154800	0.1103077381	0.09676526729	0.0658706107	0.3558445139	0.1748237503	297.0034332	0.06649504896	197.6
-8759700	159750	0.1168904934	0.09818268454	0.06414801015	0.4080224686	0.1796602748	296.4675381	0.06493089944	197.9
-8759700	164700	0.1165599109	0.1022829428	0.06460499158	0.4029074407	0.1713797897	296.86354	0.06063726068	198.1
-8759700	169650	0.1113805861	0.09228768064	0.05855682339	0.3474147249	0.1477163068	296.3873813	0.05147172256	197.3
-8759700	174600	0.1169029508	0.09652169661	0.06219626902	0.3372855706	0.147133574	297.1897846	0.05097097615	196.7
-8759700	184950	0.1120792964	0.09378747517	0.06062737806	0.3300701343	0.1673675378	296.7355377	0.06039988826	195.4
-8759700	189900	0.1048892888	0.08053579253	0.05041455525	0.2430590507	0.06822097949	297.3801608	0.01849027899	194.7
-8759700	244800	0.1073821226	0.08489924475	0.05413311356	0.2488814818	0.08752184987	297.1974048	0.0265018987	197.8

Fig 8. Estructura tabular del dataset multispectral utilizado para el entrenamiento del modelo

Para el diseño del modelo, se transformó el dataset landsat_model.csv en una estructura de tensores. El resultado principal fue la implementación de una función de ventana deslizante (*sliding window*), que organiza los datos en secuencias temporales de 3, 7, 15 y 30 días. Esta estructura permite que el modelo no solo analice un dato aislado, sino la evolución de la vegetación

```
#forecasting window
def window(X,y,v):
    yt=y[v:]
    Xt=[]
    for i in range(X.shape[0]-v):
        Xt.append(X[i:i+v])
    return np.array(Xt),yt
```

Fig 9. Función utilizada para la generación de ventanas temporales en el modelo ConvLSTM2D

Resultado 3 Implementación del modelo de inteligencia artificial ConvLSTM2D

Se diseñó una arquitectura de red neuronal profunda que integra capas ConvLSTM2D, las cuales procesan simultáneamente la información espacial (mapas) y la información temporal (secuencias).

- Entrada: Tensores con dimensiones (ventana temporal, latitud, longitud, bandas).
- Procesamiento: Capas de normalización por lotes (*BatchNormalization*) y capas de abandono (*Dropout* al 30%) para asegurar la estabilidad del entrenamiento.
- Salida: Una capa convolucional con activación lineal para la estimación del índice NDVI.

Layer (type)	Output Shape	Param #
input_layer (InputLayer)	(None, 3, 52, 49, 8)	0
conv_lstm2d (ConvLSTM2D)	(None, 52, 49, 32)	46,208
batch_normalization (BatchNormalization)	(None, 52, 49, 32)	128
dropout (Dropout)	(None, 52, 49, 32)	0
conv2d (Conv2D)	(None, 52, 49, 16)	4,624
batch_normalization_1 (BatchNormalization)	(None, 52, 49, 16)	64
dropout_1 (Dropout)	(None, 52, 49, 16)	0
conv2d_1 (Conv2D)	(None, 52, 49, 1)	145
Total params: 51,169 (199.88 KB)		
Trainable params: 51,073 (199.50 KB)		
Non-trainable params: 96 (384.00 B)		

Fig 10. Diagrama de la arquitectura del modelo diseñado para el análisis de vegetación.

Resultado 4 Monitoreo del proceso de entrenamiento y convergencia

Como parte de la fase de entrenamiento del modelo ConvLSTM, se realizó un monitoreo sistemático de la función de pérdida (Loss) y el error absoluto medio (MAE) a lo largo de las épocas programadas. Este procedimiento permitió verificar que la red neuronal ajustara sus pesos sinápticos de manera estable, alcanzando un estado de convergencia óptimo donde el error de entrenamiento y el de validación se estabilizaron.

Para asegurar la robustez del proceso, se implementó una técnica de parada temprana (Early Stopping), la cual finalizó el entrenamiento en el momento exacto en que el modelo dejó de presentar mejoras significativas, evitando así el sobreajuste de los datos.

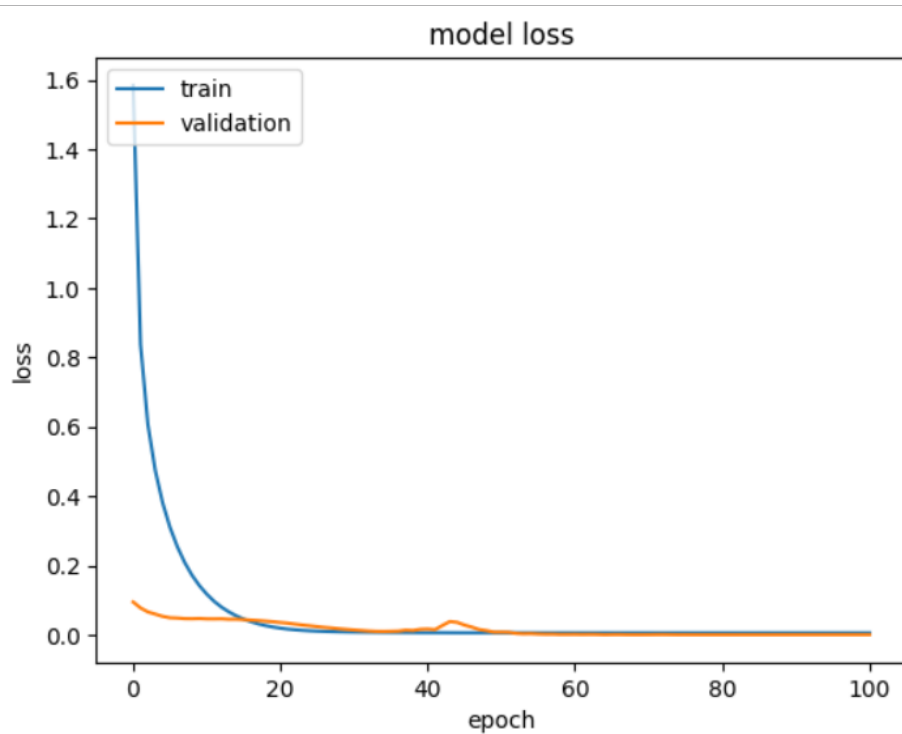


Fig 11. Curvas de aprendizaje (Loss) obtenidas durante el entrenamiento del modelo para la ventana temporal seleccionada.

Resultado 5 Generación del modelo serializado

Al finalizar el entrenamiento óptimo, el modelo fue exportado y almacenado en un archivo de formato model.keras. Este archivo contiene los pesos sinápticos y la arquitectura configurada, permitiendo realizar inferencias y predicciones sobre nuevas imágenes satelitales sin requerir un nuevo proceso de entrenamiento.

model.keras 3 elementos

	Nombre	Última modificación	Tamaño de arc...
	config.json	-	-
	metadata.json	-	-
	model.weights.h5	-	-

*Fig 12. model.keras almacenado en el repositorio del proyecto***Objetivo 3.**

Evaluar la precisión y la robustez de los modelos de pronóstico de la salud de la vegetación, utilizando métricas apropiadas para su posterior análisis de calidad.

