

ESCUELA POLITÉCNICA NACIONAL

FACULTAD DE INGENIERÍA CIVIL Y AMBIENTAL

Patrones temporales y espaciales de incendios forestales en la zona norte y sur de la región Andina del Ecuador en el periodo 2001-2018

TRABAJO DE TITULACIÓN PREVIO A LA OBTENCIÓN DEL TÍTULO DE INGENIERÍA AMBIENTAL

AUTOR: NAVARRETE NAVARRETE, ELIZABETH ABIGAIL

abigailnavarrete016@gmail.com

DIRECTOR: Ph.D ZAPATA RÍOS XAVIER EDUARDO

xavier.zapata@epn.edu.ec

Quito, enero, 2021

DECLARACIÓN

Yo, Elizabeth Abigail Navarrete Navarrete, declaro bajo juramento que el trabajo aquí descrito es de mi autoría; que no ha sido previamente presentada para ningún grado o calificación profesional; y, que he consultado las referencias bibliográficas que se incluyen en este documento.

La Escuela Politécnica Nacional puede hacer uso de los derechos correspondientes a este trabajo, según lo establecido por la Ley de Propiedad Intelectual, por su Reglamento y por la normatividad institucional vigente.

**ELIZABETH ABIGAIL NAVARRETE
NAVARRETE**

CERTIFICACIÓN

Como director del trabajo de titulación “Patrones temporales y espaciales de incendios forestales en la zona norte y sur de la región Andina del Ecuador en el periodo 2001-2018” por Elizabeth Abigail Navarrete Navarrete estudiantes de la carrera de Ingeniería Ambiental, habiendo supervisado la realización de este trabajo y realizado las correcciones correspondientes, doy por aprobada la redacción final del documento escrito para que prosiga con los trámites correspondientes a la sustentación de la Defensa oral.

Ph.D. XAVIER EDUARDO ZAPATA RÍOS

AUSPICIO

El presente trabajo de titulación contó con el auspicio financiero del proyecto PIJ-17-05 “Los patrones climáticos globales y su influencia en la respuesta temporal y espacial de índices espectrales de la vegetación del páramo en el Ecuador”, que se ejecutó en el Departamento de Ingeniería Civil y Ambiental.

AGRADECIMIENTO

En primer lugar, quiero agradecer a Dios por ayudarme a lo largo de mi vida hasta este momento. Quiero agradecer a mi madre, Nancy, por ser la fortaleza en mi vida, por acompañarme en las noches de estudio y deberes, por siempre enviarme comida para la hora que me diera hambre, por su amor y cuidado, y por dar todo de ella para su familia. Así también, agradezco a mis hermanos, Gaby y Jorge, por siempre apoyarme para continuar con mis estudios. A mis amigos Paola y Alexis, por ser amables y estar conmigo en los momentos más duros.

Mi gratitud total con el PhD Xavier Zapata Ríos, por ser una inspiración para todos sus alumnos por su calidad como maestro y también para todos aquellos quienes trabajamos con él por su calidad como profesional y jefe. Le agradezco por haberme dado la oportunidad de ser parte de este proyecto y ayudarnos a crecer profesionalmente a todos quienes estuvimos bajo su dirección.

También quiero extender mi agradecimiento a Carmen Elena López, quién fue mi jefa directa dentro del proyecto. Le agradezco por siempre tener la disposición de ayudarme dentro de todo este proceso, incluso a la distancia siempre estuvo pendiente de mis dudas. En mi mente quedará plasmada su amistad y amabilidad para ayudar a todos en la oficina.

Espero que en algún momento pueda llegar a ser tan buen profesional y empático con los que me rodean como cada uno de ustedes lo ha demostrado.

También quiero agradecer a la Escuela Politécnica Nacional por el financiamiento del proyecto PIJ-17-05 “Los patrones climáticos globales y su influencia en la respuesta temporal y espacial de índices espectrales de la vegetación del páramo en el Ecuador” Este trabajo de titulación se desarrolló en el marco de este proyecto. Además, agradezco por el estipendio brindado durante los meses enero de 2019 a marzo de 2020 cuando fui asistente de investigación del mismo.

DEDICATORIA

Le dedico este trabajo a Dios, por todo lo que él ha significado en mi vida. A mi madre y mejor amiga, por ser la mujer más valiente, generosa y fuerte, por siempre dar preferencia a sus hijos antes que a ella mismo, te dedico esto y espero que la vida me permita darte lo que te mereces y más. También, les dedico a mis hermanos, pues sin ellos no estaría aquí. A mis sobrinos, Micaela y Martín, que son la luz de mi vida. A mi mejor amiga Paola, por su apoyo incondicional. A mi mascota Batman, a quién amo con todo mi corazón.

También quiero dedicar este trabajo, a todos mis amigos de la universidad, Andrés, Winyfred, Pauli, Pao, Jeff, sin ustedes la universidad no hubiese sido tan divertida. A mi mejor amigo, quién me ha acompañado en todo este camino, a David, tu apoyo y ánimo en todo este proceso sin duda me ayudó a llegar hasta este momento. Y también quiero dedicarle esto a María Inés Rivadeneira, mi mejor amiga en la universidad, por las madrugadas de estudio y por tu alegría que mejoraba el día de todos, gracias por tanto Maru.

ÍNDICE

DECLARACIÓN	I
CERTIFICACIÓN.....	II
AUSPICIO	III
AGRADECIMIENTO.....	IV
DEDICATORIA.....	V
ÍNDICE DE TABLAS	VIII
ÍNDICE DE FIGURAS.....	IX
ÍNDICE DE ANEXOS.....	XI
RESUMEN.....	XII
ABSTRACT	XIII
CAPITULO 1 INTRODUCCIÓN.....	1
1.1. <i>Antecedentes</i>	1
1.2. <i>Justificación</i>	1
1.2.1. Justificación Teórica	1
1.2.2. Justificación Metodología.....	2
1.2.1. Justificación Práctica	3
1.3. <i>Objetivos</i>	3
1.3.1. Objetivo general	3
1.3.2. Objetivos específico.....	4
1.4. <i>Hipótesis</i>	4
1.5. <i>Alcance</i>	4
1.6. <i>Estudios Previos</i>	4
CAPITULO 2 MARCO TEÓRICO Y ESTADO DEL ARTE.....	6
2.1. <i>Marco Teórico</i>	6
2.1.1. Servicios ecosistémicos.....	6
2.1.2. Incendios forestales.....	6
2.1.3. Información Satelital usada el Estudio	13
2.1.4. Métodos estadísticos	16
2.1.5. Interpolación de la probabilidad	24
2.2. <i>Estado del Arte</i>	27
2.2.1. Incendios en Ecuador	27
2.2.2. Información de incendios Disponibles en el país	27

2.2.3. Variables asociadas a los incendios	28
CAPITULO 3 ÁREA DE ESTUDIO	30
3.1. Zona Norte	31
3.2. Zona Sur	32
CAPITULO 4 METODOLOGÍA	34
4.1. Esquema de trabajo	34
4.2. Recopilación de datos	35
4.3. Procesamiento de datos	36
4.3.1. Procesamiento de ráster incendios	38
4.4. Obtención de punto de incendios	39
4.5. Análisis de datos	40
4.5.1. Código de las variables	41
4.6. Análisis Descriptivo	42
4.7. Análisis Estadístico	42
4.7.1. Prueba de Hipótesis	42
4.7.2. Variables Significativas	43
4.8. Modelo lineal generalizado	43
4.9. Validación del modelo	43
4.10. Interpolación de probabilidad	44
4.11. Mapeo	46
CAPITULO 5 RESULTADOS Y DISCUSIÓN	47
5.1. Incendios Totales	47
5.2. Análisis descriptivo	48
5.3. Análisis Estadístico	53
5.4. Variables que se usaron en el modelo	58
5.5. Resultados del modelo logístico	59
5.5.1. Zona Norte	59
5.5.2. Zona Sur	60
5.6. Resultados del modelo	61
5.7. Resultados de la interpolación	63
5.8. Mapa de probabilidad de ocurrencia de incendios	64
5.9. Discusión	68
CAPITULO 6 CONCLUSIONES Y RECOMENDACIONES	69
6.1. Conclusiones	69
6.2. Recomendaciones	71
CAPITULO 7 BIBLIOGRAFÍA	73

ÍNDICE DE TABLAS

Tabla 1. Categorías de VCI	15
Tabla 2. Resultados de la matriz de confusión	22
Tabla 3. Resumen de los productos descargados	35
Tabla 4. Características de los productos satelitales	36
Tabla 5. Significado y tipo de variable	41
Tabla 6. Resultados de la prueba paramétrica para la variable aspecto de la zona norte	54
Tabla 7. Variables resultantes del modelo logit en la zona norte.	60
Tabla 8. Variables resultantes del modelo logit en la zona sur.	61
Tabla 9. Resultados del modelo	62
Tabla 10. Resultados del kriging	64
Tabla 11. Resumen de la probabilidad de incendio, zona norte.	66
Tabla 12. Resumen de la probabilidad de incendio, zona sur.	68

ÍNDICE DE FIGURAS

Figura 1. Factores que rigen la producción y comportamiento de los incendios forestales.	8
Figura 2. Componentes de un sistema de teledetección	10
Figura 3. Espectro electromagnético	12
Figura 4. Señales espectrales típicas para distintas cubiertas	12
Figura 5. Semivariograma	26
Figura 6. Componentes del semivariograma	27
Figura 7. Ubicación de las áreas de estudio	30
Figura 8. Variables climáticas en la zona norte	31
Figura 9. Uso de suelo, áreas protegidas y variables de infraestructura en la zona norte	32
Figura 10. Variables climáticas en la zona sur	33
Figura 11. Uso de suelo, áreas protegidas y variables de infraestructura en la zona sur	33
Figura 12. Esquema de trabajo	35
Figura 13. Procesamiento de los archivos hdf de enero 2001 en el programa mrt	37
Figura 14. Variables climáticas en la zona sur	37
Figura 15. Cálculo del condicional para convertir a valores binomiales	38
Figura 16. Áreas sin condicional y con condicional para septiembre de 2012	39
Figura 17. Diagrama del proceso para obtención de puntos de incendio	40
Figura 18. Histograma de la herramienta geostatistical analyst	44
Figura 19. Ventana para interpolación kriging	45
Figura 20. Tipos de kriging y salida de la interpolación	45
Figura 21. Puntos de incendios para cada zona	47
Figura 22. Distribución de incendios por provincias. Temporada enero-diciembre desde el 2001-2018.	48
Figura 23. Análisis temporal del comportamiento de incendios.	49
Figura 24. Análisis de la relación del uso de suelo nivel 2 con los incendios. Donde las barras muestran el total de incendios para cada zona.	50
Figura 25. Relación entre variables topográficas e incendios.	51
Figura 26. Relación entre variables de infraestructura e incendios.	52

Figura 27. Relación entre variables climáticas-biológicas e incendios.	53
Figura 28. Resultados de la prueba no paramétrica para las variables distancia a centros poblados norte.	54
Figura 29. Resultados de las pruebas MDA (a) y MDG (b) para el norte.	55
Figura 30. Resultados de las pruebas MDA (a) y MDG (b) para el sur.	56
Figura 31. Resultados de la prueba de multicolinealidad para el norte.	57
Figura 32. Resultados de la prueba de multicolinealidad para el sur.	58
Figura 33. Curva ROC	62
Figura 34. Histograma para transformación de distribución.	63
Figura 35. Mapa de probabilidad para la zona norte	65
Figura 36. Mapa de probabilidad para la zona sur	67

ÍNDICE DE ANEXOS

ANEXO 1 Pruebas no paramétricas para la zona norte	83
ANEXO 2 Pruebas no paramétricas para la zona sur	85
ANEXO 3 Pruebas paramétricas para la zona norte y sur	87

RESUMEN

Los incendios forestales causan un impacto importante sobre la dinámica de los procesos normales del ambiente, principalmente proporcionados por la vegetación. La restauración de los espacios forestales afectados, demanda un alto costo monetario, por lo que es importante estudiar estos eventos y así obtener información que ayude a las autoridades a establecer sistemas efectivos de planificación, prevención, control y combate para minimizar o evitar los impactos de un incendio. Actualmente, los incendios afectan el ecosistema ecuatoriano, especialmente en el norte y sur de su región interandina. La zona norte incluye las provincias de Carchi, Imbabura, y Pichincha, mientras que la zona sur comprende las provincias de Azuay, Loja y una porción de la provincia del El Oro.

En este sentido, el presente estudio describió los incendios en ambas zonas utilizando el producto de área quemada MODIS-MCD64A1, sistemas de información geográfica, e información geográfica de libre acceso. De este modo, las relaciones entre la ocurrencia de los incendios, y variables climáticas, biológicas, topográficas, así como de infraestructura fueron evaluadas para encontrar patrones espacio-temporales en cada zona. Las variables predictivas se escogieron en función de los criterios de bosque aleatorio (random forest) y multicolinealidad. Una vez determinadas las variables significativas, se utilizó el modelo lineal generalizado logit para establecer la probabilidad de ocurrencia de incendios y se realizaron mapas para cada zona con la información del modelo.

En los resultados se puede destacar que la zona norte tuvo un periodo de incendios que va de junio a octubre, mientras el periodo para la zona sur fue entre agosto y diciembre. De acuerdo a esto, las variables Uso de Suelo, radiación (Irradiación horizontal global), y precipitación tuvieron mayor efecto al describir la presencia de incendios. Similarmente, la actividad humana también se vio cumpliendo un rol imprescindible en ambas zonas mediante la distancia a centros poblados. Además, las condiciones de la vegetación (NDVI y VCI) también tuvieron un peso importante en la descripción de los incendios. Por otro lado, la temperatura no describió a la variable dependiente en ninguna zona.

ABSTRACT

Wildfires have an impact on the dynamics of environmental services, mainly provided by vegetation. The restoration of the forest area affected demands a high monetary cost, so it is important to study these events and thus obtain the information that helps the authorities to establish effective planning, prevention, and control and combat systems to minimize or avoid the impacts of fire. Fires are currently affecting Ecuadorian environment, especially in the north and the south of the interandean region. The northern zone included the provinces of Carchi, Imbabura, and Pichincha, while the southern zone included the provinces of Azuay, Loja and a portion of the province of El Oro.

In that sense, the current study described the fires in both zones by using the burned area product MODIS-MCD64A1, geographic information systems, and free access geographic information. Thereby, the relationships among fire production, and weather, topographic, biological and infrastructure variables were assessed to find spatial-temporal patterns in each area. The predictive variables were chosen based on the criteria of random forest and multicollinearity. Once the significant variables were determined, the probability of fire occurrence was established by using the general linear model logit, and maps were made for each area with this information from the model.

In the results section, it may be noted that the northern area has a fire period that goes from June to October, while the period of the southern zone was from August to December. According to it, land use and land cover variable had the greatest effect on describing the presence or absence of wildfires. Similarly, human activity was also playing an essential role in both areas through the distance to urban areas. In addition, the Vegetation condition (NDVI and VCI) also had an important weight in the description of fires. By contrast, the temperature data did not describe the dependent variable in any zone.

CAPITULO 1 INTRODUCCIÓN

1.1.ANTECEDENTES

Este trabajo surge del proyecto de investigación PIJ-17-05: Los patrones climáticos globales y su influencia en su respuesta temporal y espacial de índices espectrales de la vegetación del páramo en el Ecuador; del caso de estudio “Spatiotemporal patterns of burned areas, fire drivers, and fire probability across the equatorial Andes” realizado por la Ing. Carmen Elena López y dirigido por el PhD. Xavier Zapata Ríos. Como resultado de esta investigación se encontraron dos zonas donde se concentraban la mayor parte de los incendios en la región Andina. A partir de esto, se decide estudiar por separado los patrones espaciales y temporales de estas dos áreas. Es necesario señalar que las imágenes MCD64A1 fueron validadas en la investigación mencionada.

1.2.JUSTIFICACIÓN

1.2.1. JUSTIFICACIÓN TEÓRICA

Aunque es de conocimiento público el peligro y los graves efectos que causa un incendio, la información disponible y actualizada sobre modelos que predicen este fenómeno es muy escasa en el Hemisferio Sur y por lo tanto en el Ecuador (Altamirano et al., 2013). Durante la época de sequía los boques semiáridos, nublados y páramos, situados especialmente en la zona Andina, son los más vulnerables a la ocurrencia de un incendio (Ministerio del Ambiente Ecuador, 2015). Los incendios forestales han requerido la atención de varias instituciones y de la colaboración interprovincial, pues en los últimos años estos eventos se han intensificado, incluso han cruzado las fronteras provinciales (Secretaría de Gestión de Riesgos, 2018). La detección de incendios tiene el objetivo de descubrir y localizar incendios forestales, siendo esta la primera acción oportuna para la extinguir o contrarrestar los focos de fuego (Flández, 2011). Según afirman la Corporación Nacional Forestal de Chile (2019) y Montiel et al. (2009), existen diferentes sistemas de detección de incendios, y se clasifican de la siguiente manera:

1. Vigilancia terrestre fija: este sistema puede usar vigías o cámaras termográficas.
2. Vigilancia terrestre móvil: desplazamiento de vigilantes dentro de un área, el cual puede ser mediante un vehículo o a pie.
3. Vigilancia aérea: se puede usar aeronaves tripuladas o un vehículo aéreo no tripulado.
4. Redes de sensores: son diferentes tipos de sensores conectados por medio de una red inalámbrica que envía los datos obtenidos a una consola.
5. Teledetección: se usa imágenes satelitales.

