



Universidad
Nacional
de Loja

Universidad Nacional de Loja

Facultad Agropecuaria y de Recursos Naturales Renovables

Carrera de Ingeniería Ambiental

Predicción de zonas susceptibles a incendios forestales aplicando técnicas de aprendizaje automático en el cantón Zapotillo, Loja – Ecuador

Trabajo de Titulación, previo a la
obtención del título de Ingeniero
Ambiental

AUTOR:

Deiby Emersson Granda García

DIRECTOR:

Ing. Carlos Guillermo Chunchu Morocho Mg. Sc

CODIRECTOR:

Fabián René Reyes Bueno, PhD

Loja – Ecuador

2024

Certificación

Loja, 5 de marzo de 2024

Ing. Carlos Guillermo Chuncho Morocho Mg. Sc.
DIRECTOR DE TRABAJO DE TITULACIÓN

CERTIFICO:

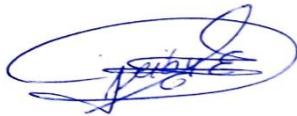
Que he revisado y orientado todo el proceso de la elaboración del Trabajo de Titulación denominado: **Predicción de zonas susceptibles a incendios forestales aplicando técnicas de aprendizaje automático en el cantón Zapotillo, Loja – Ecuador**, previo a la obtención del título de **Ingeniero Ambiental**, de la autoría del estudiante **Deiby Emersson Granda García**, con **cédula de identidad** Nro. **1900798370**, una vez que el trabajo cumple con todos los requisitos exigidos por la Universidad Nacional de Loja, para el efecto, autorizo la presentación del mismo para su respectiva sustentación y defensa.

Ing. Carlos Guillermo Chuncho Morocho Mg. Sc.
DIRECTOR DEL TRABAJO DE TITULACIÓN

Autoría

Yo, **Deiby Emersson Granda García**, declaro ser autor del presente Trabajo de Titulación y eximo expresamente a la Universidad Nacional de Loja y a sus representantes jurídicos, de posibles reclamos y acciones legales por el contenido del mismo. Adicionalmente acepto y autorizo a la Universidad Nacional de Loja la publicación de mi Trabajo de Titulación, en el Repositorio Institucional – Biblioteca Virtual.

Firma:



Cédula de identidad: 1900798370

Fecha: 11 de julio de 2024

Correo electrónico: deiby.granda@unl.edu.ec

Teléfono: 0968120858

Carta de autorización por parte del autor, para consulta, reproducción parcial o total y/o publicación electrónica del texto completo, del Trabajo de Titulación.

Yo, **Deiby Emersson Granda García**, declaro ser autor del Trabajo de Titulación denominado: **Predicción de zonas susceptibles a incendios forestales aplicando técnicas de aprendizaje automático en el cantón Zapotillo, Loja – Ecuador** como requisito para optar el título de **Ingeniero Ambiental**, autorizo al sistema Bibliotecario de la Universidad Nacional de Loja para que, con fines académicos, muestre la producción intelectual de la Universidad, a través de la visibilidad de su contenido en el Repositorio Institucional.

Los usuarios pueden consultar el contenido de este trabajo en el Repositorio Institucional, en las redes de información del país y del exterior con las cuales tenga convenio la Universidad.

La Universidad Nacional de Loja, no se responsabiliza por el plagio o copia del Trabajo de Titulación que realice un tercero.

Para constancia de esta autorización, en la ciudad de Loja, a los once días del mes de julio del dos mil veinte y cuatro.

Firma:

Autor: Deiby Emersson Granda García

Cédula: 1900798370

Dirección: Ciudadela Sol de los Andes

Correo electrónico: deiby.granda@unl.edu.ec

Teléfono: 0968120858

DATOS COMPLEMENTARIOS:

Director del Trabajo de Titulación: Ing. Carlos Guillermo Chuncho Morocho Mg. Sc.

Dedicatoria

A Dios, por haberme permitido terminar este camino gozando de salud. A mis padres, cuyo amor, apoyo incondicional y sacrificio constante han sido la base de este logro. A mis hermanos, quienes han estado presentes a lo largo de este trayecto, siendo un pilar en mi vida, brindándome ánimo para poder culminar mis estudios. Siendo mi familia lo más importante en mi vida, este trabajo es dedicado a ellos, como manera de agradecimiento por haber confiado en mí.

Deiby Emersson Granda García

Agradecimiento

Quiero expresar mi más profundo agradecimiento a la Universidad Nacional de Loja y a todos sus funcionarios, especialmente a aquellos de la carrera de Ingeniería Ambiental y a la planta docente, por su guía y sabiduría que han iluminado mi camino académico, las experiencias que me han brindado las llevaré siempre conmigo.

Quiero agradecer de manera especial a mi director de tesis, Ing. Carlos Guillermo Chuncho Morocho, Mg. Sc., al Ing. Fernando Alexis González Soto, Mg. Sc., Fabián René Reyes Bueno, PhD y a la Ing. Julia Loján Córdova, por su invaluable ayuda y mentoría, indispensables para completar mi trabajo de titulación. También extendo mi gratitud al Centro de Investigaciones Territoriales por los valiosos conocimientos que me han otorgado. A mi familia y amigos, gracias por su comprensión y compañía en cada etapa de este viaje.

Deiby Emersson Granda García

Índice de contenidos

Portada	I
Certificación	ii
Autoría	iii
Carta de autorización	iv
Dedicatoria	v
Agradecimiento	vi
Índice de contenidos	vii
Índice de tablas	x
Índice de figuras	xi
Índice de ecuaciones	xiii
Índice de Anexos	xiv
1. Título	1
2. Resumen	2
Abstract	3
3. Introducción	4
4. Marco teórico	7
4.1. Bosques secos	7
4.2. Incendios forestales	7
4.2.1. Probabilidad a que ocurran incendios forestales	8
4.2.2. Efectos de los incendios forestales	11
4.2.3. Incendios forestales en bosques secos	11
4.3. Detección de incendios forestales	12
4.3.1. Teledetección	13
4.3.2. Índices espectrales:	13
4.4. Satélites utilizados para el monitoreo de los incendios forestales	17
4.4.1. Sentinel-2	17
4.4.2. Satélite SNPP-VIIRS	18

4.5.	Inteligencia Artificial.....	19
4.5.1.	Técnicas de aprendizaje automático	19
4.5.2.	Algoritmos de aprendizaje automático supervisado	21
4.6.	Métodos para la evaluación de modelos de probabilidad.....	24
4.6.1.	Curva Característica Operativa del Receptor (ROC)	25
4.6.2.	Coeficiente de Kappa de Cohen	26
5.	Metodología.....	27
5.1.	Zona de estudio.....	27
5.2.	Softwares y plataformas	28
5.3.	Establecer una línea base de las variables predictivas para identificar la ocurrencia de incendios forestales	29
5.4.	Construir un modelo de simulación espacial de zonas susceptibles a incendios forestales.....	34
5.4.1.	Etapa 1: Preparación de la base de datos	35
5.4.2.	Etapa 2: generación y aplicación del modelo	43
5.4.3.	Etapa 3: validación del modelo	43
5.5.	Automatización del modelo en Google Earth Engine	46
6.	Resultados	48
6.1.	Línea base de las variables predictivas para identificar la ocurrencia de incendios forestales.....	48
6.1.1.	Variables topográficas	48
6.1.2.	Variables de interacción antrópica (accesibilidad).....	49
6.1.3.	Variables de interacción antrópica (distancias).....	51
6.1.4.	Variables del estado de vegetación (Índices espectrales).....	53
6.2.	Resultados del modelo de simulación espacial que estime la probabilidad de ocurrencia de incendios forestales en el cantón Zapotillo.....	55
6.3.	Automatización del modelo de simulación espacial determinando el riesgo a incendios forestales.....	58
7.	Discusión.....	63
8.	Conclusiones.....	68
9.	Recomendaciones.....	69

10.	Bibliografía.....	70
11.	Anexos.....	81

Índice de tablas

Tabla 1. Índices espectrales de condiciones del suelo y vegetación.....	14
Tabla 2. Valores de precisión del modelo de acuerdo con el AUC.....	25
Tabla 3. Valoración del índice de Kappa.....	26
Tabla 4. Variables topográficas	29
Tabla 5. Variables de interacción antrópica	30
Tabla 6. Variables del estado de vegetación.....	32
Tabla 7. Criterios para determinar incendios forestales	36
Tabla 8. Coste de movilidad según la clase de superficie	40
Tabla 9. Clases para la reclasificación de la probabilidad de ocurrencia de incendios .	43
Tabla 10. Curva de ROC e índice Kappa de Cohen en R.....	45
Tabla 11. Resultados del Factor de Inflación de Varianza en las variables	56
Tabla 12. Importancia de las variables en la modelación de probabilidad de ocurrencia de incendios forestales utilizando el algoritmo RL	57
Tabla 13. Rendimiento del modelo de Regresión Logística (RL).....	57
Tabla 14. Cuantificación del riesgo a incendios forestales, basado en un análisis espacio temporal, para el cantón Zapotillo, enero - diciembre del 2020.....	60

Índice de figuras

Figura 1. Técnicas de Aprendizaje Automático aplicadas al estudio de incendios forestales.....	21
Figura 2. Árboles de decisión logística	23
Figura 3. Splines de Regresión Adaptativa Multivariante.....	24
Figura 4. Mapa base del cantón Zapotillo	27
Figura 5. Softwares y Plataformas para el modelo.....	29
Figura 6. Fases para la generación del mapa de probabilidad de ocurrencia de incendios forestales en el cantón Zapotillo.....	35
Figura 7. Distribución geográfica de los puntos de muestreo (incendio y no incendio)	37
Figura 8. Proceso para la obtención de información de las variables.	38
Figura 9. Diagrama de flujo del proceso de cálculo de accesibilidad	38
Figura 10. Digitalización de vías.....	39
Figura 11. Proceso para la generación de la base de datos.....	41
Figura 12. Identificación de las variables que no presentan multicolinealidad.....	42
Figura 13. Matriz de confusión	44
Figura 14. Rango de susceptibilidad a incendios forestales	46
Figura 15. Proceso del análisis espacio – temporal para la automatización del riesgo de incendios.....	46
Figura 16. Variables topográficas.....	49
Figura 17. Variables de interacción antrópica (accesibilidad)	51
Figura 18. Variables de interacción antrópica (distancia)	53
Figura 19. Variables del estado de la vegetación	54
Figura 20. Mapeo de probabilidad de ocurrencia de incendios forestales para el cantón Zapotillo utilizando algoritmos de aprendizaje automático supervisado: Regresión Logística (RL), Splines de Regresión Adaptativa Multivariante (MARS) y Árboles de Decisión Logística	56
Figura 21. Número de datos de los sets de entrenamiento y de validación que fueron clasificados como correctos (aciertos) e incorrectos (errores) para el modelo de RL....	58
Figura 22. Identificación de clústeres en las imágenes ráster de probabilidad de ocurrencia de incendios forestales del cantón Zapotillo.....	59

Figura 23. Zonificación del riesgo a incendios forestales, basado en un análisis espacio temporal, para el cantón Zapotillo, enero - diciembre del 2020.....	61
Figura 24. Inhomogeneidad, distinción y calidad global de cada clúster.....	62
Figura 25. Mapeo de la calidad global, inhomogeneidad y distinción del análisis clúster para la zonificación del riesgo de incendios forestales	62

Índice de ecuaciones

Ecuación 1. Modelo de Regresión Logística	22
Ecuación 2. Modelo de MARS	24
Ecuación 3. Sensibilidad de los datos	25
Ecuación 4. Especificidad de los datos	25
Ecuación 5. Índice de Kappa	26
Ecuación 6. Factor de Inflación de varianza.....	42
Ecuación 7. Calidad Global	47

Índice de Anexos

Anexo 1. Base de datos depurada.....81

Anexo 2. Certificado de traducción del abstract82

1. Título

Predicción de zonas susceptibles a incendios forestales aplicando técnicas de aprendizaje automático en el cantón Zapotillo, Loja – Ecuador

2. Resumen

Un incendio forestal afecta grandes extensiones de ecosistemas a nivel mundial, causando daños ambientales. En Ecuador se ha presenciado esta realidad, debido a que el 38% de los eventos naturales y antrópicos que afectaron al país, entre 2010 y 2018 fueron incendios forestales. La región de la Sierra fue la más afectada por los incendios, representando el 80% a nivel nacional. En esta región, la provincia de Loja se encuentra entre las más impactadas del país, con un 19% de los eventos registrados. En particular, el cantón Zapotillo fue afectado significativamente por estos eventos, debido a sus extensas áreas de bosque seco, albergando una diversidad florística y faunística significativa. En respuesta a esta problemática, la presente investigación modeló la probabilidad de ocurrencia de incendios forestales en el cantón Zapotillo. Se aplicaron tres algoritmos de aprendizaje automático supervisado (RL, MARS y LMT) con 164 puntos (incendio y no incendio) correspondientes al periodo 2018 – 2021, que sirvieron para entrenar (75 %) y validar (25 %) los modelos. Se utilizó ocho variables de entrada asociadas a factores ambientales, topográficos y antrópicos. El modelo de RL obtuvo el mejor rendimiento con un Área Bajo la Curva (AUC) de 0,65 (set de entrenamiento) y 0,63 (set de validación). Con este modelo, se generó un mapa de probabilidad de ocurrencia de incendios forestales (enero – diciembre 2020), lo que permitió automatizar la evaluación del riesgo mediante un análisis espacio-temporal, determinando que las zonas urbanas son las que presentan mayor probabilidad a incendios.

Palabras clave: *incendios forestales, aprendizaje automático, regresión logística, evaluación del riesgo de incendio.*

Abstract

The effects of forest fires on ecosystems worldwide cause significant environmental damage. Ecuador has experienced this reality, as forest fires accounted for 38% of the natural and anthropogenic events that affected the country between 2010 and 2018. The Sierra region was the most affected by fires, representing 80% of all fires at the national level. Among the provinces of this region, the province of Loja is one of the most affected, with 19% of the events registered. This event had a particular impact on the Zapotillo canton, due to its extensive areas of dry forest, which is home to a wide variety of flora and fauna. In response to this problem, the present research models the probability of the occurrence of forest fires in Zapotillo canton. Three supervised machine learning algorithms (RL, MARS and LMT) were applied to 164 points (fire and non-fire) corresponding to the period 2018 - 2021, which served to train (75 %) and validate (25 %) the models. Eight input variables associated with environmental, topographic and anthropogenic factors were used. The RL model obtained the best performance with an Area Under the Curve (AUC) of 0.65 (training set) and 0.63 (validation set). With this model, a map of probability of occurrence of forest fires (January - December 2020) was generated, which allowed automating the risk assessment through a spatio-temporal analysis, determining those urban areas are those with the highest probability of fires.

Keywords: *forest fires, machine learning, logistic regression, fire risk assessment.*

3. Introducción

El fuego es considerado como un regulador de ecosistemas, debido a que cumple un papel en la estructura, funcionamiento y dinámica de los medios terrestres, siendo parte fundamental de la historia evolutiva de la tierra, por medio de comportamientos climáticos y a la susceptibilidad natural, en especial en temporadas de sequía (Castillo *et al.*, 2003; Villers, 2006). Cuando el fuego deja de ser controlado en los ecosistemas, se convierte en un incendio forestal, cuya magnitud, duración y propagación dependerá en gran medida del tipo de vegetación presente en el ecosistema terrestre (bosque, pastizal, humedal, matorral, sabana, etc.) (Pausas, 2020). Los incendios forestales causan diversos efectos en los ecosistemas, como la degradación del suelo (cambios en el pH, erosión, aumento de la desertificación y/o disminución de su conductividad eléctrica), contaminación de las fuentes de agua y cambios en la composición de la atmósfera (Bodí *et al.*, 2012), deterioro de la flora y fauna debido a la disminución de materia orgánica y nitrógeno en el humus como resultado de la combustión de compuestos orgánicos y sustancias nitrogenadas (Pinta *et al.*, 2023). No obstante, la magnitud de estos eventos varía en función de la intensidad, frecuencia y duración de los eventos (González, 2017).

Di Bella *et al.* (2008) mencionan que los factores que influyen en la combustión, propagación y duración de los incendios están asociados a las condiciones de susceptibilidad que favorecen la probabilidad de expansión del fuego. Estos factores incluyen las características del material combustible, las condiciones meteorológicas (como la falta de precipitación, velocidad del viento, baja cantidad de humedad y altas temperaturas) (FAO, 2021), y las actividades antrópicas como la deforestación, los incendios y los cambios de uso del suelo para la conversión de bosques en asentamientos humanos (SADMQ, 2013). Además, se debe tener en cuenta la accesibilidad a vías, la proximidad a centros urbanos, la distancia a áreas pobladas, la cercanía a cuerpos de agua y la elevación (Reyes y Loján, 2022).

Según el Ministerio del Ambiente del Ecuador (MAE, 2013) y el Sistema de Alerta del Distrito Metropolitano de Quito (SADMQ, 2013), se estima que solo el 4% de los incendios forestales en el país son de origen natural, mientras que el 96% son causados por actividades antropogénicas. En Ecuador, El Servicio Nacional de Gestión de Riesgos y Emergencias (SNGRE, 2021), comenta que, desde el año 2018, se han registrado las siguientes pérdidas de cobertura vegetal: en el año 2018 con 4143.63 ha afectadas, en el año 2019 incrementó a 8503.50 ha afectadas, para el año 2020 disminuyó a 6024.00 ha afectadas y finalmente en el

año 2021 se documentaron 782.50 ha afectadas hasta el mes de septiembre, posicionando a la provincia de Loja como una de las más impactadas a nivel nacional (SNGRE, 2021). En el mismo informe del SNGRE, el último incendio registrado en 2021 ocurrió el 23 de septiembre en el cantón Zapotillo, sector Guasimal, con un origen desconocido y clasificado como de grado 3, afectando un total de 200 ha (SNGRE, 2021).

Por lo que, el cantón Zapotillo fue seleccionado como zona de estudio para este trabajo debido a sus extensas áreas de bosques secos, con la finalidad de generar información sobre la probabilidad que tienen estas zonas de sufrir incendios forestales, ya que los bosques secos en la provincia de Loja son considerados "áreas de interés especial" para la conservación (Chiriboga y Andrade, 2005), debido a que albergan una diversidad florística y faunística significativa, así como sus altos niveles de endemismo, al encontrarse en la zona de endemismo Tumbesino (Aguirre *et al.*, 2019). Sin embargo, en los últimos años, estos bosques han experimentado sobreexplotación y degradación causada por la extracción de madera, quemas para la expansión de la frontera agrícola y el sobrepastoreo de ganado (Mendoza *et al.*, 2021). Este deterioro ha provocado un aumento significativo en la frecuencia y la gravedad de los incendios forestales, fenómeno atribuido al cambio climático. Por consiguiente, resulta fundamental prever con la mayor precisión posible las zonas con mayor susceptibilidad a incendiarse. Esta información es fundamental para las autoridades encargadas de la gestión y planificación forestal, la asignación de recursos, los servicios de emergencia y los sistemas de alerta temprana (Tien Bui *et al.*, 2017).

Para llevar a cabo el estudio de la probabilidad de incendios forestales, es fundamental emplear la teledetección, dado que esta técnica posibilita la vigilancia de las condiciones de la vegetación tanto antes como después de un incendio mediante el uso de imágenes satelitales. Esto facilita la evaluación de la severidad del incendio, la tasa de pérdida de la cobertura vegetal y el proceso de recuperación de la misma (Di Bella *et al.*, 2008). El uso de imágenes satelitales es crucial para el análisis de los incendios forestales, volviéndose importantes en la implementación de modelos de aprendizaje automático basados en índices espectrales para el monitoreo de estos eventos (Gholinejad y Khesali, 2021).

Diversas investigaciones han abordado la predicción de incendios forestales mediante la aplicación de algoritmos de aprendizaje automático de Máxima Entropía (MaxEnt), Redes Neuronales Artificiales (ANN), Bosques Aleatorios (RF) y Máquinas de Vectores de Soporte (Valdez *et al.*, 2017). Además, Reyes y Loján (2022), proponen utilizar técnicas de aprendizaje

supervisado como árboles de decisión (LMT), regresión logística (RL) y splines de regresión adaptativa multivariante (MARS).

Por lo tanto, para abordar esta problemática se ha planteado la siguiente interrogante:

¿Qué tan eficiente es el modelo de predicción para identificar y mapear las zonas susceptibles a incendios forestales, utilizando herramientas de aprendizaje automático (Machine Learning) en el cantón Zapotillo, Loja – Ecuador?

Bajo este contexto, el objetivo general de este estudio es desarrollar un modelo de simulación espacial mediante la aplicación de técnicas de Machine Learning para identificar zonas susceptibles a incendios forestales en el cantón Zapotillo, Loja – Ecuador. Para lograr este propósito, se plantean los siguientes objetivos específicos: en primer lugar, establecer una línea base de las variables predictivas para identificar la ocurrencia de incendios forestales en el cantón Zapotillo; en segundo lugar, construir un modelo de simulación espacial utilizando técnicas de Machine Learning, que estime la probabilidad de ocurrencia de incendios forestales en diferentes áreas del cantón Zapotillo; y finalmente, automatizar el modelo de simulación espacial para identificar y delimitar áreas con diferentes niveles de riesgo a incendios forestales en el cantón Zapotillo.

4. Marco teórico

4.1. Bosques secos

Los bosques secos son áreas donde se asientan las poblaciones debido a que cuenta con suelos adecuados para la agricultura. Debido a esto, estos bosques han venido experimentando una mayor intervención y destrucción en comparación con los bosques húmedos (Aguirre *et al.*, 2006). Estos bosques son ecosistemas vegetales caracterizados por una precipitación anual que varía entre 350 y 800 mm, y una temporada seca que dura al menos cinco a seis meses (Aguirre *et al.*, 2014).

Los bosques secos en la región occidental de Ecuador y el noroeste de Perú, conforman el “Centro de Endemismo Tumbesino”, abarcando una extensión de 135 000 km², ubicados desde Manabí hasta la costa del norte de Perú, siendo un área reconocida a nivel mundial por su alta biodiversidad y la presencia de especies endémicas, lo que la convierte en una de las regiones más destacadas en términos de conservación biológica. Además, se considera uno de los hotspots de biodiversidad del planeta, debido a la concentración de especies únicas y amenazadas que alberga (Aguirre *et al.*, 2014; Paladines, 2003). Asimismo, el Centro de endemismo Tumbesino, cuenta con un alto número de especies de aves endémicas registradas, además posee vegetación de tipo costera, con flora adaptada a ecosistemas en condiciones de aridez, considerándose el 31% de la superficie total de la provincia de Loja, como bosque seco (Lozano, 2002).

Según Ramírez *et al.* (2018) la protección y conservación del bosque seco son fundamentales tanto para los organismos que lo habitan como para la vida humana. Es importante reconocer los servicios ecosistémicos que este ecosistema nos brinda, es decir, los beneficios que la sociedad puede obtener de él. Estos servicios incluyen el suministro de alimentos, especies ornamentales y materiales de construcción, la regulación del clima, la calidad del agua y la prevención de la erosión del suelo, así como aspectos culturales como elementos ceremoniales, ecoturismo e identidad. Por lo tanto, estas zonas se consideradas como imprescindibles para el desarrollo en la vida del planeta. Sin embargo, todos estos procesos están amenazados por actividades antropogénicas como la degradación de suelo, deforestación, incendios forestales, expansión agrícola, desarrollo urbano, entre otros (CNF, 2010).

4.2. Incendios forestales

Los incendios forestales se generan cuando la vegetación del área se quema debido a un fuego que puede ser causado tanto por factores naturales como por actividades antrópicas.

Se entiende como un incendio forestal a la propagación descontrolada del fuego sobre la vegetación aledaña, incluyendo árboles, arbustos, pastizales y/o cultivos. Cabe recalcar que en la actualidad se usan de manera indistinta los términos "fuego" e "incendio", el primero se refiere al elemento en sí, mientras que el segundo es una expresión que describe su manifestación (Plana *et al.*, 2016).

Los incendios forestales son comunes en todas las áreas de Ecuador, pero la región de la Sierra es la más afectada. Esto se debe a la presencia de extensas áreas de bosques de *Eucalyptus* introducidos, lo que proporciona una gran cantidad de combustible para el desarrollo de incendios forestales. Por lo tanto, la región de la Sierra es altamente susceptible a este tipo de amenaza (Anchaluisa, 2013).

Los incendios forestales pueden ser superficiales, subterráneos o de copa. En el primer caso, se da cuando la propagación del fuego se produce de manera horizontal sobre la superficie del terreno alcanzando una altura de hasta metro y medio. Estos incendios afectan a los combustibles vivos y muertos presentes en el área, como pastizales, hojas, ramas, ramillas, arbustos, pequeños árboles en regeneración natural o plantados, troncos y humus, entre otros (CNF, 2010). Los incendios subterráneos ocurren cuando un incendio de superficie se propaga hacia el subsuelo. Durante este proceso, la ignición afecta la materia orgánica acumulada y en las raíces, pudiendo incluso llegar hasta los afloramientos rocosos. En general, estos incendios carecen de llamas visibles y presentan una baja emisión de humo (CNF, 2010).

Finalmente, los incendios de copa, son considerados los más devastadores peligrosos y difíciles de contener, debido a que consumen toda la vegetación en su trayectoria. Estos también comienzan de manera superficial, pero en este caso, las llamas se propagan verticalmente desde el nivel del suelo, escalando la vegetación hacia arriba que actúa como combustible en forma de escalera hasta llegar a las copas de los árboles (CNF, 2010).

4.2.1. Probabilidad a que ocurran incendios forestales

Según Pérez-Verdín *et al.* (2013), los factores que influyen en la magnitud de los incendios se pueden dividir en dos categorías principales: ambientales y antropogénicos. Los factores ambientales incluyen la precipitación, la temperatura y la altitud. Estos elementos pueden afectar la disponibilidad de combustible y la propagación del fuego. Por otro lado, los factores antropogénicos están relacionados con la actividad humana y pueden incluir la distancia a caminos, la distancia a localidades y la distancia a zonas deforestadas. Estos factores pueden influir en la probabilidad de inicio de incendios y en la capacidad de respuesta y control de los mismos.

Verdú y Salas (2011) mencionan que la frecuencia de los incendios forestales está influenciada por una variedad de factores ambientales y humanos, los cuales determinan la probabilidad de inicio y propagación del fuego. Estos factores pueden ser agrupados en cuatro categorías principales: condiciones meteorológicas, tipo y estado de la vegetación, variables topográficas y actividades humanas. Así mismo, Oliveira *et al.* (2012) menciona que los factores físicos representan las características naturales del territorio, mientras que los factores humanos son variables significativas debido a su influencia en la aparición de incendios forestales.

Las variables comúnmente utilizadas para la determinación de la probabilidad de ocurrencia de incendios forestales implican elementos ambientales, factores físicos y humanos, cuya influencia se describe a continuación.

4.2.1.1. Factores ambientales que facilitan en la ocurrencia y propagación de incendios forestales

Para que se inicie un incendio, se requiere la presencia simultánea de tres componentes esenciales: una fuente de calor, un material combustible y oxígeno. Una vez que se ha iniciado, la propagación del fuego se ve afectada por tres factores principales: las características del combustible, las condiciones climáticas y la topografía (Díaz *et al.*, 2013).

En base a lo que comentan Díaz *et al.*, (2013) el combustible de un incendio forestal consiste en materiales vegetales, tanto vivos como muertos, que son capaces de encenderse y arder cuando se exponen a una fuente de calor (e.g. hojas, acículas, ramas y humus), que se acumulan de forma natural en el ecosistema debido al proceso de descomposición de la vegetación. Sin embargo, también puede haber combustibles generados por actividades humanas, como la deforestación, que contribuyen a la disponibilidad de material inflamable en el entorno forestal. Lo que respecta a fuente de calor, se refiere a que en un bosque no puede ocurrir una combustión espontánea, es decir, no se produce un incendio sin una fuente externa de calor. Para que se produzca un incendio en un bosque, se necesita una fuente de incandescencia externa que tenga una temperatura mayor a 200°C (CNF, 2010).

El análisis del clima es fundamental para comprender los incendios forestales teniendo en cuenta variables como la humedad relativa, la temperatura del aire, la precipitación y la velocidad del viento. Estas variables meteorológicas influyen en todos los aspectos de los incendios forestales, incluyendo su inicio, propagación e intensidad. Este tipo de eventos ocurren cuando se combinan altas temperaturas, baja humedad relativa y niveles de precipitación que han secado la vegetación hasta un punto inflamable. Además, se requiere una

fuelle de combustión, ya sea natural o causada por actividades humanas (Pazmiño, 2019). La velocidad del viento es un estimulante en la intensidad de un incendio forestal debido a que reduce la cantidad de humedad en la vegetación, mientras provee de oxígeno para que se dé la combustión, además, la inflamabilidad y la distribución de la vegetación son controladas por la precipitación de la zona, los cuales influyen en la intensidad de un incendio forestales (Haynes Bradstock *et al.*, 2012).