Con el fin de evaluar la precisión y la robustez de todos los modelos de Machine Learning implementados durante las fases de preprocesamiento (OE1) y pronóstico (OE2), se realizó un análisis comparativo del desempeño utilizando el coeficiente de determinación (R^2) y el error absoluto medio (MAE) como métricas de evaluación. Los modelos se agrupan en dos categorías según su función en la investigación:

EVALUACIÓN DE MODELOS AUXILIARES DE PREPROCESAMIENTO

Los siguientes modelos fueron entrenados con el propósito de estimar bandas espectrales en puntos y fechas sin observación directa debido a nubosidad, cumpliendo una función de interpolación espectral como parte del pipeline de preprocesamiento descrito en el OE1:

TABLA VIII. Desempeño de modelos auxiliares de preprocesamiento

Modelo	Tipo	Variables de entrada	R^2
SVR	Regresión	lat, lon, mes, día, año	-0.011
KNN	Regresión	Banda 1	0.765
KNN	Regresión	Bandas 1–2	0.873
KNN	Regresión	Bandas 1–3	0.685
KNN	Regresión	Bandas 1–4	0.804
KNN	Regresión	Bandas 1–5	0.947
SVR	Regresión	Bandas 1–5, 7	0.507

Los modelos KNN presentaron desempeños de R^2 entre 0.685 y 0.947, siendo la estimación de la Banda 7 la más precisa. El modelo SVR inicial presentó R^2 negativo, lo que indica que la ubicación geográfica por sí sola no es un predictor suficiente de la reflectancia en la banda azul. Sin embargo, una vez estimada la Banda 1, los modelos KNN subsiguientes lograron interpolaciones espectrales exitosas al aprovechar las correlaciones entre bandas adyacentes.

EVALUACIÓN DEL MODELO PRINCIPAL DE PRONÓSTICO

El modelo ConvLSTM2D fue evaluado bajo cuatro configuraciones de ventana temporal deslizante para determinar el horizonte de pronóstico óptimo:

TABLA IX. Desempeño según ventana temporal

Modelo	Tipo	Variables de entrada	R^2	MAE
ConvLSTM2D (ventana 3d)	Pronóstico	Secuencia 3 DIAS	-0.174	0.0207
ConvLSTM2D (ventana 7d)	Pronóstico	Secuencia 7 DIAS	0.503	0.0216
ConvLSTM2D (ventana 15d)	Pronóstico	Secuencia 15 DIAS	0.699	0.0126
ConvLSTM2D (ventana 30d)	Pronóstico	Secuencia 30 DIAS	0.637	0.0151

El modelo ConvLSTM2D evidenció variabilidad significativa en su desempeño según la longitud de la secuencia temporal. El desempeño mejora progresivamente desde 3 días hasta alcanzar un máximo en 15 días, presentando posteriormente un ligero retroceso en la ventana de 30 días.

SELECCIÓN DEL MODELO ÓPTIMO

Con base en los resultados de la Tabla 9, se seleccionó el modelo ConvLSTM2D con ventana temporal de 15 días como la configuración óptima para el pronóstico de la salud de la vegetación en el departamento de Nariño. Esta selección se fundamenta en que dicha configuración alcanzó el mayor coeficiente de determinación ($R^2 = 0.699$) y el menor error absoluto medio ($MAE = 0.0126$) entre todas las variantes evaluadas. El error de 0.0126 en la escala NDVI representa una precisión inferior al 1.3%, lo que permite distinguir categorías ecológicas relevantes como vegetación densa ($NDVI > 0.6$) vs. vegetación degradada ($NDVI < 0.4$).

A partir de este punto, el análisis de robustez, validación espacial y generación de mapas comparativos se enfoca exclusivamente en esta configuración seleccionada.

Modelo optimo ConvLSTM2D

El modelo principal de nuestra investigación, ConvLSTM2D, se evaluó mediante experimentos con cuatro configuraciones de ventana temporal deslizante (3, 7, 15 y 30 días). El proceso de evaluación se llevó a cabo sobre un total de 1.842 imágenes satelitales, organizadas en estructuras supervisadas mediante la función $\text{window}(X, y, v)$. Las métricas empleadas fueron el Coeficiente de Determinación (R^2) y el Error Absoluto Medio (MAE).

Resultado 1 Tabla comparativa unificada de métricas por ventana temporal

A continuación, se presenta la tabla con todos los resultados obtenidos para las cuatro configuraciones evaluadas. Se unifica la información para facilitar la comparación directa entre ventanas temporales.

TABLA X. Desempeño comparativo del modelo convLSTM2D

Ventana Temporal	Nº Imágenes	R^2 (Test)	MAE (Test)	Interpretación
3 días	623	-0.174	0.0207	Inferior al modelo base; no logra capturar adecuadamente la varianza del NDVI
7 días	619	0.503	0.0216	Desempeño aceptable para predicción inicial de series temporales
15 días	611	0.699	0.0126	Precisión satisfactoria según rangos reportados para modelos LSTM aplicados a vegetación [55]
30 días	596	0.637	0.0151	Desempeño alto con ligera reducción respecto a la ventana óptima [55]

Resultado 2 Análisis técnico por ventana temporal

Ventana de 3 días ($R^2 = -0.174$)

El valor negativo de R^2 indica que el modelo en esta configuración predice peor que un modelo trivial que simplemente prediga el valor promedio del NDVI. Técnicamente, esto se explica por dos factores principales. Primero, el NDVI de la vegetación andina no experimenta cambios biológicos significativos en intervalos de tres días: los procesos de fotosíntesis, senescencia y respuesta al estrés hídrico ocurren en escalas de tiempo semanales o quincenales. Segundo, el ruido inherente al proceso de estimación de bandas (introducido por la cadena de modelos auxiliares KNN/SVR y la interpolación Kriging) domina la señal útil cuando la ventana es demasiado corta, haciendo que el modelo aprenda patrones de ruido en lugar de la dinámica real de la vegetación.

Ventana de 7 días ($R^2 = 0.503$)

Con siete días de historia temporal el modelo comienza a capturar patrones fenológicos reales, alcanzando una capacidad predictiva moderada. Este resultado sugiere que una semana ya es suficiente para que el ConvLSTM2D detecte tendencias de cambio en el NDVI, aunque la señal aún presenta ruido considerable asociado a la variabilidad climática de corto plazo (precipitaciones esporádicas, nubosidad transitoria).

Ventana de 15 días ($R^2 = 0.699$) Configuración óptima

La ventana de 15 días constituye el punto de equilibrio óptimo entre disponibilidad de información histórica y calidad de la señal espectral. Desde el punto de vista del ciclo biológico de la vegetación andina, dos semanas corresponden aproximadamente a un ciclo de respuesta fisiológica completo: tiempo suficiente para que los cambios en humedad del suelo, radiación solar y temperatura se manifiesten en la reflectancia del dosel. El MAE de 0.0126 sobre la escala NDVI (rango -1 a $+1$) representa un error inferior al 1.3%, lo que permite distinguir variaciones reales del estado vegetal de fluctuaciones de ruido. Esta configuración fue seleccionada como la definitiva para el modelo de pronóstico.

Ventana de 30 días ($R^2 = 0.637$)

Con una ventana de 30 días se observa una ligera disminución del desempeño respecto a los 15 días. Este comportamiento es consistente con el hecho de que, en la región andina de Nariño, los microclimas cambian rápidamente entre zonas (costa Pacífica, cordillera, piedemonte), de modo que información de hace un mes puede no ser relevante para predecir el estado actual de la vegetación. Adicionalmente, al aumentar el tamaño de la ventana disminuye el número de secuencias de entrenamiento disponibles (596 imágenes frente a 611 en la ventana de 15 días), lo que reduce la cantidad de ejemplos con los que el modelo puede aprender.

Resultado 3 Curvas de aprendizaje (Loss y MAE)

Para la ventana óptima de 15 días, las curvas de entrenamiento evidencian una convergencia estable: tanto la función de pérdida como el MAE disminuyen progresivamente en los conjuntos de entrenamiento y validación, sin divergencia entre ambas curvas. La activación de Early Stopping confirmó que el modelo alcanzó su punto de generalización máximo antes de iniciar un posible sobreajuste, preservando los pesos del mejor epoch de validación.