1.2.2. JUSTIFICACIÓN METODOLOGÍA

Los Sistemas de Información Geográfica (SIG) pueden ser aprovechados para realizar monitoreos, mapas, base de datos, reportes, etc. Esto constituye a los SIGs como una herramienta importante dentro de la gestión ambiental, pues además de su utilidad, el procesamiento de los datos tienen un bajo costo y existe una gran disponibilidad temporal de la información en la web (Kapetsky & Aguilar, 2009). Debido a que la recopilación de información en campo sobre incendios es impredecible, costosa y muy riesgosa, se utilizará información satelital (Sabuco, 2013). La información satelital seleccionada para los incendios es el producto MODIS MCD64A1-Área quemada, versión 6 y resolución 500m. Este producto se obtiene por la combinación de los datos de los satélites de la NASA, Aqua y Terra (L Giglio et al., 2015). La información de las imágenes se obtiene mediante la reflectancia de la superficie y observaciones de fuego (L Giglio et al., 2015). La información satelital también se usará para determinar las variables climáticas, topográficas y biológicas. Mientras que los datos para las variables socio-económicas tales como uso de suelo, distancia a vías, distancia a centros poblados se obtendrán de las entidades del estado.

Una vez que los datos estén corregidos (según la guía de cada producto) y procesados se realiza un análisis estadístico para desarrollar el modelo de predicción, en función de las variables más significativas. Métodos similares han sido utilizados en distintos países, donde este fenómeno es recurrente como:

España por Martínez et al., (2004). Chile por Altamirano et al. en (2013). Sur de Europa por Vilar del Hoyo et al. (2011). Irán por Mohammadi et al. (2014).

La probabilidad de la ocurrencia de incendios se estimó mediante un análisis de regresión logística (Chang et al., 2013; Garcia et al., 1995; Guo et al., 2016). Este tipo de análisis se ha utilizado por distintos estudios sobre estos eventos. Además, este análisis se enmarca dentro de los modelos líneas generalizados para variables dependientes dicotómicas, y que por lo tanto no tienen distribución normal (Dobson & Barnett, 2018; McCullagh & Nelder, 2019). Para la variable dicotómica se establece el valor de 1 para los lugares donde ocurrieron los incendios y 0 para los lugares donde no se produjeron incendios.

1.2.1. JUSTIFICACIÓN PRÁCTICA

La detección oportuna y confiable de nuevos estallidos fue particularmente crucial para el manejo efectivo de incendios forestales (Sifakis et al., 2011). La coordinadora del programa Amazonía sin fuego-zona norte, Andrea Bustos (como se citó en Rosales, 2019) afirmó que cerca 99% de los incendios son provocados por acción humana. Además, según describe Ecuador Forestal (2013) a través de su página web, en el 2013 el Distrito Metropolitano de Quito calculó que 51 millones de dólares fueron invertidos en la restauración de los espacios forestales afectados por los incendios. Debido a las consecuencias económicas y ambientales que conllevan estos eventos, se vuelve clave gestionar las acciones de manejo para la toma de futuras decisiones (Altamirano et al., 2013). Por ello es importante la identificación de los patrones de incendios (Altamirano et al., 2013).

1.3.OBJETIVOS

1.3.1. OBJETIVO GENERAL

Identificar los patrones espacio-temporales que determinan la producción de incendios en la zona norte y sur de la región Andina del Ecuador, mediante el uso de imágenes satelitales y métodos estadísticos, para establecer un modelo de probabilidad de ocurrencia para cada zona.

1.3.2. OBJETIVOS ESPECÍFICO

- Procesar los datos cuantitativos y cualitativos, con la ayuda de las herramientas del programa ArcGIS, para generar la base de datos con la cual se trabajará.
- Determinar las variables significativas de cada zona, usando pruebas no paramétricas y de distribución, además de los criterios de árboles aleatorios y de multicolinealidad, para la elaboración del modelo.
- Elaborar un modelo de probabilidad de ocurrencia de incendio, a través de un modelo lineal generalizado, para obtener información con la cual se realizará un mapa de probabilidad para cada zona.

1.4. HIPÓTESIS

La producción de un incendio difiere entre la zona norte y sur de la región Andina del Ecuador, debido a que cada zona tiene sus propias variables que rigen el comportamiento de este fenómeno.

1.5. ALCANCE

El alcance de este estudio es investigar cuales son los patrones espacio-temporales de los incendios forestales en dos zonas de la región interandina, que incluyen 6 provincias, dentro de un periodo de 18 años. Además, para poder entender estos patrones se analizaron variables climáticas, topográficas, biológicas y de infraestructura. A partir de esto, entregar un mapa para cada zona que permita la visualización de la probabilidad de ocurrencia de un incendio forestal.

1.6. ESTUDIOS PREVIOS

El producto principal dentro de esta investigación es el MCD64A1-Área Quemada (Burned Area) del instrumento MODIS de la NASA. La Escuela Politécnica Nacional, tiene un estudio previo del producto, en una versión anterior a la mencionada. Este estudio lo realizó Esteban Ospina como proyecto de titulación para obtener su título como Ingeniero Ambiental. Este estudio se realizó en el año del 2017 y consistió en el mapeo de áreas quemadas en todo el Ecuador con el producto MCD45A1 versión 5. Ospina (2017) presenta la distribución espacio

temporal de los incendios pero, no construyó un modelo explicativo de los mismos. Además, se puede mencionar que el producto usado en esta investigación (versión 6) tiene mejoras en relación a la versión 5 utilizada por Ospina. La Universidad de Maryland (2017) menciona que el nuevo algoritmo de la versión 6 es extremadamente tolerante a las nubes y a la contaminación por aerosoles, esto significa que el MCD64A1 tiene una mejor detección para los incendios pequeños y tiene una reducción de la incertidumbre temporal de la fecha de incendio en comparación con la versión 5.

CAPITULO 2 MARCO TEÓRICO Y ESTADO DEL ARTE

2.1.MARCO TEÓRICO

2.1.1. SERVICIOS ECOSISTÉMICOS

Los servicios ecosistémicos también conocidos como servicios ambientales hacen referencia a la capacidad de los ecosistemas, naturales o intervenidos, para generar beneficios que satisfagan las necesidades humanas. Estos beneficios se originan a partir de la conservación, mantenimiento y regeneración de las funciones ecológicas, y ciclos de la naturaleza (De Groot, 1992; Ministerio del Ambiente Ecuador, 2019). En el Ecuador se reconoce cuatro tipos de servicios ecosistémicos, estos son mencionados en el Art. 240 del Reglamento al Código Orgánico Ambiental del Ministerio del Ambiente Ecuador (2019), los cuales se mencionan a continuación:

- a) Servicios de aprovisionamiento: son bienes tangibles que pueden pertenecer a los recursos renovables como a los no renovables. Ejemplo de estos son: alimentos, madera, agua, entre otros.
- b) Servicios de regulación: son derivados del funcionamiento de los procesos ecosistémicos y son vitales para la supervivencia humana. Por ejemplo: calidad del aire, fertilidad de suelos, control de inundaciones, etc.
- c) Servicios culturales: beneficios no materiales. Tales como: reflexión, recreación, enriquecimiento espiritual, etc.
- d) Servicios de soporte: procesos básicos que aseguran el buen funcionamiento del resto de servicios ecosistémicos. Entre los cuales se encuentran: fotosíntesis, formación del suelo, ciclos biogeoquímicos, entre otros.

Los incendios interactúan con varios componentes en los ecosistemas, por lo que todos estos servicios se encuentran comprometidos ante estos eventos.

2.1.2. INCENDIOS FORESTALES

En varias regiones del mundo, los incendios forestales constituyen un proceso natural, pues estos ayudan a mantener la estructura y composición característica de los ecosistemas de estas regiones (Hirschberger, 2017). Es así que cerca del 46% de las ecorregiones dependen o tienen influencia del fuego en sus procesos

naturales (Hirschberger, 2017). Por lo tanto, cuando este régimen es alterado afecta a los procesos dinámicos de los ecosistemas.

Anualmente de 3 a 5 millones de km² de superficie terrestre se incendian en todo el mundo (Jones et al., 2019). Además, Hirschberger (2017) argumenta que de todos los incendios producidos en el mundo, solamente el 4% son por causas naturales, y el resto de los eventos son por responsabilidad humana directa o indirecta. Se puede evidenciar que el ecosistema ha sido perturbado por actividades humanas cuando el régimen de fuego ha cambiado y se producen incendios en el lugar equivocado, en un momento inusual o con mayor frecuencia. Entonces, una vez que el fuego ha sobrepasado la capacidad de resistencia del ecosistema, trae consigo graves consecuencias ecológicas, sociales y económicas (Food and Agriculture Organization, 2009; Hirschberger, 2017). Por ejemplo, existen impactos sobre la calidad del agua y del aire, también en la cantidad de agua, en los ciclos biogeoquímicos, en la destrucción de biodiversidad, degradación del suelo, pérdidas de vidas humanas, y además constituye una fuente importante de contaminación que contribuye al calentamiento global (Cruz & Rainer, 2012; Hirschberger, 2017; Rowell & Moore, 2000). van Der Werf et al. (2017) sostienen que entre el periodo de 1997-2016, un promedio de 2.2 Pg C (1Pg C=1x10¹⁵g de carbono) al año fueron emitidos, globalmente, a la atmósfera debido a incendios forestales, incendios de deforestación y quemadas agrícolas.

El aumento de temperatura debido al calentamiento global hace que los bosques estén más secos, por lo cual la vulnerabilidad a incendios crece, especialmente en las zonas que poseen una gran área forestal (Food and Agriculture Organization, 2009; Hirschberger, 2017). Además, el cambio climático también genera condiciones inusuales en el ambiente, lo que ocasiona que las consecuencias, ya provocadas por el mismo, se magnifiquen (Escobar, 2017). Debido a esto, un incendio tiene mayor probabilidad, no solo de producirse, sino también de propagarse y/o mantenerse activo.

Existen factores que rigen la producción y comportamiento de los incendios forestales, entre ellos tenemos variables climáticas, topográficas, de uso de suelo y actividad humana (Altamirano et al., 2013). Tal como se muestra en la *Figura 1*.

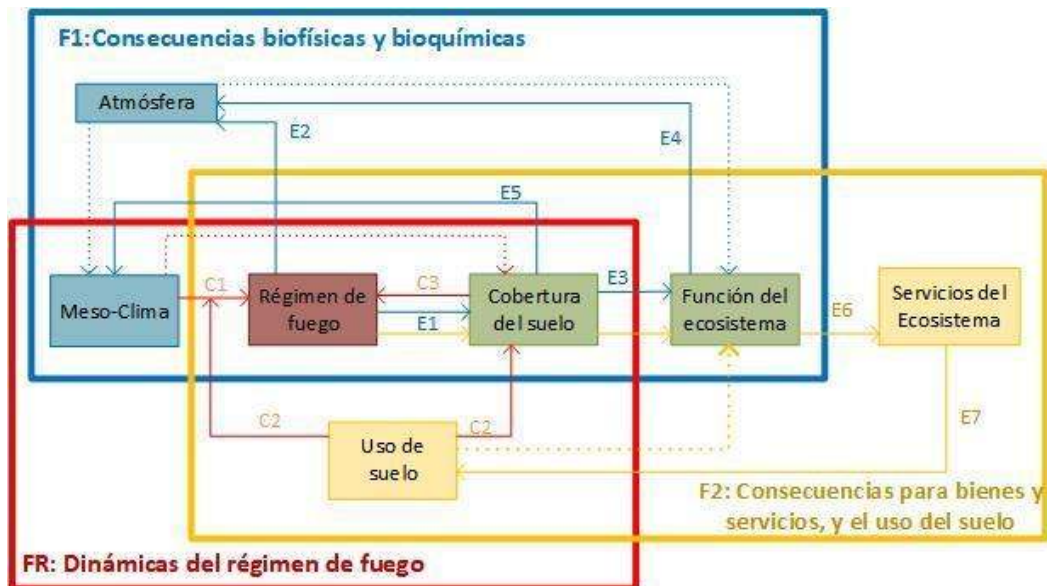


Figura 1. Factores que rigen la producción y comportamiento de los incendios forestales. Esta figura resume las relaciones entre los diferentes factores del sistema humano-ambiente implicados en las causas y efectos de los eventos de fuego. Tiene 3 ciclos, el primero es el ciclo de control, llamado régimen de fuego (FR, cuadro y flechas rojo), los dos ciclos restantes es de retroalimentación, el primero impulsa las consecuencias del fuego para los procesos biofísicos y bioquímicas (F1, cuadro y flechas azul) y el último ciclo expresa las consecuencias del fuego para servicios de los ecosistemas y el uso de la tierra. Las flechas completas indican temas de interés directo para la investigación de Lavorel, Flannigan, Lambin, & Scholes, mientras que las flechas punteadas representan efectos importantes pero indirectos. Donde E es Efecto y C es Causa

Fuente: (Lavorel, Flannigan, Lambin, & Scholes, 2007)

Modificado: Abigail Navarrete

- **Variables Climáticas:** estas variables son de gran importancia debido a que acondicionan el ecosistema para que un incendio se produzca (Swetnam, 1993). Así mismo, el clima influye en la disponibilidad del combustible, en la inflamabilidad, propagación y muchas veces en la ignición (Lavorel et al., 2007). Estas variables pueden ser temperatura, precipitación, humedad relativa, radiación y velocidad del viento, etc.
- **Variables Topográficas:** la variabilidad de la superficie contribuye en la propagación del fuego e influye en la agilidad de una pronta respuesta y control ante un incendio (Stambaugh & Guyette, 2008).

- Variables de uso de suelo: el uso de suelo influye, junto con el clima, en la propagación del fuego. Además, cada tipo de cobertura tiene diferentes puntos de inflamabilidad, esto dependerá de la composición, la edad y la densidad de la especie vegetal, y otro factor que interviene es la abundancia de vegetación inflamable (Lavorel et al., 2007; Turner & Romme, 1994).
- Variables de actividad humana: el fuego es una herramienta importante para actividades como la caza de animales, tala de bosques, mantenimiento de pastizales, control de pestes y mantenimiento de cultivos (Eva & Lambin, 2000; Hirschberger, 2017).

2.1.2.1. Teledetección

Pérez & Muñoz (2002) definen a la Teledetección como el procedimiento que permite adquirir información, a pesar de no tener un contacto con esta y además su objeto es la captura, tratamiento y análisis de imágenes tomadas por satélites artificiales. La información presentada por los satélites a través de sus observaciones de la Tierra, se ha convertido en una herramienta beneficiosa para el estudio y seguimiento de los procesos ambientales (Chuvieco, 2008). Además, el funcionamiento de esta técnica se basa en la interacción energética, ya sea por reflexión del sol o por un haz energético artificial, que se da entre el suelo y el sensor (*Figura 2*). Asimismo, un sistema de teledetección comprende los siguientes elementos, sugeridos en el libro de Chuvieco (2008) (*Figura 2*):

- Fuente de Energía: se refiere a la radiación electro-magnética detectada por el sensor. Este tipo de energía puede provenir del sol (foco externo-teledetección pasiva) u originarse de un haz energético (fuente artificial-teledetección activa) (Chuvieco, 2008).
- Cubierta terrestre: describe cada objeto de la superficie terrestre tal como vegetación, suelos, construcciones, entre otras, los cuales reaccionan de distintas formas ante el flujo de energía, esto debido a las características físicas de cada objeto (Chuvieco, 2008)(Chuvieco, 2008).
- Sistema sensor: está formado por un sensor y un soporte que sirve como plataforma para sostener el sensor. Este equipo es el que capta la energía electro-magnética procedente de los objetos en la superficie, la codifica, graba y envía la información a receptores terrestres.

- Sistema de recepción-comercialización: recibe la información que el sensor envió, luego graba datos en un formato apropiado, si es necesario realiza correcciones y finalmente se distribuye la información (Chuvienco, 2008).
- Intérprete: los datos son tratados visual o digitalmente mediante programas informáticos para obtener cartografía temática (Chuvienco, 2008).
- Usuario final: en este paso, las personas pueden revisar y adquiere la información proporcionada por los sensores (Chuvienco, 2008).