Por último, están los aspectos topográficos (elevación, pendiente y aspecto). La pendiente es un factor que influye en la velocidad y dirección de la propagación de un incendio (Castellnou *et al.*, 2009), además, tiene relación con el tamaño del fuego, si el incendio desciende por pendientes pronunciadas la velocidad de propagación del fuego disminuye, por otro lado, si el incendio asciende, afectará más extensiones de terreno, esto se debe a que la disposición típica de las flamas en altura y la tendencia del aire caliente a ascender ocasionan un impacto más directo en la vegetación de las partes altas (Ruiz, 2004). Sin embargo, el resultado puede variar dependiendo de otros factores y fuerzas involucradas en el fenómeno (Castellnou *et al.*, 2009). La elevación y el aspecto del terreno influye en la distribución y composición de la vegetación, además de la humedad del material combustible, condicionando el estado del combustible, que se seca más rápido en las laderas expuestas al sol, apareciendo mayor cantidad de combustible (Verdú y Salas, 2011).

4.2.1.2. Factores antrópicos que contribuyen en la ocurrencia y propagación de incendios forestales

Según Reyes y Córdova (2021) los incendios forestales se dan a causa de actividades antrópicas, las cuales, en su mayoría están cerca de las poblaciones. La FAO (2021) menciona que la gran parte de los incendios son causados por la mano del hombre. Así mismo, actualmente, la mayoría de los incendios forestales en Ecuador son de naturaleza antropogénica, en el que más del 95% de estos eventos son ocasionados por las actividades del hombre, siendo el cambio y uso de suelo a agrícola la principal actividad causante de los incendios, seguidamente del aumento de asentamientos humanos, el aumento de la deforestación y la pérdida de conocimientos tradicionales sobre el manejo del fuego en áreas agrícolas (ESA, 2015; Pazmiño, 2019).

Un estudio realizado en España, menciona que gran parte de su territorio se vio afectada por el incremento de estructura poblacional, los usos agrícolas y ganaderos por motivos económicos, la quema extensiva de restos forestales, pastizales y matorral (Pernas *et al.*, 2011). Bajo este contexto, los investigadores han considerado importante el uso de variables

antrópicas en los estudios de probabilidad a incendios forestales para tener una mejor precisión en base al territorio. Un caso sería de Vilar del Hoyo (2008), el cual en su estudio usó variables de infraestructura (carreteras, líneas eléctricas, pistas, campos de tiro, vías de ferrocarril), y socioeconómicas (variación de la población, hoteles, áreas recreativas).

4.2.2. Efectos de los incendios forestales

Los efectos en la vegetación después de un incendio son la modificación de su estructura, composición y servicios ecosistémicos, provocando el reemplazo de especies nativas por especies invasoras/exóticas, una vez que se dan estos reemplazos de especies, puede resultar en una menor cantidad de precipitación en la zona, cambios en la temperatura y humedad del suelo (González, 2017).

Según González (2017) los incendios forestales pueden afectar a los animales vertebrados e invertebrados de diversas maneras, como la muerte de individuos, la pérdida de hábitat, territorios, refugio y alimentación, ocasionando desequilibrios en los ecosistemas, como la destrucción de vegetación y cambios en la composición de especies. Además, provocan migraciones de animales, aves, insectos y microorganismos, así como desequilibrios ecológicos y rupturas en la cadena biológica.

Los incendios forestales también tienen un impacto negativo en la actividad biológica del suelo, lo que provoca una alteración en el ciclo de nutrientes. Esta alteración se debe a la reducción de los microorganismos que descomponen los compuestos orgánicos y facilitan la disponibilidad de nutrientes para las plantas (González, 2017). Durante un incendio se puede alcanzar temperaturas de 1200 a 1400 °C dentro de la masa de combustión, estas incidencias de calor, modifican las propiedades físico-químicas y biológicas del suelo. En el caso de que existan varios incendios en una misma zona, se incrementa la erosión, disminuye su fertilidad, dando como resultado un suelo pobre en nutrientes (Cuesta y Giraldo, 2013).

4.2.3. Incendios forestales en bosques secos

Los incendios forestales ocurren en la mayoría de ecosistemas del planeta, aunque su propagación depende de la cantidad de combustible que se dispone para que ocurra el fuego y la cantidad de humedad que presenta dicho combustible (Mondal y Sukumar, 2016). Los bosques secos son uno de los ecosistemas forestales más vulnerables del planeta, debido a la fuerte alteración causada por la actividades naturales y actividades antropogénicas. Estos bosques tienen una estación seca que puede extenderse hasta 8 meses, volviéndolos

susceptibles a incendios forestales todo el año, debido a que tienen una precipitación media por mes inferior a 100 mm (Calvo *et al.*, 2013).

Durante la estación seca se produce una acumulación de hojarasca y restos de madera en el suelo del bosque debido a la caída de las hojas de los árboles caducifolios y los vientos intensos, convirtiéndose en combustible, creándose un entorno propicio para la propagación de incendios forestales (Otterstrom *et al.*, 2006). Los incendios forestales en bosques secos son ocasionados en gran medida de manera intencional por agricultores con la intención de expandir sus campos agrícolas o realizando limpieza de los pastos, esto debido a que las condiciones climáticas y los suelos de estos bosques presentan alta fertilidad y buenas propiedades físicas, volviéndolas atractivas para el desarrollo de actividades antropogénicas (Vargas y Campos, 2018).

Los informes del Servicio Nacional de Gestión del Riesgo y Emergencia indican que en Ecuador entre 2010 y 2018 los incendios forestales representaron el 38% de los fenómenos naturales que generan riesgo en el país. La región sierra es la más afectada por los incendios, en el que se registró que un 80% de los incendios a nivel Nacional ocurrieron dentro de esta región. Por otro lado, la región costa sufrió un 18% de los incendios forestales, mientras que la región amazónica e insular solo experimentó un 2% de afectación (Banco de Desarrollo de América Latina, 2020). Estos datos reflejan la distribución desigual de los incendios forestales debido a la presencia de bosques secos en diferentes regiones del país.

4.3. Detección de incendios forestales

Los incendios forestales representan una importante preocupación en términos ambientales, debido a que, si no se logra detectar el evento, genera impactos negativos irreversibles en los ecosistemas, la economía y la vida humana. Por lo tanto, la detección temprana de estos incendios se convierte en una herramienta fundamental para combatir de manera más efectiva esta problemática (Cárdenas *et al.*, 2015).

El análisis de incendios forestales es un procedimiento de alta complejidad debido a la amplia gama de factores que afectan su aparición y comportamiento. Por lo tanto, la utilización de tecnologías y aplicaciones predictivas resultan ser esenciales para generar información relevante en la gestión de incendios forestales. Este método de detección se basa en un conjunto de procedimientos, para descubrir y anticipar la localización de un posible incendio, por medio del uso de focos de calor, índices de vegetación y teledetección.

La detección de los focos de calor se realiza utilizando sensores remotos que capturan la energía emitida por la superficie en las longitudes de onda del infrarrojo medio y térmico.

Los umbrales para identificar los focos de calor se establecen en función de las temperaturas alcanzadas por los incendios, que pueden variar entre 300 °C y 1500 °C dependiendo del tipo de vegetación, la cantidad de combustible acumulado y las condiciones ambientales, entre otros factores (Di Bella *et al.*, 2008).

4.3.1. Teledetección

La teledetección se basa en la observación de objetos desde una distancia en la que no exista contacto directo con ellos. Los sensores remotos pasivos están posicionados en diferentes plataformas, y se encargan de capturar la energía electromagnética reflejada por el sol y emitida por la superficie terrestre en diferentes longitudes de onda. Estas longitudes de onda incluyen el espectro visible (0,4-0,7 μm), el infrarrojo cercano (0,7-1,3 μm), el infrarrojo medio (1,3-8 μm) y el infrarrojo térmico (8-14 μm) (Di Bella *et al.*, 2008). Para identificar la susceptibilidad a incendios forestales se han utilizado diferentes métodos y algoritmos a diferentes escalas, ecosistemas y localidad a través de los sistemas de información geográfica y técnicas de teledetección. En la actualidad se ha viralizado el uso de técnicas de minería de datos, debido a que permiten el análisis de problemas complejos que contienen múltiples variables, construyendo modelos predictivos a partir de extensas bases de datos, también conocidos como técnicas de aprendizaje automático (Reyes y Balcázar, 2021).

La precisión para detectar y estimar los incendios forestales está influenciada por la calidad de la imagen y la metodología empleada. Por esta razón, los investigadores desarrollan metodologías que permiten transformar los datos originales de los satélites con el objetivo de mejorar la capacidad de discriminación de las áreas afectadas por el fuego. Una de las transformaciones más comunes utilizadas son los índices espectrales (Gómez y Martín, 2008).

4.3.2. Índices espectrales:

Los índices espectrales son una herramienta que proporciona información de forma rápida y no destructiva sobre extensiones territoriales grandes, con una resolución espacial moderada y una alta frecuencia temporal. Esto los hace útiles para el monitoreo continuo de cambios en la vegetación, el suelo y otros aspectos del paisaje a lo largo del tiempo (Buzzi *et al.*, 2017). Por lo tanto, los índices espectrales en un incendio forestal, tienen la función de resaltar las áreas afectadas por incendios en comparación con otras coberturas, ya sea la vegetación existente antes del incendio (cuando se utiliza un enfoque de análisis multitemporal) o la vegetación que surge después del incendio (Gómez y Martín, 2008).

Los índices de vegetación se los usa en el ecosistema como indicadores cuantitativos, esto se debe a que involucra el verdor de las plantas, la tierra sin vegetación y la presencia de agua que se basa en datos obtenidos por satélites (Carbajal y Onofre, 2020). Por medio de la teledetección se puede obtener imágenes multitemporales que son usadas para la evaluación y monitoreo de la vegetación a diferentes niveles (global, regional, nacional y local) (Muñoz, 2013). Por lo tanto, un Índice de Vegetación se lo puede considerar como un parámetro calculado a partir de reflectancia en diferentes longitudes de onda, y que muestra una sensibilidad especial hacia la vegetación que cubre el área (Muñoz, 2013).

Algunos de los índices espectrales utilizados en estudios de condiciones de vegetación y suelo, mostrados en la Tabla 1, incluyen el Índice de Vegetación de la Diferencia Normalizada (NDVI), el Índice de Suelo Desnudo (BSI), el Índice de Humedad (Moisture), el Índice de Vegetación Ajustada al Suelo (SAVI), el Índice de Estrés Hídrico (MSI), el Índice de Calcinación Normalizada (NBR), el Índice de Diferencia Normalizada de Humedad (NDMI), el Índice de Vegetación Atmosféricamente Resistente (ARVI), el Índice de Vegetación de Diferencia Normalizada Verde (GNVI) y el Índice de Vegetación Normalizada Verde (NDWI). Estos índices proporcionan información valiosa para el monitoreo y análisis de incendios forestales (De La Casa y Ovando, 2006; Gilabert, 1997; González, 2020; Reyes y Loján, 2022; Solis *et al.*, 2021).

Tabla 1. Índices espectrales de condiciones del suelo y vegetación

Índice	Descripción	Rango	Interpretación
Índice de Vegetación de la Diferencia Normalizada (NDVI)	El NDVI es capaz de proporcionar una estimación de la cantidad, calidad y desarrollo de la vegetación al medir la biomasa verde, que se determina a través del área foliar y el contenido de clorofila.	-1 a -0.1	Superficies con escasa cubierta vegetal
		-0.1 a 0.1	Superficies con vegetación media
		0.1 a 1	Superficies con alta cubierta vegetal

Nota: Adaptado de Informe de variables que se utilizan para la generación de modelos de susceptibilidad de incendios con la identificación de sus fuentes para descarga por Loján (2020).

Índice	Descripción	Rango	Interpretación
Índice de Diferencia Normalizada de Humedad (NDMI)	El NDMI permite determinar el contenido de agua de la vegetación, lo que a su vez posibilita la identificación de áreas con problemas de estrés hídrico.	-1 a -0.8	Suelo desnudo
		-0.8 a -0.6	Cubierta de dosel casi ausente
		-0.6 a -0.4	Cubierta de dosel muy baja
		-0.4 a 0	Cubierta de dosel seca
		0 a 0.2	Cubierta de dosel medio baja, bajo estrés hídrico
		0.2 a 0.4	Cubierta de dosel media alta, estrés hídrico medio
		0.4 a 0.6	Cubierta de dosel alta, sin estrés hídrico
		0.6 a 0.8	Cubierta de dosel muy alta, sin estrés hídrico
		0.8 a 1	Cubierta total de dosel, sin estrés hídrico
Índice de Vegetación Atmosféricamente Resistente (ARVI)	El ARVI es una versión ajustada del NDVI que tiene en cuenta la dispersión atmosférica en la banda de reflectancia roja mediante el uso de mediciones en longitudes de onda azules.	-1 a -0.1	Cuerpos de agua
		-0.1 a 0.1	Áreas rocosas, arenosas o nieve
		0.1 a 0.4	Hierbas y arbustos
		0.4 a 1	Árboles
Índice de Vegetación de Diferencia Normalizada verde (GNDVI)	El GNDVI se utiliza para evaluar la cantidad y salud de la vegetación en un área	-1 a -0.2	Suelo desnudo

determinada, es más sensible a la variación de la clorofila que el NDVI.	-0.2 a 0.2	Vegetación moderada
	0.2 a 1	Vegetación sana

Nota: Adaptado de Informe de variables que se utilizan para la generación de modelos de susceptibilidad de incendios con la identificación de sus fuentes para descarga por *Loján (2020)*.

Índice	Descripción	Rango	Interpretación
Índice de Suelo Desnudo (BSI)	El BSI identifica zonas con escasa vegetación, usando un indicador numérico que combina bandas espectrales azules, rojas, infrarrojas cercanas e infrarrojas de onda corta para capturar las variaciones del suelo.	-1 a 0	Vegetación arbustiva
		0 a 0.6	Vegetación herbácea
		0.6 a 1	Suelo desnudo
Índice de Calcinación Normalizada (NBR)	El NBR detecta las áreas quemadas y monitorea la recuperación del ecosistema por medio de bandas espectrales de infrarrojo cercano e infrarrojo de onda corta, que son sensibles a las variaciones en la vegetación.	< -0.25	Alto crecimiento de vegetación posterior al fuego
		-0.25 a -0.1	Bajo crecimiento de vegetación posterior al fuego
		-0.1 a 0.1	Zonas estables o sin quemar
		0.1 a 0.27	Cubierta de dosel seca
		0.27 a 0.44	Cubierta de dosel seca
		0.44 a 0.66	Zonas quemadas con gravedad moderada-baja
		> 0.66	Zonas quemadas con gravedad alta
Índice de Humedad (Moisture)	El Moisture es una medida que indica el nivel de humedad presente en el suelo y la vegetación	< 0.2	Suelo desnudo
		0.2 a 0.8	Vegetación Sana

		> 0.8	Suelo cubierto por vegetación
		-1 a -0.3	Superficies sin agua, secúas
Índice de Agua de Diferencia Normalizada (NDWI)	El NDWI detecta y controla los cambios leves en la cantidad de agua en las masas de agua, realzando estas cantidades en las imágenes satelitales	-0.3 a 0	Sequía moderada
		0 a 0.2	Inundación, humedad alta
		0.2 a 1	Superficie del agua

Nota: Adaptado de Informe de variables que se utilizan para la generación de modelos de susceptibilidad de incendios con la identificación de sus fuentes para descarga por *Loján (2020)*.

Índice	Descripción	Rango	Interpretación
Índice de Vegetación Ajustada al Suelo (SAVI)	SAVI es una medida de la vegetación que corrige el NDVI para tener en cuenta la influencia del suelo	-1 a -0.1	Superficies con escasa cubierta vegetal
		-0.1 a 0.1	Superficies con vegetación media
		0.1 a 1	Superficies con alta cubierta vegetal

Nota: Adaptado de Informe de variables que se utilizan para la generación de modelos de susceptibilidad de incendios con la identificación de sus fuentes para descarga por *Loján (2020)*.

4.4. Satélites utilizados para el monitoreo de los incendios forestales

En esta sección se detalla el uso de satélites como medios para monitorear los incendios forestales, incluyendo una descripción de los satélites específicos utilizados con este propósito y resaltando su importancia en la detección temprana, seguimiento y evaluación de la magnitud de los incendios forestales.

4.4.1. Sentinel-2

Este satélite fue creado con el propósito de recolectar información en diversas franjas del espectro electromagnético, que van desde la luz visible hasta el infrarrojo térmico en un tiempo de revisión de 10 días. Emplea tecnología avanzada de radar y óptica de alta precisión

para obtener imágenes minuciosas de la superficie del planeta, lo cual facilita la supervisión y el estudio de distintos fenómenos y alteraciones medioambientales (NASA EARTHDATA, 2020). Desde 2015, se han empleado imágenes satelitales de las misiones Sentinel-2 para detectar incendios en diversas regiones del mundo casi en tiempo real. Este satélite dispone de las combinaciones de bandas necesarias para aplicar varios índices que permiten identificar intentos de incendios y analizar variables como la severidad (González, 2020), debido a que cuenta con una cámara multispectral de alta resolución, con trece bandas espectral, para poder tener una nueva perspectiva de la superficie terrestre y de la vegetación (ESA, 2015).

El uso de los satélites Sentinel-2 ha demostrado ser extremadamente útil en una amplia variedad de aplicaciones. Su capacidad para proporcionar datos precisos y actualizados a nivel mundial ha mejorado significativamente la capacidad de los científicos, los responsables políticos y otros usuarios para tomar decisiones informadas en relación con el medio ambiente y los recursos naturales (Gascón *et al.*, 2017).

Según la ESA (2015) y Pérez *et al.* (2020), el satélite Sentinel-2A resultó de una colaboración estrecha entre la Agencia Espacial Europea (ESA), la Comisión Europea, la industria, proveedores de servicios y usuarios de datos. Donde la construcción y el desarrollo se llevó a cabo mediante un consorcio de 60 compañías, encabezado por Airbus Defence and Space, con el respaldo de la agencia espacial francesa CNES para mejorar la calidad de las imágenes, y del Centro Aeroespacial Alemán DLR para optimizar la recuperación de datos a través de comunicaciones ópticas. Además, el satélite Sentinel-2A se distingue por tener bandas multispectrales mejoradas, una resolución espacial de 10 metros para las bandas del espectro visible RGB y un intervalo de revisita de 5 días (Brovkina *et al.*, en 2020).

4.4.2. Satélite SNPP-VIIRS

En el satélite Suomi National Polar-orbiting Partnership (SNPP) está presente la herramienta Visible Infrared Imaging Radiometer Suite (VIIRS), siendo el instrumento más grande a bordo de este satélite, el cual recopila imágenes radiométricas en longitudes de onda visibles e infrarrojas de la tierra, la atmósfera, el hielo y el océano. Los datos recopilados por VIIRS, provenientes de 22 canales que abarcan todo el rango del espectro electromagnético, se emplean para la observación de incendios en curso, vegetación, tonalidades del océano, temperatura de la superficie marina y otras características de la superficie (Keck *et al.*, 2011). El satélite SNPP-VIIRS cuenta con el sensor DNB (Day/Night Band), volviéndolo más completo, debido a que cuenta con 11 bandas espectrales de noche y 21 de día. Estas bandas posibilitan una mayor distinción de niveles bajos de luminosidad y la detección y

discriminación de fuentes térmicas de emisión de luz, como los incendios o la quema de biomasa (Checa *et al.*, 2016).

Schroeder y Giglio (2017) y Zhang *et al.* (2023) comentan que los datos de incendios obtenidos a través de VIIRS han demostrado una mejora significativa en la resolución espacial, el alcance de observación y la calidad de los datos, en comparación con sus predecesores, lo que ha resultado en una capacidad mejorada de los datos de incendios de VIIRS para respaldar la gestión de incendios, los sistemas de alerta de incendios forestales casi en tiempo real y otras aplicaciones científicas que requieren una mayor precisión en la cartografía.

4.5. Inteligencia Artificial

La Inteligencia Artificial (IA) busca analizar y reproducir el comportamiento humano, incluyendo la comprensión, percepción, resolución de problemas y toma de decisiones, con la ayuda de computadoras. Esto implica la capacidad de imitar funciones cognitivas humanas como la creatividad, la sensibilidad, el aprendizaje, la comprensión, la percepción del entorno y el uso del lenguaje, con la posibilidad de incluso mejorar las capacidades humanas en estos aspectos. (Milla, 2014; Tortajada, 2023). Una rama de la inteligencia artificial, son las técnicas de aprendizaje automático, que se centra en el desarrollo de algoritmos y modelos que permiten a las máquinas aprender patrones a partir de datos y mejorar su rendimiento con la experiencia, sin una programación explícita (Pfeiffer, 2021).

4.5.1. Técnicas de aprendizaje automático

En la actualidad, el progreso de las tecnologías de la información y la comunicación (TIC) ha tenido un impacto significativo en diversas áreas. Esto ha llevado a un aumento considerable en la disponibilidad de información, la cual debe ser procesada y gestionada para obtener una extensa cantidad de datos, siendo útiles para la toma de decisiones. Con esta finalidad, es vital el uso de técnicas de aprendizaje que permitan operar los datos y obtener información de ellos (Sedeño *et al.*, 2022).

Se conoce a las técnicas de aprendizaje automático (Machine Learning) como una inteligencia artificial que le permite a un sistema extraer y aprender conocimiento de datos proporcionados, sin la necesidad de programarlo (Villegas *et al.*, 2022). Para emplearlo, es necesario suministrarle una base de datos sustancial, a partir de la cual aprenderá y generará predicciones sobre nuevos datos o datos que aún no han sido detectados. (Sandoval, 2018).

En el contexto de la aplicación de técnicas de aprendizaje supervisado, se parte de un conjunto de datos iniciales que consisten en entradas y salidas observadas para una serie de

eventos específicos. Estas entradas representan los factores que influyen en la ocurrencia de incendios forestales, mientras que las salidas corresponden a los resultados asociados a esos eventos. Mediante la utilización de estos datos, es posible construir un modelo que permitirá predecir los resultados para nuevos eventos. Este modelo se basa en los patrones identificados en los datos de entrenamiento y se utiliza para hacer predicciones sobre eventos futuros (Cárdenas *et al.*, 2015). Estas técnicas se pueden clasificar en dos grupos principales: aprendizaje no supervisado, aprendizaje supervisado.

- Aprendizaje no supervisado: en esta técnica, se le proporciona datos de entrada (características) al algoritmo, mas no las etiquetas, debido a que deseamos que el algoritmo agrupe lo datos que le proporcionamos según sus características, buscando patrones en los datos de entrada, por lo tanto, el algoritmo solo infiere que, al compartir ciertas características, es probable que pertenezcan al mismo grupo, además, puede ayudar a predecir posibles datos futuros del mismo fenómeno (Jain *et al.*, 2020; Sancho, 2020; Sandoval, 2018).
- Aprendizaje supervisado: En esta técnica, entrenamos el modelo proporcionándole preguntas (características) y respuestas (etiquetas), lo que implica la predicción de resultados futuros a partir de datos de entrada y salida conocidos para deducir una función que permita predecir los valores correspondientes a los datos de entrada del modelo después del entrenamiento (Jain *et al.*, 2020; Moncayo, 2013; Sandoval, 2018)

Las técnicas de aprendizaje automático cuentan con dos fases de desarrollo, también conocido como validación cruzada, en la que se divide la base de datos para entrenar y validar el modelo:

- Fase de entrenamiento: Durante esta etapa, se utiliza gran parte de la base de datos, de los cuales se separa una porción para entrenar al algoritmo, proporcionándole toda esta información para que identifique los patrones necesarios y pueda realizar predicciones posteriormente (Sandoval, 2018).
- Fase de validación: Los datos restantes se reservarán para realizar pruebas. De esta manera, podemos plantear preguntas al algoritmo y evaluar la precisión de sus respuestas para determinar si está aprendiendo de manera efectiva. Si observamos que las respuestas no coinciden con los datos, será necesario agregar más información o modificar el método utilizado. Sin embargo, si notamos que entre el 80% y el 90% de

las respuestas son correctas, podemos concluir que el algoritmo ha aprendido de manera satisfactoria y está listo para su uso (Sandoval, 2018).

Entre los algoritmos más comunes del aprendizaje supervisado en machine learning, se encuentran los Random Forest (RF), Naive Bayes (NB), Máxima Entropía (MaxEnt), Árboles de Decisión (DT), Red Neuronal Profunda (DNN), Redes Neuronales Artificiales (ANN), Árboles de Decisión Logística (LMT), Regresión Logística (LR), y Splines de Regresión Adaptativa Multivariante (MARS).

4.5.2. Algoritmos de aprendizaje automático supervisado

En la Figura 1, se observa los diversos algoritmos de aprendizaje automático, empleados para determinar la ocurrencia a incendios forestales.

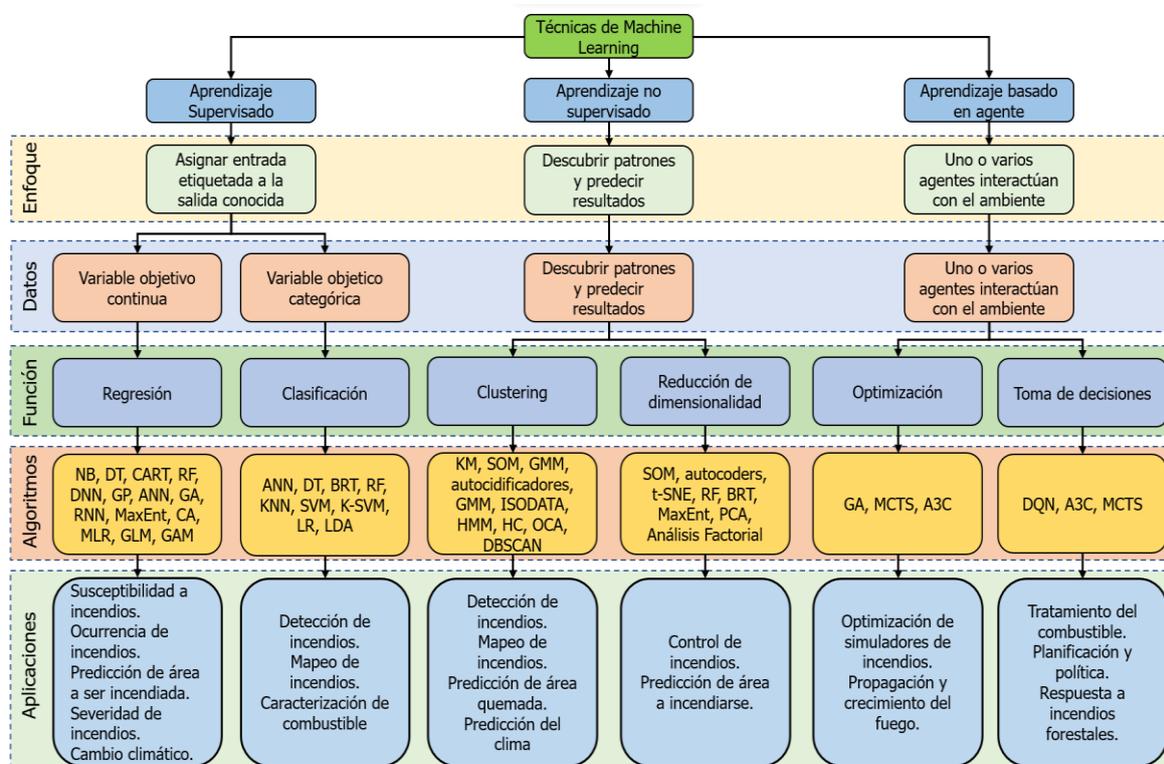


Figura 1. Técnicas de Aprendizaje Automático aplicadas al estudio de incendios forestales.

Nota: Adaptado de A review of machine learning applications in wildfire science and management. *Environmental Reviews* por Jain, P., Coogan, S. C. P., Subramanian, S. G., Crowley, M., Taylor, S., y Flannigan, M. D. (2020).

En este estudio, utilizamos tres algoritmos: regresión logística (RL), los árboles de decisión logística (LMT) y los splines de regresión adaptativa multivariante (MARS).

4.5.2.1. Regresión Logística (RL)

Los modelos de RL explican cuál es la relación que existe entre una variable dependiente y un conjunto de variables independientes continuas o categóricas, así como

cuantificar las relaciones y clasificar. Se lo considera como uno de los algoritmos de aprendizaje automático más sencillos y utilizados. Los modelos de RL son lineales, por lo que su desempeño es óptimo cuando los datos son linealmente separables a través de una frontera de decisión (Ibáñez, 2019; Vilar del Hoyo *et al.* 2002).