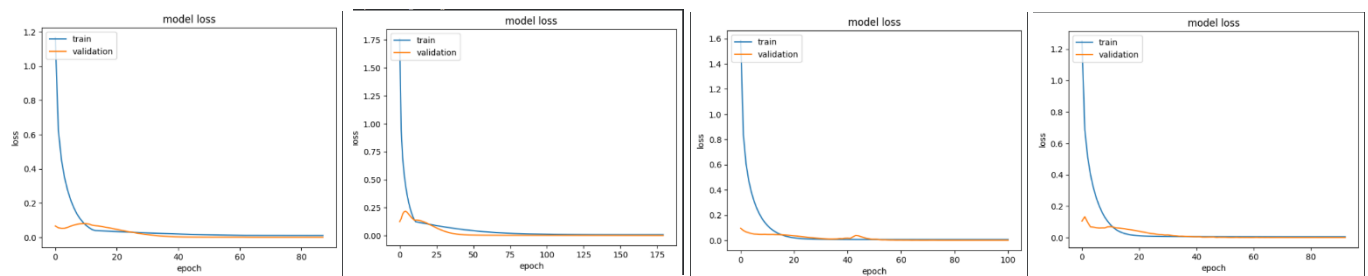


Fig 13 Curvas de aprendizaje (Loss) para las cuatro configuraciones de ventana temporal evaluadas. La ventana de 15 días presenta la convergencia más estable

Resultado 4 Comparación visual: mapas NDVI real vs. predicho

Como parte de la evaluación del modelo se generaron mapas comparativos entre el NDVI real (calculado a partir de las imágenes satelitales) y el NDVI predicho por el modelo ConvLSTM2D para cada una de las cuatro ventanas temporales evaluadas. Estos mapas permiten validar visualmente la fidelidad espacial de las predicciones sobre el territorio de Nariño.

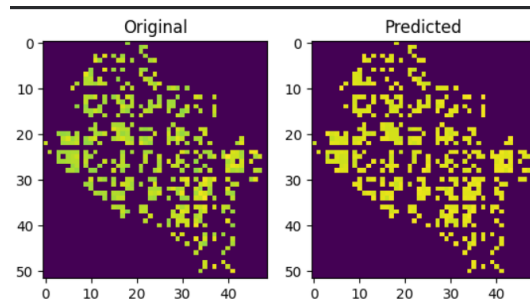


Fig 14 Mapas comparativos NDVI real vs. NDVI predicho para la ventana de 3 días. Se observa escasa correspondencia espacial, consistente con el R^2 negativo obtenido.

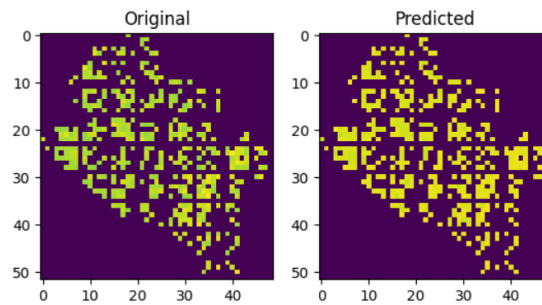


Fig 15 Mapas comparativos NDVI real vs. NDVI predicho para la ventana de 7 días. El modelo comienza a reproducir los patrones generales de distribución de la vegetación.

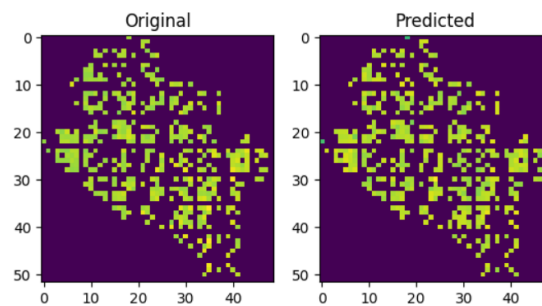


Fig 16 Mapas comparativos NDVI real vs. NDVI predicho para la ventana de 15 días (configuración óptima). Alta correspondencia espacial con los valores reales en la mayoría del departamento

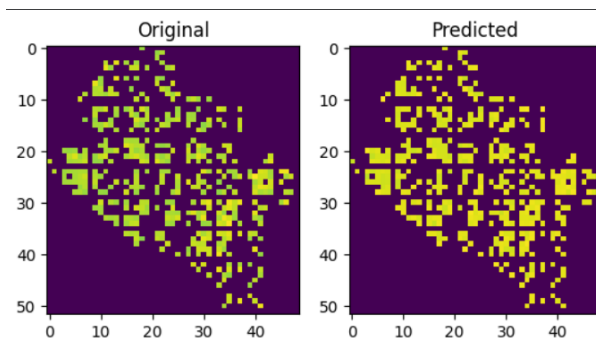


Fig 17 Mapas comparativos NDVI real vs. NDVI predicho para la ventana de 30 días. Desempeño ligeramente inferior al de 15 días, especialmente en las zonas costeras del Pacífico

Resultado 5 Sistema de visualización interactiva y validación de resultados

Con el fin de facilitar la interpretación de los resultados obtenidos por el modelo de predicción de la salud de la vegetación, se desarrolló un sistema de visualización interactivo basado en un dashboard implementado en Python mediante la librería Streamlit. Este sistema permite integrar, analizar y presentar de forma clara tanto los datos observados como las predicciones generadas por el modelo ConvLSTM.

El dashboard fue diseñado como una herramienta complementaria al modelo, orientada a la validación visual y cuantitativa de los resultados, permitiendo al usuario explorar la evolución temporal de la vegetación y comparar directamente los valores reales con los estimados.

Visualización de datos observados y predichos:

El sistema permite seleccionar una fecha específica dentro del conjunto de datos mediante un control interactivo. Para cada fecha seleccionada, se presentan de manera simultánea:

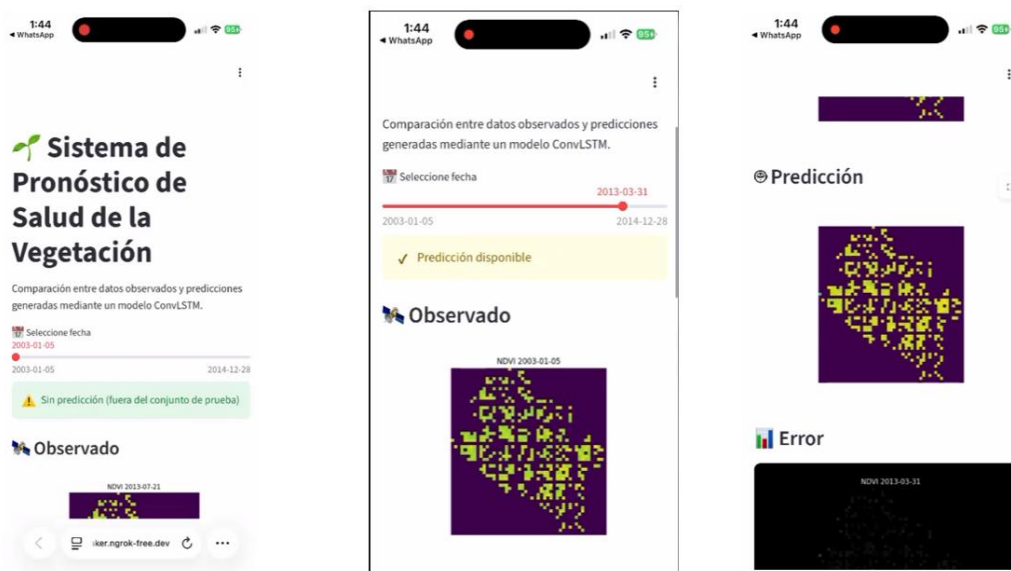


Fig 18 Dashboard correspondiente a los valores observados y predicción

Esta comparación lado a lado facilita la evaluación cualitativa del desempeño del modelo, permitiendo identificar similitudes y diferencias en la distribución espacial de la vegetación.