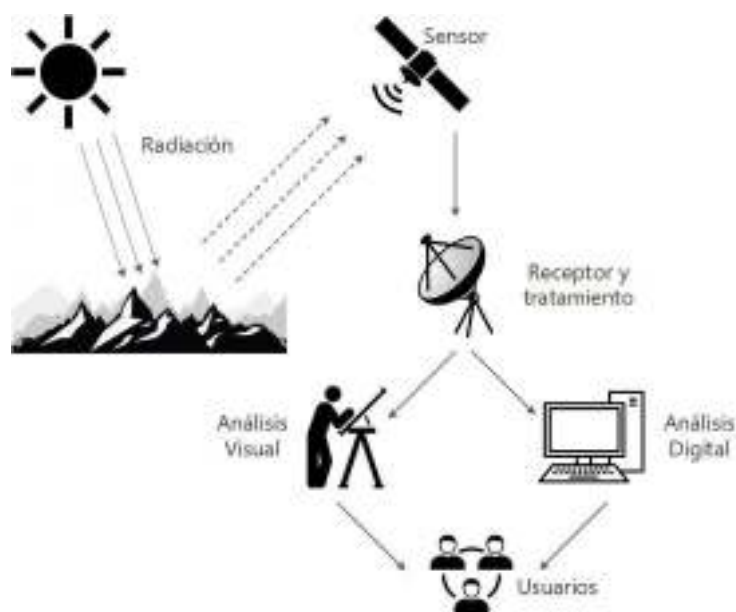


Figura 2. Componentes de un Sistema de Teledetección

Fuente: (Chuvienco, 2008)

Modificado: Abigail Navarrete

Como se mencionó anteriormente la energía es un elemento importante en la teledetección porque constituye la base del sistema. Una de las formas de la energía electromagnética es el flujo energético que se mueve en la atmósfera terrestre y su comportamiento se puede estudiar a través de una serie de bandas de longitudes de onda, que se denominan espectro electro-magnético (Chuvienco, 2008; Pérez & Muñoz, 2002), tal como se puede observar en la *Figura 3*. De acuerdo a lo que sugiere Chuvienco (2008), el espectro electro-magnético, comprende las longitudes desde:

- Espectro Visible: tal como dice su nombre es la única radiación visible a los ojos humanos, este se extiende desde los 0.4 a 0.7 μ m, este rango se divide

en las siguientes bandas, Azul (0.4-0.5 μm), Verde (0.5-0.6 μm) y Rojo (0.6-0.7 μm) (Chuvienco, 2008).

- Infrarrojo cercano (IRC): describe el rango de longitud de onda más corta del espectro infrarrojo, esta entre 0.7 y 1.3 μm y además la luz de esta región no es térmica. Estas características son las que permiten distinguir masas vegetales y concentraciones de humedad (Chuvienco, 2008).
- Infrarrojo medio (IRM): en esta longitud de onda, la reflexión de la luz solar y las emisiones que vienen desde la superficie terrestre, se mezclan, por lo tanto se distinguen 2 bandas (Chuvienco, 2008). La primera es la infrarrojo de onda corta la, cual se sitúa entre 1.3 a 2.5 μm ; la segunda banda es apropiada para la detección de focos de alta temperatura, como los que se producen en un incendio o en volcanes activos, se la conoce como infrarrojo medio y está comprendida entre los 3 y 5 μm (Chuvienco, 2008).
- Infrarrojo lejano (IRT): explica la energía térmica (calor), con esta banda se puede estudiar el calor que es emitido por la mayoría de las superficies sobre el planeta y se ubica entre los 8 a 14 μm (Chuvienco, 2008).
- Micro-ondas: esta banda está por sobre el 1mm, se caracteriza por ser energía muy transparente. Se usa para las telecomunicaciones como la radio y televisión (Chuvienco, 2008).

En la *Figura 4* se puede visualizar el comportamiento del espectro electromagnético en relación a las coberturas terrestres. Además, es necesario mencionar que la captura de la información se ve limitada por sustancias (ozono, vapor de agua y dióxido de carbono) presentes en la atmósfera, el efecto absorbente que estas producen, evita y obstruye el paso de determinadas longitudes de onda dentro del espectro (Pérez & Muñoz, 2002).

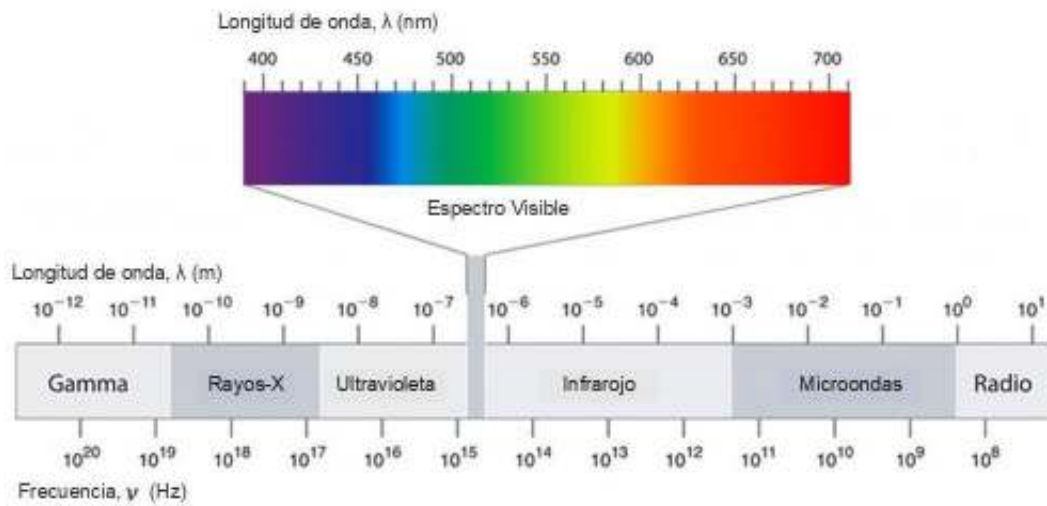


Figura 3. Espectro Electromagnético

Fuente: (Pérez & Muñoz, 2002)

Modificado: Abigail Navarrete

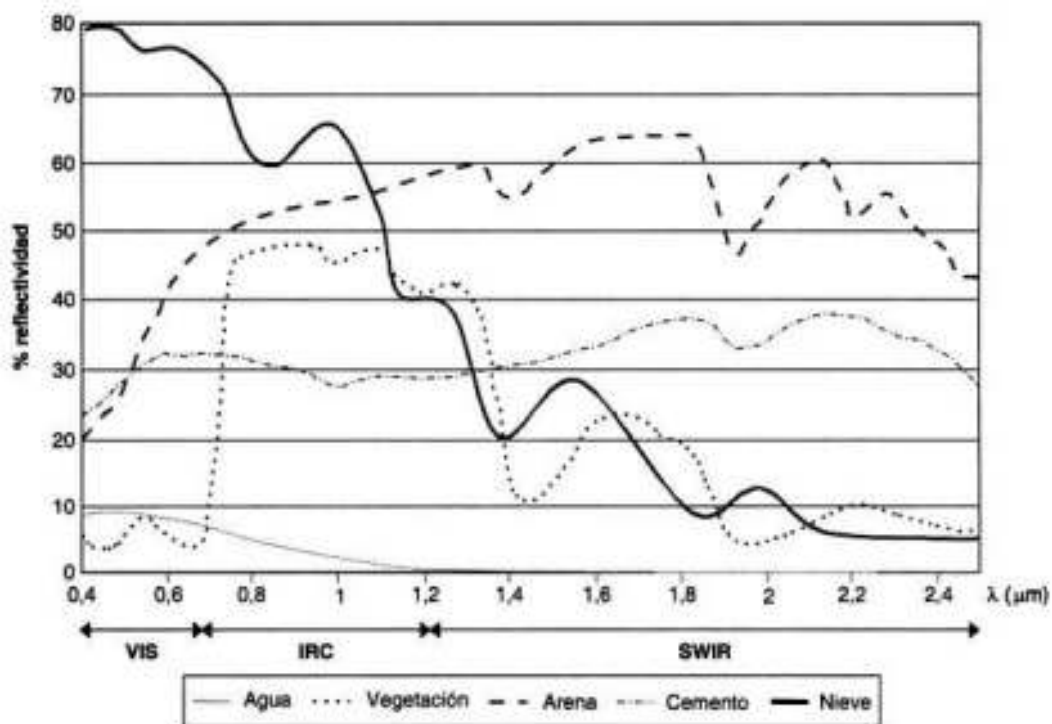


Figura 4. Señales espectrales típicas para distintas cubiertas

Fuente: (Chuvienco, 2008)

Pérez & Muñoz (2002) plantean que, para evaluar los rangos de precisión de estos sistemas, se usa cinco resoluciones distintas, y son las siguientes:

- Espacial: hace referencia al objeto más pequeño y distinguible dentro de la imagen, al cual se lo llama píxel. Además, el tamaño del píxel dependerá de

la altura del sensor, ángulo de visión, velocidad al escanear y características ópticas (Pérez & Muñoz, 2002).

- Espectral: es el ancho y número de bandas que puede discriminar o registrar un sensor. Se dice que la resolución es mayor cuando las bandas son más estrechas (Pérez & Muñoz, 2002).
- Temporal: se refiere a la capacidad que tiene un sensor para poder registrar cambios, esto tiene que ver con el tiempo que tarda el satélite en pasar sobre la vertical de un punto. Este tipo de resolución se ve condicionada por el tipo de órbita, el ángulo de barrido y la velocidad del satélite (Pérez & Muñoz, 2002).
- Radiométrica: se refiere a la cantidad de niveles de color que discrimina el sensor. La radiación recibida por el sensor se transforma a formato binario para ser codificada; a mayor número de bits mayor resolución (Pérez & Muñoz, 2002).

La teledetección es una de las fuentes de información global y multitemporal. Además, los sensores ópticos electrónicos permiten acceder a lecturas en las bandas del infrarrojo medio, térmico y las microondas. Esta información es valiosa para estudios ambientales (Chuvienco, 2008). Sin embargo, también tiene limitaciones como por ejemplo no puede detectar información de profundidad de agua y suelo, asimismo las nubes reducen la visibilidad del área observada. Algunas de las aplicaciones en las que usa esta información son:

- Estudios de escorrentía y erosión
- Estudios de aguas superficiales y humedales
- Estudios para control de riesgos
- Estudios sobre impactos ambientales
- Estudios de uso de suelo
- Cálculos de índices de vegetación
- Entre otros

2.1.3. INFORMACIÓN SATELITAL USADA EL ESTUDIO

Los productos descargados tienen un periodo de 18 años que comprenden desde el 2001 hasta 2018.

2.1.3.1. Datos satelitales

- Área Quemada (MCD63A1): en un estudio comparativo, (Moreno et al., 2013) afirman que este producto obtuvo los mejores resultados generales de área quemada. El MCD63A1 combina datos de los satélites Terra-Aqua del instrumento MODIS, es un producto mensual de área quemada con cuadrícula de 500m. El algoritmo usa un índice de vegetación (IV) sensible a las quemaduras, este IV usa reflectancia superficial infrarroja de onda corta y el registro es la fecha de incendio en la celda (L Giglio et al., 2015).
- Índice de vegetación de diferencia normalizada (MOD13A3) o NDVI: con este producto se puede estimar distintos factores, como por ejemplo la cantidad, calidad y desarrollo de la vegetación (Borowik et al., 2013). Este producto usa la energía que se emite desde la banda infrarroja cercana y roja visible, que va desde 0.6 a 1.3 μm (Borowik et al., 2013). Además, puede tomar valores entre el rango de -1 a 1, donde los valores entre -1 a 0.19 corresponden a agua, áreas áridas y carreteras, valores de 0.2-0.5 representan arbustos y pastizales, y desde 0.501 a 1.0 pertenece a bosques templado y tropicales (Hashim, Abd, & Adnan, 2019). El MOD13A3 son imágenes del satélite Terra del instrumento MODIS, que mensualmente entrega un producto con una resolución de 1km, el NDVI se calcula por dispersión molecular, absorción de ozono y aerosoles (Didan, 2015). También, a partir de la información de NDVI se generó la base de datos del índice de condición vegetal.
- Índice de condición vegetal (VCI): este índice se usa para determinar el estrés de la vegetación que ocurre durante una sequía (Kogan, 1997). Mataveli et al. (2018) postulan que el VCI es un indicador de la susceptibilidad de la vegetación a la ignición. Además, este índice se expresa mediante porcentaje donde 0% indica condición de sequía y 100% condiciones óptimas de vegetación (CAZALAC, 2018). Este índice compara los valores actuales de NDVI con sus valores mínimos y máximos registrados en los años anteriores, y se calculó de la siguiente forma, Ecuación 1:

$$VCI = \frac{NDVI - NDVI_{Min}}{NDVI_{Max} - NDVI_{Min}} \times 100 \quad (1)$$

Vallejo-Villalta, Rodríguez-Navas, & Márquez-Pérez (2019) proponen la siguiente clasificación de los valores de VCI:

Tabla 1.
Categorías de VCI

Rango de VCI	Nivel de Sequía
0 - 20	Extrema
20 - 40	Severa
40 - 60	Moderada
60 - 80	Ligero
80 - 100	Muy ligero

- Temperatura (MOD11B3): son imágenes del satélite Terra del instrumento MODIS, que se determinan mediante el algoritmo LST a través de observaciones de pares día/noche. Estos productos se entregan mensualmente con una resolución espacial de 6km y la unidad de temperatura es Kelvin (Wan, Hook, & Hulley, 2015). La capa que se usó para la base de datos fue diurno mensual.
- Precipitación: estos datos fueron recuperados del Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS). Las estimaciones de lluvia se hacen a partir de pluviómetros y observaciones satelitales (Climate Hazards Center, 2020). Además, ofrece una resolución de 0.05° que equivale aproximadamente a 5.55km.
- Irradiación global horizontal (GHI): este tipo de radiación es el incidente total de la radiación solar en una superficie horizontal (Brownson, 2014). SOLARGIS es una empresa que proporciona datos sobre energía solar y meteorológica. El producto entregado de GHI es un mapa del promedio del periodo 1999-2018, con unidades de kWh/m² (SOLARGIS, 2019).
- Modelo de elevación digital (DEM): el DEM fue recuperado de las imágenes STRM con una resolución espacial de 30m. Como se mencionó antes, la variabilidad topografía está relacionada con la propagación terrestre de los incendios (Oliveira, Oehler, San-Miguel-Ayanz, Camia, & Pereira, 2012). Por ello, se calculó la pendiente y el aspecto a partir del DEM con las herramientas de Análisis Espacial de ArcGIS

2.1.3.2. Datos de fuentes nacionales:

- Cobertura de uso de suelo: los datos de la distribución del paisaje vegetal y su uso se obtuvo del Ministerio de Agricultura, Ganadería, Acuacultura y Pesca (MAGAP), esta información está disponible a través del Sistema Nacional de Información (<https://sni.gob.ec/mapa-cobertura-uso>).
- Carreteras, áreas urbanas y cuerpos de agua: la mayor parte de la producción de incendios es debido a factores humanos, por ello este tipo de información es fundamental para evaluar el modelo. Maingi & Henry (2007) señalan que encuentran vínculos entre la ubicación de incendios y la proximidad de carreteras y asentamientos. La información se obtuvo del portal del Sistema Nacional de Información.

2.1.4. MÉTODOS ESTADÍSTICOS

2.1.4.1. Variables

- Variable Independiente: conocida como variable estímulo, causa, entrada, input. Este tipo de variable no es afectada por el experimento o por otras variables que intervengan en el estudio, pero si influye en el resultado de la variable dependiente. Además, se la representa con “x” (Achaerandio, 2001; Buendía, Colás, & Hernández, 2001).
- Variable dependiente: también llamada variable respuesta o salida. Es el factor que se observa o se estudia para determinar el efecto de la variable x, sus valores van a depender de los valores de la variable causa y se representa con “y” (Buendía et al., 2001).

Las variables también pueden clasificarse de la siguiente manera:

- Variable Cualitativa: son datos que no tienen significado aritmético y representan cualidades. (Salazar & Del Castillo, 2018).
- Variable Cuantitativa: son datos con significado aritmético, por lo tanto se expresa mediante un número (Salazar & Del Castillo, 2018).
- Variable Categórica: también llamada discreta, son variables que tienen una valorización numérica. Sin embargo, tales valores no tienen un orden jerárquico sino que simbolizan una clase o categoría (Achaerandio, 2001).

Estas variables pueden ser “dicotómicas” (dos valores) o “politémica” (tres o más valores).

2.1.4.2. Análisis de variables significativas

La selección de variables es muy importante, especialmente cuando el conjunto de datos es muy grande. Ciertas características, que son adicionales o que se repiten, hacen que el modelo sea más complejo y por lo tanto sus resultados pueden ser deficientes, a esto también se denomina alta varianza (overfitting) (Frost, 2019). Debido a que el modelo se va entrenando aleatoriamente, se necesita escoger variables que sean significativas para la predicción y así evitar que el modelo aprende del ruido provocado por las variables adicionales (Dietterich, 1995; Frost, 2019). Los métodos presentados son los que se utilizaron en este estudio para escoger las variables predictoras.