La RL ha sido empleada para predecir la ocurrencia de incendios forestales y para analizar los factores que influyen en este evento, en diversas regiones y escalas (Guo *et al.*, 2016). La regresión logística tiene la siguiente expresión (Ecuación 1):

Ecuación 1. Modelo de Regresión Logística

$$\text{logit}(P_i) = \ln\left(\frac{P_i}{1-P_i}\right) = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n$$

en donde p es la probabilidad de ocurrencia de incendios forestales, n es el número de covariables, $(\alpha_1, \alpha_2, \dots, \alpha_n)$ son los coeficientes para cada variable y (x_1, x_2, \dots, x_n) son las variables que influyen en la ocurrencia de incendios forestales.

4.5.2.2. Árboles de decisión logística (LMT)

Según Chen *et al.* (2017), el modelo de clasificación que combina los métodos de aprendizaje de árboles de decisión y regresión logística se basa en el algoritmo LogitBoost. En este modelo, el algoritmo LogitBoost se utiliza para generar un modelo de regresión logística en cada nodo del árbol, así como se muestra en la Figura 2. El ajuste del árbol se realiza mediante el algoritmo CART (Classification and Regression Trees). Asimismo, Charris *et al.* (2014) señalan que el modelo LMT (Logistic Model Tree) ofrece una representación efectiva de los datos. Este modelo combina una estructura de árbol de decisión con funciones de regresión logística en las hojas. Al igual que en los árboles de decisión convencionales, cada nodo interno del árbol LMT se asocia con una prueba sobre un atributo específico.

Para evitar el sobreajuste de los datos de entrenamiento, se utiliza una técnica de validación cruzada. Esto implica encontrar un número óptimo de iteraciones que permita obtener un equilibrio entre el rendimiento del modelo y la generalización a nuevos datos. La validación cruzada ayuda a evaluar el desempeño del modelo en diferentes subconjuntos de datos y seleccionar el mejor número de iteraciones para evitar el sobreajuste (Chen *et al.*, 2017).

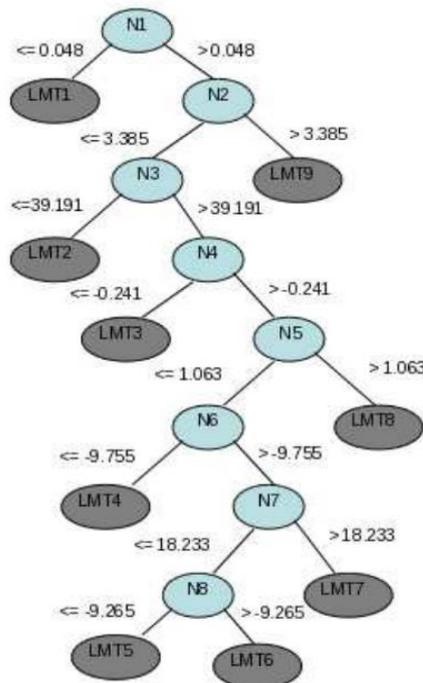


Figura 2. Árboles de decisión logística

Nota: Obtenido de Modelo basado en meta-clasificadores para diagnóstico en marcha patológica mediante análisis cinético por *Aguilera et al. (2010)*

4.5.2.3. Splines de regresión adaptativa multivariante (MARS)

El algoritmo MARS es una herramienta adaptable que automatiza la creación de modelos predictivos, permitiendo la selección de variables pertinentes, la transformación de las variables predictoras, el establecimiento de interacciones entre las variables predictoras, y el manejo de los valores faltantes, además de realizar un autotest para prevenir el sobreajuste. Además, puede identificar patrones y relaciones que resultan difíciles, o incluso imposibles, de descubrir mediante otros métodos (Vanegas y Vásquez, 2017).

El algoritmo MARS se utiliza para crear regresiones lineales en todo el rango de valores predictores (Figura 3). Este modelo funciona en dos fases distintas: en la primera fase, conocida como la fase de creación de funciones base, el rango de valores se divide en varios grupos y se modelan regresiones lineales con pendientes distintas para cada grupo. Esto permite capturar las relaciones no lineales entre las variables predictoras y la variable objetivo. En la segunda fase, llamada fase de retroceso o poda, se revisa cada una de las funciones base creadas en la fase anterior. Aquellas funciones que no aportan un rendimiento significativo al modelo son eliminadas, lo que ayuda a simplificar y mejorar la interpretación del modelo final (Dobilas, 2020).

El modelo puede escribirse como en la Ecuación 2, obtenido de Vanegas y Vásquez (2017).

Ecuación 2. Modelo de MARS

$$y_t = f(x_t) = \beta_0 + \sum_{i=1}^k \beta_i B(x_{it})$$

donde y_t es la variable respuesta en el instante t y β_i son los parámetros del modelo para las respectivas variables x_{it} , que van de $i = 1, \dots, k$. El valor β_0 representa al intercepto, las funciones bases $B(x_{it})$ son funciones que dependen de las respectivas variables x_{it} , en donde cada $B(x_{it})$ puede escribirse como $B(x_{it}) = \max(0, x_{it} - c)$ o $B(x_{it}) = \max(0, c - x_{it})$, c es un valor umbral y k representa el número de explicativas, que incluye interacciones de las variables predictoras.

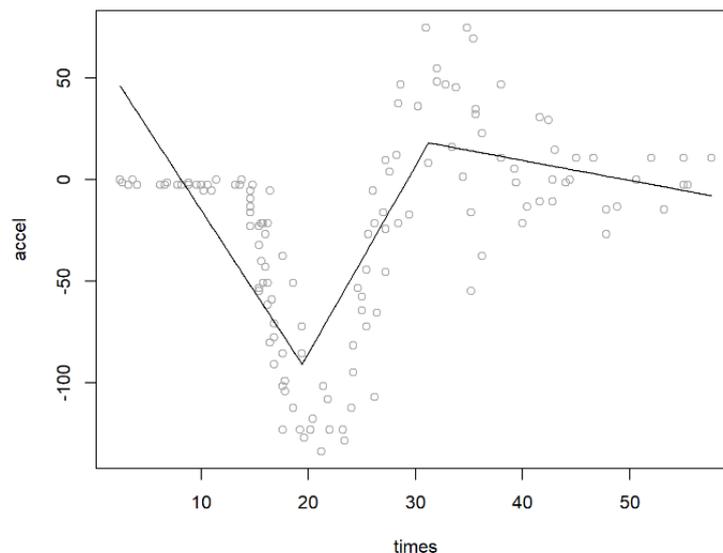


Figura 3. Splines de Regresión Adaptativa Multivariante

Nota: Obtenido de Data Science con R. Análisis de datos y algoritmos de predicción por Paredes (2020).

4.6. Métodos para la evaluación de modelos de probabilidad

Los resultados obtenidos mediante la implementación de técnicas de aprendizaje automático pueden ser evaluados utilizando medidas de rendimiento, como la curva Característica Operativa del Receptor (ROC), utilizada para modelos de clasificación, y el índice Kappa para poder evaluar la concordancia de las variables.

4.6.1. Curva Característica Operativa del Receptor (ROC)

Cerda y Cifuentes (2012) explican que la curva ROC (Receiver Operating Characteristic) es una herramienta utilizada para evaluar el rendimiento de un modelo en la clasificación de datos. Esta curva representa la relación entre la sensibilidad (verdaderos positivos) y la especificidad (verdaderos negativos).

La sensibilidad, también conocida como tasa de verdaderos positivos, es la proporción de datos calificados correctamente como “verdaderos positivos” (VP) sobre el número total de datos clasificados como positivos: verdaderos positivos (VP) y falsos positivos (FP) (Hern *et al.* (2019), y se la calcula por medio de la ecuación 3.

Ecuación 3. Sensibilidad de los datos

$$\text{Sensibilidad} = \frac{VP}{VP+FP}$$

Así mismo, Hern *et al.* (2019) menciona que la especificidad es la proporción de verdaderos negativos (VN) sobre el número total de datos clasificados como negativos, incluyendo los falsos positivos (FP) y los verdaderos negativos (VN), volviéndola una medida que distingue la capacidad de un modelo de clasificación para identificar correctamente los casos negativos, y se la calcula por medio de la ecuación 4.

Ecuación 4. Especificidad de los datos

$$\text{Especificidad} = \frac{VN}{FP + VN}$$

El área bajo la curva (AUC) de ROC, permite determinar el grado de fiabilidad de un modelo, indicando la probabilidad de aciertos previstos para el modelo analizado (Hern *et al.*, 2019). Por lo general, los valores superiores a 0.70 son considerados como buenos predictores, ya que el modelo es capaz de diferenciar entre clases positivas y negativas. Los valores para la interpretación del AUC se muestran en la Tabla 2:

Tabla 2. Valores de precisión del modelo de acuerdo con el AUC

Valor de AUC	Nivel de precisión
0.50 – 0.60	Insuficiente
0.60 – 0.70	Pobre
0.70 – 0.80	Regular
0.90 – 0.90	Bueno
0.90 – 1.00	Excelente

Nota: Adaptado de una propuesta metodológica para la selección de umbrales de clasificación en procesos de seroconversión y transmisión de anticuerpos a recién nacidos en Bogotá-Colombia, por Hern, E. R., Lemus, D. F., y Pineda, W. D. (2019).

4.6.2. Coeficiente de Kappa de Cohen

El coeficiente Kappa o índice de kappa es un instrumento estadístico que estima la concordancia entre dos observadores o herramientas de medición para un mismo fenómeno, con el propósito de ajustar el efecto del azar en la proporción de la concordancia esperada (Galparsoro y Pita, 2010; Landa *et al.*, 2014), y se la calcula por medio de la ecuación 5.

Ecuación 5. Índice de Kappa

$$Kappa = \frac{P_0 - P_e}{1 - P_e}$$

Donde:

P_0 = proporción de concordancia observada

P_e = proporción de concordancia esperada por el azar

Los valores del índice de kappa pueden ir de 0 y 1, para poder evaluar la concordancia de las variables. Los valores para la interpretación del índice de Kappa se muestran en la Tabla 3.

Tabla 3. Valoración del índice de Kappa

Valor de k	Fuerza de la concordancia
< 0.20	Pobre
0.21 – 0.40	Débil
0.41 – 0.60	Moderada
0.61 – 0,80	Buena
0,80 – 1,00	Muy buena

Nota: Obtenido de Medidas de concordancia: el índice Kappa de Galparsoro, L. de U., y Pita, F. S. (2010).

5. Metodología

5.1. Zona de estudio

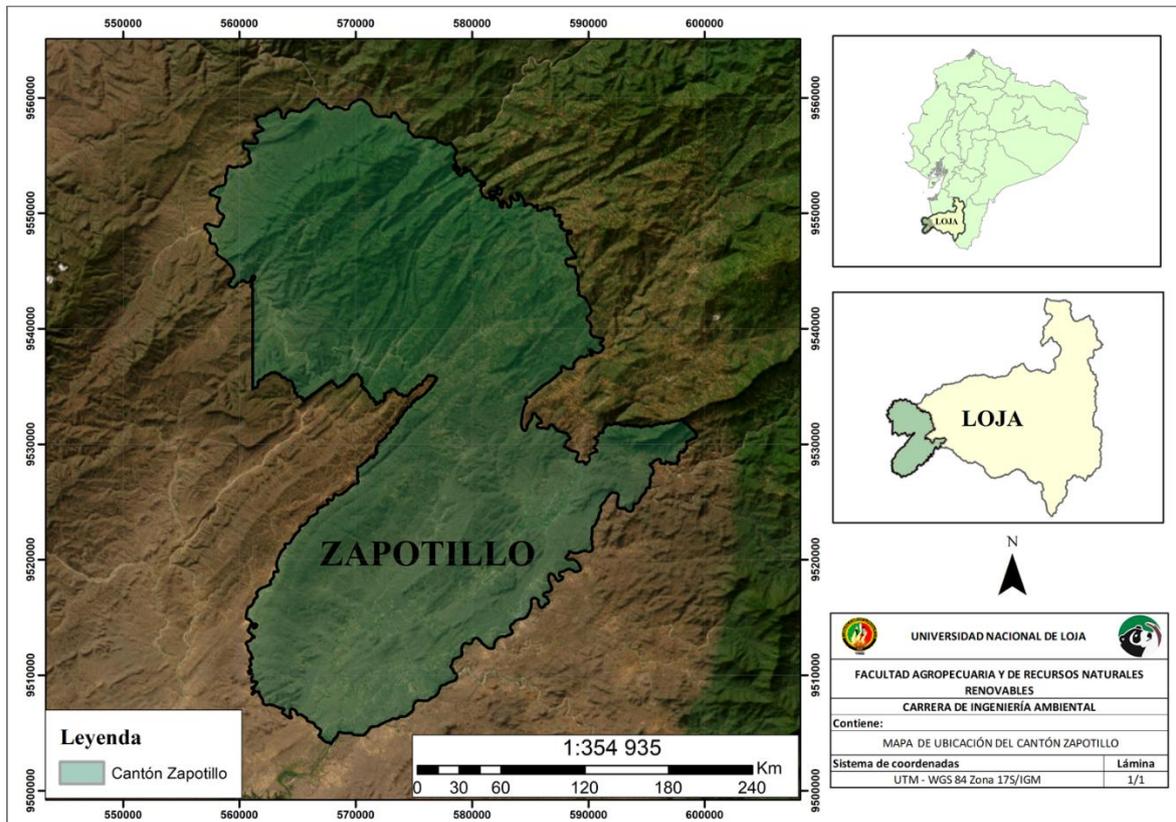


Figura 4. Mapa base del cantón Zapotillo

El cantón Zapotillo se encuentra ubicado al sur occidental de la provincia de Loja, siendo sus coordenadas geográficas son $04^{\circ} 15'$ y $04^{\circ} 29'$ de latitud sur, y $80^{\circ} 22' 15''$ y $80^{\circ} 23' 3''$ de longitud oeste. Este cantón limita al norte con la República del Perú y el cantón Puyango, al sur con la República del Perú, al este con los cantones Puyango, Celica, Macará, Pindal y al oeste con la república del Perú (PDOT, 2015).

Zapotillo cuenta con una superficie total de 121.157,80 ha, las cuales son usadas para conservación, agrícola, pecuario, forestal, erial, antrópico y cuerpos de agua. Además, esta zona se destaca por la presencia de diversos ecosistemas, siendo el bosque seco su prioridad en términos de conservación. El ecosistema más representativo es el bosque deciduo de tierras bajas del Jama-Zapotillo, que abarca el 37,57% del territorio con una superficie de 45.523,12 ha. Le sigue el bosque deciduo Pie-montano del Catamayo-Alamor, que representa el 20,61% con una superficie de 24.965,86 ha. También se encuentra el bosque semideciduo Pie-montano

del Catamayo-Alamor, que ocupa el 13,70% con una superficie de 16.603,53 ha. Otro ecosistema importante es el bosque bajo y arbustal deciduo de tierras bajas del Jama y Zapotillo, que abarca el 11,04% con una superficie de 13.378,60 ha. Por último, es relevante mencionar que el porcentaje de intervención humana en el territorio es del 16,50%, cubriendo una superficie de 19.990,37 ha (PDOT, 2019).

Según la información otorgada por PDOT (2019), el cantón experimenta periodos secos que se extienden desde julio hasta diciembre, con una temperatura promedio anual de 24°C. Por otro lado, los meses con las mayores precipitaciones del año son enero, febrero y marzo, siendo considerados como los más lluviosos en la región. Durante este periodo, el cantón de Zapotillo registra valores representativos de precipitación que oscilan entre 500 y 800 mm. Además, la altitud que presente el cantón Zapotillo oscila entre los 193 y 1.436 m.s.n.m. (INEC, 2010).

5.2. Softwares y plataformas

Los softwares y las plataformas son mostradas en la Figura 5.

- De escritorio
 1. **QGIS.** - es un SIG de código abierto y multiplataforma, ampliamente utilizado para visualizar, analizar y gestionar datos geoespaciales, incluyendo formatos vectoriales y raster.
 2. **SAGA GIS.** - es otro SIG de código abierto, especializado en análisis avanzados de terreno y procesamiento de imágenes ráster, ideal para estudios de geomorfología y análisis hidrológicos.
 3. **R.** - Es una herramienta que sirve como lenguaje de programación utilizado para estadísticas y análisis de datos, con amplias capacidades de manipulación y visualización de datos geoespaciales.
- Online
 4. **Google Earth Engine.** - es una plataforma en la nube que permite el análisis de imágenes satelitales a gran escala y ofrece herramientas avanzadas para el monitoreo y análisis de cambios ambientales.
 5. **Geoportales y Visores Geográficos.** - cómo ArcGIS Online o OpenStreetMap, brindan acceso a datos geoespaciales y mapas interactivos, facilitando la visualización y análisis de información geográfica sin necesidad de software especializado.

Softwares y Plataformas

DE ESCRITORIO (proceso de instalación)

ONLINE (proceso de registro)



QGIS en 3 versiones (with GRASS): 3.16.1. / 2.18.28 / 2.14
<https://qgis.org/es/site/forusers/download.html>



SAGA GIS 2.3.2
<https://sourceforge.net/projects/saga-gis/>



Rstudio + Weka
<https://www.rstudio.com/products/rstudio/download/>



Google Earth Engine

<https://earthengine.google.com>

Tutorial de registro: <https://youtu.be/vZTMOXLRBmo>

<https://youtube.com/playlist?list=PLnOWTRns6TAtUG9Xr9ER83JQf9m76R9Ku>



Geoportales y Visores Geográficos

<https://sni.gob.ec/geoportales-y-visores>

Ejemplo de descargas: <https://youtu.be/a07PxKcCxbA>

Figura 5. Softwares y Plataformas para el modelo.

5.3. Establecer una línea base de las variables predictivas para identificar la ocurrencia de incendios forestales

La generación de la línea base implica la recopilación de información sobre variables relacionadas con la ocurrencia de incendios forestales. En este estudio, se utilizó la metodología de Reyes y Córdova (2021), en la que emplean variables topográficas, variables de interacción antrópica e índices espectrales obtenidos del satélite "Sentinel 2A", en total se utilizó 23 variables independientes, las cuales se detallan en las Tablas 4, 5 y 6.

Variables topográficas

Tabla 4. Variables topográficas

Variable	Descripción	Proceso de obtención	Metadatos de la capa
1 Elevación	Altura sobre el nivel del mar.	Se importa un archivo “.kml” con los límites del cantón de interés y se realiza la búsqueda de Modelos de Elevación Digital (DEM) disponibles en la web de ALOS PALSAR para la zona. Una vez descargados los archivos en <i>high resolution</i> se utiliza la herramienta de QGIS para combinar las teselas (Ráster, Miscelánea, Combinar), posteriormente se recorta a los límites de nuestra zona de interés (Ráster, Extracción, Cortar ráster por capa de máscara)	SRC: EPSG:32717 Resolución espacial: 10 x 10 metros Rango de valores: 0 – 3668 Unidades: m.s.n.m.

2	Pendiente	Grado de inclinación del terreno, puede medirse en grados o porcentaje.	Al DEM obtenido en la web de ALOS PALSAR, se le aplica un filtro de suavizado en QGIS herramienta SAGA (Simple Filter, Smooth 3). Después de aplicar el filtro de suavizado, se utiliza la herramienta GRASS (r.slope.aspect, percent)	SRC: EPSG:32717 Resolución espacial: 10 x 10 metros Rango de valores: 0 – 375 Unidades: porcentaje (%)
3	Orientación	Dirección de pendiente descendente que presenta cada celda, tomando como referencia los puntos cardinales (Norte, Sur, Este, Oeste)	Al DEM obtenido en la web de ALOS PALSAR, se le aplica un filtro de suavizado en QGIS herramienta SAGA (Simple Filter, Smooth 3). Después de aplicar el filtro de suavizado, se utiliza la herramienta GRASS (r.slope.aspect)	SRC: EPSG:32717 Resolución espacial: 10 x 10 metros Rango de valores: 0 – 360 Unidades: grados (°)

Nota: Adaptado de Informe de variables que se utilizan para la generación de modelos de susceptibilidad de incendios con la identificación de sus fuentes para descarga por *Loján (2020)*.

Variables de interacción antrópica

Tabla 5. Variables de interacción antrópica

Variable	Descripción	Proceso de obtención	Metadatos de la capa	
4	Accesibilidad al centro económico	Facilidad de movimiento desde cualquier punto del cantón hasta el centro económico.	Utilizando la herramienta GRASS r.cost.full, se configura como punto(s) la capa de centro económico descargada en el portal del IGM.	SRC: EPSG:32717 Resolución espacial: 10 x 10 metros Rango de valores: 0 – 140 Unidades: minutos
5	Accesibilidad a áreas protegidas	Facilidad de movimiento desde cualquier punto del cantón hasta áreas protegidas.	Utilizando la herramienta GRASS r.cost.full, se configura como punto(s) objetivo(s) la capa de áreas protegidas descargada en el Mapa Interactivo Ambiental. Las categorías consideradas de áreas protegidas son: - Sistema Nacional de Áreas Protegidas Bosque y Vegetación Protectora - Área Bajo Conservación - Patrimonio Forestal del Estado - Áreas de Protección Hídrica	SRC: EPSG:32717 Resolución espacial: 10 x 10 metros Rango de valores: 0 – 140 Unidades: minutos

6	Accesibilidad a zonas pobladas	Facilidad de movimiento desde cualquier punto del cantón hasta poblados.	Utilizando la herramienta GRASS r.cost.full, se configura como punto(s) objetivo(s) la capa de poblados descargada en el portal del IGM.	SRC: EPSG:32717 Resolución espacial: 10 x 10 metros Rango de valores: 0 – 130 Unidades: minutos
7	Accesibilidad a vías	Facilidad de movimiento desde cualquier punto del cantón hasta vías.	Utilizando la herramienta GRASS r.cost.full, se configura como punto(s) objetivo(s) la capa de vías digitalizada de manera manual en Google Maps Satellite.	SRC: EPSG:32717 Resolución espacial: 10 x 10 metros Rango de valores: 0 – 145 Unidades: minutos
8	Accesibilidad a zonas antrópicas	Facilidad de movimiento desde cualquier punto del cantón hasta zonas pobladas o zonas agropecuarias.	Utilizando la herramienta GRASS r.cost.full, se configura como punto(s) objetivo(s) la capa de cobertura vegetal descargada en el Mapa Interactivo Ambiental.	SRC: EPSG:32717 Resolución espacial: 10 x 10 metros Rango de valores: 0 – 120 Unidades: minutos
9	Distancia desde el centro económico	Distancia desde el centro económico.	Utilizando la herramienta GRASS r.grow.distance, se configura como punto(s) objetivo(s) la capa de centro económico descargada en el portal del IGM.	SRC: EPSG:32717 Resolución espacial: 10 x 10 metros Rango de valores: 0 – 20 597 Unidades: metros
10	Distancia a ríos	Distancia desde ríos.	Utilizando la herramienta GRASS r.grow.distance, se configura como punto(s) objetivo(s) la capa de ríos generada a partir del DEM.	SRC: EPSG:32717 Resolución espacial: 10 x 10 metros Rango de valores: 0 – 1141 Unidades: metros
11	Distancia a vías	Distancia desde vías	Utilizando la herramienta GRASS r.grow.distance, se configura como punto(s) objetivo(s) la capa de vías digitalizada de manera manual en Google Maps Satellite.	SRC: EPSG:32717 Resolución espacial: 10 x 10 metros Rango de valores: 0 – 4157 Unidades: metros

12	Distancia desde áreas protegidas	Distancia desde áreas protegidas	Utilizando la herramienta GRASS r.grow.distance, se configura como punto(s) objetivo(s) la capa áreas protegidas descargada en el Mapa Interactivo Ambiental. Las categorías consideradas de áreas protegidas son: Sistema Nacional de Áreas Protegidas Bosque y Vegetación Protectora Área Bajo Conservación Patrimonio Forestal del Estado Áreas de Protección Hídrica	SRC: EPSG:32717 Resolución espacial: 10 x 10 metros Rango de valores: 0 – 12322 Unidades: metros
13	Distancia a zonas antrópicas	Distancia desde zonas pobladas o zonas agropecuarias.	Utilizando la herramienta GRASS r.grow.distance, se configura como punto(s) objetivo(s) la capa zonas pobladas o zonas agropecuarias.	SRC: EPSG:32717 Resolución espacial: 10 x 10 metros Rango de valores: 0 – 3820 Unidades: metros

Nota: Adaptado de Informe de variables que se utilizan para la generación de modelos de susceptibilidad de incendios con la identificación de sus fuentes para descarga por *Loján (2020)*.

Variables del estado de vegetación (Índices espectrales)

Tabla 6. Variables del estado de vegetación

Variable	Descripción	Proceso de Obtención	Metadatos de la capa
14	<p>ARVI</p> <p><i>Atmospherically Resistant Vegetation Index</i></p> <p>NDVI corregido para efectos de dispersión atmosférica</p> $ARVI = \frac{NIR - (2 * Red) + Blue}{NIR + (2 * Red) + Blue}$ <p>NIR (Near-Infrared Region) → B8 Red (Rojo, espectro visible) → B11 Blue (Azul, espectro visible) → B2</p>	Los índices espectrales o índices de vegetación son descargados mediante un script de Google Earth Engine, para ello se usa imágenes satelitales de Sentinel 2A.	SRC: EPSG:32717 Resolución espacial: 10 x 10 metros Rango de valores: -1 a 1
15	<p>BSI</p> <p><i>Bare Soil Index</i></p> <p>Identificar zonas sin cobertura vegetal</p> $BSI = \frac{(Red + SWIR) - (NIR + Blue)}{(Red + SWIR) + (NIR + Blue)}$ <p>Red (Rojo, espectro visible) → B11 SWIR (Shortwave Infrared) → B4 NIR (Near-Infrared Region) → B8 Blue (Azul, espectro visible) → B2</p>		SRC: EPSG:32717 Resolución espacial: 10 x 10 metros Rango de valores: -1 a 1

16	GNDVI	<p>Green Normalized Difference Vegetation Index</p> <p>Versión del NDVI más sensible a la variación del contenido de clorofila y nitrógeno en el cultivo.</p> $GNDVI = \frac{NIR - SWIR}{NIR + SWIR}$ <p>SWIR (Shortwave Infrared) → B3 NIR (Near-Infrared Region) → B8</p>	<p>SRC: EPSG:32717</p> <p>Resolución espacial: 10 x 10 metros</p> <p>Rango de valores: -1 a 1</p>
17	MSI	<p>Moisture Stress Index</p> <p>Determinar el estrés hídrico de la vegetación</p> $MSI = \frac{MidIR}{NIR}$ <p>MidIR (Shortwave Infrared) → B11 NIR (Near-Infrared Region) → B8</p>	<p>SRC: EPSG:32717</p> <p>Resolución espacial: 10 x 10 metros</p> <p>Rango de valores: 0 a 3+</p>
18	NBR	<p>Normalized Burn Ratio</p> <p>Estimar la cantidad y desarrollo de áreas incendiadas</p> $NBR = \frac{NIR - SWIR}{NIR + SWIR}$ <p>SWIR (Shortwave Infrared) → B12 NIR (Near-Infrared Region) → B8</p>	<p>SRC: EPSG:32717</p> <p>Resolución espacial: 10 x 10 metros</p> <p>Rango de valores: -1 a 1</p>
19	NDMI	<p>Normalized Difference Moisture Index</p> <p>Determinar el contenido de agua de la vegetación</p> $NDMI = \frac{NIR - SWIR}{NIR + SWIR}$ <p>SWIR (Shortwave Infrared) → B11 NIR (Near-Infrared Region) → B8</p>	<p>SRC: EPSG:32717</p> <p>Resolución espacial: 10 x 10 metros</p> <p>Rango de valores: -1 a 1</p>
20	NDWI	<p>Normalized Difference Water Index</p> <p>Identificar masas de agua y zonas de elevada humedad.</p> $NDWI = \frac{NIR - Green}{NIR + Green}$ <p>NIR (Near-Infrared Region) → B8 Green (Verde, espectro visible) → B3</p>	<p>SRC: EPSG:32717</p> <p>Resolución espacial: 10 x 10 metros</p> <p>Rango de valores: -1 a 1</p>
21	NDVI	<p>Normalized Difference Vegetation Index</p> <p>Facilidad con la que el fuego se propaga en la vegetación.</p> $NBR = \frac{NIR - Red}{NIR + Red}$ <p>NIR (Near-Infrared Region) → B8 Red (Rojo, espectro visible) → B4</p>	<p>SRC: EPSG:32717</p> <p>Resolución espacial: 10 x 10 metros</p> <p>Rango de valores: -1 a 1</p>

22	SAVI	<p>Soil Adjusted Vegetation Index Analizar la vegetación en etapas de crecimiento inicial o dispersa.</p> $NBR = \frac{NIR - R}{NIR + R + L} * (1 + L)$ <p>NIR (Near-Infrared Region) → B8 Red (Rojo, espectro visible) → B4 Factor de corrección de la luminosidad del suelo → L</p>	<p>SRC: EPSG:32717</p> <p>Resolución espacial: 10 x 10 metros</p> <p>Rango de valores: -1 a 1</p>
23	Moisture Index	<p>Normalized Difference Moisture Index Estimar el contenido de humedad de la vegetación.</p> $Moisture = \frac{NIR - SWIR}{NIR + SWIR}$ <p>SWIR (Shortwave Infrared) → B11 NIR (Near-Infrared Region) → B8A</p>	<p>SRC: EPSG:32717</p> <p>Resolución espacial: 10 x 10 metros</p> <p>Rango de valores: -1 a 1</p>

Nota: Adaptado de Informe de variables que se utilizan para la generación de modelos de susceptibilidad de incendios con la identificación de sus fuentes para descarga por *Loján (2020)*.