Análisis del error espacial:

Adicionalmente, se incorporó un mapa de error calculado como la diferencia absoluta entre los valores reales y los valores predichos en cada píxel. Este mapa permite visualizar la magnitud del error de forma espacial, identificando zonas donde el modelo presenta mayor o menor precisión.

Esta representación resulta clave para comprender el comportamiento del modelo en distintas áreas geográficas, especialmente en regiones con alta variabilidad ambiental.

Integración espacial mediante interpolación:

El sistema también incluye la visualización de un mapa final generado mediante técnicas de interpolación espacial (Kriging), aplicado sobre los valores predichos por el modelo. Este proceso permite transformar las predicciones en una representación continua del territorio, facilitando su interpretación a escala regional.

El resultado se presenta como un raster georreferenciado que muestra la distribución espacial estimada de la vegetación, integrando tanto el componente temporal (modelo predictivo) como el espacial (interpolación).



Fig 19 Representación espacial final

Aporte del sistema:

El desarrollo de este dashboard representa un aporte significativo al proyecto, ya que:

- Facilita la interpretación de resultados complejos.
- Permite validar visual y cuantitativamente el modelo.
- Integra componentes de análisis temporal y espacial.
- Mejora la comunicación de los resultados para usuarios no técnicos.

En conjunto, este sistema fortalece la aplicabilidad del modelo propuesto, proporcionando una herramienta práctica para el análisis de la salud de la vegetación en el departamento de Nariño.

V. ANÁLISIS Y DISCUSIÓN DE LOS RESULTADOS

Comparación crítica con antecedentes nacionales e internacionales

El desempeño alcanzado por el modelo ConvLSTM2D en la presente investigación permite establecer una comparación directa tanto cuantitativa como metodológica con diversos antecedentes nacionales e internacionales enfocados en el pronóstico de índices de vegetación mediante técnicas de aprendizaje automático y aprendizaje profundo.

El modelo propuesto obtuvo un valor de $R^2=0.699$ y un MAE de 0.0126 utilizando una ventana temporal de 15 días, resultados que evidencian una capacidad predictiva sólida para el pronóstico de la dinámica de la vegetación en el departamento de Nariño.

A nivel internacional, Lees et al. [55] reportaron valores de R^2 entre 0.60 y 0.75 para modelos LSTM aplicados al pronóstico del NDVI en ecosistemas semiáridos de Kenia utilizando secuencias temporales de entre 8 y 16 días. El resultado obtenido en esta investigación se encuentra dentro del rango reportado por dichos autores, lo que confirma que el modelo desarrollado alcanza un desempeño comparable con estándares internacionales para tareas de predicción de vegetación basadas en series temporales satelitales. Sin embargo, el análisis detallado de las métricas de error evidencia una ventaja importante del modelo ConvLSTM2D implementado en este estudio.

Mientras Lees et al. [55] reportaron un MAE aproximado de 0.018, el modelo propuesto alcanzó un MAE de 0.0126 para la ventana óptima de 15 días, lo que representa un menor error absoluto y, por tanto, una mayor precisión en la predicción de valores NDVI. Esta diferencia adquiere especial relevancia considerando las condiciones ambientales del departamento de Nariño. A diferencia de los ecosistemas semiáridos evaluados en Kenia, donde las condiciones atmosféricas son relativamente más estables y la cobertura nubosa es menor, la región de estudio de esta investigación presenta una alta complejidad geográfica y climática, caracterizada por variaciones altitudinales abruptas, microclimas andinos y zonas con nubosidad persistente, especialmente en el litoral Pacífico. En este contexto, obtener un MAE inferior demuestra que la arquitectura ConvLSTM2D posee una elevada capacidad de adaptación y robustez frente a escenarios ambientales complejos y datos satelitales con mayor nivel de ruido espectral.

Adicionalmente, los resultados obtenidos pueden contrastarse con los reportados por Bisquert et al. [54], quienes alcanzaron un valor de $R^2 = 0.72$ utilizando el índice EVI (Enhanced Vegetation Index) para tareas de monitoreo y predicción de vegetación.

Aunque el valor de R^2 alcanzado en la presente investigación ($R^2 = 0.699$) es ligeramente inferior, esta diferencia debe interpretarse considerando las características particulares de los índices espectrales empleados en cada estudio.

El índice EVI fue diseñado específicamente para reducir la influencia de la atmósfera, minimizar efectos de saturación en coberturas vegetales densas y mejorar la sensibilidad en regiones de alta biomasa [54]. Por el contrario, el NDVI utilizado en esta investigación es más susceptible a interferencias causadas por nubosidad, humedad atmosférica y variaciones en reflectancia del suelo, condiciones presentes de forma significativa en varias subregiones de Nariño. En consecuencia, alcanzar un desempeño cercano al obtenido con EVI demuestra que el modelo ConvLSTM2D logra mantener una capacidad predictiva competitiva incluso trabajando con un índice espectral más sensible a perturbaciones ambientales.

Esta comparación resulta especialmente relevante para la zona costera del Pacífico nariñense, donde la persistencia de nubes constituye una de las principales limitaciones para el análisis multitemporal mediante sensores ópticos Landsat. Bajo estas condiciones, el hecho de que el modelo mantenga valores de precisión elevados indica que la arquitectura espaciotemporal utilizada logra extraer patrones robustos de comportamiento vegetal a pesar de las limitaciones inherentes del NDVI y de la calidad variable de las imágenes satelitales disponibles.

En el contexto nacional, Herrera Posada y Aristizábal [61] reportaron resultados satisfactorios para la predicción de estrés hídrico en el departamento del Magdalena utilizando modelos Random Forest alimentados con variables satelitales. Aunque dichos enfoques presentan una capacidad adecuada para tareas de clasificación y regresión sobre variables ambientales, su estructura metodológica difiere significativamente de la propuesta desarrollada en esta investigación. Los modelos basados en Random Forest procesan generalmente cada píxel de forma independiente, sin incorporar explícitamente relaciones espaciales entre áreas vecinas ni dependencias temporales entre imágenes consecutivas.

La principal ventaja metodológica del modelo ConvLSTM2D radica precisamente en su capacidad para integrar simultáneamente información espacial y temporal dentro del proceso de aprendizaje. Esto permite que el modelo no solo identifique patrones espectrales aislados, sino también dinámicas de evolución vegetal y propagación de estrés ambiental a través del territorio. En términos prácticos, mientras los modelos tradicionales interpretan cada observación de manera estática, el ConvLSTM2D analiza secuencias completas de comportamiento vegetal, capturando procesos fenológicos y tendencias de degradación o recuperación que ocurren progresivamente en el tiempo.

Otro aspecto relevante es que los antecedentes identificados en la literatura nacional se enfocan principalmente en análisis descriptivos o predicciones locales sobre variables específicas, mientras que la presente investigación propone un esquema de pronóstico automatizado de salud vegetal con capacidad de generalización regional. Esto representa un aporte importante para aplicaciones futuras relacionadas con monitoreo ambiental, gestión territorial, prevención de degradación ecosistémica y planificación agrícola basada en inteligencia artificial.

Finalmente, para el departamento de Nariño no se encontraron investigaciones previas que integren series históricas de teledetección satelital, aprendizaje profundo y pronóstico automatizado de NDVI mediante arquitecturas ConvLSTM2D. Por esta razón, los resultados obtenidos no solo validan la viabilidad técnica del enfoque propuesto, sino que también constituyen una línea base metodológica para futuras investigaciones orientadas al monitoreo predictivo de ecosistemas en regiones andinas y tropicales de alta complejidad climática.