- **Prueba de Hipótesis:** este tipo de prueba permite probar si los resultados de un experimento son significativos, es decir que estos resultados no se hayan obtenido por casualidad (Good, 2005; Privitera, 2014). Se necesita proponer dos hipótesis, la primera es la hipótesis nula que por lo general es la que no tiene efecto o diferencia, y la segunda es la hipótesis alternativa, esta será la que se desea probar y por lo tanto debe ser contraria a la hipótesis nula (Good, 2005). Como resultado de la significancia estadística se usa el valor p. Entonces, para refutar la hipótesis nula, se debe obtener un valor p menor a 0.05 (Privitera, 2014).
- **Prueba No Paramétrica:** esta prueba es específica para variables cuantitativas. Se usó la estimación de la Densidad de Kernel, que compara dos grupos a través de una prueba de hipótesis de igualdad de densidad (Müller, 2006). La ecuación general es:

$$\hat{f}(x) = \frac{1}{nh} \sum_{a=1}^n K\left(\frac{x - x_i}{h}\right) \quad (2)$$

Donde x es el punto que se trata de estimar; x_i es el valor de la variable y puede ir desde $a=1, \dots$, hasta n número de variables; K es el símbolo de Kernel; y el ancho de la ventana o del parámetro suavizado se representa con la letra h (Jiménez Moreno, 1991).

- **Prueba Paramétrica:** esta prueba es para variables cualitativas-categorías. Se usó una prueba de proporciones, así como dice su nombre se usan proporciones o porcentajes para poder describir a la variable, estas proporciones se comparan mediante un contraste de hipótesis, entonces si el valor p resuelve la certeza de la hipótesis nula con los mismos valores de p, planteados en la prueba de hipótesis (Laguna, 2014).
- **Bosques Aleatorios:** es un tipo de clasificación supervisada que tiene la misma estructura que un árbol, trabaja con un vector aleatorio de la muestra lo que permite que las colecciones de árboles no están correlacionados entre sí (Medina & Ñique, 2017) . La función *importance*, del paquete randomForest de R, extrae la medida de importancia de las variables predictoras y una medida de la estructura interna de los datos (Liaw & Wiener, 2002). La importancia de una variable se estima a través de cuánto aumente el error de predicción cuando los datos que están fuera de la bolsa (out-of-bag) de una variable, se permutan mientras el resto de datos no se modifican (Breiman, Culter, Liaw, & Wiener, 2018; Liaw & Wiener, 2002). Mide 2 parámetros, el primero es la Disminución Media de la Precisión, cuanto más disminuye la precisión del bosque por la exclusión de una variable más importantes es la misma; el segundo parámetro es la Disminución Media de la Impureza del nodo o MDG (por sus siglas en inglés), medida por el índice de Gini, cuanto mayor sea el valor de MDG más pura será la variables (Breiman et al., 2018; Han, Guo, & Yu, 2016).
- **Análisis de Multicolinealidad:** la multicolinealidad ocurre cuando las variables independientes están correlacionadas. La dirección y fuerza de una relación lineal, está dada por el coeficiente de correlación de Pearson y sus valores pueden variar de -1 a +1 (Frost, 2019). Donde los extremos (-1, +1) representa una relación fuerte perfectamente lineal, el coeficiente igual a cero indica que no existe ninguna relación lineal (Frost, 2019).

2.1.4.3. Modelo lineal generalizado

El modelo lineal generalizado o también conocido como GLM (por sus siglas en inglés) es una extensión de los modelos lineales utilizado cuando la variable dependiente tiene una distribución diferente a la normal (Rodríguez, 2007). Al no

existir una relación lineal entre la información de la variable dicotómica-salida y las variables de entrada (Dobson & Barnett, 2018). Debido a que no existe tal relación lineal entre nuestras variables (independientes-dependiente), no se puede usar una regresión lineal como tal. Fox (2015) menciona 3 componentes del GLM:

- Componente aleatorio: este componente es aquel que detalla la distribución condicional de la variable salida, es miembro de la familia exponencial como Gaussiana, Gaussiana inversa, Binomial, de Poisson o Gamma (Fox, 2015).
- Un predictor lineal: es una función lineal de regresores (Fox, 2015).
- Una función de enlace: transforma la expectativa de la variable salida, en el predictor lineal (Fox, 2015).

López-González & Ruiz-Soler (2011) explican los dos últimos componentes con las siguientes ecuaciones.

$$Y_i = \sum_j \beta_j X_{ij} + \varepsilon \quad (3)$$

$$\mu_i = E(Y_i) \quad (4)$$

$$\eta_i = \sum_j \beta_j X_{ij} \quad (5)$$

$$g(\mu_i) = \eta_i \quad (6)$$

Donde Y es la variable respuesta, X representa un conjunto de variables independientes, ε es el componente aleatorio, β_j es el vector de parámetros, μ valores ajustados, η es el predictor lineal, la ecuación 6 representa la función de enlace. El valor esperado μ será mayor o igual a cero, en tanto que el predictor lineal η puede ser igual a cualquier valor positivo o negativo. Esto quiere decir que, tanto μ como η tienen diferentes escalas de medida. Entonces, la función de enlace ayuda a transformar el valor de μ a la misma escala de η , esto es igual a la ecuación 5 (López-González & Ruiz-Soler, 2011).

La naturaleza de la variable de salida es fundamental porque en función de esto se escoge el tipo de GLM (López-González & Ruiz-Soler, 2011). En esta investigación se trabajó con una variable dependiente dicotómica, es por ello que se utilizó la

regresión logística. Esta regresión usa una función de enlace llamada Logit (López-González & Ruiz-Soler, 2011). Y esto se expresa de la siguiente manera (Guo et al., 2016, 2017):

$$\ln\left(\frac{P}{1-P}\right) = \beta_0 + \sum_{j=1}^p \beta_j X_j \quad (7)$$

Donde β_0 coeficiente independiente, β_j coeficiente de la variable independiente, X_j variable independiente. Y puede tomar los valores de 1 (Incendio) y 0 (No incendio). Para la ecuación (7) se asumió que la probabilidad de ocurrencia de incendio está representada con la letra P cuando $Y = 1$ y la probabilidad de no ocurrencia de incendio será $(1 - P)$ cuando $Y = 0$ (Guo et al., 2016, 2017)

$$P = \frac{1}{(1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_m X_m)})} \quad (8)$$

Para la ecuación (8) m es el número de covariables, X_1, X_2, \dots, X_m son aquellas las variables predictivas que influyen en la ocurrencia de incendios, y los coeficientes de correlación de los predictores están representados por $\beta_1, \beta_2, \dots, \beta_m$ (Guo et al., 2016, 2017).

2.1.4.4. Validación del modelo lineal generalizado

Una vez que se hayan elegido las variables predictoras y además se haya corrido el modelo de predicción, se procede a evaluar la capacidad del modelo, esto nos permitirá saber si el modelo y las variables independientes ofrecen datos satisfactorios de salida. Para lo mencionado anteriormente se pueden usar los siguientes métodos de validación:

- **Test de Wald:** esta prueba determina si el coeficiente de una variable predictora o explicativa es significativo, es decir si está aportando al modelo de probabilidad. Esta prueba se da para $Y = 1$ con el siguiente contraste de hipótesis: $H_0: \beta_k = \beta_{k0}$ y $H_1: \beta_k \neq \beta_{k0}$ (Beltrán & Ponte, 2015). Hauck & Donner (1977) proponen la siguiente ecuación de la prueba de Wald:

$$W = \frac{(b_j - \beta_{k0})^2}{H_{kk}} \approx (X_1)^2 \quad (9)$$

La evaluación de la máxima verosimilitud de los valores β_k , está representado con b_j , β_{k0} son los valores de los parámetros propuestos, H_{kk} es la varianza estimada de b_j , y $(X_1)^2$ es una distribución asintótica igual a $(-2 \ln \lambda)$ con un grado de libertad. Además, si se tiene un p-valor < 0.05 la hipótesis nula H_0 no es aceptada, esto indicaría que la variable es significativa, caso contrario la variable puede ser separada del modelo sin alterar el mismo (Beltrán & Ponte, 2015; Hauck & Donner, 1977).

- **Validación Cruzada:** este método estima dos tipos de errores que evalúan la capacidad predictiva del modelo (Amat, 2016). Entonces, para realizar esta evaluación se divide en dos grupos el conjunto de datos, donde un grupo será utilizado para entrenar al modelo (training data set) y el otro para evaluar al mismo (Test data set) (Amat, 2016; Dobson & Barnett, 2018; Frost, 2019). El primer error es la tasa de error de entrenamiento o training error rate, este estima el error en la predicción del modelo al usar los datos de entrenamiento; al segundo error se lo conoce como la tasa de error de prueba (Test error rate), este evalúa el error de los datos de prueba (Amat, 2016; Frost, 2019). Frost (2019) menciona que la tasa de error de entrenamiento, generalmente, puede estar sobrevalorando la capacidad del modelo, debido a que se utiliza datos que ya ha examinado el modelo; por lo tanto, el error de prueba evalúa mejor la capacidad del modelo porque predice nuevas observaciones.

El tipo de evaluación cruzada que se usó en este estudio fue K-Fold. Este método es iterativo, divide los datos en grupos representativos y del mismo tamaño, el primer grupo entrena y el segundo se emplea como test, este proceso se repite k veces usando grupos diferentes en cada iteración, donde k puede tomar el valor 5 y 10 (Amat, 2016; Refaeilzadeh, Tang, & Huan, 2009). Cada iteración tiene su error y para el resultado final se promedian todos los errores. Amat (2016) Postula la siguiente ecuación para los datos de prueba:

$$CV_{(k)} = \frac{1}{k} \sum_{i=1}^k (Err_i) \quad (10)$$

Donde CV representa el valor de la validación cruzada, k es el número de iteraciones, Err_i simboliza el error de prueba en cada iteración.

- **Curva ROC:** ROC es las siglas en inglés de las características del funcionamiento del receptor (Receiver Operating Characteristic), la gráfica que ofrece este tipo de prueba permite organizar clasificadores y visualizar rendimiento (Fawcett, 2006). Aunque esta prueba es mayormente conocida en decisiones médicas, en los últimos años se han utilizado en el aprendizaje automático y en la investigación de minería de datos (Fawcett, 2006; López & Pita, 1998). Fawcett, (2006) argumenta que, para representar la compensación entre tasa de aciertos y tasas de falsas alarmas de los clasificadores, es necesario primero usar una matriz de confusión, la cual brinda 4 resultado posibles (**Tabla 2**).

Tabla 2.

Resultados de la Matriz de Confusión

		Predicción	
		POSITIVO	NEGATIVO
Observación	POSITIVO	Verdadero Positivo (VP)	Falso Positivo (FP)
	NEGATIVO	Falso Negativo (FN)	Verdadero Negativo (VN)
Total		$VP + FN$	$FP + VN$

Entonces, a partir de estos resultados se analiza la exactitud diagnóstica a través de 2 términos asociados a ROC, el primero la Sensibilidad que es la proporción de elementos clasificados correctamente como Positivo; el segundo término representa la proporción que hay de los elementos negativos que fueron correctamente clasificados y se los conoce como Especificidad (López & Pita, 1998). A partir de estas definiciones se expresan las siguientes ecuaciones:

Fracción de Verdaderos Positivos (FVP):

$$Sensibilidad = \frac{VP}{VP + FN} \quad (11)$$

Fracción de Falso Positivo (FFP)

$$FFP = \frac{FP}{FP + VN} \quad (12)$$

Fracción de Verdaderos Negativos (FVN)

$$Especificidad = \frac{VN}{VN + FP} \quad (13)$$

$$Especificidad = 1 - FFP \quad (14)$$

El gráfico ROC describe la relación entre los beneficios de los VP (eje de las y) y los costos de los FP (eje de las x). Cuando los valores de los Verdaderos Positivos son iguales a los Falsos Positivos, la relación se representa mediante una diagonal imaginaria que va desde la esquina superior derecha hasta la esquina inferior izquierda (López & Pita, 1998). Entonces, mientras más alejada esté la curva de esta diagonal, hacia la esquina superior izquierda, la exactitud de la prueba es mayor (López & Pita, 1998).

- **Área Bajo la Curva ROC:** representado como AUC (area under curve), por sus siglas en inglés. Este es un índice que revela la exactitud global de la prueba. El valor del área de la curva más cercano a 1 tiene mayor exactitud y el valor mínimo aceptable es de 0.5 (López & Pita, 1998).
- **Pseudo R cuadrado:** en la regresión logística no existe R-cuadrado. Sin embargo, se puede usar el Pseudo R-cuadrado porque está en escala similar al R-cuadrado. Esta es una estadística generada en la regresión de mínimos cuadrados ordinarios (MCO), que generalmente se usa para la bondad de ajuste (Statistical Consulting Group, 2020). Esta prueba revela qué tan bien explica el modelo la variable dependiente. Se han desarrollado diferentes Pseudo R-cuadrado (Statistical Consulting Group, 2020; Williams, 2019). El Pseudo R-cuadrado que se usó en esta investigación es Contar (Count, en inglés) el cuál es usado en casos binarios, no está basado en MCO, y usa la siguiente ecuación (Long & Freese, 2001; Statistical Consulting Group, 2020):

$$R^2 = \frac{\text{Número de aciertos}}{\text{Número total de conteo}} \quad (15)$$

Entonces, se cuentan los 1 pronosticados que coinciden con los 1 reales, así también los ceros, la suma de ambos es el número de aciertos y el número total es las pruebas totales hechas (Statistical Consulting Group, 2020).

- **Pruebas de colinealidad VIF:** el Factor de Inflación de Varianza (Variance Inflation Factor, en inglés), evalúa cuánto se infla la varianza de un coeficiente dentro de una regresión, esto en función a la multicolinealidad en el modelo (Departament of Statistics Online Program, 2018). Para el cálculo se necesita obtener un R-cuadrado para cada predictor, el cual se obtiene realizando una regresión de la variable independiente elegida frente a los demás predictores del modelo, estos valores se introducen en la siguiente fórmula:

$$VIF = \frac{1}{1 - R_i^2} \quad (16)$$

Donde R_i es el R-cuadrado de cada predictor. El valor numérico decimal sobre 1 indica el porcentaje de varianza que se infla para cada coeficiente. Se sugiere que valores iguales a 1 significan que no hay correlación, valores entre 1 a 4 indican correlación moderada, mientras que valores superiores a 10 delatan correlaciones muy altas (Departament of Statistics Online Program, 2018).

2.1.5. INTERPOLACIÓN DE LA PROBABILIDAD

La interpolación supone que los objetos que están distribuidos de forma espacial están correlacionados. Es por ello que esta es una buena herramienta para obtener valores que eran desconocidos. La interpolación Kriging es un método geoestadístico, para datos que siguen una distribución normal, esta se basa en la autocorrelación, permitiendo producir una superficie de predicción y además provee una medida de certeza de las predicciones (ArcGIS for Desktop, 2016; Burrough & McDonnell, 1986; Oliver & Webster, 1990). Para ArcGIS 10.4 existen varios métodos Kriging, entre los más usados se tiene, el ordinario que es el predeterminado y supone que el valor medio es constante, y el segundo llamado Kriging universal que se utiliza si se conoce la tendencia de los datos (ArcGIS for

Desktop, 2016; Oliver & Webster, 1990). La fórmula general del método es la suma ponderada de los datos:

$$\hat{Z}(S_0) = \sum_{j=1}^M \lambda_j Z(S_j) \quad (17)$$

Donde $Z(S_j)$ representa el valor medio en la ubicación j (este cálculo se basa en la distancia que existe entre los puntos medidos, la ubicación de la predicción y en la disposición espacial general de los puntos medidos), λ_j es el símbolo para una ponderación que es desconocida para el valor medido en la ubicación j , S_0 es la ubicación de la predicción, y finalmente M es la cantidad de valores medidos (ArcGIS for Desktop, 2016). Además, la interpolación Kriging tiene dos pasos importantes, la creación de variogramas y la predicción, los cuales se detallan seguidamente:

- Creación de variogramas: en la variografía de puntos medios, se comienza con un semivariograma empírico que se calcula mediante la fórmula que se plantea en ArcGIS for Desktop (2016).

$$\text{Semivariogram}(distance_h) = 0.5 \times \text{average}((value_i - value_j)^2) \quad (18)$$

Donde h es la distancia de separación de los pares de ubicaciones. La fórmula calcula el semivariograma mediante la diferencia cuadrada de los valores estudiados en cada una de las ubicaciones donde se hizo muestreo (ArcGIS for Desktop, 2016). El semivariograma empírico es la gráfica donde el eje X es la distancia y el eje Y son los valores de semivarianza de los distintos pares de ubicaciones (*Figura 5*), este también proporciona información sobre la autocorrelación espacial del conjunto de datos (ArcGIS for Desktop, 2016).