5.4. Construir un modelo de simulación espacial de zonas susceptibles a incendios forestales

El proceso para crear el modelo de simulación espacial de zonas susceptibles a incendios forestales en el cantón Zapotillo constó de tres etapas, se visualiza en la Figura 6: (I) preparación de la base de datos en la que se realizó una depuración y análisis de datos ambientales y antrópicos que conformaron las variables de importancia del modelo; (II) generación y aplicación del modelo en la que se emplearon tres algoritmos de aprendizaje automático supervisado: árboles de decisión (LMT)); regresión logística (LR), splines de regresión adaptativa multivariante (MARS); (III) análisis del rendimiento del modelo en la que se evalúa el modelo utilizando dos medidas de rendimiento: el área bajo la curva (AUC) de ROC, que indica la capacidad predictiva del modelo; y el índice kappa de Cohen que evalúa la sensibilidad y reproducibilidad.

A continuación, se explican las actividades realizadas en cada fase de este método.

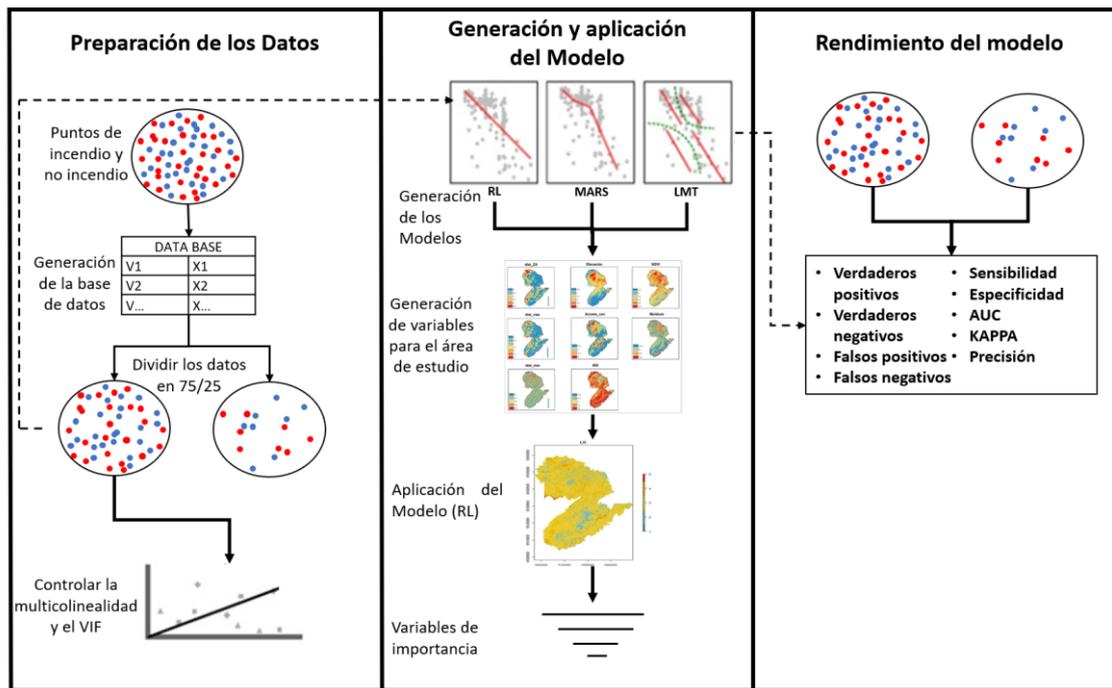


Figura 6. Fases para la generación del mapa de probabilidad de ocurrencia de incendios forestales en el cantón Zapotillo

Nota: Adaptado de “Assessment oh Three Machine Learning Techniques with Open-Acess Geographic Data for Forest Fire Susceptibility Monitoring-Evidence form Southern Ecuador” por Reyes y Loján (2022).

5.4.1. Etapa 1: Preparación de la base de datos

5.4.1.1. Obtención de la información de puntos de muestreo

Los datos se recolectaron de una fuente confiable, siendo la plataforma *Fire Information for Resource Management System* de la NASA (FIRMS), debido a que esta plataforma brinda datos de focos de calor obtenidos por el satélite SUOMI NPP, la misma que cuenta con una herramienta llamada VIIRS, además también se utilizó información disponible para la región del satélite “Sentinel 2A”. Se recopilaron 657 focos de calor (incendios) en un periodo de diciembre 2018 hasta diciembre del 2021 en un área de influencia dentro del cantón. Este período de datos fue seleccionado debido a que las imágenes satelitales de Sentinel 2A, están disponibles para la zona ecuatorial desde el año 2018. Asimismo, esta temporalidad se alineó con la base de datos geográfica emitida por la Secretaría Nacional de Gestión de Riesgos, que abarca hasta 2022. Además, el último incendio registrado que tuvo gran impacto en el cantón Zapotillo, ocurrió en el año 2021, afectando un total de 200 ha.

5.4.1.2. Preparación de los datos

5.4.1.2.1. Escoger los focos de calor considerados como incendios

Para iniciar esta etapa, se inició con la identificación de los incendios forestales presentes en la base de datos de la herramienta “VIIRS” con la ayuda del software GIS v.3.22., además, se utilizó imágenes de severidad y color natural que otorga el satélite “Sentinel 2A” y la plataforma “Google Earth Engine”, esto con base en la fecha en la que salió la imagen satelital, la cual se va actualizando cada 5 días. En la comparación se observa si el punto de calor está ubicado dentro o cerca de una zona roja en la imagen de severidad, demostrando que existió un incendio forestal de acuerdo a la fecha de registro del foco, si no es el caso, se depura la información. Esta depuración de datos es necesaria debido a que no todos los focos de calor registrados por esta herramienta son incendios forestales, por lo que debemos ir eliminando estos datos fuera de nuestra necesidad.

El tipo de muestreo que se utilizó es un muestreo no probabilístico, específicamente un muestreo por juicio con múltiples criterios, en el que el investigador utiliza su criterio o experiencia para seleccionar intencionalmente los elementos de la muestra, basándose en su conocimiento del fenómeno en estudio. En este caso particular, se realizó la selección de datos de los focos de calor como incendios forestales que exhiban un proceso de ignición identificados visualmente y que contengan información de Potencia Radiativa del Fuego (FRP), el mismo nos proporcionó información para evaluar la intensidad y el tamaño de un incendio por medio de la energía radiante emitida por este evento en un periodo de tiempo dentro de un pixel (NASA EARTHDATA, 2020). Esta selección específica se realizó con el propósito de enfocarse en aquellos focos de calor considerados como más representativos para el análisis de la susceptibilidad de incendios en el cantón Zapotillo. Los criterios que se tomaron en cuenta son mencionados en la tabla 7, estos en el orden de importancia que debe tener cada uno.

Tabla 7. Criterios para determinar incendios forestales

Criterios para determinar incendios forestales	
“ACQ_TIME”	¿Cuál sucedió primero?
“CONFIDENCE”	n, l, h
“FRP”	FRP > 9
Tonalidad	Tonalidades oscuras e intensas (marrones, rojiza, naranjas, negras)

En el caso de que un foco de calor es considerado un incendio, se crea un punto cercano no más de 2 km de distancia en una zona no afectada, con la finalidad de que el algoritmo funcione correctamente, identificando zonas que fueron afectadas por el evento y zonas que no fueron afectadas (50% de puntos de ocurrencia y 50% de no ocurrencia). Además, se excluyeron los focos de calor localizados en áreas con presencia de nubes antes de la fecha del incendio, ya que con la herramienta Point Sampling Tool no se podrá extraer ningún tipo de información si existe nubosidad. Una vez que se realizó la depuración, obtuvimos 82 focos de calor considerados como incendios y se añadió el mismo número de puntos que identificaron áreas donde no hubo incendios (Figura 7). Estas áreas fueron identificadas con la ayuda de imágenes Sentinel 2A posteriores a la fecha de registro del foco de calor.

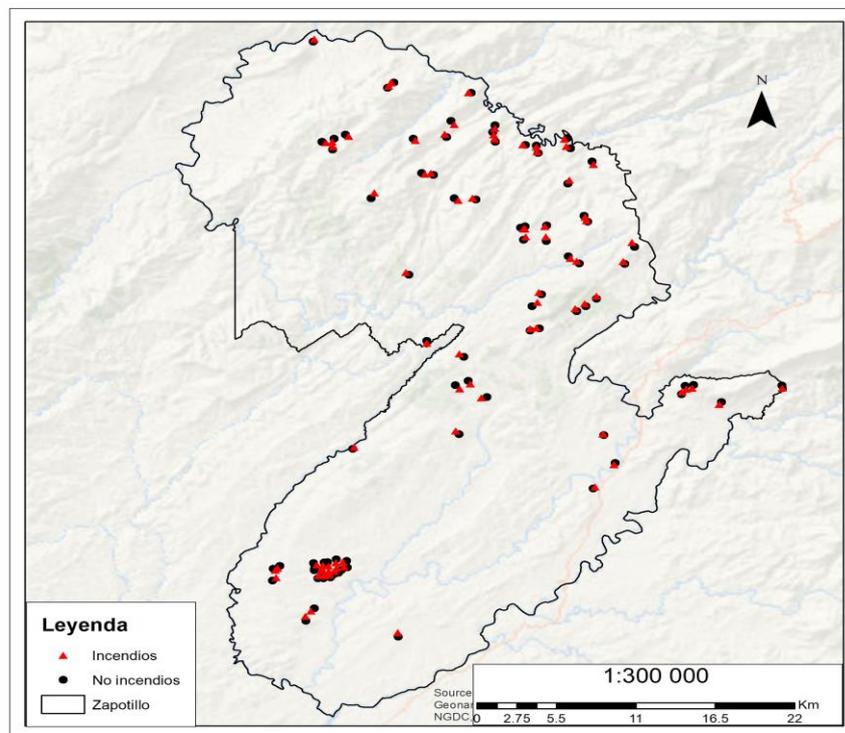


Figura 7. Distribución geográfica de los puntos de muestreo (incendio y no incendio)

5.4.1.2.2. Extracción de información de las variables independientes

En este punto se abordará el proceso que se ejecutó para obtener la información de las variables que se consideran para el modelo (Figura 8), esto con la finalidad de construir una base de datos, además de la información de las variables que se extrajo de estos puntos de muestreo (focos de calor) a partir de la herramienta Point Sampling Tool.



Figura 8. Proceso para la obtención de información de las variables.

5.4.1.2.2.1. Obtención de la capa de accesibilidad

En base a la metodología propuesta por Reyes (2017) en el “Manual de Prácticas de Ordenación Territorial”, el cálculo de la superficie de fricción y accesibilidad se observa en la Figura 9.

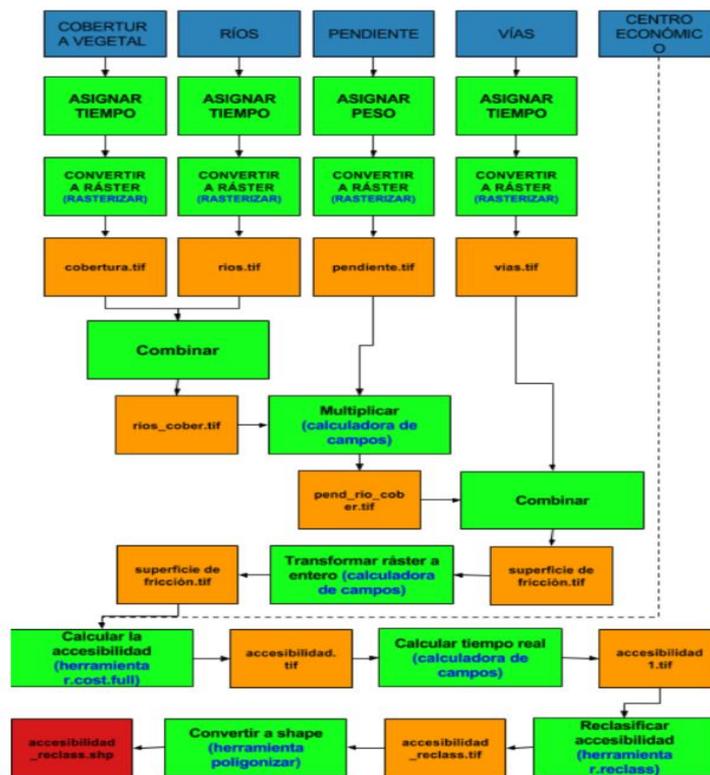


Figura 9. Diagrama de flujo del proceso de cálculo de accesibilidad

Nota: Obtenido del Manual de Prácticas de ordenación Territorial por Reyes (2017).

Las capas necesarias para calcular la accesibilidad al centro económico, se obtuvieron del mapa interactivo del MAATE (para la cobertura vegetal). Las capas obtenidas fueron: la cobertura vegetal, áreas protegidas, y del Geoportal del Instituto Geográfico Militar (IGM): la ubicación de los centros económicos y las zonas pobladas. Pero en el caso de las vías, fue

necesario digitalizarlas de manera manual en base a lo que se muestra en Google Maps satélite, además, se clasificó los elementos de la capa de vías en tres categorías: vía primaria (Pavimentado), vía secundaria (Lastrado) y vía terciaria (Tierra) mostrado en la Figura 10.

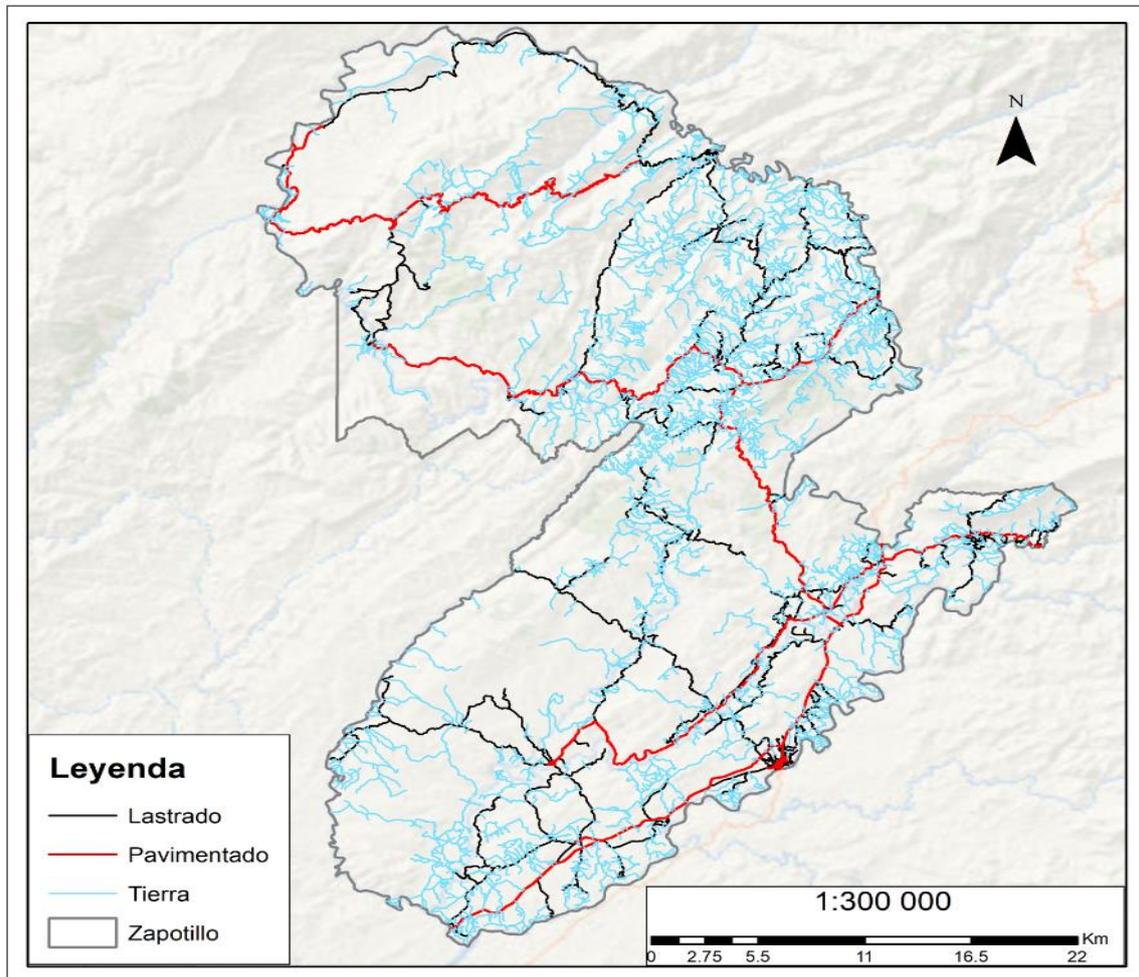


Figura 10. Digitalización de vías

Para generar información sobre la accesibilidad es necesario calcular una superficie de fricción, el cual es un procedimiento para conocer las interacciones entre un objeto y la superficie, de una manera geográfica en el que asigna un coste o un valor de movilización el cual es medido en minutos. Según Reyes (2017), primero se debe definir la velocidad requerida para atravesar cada punto en el territorio, considerando que la cobertura vegetal, la pendiente, las vías y los ríos limitan el movimiento. En la Tabla 8 se detalla la asignación del tiempo para cada capa (derivado de la velocidad necesaria para desplazarse).

Tabla 8. *Coste de movilidad según la clase de superficie*

Capa	Clase	Coste
Cobertura	Área Poblada, Área sin cobertura vegetal, Infraestructura	1 200
	Cultivo anual, Cultivo semipermanente, Mosaico agropecuario, Pastizal, Plantación forestal, Vegetación herbácea	1 500
	Bosque nativo, Natural, Vegetación arbustiva	2 400
Ríos	Orden 1, 2 y 3	1 200
Vías	Vía primaria	80
	Vía secundaria	120
	Vía terciaria	200

Para la obtención de las variables de distancia (Figura 8) para las capas de ríos, vías, zonas antrópicas, áreas protegidas, centro económico, se utilizó la herramienta “r.grow.distance” de QGIS 3.20.2.

En el caso de los índices espectrales, se obtuvieron a partir de las imágenes satelitales Sentinel 2A a través de un código en Google Earth Engine (GEE) creado Reyes y Loján (2022), en su trabajo sobre susceptibilidad a incendios forestales en el cantón Loja, el cual fue adaptado a nuestra zona de estudio, considerando los siguientes aspectos: la delimitación del área de interés, modelo de probabilidad, variables de importancia, las fechas de interés y el cálculo de los índices espectrales.

Finalmente, se generó la base de datos, con la extracción de la información que obtuvimos anteriormente, usando el software libre QGIS v.3.22. con su complemento “Point Sampling Tool”, el cual nos permite seleccionar puntos de interés en un mapa y extraer información asociada a esos puntos de muestreo, como valores de atributos o características específicas, que en este caso específico son los focos de calor que determinamos como incendios forestales. La información de las variables que esté alrededor de los focos de calor, se la va a extraer y crear la base de datos en formato vectorial (.shp), para después ser exportado en formato delimitado por comas (.csv). Para ilustrar mejor este proceso, se lo puede entender en la Figura 11.

De las 23 variables contenidas en la base de datos, 13 son de carácter permanente (variables de interacción antrópica y variables topográficas) y 10 son actualizables, debido a

que reflejan cambios en la cobertura vegetal a lo largo del tiempo (variables del estado de vegetación), por lo que su tratamiento de datos es diferente.

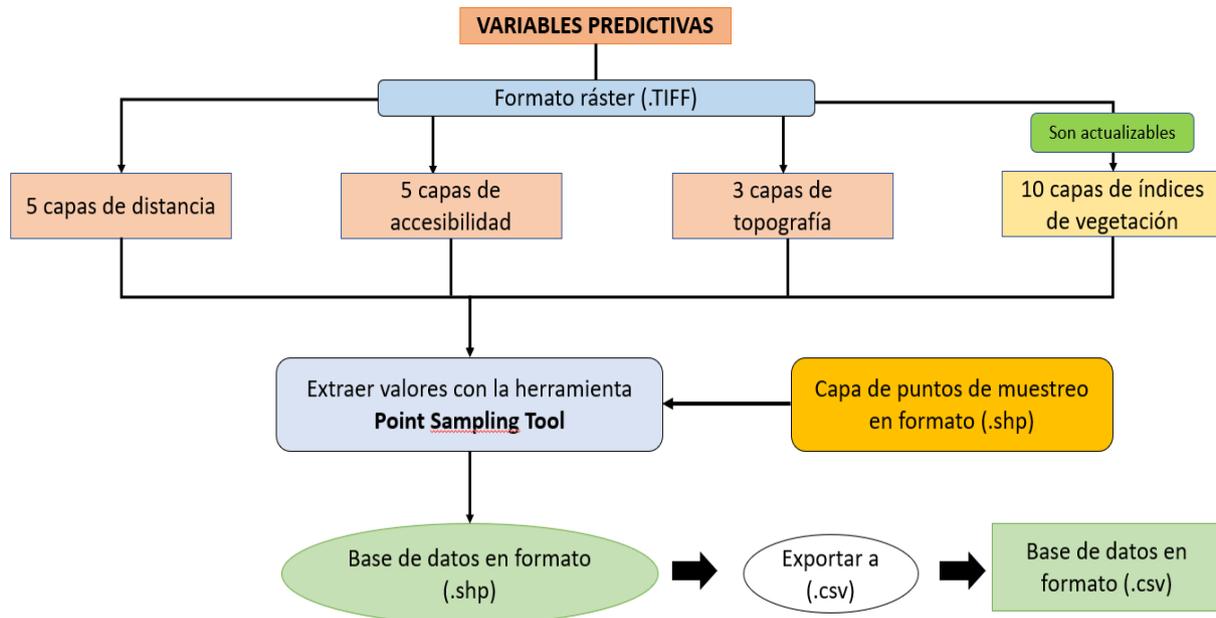


Figura 11. Proceso para la generación de la base de datos.

Nota: Adaptado de Reyes y Córdova, (2021) además las fuentes de descarga para los archivos ráster y vectorial se muestran en las Tabla 4, 5, y 6. El proceso detallado para la obtención de capas de accesibilidad se muestra en la Figura 9 y la obtención de los índices de vegetación por fechas se obtuvo utilizando la herramienta GEE.

5.4.1.2.3. Variables que no presentan multicolinealidad

Las variables empleadas en mapas de probabilidad de incendios forestales cambian dependiendo de los objetivos del investigador y los métodos utilizados en cada investigación. Sin embargo, las variables causantes de la ignición están relacionados con agentes naturales (condiciones topográficas, meteorológicas y climáticas), y agentes antrópicos de la zona analizada (Llorens *et al.*, 2021; Paucar *et al.*, 2021).

En la Figura 12 se muestran los elementos ambientales y antrópicos que se consideraron para la obtención de variables seleccionadas de acuerdo a una revisión bibliográfica, disponibilidad de datos, un test de multicolinealidad en el que se aplica un modelo lineal en R, debido a que la colinealidad entre variables reduce la capacidad estadística del modelo, lo que conlleva a problemas de ajuste excesivo y aumenta la incertidumbre en los resultados (Rojas *et al.*, 2023). Además, se aplicó el Factor de Inflación de Varianza (VIF) para evitar que las variables independientes presenten una correlación entre ellas (Reyes y Loján, 2022).

Ecuación 6. Factor de Inflación de varianza

$$VIF = \frac{1}{1-R^2_d}$$

Donde:

R^2 : es el factor determinante para la regresión de variables explicativas

d : Se refiere a todas las demás variables explicativas

Cuanto más alto sea el valor del VIF, mayor probabilidad existe de que haya multicolinealidad, por lo que el valor del VIF aumenta con el incremento de la multicolinealidad. En la prueba VIF, los valores superiores a 10 se consideran críticos (Reyes y Loján, 2022).

Además, en la Figura 13, se exhiben las fuentes de adquisición las variables generadas tras el procesamiento geoespacial en QGIS (para las variables de elevación, distancia y accesibilidad) y mediante el uso de la herramienta Google Earth Engine (GEE) para el cálculo de los índices espectrales. El Modelo Digital de Elevación (DEM) necesario para este propósito fue descargado del satélite Alos Palsar a través de la plataforma Earth Data de la NASA.

Los datos de estas variables (elevación, distancia, accesibilidad e índices espectrales) se obtuvieron para cada punto de muestreo (foco de calor y puntos de control), teniendo en cuenta las imágenes del satélite Sentinel 2A anteriores a la ocurrencia del incendio.

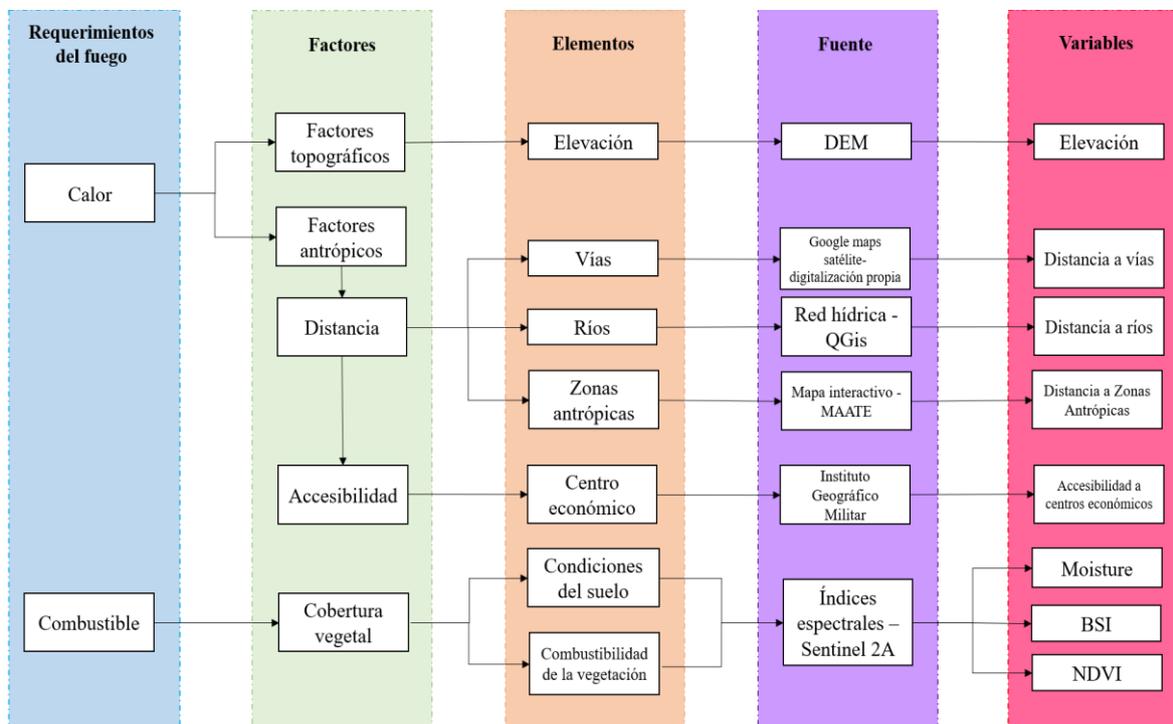


Figura 12. Identificación de las variables que no presentan multicolinealidad

5.4.2. Etapa 2: generación y aplicación del modelo

En base a la metodología de Reyes (2017), los modelos se generaron utilizando tres algoritmos de aprendizaje automático supervisado que son usados generalmente en minería de datos y predicciones: regresión logística (RL), splines de regresión adaptativa multivariante (MARS) y árboles de decisión logística (LMT).

El procedimiento se llevó a cabo en el software RStudio, con la asistencia del paquete "caret", donde se cargó la base de datos obtenida y se dividió en dos partes, asignando el 75 % para el entrenamiento y el 25 % para la validación del modelo, esto con la intención de evitar el sobreajuste de los datos en los modelos y surja un error.

Se emplearon los paquetes "glm" para obtener resultados del modelo de regresión logística (RL), "weka" para el modelo de árboles de decisión logística (LMT) y "earth" para el modelo de splines de regresión adaptativa multivariante (MARS) en la aplicación de RStudio. Con el 75% de los datos se entrenó el modelo, y con el 25% restante se lo valida. Para los 3 modelos el número de interacciones se determinó mediante validación cruzada con el fin de evaluar y seleccionar el modelo de aprendizaje automático que mejor se ajustó, asegurando que los resultados sean fiables.