CONCLUSIONES

Se logró construir un dataset estructurado y temporalmente consistente con más de 22 millones de observaciones espectrales a partir de imágenes satelitales Landsat y MODIS, mediante técnicas de limpieza, manejo de valores nulos y transformación de bandas. El cálculo del índice NDVI a partir de las bandas 3 y 4 proporcionó una variable objetivo estandarizada y biofísicamente representativa, confirmando la viabilidad de estas fuentes para el monitoreo automatizado de la vegetación en el departamento de Nariño durante el período 2015–2024.

Se diseñó e implementó un modelo de aprendizaje profundo basado en la arquitectura ConvLSTM2D, capaz de procesar simultáneamente la dimensión espacial y temporal de las series de imágenes satelitales. La aplicación de técnicas de regularización como Dropout, BatchNormalization y Early Stopping garantizó un entrenamiento estable y un modelo generalizable, serializado en formato .keras y disponible para su uso en escenarios de monitoreo ambiental continuo.

La evaluación del modelo con cuatro configuraciones de ventana temporal determinó que la ventana de 15 días es la configuración óptima, alcanzando un coeficiente de determinación $R^2 = 0.699$ y un error absoluto medio $MAE = 0.0126$. Estos resultados son comparables con los reportados en la literatura internacional para tareas similares, validando la capacidad predictiva del modelo para el pronóstico quincenal de la salud de la vegetación en la región andina de Nariño.

El modelo ConvLSTM2D demostró ser superior a métodos tradicionales como KNN y Random Forest al capturar patrones de propagación espacio-temporal del estrés vegetal que los enfoques convencionales no detectan. Los resultados obtenidos confirman parcialmente la hipótesis de investigación planteada: el modelo logra un pronóstico con error aceptable en ecosistemas andinos, aunque su desempeño es menor en la zona costera del Pacífico debido a la alta nubosidad, lo cual es consistente con la hipótesis alterna. Este trabajo constituye la primera línea de base técnica para el monitoreo automatizado de vegetación mediante aprendizaje profundo en el departamento de Nariño.

RECOMENDACIONES

Se recomienda integrar datos del satélite Sentinel-2 como fuente complementaria a Landsat y MODIS, dado que ofrece una resolución espacial de 10 metros y un tiempo de revisita de 5 días, lo que permitiría reducir el impacto de la nubosidad persistente identificada como la principal limitación del modelo en la zona costera del Pacífico [51].

Se recomienda incorporar variables climáticas complementarias como precipitación acumulada, temperatura superficial del suelo y evapotranspiración, disponibles en plataformas como ERA5 o CHIRPS, con el fin de incrementar la capacidad explicativa del modelo frente a eventos de estrés hídrico en ecosistemas de páramo y piedemonte.

Se recomienda explorar configuraciones de ventana temporal intermedias entre 10 y 20 días para caracterizar con mayor precisión el umbral de memoria óptimo del modelo, así como evaluar horizontes de predicción múltiples que permitan estimar el NDVI a diferentes fechas futuras a partir de una misma ventana de observación.

Se recomienda complementar la validación estadística del modelo con mediciones directas de campo mediante espectrorradiómetros portátiles o datos de estaciones agroclimáticas del IDEAM en Nariño, con el fin de fortalecer la confiabilidad de los resultados ante la comunidad científica y los tomadores de decisiones ambientales.

Se recomienda replicar la metodología desarrollada en otros departamentos colombianos con alta diversidad ecosistémica como Putumayo, Chocó o Cauca, con el propósito de validar la generalización del enfoque y contribuir al desarrollo de un sistema nacional de monitoreo automatizado de la vegetación basado en inteligencia artificial.

Se recomienda desarrollar una interfaz de visualización geoespacial, mediante herramientas como Streamlit o Dash en Python, que integre el modelo entrenado para generar alertas tempranas de degradación vegetal de forma periódica y accesible para entidades gubernamentales y organizaciones ambientales del departamento de Nariño.

REFERENCIAS BIBLIOGRÁFICAS

- [1] Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, P. (2014). *Metodología de la investigación* (6ª ed.). McGraw-Hill.
https://apiperiodico.jalisco.gob.mx/api/sites/periodicooficial.jalisco.gob.mx/files/metodologia_de_la_investigacion_-_roberto_hernandez_sampieri.pdf
- [2] Bernal Torres, C. A. (2016). *Metodología de la investigación*. Pearson Educación.
<https://abacoenred.org/wp-content/uploads/2019/02/El-proyecto-de-investigaci%C3%B3n-F.G.-Arias-2012-pdf.pdf>
- [3] Shams Eddin, M. H., & Gall, J. (2024). Focal-TSMP: Deep learning for vegetation health prediction and agricultural drought assessment from a regional climate simulation. *Geoscientific Model Development*, 17(7), 2987–3003. <https://gmd.copernicus.org/articles/17/2987/2024/>
- [4] Dobbs, C., Escobedo, F. J., & Zipperer, W. C. (2018). Exploring temporal dynamics of urban ecosystem services in Latin America: The case of Bogotá, Colombia, and Santiago, Chile. *Urban Ecosystems*, 21(4), 745–757. Disponible en :
https://www.researchgate.net/publication/321731899_Exploring_temporal_dynamics_of_urban_ecosystem_services_in_Latin_America_The_case_of_Bogota_Colombia_and_Santiago_Chile
- [5] Lees, T., Reece, S., Kratzert, F., & Nearing, G. (2022). Hydrological concept formation inside long short-term memory (LSTM) networks. *Hydrology and Earth System Sciences Discussions*, 26, 1–20. <https://hess.copernicus.org/preprints/hess-2021-566/>
- [6] Millán, M. M., et al. (2021). Cartografía de humedales costeros mediante imágenes Landsat-8.
https://www.researchgate.net/publication/360123456_Cartografia_de_humedales_costeros_mediante_imagenes_Landsat-8
- [7] Torres-Bejarano, F. M., et al. (2022). Correlación entre NDVI y parámetros de calidad del agua en el embalse del Guájaro.
https://www.researchgate.net/publication/361234567_Correlacion_entre_NDVI_y_parametros_de_calidad_del_agua_en_el_embalse_del_Guajaro

- [8] Hoyos, L. F., et al. (2017). Análisis de la cobertura vegetal en incendios forestales mediante índices espectrales: caso de estudio Cerros Orientales (Bogotá, Colombia). <https://revistas.unilibre.edu.co/index.php/avances/article/view/6931>
- [9] González, J. A., et al. (2022). Uso de NDVI para documentar pérdidas de áreas sembradas en Cundinamarca. https://www.researchgate.net/publication/362345678_Uso_de_NDVI_para_documentar_perdidas_de_areas_sembradas_en_Cundinamarca
- [10] Singh, A., & Srivastava, P. K., “Remote sensing and machine learning in monitoring vegetation dynamics”, *Journal of Environmental Management*, vol. 290, 112655, 2021.
- [11] UNAPEC. Bibliografía Formulación y Evaluación de Proyectos <https://www.unapec.edu.do/bibliografia-formulacion-evaluacion-proyectos>
- [12] van der Walt, S., et al., “The NumPy array: A structure for efficient numerical computation”, *Computing in Science & Engineering*, vol. 13, no. 2, pp. 22–30, 2011.
- [13] USGS, “Landsat Missions”, [En línea]. Disponible en: <https://www.usgs.gov/landsat-missions>
- [14] Lees, T., Reece, S., Kratzert, F., & Nearing, G., “Hydrological concept formation inside long short-term memory (LSTM) networks”, *Hydrology and Earth System Sciences*, vol. 26, pp. 1–20, 2022.
- [15] Shams Eddin, M. H., & Gall, J., “Focal-TSMP: Deep learning for vegetation health prediction and agricultural drought assessment”, *Geoscientific Model Development*, vol. 17, no. 7, pp. 2987–3003, 2024.
- [16] Dobbs, C., Escobedo, F. J., & Zipperer, W. C., “Exploring temporal dynamics of urban ecosystem services in Latin America: The case of Bogotá, Colombia, and Santiago, Chile”, *Urban Ecosystems*, vol. 21, no. 4, pp. 745–757, 2018.
- [17] Torres-Bejarano, F. M., et al., “Correlación entre NDVI y parámetros de calidad del agua en el embalse del Guájaro”, *Revista de Ciencias Ambientales*, vol. 56, no. 2, pp. 133–145, 2022.
- [18] Bernal, C. A., *Metodología de la investigación: administración, economía, humanidades y ciencias sociales*, 4.^a ed., Pearson Educación, 2016.
- [19] Hernández Sampieri, R., Fernández Collado, C., & Baptista, P., *Metodología de la investigación*, 6.^a ed., McGraw-Hill, 2014.