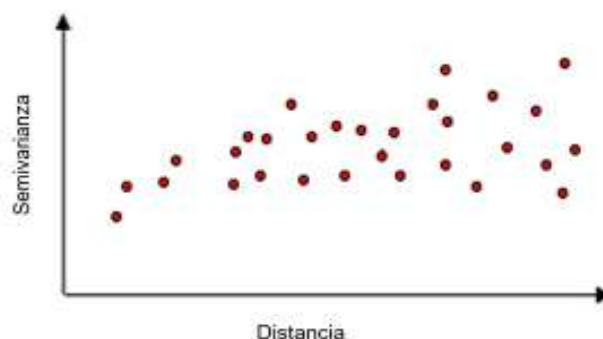


Figura 5. Semivariograma
Fuente: (ArcGIS for Desktop, 2016)
Modificado: Abigail Navarrete

Entonces, mientras los pares de ubicaciones estén más cerca tienen valores similares y a medida que crece en ambos ejes, los pares de ubicaciones son más distintos y con un diferencia cuadrada mayor (ArcGIS for Desktop, 2016). El semivariograma empírico puede cambiar de acuerdo al modelo elegido para el ajuste y cuando la curva está más cerca al origen y además tiene cambios suaves significa que, existirá mayor influencia de los datos para predecir un punto sin información; los modelos pueden ser Circular, Esférico, Exponencial, Gaussiana y lineal (ArcGIS for Desktop, 2016; McBratney & Webster, 1986).

Los semivariogramas tienen 3 componentes que se usan para describir los modelos (*Figura 6*):

- Rango: se denomina rango a la distancia, en el eje X, donde la curva comienza a nivelarse. Entre más pequeñas sean las distancias entre los pares, en comparación con el rango, la información de los pares estará correlacionada, y aquellos pares que tienen una distancia mayor a la del rango no lo están (ArcGIS for Desktop, 2016).
- Meseta: es la distancia, en el eje Y, donde la curva comienza a nivelarse (ArcGIS for Desktop, 2016).
- Nugget: es una discontinuidad en el origen de la gráfica; este efecto se lo puede atribuir a los errores de medición u otros (ArcGIS for Desktop, 2016).

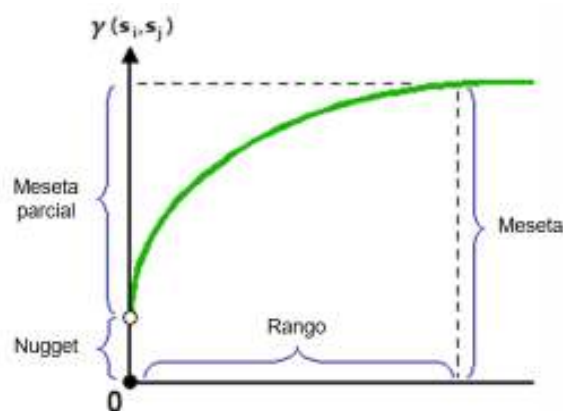


Figura 6. Componentes del Semivariograma

Fuente: (ArcGIS for Desktop, 2016)

Modificado: Abigail Navarrete

- **Predicción:** el paso dos es la predicción, esta es la aplicación principal de la herramienta Kriging. Se encarga de la predicción de aquellos valores de los puntos que no fueron muestreados y que por lo tanto no tienen información (ArcGIS for Desktop, 2016). Este proceso se da a través de ponderaciones de los valores cercanos, donde se crea una superficie continua con predicciones para cada centro de celda; además, este cálculo está basado en el semivariograma que se realizó en el paso uno y en la distribución espacial de los datos muestreados (ArcGIS for Desktop, 2016).

2.2.ESTADO DEL ARTE

2.2.1. INCENDIOS EN ECUADOR

En el Ecuador los incendios forestales se producen con mayor intensidad a partir del inicio del verano, y las causas se distribuye de la siguiente manera, el 70% son por causa antrópica, el 25% por las actividades agrícolas y solamente el 5% se producen por causas naturales (Ministerio del Ambiente y Agua, 2014). Además, desde el 2016 al 2019 solamente en los meses de agosto a octubre se produjeron 3672 eventos, estos incendios afectaron a un área de 51540.83 hectáreas (Dirección de Monitoreo de Eventos Adversos, 2016, 2017, 2018, 2019). Para octubre del 2019 las provincias más afectadas por incendios fueron Pichincha, Imbabura, Guayas, Chimborazo, Guayas, Azuay, Loja, El Oro (Dirección de Monitoreo de Eventos Adversos, 2019).

2.2.2. INFORMACIÓN DE INCENDIOS DISPONIBLES EN EL PAÍS

Los organismos estatales que se encargan de estudiar, prevenir e informar sobre los eventos de incendios forestales en el país, son los siguientes:

- Servicio Nacional de Gestión de Riesgos y Emergencias (SNGR): esta institución se encarga de garantizar la protección ante desastres naturales o antrópicos a nivel nacional
- Geoportal del SNGR - Incendios forestales: este sitio ofrece información geográfica del periodo 2013-2018, sobre el análisis de susceptibilidad a incendios forestales, escenarios de afectación, focos de calor y eventos registrados en el territorio nacional (SNGR, 2017)
- Ministerio del Ambiente y Agua: esta entidad pública es la encargada de realizar los planes de contingencia de Incendios forestales, los cuales tienen como propósito evitar que este fenómeno cause grandes daños en las zonas del Sistema Nacional de Áreas Protegidas (también conocido como SNAP), y en otras áreas que también son vulnerables pero que no son parte del SNAP, como los páramos y bosques (Ministerio del Ambiente y Agua, 2014).

Existen otras entidades que también ofrecen y aportan con información de estos eventos:

- Bomberos: son organismos públicos destinados a defender a personas y a bienes contra el fuego.
- DesInventar: ofrece información sobre ocurrencia de desastres en América Latina, esta base de datos se crea a través de reportes, y fuentes hemerográficas.
- Observatorio Social El Batefuegos de Oro: es una asociación que ofrece información sobre gestión forestal y análisis de los eventos de incendios, esta información está abierta para el público mediante su portal.
- Instituciones Educativas: las instituciones universitarias que realizan investigaciones sobre este fenómeno.

2.2.3. VARIABLES ASOCIADAS A LOS INCENDIOS

Ecuador es muy rico en vegetación, la cual en verano se convierte en un buen combustible, pero este no es el único factor que provoca un incendio, sino que existen más variables que ayudan a producir estos eventos. Así tenemos, la

radiación solar que sumada a las grandes alturas especialmente de la Sierra ecuatoriana, aceleran la pérdida de humedad relativa provocando que cualquier chispa produzca un incendio (Bustos et al., 2019). En algunas investigaciones sobre incendios se ha usado variables climáticas como temperatura de la superficie del mar, temperatura del aire, humedad relativa, velocidad del viento, esto ha servido para caracterizar las respectivas áreas de estudio (Kitzberger et al., 2006; Pazmiño, 2019; Tan & Panda, 2011).

Martinez et al. (2011) mencionan que, otros factores que (a parte de los que ya se han mencionado) influyen en las características del clima andino son, los bosques andinos porque producen grandes cantidades de vapor de agua que influyen en la Cordillera Oriental; las corrientes oceánicas debido a la cantidad de humedad que llevan hacia la Cordillera Occidental, esto se da por la advección de masas de aire que vienen desde el Océano Pacífico. Por otra parte, también se recomienda aportar al modelo con información socio-económica como sectores económicos, crecimiento urbano, actividad económica, uso de suelo, etc., pues la mayoría de los eventos son provocados por actividad humana (Altamirano et al., 2013; Martínez et al., 2004)

CAPITULO 3 ÁREA DE ESTUDIO

El área de estudio (AE) se definió a partir de los resultados que se obtuvieron del análisis de la investigación de la EPN “Spatiotemporal patterns of burned areas, fire drivers, and fire probability across the equatorial Andes”, a esta investigación se la llamará “AE PIJ”. El área en la que se centró la investigación de esta tesis, comprende la zona ubicada en el norte y en el sur de la región Interandina del Ecuador más una pequeña porción de la región de la Costa, así como se presenta en la *Figura 7*.

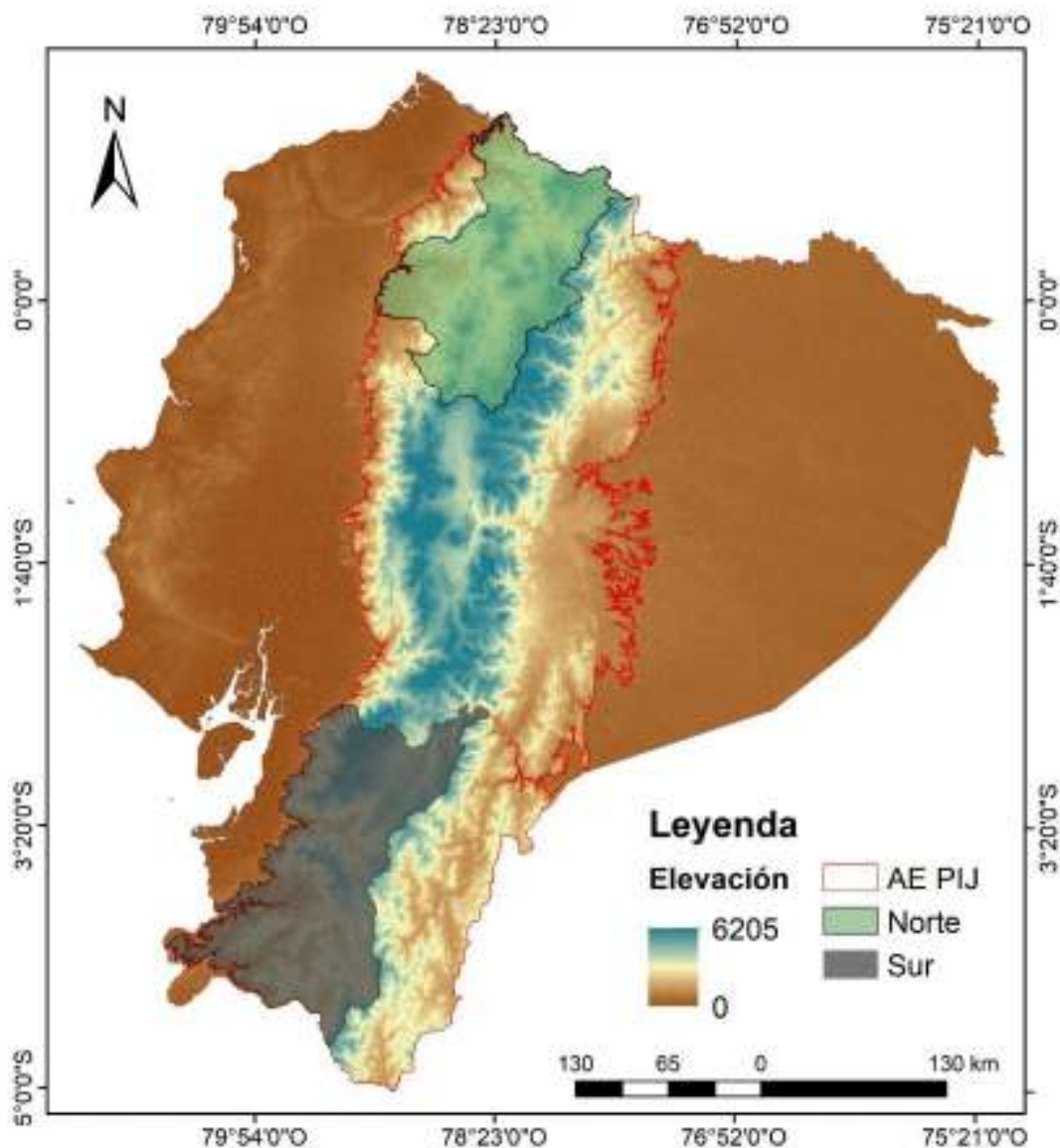


Figura 7. Ubicación de las áreas de estudio
Elaborado: Abigail Navarrete

3.1.ZONA NORTE

Incluye las provincias de Carchi, Imbabura, y Pichincha, abarca 20 cantones que ocupan un área aproximada de 16 611.56 km², la cual representa un 6.7% de la superficie continental de Ecuador. Comprende elevaciones que van desde 233 a 5876 m.s.n.m. La zona está dominada por los bioclimas Pluvial y Pluvial estacional, esto implica que todo el año existirá vegetación, la cual es casi en su totalidad siempre verde estacional (MAE, 2012). De acuerdo, a los datos de cobertura y uso de suelo que ofrece el Ministerio de Agricultura, Ganadería, Acuacultura y Pesca del Ecuador (MAGAP), encontramos que del total de la superficie del norte el 39.63% es Tierra Agropecuaria, el 30.10% es Bosque, el 25.80% es Vegetación Arbustiva y Herbácea, el 3.38% es Zona Antrópica, el 0.30% son Cuerpos de Agua, y el 0.80% son Otras Áreas. Además, el área tiene una precipitación media anual igual a 1345 mm, Irradiación global horizontal media anual de 1615 kwh/m² y una temperatura media mensual de la superficie del suelo igual a 21°C. Asimismo, la población es de 3 105 686 habitantes, según el censo que publicó el Instituto Nacional de Estadística y Censos del Ecuador (INEC) para el año 2010 (*Figura 8 y 9*)

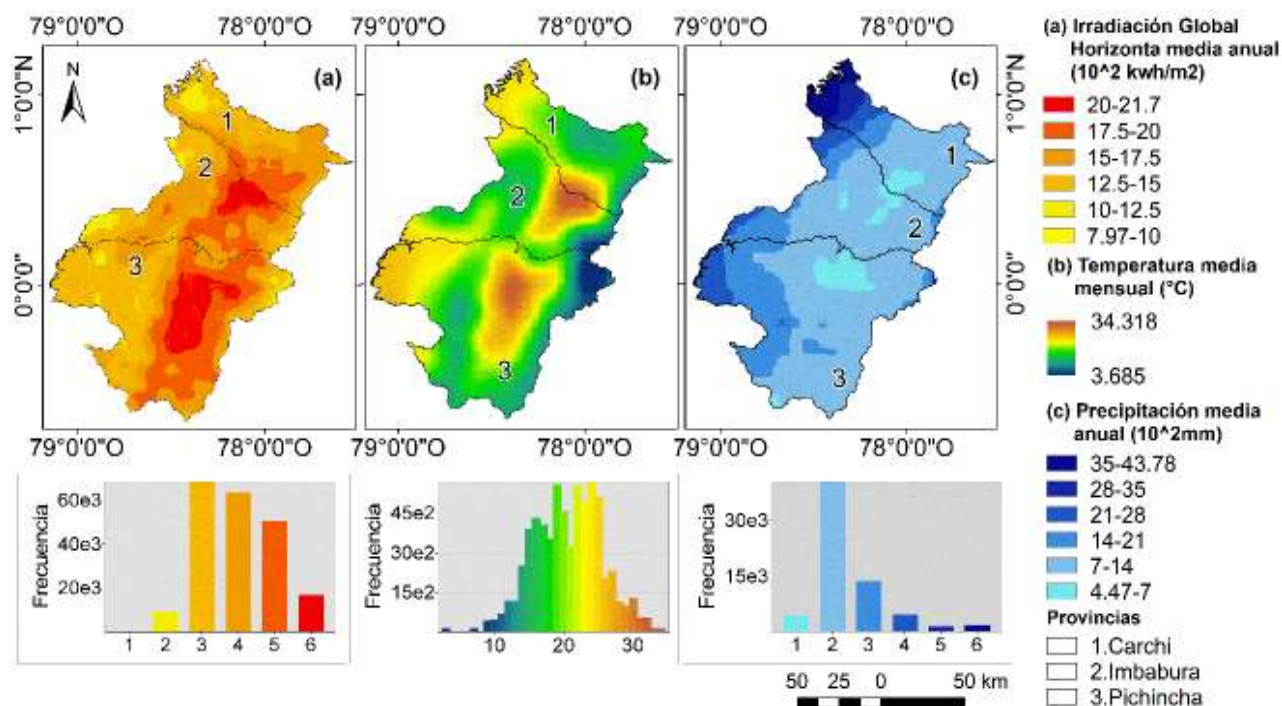


Figura 8. Variables climáticas en la Zona Norte

Elaborado: Abigail Navarrete

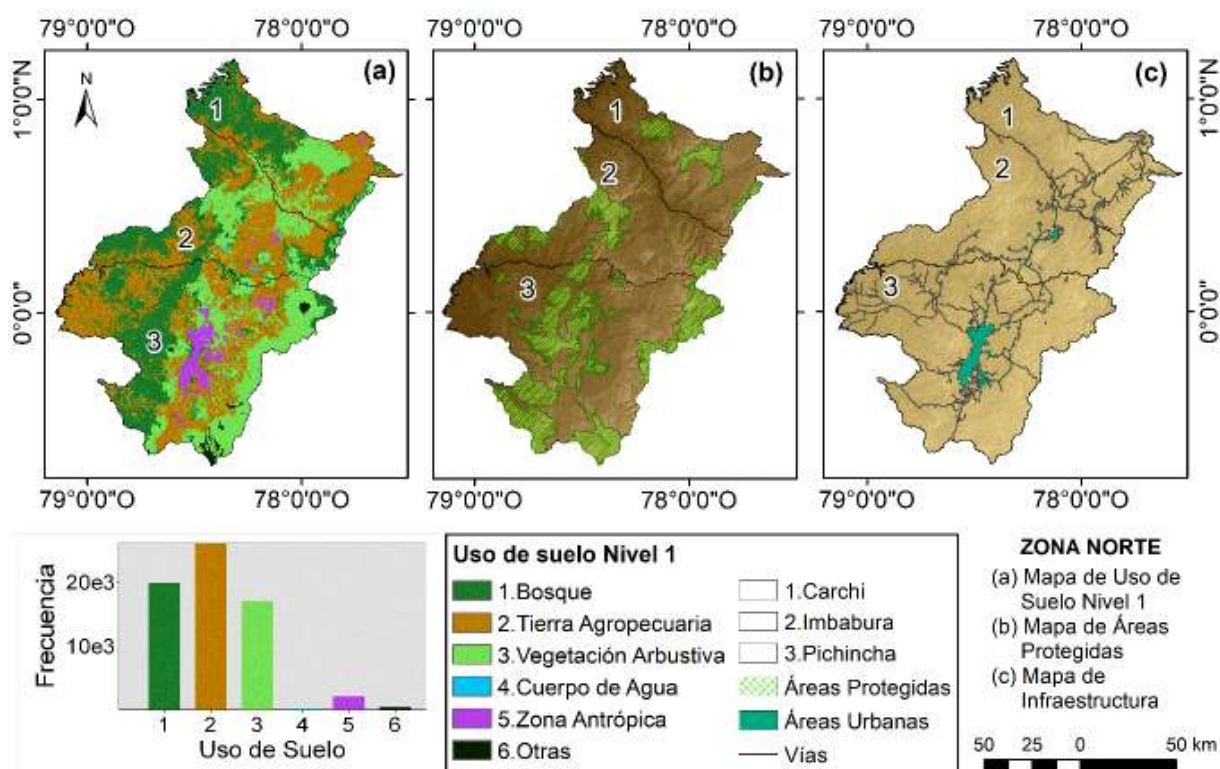


Figura 9. Uso de suelo, Áreas Protegidas y variables de Infraestructura en la Zona Norte
 Elaborado: Abigail Navarrete