Finalmente, se empleó el comando "varImp" del paquete "caret" para calcular la importancia de las variables. La función "predict" permitió la creación de un ráster con predicciones basadas en los modelos ajustados. Por último, los valores generados en los modelos se reclasificaron del uno al cinco, conforme al número de clases de probabilidad de incendios (Tabla 9), y se representaron gráficamente en formato (.tif).

Tabla 9. Clases para la reclasificación de la probabilidad de ocurrencia de incendios

Probabilidad		Clase
$X \leq 0,2$	1	Muy baja
$0,2 > X \leq 0,4$	2	Baja
$0,4 > X \leq 0,6$	3	Media
$0,6 > X \leq 0,8$	4	Alta
$X > 0,8$	5	Muy alta

5.4.3. Etapa 3: validación del modelo

Validar el modelo de probabilidad es crucial para garantizar su precisión y confiabilidad, evitando el sobreajuste y asegurando que generalice bien a nuevos datos, permitiendo evaluar el desempeño del modelo a través de métricas objetivas, mejorando su robustez y estabilidad

frente a variaciones. Además, incrementa la confianza en las predicciones y asegura que el modelo sea útil y aplicable en situaciones del mundo real.

Para ello, todo este proceso se lo realiza en la aplicación Rstudio, en la que inicia con el 75% de la base de datos el cual fue usado para poder entrenar el modelo y una vez entrenados los algoritmos de aprendizaje automático (RL, MARS, LMT) fueron evaluados con el 25% restante. En ambos casos, se consideraron medidas estadísticas de: área bajo la curva de ROC e índice Kappa de Cohen para poder evaluar la sensibilidad de los modelos y su capacidad como clasificadores. Por medio de estas dos métricas de evaluación, se determinó la técnica de aprendizaje que mejor se adaptó a la base de datos.

El área bajo la curva de ROC permitió conocer las probabilidades por las cuales se clasifica la observación, evaluando el rendimiento de un modelo en la clasificación de datos. Los valores de AUC varían de 0,5 a 1,0 en donde los valores de 0,9 a 1,0 indican que el modelo tiene un nivel de precisión excelente, generalmente, los valores superiores a 0,70 son considerados como buenos predictores, con un nivel de precisión moderada, mientras que valores bajos (0,5 a 0,6) indican un nivel de precisión insuficiente (Arabameri *et al.*, 2020).

El índice de Kappa se obtiene al asociarlo con la matriz de confusión (Figura 13), y permite visualizar el rendimiento de un algoritmo y las clases reales de los datos, por lo general, de aprendizaje. En una matriz de confusión, cada fila representa las predicciones de una clase específica, mientras que cada columna representa las instancias reales de esa misma clase (o viceversa). Su nombre se debe a que facilita la identificación de confusiones entre clases, es decir, la tendencia del sistema a etiquetar incorrectamente una clase como otra. En el caso de clasificaciones binarias, donde el resultado puede tomar solo dos clases, se utiliza esta matriz para evaluar el rendimiento del modelo (Paredes, 2020).

	Realmente es positivo	Realmente es negativo
Predicho como positivo	Verdaderos Positivos (VP)	Falsos Positivos (FP)
Predicho como negativo	Falsos Negativos (FN)	Verdaderos Negativos (VN)

Figura 13. Matriz de confusión

Nota: Obtenido de Data Science con R. Análisis de datos y algoritmos de predicción por Paredes (2020).

En la Tabla 10 se observan las principales funciones que se ejecutaron en R para obtener dichas medidas.

Tabla 10. Curva de ROC e índice Kappa de Cohen en R

Descripción	Función en R	Argumentos
Curva de ROC: paquete (pROC)	<code>roc(predictor, response)</code>	Predictor: probabilidades predichas para cada clase Response: un valor de factor numérico o de caracteres de respuesta (codificado como 0 para controles y 1 para casos). Clase verdadera a la que pertenece cada observación.
Área bajo la curva AUC	<code>auc(roc)</code>	Roc: objeto de la función ROC.
Matriz de confusión	<code>confusionMatrix(data_prediccion, output_real, positivo = "")</code>	Data_prediccion: es un objeto o factor de clases previstas. Output_real: en un objeto o factor de clases verdadera a la que pertenece cada observación. Positivo: una cadena de caracteres opcional para el nivel de factor que corresponde a un resultado "positivo".

Cabe recalcar que este modelo usa imágenes satelitales proporcionadas por el satélite “Sentinel 2A”, los resultados obtenidos fueron presentados en mapas temáticos en Google Earth Engine, los cuales se actualizarán automáticamente cada 5 días, permitiendo una comprensión visual efectiva y actualizada de las zonas de mayor riesgo.

Para visualizar los resultados, se pueden utilizar las herramientas de visualización disponibles, las cuales permiten analizar los resultados del modelo y observar las áreas con mayor riesgo de incendios forestales en base a un rango de susceptibilidad (Figura 14). Además, el modelo puede ser actualizado y mejorado continuamente a medida que se disponga de nuevos datos, lo que permite una mayor precisión y capacidad predictiva a lo largo del tiempo. Estas predicciones pueden ser utilizadas para tomar decisiones informadas y planificar acciones futuras.

GRADO DE SUSCEPTIBILIDAD

	RANGO	CATEGORIA	DESCRIPCIÓN
	0 a 0.20	Muy baja	Escasa o nula probabilidad de ocurrencia de incendios forestales.
	0.20 a 0.40	Baja	Incendios forestales pueden ocurrir bajo comportamiento natural anómalo muy puntual.
	0.40 a 0.60	Media	Ocurrencia de incendios forestales responde la ciclicidad normal de convergencia de variables naturales asociadas.
	0.60 a 0.80	Alta	Todas las condiciones estáticas necesarias para la ocurrencia de incendios forestales.
	0.80 a 1	Muy alta	Todas las condiciones estáticas necesarias para la ocurrencia de incendios forestales. Procesos de ignición latentes.
		Sin datos	Zonas con alta nubosidad. Información satelital no disponible.

Figura 14. Rango de susceptibilidad a incendios forestales

5.5. Automatización del modelo en Google Earth Engine

En el presente estudio la automatización del riesgo de incendios forestales se realizó mediante clustering para identificar patrones espaciales y temporales de los datos, la Figura 15, esquematiza el proceso.

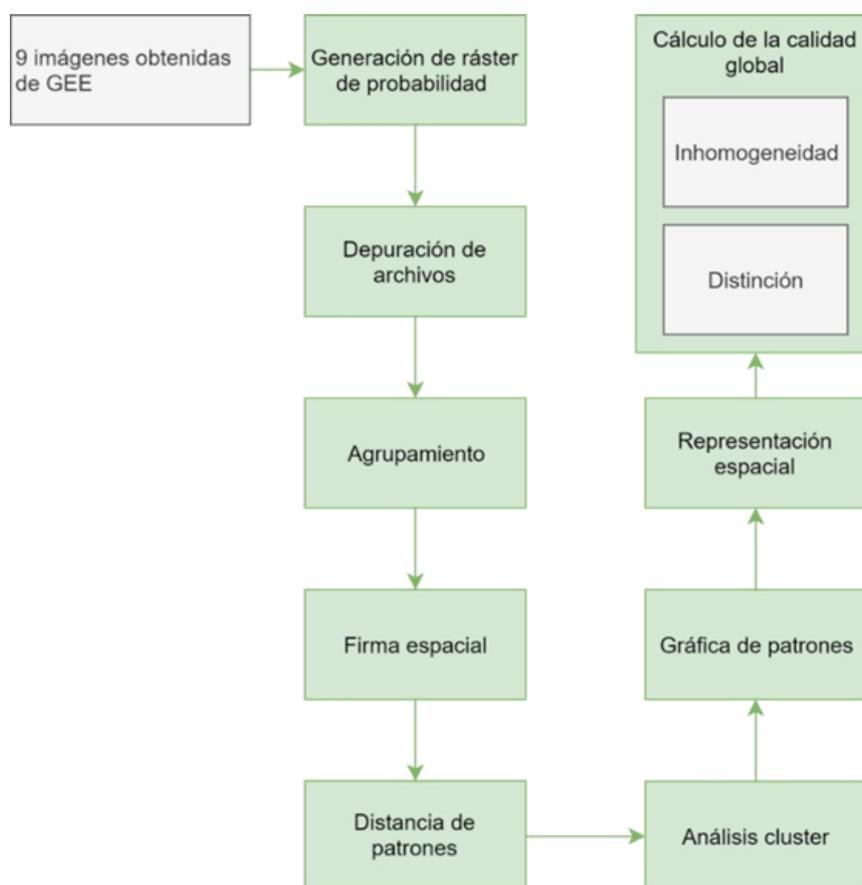


Figura 15. Proceso del análisis espacio – temporal para la automatización del riesgo de incendios

A partir del modelo de Regresión Logística de probabilidad, el cual demostró la mayor especificidad y rendimiento en este trabajo, se llevó a cabo la identificación de los patrones en imágenes de susceptibilidad a incendios forestales reclasificadas para cada mes del año (Enero – diciembre), debido a que este año presenta menor porcentaje de nubosidad con respecto a años anteriores y posteriores al mismo. Estas imágenes fueron obtenidas por medio del código creado por Reyes y Loján (2022), en la plataforma Google Earth Engine (GEE), el cual fue adaptado a nuestra zona de estudio. Es importante mencionar que para las nueve imágenes descargadas se procedió a reemplazar los valores correspondientes a la presencia de nubes por valores nulos utilizando el software QGIS.

Para el análisis clúster, los archivos ráster de las imágenes, fueron agrupados cronológicamente utilizando el paquete "motif" de R, el cual posibilita la identificación de patrones en series temporales (Nowosad, 2021). Para calcular la firma espectral para una ventana de 25 píxeles, se determinó la distancia entre los patrones de cada zona mediante el método de divergencia de Jensen-Shannon. Posteriormente, se llevó a cabo un análisis de clúster de manera jerárquica, utilizando el método Ward, que minimiza la varianza total dentro del clúster, en la que se especificó 5 clústeres que reflejan la mayor y menor susceptibilidad de los patrones.

Se realizó el cálculo de la calidad global de la agrupación con el objetivo de evaluar la efectividad del agrupamiento de los clústeres. Este cálculo se basó en la inhomogeneidad y la distinción de los clústeres, tal como se muestra en la Ecuación 7.7.

Ecuación 7. Calidad Global

$$Calidad\ global = 1 - \left(\frac{inhomogeneidad}{distinción} \right)$$

La inhomogeneidad es una medida que cuantifica la distancia mutua entre los objetos dentro de un clúster, con valores que oscilan entre 0 y 1. Valores cercanos a 0 indican una mayor consistencia en los patrones dentro del clúster. Por otro lado, la distinción proporciona una medida promedio de la distancia entre el clúster en cuestión y los demás, valores cercanos a 1 indican que el clúster específico se diferencia en mayor medida de los demás. Por lo tanto, la calidad global del clúster mejora a medida que el valor se acerca a 1 (Nowosad, 2021).

6. Resultados

6.1. Línea base de las variables predictivas para identificar la ocurrencia de incendios forestales

Se construyó una base de datos espacial para el periodo 2018-2021 (Anexo 1), siguiendo la metodología propuesta por Reyes y Loján (2022), que clasifica las variables en tres grupos: variables topográficas, variables de interacción antrópica y variables del estado de la vegetación, también conocidos como índices espectrales.

6.1.1. Variables topográficas

Las variables topográficas son características físicas del terreno que describen su forma, estructura y relación con otras características geográficas, permitiendo un mejor entendimiento de en el comportamiento de un incendio forestal. A continuación, se explican en detalle las variables mostradas en la Figura 16:

- La pendiente del territorio de Zapotillo varía significativamente, predominando una pendiente baja de 0% a 15%, que abarca el 69.69% del total de la zona de estudio. Asimismo, las pendientes con mayor rango, de 25% a 70%, constituyen un 8.17% del total del territorio. En pendientes altas, un incendio se dispersa de manera más rápida en comparación con las pendientes bajas. Por lo tanto, es importante tener en cuenta esta variabilidad para implementar medidas de mitigación adecuadas. En el caso del cantón Zapotillo, las áreas con pendientes bajas son más extensas, lo que indica que un incendio tiende a propagarse de manera más lenta en la mayor parte de esta región.
- La orientación del terreno está distribuida uniformemente, sin variaciones considerables entre los diferentes rangos de orientación, ya que cada uno cubre aproximadamente una cuarta parte del territorio total. Cabe recalcar que el rango de orientaciones entre 180° y 360° corresponde a las horas de la tarde, cuando el sol incide de manera más directa, aumentando significativamente la temperatura y, por ende, la probabilidad de incendios forestales debido a las condiciones más secas y calurosas del territorio.
- La elevación del terreno en el cantón Zapotillo muestra que gran parte se encuentra entre 140 y 320 metros sobre el nivel del mar (m.s.n.m.), abarcando un 37.86% del territorio total. Estas áreas son generalmente más secas y pueden presentar mayor cantidad de oxígeno, factores que pueden influir en la propagación y el comportamiento del fuego durante un incendio forestal. Por otro lado, las áreas de mayor altitud, que representan

el 8% del territorio y se encuentran en el rango de 670 m.s.n.m. a 1080 m.s.n.m., presentan un terreno más irregular, representando un desafío adicional en la mitigación del fuego debido a la dificultad para acceder a estas zonas y para controlar la propagación del fuego en terrenos más empinados.

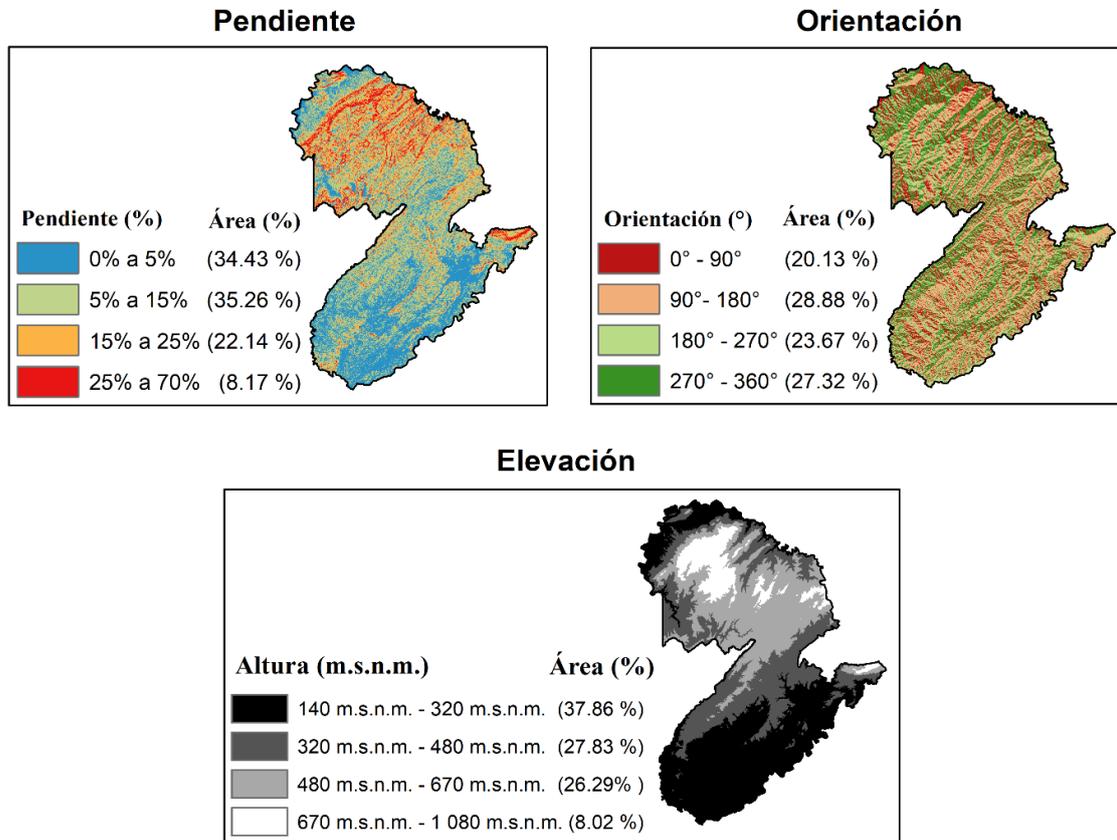


Figura 16. Variables topográficas

6.1.2. Variables de interacción antrópica (accesibilidad)

Las variables de accesibilidad son fundamentales para evaluar la proximidad medida en tiempo hacia diferentes zonas, lo cual puede influir significativamente en la rapidez de las respuestas de emergencia contra incendios por parte del cantón. A continuación, se explican en detalle las variables mostradas en la Figura 17:

- **Accesibilidad a áreas protegidas:** El acceso a estas áreas en menos de 40 minutos corresponde al 80% del cantón, lo que garantiza una respuesta rápida en caso de ignición por parte de los servicios de emergencia en áreas protegidas. Por otro lado, solo el 5% del área del cantón Zapotillo tiene una accesibilidad de 70 a 140 minutos, lo que significa que en estas pequeñas zonas la respuesta de los servicios de emergencia será más tardía.

- **Accesibilidad a poblados:** Se observa que un 53.26% del área total del cantón está cercano a los poblados con un tiempo de accesibilidad de 0 a 20 minutos. Esto presenta un contraste: por un lado, estas áreas pueden beneficiarse de una respuesta rápida y efectiva para la mitigación de incendios debido a la proximidad a los servicios de emergencia; por otro lado, la mayor interacción humana en las zonas rurales incrementa la probabilidad de que se produzcan incendios. Asimismo, un 5.5% del área total del cantón está a entre 70 y 130 minutos de los poblados, lo que dificulta una respuesta rápida en caso de emergencia de ignición.
- **Accesibilidad a zonas antrópicas:** El 63% del área total del cantón Zapotillo está entre 0 y 15 minutos de accesibilidad a zonas antrópicas, lo que permite una rápida respuesta de mitigación en caso de incendio por parte de los servicios de emergencia. Por otro lado, existen pequeñas zonas que corresponden al 3.8% del área total del cantón, con una accesibilidad de entre 60 y 120 minutos, lo que dificulta y retrasa la llegada de los servicios de emergencia. Es importante mencionar que las zonas cercanas a áreas antrópicas tienen una mayor probabilidad de incendios forestales debido a las actividades humanas, mientras que las áreas más alejadas tienen menos probabilidad de sufrir incendios.
- **Accesibilidad a centros económicos:** Se observa que un 45.6% del territorio total está entre 0 y 20 minutos de un centro económico, lo cual facilita la mitigación en caso de incendio debido a la cercanía a la infraestructura antrópica. Sin embargo, es importante tener en consideración que las zonas más alejadas de los centros económicos, están en un rango de 70 a 140 minutos de accesibilidad, corresponden a un 6% del área total, es decir la respuesta de los centros de emergencia será tardío en respuesta a un evento de ignición. Además, los centros económicos corresponden a áreas rurales con actividades antrópicas, lo que aumenta el riesgo de incendios forestales.
- **Accesibilidad a vías:** Se evidencia que un 41.5% del área total del cantón está a entre 0 y 30 minutos de las vías, lo que facilita una respuesta rápida de los servicios de emergencia en caso de incendios. Asimismo, un 6.6% del área total está entre 80 y 145 minutos de las vías, dificultando el acceso para mitigar estos eventos. No obstante, la proximidad a la red vial principal se lo considera como un factor que incrementa la probabilidad de incendios debido a la mayor presencia de actividades humanas.

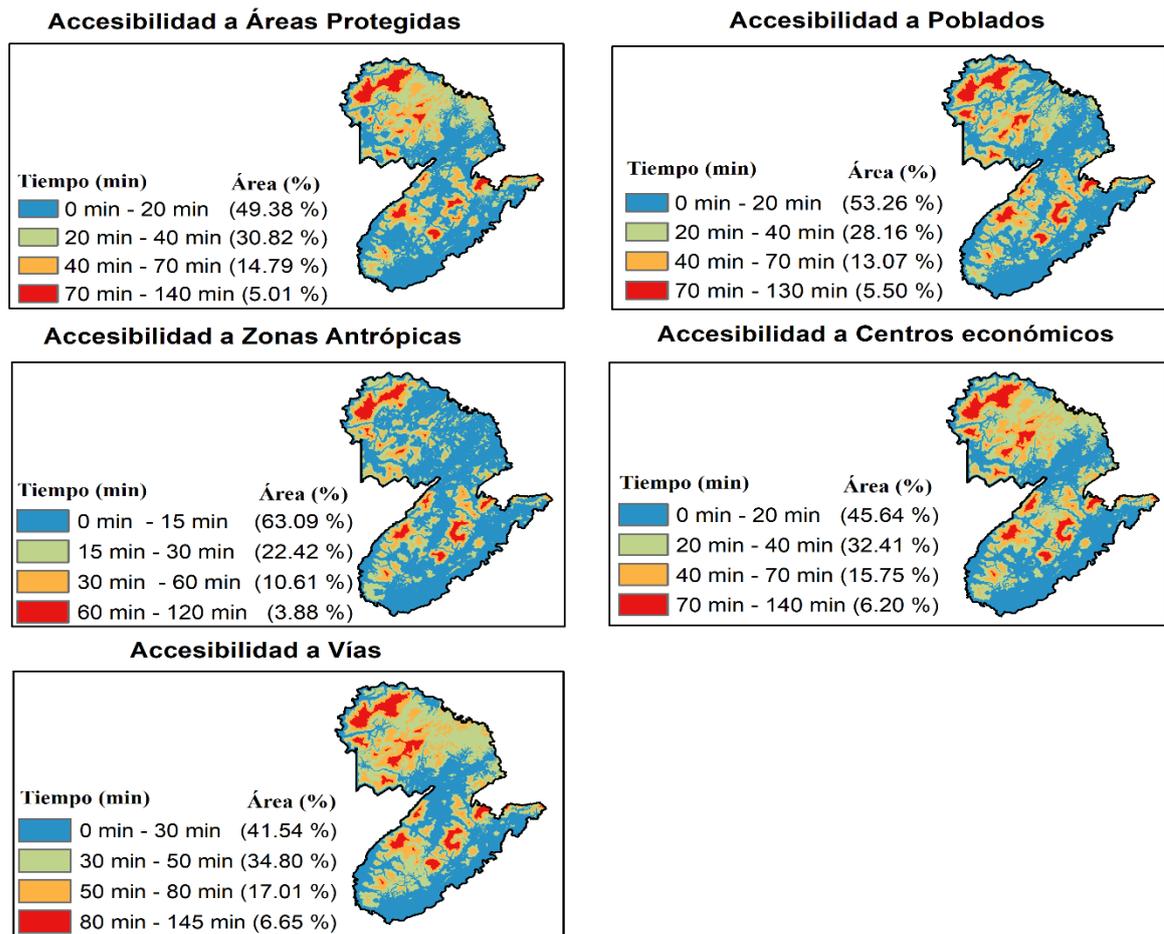


Figura 17. Variables de interacción antrópica (accesibilidad)

6.1.3. Variables de interacción antrópica (distancias)

Las variables de distancia son esenciales para evaluar la cercanía medida en metros hacia distintas zonas, lo cual puede tener un impacto significativo en la velocidad de respuesta de emergencia frente a los incendios forestales por parte del cantón Zapotillo. A continuación, se explican en detalle las variables mostradas en la Figura 18:

- **Distancia desde zonas antrópicas:** El 58.59% del territorio está cerca de zonas antrópicas, con distancias que van desde 0 hasta 350 metros, lo cual facilita una respuesta rápida por parte de los servicios de emergencia. Asimismo, existen pequeñas zonas que corresponden al 4.43% del área total, ubicadas a una distancia de 1900 a 3800 metros de las zonas antrópicas, lo que dificulta una intervención rápida por parte de los servicios de emergencia.
- **Distancia desde centros económicos:** El 27.45% del territorio total del cantón se encuentra ubicado a una distancia de 0 a 5000 metros de los centros económicos, lo que asegura una respuesta rápida de los servicios de emergencia ante eventos de ignición,

ya que están cerca de zonas rurales y/o urbanas. Por otro lado, las zonas más alejadas de los centros económicos cuentan una distancia de 13000 a 20600 metros, representando el 9.9% del territorio, llegando a experimentar retrasos significativos de respuesta por parte de los servicios de emergencia.

- **Distancia desde vías:** El 61.46% del área total del territorio está ubicado a una distancia de 0 a 300 metros desde las vías, facilitando el acceso rápido de los servicios de emergencia en caso de incendios. Sin embargo, las zonas más alejadas desde las vías, comprenden un 3.8% del área total (1600 a 4100 metros), por lo que puede llegar a presentarse mayores dificultades para los servicios de emergencia.
- **Distancia desde áreas protegidas:** Gran parte del territorio, que representa el 42.3%, se encuentra ubicado a una distancia de 0 a 1500 metros desde las áreas protegidas. Estas áreas, al ser zonas protegidas, tienen menos intervenciones humanas, pero son de alta prioridad para la protección ante eventos de incendios forestales por parte de los servicios de emergencia. Por otro lado, el 6.7% del territorio está ubicado a una distancia más alejada, en un rango que va desde 7100 a 12300 metros desde las áreas protegidas, lo que dificulta el acceso y la rápida respuesta de los servicios de emergencia ante eventos de ignición en estas zonas distantes.
- **Distancia desde ríos:** El 38% del territorio está ubicado a una distancia cercana a los ríos, en un rango desde 0 a 130 metros, facilitando el acceso a los recursos hídricos a los servicios de emergencia para mitigar los incendios. Por otro lado, las zonas más alejadas cuentan con el 6.7% del territorio, con una distancia de 500 a 1150 metros desde los ríos, disminuyendo el rápido acceso a fuentes de agua durante las emergencias. La cercanía a los ríos puede reducir el riesgo de inicio de incendios debido a la mayor humedad del área, mientras que las áreas más distantes pueden ser más propensas a incendios por la sequedad en la vegetación.

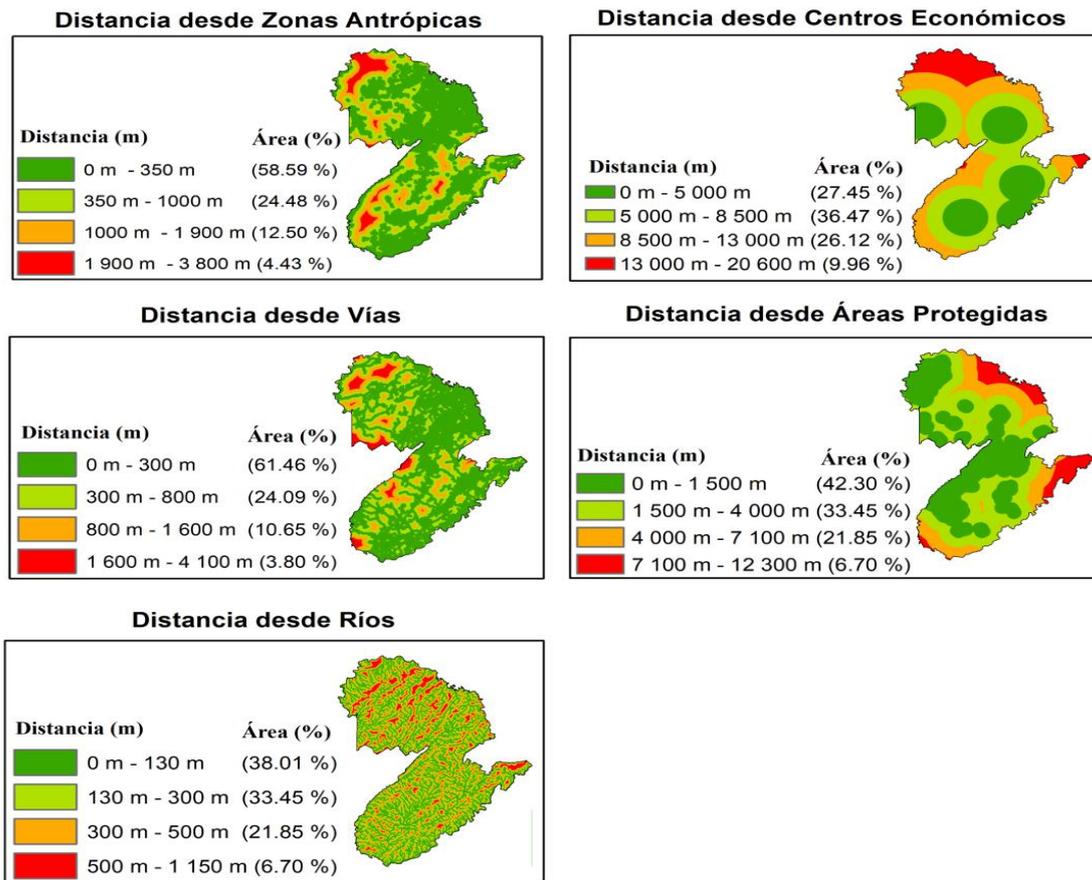


Figura 18. Variables de interacción antrópica (distancia)

6.1.4. Variables del estado de vegetación (Índices espectrales)

La Figura 19 corresponde a un ejemplo del estado de vegetación del cantón, correspondiente a un día de evaluación (14 de septiembre del 2019) de los Índices espectrales de condiciones del suelo y vegetación (Tabla 1 y Tabla 6). No se muestra más imágenes debido a la actualización periódica de las variables (5 días), es decir, no son estáticas, ya que reflejan los cambios que sufre el estado de vegetación de la zona, por lo que estas variables le sirven al modelo de probabilidad para poder identificar el tipo de vegetación en la zona de estudio.