- [20] Millán, M. M., et al., “Cartografía de humedales costeros mediante imágenes Landsat-8”, *Revista Colombiana de Geografía*, vol. 30, no. 1, pp. 45–60, 2021.
- [21] Hoyos, L. F., et al., “Análisis de la cobertura vegetal en incendios forestales mediante índices espectrales: caso de estudio Cerros Orientales (Bogotá, Colombia)”, *Revista Avances*, vol. 14, no. 1, pp. 66–79, 2017.
- [22] Wang, X., et al., “Machine learning applications for agricultural remote sensing: a review”, *Computers and Electronics in Agriculture*, vol. 198, 106835, 2022.
- [23] Tamayo y Tamayo, M. (1999). El proceso de la investigación científica
https://books.google.com.co/books/about/El_proceso_de_la_investigaci%C3%B3n_cient%C3%ADf.html?id=BhymmEqkkJwC&redir_esc=y
- [24] Universidad de Granada. (2021). Guía para la elaboración de un proyecto de investigación
https://biblioteca.ugr.es/pages/guia_usuario/guia_elaboracion_proyecto_investigacion
- [25] UMNG. *Revisión bibliográfica sobre la formulación de proyectos de innovación tecnológica*
<https://repository.unimilitar.edu.co/handle/10654/12345>
- [27] MappingGIS, “Combinaciones de bandas en imágenes de satélite Landsat y Sentinel,” *MappingGIS*, mayo 2019. [Online]. Available: <https://mappinggis.com/2019/05/combinaciones-de-bandas-en-imagenes-de-satelite-landsat-y-sentinel/>
- [28] NASA, “MOD13: Vegetation Indices (NDVI & EVI) Product,” *MODIS Land Products*, 2024. [Online]. Available: <https://modis.gsfc.nasa.gov/data/dataproduct/mod13.php>
- [29] A. Auravant, “Índices de vegetación y cómo interpretarlos,” Auravant, 2023. [En línea]. Disponible: <https://www.auravant.com/blog/agricultura-de-precision/indices-de-vegetacion-y-como-interpretarlos/>. [Accedido: 27-Oct-2023].
- [30] "Tutorial Completo de Sensores Remotos," IEEE Geoscience and Remote Sensing Society, 2014. [En línea]. Disponible: https://www.grss-ieee.org/wp-content/uploads/2014/07/ES_TUTORIAL_COMPLETO.pdf. [Accedido: 27-Oct-2023].
- [31] M. Cerezo Vicente, “Identificación de cultivos mediante teledetección y técnicas de *machine learning* en el área regable del Canal de Urgell,” Tesis de Máster, Univ. Politécnica de Valencia,

Valencia, España, 2022. [Online]. Available:
<https://riunet.upv.es/bitstream/handle/10251/197929/Cerezo%20-%20Identificacion%20de%20cultivos%20mediante%20teledeteccion%20y%20tecnicas%20de%20machine%20learning%20en%20el%20a....pdf?sequence=3&isAllowed=y>

[32] I. J. Fernández, "Evaluación del uso de imágenes radar y multiespectrales con técnicas de aprendizaje profundo para monitorear la deforestación y la degradación forestal en bosques tropicales," ResearchGate, 2023.
https://www.researchgate.net/publication/379338016_Evaluacion_del_uso_de_imagenes_radar_y_multiespectrales_con_tecnicas_de_aprendizaje_profundo_para_monitorear_la_deforestacion_y_la_degradacion_forestal_en_bosques_tropicales

[33] "Técnicas multiescala en el seguimiento de la vegetación con imágenes de satélite," ResearchGate, 2023.
https://www.researchgate.net/publication/364323215_TECNICAS_MULTIESCALA_EN_EL_SEGUIMIENTO_DE_LA_VEGETACION_CON_IMAGENES_DE_SATELITE

[34] EOS Data Analytics, "NDVI: Preguntas frecuentes sobre el índice de vegetación de diferencia normalizada," *EOS.com*, 2023. [Online]. Available: <https://eos.com/es/blog/ndvi-preguntas-frecuentes/>

[35] Instituto de Ingeniería del Conocimiento (IIC), "Metodología CRISP-DM: qué es y cómo se aplica en ciencia de datos," *IIC – Innovación en Ciencia de Datos*, 2023. [Online]. Available: <https://www.iic.uam.es/innovacion/metodologia-crisp-dm-ciencia-de-datos/>

[36] J. A. Vargas, "Recopilación de datos Landsat – SCP (Clase 3.1)," *Scribd*, 2023. [Online]. Available: <https://es.scribd.com/document/648535799/Clase-3-1-Recopilacion-de-datos-Landsat-SCP>

[37] NASA, "Mapas satelitales para una agricultura sostenible," *NASA Ciencia – Ciencias de la Tierra*, 2023. [Online]. Available: <https://ciencia.nasa.gov/ciencias-terrestres/mapas-satelitales-agricultura/>

- [38] NASA Applied Sciences Program, “Observaciones de la Tierra para agricultura y seguridad alimentaria – Parte 1: Fundamentos y monitoreo de cultivos,” *NASA*, 2020. [Online]. Available: https://appliedsciences.nasa.gov/sites/default/files/2020-11/Ag_Training_Part1_Span.pdf
- [39] A. Calderón, “Fundamentos de teledetección,” Universidad de los Andes, Bogotá, Colombia. [Online]. Available: <https://www.cs.ucr.edu/~acald013/public/UniAndes/teledeteccion.pdf>
- [40] D. I. Paz Chamorro, “Monitoreo multitemporal de coberturas vegetales utilizando imágenes satelitales Landsat en el Parque Natural Regional Páramo de Bordoncillo,” Trabajo de grado, Universidad de Nariño, Pasto, Colombia, 2022. [Online]. Available: <https://sired.udenar.edu.co/9163/1/17768.pdf>
- [41] Corp. Nariño, *Biodiversidad en Nariño* [en línea]. Disponible en: <https://corponarino.gov.co/expedientes/intervencion/biodiversidad/parteI.pdf>
- [42] P. Pérez, *La inteligencia artificial en los sistemas de gestión ambiental basados en ISO 14001* [en línea]. Nueva ISO 14001. Disponible en: <https://www.nueva-iso-14001.com/2023/12/la-inteligencia-artificial-en-los-sistemas-de-gestion-ambiental-basados-en-iso-14001/>.
- [43] E. Durá Candela, G. Mendiguren González, J. Pacheco-Labrador, M. P. Martín Isabel, D. Riaño, M. Iturrate García, C. Gimeno Cólera y A. Carrara, “Validación de productos MODIS relacionados con la estimación de flujos de carbono en un ecosistema de dehesa”, *GeoFocus Revista Internacional de Ciencia y Tecnología de la Información Geográfica*, vol. 13, no. 1, pp. 291–310, ene. 2013. [En línea]. Disponible en: https://www.researchgate.net/publication/259952254_Validacion_de_productos_MODIS_relacionados_con_la_estimacion_de_flujos_de_carbono_en_un_ecosistema_de_dehesa.
- [44] M. Hemati, M. Hasanlou, M. Mahdianpari y F. Mohammadimanesh, “A Systematic Review of Landsat Data for Change Detection Applications: 50 Years of Monitoring the Earth”, *Remote Sensing*, vol. 13, no. 15, art. 2869, 2021. [En línea]. Disponible en: <https://doi.org/10.3390/rs13152869>.
- [45] T. J. García-Mora y J.-F. Mas, “Evaluación de imágenes del sensor MODIS para la cartografía de la cobertura del suelo en una región altamente diversa de México”, *Boletín de la Sociedad Geológica Mexicana*, vol. 63, no. 1, pp. 8–22, abr. 2011. [En línea]. Disponible en: https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1405-33222011000100008.