3.2.ZONA SUR

Comprende las 3 provincias del sur de la región interandina, Azuay, Loja y El Oro. Este barca 43 cantones, posee una superficie igual a 20530.54 km² representando el 8.3% de la superficie de Ecuador continental. En esta zona se puede encontrar elevaciones que van desde 289 hasta 4532 m.s.n.m. También, predominan los bioclimas Pluvial, Pluvial estacional y xérico, esto implica que todo el año existe vegetación, pero también la sequía se hará presente en ciertos sectores de esta zona. Además, la vegetación que sobresale en esta área es semideciduo (MAE, 2012). De acuerdo a los datos de cobertura y uso de suelo del MAGAP de Ecuador, se encontró que el 43.70% del total del área es Tierra Agropecuaria, el 29.96% es Vegetación Arbustiva y Herbácea, el 24.60% es Bosques, el 0.57% es Zona Antrópica, el 0.31% son Cuerpos de Agua y el 0.85% son Otras Áreas. Esta zona tiene una precipitación media anual igual a 877 mm, irradiación global horizontal media anual 1537 kwh/m² y una temperatura media mensual de 24°C. Por último, para el censo del 2010 que realizó el INEC de Ecuador, la población para esta zona fue de 1 248 800 habitantes (*Figura 10 y 11*).

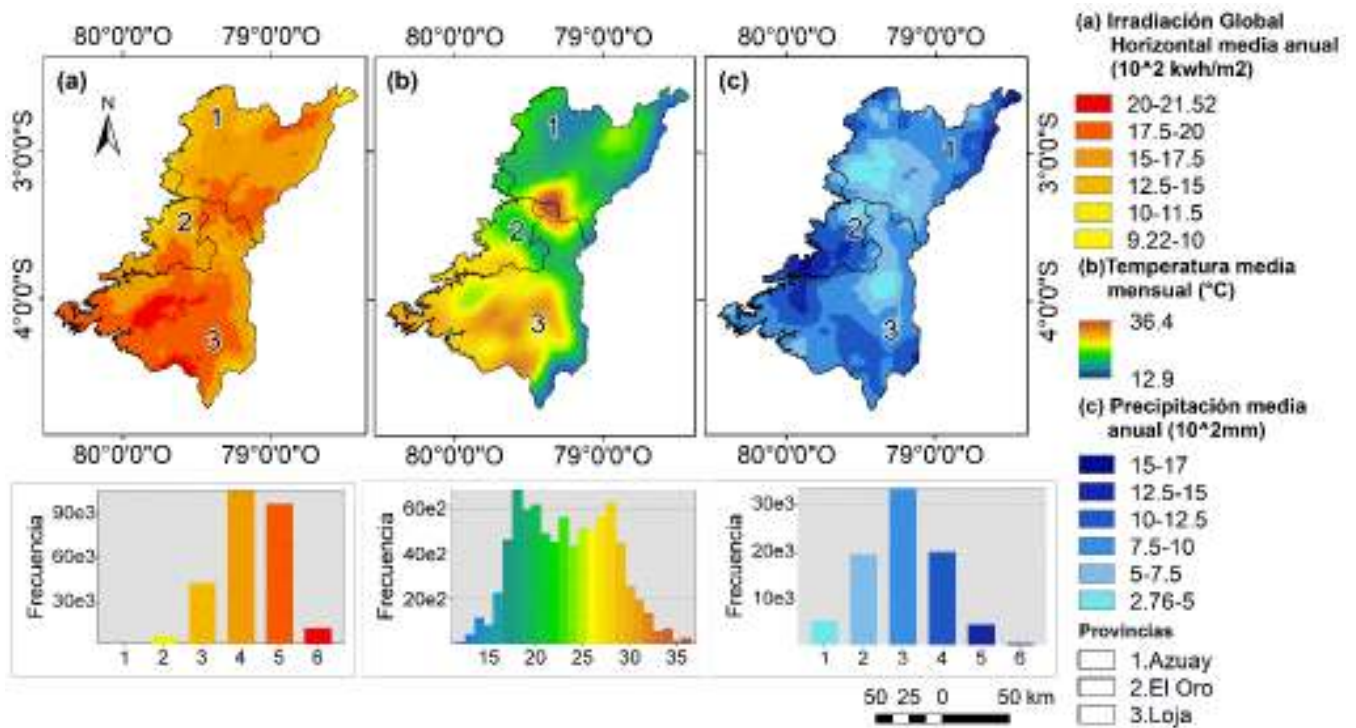


Figura 10. Variables climáticas en la Zona Sur
Elaborado: Abigail Navarrete

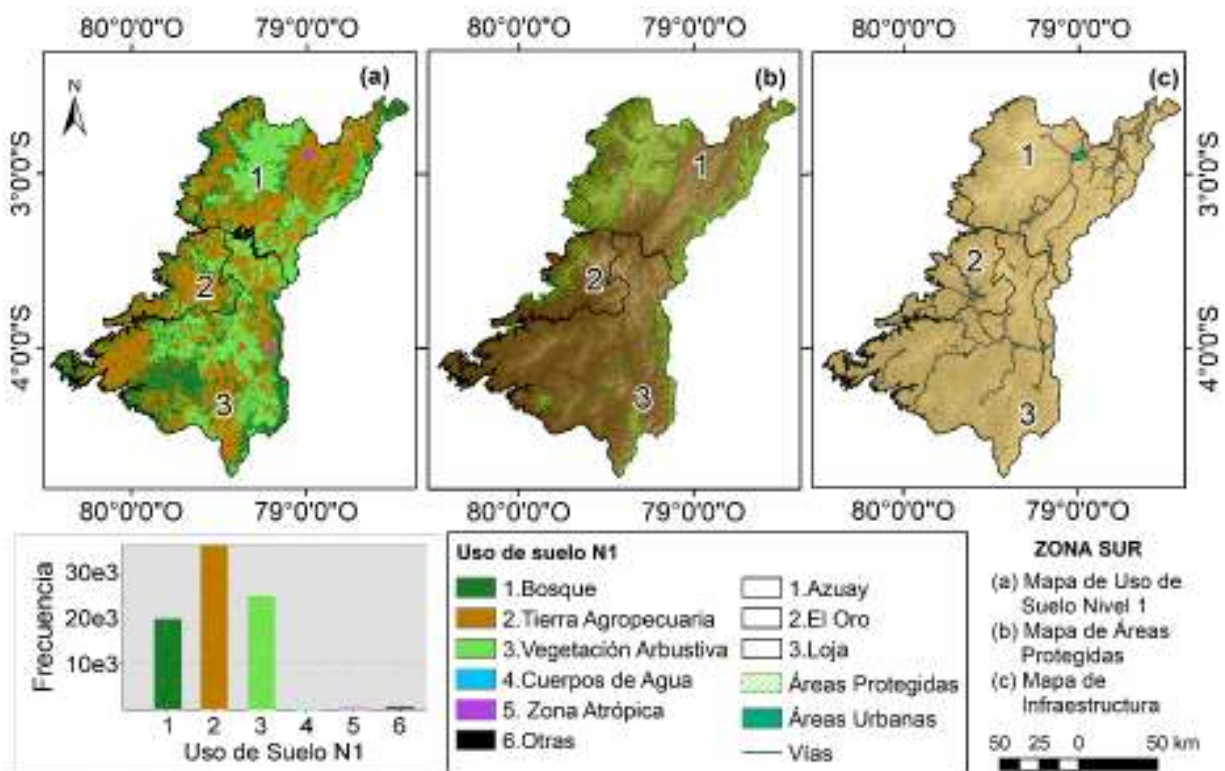


Figura 11. Uso de suelo, Áreas Protegidas y variables de infraestructura en la Zona Sur
Elaborado: Abigail Navarrete

CAPITULO 4 METODOLOGÍA

4.1.ESQUEMA DE TRABAJO

Se trabajó en 4 etapas, tal como muestra la *Figura 12*. La primera etapa abarcó todo lo relacionado con la descarga de datos, el tratamiento de los mismos como mosaico y reproyección de imágenes crudas (según correspondía), se realizó un análisis de promedios mensuales y anuales de los productos, y se trató el área quemada, aquí se estableció el comportamiento espacial de los incendios, finalmente se extrajo el área de estudio. La segunda etapa implicó un análisis de la información de forma separada por lo que se dividió la base de datos para las dos zonas de estudio; una vez dividida esta base se procedió a realizar un análisis descriptivo, el cual implicaba en visualizar el comportamiento de las variables independientes en función del área quemada, esto nos permitió establecer la temporalidad de los incendios. Para la tercera etapa se realizó un análisis estadístico, mediante pruebas estadísticas que permitieron establecer las variables significativas; se dividió los datos de cada zona, el primer grupo de datos se usó para entrenar el modelo, se elaboró el modelo logístico, y se valida este entrenamiento. En la etapa 4 se usó el segundo grupo de los datos para evaluar al GLM y se validó al modelo mediante la curva ROC y El área Bajo la curva ROC, se interpoló la probabilidad obtenida del GLM y se realizaron mapas de probabilidad para cada zona.

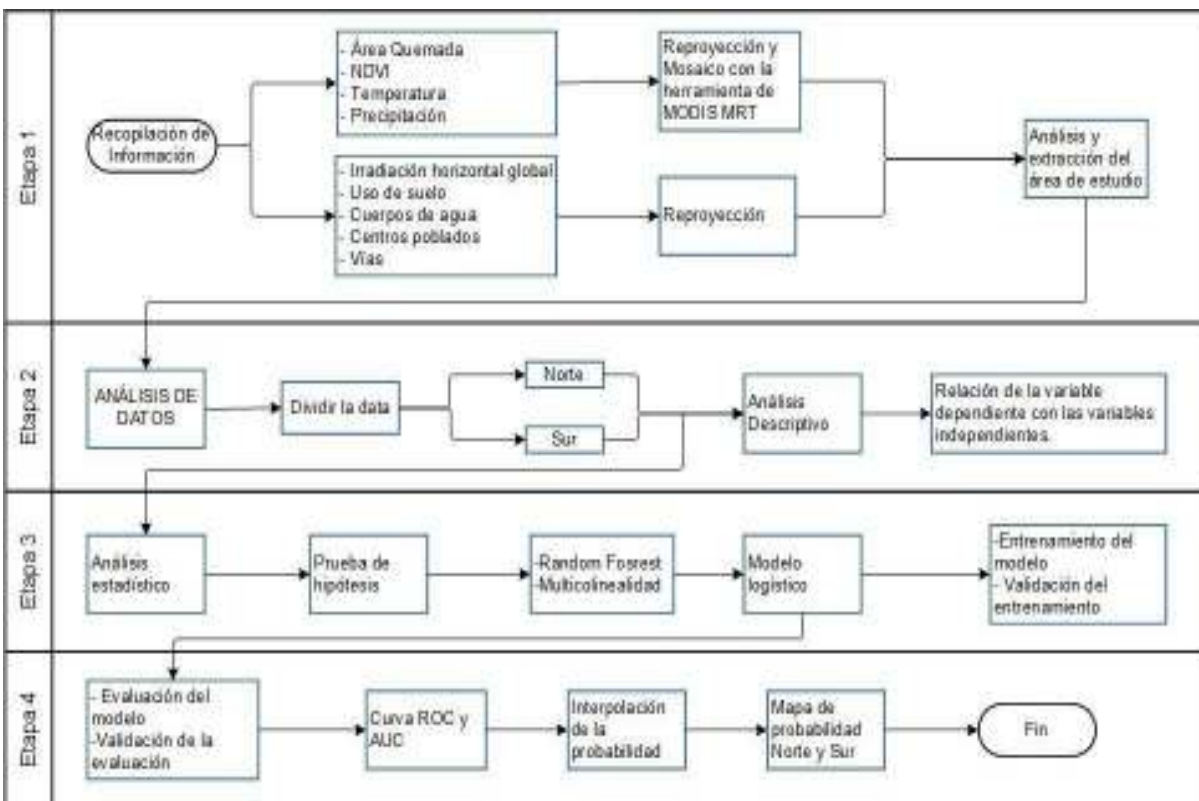


Figura 12. Esquema de trabajo
Elaborado: Abigail Navarrete

4.2.RECOPIACIÓN DE DATOS

La información que se procesó para la base de datos, se resume en la tabla que se presenta en las siguientes **Tablas 3 y 4**:

Tabla 3.

Resumen de los productos descargados

Producto	Instrumento	Empresa	Enlace
Área Quemada	MODIS (Terra-Aqua)	NASA	https://search.earthdata.nasa.gov/search?q=MCD64A1
NDVI	MODIS (Terra)	NASA	https://search.earthdata.nasa.gov/search?q=MOD13A3
Temperatura	MODIS (Terra)	NASA	https://search.earthdata.nasa.gov/search?q=MOD11B3
Precipitación	Datos globales de estación y CHPclim, imágenes satelitales	Climate Hazards Group, USAID	https://www.chc.ucsb.edu/data

Irradiación horizontal global	Privado	SolarGis	https://solargis.com/es/maps-and-gis-data/download/world
Modelo de elevación digita	SRTM		http://srtm.csi.cgiar.org/srtmdata/

Tabla 4.
Características de los productos satelitales

Producto	Periodo	Resolución Espacial	Resolución Temporal	Empresa	Formato
Área Quemada	2001-2018	500m	Mensual	NASA	HDF
NDVI	2001-2018	1000m	Mensual	NASA	HDF
Temperatura	2001-2018	6000m	Mensual	NASA	HDF
Precipitación	2001-2018	0.05°	Mensual	Climate Hazards Group, USAID	HDF
Irradiación horizontal global	1999-2018	9 arc-sec	Promedio anual	SolarGis	TIFF
Modelo de elevación digita		30 m			Ráster
Uso de suelo				MAGAP	Polígono
Cuerpos de agua				SNI	Polígono
Centros urbanos				SNI	Polígono
Vías				SNI	Polígono

4.3.PROCESAMIENTO DE DATOS

Por cada fecha de los diferentes productos de la NASA se descargó 4 archivos HDF sin proyección y se realizó el siguiente proceso:

Los datos se procesaron mediante la herramienta MRT (Modis Reprojection Tool), esta herramienta de MODIS permitió que las imágenes fuesen proyectadas a WGS 1984 UTM Zone 17N y se hizo un mosaico, es así que como información de salida se obtuvo 1 archivo por cada fecha en formato TIF (*Figura 13-14*).