Además, la Figura 19 muestra que el Índice de Vegetación de Diferencia Normalizada (NDVI) identifica regiones con alta biomasa, con áreas con valores de 0 a 1 (azul) representando mayor densidad de vegetación, y zonas con valores de -1 a 0 (verde) indicando menor densidad. El Índice de Relación de Normalización (NBR) revela que las regiones con valores de 0 a 1 (marrón oscuro) tienen mayor severidad de quemado, mientras que las áreas con valores de -1 a 0 (amarillo) son menos afectadas por incendios recientes. El Índice de Agua de Diferencia Normalizada (NDWI) muestra que las áreas con valores de 0 a 1 (verde oscuro) tienen mayor disponibilidad de agua, reduciendo la inflamabilidad, mientras que las áreas con valores de -1

a 0 (verde claro) tienen menor contenido de agua. El Índice de Vegetación de Diferencia Normalizada de Verde (GNDVI) indica que las áreas con valores de 0 a 1 (rojo) tienen mayor contenido de clorofila y vegetación más saludable, mientras que las zonas con valores de -1 a 0 (azul) tienen vegetación menos densa. El Índice de Suelo de Diferencia Normalizada (BSI) muestra áreas con valores de 0 a 1 (naranja oscuro) con mayor presencia de suelo expuesto, aumentando el riesgo de erosión y disminuyendo la humedad del suelo, mientras que las áreas con valores de -1 a 0 (naranja claro) presentan menor presencia de suelo expuesto. El Índice de Humedad de Diferencia Normalizada (NDMI) evalúa la humedad del suelo, con áreas con valores de -1 a 0 (marrón) indicando suelos más secos y áreas con valores de 0 a 1 (blanco) suelos más húmedos. El Índice de Humedad (Moisture) indica que las áreas con valores de 0 a 1 (azul oscuro) tienen mayor humedad, mientras que las áreas con valores de -1 a 0 (azul claro) tienen menor humedad, aumentando la susceptibilidad al fuego. El Índice de Vegetación Ajustado al Suelo (SAVI) muestra que las áreas con valores de 0 a 1 (violeta oscuro) tienen alta densidad de vegetación, y las áreas con valores de -1 a 0 (rosa claro) menor densidad. Finalmente, el Índice de Vegetación de Relación Ajustada (ARVI) corrige efectos atmosféricos, con áreas con valores de 0 a 1 (azul oscuro) indicando vegetación más densa y saludable, y áreas con valores de -1 a 0 (violeta claro) indica vegetación menos densa.

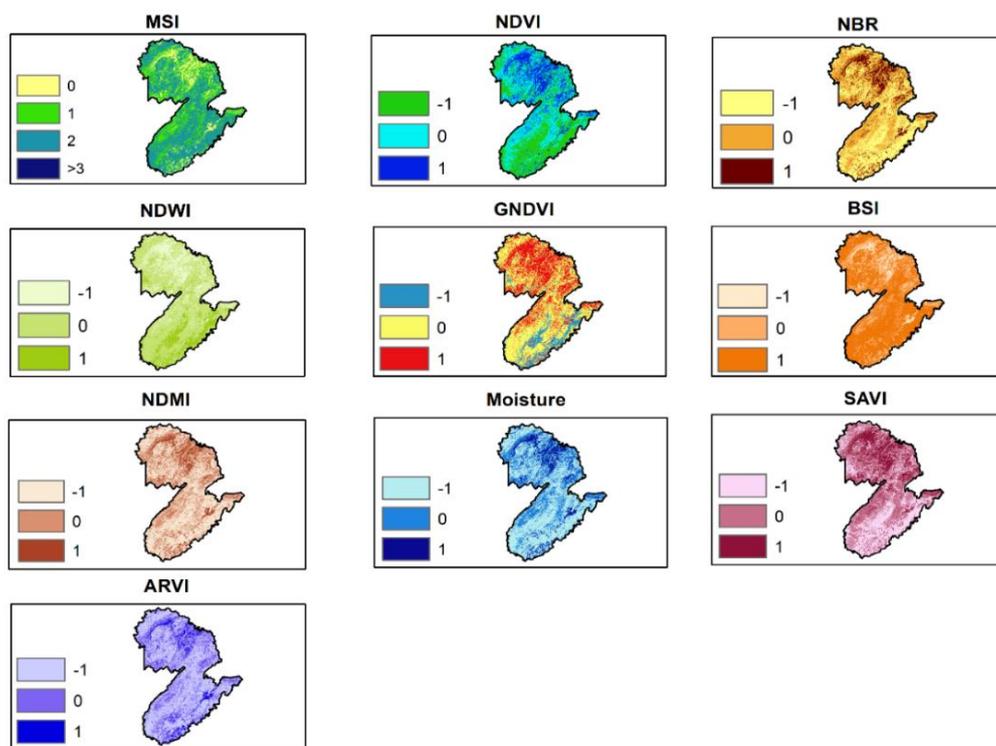


Figura 19. Variables del estado de la vegetación

Debido a la dificultad, el costo y la limitante de acceder a datos climáticos como la precipitación, la temperatura y el viento, se optó por utilizar los índices espectrales como una alternativa para estimar estas variables, ya que son de acceso libre y actualización constante.

Las variables espectrales que se relacionan con la variable “viento” son: NDVI, NDMI, GNDVI, ARVI y SAVI. Debido a que los índices espectrales son medidas que se obtienen a partir de una imagen satelital de Sentinel 2A por medio de la combinación de las diferentes bandas espectrales en la que se obtuvo el estado de la vegetación cada 5 días. Estos índices permiten determinar el estado de la vegetación y su humedad, estando relacionados con el potencial de ignición. El “viento” influye en el comportamiento del fuego, en la evaporación de la humedad, dirección de la ignición y velocidad de propagación del mismo.

Las variables que se relacionan con la variable “precipitación” son: MSI, GNDVI, NDVI, NDMI, Moisture. Estos índices nos permiten determinar el contenido de humedad de la vegetación, por lo que está relacionado directamente con el potencial de ignición, debido a que entre menos humedad exista, mayor probabilidad de que inicie un incendio y se propague con mayor intensidad.

Las variables espectrales que se relacionan con la variable “temperatura” son: NDVI, BSI, GNDVI, NBR, Moisture. Estos índices se relacionan con la variable de temperatura debido a que estos determinan el grado de sequedad de la vegetación y del suelo y su inflamabilidad, debido a que, a mayor temperatura, mayor probabilidad de producir un incendio.

6.2. Resultados del modelo de simulación espacial que estime la probabilidad de ocurrencia de incendios forestales en el cantón Zapotillo

Una vez empleado el análisis de multicolinealidad y el Factor de Inflación de Varianza (VIF) en el modelo de Regresión Logística (RL) en el cantón Zapotillo, se adquirió ocho variables, mostrados en la Tabla 11, que obtuvieron valores menores a 10, descartándolas de ser consideradas críticas, es decir, que no presentan problemas de multicolinealidad entre ellas, los cuales fueron: distancia a Zonas Antrópicas, distancia a Vías, distancia a Ríos, elevación, accesibilidad a Centros económicos, y los índices espectrales (BSI, NDVI y Moisture).

Cuanto más alto sea el valor del VIF, mayor probabilidad existe de que posea multicolinealidad, por lo que el valor del VIF aumenta con el incremento de la multicolinealidad. En la prueba VIF, los valores superiores a 10 se consideran críticos (Reyes y Loján, 2022).

Tabla 11. Resultados del Factor de Inflación de Varianza en las variables

Variabes	VIF
Distancia a ríos	1.211217
Accesibilidad a centros económicos	1.555325
Distancia a vías	1.591270
Distancia a zonas antrópicas	1.766030
NDVI	1.900793
Elevación	1.905604
BSI	6.535660
Moisture	8.743026

Los resultados (Figura 20) mostraron que los modelos (LMT y MARS) no se ajustaron a los datos, por lo tanto, no arrojaron resultados sobre las zonas con probabilidad de ocurrencia de incendios. Por otro lado, el modelo de RL si logró ajustarse a los datos, mostrando zonas con mayor y menor probabilidad de ocurrencia, aunque el modelo no es muy eficiente, esto debido a la falta de datos en la zona de estudio, ya que en el tiempo escogido de muestreo (2018 al 2021), se observó que existió una gran cantidad de nubes, limitando nuestra información. La escala del 1 al 5, clasifica los valores de los píxeles en clases de probabilidad de ocurrencia, desde muy baja a muy alta respectivamente.

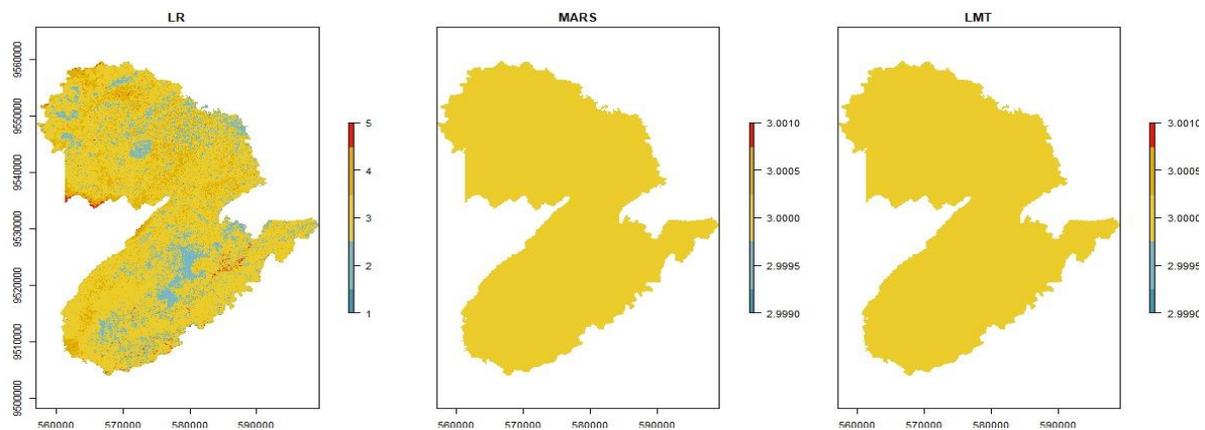


Figura 20. Mapeo de probabilidad de ocurrencia de incendios forestales para el cantón Zapotillo utilizando algoritmos de aprendizaje automático supervisado: Regresión Logística (RL), Splines de Regresión Adaptativa Multivariante (MARS) y Árboles de Decisión Logística

Como se observa en la Tabla 12, para Regresión Logística (RL) las variables más importantes fueron BSI e índice de humedad (Moisture); las variables medianamente importantes fueron NDVI, elevación, distancia a vías y distancia a zonas antrópicas, finalmente, las variables de menor importancia fueron, distancia a ríos y accesibilidad a centro económicos. Donde los índices espectrales son variables que fluctúan en función de las condiciones del suelo

y la vegetación. Por lo tanto, las variables mencionadas se pueden ir actualizando. Además, las cinco variables restantes (distancia a zonas antrópicas, distancia a ríos, distancia a vías, elevación, y accesibilidad a centros económicos) se consideraron variables estáticas, dado que su variación temporal es mínima.

Tabla 12. Importancia de las variables en la modelación de probabilidad de ocurrencia de incendios forestales utilizando el algoritmo RL

Variables	RL
BSI	1.10881831
Moisture	1.05248684
NDVI	0.48195506
Elevación	0.42405001
Distancia a vías	0.35627438
Distancia a zonas antrópicas	0.22905948
Distancia a ríos	0.08978006
Accesibilidad a centros económicos	0.02813539

Nota: Para obtener la importancia de las variables se utilizó el comando varImp del paquete Caret en R.

Los resultados que se obtuvo cuando se validó el modelo de probabilidad escogido, se muestra en la Tabla 13, en el que se observa que para el set de entrenamiento el modelo RL presentó un nivel de precisión de 0.56, sin embargo, para el set de validación presentó un nivel de precisión mayor de 0.71.

En base a los resultados del área bajo la curva, el modelo RL presenta un AUC = 0.65 para el set de entrenamiento y un AUC = 0.63 para el set de validación, por lo tanto, no se lo consideraría como buen predictor ya que se necesita un AUC > 0.7, pero en este caso, es el modelo que tiene mayor capacidad para clasificar los datos, en comparación a los demás.

Con respecto al índice de Kappa, para el set de entrenamiento el modelo RL presenta una concordancia débil y pobre (Kappa = 0.12), aunque para el set de validación presentó un valor mayor (Kappa = 0.34) sigue presentando una concordancia débil.

Las medidas de precisión, sensibilidad y especificidad se obtuvieron de la matriz de confusión, tomando en cuenta, la relación entre los datos clasificados como verdaderos positivos o verdaderos negativos. El modelo de RL presentó una capacidad moderada como clasificador por su especificidad.

Tabla 13. Rendimiento del modelo de Regresión Logística (RL)

Medias estadísticas	Set de entrenamiento	Set de validación
----------------------------	-----------------------------	--------------------------

Modelo de ML	RL	RL
Área bajo la curva (AUC)	0.65	0.63
Precisión	0.56	0.71
Índice Kappa	0.12	0.34
Sensibilidad	0.63	0.43
Especificidad	0.49	0.89

Nota: Los resultados se obtuvieron en R mediante la aplicación del comando AUC del paquete pROC y confusionMatrix del paquete caret.

En la Figura 21, se puede observar el número total de datos clasificados correctamente como puntos de incendio y los puntos de no incendio o control. Esta información fue obtenida por las matrices de confusión generadas en R como resultado del entrenamiento y validación del modelo RL. Para el set de entrenamiento, RL clasificó 43 datos correctamente, y 32 datos clasificados como errores, mientras que, para el set de validación, RL clasificó 48 datos correctamente y 20 datos clasificados como errores.

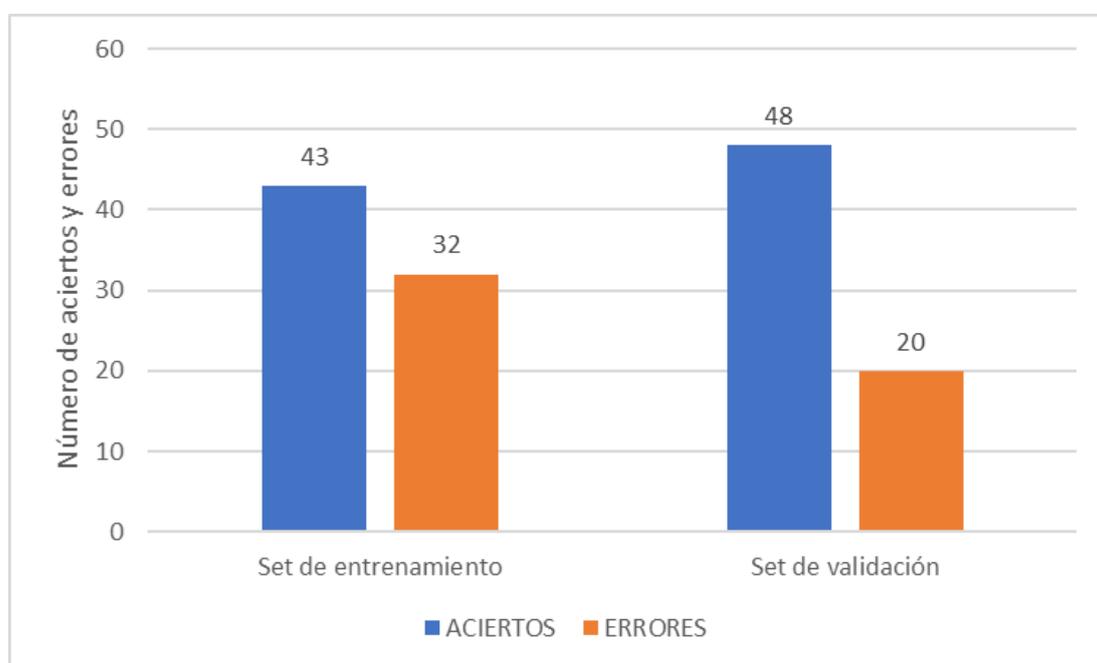


Figura 21. Número de datos de los sets de entrenamiento y de validación que fueron clasificados como correctos (aciertos) e incorrectos (errores) para el modelo de RL

6.3. Automatización del modelo de simulación espacial determinando el riesgo a incendios forestales

El reconocimiento de patrones de agrupamiento se realizó por medio del cálculo de firmas espectrales de un mosaico con nueve imágenes de ráster de probabilidad a incendios forestales, que fueron generados en GEE, para enero a diciembre del año 2020, en el cual se aplicó el modelo RL, debido a que tiene mayor capacidad de predicción que los demás modelos.

El número en el que se dio el agrupamiento se definió en 5 clases, dando a una buena separación entre los datos (Figura 22).

La Figura 22 muestra la identificación de los grupos de píxeles para las firmas espectrales. El clúster con la mayor cantidad de píxeles es el clúster dos, con 8030 píxeles, que cubre el 40.9% (50187.5 ha) de la superficie del cantón; seguido por el clúster uno, con 7633 píxeles que cubre el 38.9% (47706.25 ha) de la superficie del cantón.

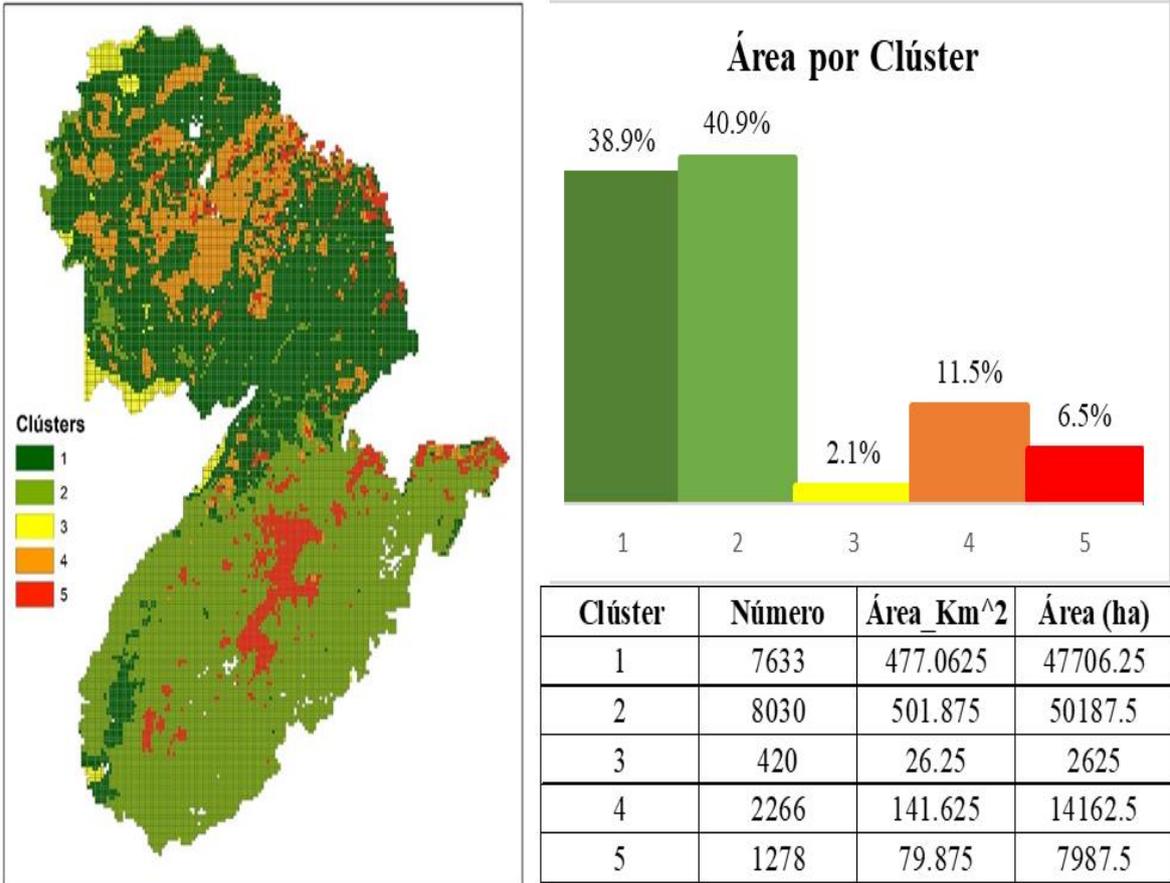


Figura 22. Identificación de clústeres en las imágenes ráster de probabilidad de ocurrencia de incendios forestales del cantón Zapotillo

Nota: Gráfica obtenida en R utilizando la función “lsp_add_clústeres”. El clustering fue jerárquico utilizando el método de Ward.

La Tabla 14 muestra la cuantificación de las áreas con distintas probabilidades de incendios forestales en el cantón Zapotillo, basado en un análisis espacio-temporal de clústeres para el año 2020. El análisis indica que, de enero a junio, la mayor parte del territorio de Zapotillo tiene una probabilidad muy baja de incendios (clúster 1). Sin embargo, entre julio y noviembre, la probabilidad de incendios aumenta ligeramente, predominando áreas con

probabilidad baja (clúster 2), lo que sugiere que, durante la mayor parte del año, el riesgo de incendios es bajo. Por otro lado, en diciembre, se observa un incremento en el porcentaje de áreas con una probabilidad muy alta de incendios (clúster 5), destacando un mayor riesgo en este mes, especialmente en las zonas cercanas a las áreas urbanas del cantón.

Tabla 14. Cuantificación del riesgo a incendios forestales, basado en un análisis espacio temporal, para el cantón Zapotillo, enero - diciembre del 2020

Imagen 1		Imagen 2		Imagen 3	
Enero - Febrero		Febrero - Marzo		Marzo - Mayo	
Probabilidad	Área (%)	Probabilidad	Área (%)	Probabilidad	Área (%)
Muy baja	47.96	Muy baja	29.75	Muy baja	18.06
Baja	44.58	Baja	45.64	Baja	66.11
Media	3.67	Media	22.11	Media	13.58
Alta	0.18	Alta	0.30	Alta	0.28
Muy alta	3.61	Muy alta	2.21	Muy alta	1.97
Imagen 4		Imagen 5		Imagen 6	
Mayo - Junio		Junio - Julio		Julio - Septiembre	
Probabilidad	Área (%)	Probabilidad	Área (%)	Probabilidad	Área (%)
Muy baja	52.18	Muy baja	25.54	Muy baja	8.72
Baja	41.77	Baja	64.28	Baja	77.25
Media	4.18	Media	8	Media	12.06
Alta	0.18	Alta	0.2	Alta	0.15
Muy alta	1.69	Muy alta	1.98	Muy alta	1.82
Imagen 7		Imagen 8		Imagen 9	
Septiembre - Octubre		Octubre - Noviembre		Noviembre - Diciembre	
Probabilidad	Área (%)	Probabilidad	Área (%)	Probabilidad	Área (%)
Muy baja	5.60	Muy baja	4.56	Muy baja	5.86
Baja	76.51	Baja	73.69	Baja	69.78
Media	15.4	Media	19.35	Media	13.79
Alta	0.09	Alta	0.10	Alta	0.19
Muy alta	2.39	Muy alta	2.30	Muy alta	10.38

La Figura 23 muestra la representación gráfica de la Tabla 14. Se evidencia en las parroquias Zapotillo, Limones, Garzareal y Mangahurco, son zonas que presentan una probabilidad de incendios forestales, que va de medianamente alta a permanentemente alta durante el periodo de julio a diciembre, lo cual está relacionado con la estacionalidad seca del cantón.

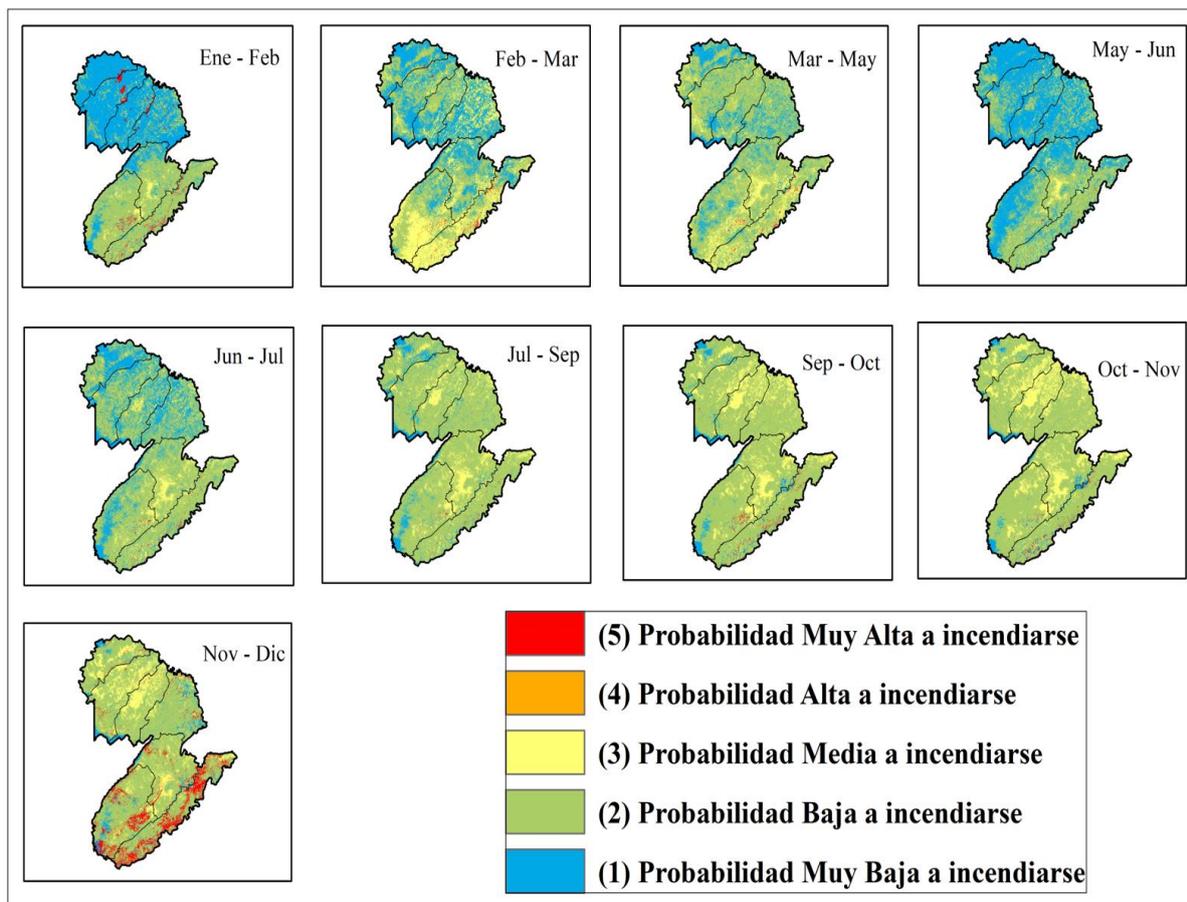


Figura 23. Zonificación del riesgo a incendios forestales, basado en un análisis espacio temporal, para el cantón Zapotillo, enero - diciembre del 2020

Nota: Gráfica generada en Arcgis, a partir de las imágenes reclasificadas (modelo RL) obtenidas de GEE en donde se eliminaron los píxeles (función “cloud.mask”) que presentan una probabilidad > 40 % de presencia de nubes. Las imágenes se reclasifican en 5 clases de probabilidad de ocurrencia de incendios.

La calidad global por cada clúster se observa en la Figura 24, donde la inhomogeneidad es baja, teniendo valores que van de 0.29 a 0.24 para todos los clústeres, demostrando que existe una buena consistencia o agrupamiento, debido a que son valores cercanos a 0. La distinción muestra valores de 0.45 a 0.82, por ende, la separación de cada clúster es moderada. El análisis de la validación del clustering se muestra de forma gráfica en la Figura 25, en donde se observa que la calidad global de los clústeres varía de 0.45 a 0.68, indicando una calidad de agrupación moderada.