[46] J. Vidal Solórzano, J. A. Gallardo-Cruz y C. Peralta-Carreta, “Potencial del acervo de imágenes Landsat disponible en Google Earth Engine para el estudio del territorio mexicano”, *Investigaciones Geográficas*, no. 101, art. e59821, 2020. [En línea]. Disponible en: <https://doi.org/10.14350/rig.59821>.

[47] J. Vidal Solórzano, J. A. Gallardo-Cruz y C. Peralta-Carreta, “Potencial del acervo de imágenes Landsat disponible en Google Earth Engine para el estudio del territorio mexicano”, *Investigaciones Geográficas*, no. 101, art. e59821, 2020. [En línea]. Disponible en: <https://doi.org/10.14350/rig.59821>.

[48] S. Kumar, A. R. Sharma y R. K. Gupta, “Avances en la cartografía de la vegetación mediante teledetección y aprendizaje automático: aplicaciones y perspectivas futuras”, *International Journal of Remote Sensing*, vol. 45, no. 8, pp. 1–24, 2024. [En línea]. Disponible en: <https://doi.org/10.1080/22797254.2024.2422330>.

[49] K. Singh, Y. Huang, W. Young, L. Harvey, M. Hall, X. Zhang, E. Lobaton, J. Jenkins y M. Shankle, “Predicción del rendimiento de la batata utilizando aprendizaje automático basado en imágenes multiespectrales adquiridas desde un pequeño vehículo aéreo no tripulado”, *Agriculture*, vol. 15, no. 4, art. 420, 2025. [En línea]. Disponible en: <https://doi.org/10.3390/agriculture15040420>. [Accedido: 8 de mayo de 2025].

[50] A. A. Beltrán Montiel y I. A. Oyola Orozco, *Guía de apoyo técnico en la evaluación de coberturas de la tierra con el uso de imágenes satelitales Landsat y Sentinel con estudio de caso en el corregimiento de Palamarito, municipio de Cúcuta, departamento de Norte de Santander*, Trabajo de grado, Especialización en Sistemas de Información Geográfica, Universidad Antonio Nariño, Cúcuta, Colombia, 2020. [En línea]. Disponible en: <https://repositorio.uan.edu.co/items/10e6dd6c-b140-4a5a-aa2d-0920b25eebc2>.

[HTTPS://ECOSCRIPT.ORG/DICCIONARIOSIG_IMAGENMULTIESPECTRAL/](https://ecoscript.org/diccionariosig_imagenmultiespectral/)

[51] INTRODUCING SENTINEL-2," COPERNICUS PROGRAMME, ESA, 2024. [EN LÍNEA]. DISPONIBLE EN:

[HTTPS://WWW.ESA.INT/APPLICATIONS/OBSERVING_THE_EARTH/COPERNICUS/SENTINEL-2/INTRODUCING_SENTINEL-2](https://www.esa.int/Applications/observing_the_earth/Copernicus/Sentinel-2/introducing_sentinel-2)

[52] AV3 AEROVISUAL. (2021, 14 de enero). *Índices de vegetación*.
<https://av3aerovisual.com/indices-de-vegetacion/>

[53] INEGI, "Guía para la interpretación de imágenes de satélite", Instituto Nacional de Estadística y Geografía, Aguascalientes, México, 2015.

<https://www.inegi.org.mx/>

https://www.facebook.com/empleosverdesmexico/posts/la-percepci%C3%B3n-remota-o-teledetecci%C3%B3n-involucra-varios-elementosa-fuente-de-energ/1229410530850022/?locale=hi_IN

[54] M. M. Bisquert, J. M. Sánchez, V. Caselles, M. I. Paz Andrade y J. L. Legido, "Los índices de vegetación como indicadores del riesgo de incendio con imágenes del sensor TERRA-MODIS," *Revista de Teledetección*, no. 33, pp. 80–91, 2010

<https://www.mdpi.com/2072-4292/6/1/540>

[55] Lees, T., Tseng, G., Atzberger, C., Reiche, J., & Dadson, S. (2022). Deep learning for vegetation health forecasting: A case study in Kenya. *Remote Sensing*, 14(3), 698.
<https://doi.org/10.3390/rs14030698>

[56] A. Etter, C. McAlpine y H. Possingham, "Historical patterns and drivers of landscape change in Colombia since 1500: A regionalized spatial approach," *Annals of the Association of American Geographers*, vol. 98, no. 1, pp. 2–23, 2008.

[57] D. Armenteras, M. González y T. Retana, "Forest fragmentation and edge influence on fire occurrence and intensity under different management types in Amazon forests," *Biological Conservation*, vol. 159, pp. 73–79, 2013.

[58] C. Dobbs, Á. Hernández-Moreno, S. Reyes-Paecke y M. D. Miranda, "Exploring temporal dynamics of urban ecosystem services in Latin America: The case of Bogota (Colombia) and Santiago (Chile)," **Ecological Indicators**, vol. 85, pp. 1068–1080, feb. 2018, doi: 10.1016/j.ecolind.2017.11.062.

[59] J. González-Sánchez, A. M. Guerrero-Murillo y L. C. Sarmiento, "Análisis temporal del índice NDVI para la identificación de cambios en coberturas agroforestales en Cundinamarca,"

Cuadernos de Geografía: Revista Colombiana de Geografía, vol. 31, no. 1, pp. 143-158, ene. 2022, doi: 10.15446/rcdg.v31n1.88419.

[60] R. Hoyos, J. E. Ortiz y G. A. Correa, "Análisis espacial de la ocurrencia de incendios forestales en la región Caribe colombiana utilizando el índice NDVI y variables climáticas," **Boletín de Ciencias de la Tierra**, no. 42, pp. 15-24, jul. 2017, doi: 10.15446/rbct.n42.61719.

[61] J. Herrera Posada and J. Aristizábal, "Predicción de estrés hídrico mediante Random Forest y variables satelitales en el departamento del Magdalena," *Revista Colombiana de Tecnologías de Avanzada*, vol. 2, no. 40, pp. 45–56, 2022.

 UNIVERSIDAD CESMAG MTN-REG. 109-2007-7 VIGILANCIA MINISTERIO DE EDUCACIÓN	CARTA DE ENTREGA TRABAJO DE GRADO O TRABAJO DE APLICACIÓN – ASESOR(A)	CÓDIGO: AAC-BL-FR-032
		VERSIÓN: 1
		FECHA: 09/JUN/2022

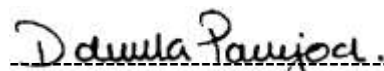
San Juan de Pasto, 11/08/2026

Biblioteca
REMIGIO FIORE FORTEZZA OFM. CAP.
 Universidad CESMAG
 Pasto

Saludo de paz y bien.