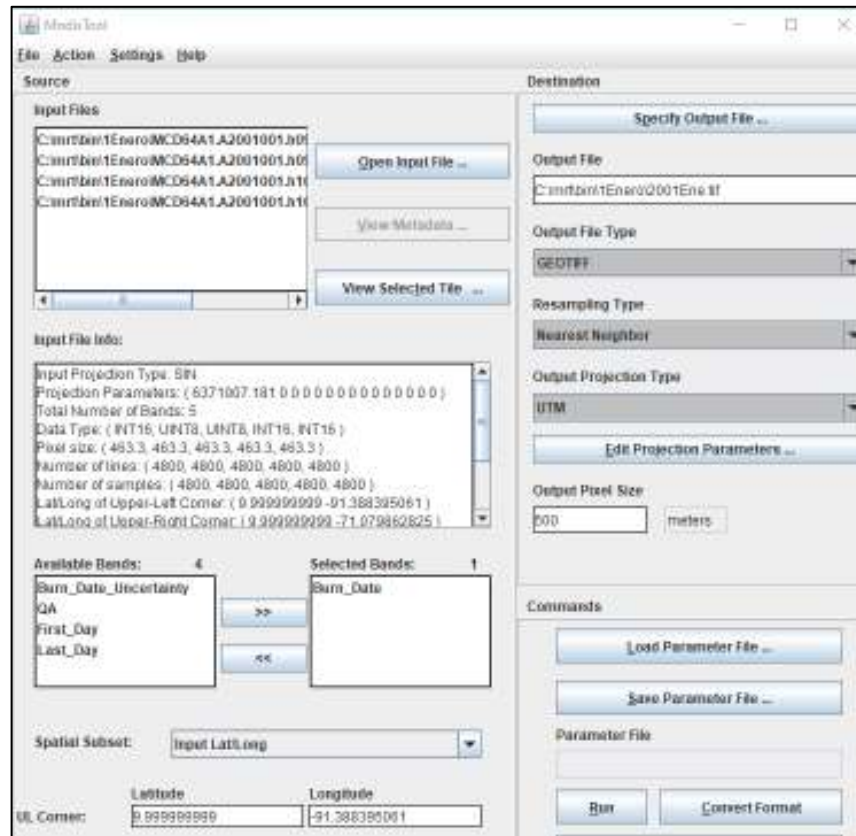


Figura 13. Procesamiento de los archivos HDF de enero 2001 en el programa MRT
Elaborado: Abigail Navarrete

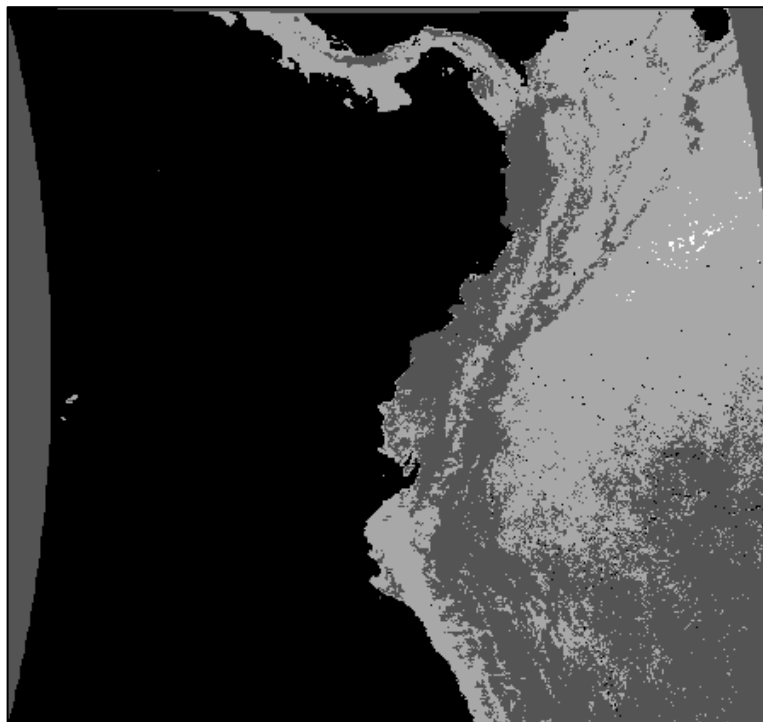


Figura 14. Variables climáticas en la Zona Sur
Elaborado: Abigail Navarrete

Luego, de cada mosaico se extrae el área de estudio y las imágenes se analizan en el programa ArcGIS. Para la información externa a los productos de la NASA se procedió a re proyectar y extraer el área de estudio en ArcGIS.

4.3.1. PROCESAMIENTO DE RÁSTER INCENDIOS

Como se escogió la banda Burn_Date, los pixeles de cada imagen mostrarán los siguiente valores, fecha juliana (1-366) es la fecha de incendio, el valor 0 indica tierra sin quemar, -1 significa que no se pudo mapear debido a datos insuficientes, y -2 es agua (Louis Giglio, Boschetti, Roy, Hoffman, & Humber, 2016). Después de obtener el área de estudio del producto MCD64A1, se procedió a reemplazar los valores antes mencionada, donde hay fecha (juliana) de área quemada el pixel será igual a 1 y 0 donde los valores van de -2 a 0. Esto se hizo mediante el condicional $OutRas = Con("mes", 1, 0, "VALUE > 0")$ para cada mes de cada año en la Calculadora Ráster de ArcMap, tal como se muestran en las *Figuras 15-16* para septiembre de 2012:

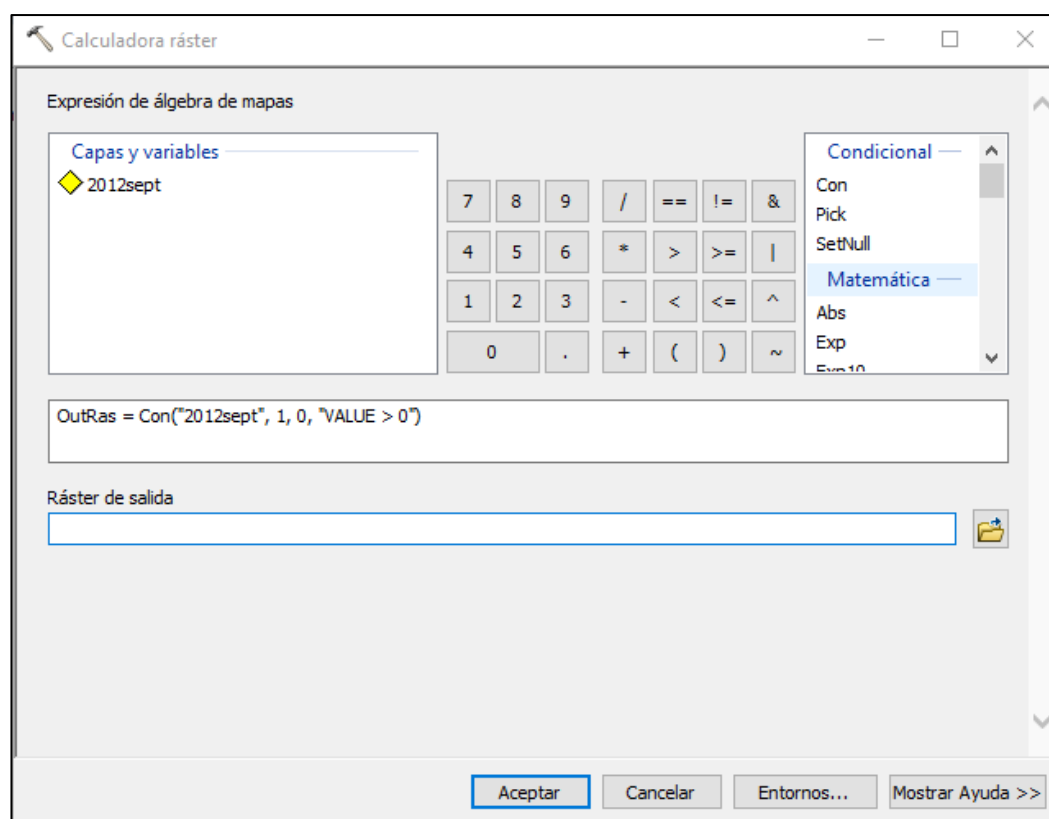


Figura 15. Cálculo del condicional para convertir a valores binomiales
Elaborado: Abigail Navarrete

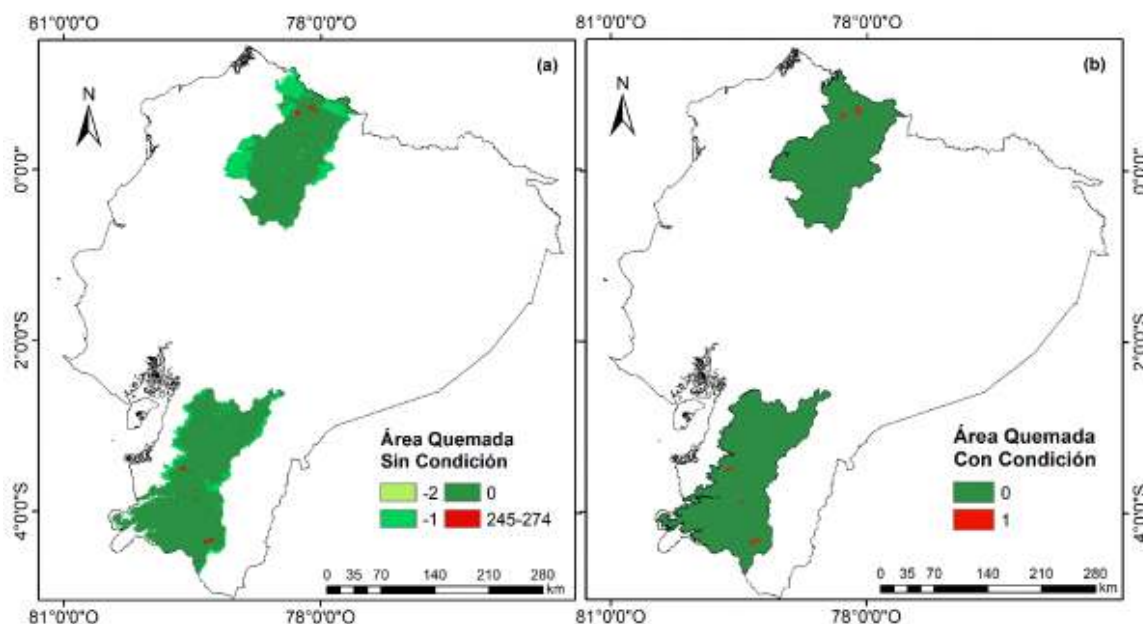


Figura 16. Áreas sin condicional y con condicional para septiembre de 2012
Elaborado: Abigail Navarrete

Luego de obtener todas las imágenes como dato binomial, se procedió a sumar de forma mensual y finalmente de forma anual, mediante la Calculadora Ráster, para tener el número total de incendios en todo el periodo.

4.4. OBTENCIÓN DE PUNTO DE INCEDIOS

Para obtener los puntos de incendios, primero se transformó en una cobertura vectorial de tipo polígono a cada imagen Ráster de Área Quemada. Luego, se procedió a exportar los datos con fecha juliana como polígonos independientes. Entonces, se transformó a puntos los polígonos independientes a través de la herramienta “Feature To Point” y finalmente se hizo una fusión de todos los puntos de incendio con la herramienta “Merge”. Todo esto se trabajó en el programa ArcGIS y el procedimiento se resumió en el siguiente diagrama (*Figura 17*).

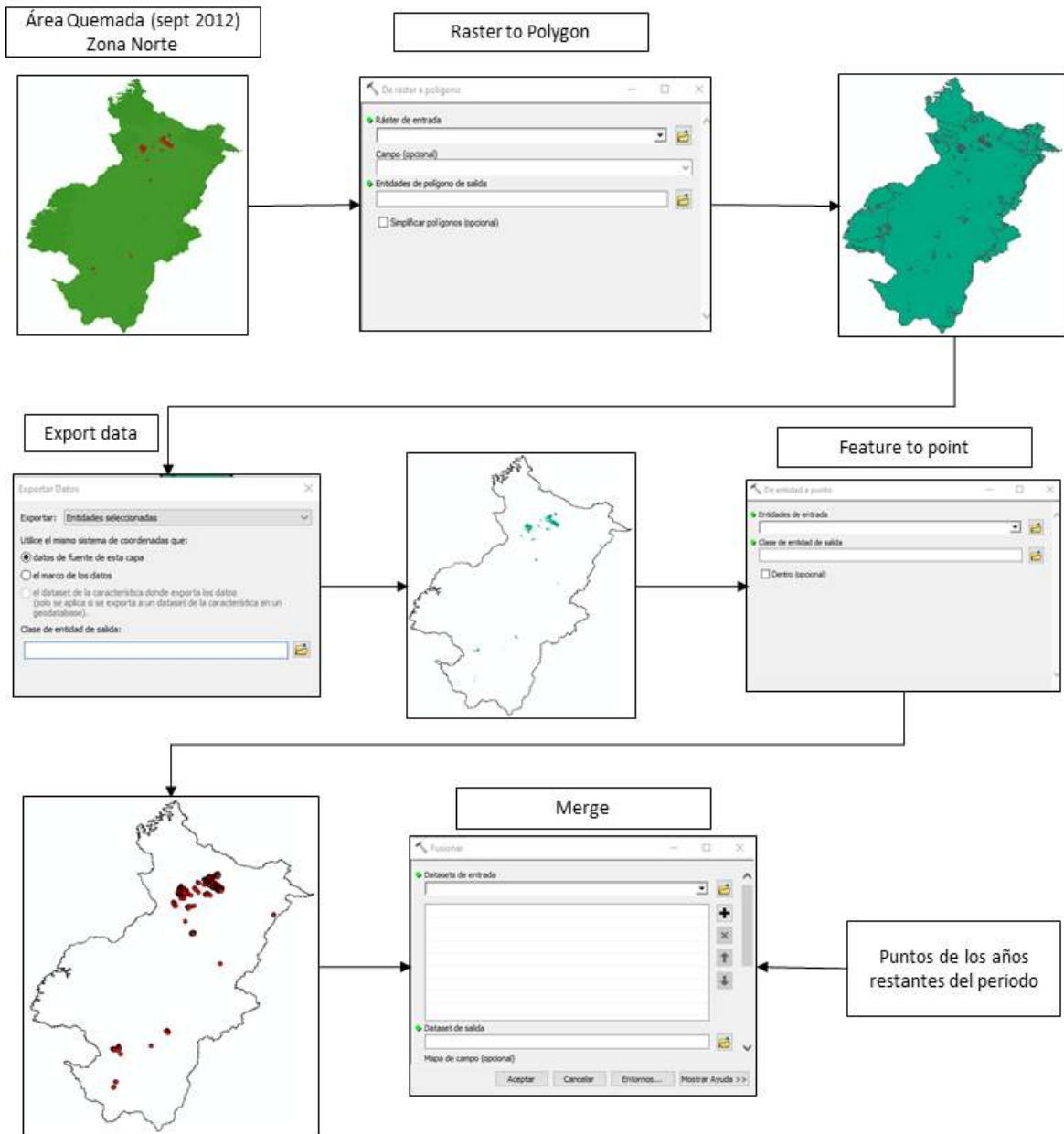


Figura 17. Diagrama del proceso para obtención de puntos de incendio
Elaborado: Abigail Navarrete

4.5. ANÁLISIS DE DATOS

Dentro del análisis, lo primero que se hizo es retirar los datos atípicos como incendios en cuerpos de agua y aquellos que no aportaban con información al modelo como incendios en zonas antrópicas, áreas sin información y otras áreas. A partir de esta información se realizó dos análisis uno descriptivo y otro estadístico.

4.5.1. CÓDIGO DE LAS VARIABLES

Los códigos de las variables se colocó por su nombre en inglés y se añadió después del guion bajo (_) información adicional, que fue específica de cada variable como por ejemplo la unidad de medida o la temporada en la que se promediaron los datos, y en el caso del uso de suelo se colocó el nivel de uso de suelo con el que se trabajó. Para la temporalidad mjj significa temporada antes de los incendios y tot significa que comprende todo el periodo de estudio (**Tabla 5**).

Tabla 5.

Significado y tipo de variable

Código	Significado	Tipo de variable
Class	Variable dicotómica de incendio	dependiente-dicotómica
urban_m	Distancia a centros en metros	Infraestructura-independiente
road_m	Distancia a vías en metros	Infraestructura-independiente
river_m	Distancia a ríos en metros	Infraestructura-independiente
slope	Pendiente en porcentaje	Topográfica-independiente
dem_m	Elevación de modelo digital en metros	Topográfica-independiente
aspect_1	Aspecto en grados	Topográfica-independiente
GHI	Irradiación horizontal global	Climática-independiente
prec_mm	Media mensual de la precipitación en mm	Climática-independiente
prec_ma	Media anual de la precipitación en mm	Climática-independiente
prec_mjj	Media mensual de la precipitación de la temporada antes de los incendios	Climática-independiente
temp_mm	Media mensual de la temperatura °C	Climática-independiente
temp_mjj	Media mensual de la temporada antes de los incendios °C	Climática-independiente
ndvi_tot	Índice de vegetación de diferencia normalizada total	Biológica-independiente
ndvi_mjj	Índice de vegetación de diferencia normalizada de la temporada antes de los incendios	Biológica-independiente
vci_tot	Índice de condición vegetal de todo el periodo en porcentaje	Biológica-independiente
Cus_n1	Cobertura de uso de suelo del nivel 1	Biológica-independiente

4.6. ANÁLISIS DESCRIPTIVO

En este análisis se examina la interacción de los incendios, de enero a diciembre, con las variables independientes (Tabla 5). Para esto se intersecaron los puntos con la información vectorial mediante la herramienta *intersect* de ArcGIS, mientras que para la información ráster se usó la herramienta *extract multi values to point* de ArcGIS. Una vez que se obtuvo una tabla con la información de las variables independientes para cada punto de incendio, se procedió a realizar gráficos de barra y línea para observar el comportamiento de los incendios. Además, este es un primer paso para observar los datos que benefician al modelo logístico.