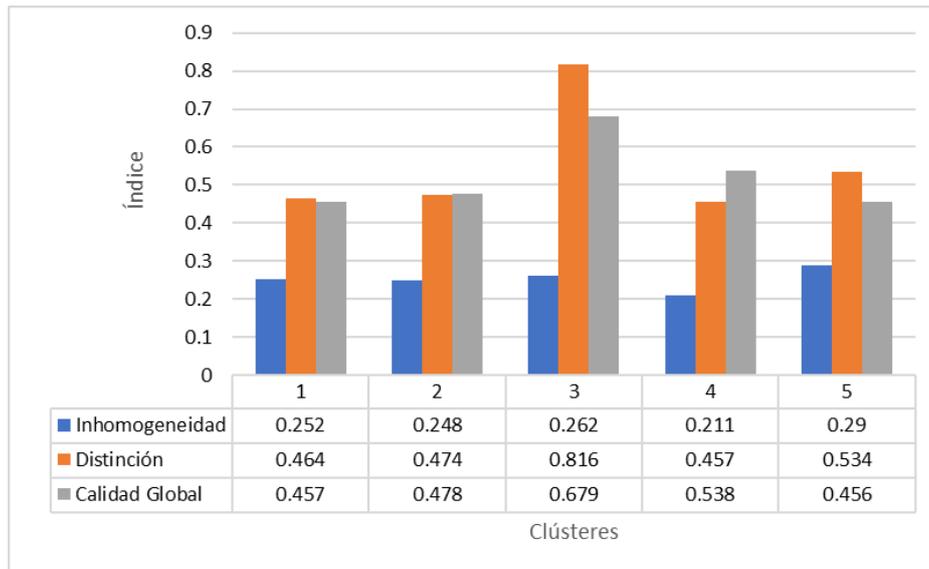


Figura 24. Inhomogeneidad, distinción y calidad global de cada clúster

Nota: La información se generó en R utilizando el paquete “motif”

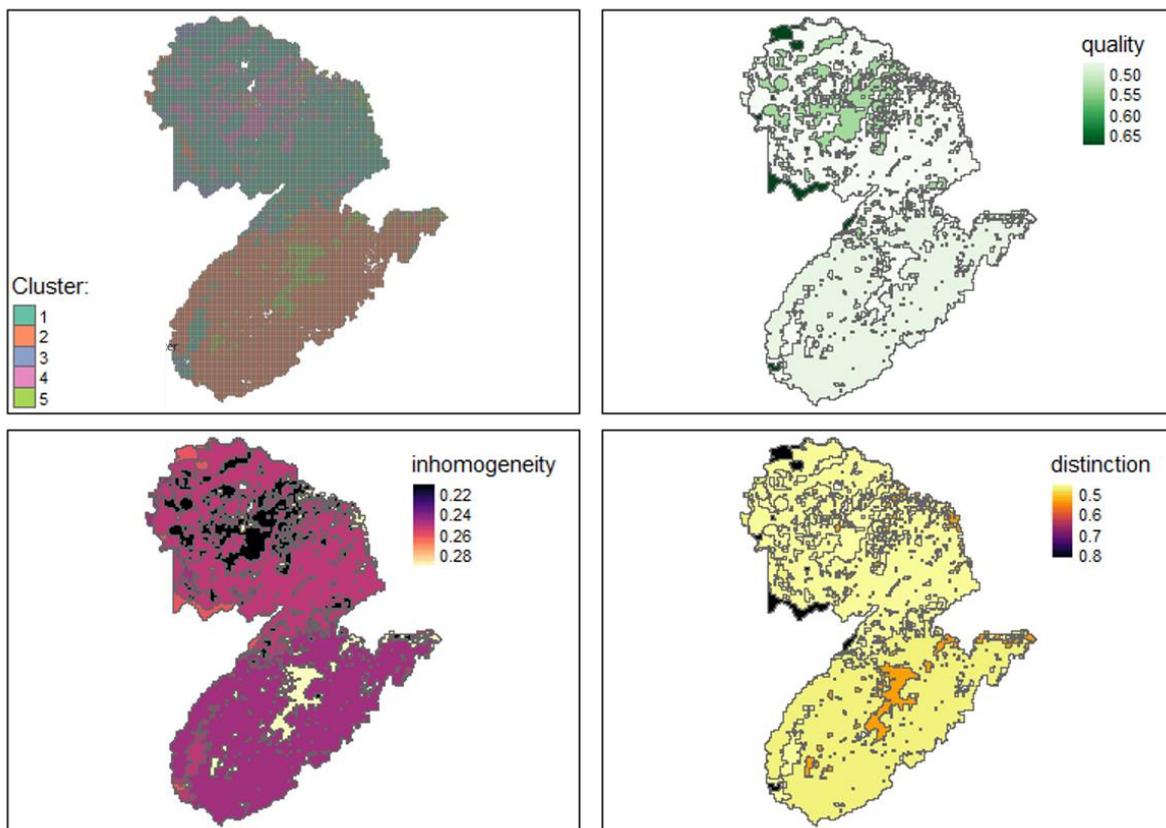


Figura 25. Mapeo de la calidad global, inhomogeneidad y distinción del análisis clúster para la zonificación del riesgo de incendios forestales

7. Discusión

La revisión de los estudios, concuerdan en que los incendios son generados por diferentes factores, pero hay que recalcar que las variables meteorológicas son las que más se destacan (Voltas, 2021), debido a la disponibilidad y facilidad de acceso de otros países a los datos espaciales (Rojas *et al.*, 2023); por lo que, son de mayor importancia para los modelos de probabilidad a incendios forestales, recibiendo hasta un 60% de importancia en estos eventos (Muñoz *et al.*, 2005). Sin embargo, en Ecuador, esta información espacial con buena resolución, y de actualización constante es limitada y tiene un costo. Por lo tanto, en el presente estudio, se consideró el uso de índices espectrales como indicadores de condiciones climáticas, con la finalidad de evaluar la vegetación por medio de la combinación de dos o más bandas, contribuyendo a la mejora del análisis de las propiedades de la vegetación y facilita la obtención de la distribución espacial del contenido de clorofila y otros componentes, así como la estructura de la vegetación de manera precisa y eficiente (Aponte y Ospina, 2019), logrando obtener una comparación multi-temporal entre imágenes antes y después del incendio (Opazo y Rodríguez, 2013).

Estudios previos sobre la predicción de incendios forestales por medio del Machine Learning mencionan que para poder predecir estos eventos se utilizan todas las variables que son parte del Triángulo del Comportamiento del Fuego, las cuales son la topografía, el combustible y la meteorología. Además, el índice NDVI ha sido utilizado en varios estudios debido a que indica la salud de la vegetación (Ibáñez, 2019), lo cual tiene relación a lo que menciona González *et al.* (2023) en la que un incendio se da por dos factores que se relacionan, las variables de ignición (causas naturales y las actividades antrópicas), y las variables de propagación (pendiente, viento, vegetación, humedad e inflamabilidad).

Los estudios revisados mostraron que gran parte de las variables utilizadas en el presente estudio, concuerdan con las variables utilizadas por Carrillo *et al.* (2012); González *et al.* (2023); Oliveira *et al.* (2012); Rojas *et al.* (2023); Sari (2021); Usmadi (2023); Voltas (2021), para determinar la probabilidad a incendios forestales con diferentes modelos. Las variables más utilizadas fueron: accesibilidad, pendiente, orientación, elevación, dirección y velocidad del viento, humedad, precipitación, temperatura, distancia a cuerpos de agua, distancia a carreteras, densidad poblacional, densidad forestal, distancia a zonas urbanas, uso

del suelo (ganadería y agricultura), contenido de humedad y grado de inflamabilidad. De estas, seis son variables antrópicas, dando a entender lo importante de considerar los factores antrópicos para comprender el comportamiento de un incendio forestal.

Este estudio optó por considerar el número de incendios forestales como la variable dependiente, mientras que investigaciones anteriores, como las de Adab *et al.* (2013), Bar *et al.* (2021), Nami *et al.* (2018) y Valdez *et al.* (2017), obtuvieron datos de incendios del Moderate Resolution Imaging Spectrometer (MODIS), este trabajo se basa en el enfoque metodológico de Reyes y Loján (2022) y Zhang *et al.* (2023), quienes emplearon datos de incendios obtenidos a través del Visible Infrared Imaging Radiometer Suite (VIIRS). Este método ha demostrado una mejora significativa en la resolución espacial, el alcance de observación y la calidad de los datos, lo que resulta en una capacidad mejorada de los datos de incendios de VIIRS para respaldar la gestión de incendios, los sistemas de alerta de incendios forestales casi en tiempo real y otras aplicaciones científicas que requieren una mayor precisión en la cartografía.

De igual manera, los artículos de Reyes y Loján (2022); Sari (2024); Zhang *et al.* (2023) utilizaron el Factor de Inflación de la Varianza (VIF), con la finalidad de eliminar las variables independientes que presenten valores superiores a diez, siendo así, variables con alta colinealidad; esta misma metodología es efectuada en este estudio. A pesar de esto, Vilar del Hoyo (2008) menciona que para ser considerados valores problemáticos o críticos deben ser superiores a treinta.

Una vez realizada el análisis de multicolinealidad y Factor de Inflación de la Varianza (VIF), identificamos las variables de importancia para este trabajo, que también han sido consideradas como significativas en otros estudios, como son los casos de Usmadi (2023) en cuanto a distancia a zonas antrópicas; Rojas *et al.* (2023) para distancia a vías y ríos; Zhang *et al.* (2023) para elevación y NDVI; Adab *et al.* (2013) en cuanto a Moisture; Reyes y Loján (2022) para accesibilidad a centros económicos y BSI. Por tanto, se demuestra lo significativo que llegan a ser las variables de distancia, accesibilidad, topográficos e índices espectrales para determinar la probabilidad a incendios forestales.

En base a lo que mencionan Galparsoro y Pita (2010) sobre la concordancia del índice de Kappa, los resultados obtenidos en este estudio revelan un desempeño pobre y débil del modelo de regresión logística, ya que el índice de Kappa arroja valores inferiores a 0.4 (0.12 –

0.34) tanto para el set de entrenamiento como para el de validación, respectivamente. Estos hallazgos contrastan con los resultados del desempeño del mismo modelo (RL) reportados en el estudio de Montealegre *et al.* (2017), en el que asume un desempeño bueno para el set de entrenamiento (Kappa = 0.68) y un desempeño moderado para el set de validación (Kappa = 0.57).

A pesar de que se conoce que el algoritmo de Regresión Logística es considerado como un creador de modelos con un buen nivel de rendimiento predictivo, y al requerir la conversión de la variable dependiente en dicotómica, impide la obtención de una estimación directa del número de incendios, permitiendo únicamente la estimación de la probabilidad de ocurrencia o no ocurrencia (Vilar del Hoyo *et al.*, 2002). Como en el estudio de Montealegre *et al.* (2017) y Sari (2024), quienes aplicaron un modelo de RL para determinar la susceptibilidad de los incendios forestales, en el que mostraron valores de AUC mayores a 0.9, siendo considerado un modelo con un alto rendimiento predictivo a la ignición. Pese a esto, el presente estudio mostró valores de AUC menores a 0.7, demostrando que el modelo presenta un rendimiento predictivo pobre para la zona de estudio, esto se dio por la alta presencia de nubosidad en Zapotillo del 2018 a 2021, lo cual queda registrado en las imágenes satelitales de Sentinel 2A, generando que los focos de calor detectados por VIIRS sean depurados, limitando la base de datos, causando una caída en el nivel de eficiencia del modelo.

Cabe recalcar que existen algoritmos similares de predicción de incendios como: Modelos Lineales Generalizados (GLM), con un AUC de 0.81, indicando ser un modelo con un rendimiento predictivo moderado (Vilar *et al.*, 2016); Redes Neuronales Artificiales (ANN) con un AUC de 0.84, indicando ser un modelo con un rendimiento predictivo moderado (Yilmaz, 2009); Máxima Entropía (MaxEnt) presentando un AUC de 0.94, demostrando ser un modelo con un rendimiento predictivo excelente (Sari, 2024); Random Forest (RF), presentando un AUC de 0.87 y 0.97 (Cáceres, 2023; Valdez *et al.*, 2017) respectivamente, demostrando ser un modelo con un rendimiento predictivo excelente. Es importante mencionar que después de una revisión de los artículos referente a modelos predictivos de incendios forestales, el modelo que mayor presencia y rendimiento presentó fue Random Forest (RF), siendo un algoritmo con mejor capacidad predictiva por encima del modelo de regresión logística (Montealegre *et al.*, 2017; Sari, 2024) que también es aplicado en este estudio.

El estudio de Jasiewicz *et al.* (2018) y Nowosad (2021) mencionan que la Inhomogeneidad mide el grado de diferencia o disimilitud mutua entre todos los objetos en un clúster, en el que su valor oscila entre 0 y 1, donde un valor bajo indica que todos los objetos en el clúster representan patrones consistentes, lo que sugiere que el clúster es homogéneo en cuanto a sus características. El presente estudio presenta una inhomogeneidad baja (0.29 a 0.24), demostrando tener un agrupamiento bueno entre los clústeres, ya que presenta patrones consistentes.

Por otro lado, Nowosad (2023) también comenta que la Distinción es una medida de la distancia media entre el clúster de interés y todos los demás clústeres. Este valor varía entre 0 y 1, donde un valor alto indica que el clúster se diferencia significativamente del resto de los clústeres. El presente estudio exhibe una distinción que se muestra desde 0.45 a 0.82, exponiendo que la separación de cada clúster es moderada. Estos resultados son compatibles, debido a que la inhomogeneidad y la distinción muestran tendencias opuestas: a medida que aumenta el número de clústeres, la inhomogeneidad disminuye, pero la separación o distinción entre clústeres aumenta (Amat, 2017).

De igual manera Nowosad (2021) menciona que la calidad global (quality) es una medida que combina la inhomogeneidad y la distinción para evaluar la calidad de la agrupación, es decir, es una forma de obtener la coherencia dentro de los clústeres y la separación entre ellos en una sola medida, donde valores más altos indican una agrupación de mayor calidad. En el presente estudio, presenta una calidad global que va desde 0.45 a 0.68, indicando una calidad de agrupación moderada.

Esta agrupación se la puede evidenciar en la representación gráfica del análisis espaciotemporal del clúster para cada periodo evaluado en el año 2020 en la zona de estudio, mostrando cinco clústeres. En la clase 1 se representan las áreas con baja probabilidad de ocurrencia, mientras que en la clase 5 se identifican las zonas pequeñas con alta probabilidad de ocurrencia. Esto sugiere que, en el cantón Zapotillo, los primeros meses del año presentan una menor probabilidad de incendios debido a que son las épocas con mayor precipitación, mientras que los periodos secos, que se extienden desde julio hasta diciembre, muestran una mayor probabilidad de incendios. Este hallazgo se respalda con el estudio de Samaniego *et al.* (2015), y el PDOT (2019) de Zapotillo, en el que están de acuerdo que, a nivel del sur de Loja (Zapotillo, Macará, parte baja de Puyango, Pindal) presentan un periodo de lluvias que

predominan el primer y segundo trimestre, y una estación seca que se presentan el tercer y cuarto trimestre del año.

Cada clúster exhibe un patrón de comportamiento único durante el período de tiempo analizado, el cual está vinculado a las variables: NDVI, BSI, Moisture, influenciadas por condiciones climáticas y características del entorno vegetal. Los clústeres con patrones de comportamiento menos cambiantes están espacialmente asociados a áreas donde las variables, especialmente la distancia a zonas antrópicas y la accesibilidad al centro económico, son favorables, lo que respalda la afirmación de Voltas (2021) de que el riesgo de incendios se concentra principalmente en zonas urbanas.

La validación de la automatización del riesgo a incendios en este análisis se respalda con la información histórica de incendios que han ocurrido en el cantón Zapotillo, siendo las parroquias: Zapotillo, Paletillas y Mangahurco (Cazaderos) las zonas con mayor riesgo de incendios forestales que para el año 2020 sumaron 366 ha quemadas (SNGRE, 2021).

8. Conclusiones

Los factores que influyen en la probabilidad de ocurrencia de incendios forestales son los ambientales, meteorológicos y antrópicos, en la que destacan las variables de cobertura vegetal como los índices espectrales (humedad de la vegetación, condiciones del suelo y de la vegetación), logrando tener mayor importancia, por encima de las variables topográficas y antrópicas. Cada análisis de las variables debe ser abordado teniendo en cuenta las particularidades y condiciones específicas del territorio, por lo que la relación de las variables con los modelos no es definitiva y no debe ser empleada como un criterio absoluto para la selección de variables de entrada en un modelo de aprendizaje automático.

Los modelos de Regresión Adaptativa Multivariante (MARS) y Árboles de Decisión Logística (LMT) fueron descartados debido a su falta de ajuste a los datos de Zapotillo. Aunque el modelo de Regresión Logística se ajustó a los datos, su precisión predictiva resultó ser baja. En este estudio, se tomaron en cuenta ocho variables en base a un análisis de multicolinealidad y de importancia, donde las variables más importantes para el modelo de Regresión Logística (RL) fueron BSI y Moisture; las variables medianamente importantes fueron NDVI, elevación, distancia a vías y distancia a zonas antrópicas, por último, las variables de menor importancia fueron distancia a ríos y la accesibilidad al centro económico.

Los resultados de la automatización del riesgo de incendios forestales revelan que las parroquias Zapotillo, Limones, Garzareal y Mangahurco son zonas con una probabilidad de incendios forestales que va de medianamente alta a permanentemente alta durante el periodo de julio a diciembre, lo cual está relacionado con la estacionalidad seca del cantón. Asimismo, el análisis automatizado indica que, durante el mismo periodo, las zonas urbanas presentan una alta probabilidad de ignición. Esta información es crucial para la gestión y prevención de estos eventos, y será proporcionada a las autoridades locales con el fin de desarrollar estrategias de prevención y mitigación en las áreas más susceptibles a incendios.

9. Recomendaciones

Uno de los desafíos al utilizar imágenes satelitales para el monitoreo de incendios forestales es la alta presencia de nubes en las imágenes, lo que dificulta el uso completo de los datos obtenidos por VIIRS. En este estudio, se tuvo que recurrir a la depuración de los mismos, y poder utilizar solo aquellas que no afecten la precisión de las predicciones. Por lo tanto, se sugiere investigar y aplicar nuevas técnicas que permitan un análisis exhaustivo sin la necesidad de descartar datos que podrían ser importantes para un análisis acumulativo.

Además, ampliar el período de recolección de datos, debido a que en este estudio se recopilaban datos de incendios forestales durante tres años (2018 – 2021). Por lo tanto, se sugiere una recopilación de información durante un período de 5 a 10 años, al igual que otros estudios, para mejorar la capacidad predictiva del modelo.

10. Bibliografía

- Adab, H.; Kanniah, K.D.; Solaimani, K. (2013). Modeling forest fire risk in the northeast of Iran using remote sensing and GIS techniques. *Natural Hazards* 65, 1723–1743.
- Aguilera, A. I., Cala, L. D., y Subero, A. R. (2010). Modelo basado en metaclassificadores para diagnóstico en marcha patológica mediante análisis cinético. 7–16.
- Aguirre Mendoza, Z., Buri Sivisaca, D., Geada López, G., y Betancourt Figueras, Y. (2014). Composición florística, estructura y endemismo en una parcela permanente de bosque seco en Zapotillo, provincia de Loja, Ecuador.
- Aguirre, Z., Kvist, L. P., y Sánchez, O. (2006). Bosques secos en Ecuador y su diversidad. *Botánica económica de los Andes Centrales*, 2006, 162-187.
- Aguirre, Z., Moran, M. E. R., y Moser, V. G. (2019). Productos forestales no maderables de los bosques secos de Zapotillo, Loja, Ecuador. *Arnaldoa Universidad Nacional de Loja*, 26(2), 575–594. <http://www.scielo.org.pe/pdf/arnal/v26n2/a04v26n2.pdf>
- Amat, J. (2017). Clustering y heatmaps: aprendizaje no supervisado. https://www.cienciadedatos.net/documentos/37_clustering_y_heatmaps#Calidad_de_los_clústeres
- Anchaluisa, S. (2013). Efecto del fuego sobre la estructura, microclima y funciones ecosistémicas de los bosques introducidos de eucalipto (*Eucalyptus globulus*, Mytaceae) en el Distrito Metropolitano de Quito. Universidad San Francisco de Quito. Disponible en: <http://repositorio.usfq.edu.ec/bitstream/23000/2899/1/109458.pdf>.
- Arabameri, A., Nalivan, O. A., Saha, S., Roy, J., Pradhan, B., Tiefenbacher, J. P., y Ngo, P. T. T. (2020). Novel ensemble approaches of machine learning techniques in modeling the gully erosion susceptibility. *Remote Sensing*, 12(11).
- Aponte, S. J., y Ospina, N. J. E. (2019). Evaluando el desempeño de índices espectrales para identificar humedales alto andinos. *Revista de Teledetección*, 53, 59. <https://doi.org/10.4995/raet.2019.10580>
- Banco de Desarrollo de América Latina (CAF). (2020). Ecuador Fortalecerá la Prevención de Incendios Forestales en las Zonas más Vulnerables. <https://www.caf.com/es/actualidad/noticias/2020/05/ecuador-fortalecera-la-prevencion-deincendios-forestales-en-las-zonas-mas-vulnerables/>
- Bar, S., Parida, B. R., Roberts, G., Pandey, A. C., Acharya, P., y Dash, J. (2021). Spatio-temporal characterization of landscape fire in relation to anthropogenic activity and

- climatic variability over the Western Himalaya, India. *GIScience and Remote Sensing*, 58(2), 281–299. <https://doi.org/10.1080/15481603.2021.1879495>
- Bodí, M. B., Cerdá, A., Mataix-Solera, J., y Doerr, S. H. (2012). Efectos de los incendios forestales en la vegetación y el suelo en la cuenca mediterránea: revisión bibliográfica. *Boletín de la asociación de Geógrafos Españoles*.
- Brovkina, O., Stojanović, M., Milanović, S., y Latypov, I. (2020). Monitoring of post-fire forest scars in Serbia based on satellite Sentinel-2 data. *Geomatics, Natural Hazards and Risk*, 11(1), 2315–2339. <https://doi.org/10.1080/19475705.2020.1836037>
- Buzzi, M. A., Rueter, B. L., y Ghermandi, L. (2017). Múltiples índices espectrales para predecir la variabilidad de atributos estructurales y funcionales en zonas áridas. *Ecología Austral*, 27(1), 055–062. <https://doi.org/10.25260/ea.17.27.1.0.315>
- Cáceres Coello, J. (2023). Comparación y análisis estadístico entre modelos de probabilidad de ocurrencia de incendios forestales para Honduras. Año 2019. In *Ciencias Espaciales* (Vol. 14, Issue 2, pp. 19–37). <https://doi.org/10.5377/ce.v14i2.16918>
- Calvo-Alvarado, J., Sánchez-Azofeifa, A., y Portillo-Quintero, C. (2013). Neotropical Seasonally Dry Forests. *Encyclopedia of Biodiversity*, 5, 488–500. doi: 10.1016/B978-0-12-384719-5.00354-3
- Carbajal-Morán, H. I., y Onofre-Enriquez, F. I. (2020). Determinación con imágenes satelitales del índice de vegetación de diferencia normalizada del valle de Pampas-Tayacaja Determination with satellite images of the normalized difference vegetation index of the Pampas-Tayacaja valley Determinação com imagens de satélite do índice de vegetação de diferença normalizada do vale Pampas-Tayacaja *Ciencias técnica y aplicada* Artículo de investigación. 5(1), 228–245. <https://doi.org/10.23857/pc.v5i1.1893>
- Cárdenas, M., Medel, R., Castillo, J., Vázquez, J. C., y Casco, O. (2015). Modelos de Aprendizaje Supervisados: aplicaciones para la predicción de incendios forestales en la provincia de Córdoba. XVII Workshop de Investigadores En Ciencias de La Computación, 1–5. <http://sedici.unlp.edu.ar/handle/10915/45467%0Ahttp://hdl.handle.net/10915/45467>
- Carrillo García, R. L., Rodríguez Trejo, D. A., Tchikoué, H., Monterroso Rivas, A. I., y Santillan Pérez, J. (2012). Análisis espacial de peligro de incendios forestales en Puebla, México. *Interciencia*, 37, 678–683. <https://www.redalyc.org/pdf/339/33925502012.pdf>

- Castellnou, M., Pagés, J., Miralles, M., y Piqué, M. (2009). Tipificación de los incendios forestales de Cataluña. Elaboración del mapa de incendios de diseño como herramienta para la gestión forestal. 5o Congreso Forestal Avila, 1–15
- Castillo, M., Pedernera, P., y Pena, E. (2003). Incendios forestales y medio ambiente: una síntesis global. *Revista Ambiente y Desarrollo*, 9(3), 44-53.
- Cerda, Jaime, y Cifuentes, Lorena. (2012). Uso de curvas ROC en investigación clínica: Aspectos teórico-prácticos. *Revista chilena de infectología*, 29(2), 138-141. <https://dx.doi.org/10.4067/S0716-10182012000200003>
- Charris, L., Henriquez, C., Hernandez, S., Jimeno, L., Guillen, O., y Moreno, S. (2014). Análisis comparativo de algoritmos de árboles de decisión en el procesamiento de datos biológicos. *Revista I+D En TIC*, 9(1), 26–34. <http://revistas.unisimon.edu.co/index.php/identific>
- Checa, J., Gomà, A., López, J., Martín, J., y Nel-lo, O. (2016). Energía y proceso de urbanización en España, Nuevas perspectivas de investigación a partir de la imagen satelital nocturna de la Tierra. *Crisis, Globalización y Desequilibrios Territoriales En España. Aportación Española Al 33 Congreso de Beijing*.
- Chen, W., Xie, X., Wang, J., Pradhan, B., Hong, H., Bui, D. T., Duan, Z., y Ma, J. (2017). A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility. *Catena*, 151, 147–160. <https://doi.org/10.1016/j.catena.2016.11.032>
- Chiriboga, C., y Andrade, K. (2005). Diagnóstico socioeconómico de las poblaciones usuarias de los remanentes de bosque seco de los cantones Alamor y Zapotillo en la provincia de Loja, Ecuador. *Biodiversidad en los bosques secos de la zona de Cerro Negro-Cazaderos, occidente de la provincia de Loja: un reporte de las evaluaciones ecológicas y socioeconómicas rápidas*, 91-116.
- CNF, 2010. *Incendios Forestales, Guía Práctica para Comunicadores*. Comisión Nacional Forestal. Tercera Edición. Zapopan, Jalisco, México.
- Cuesta, J. R., y Giraldo, I. O. (2013). Efectos de los incendios forestales en las propiedades del suelo. *Ingenium*, 7(16), 11. <https://doi.org/10.21774/ing.v7i16.264>
- De La Casa, A., y Ovando, G. (2006). Relación entre la precipitación e Índices de Vegetación durante el comienzo del ciclo anual de lluvias en la Provincia de Córdoba, Argentina. *35*, 67–86. <https://www.redalyc.org/pdf/864/86435105.pdf>

- Di Bella, C.M., Posse, G., Beget, M.E., Fischer, M.A., Mari, N., Veron, S. (2008). La teledetección como herramienta para la prevención, seguimiento y evaluación de incendios e inundaciones. *Ecosistemas* 17(3):39-52.
- Díaz, E., González, M., Jiménez, J., Treviño, E., y Ávila, D. (2013). Caracterización de Combustibles Forestales Mediante un Muestreo Directo en Plantaciones Forestales. *Memorias Del Cuarto Simposio Internacional Sobre Políticas, Planificación y Economía de Los Incendios Forestales: Cambio Climático e Incendios Forestales*, 426–436. <https://www.fs.usda.gov/treearch/pubs/44565>
- Dobilas, S. (2020, November 28). MARS: Splines de regresión adaptativa multivariante: ¿cómo mejorar la regresión lineal? . <https://towardsdatascience.com/mars-multivariate-adaptive-regression-splines-how-to-improve-on-linear-regression-e1e7a63c5eae>
- ESA. (2015). Sentinel 2. Obtenido de The European Space Agency: https://www.esa.int/Space_in_Member_States/Spain/SENTINEL_2
- FAO (2021). Evaluación de los recursos forestales mundiales 2020 - Informe principal. Roma. <https://doi.org/10.4060/ca9825es>
- FAO. (2018). Global Forest Resources Assessment 2015: How are the world's forests changing? Food and Agriculture Organization of the United Nations.
- Galparsoro, L. de U., y Pita, F. S. (2010). Medidas de concordancia: el índice Kappa. *Fisterra*, 6, 1–7.
- Gascón, F., Bouzinac, C., Thépaut, O., Jung, M., Francesconi, B., Louis, J y Fernandez, V. (2017). Estado de calibración y validación de productos de Copernicus Sentinel-2A. *Teledetección*, 9(6), 584.
- Gholinejad, S., y Khesali, E. (2021). An automatic procedure for generating burn severity maps from the satellite images-derived spectral indices. *International Journal of Digital Earth*, 14(11), 1659–1673. <https://doi.org/10.1080/17538947.2021.1966525>
- Gilabert, M. A. (1997). Acerca de los índices de vegetación. May 2014.
- Gómez, I., y Martín, I. (2008). Estudio Comparativo de Índices Espectrales para la Cartografía de Áreas Quemadas con Imágenes MODIS. *Revista de Teledetección: Revista de La Asociación Española de Teledetección*, 29, 15–24.
- González Gutiérrez, I., Farfán Gutiérrez, M., Morales Manilla, L. M., Pérez Salicrup, D. R., Garza Saldaña, J. J., y Medina Puente, A. (2023). Índices y modelos para la predicción