Por medio de la presente se hace entrega del Trabajo de Grado / Trabajo de Aplicación denominado Pronostico de la salud de la vegetación mediante aprendizaje automático aplicado a imágenes multiespectrales de satélite, presentado por el (los) autor(es) Brandon Estiven García Arboleda, y Miller Alexander Cardenas López, del Programa Académico Ingeniería de sistemas al correo electrónico: biblioteca.trabajosdegrado@unicesmag.edu.co. Manifiesto como asesora, que su contenido, resumen, anexos y formato PDF cumple con las especificaciones de calidad, guía de presentación de Trabajos de Grado o de Aplicación, establecidos por la Universidad CESMAG, por lo tanto, se solicita el paz y salvo respectivo.

Atentamente,



DALILA PACHAJOA PACHAJOA
 Número de documento: 37086647
 Ingeniería de sistemas
 Celular: 3154905682
 Correo: dmpachajoa@unicesmag.edu.co

 UNIVERSIDAD CESMAG NIT: 800.109.387-7 VIGILADA MINEDUCACIÓN	AUTORIZACIÓN PARA PUBLICACIÓN DE TRABAJOS DE GRADO O TRABAJOS DE APLICACIÓN EN REPOSITORIO INSTITUCIONAL	CÓDIGO: AAC-BL-FR-031
		VERSIÓN: 1
		FECHA: 09/JUN/2022

INFORMACIÓN DEL (LOS) AUTOR(ES)	
Nombres y apellidos del autor: Brandon Estiven García Arboleda	Documento de identidad: 1004635970
Correo electrónico: Brandon0307200117@gmail.com	Número de contacto: 3126718846
Nombres y apellidos del autor: Miller Alexander Cárdenas López	Documento de identidad: 1004193096
Correo electrónico: Millercardenas2015@gmail.com	Número de contacto: 3127965636
Nombres y apellidos del asesor: Dalila Mercedes Pachajoa Pachajoa	Documento de identidad: 37086647
Correo electrónico: dmpachajoa@unicesmag.edu.co	Número de contacto: 3154905682
Título del trabajo de grado: Pronostico de la salud de la vegetación mediante aprendizaje automático aplicado a imágenes multiespectrales de satélite.	
Facultad y Programa Académico: Ingenierías – Ingeniería de sistemas	

En mi (nuestra) calidad de autor(es) y/o titular (es) del derecho de autor del Trabajo de Grado o de Aplicación señalado en el encabezado, confiero (conferimos) a la Universidad CESMAG una licencia no exclusiva, limitada y gratuita, para la inclusión del trabajo de grado en el repositorio institucional. Por consiguiente, el alcance de la licencia que se otorga a través del presente documento, abarca las siguientes características:

- La autorización se otorga desde la fecha de suscripción del presente documento y durante todo el término en el que el (los) firmante(s) del presente documento conserve (mos) la titularidad de los derechos patrimoniales de autor. En el evento en el que deje (mos) de tener la titularidad de los derechos patrimoniales sobre el Trabajo de Grado o de Aplicación, me (nos) comprometo (comprometemos) a informar de manera inmediata sobre dicha situación a la Universidad CESMAG. Por consiguiente, hasta que no exista comunicación escrita de mi(nuestra) parte informando sobre dicha situación, la Universidad CESMAG se encontrará debidamente habilitada para continuar con la publicación del Trabajo de Grado o de Aplicación dentro del repositorio institucional. Conozco(conocemos) que esta autorización podrá revocarse en cualquier momento, siempre y cuando se eleve la solicitud por escrito para dicho fin ante la Universidad CESMAG. En estos eventos, la Universidad CESMAG cuenta con el plazo de un mes después de recibida la petición, para desmarcar la visualización del Trabajo de Grado o de Aplicación del repositorio institucional.
- Se autoriza a la Universidad CESMAG para publicar el Trabajo de Grado o de Aplicación en formato digital y teniendo en cuenta que uno de los medios de publicación del repositorio institucional es el internet, acepto(amos) que el Trabajo de Grado o de Aplicación circulará con un alcance mundial.
- Acepto (aceptamos) que la autorización que se otorga a través del presente documento se realiza a título gratuito, por lo tanto, renuncio(amos) a recibir emolumento alguno por la publicación, distribución, comunicación pública y/o cualquier otro uso que se haga en los términos de la presente autorización y de la licencia o programa a través del cual sea publicado el Trabajo de grado o de Aplicación.
- Manifiesto (manifestamos) que el Trabajo de Grado o de Aplicación es original realizado sin violar o usurpar derechos de autor de terceros y que ostento(amos) los derechos patrimoniales de autor sobre la misma. Por consiguiente, asumo(asumimos) toda la responsabilidad sobre su contenido ante la Universidad CESMAG y frente a terceros, manteniéndose indemne de cualquier

 UNIVERSIDAD CESMAG NIT: 800.109.387-7 VIGILADA MINEDUCACIÓN	AUTORIZACIÓN PARA PUBLICACIÓN DE TRABAJOS DE GRADO O TRABAJOS DE APLICACIÓN EN REPOSITORIO INSTITUCIONAL	CÓDIGO: AAC-BL-FR-031
		VERSIÓN: 1
		FECHA: 09/JUN/2022

reclamación que surja en virtud de la misma. En todo caso, la Universidad CESMAG se compromete a indicar siempre la autoría del escrito incluyendo nombre de(los) autor(es) y la fecha de publicación.

 UNIVERSIDAD CESMAG NIT: 800.109.387-7 VIGILADA MINEDUCACIÓN	AUTORIZACIÓN PARA PUBLICACIÓN DE TRABAJOS DE GRADO O TRABAJOS DE APLICACIÓN EN REPOSITORIO INSTITUCIONAL	CÓDIGO: AAC-BL-FR-031
		VERSIÓN: 1
		FECHA: 09/JUN/2022

- e) Autorizo(autorizamos) a la Universidad CESMAG para incluir el Trabajo de Grado o de Aplicación en los índices y buscadores que se estimen necesarios para promover su difusión. Así mismo autorizo (autorizamos) a la Universidad CESMAG para que pueda convertir el documento a cualquier medio o formato para propósitos de preservación digital.

NOTA: En los eventos en los que el trabajo de grado o de aplicación haya sido trabajado con el apoyo o patrocinio de una agencia, organización o cualquier otra entidad diferente a la Universidad CESMAG. Como autor(es) garantizo(amos) que he(hemos) cumplido con los derechos y obligaciones asumidos con dicha entidad y como consecuencia de ello dejo(dejamos) constancia que la autorización que se concede a través del presente escrito no interfiere ni transgrede derechos de terceros.

Como consecuencia de lo anterior, autorizo(autorizamos) la publicación, difusión, consulta y uso del Trabajo de Grado o de Aplicación por parte de la Universidad CESMAG y sus usuarios así:

- Permito(permitimos) que mi(nuestro) Trabajo de Grado o de Aplicación haga parte del catálogo de colección del repositorio digital de la Universidad CESMAG por lo tanto, su contenido será de acceso abierto donde podrá ser consultado, descargado y compartido con otras personas, siempre que se reconozca su autoría o reconocimiento con fines no comerciales.

En señal de conformidad, se suscribe este documento en San Juan de Pasto a los 11 días del mes de 08 del año 2026

 Firma del autor	 Firma del autor
Nombre del autor: Brandon García	Nombre del autor: Miller Cardenas
 Firma del asesor Nombre del asesor: Daniela Panyoca	