- de la ocurrencia de incendios forestales: una revisión para México. *Revista Geográfica de América Central*, 2(71), 193–219. <https://doi.org/10.15359/rgac.71-2.7>
- González Ulibarry, P. (2017). Impacto de los incendios forestales en suelo, agua, vegetación y fauna. *Biblioteca Del Congreso Nacional De Chile*, 32(1), 3–18. <https://www.camara.cl/verDoc.aspx?prmTIPO=DOCUMENTOCOMUNICACIONC UENTAyprmID=39186>
- González, A. J. (2020). Identificación de Áreas Quemadas por medio de Índices Espectrales utilizando imágenes SENTINEL 2 en el Municipio de Orocué, Casanare. 21. <https://repository.unimilitar.edu.co/bitstream/handle/10654/38042/GonzalezAlixJohana2020.pdf?sequence=1&isAllowed=y>
- Guo, F., Wang, G., Su, Z., Liang, H., Wang, W., Lin, F., y Liu, A. (2016). What drives forest fire in Fujian, China? Evidence from logistic regression and Random Forests. *International Journal of Wildland Fire*, 25(5), 505–519. <https://doi.org/10.1071/WF15121>
- Haynes Bradstock, R. A., Gill, A. M. and Williams, R. J. B. (2012). *Flammable Australia: Fire Regimes, Biodiversity and Ecosystems in a Changing World*. CSIRO Publishing.
- Hern, E. R., Lemus, D. F., y Pineda, W. D. (2019). Una propuesta metodológica para la selección de umbrales de clasificación en procesos de seroconversión y transmisión de anticuerpos a recién nacidos en Bogotá-Colombia. 12(2), 193–210.
- Ibáñez Martín, A. (2019). *Modelos de Machine Learning: Optimización y Aplicaciones*. 16–25.
- INEC (2010). *Carta topográfica Zapotillo. Censo de población y vivienda*.
- Jain, P., Coogan, S. C. P., Subramanian, S. G., Crowley, M., Taylor, S., y Flannigan, M. D. (2020). A review of machine learning applications in wildfire science and management. *Environmental Reviews*, 28(4), 478–505. <https://doi.org/10.1139/er-2020-0019>
- Jasiewicz, J., Stepinski, T., y Niesterowicz, J. (2018). Multi-scale segmentation algorithm for pattern-based partitioning of large categorical rasters. *Computers and Geosciences*, 118, 122–130. <https://doi.org/10.1016/j.cageo.2018.06.003>
- Keck, A., Gray, E., Lynch, P., O'Carroll, C., y Voiland, A. (2011). *NPPNPOESS Preparatory Project Building a Bridge to a New Era of Earth Observations*. 15. https://www.nasa.gov/pdf/596329main_NPP_Brochure_ForWeb.pdf

- Landa, F., Cruz, N., Coria, G., Rebolledo, G., y Hernández, S. (2014). El Índice Kappa de Cohen como Herramienta para la Reducción de Ruido de Información Cerebral en la Identificación de Niveles de Valencia. Puebla.
- Llorens, R., Sobrino, J. A., Fernández, C., Fernández-Alonso, J. M., y Vega, J. A. (2021). A methodology to estimate forest fires burned areas and burn severity degrees using Sentinel-2 data. Application to the October 2017 fires in the Iberian Peninsula. *International Journal of Applied Earth Observation and Geoinformation*, 95(October 2020). <https://doi.org/10.1016/j.jag.2020.102243>
- Loján, J. (2020). Informe de variables que se utilizan para la generación de modelos de susceptibilidad de incendios con la identificación de sus fuentes para descarga.
- Lozano, P. (2002). Los tipos de bosque en el sur del Ecuador. *Bótanica Austroecuatoriana*, 2002, 29–49.
- MAE (2013). Plan de contingencia de incendios forestales.
- María, P., Sofía, S., y Puchaicela, M. (2023). *Los incendios forestales en el bosque seco del sur del Ecuador: causas y efectos*. https://www.researchgate.net/publication/375594759_Los_incendios_forestales_en_el_bosque_seco_del_sur_del_Ecuador_causas_y_efectos
- Mendoza, Z. A., Córdova, C. A., y Coronel, W. Q. (2021). Bosque seco de la parroquia Mangahurco, Zapotillo, Loja, estudio de su composición florística, estructura y endemismo. In *Ciencia Latina Revista Científica Multidisciplinar* (Vol. 5, Issue 5, pp. 7162–7182). https://doi.org/10.37811/cl_rcm.v5i5.838
- Milla Lostaunau, L. (2014). La Inteligencia Artificial. *Quipukamayoc*, 11(22), 109. <https://doi.org/10.15381/quipu.v11i22.5480>
- Moncayo, L. (2013). Aprendizaje supervisado: Algoritmos de clasificación y regresión.
- Mondal, N., y Sukumar, R. (2016). Fires in seasonally dry tropical forest: Testing the varying constraints hypothesis across a regional rainfall gradient. *PLoS ONE*, 11(7), 1–16. [doi:10.1371/journal.pone.0159691](https://doi.org/10.1371/journal.pone.0159691)
- Montealegre, A. L., Lamelas, M. T., Tanase, M. A., y De la Riva, J. (2017). Estimación de la gravedad de los incendios forestales basada en los datos LiDAR-PNOA y los valores del Composite Burn Index. *Teledetección*, 49, 1–16. <https://doi.org/10.4995/raet.2017.7371>

- Muñoz C.A., Treviño E.J., Verástegui J., Jiménez J., Aguirre O.A., (2005). Desarrollo de un modelo espacial para la evaluación del peligro de incendios forestales en la sierra madre oriental de México. *Inv. Geogr.* 56: 101-117.
- Muñoz, P. (2013). Apuntes de Teledetección: Índices de vegetación. Centro de Información de Recursos Naturales, 15. [http://bibliotecadigital.ciren.cl/bitstream/handle/123456789/26389/Tema Índices de vegetación%2C Pedro Muñoz A.pdf?sequence=1&isAllowed=y](http://bibliotecadigital.ciren.cl/bitstream/handle/123456789/26389/Tema%20Indices%20de%20vegetaci%C3%B3n%20Pedro%20Mu%C3%B1oz%20A.pdf?sequence=1&isAllowed=y)
- Nami, M.; Jaafari, A.; Fallah, M.; Nabiuni, S., (2018). Spatial prediction of wildfire probability in the Hyrcanian ecoregion using evidential belief function model and GIS. *Int. J. Environ. Sci. Technol.* 15, 373–384.
- NASA EARTHDATA. (2020). FIRMS FAQ. Recuperado de: <https://earthdata.nasa.gov/faq/firms-faq#ed-viirs-375m-product>
- Nowosad, J. (2021). Motif: an open-source R tool for pattern-based spatial analysis. *Landscape Ecology*, 36(1), 29–43. <https://doi.org/10.1007/s10980-020-01135-0>
- Nowosad, J. (2023). Spatial patterns’ clustering. Obtenido de motif: https://jakubnowosad.com/motif/articles/v5_cluster.html
- Oliveira, S., Oehler, F., San-Miguel-Ayanz, J., Camia, A., y Pereira, J. M. C. (2012). Modeling spatial patterns of fire occurrence in Mediterranean Europe using Multiple Regression and Random Forest. *Forest Ecology and Management*, 275, 117–129. <https://doi.org/10.1016/j.foreco.2012.03.003>
- Opazo Saldivia, S., y Rodríguez-Verdú, F. (2013). Variación espacial de índices espectrales sobre áreas quemadas en Sudamérica. *Cuadernos de Investigación Geográfica*, 33, 39–57. <https://doi.org/10.18172/cig.1188>
- Otterstrom, S., Schwartz, M., y Velázquez-Rocha, I. (2006). Responses to Fire in Selected Tropical Dry Forest Trees. *Biotropica*, 38(5), 592–598. Recuperado de: <http://onlinelibrary.wiley.com/doi/10.1111/j.1744-7429.2006.00188.x/full>
- Paladines, R. (2003). Propuesta de conservación del Bosque seco en el Sur de Ecuador. *Lyonia*, 4(2), 183-186.
- Paredes, D. (2020). Data Science con R. Análisis de datos y algoritmos de predicción. <https://bookdown.org/dparedesi/data-science-con-r/>
- Paucar, A., Vallejo, M., Barragán, G., y Villacis, L. (2021). Factores condicionantes y climáticos en la amenaza de incendio forestal de la microcuenca del Río Chazo Juan, Ecuador. *Journal of Science and Research*, 6, 139–165.

- Pausas, J. G. (2020). *Incendios forestales*. Los libros de la Catarata.
- Pazmiño, D. (2019). Peligro de incendios forestales asociado a factores climáticos en Ecuador. *FIGEMPA: Investigación y Desarrollo*, 7(1), 10-18.
- PDOT (2015). Actualización del Plan de Desarrollo y Ordenamiento Territorial del cantón Zapotillo. Actualización 2015 - 2019.
- PDOT (2019). Plan de Desarrollo y Ordenamiento Territorial (PDOT) 2019-2023. Zapotillo, Ecuador. <https://gobiernodezapotillo.gob.ec/wp-content/uploads/2020/05/Resumen-Zapotillo.pdf>
- Pérez, B., Serna, A., Delgado, J., Caballero, M., y Villa, G. (2020). El programa Copernicus aplicado a la producción y gestión de la información geoespacial. *Copernicus*, 2, 47.
- Pérez-Verdín, G., Márquez-Linares, M. A., Cortés-Ortiz, A., y Salmerón-Macías, M. (2013). Análisis espacio-temporal de la ocurrencia de incendios forestales en Durango, México. *Madera Bosques*, 19(2), 37–58. <https://doi.org/10.21829/myb.2013.192339>
- Pernas, J. B., García, M. L., y Martín, J. P. (2011). Estudio de la Incidencia de los incendios en Galicia: una perspectiva socioeconómica. *Galega de Economía*, 20, 1–20
- Pfeiffer, S. (2021). *Contextualizing algorithms: Artificial intelligence, machine learning, and employee competencies*. 61–75.
- Pinta, M., Silva, S., y Puchaicela, M. (2023). *Los incendios forestales en el bosque seco del sur del Ecuador: causas y efectos*. https://www.researchgate.net/publication/375594759_Los_incendios_forestales_en_el_bosque_seco_del_sur_del_Ecuador_causas_y_efectos
- Plana, E., Font, M., y Serra, M. (2016). Los incendios forestales, guía para comunicadores y periodistas. Proyecto eFIRECOM. Ediciones CTFC. 32pp. Obtenido de http://efirecom.ctfc.cat/docs/efirecomperiodistes_es.pdf
- Ramírez Flores, Vania Alejandra Aranda Delgado, L., y Rico Gray, V. (2018). La resiliencia del bosque seco tropical: un seguro de vida para su conservación. *CONABIO, Biodiversitas*, 137: 12–16.
- Reyes-Bueno, F. (2017). Manual de Prácticas de Ordenación Territorial.
- Reyes-Bueno, F. (2021). Manual de aplicación susceptibilidad a incendios forestales. Caso: Cantón Loja.

- Reyes-Bueno, F., y Balcázar-Gallegos, C. (2021). Factores que inciden en la probabilidad de ocurrencia de incendios forestales en Ecuador. *FIGEMPA: Investigación y Desarrollo*, 11(1), 50–60. <https://doi.org/10.29166/revfig.v11i1.2634>
- Reyes-Bueno, F., y Córdova, J. L. (2021). Documento explicativo de los modelos de susceptibilidad a incendios forestales.
- Reyes-Bueno, F., y Loján-Córdova, J. (2022). Assessment of Three Machine Learning Techniques with Open-Access Geographic Data for Forest Fire Susceptibility Monitoring—Evidence from Southern Ecuador. *Forests*, 13(3), 474. <https://doi.org/10.3390/f13030474>
- Rojas, N., Soto, C., Tuesta, S., Silva, J., y Sáenz, G. (2023). Probabilidad de ocurrencia de incendios forestales en Apurímac. *Hatun Yachay Wasi* 2(2), 85–97. <https://doi.org/10.57107/hyw.v2i2.49>
- Ruiz Gallardo, J. R. (2004). Teledetección y SIG en la asistencia de la actuación forestal postincendio. <http://www.tdx.cat/handle/10803/48941>
- SADMQ, (2013). Los incendios forestales en el Distrito Metropolitano de Quito evaluación 2012.
- Samaniego, N., Eguiguren, P., Maita, J., y Aguirre, N. (2015). Diversidad de briófitos de los Páramos de Cajanuma del Parque Nacional Podocarpus. In *Cambio climático y Biodiversidad: Estudio de caso de los páramos del Parque Nacional Podocarpus, Ecuador*. <https://doi.org/10.13140/RG.2.1.5009.1921>
- Sancho, F. (2020). Aprendizaje Supervisado y No Supervisado. <http://www.cs.us.es/~fsancho/?e=77>
- Sandoval, J. L. (2018). Algoritmos de aprendizaje automático para análisis y predicción de datos. *Revista Tecnológica ITCA-FEPADE*, 11, 36–40.
- Sari, F. (2021). Forest fire susceptibility mapping via multi-criteria decision analysis techniques for Mugla, Turkey: A comparative analysis of VIKOR and TOPSIS. *Forest Ecology and Management*, 480(2020), 118644. <https://doi.org/10.1016/j.foreco.2020.118644>
- Sari, F. (2024). Assessment of the effects of different variable weights on wildfire susceptibility. *European Journal of Forest Research*, 0123456789. <https://doi.org/10.1007/s10342-023-01643-z>
- Schroeder, W., y Giglio, L. (2017). Visible Infrared Imaging Radiometer Suite (VIIRS) 750 m Active Fire Detection and Characterization Algorithm Theoretical Basis Document. 0–21.

- Sedeño, B., Armando, E., y Quintana, M. (2022). Aprendizaje Automático como herramienta para el manejo integrado de recursos naturales en un contexto de cambio climático Machine Learning as Tool for the Integrated Management of Natural Resources in a Climate Change Context. *Revista Cubana de Transformación Digital*, 5, 1–10.
- SNGRE. (2021). Informe de Situación No. 23 de Incendios Forestales a nivel Nacional. Servicio Nacional de Gestión de Riesgos y Emergencias (SNGREb), 10, 1–4. https://www.gestionderiesgos.gob.ec/wp-content/uploads/2021/09/SITREP-No-10-Incendios_Forestales-21092021.pdf
- Solis, A. F., Revelo luna, D. A., Campo Ceballos, D. A., y Gaviria López, C. A. (2021). Correlación del contenido de clorofila foliar de la especie Coffea arabica con índices espectrales en imágenes. *Biotecnología En El Sector Agropecuario y Agroindustrial*, 19(2), 51–68. <https://doi.org/10.18684/bsaa.v19.n2.2021.1536>
- Tien Bui, Dieu; Bui, Quang-Thanh; Nguyen, Quoc-Phi; Pradhan, Biswajeet; Nampak, Haleh; Trinh, Phan Trong (2017). A hybrid artificial intelligence approach using GIS-based neural-fuzzy inference system and particle swarm optimization for forest fire susceptibility modeling at a tropical area. *Agricultural and Forest Meteorology*, 233(), 32–44. doi:10.1016/j.agrformet.2016.11.002
- Tortajada, P. E. (2023). Inteligencia Artificial. *Revista de Derecho Civil*, 10(2), 1–2. <https://doi.org/10.29057/xikua.v3i5.1271>
- Usmadi, D. (2023). Maximum entropy application in predicting the vulnerability of land and forest fires in South Sumatra Province, Indonesia. *IOP Conference Series: Earth and Environmental Science*, 1183(1). <https://doi.org/10.1088/1755-1315/1183/1/012105>
- Valdez, M. C., Chang, K. T., Chen, C. F., Chiang, S. H., y Santos, J. L. (2017). Modelling the spatial variability of wildfire susceptibility in Honduras using remote sensing and geographical information systems. *Geomatics, Natural Hazards and Risk*, 8(2), 876–892. <https://doi.org/10.1080/19475705.2016.1278404>
- Vanegas, J., y Vásquez, F. (2017). Multivariate Adaptive Regression Splines (MARS), una alternativa para el análisis de series de tiempo. *Gaceta Sanitaria*, 31(3), 235–237. <https://doi.org/10.1016/j.gaceta.2016.10.003>
- Vargas-Sanabria, D., y Campos-Vargas, C. (2018). Modelo de vulnerabilidad ante incendios forestales para el Área de Conservación Guanacaste, Costa Rica. *UNED Research Journal*, 10(2), 435–446. <https://doi.org/10.22458/urj.v10i2.2173>

- Verdú, F., y Salas, J. (2011). Caracterización De Variables Biofísicas En Los Incendios Forestales Mayores De 25 Ha. 79–100.
- Vilar del Hoyo, L. (2008). Empleo de regresión logística para la obtención de modelos de riesgo humano de incendios forestales. *Boletín de La A.G.E.*, 47, 5–29.
- Vilar del Hoyo, L., Isabel, P. M., y Vega, J. M. (2002). Empleo de técnicas de regresión logística para la obtención de modelos de riesgo humano de incendio forestal a escala regional. *Pacific Health Dialog: A Publication of the Pacific Basin Officers Training Program and the Fiji School of Medicine*, 9(1), 5–29.
- Vilar, L., Gómez, I., Martínez-Vega, J., Echavarría, P., Riaño, D., y Martín, M. P. (2016). Multitemporal modelling of socio-economic wildfire drivers in central Spain between the 1980s and the 2000s: Comparing generalized linear models to machine learning algorithms. *PLoS ONE*, 11(8), 1–17. <https://doi.org/10.1371/journal.pone.0161344>
- Villegas Cubas, J. E., Capuñay Uceda O. E., Maquen Niño G. L. E., Fuentes Adrianzén D. J., y Bustamante Quintana, P. H. (2022). Sistema de detección de phishing basado en machine learning. *Biblioteca Colloquium*. Recuperado a partir de <https://www.colloquiumbiblioteca.com/index.php/web/article/view/122>
- Villers Ruíz, M. D. L. (2006). Incendios forestales. *Ciencias*, (081).
- Voltas, X. R. (2021). Creación de un modelo de predicción de riesgos de incendios forestales usando una red neuronal convolucional sobre datos históricos de meteorología de California. *Ciencias de Datos*.
- Yilmaz, I. (2009). Comparison of landslide susceptibility mapping methodologies for Koyulhisar, Turkey: conditional probability, logistic regression, artificial neural networks, and support vector machine. *Environmental Earth Sciences*, 61(4), 821–836. doi:10.1007/s12665-009-0394-9
- Zhang, X., Lan, M., Ming, J., Zhu, J., y Lo, S. (2023). Spatiotemporal Heterogeneity of Forest Fire Occurrence Based on Remote Sensing Data: An Analysis in Anhui, China. *Remote Sensing*, 15(3). <https://doi.org/10.3390/rs15030598>

11. Anexos

Anexo 1. Base de datos depurada

	ACQ_DATE	ACQ_TIME	Incendio	Sent Indic	ARVI	BSI	GNDVI	Moisture	MSI	NBR	NDMI	NDVI	NDWI	SAVI
1	2019-11-18	1900	1	2019-11-18	-0,354	0,19	0,409	-0,119	1,356	0,026	-0,151	0,291	-0,409	0,437
2	2019-11-19	1842	1	2019-11-18	-0,313	0,155	0,416	-0,077	1,250	0,085	-0,111	0,3	-0,416	0,449
3	2019-11-19	1842	1	2019-11-18	-0,21	-0,028	0,637	0,107	0,884	0,315	0,062	0,656	-0,637	0,984
4	2019-11-27	1754	1	2019-11-18	-0,36	0,182	0,45	-0,08	1,377	0,038	-0,159	0,334	-0,45	0,501
5	2019-12-05	1842	1	2019-12-03	-0,393	0,232	0,275	-0,262	1,669	-0,132	-0,251	0,2	-0,275	0,3
6	2019-12-09	1906	1	2019-12-03	-0,417	0,235	0,355	-0,195	1,673	-0,063	-0,252	0,289	-0,355	0,434
7	2019-12-09	1906	1	2019-12-03	-0,409	0,246	0,366	-0,188	1,631	-0,084	-0,24	0,227	-0,366	0,34
8	2019-12-09	1906	1	2019-12-03	-0,4	0,247	0,293	-0,193	1,682	-0,117	-0,254	0,174	-0,293	0,261
9	2019-12-09	1906	1	2019-12-03	-0,358	0,182	0,359	-0,129	1,460	0,011	-0,187	0,302	-0,359	0,453
10	2019-12-13	1754	1	2019-12-03	-0,379	0,225	0,328	-0,168	1,553	-0,091	-0,217	0,207	-0,328	0,311
11	2019-12-13	1754	1	2019-12-03	-0,408	0,238	0,329	-0,17	1,695	-0,109	-0,258	0,234	-0,329	0,35
12	2019-12-13	1930	1	2019-12-03	-0,387	0,242	0,348	-0,132	1,522	-0,048	-0,207	0,187	-0,348	0,281
13	2019-12-13	1930	1	2019-12-03	-0,249	0,125	0,159	-0,141	1,387	-0,051	-0,162	0,14	-0,159	0,209
14	2019-12-14	1912	1	2019-12-03	-0,282	0,157	0,204	-0,124	1,398	-0,009	-0,166	0,143	-0,204	0,214
15	2019-12-14	1912	1	2019-12-03	-0,336	0,125	0,457	-0,09	1,276	0,043	-0,121	0,477	-0,457	0,715
16	2019-12-14	1912	1	2019-12-03	-0,334	0,15	0,395	-0,067	1,364	0,044	-0,154	0,352	-0,395	0,529
17	2019-12-15	1854	1	2019-12-03	-0,41	0,258	0,298	-0,216	1,697	-0,142	-0,259	0,17	-0,298	0,254
18	2019-12-27	1830	1	2019-12-23	-0,445	0,281	0,386	-0,184	1,658	-0,076	-0,248	0,242	-0,386	0,363
19	2020-01-16	1854	1	2020-01-12	-0,384	0,206	0,442	-0,128	1,449	0,012	-0,183	0,325	-0,442	0,487
20	2020-01-16	1854	1	2020-01-12	-0,376	0,22	0,356	-0,134	1,475	-0,017	-0,192	0,236	-0,356	0,354
21	2020-05-22	1912	1	2020-05-21	-0,27	0,129	0,356	0,01	1,152	0,123	-0,071	0,276	-0,356	0,414
22	2020-06-04	1830	1	2020-05-31	0,065	-0,264	0,622	0,365	0,485	0,561	0,347	0,727	-0,622	1,090
23	2020-10-25	1848	1	2020-10-23	-0,409	0,24	0,365	-0,168	1,650	-0,093	-0,245	0,242	-0,365	0,362
24	2020-10-27	1812	1	2020-10-23	-0,343	0,206	0,205	-0,232	1,558	-0,112	-0,218	0,14	-0,205	0,211
25	2020-10-27	1812	1	2020-10-23	-0,423	0,275	0,272	-0,213	1,804	-0,196	-0,287	0,128	-0,272	0,191
26	2020-10-27	1812	1	2020-10-23	-0,403	0,277	0,272	-0,238	1,678	-0,191	-0,283	0,124	-0,272	0,184

access_cen	access_pob	access_AP	access_via	access_ZA	dist_CE	dist_rios	dist_vias	dist_ZA	dist_AP	orientacio	elevacion	pendiente
29,229	18,515	20,199	46,651	7,063	10,935,118	338,364	109,992	0	7,215,842	105,611	424,600	15,771
32,873	23,675	23,842	50,295	9,308	10,466,654	254,937	127,273	85,437	7,098,010	16,049	370,600	8,638
39,295	30,097	30,264	56,717	7,141	10,827,525	0	302,312	106,296	7,405,812	48,762	341,047	1,327
31,913	20,884	32,044	41,242	1,273	12,819,908	475,060	222,020	0	7,784,783	129,440	527,800	17,803
2,986	2,118	6,524	17,440	2,646	1,832,077	10,000	22,359	0	2,997,535	355,834	583,198	2,007
23,624	22,896	8,730	21,241	0	8,710,676	393,928	250,193	0	139,274	72,646	460,800	11,374
17,016	4,813	6,456	5,530	1,149	8,680,155	9,999	80,618	0	819,976	118,610	425,400	5,724
15,512	7,762	10,214	6,108	2,349	9,810,716	366,175	107,696	0	197,980	86,878	529,000	6,287
9,797	9,733	9,928	6,693	2,100	7,321,618	136,007	42,424	0	0	62,282	546,800	7,593
25,888	9,379	21,722	35,217	0,424	12,367,864	361,229	31,622	0	8,188,216	192,665	461,000	20,046
14,339	6,030	14,470	23,668	0,276	15,554,988	49,997	31,621	9,999	10,572,322	356,424	465,000	14,386
15,541	6,792	9,244	6,137	2,430	9,144,021	0	79,998	0	615,836	44,554	519,074	1,327
12,559	6,627	12,551	26,587	1,060	8,458,284	197,223	14,141	0	4,372,605	318,180	687,000	5,823
5,976	4,682	9,514	19,989	2,800	2,348,140	269,980	59,996	0	2,179,870	246,250	592,200	3,127
31,815	16,981	20,586	41,144	6,108	7,281,748	352,258	121,646	0	2,751,001	192,456	684,800	27,022
35,045	20,211	25,551	44,374	5,091	7,622,432	72,108	136,009	39,997	3,177,018	342,050	559,400	16,622
29,785	19,619	33,052	47,207	8,494	10,706,116	225,604	221,344	0	8,867,640	243,812	425,200	15,212
7,193	6,629	15,566	25,190	0,838	3,651,156	44,720	9,999	0	3,372,243	66,037	267,600	2,256
42,944	38,690	43,075	52,273	28,446	6,865,392	53,850	591,311	763,768	3,187,117	90,379	238,800	16,805
50,495	49,226	50,626	59,824	44,969	6,351,103	86,018	1,071,931	932,941	2,611,854	33,311	194,600	4,788
5,815	5,653	10,028	19,226	2,076	2,029,338	169,697	78,098	0	4,859,867	54,229	245,000	1,010
6,103	3,599	14,494	24,101	0,724	2,911,445	262,474	166,426	0	5,408,063	247,511	240,904	0,108
25,688	14,659	25,819	35,017	6,042	13,332,504	369,988	94,334	59,996	8,830,722	90,466	414,000	13,821
7,643	7,579	7,774	16,972	4,958	8,193,867	304,122	19,999	1,688,767	1,613,783	320,856	359,200	6,328
20,364	18,527	20,495	29,693	17,543	8,252,198	84,849	277,838	1,989,543	2,127,910	193,799	305,000	13,212
11,174	11,110	11,205	20,502	8,618	8,227,380	263,670	84,840	1,677,550	1,838,432	246,161	227,000	6,374

Anexo 2. Certificado de traducción del abstract



Mg. Yanina Quizhpe Espinoza
Licenciada en Ciencias de Educación mención Inglés
Magister en Traducción y mediación cultural

Celular: 0989805087
Email: yaniques@icloud.com
Loja, Ecuador 110104

Loja, 19 de junio de 2024

Yo, Lic. Yanina Quizhpe Espinoza, con cédula de identidad 1104337553, docente del Instituto de Idiomas de la Universidad Nacional de Loja, y con master en Traducción, con registro 724187576 en la Senescyt, certifico:

Que tengo el conocimiento y dominio de los idiomas español e inglés y que la traducción del resumen del Trabajo de Integración Curricular, **Predicción de zonas susceptibles a incendios forestales aplicando técnicas de aprendizaje automático en el cantón Zapotillo, Loja – Ecuador**, cuya autoría del estudiante Deiby Emersson Granda García, con cédula 1900798370, estudiante de la Carrera de Ingeniería Ambiental, perteneciente a la Facultad Agropecuaria y de Recursos Naturales Renovables la Universidad Nacional de Loja, es verdadero y correcto a mi mejor saber y entender.

Atentamente

YANINA
BELEN
QUIZHPE
ESPINOZA
Firmado digitalmente por
YANINA BELEN
QUIZHPE
ESPINOZA
Fecha: 2024.06.19
12:11:41 -0500

Mg. Yanina Quizhpe Espinoza.

Traductora freelance

Full text translator: servicios de traducción