4.7. ANÁLISIS ESTADÍSTICO

En esta sección se explora y describe la relación de las variables y el comportamiento estadístico individual de las mismas, esto es para crear el modelo probabilístico. Aquí se usó la información de los eventos que sucedieron en la temporada de incendio para su respectiva zona. Además, para satisfacer la variable dicotómica se crearon puntos aleatorios de no incendio en la proporción 1:1.5 (Catry, Rego, Bação, & Moreira, 2009; Guo et al., 2016). Estos puntos se crearon a un 1km de distancia de los puntos de incendios y se les asignó el valor de cero, mientras que a los incendios se les designa uno (Guo et al., 2016, 2017).

4.7.1. PRUEBA DE HIPÓTESIS

Esta prueba se realizó para las variables independientes. Entonces, para las variables cuantitativas se realizó una prueba no paramétrica, para la cual se usó la función `sm.density.compare` del paquete `sm` del programa R, esta función permite comparar la distribución de incendios y la de no incendios, además produce una banda de referencia, que se muestra en color celeste, para una prueba de hipótesis de arranque de igualdad. Por otro lado, para las variables cualitativas se usó una prueba paramétrica, se asume como hipótesis nula que las proporciones en cada grupo son iguales (Müller, 2006); se usó la función `prop.test` y el resultado se presenta como `p_value`, al 95% de confianza. Lo que las pruebas muestran es que, si la distribución de incendios y no incendios es similar, la variable no aporta con información al modelo y puede ser eliminada.

4.7.2. VARIABLES SIGNIFICATIVAS

4.7.2.1. Árboles de decisiones

Como primer paso para esta prueba, se hizo una gráfica para establecer un número determinado de árboles estables, ambas zonas mostraron estabilidad para un número aproximado de 200 árboles. De acuerdo con este número se realizó la prueba de Mean Decrease Gini (MDG) y Mean Decrease Accuracy (MDA), donde la significancia de las variables disminuye conforme se descende en la gráfica. Para esto se utiliza la función *importance*, del paquete *randomForest* de R.

4.7.2.2. Multicolinealidad

En esta prueba no se usó la variable discreta (Uso de Suelo) como tal, sino que se utilizó un valor numérico que representa a la categoría. La función que genera este análisis en R fue *cor()* del paquete *stats*.

4.8.MODELO LINEAL GENERALIZADO

Se dividió los datos en dos secciones, la primera fue la sección de entrenamiento y la segunda sección fue de prueba, cada sección contó con la representatividad de la muestra, es decir una proporción 0.4:0.6 entre incendios y no incendios, respectivamente. Luego se escogieron las variables independientes que alimentarían al modelo, esto fue en función de las pruebas de significancia, a la combinación de estas variables predictivas la llamamos receta. El modelo fue automatizado en R mediante el paquete *stats* y la función *glm*, en la cual se usó la familia binomial. Lo que se obtiene de este análisis es la probabilidad de ocurrencia.

4.9.VALIDACIÓN DEL MODELO

La validación empieza desde la división de los datos en la sección de entrenamiento y en la sección de prueba. Para la validación cruzada el k-fold fue igual a 10 iteraciones, para lo cual se utilizó la función *folds*. También, se evaluó la prueba de Wald con la función *regTermTest* del paquete *survey*, y para el VIF se utilizó la función *VIF* del paquete *car*. La función *roc* se usó para obtener la curva ROC y finalmente para el área bajo la curva se usó la función llamada *auc*, ambas funciones del paquete *pROC*, todo esto en el programa R.

4.10. INTERPOLACIÓN DE PROBABILIDAD

Para tener la probabilidad en toda el área de estudio, se usó la interpolación Kriging a través de la herramienta Geostatistical Analyst en ArcGIS. Se precedió de la siguiente forma:

1. Se analizó la distribución de la probabilidad que se obtuvo del análisis estadístico en R. Debido a que la distribución no era gaussiana se aplicó transformaciones para acercar los datos a este tipo de distribución. La transformación que se usó fue Box-Cox (*Figura 18*).

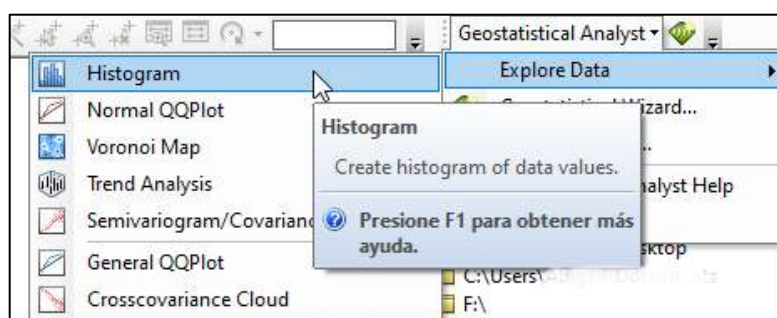


Figura 18. Histograma de la herramienta Geostatistical Analyst
Elaborado: Abigail Navarrete

2. Luego se analizó la tendencia de los datos con la herramienta Trend Analyst del Geostatistical Analyst. Dando como resultado para ambas zonas una tendencia polinómica de segundo orden.
3. Después se procedió a la interpolación Kriging mediante Geostatistical Wizard, ambas zonas tuvieron el mismo proceso, excepto en el modelo de ajuste. Y se usó la media para aquellos valores que se superponen (*Figura 19*).

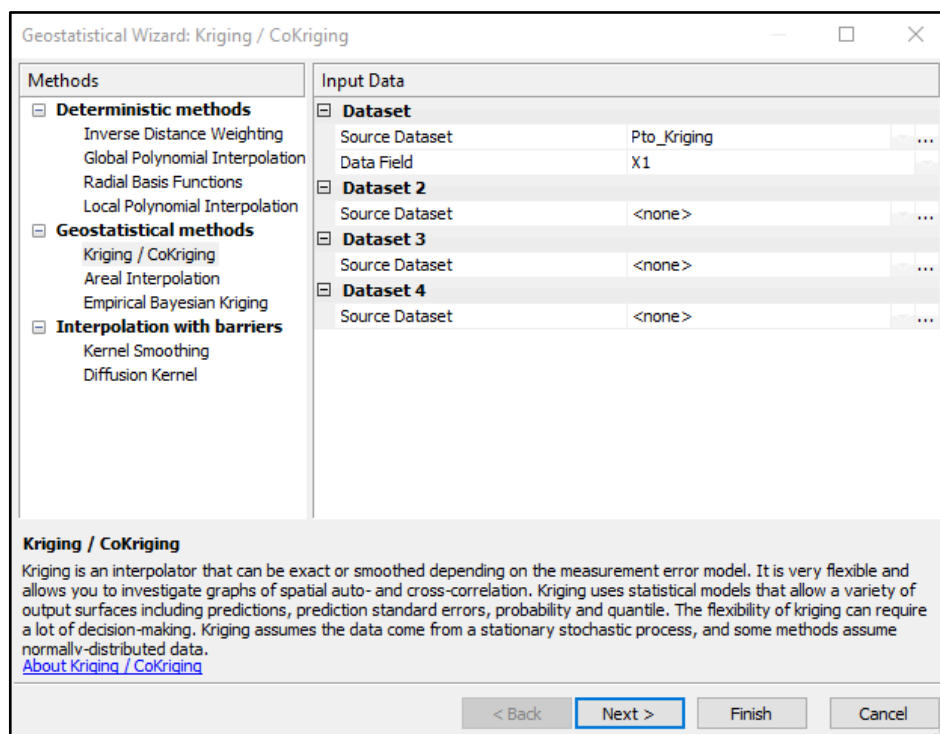


Figura 19. Ventana para interpolación Kriging

Elaborado: Abigail Navarrete

4. En el tipo de Kriging se escogió Ordinario y en el tipo de salida Predicción. Además, aquí se coloca la información que se obtuvo en el paso 1 (*Figura 20*).

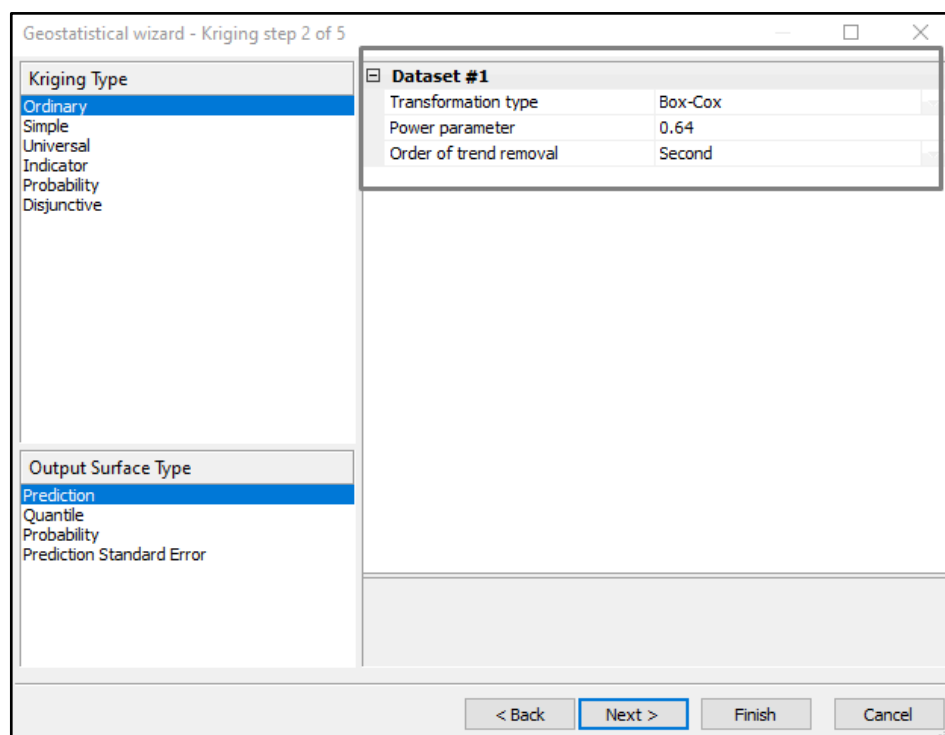


Figura 20. Tipos de Kriging y salida de la interpolación

Elaborado: Abigail Navarrete

5. Luego se escoge el tipo de Kriging para cada zona. Para este estudio, el tipo de Kriging para el Norte fue el Circular y para el Sur se usó el Esférico, estos fueron los modelos en los cuales se obtuvo menor error de predicción. En ambos modelos se usó la opción de optimización, esto ayudó a que la porción ascendente sea más pronunciada y a reducir el Nugget.
6. Luego se exportó el Kriging interpolado como ráster con tamaño celda igual a 500.

4.11. MAPEO

Este corresponde al paso final en el trabajo, para esta tarea se usó el programa ArcMap. Se recortó el ráster interpolado para su respectiva área de estudio y se realizó la gráfica de probabilidad de que ocurran incendios forestales, esto para ambas zonas. Es necesario resaltar que las esquinas de las dos áreas de estudio quedaron sin información de probabilidad, pues no existieron datos para estos sitios, ya sea que no se encontró puntos de incendio o que no se crearon puntos aleatorios de no incendio para estos lugares.

CAPITULO 5 RESULTADOS Y DISCUSIÓN

5.1. INCENDIOS TOTALES

A continuación, se muestra la distribución de los puntos incendios, que se obtuvieron del producto MCD64A1-Área Quemada para cada una de las zonas estudiadas. Después de la depuración de los datos, se registró para las 6 provincias, en el periodo de estudio (2001-2018), un total de 5109 eventos entre enero y diciembre (*Figura 21*).

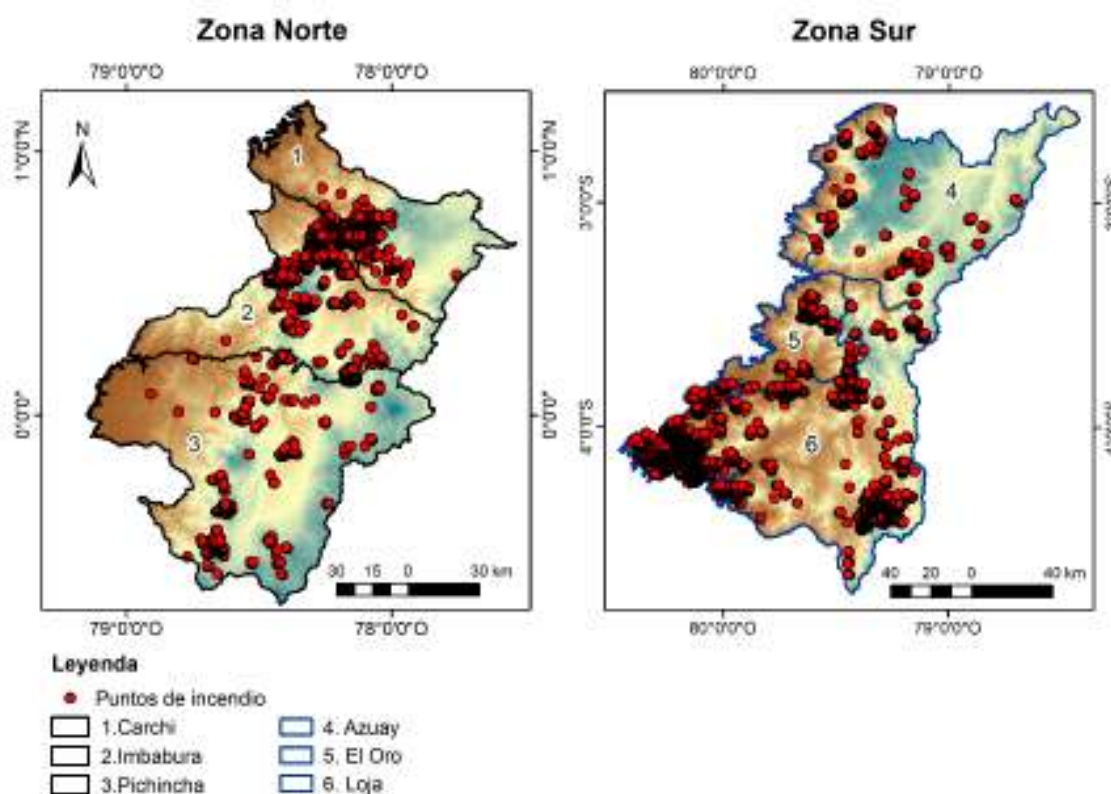


Figura 21. Puntos de Incendios para cada zona
Elaborado: Abigail Navarrete

Los incendios para cada zona se distribuyen de la siguiente forma, para el norte se obtuvo 1545 y para el sur 3564 incendios. La *Figura 22* muestra el porcentaje de incendios en cada zona y se puede observar que en el norte la provincia que más acumula incendios es Imbabura, mientras que para el sur los incendios se agrupan en Loja

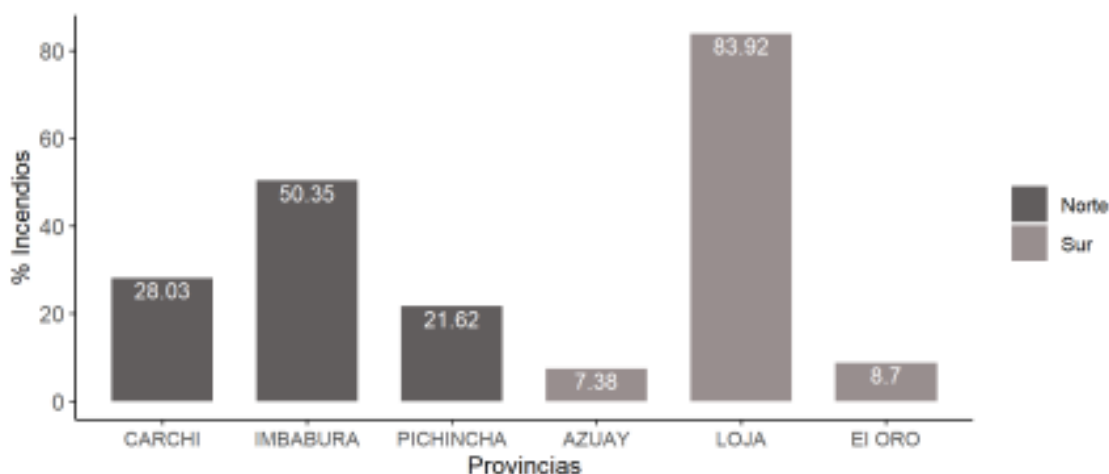


Figura 22. Distribución de incendios por provincias. Temporada enero-diciembre desde el 2001-2018.

Elaborado: Abigail Navarrete

5.2. ANÁLISIS DESCRIPTIVO

El análisis temporal reveló que el comportamiento mensual y anual de los incendios es diferente para cada zona. En la *Figura 23a*, se registra que la presencia de incendios para el norte se concentra entre Junio-Octubre; mientras que, para la zona sur entre Agosto-Diciembre, a estos periodos los llamamos “Temporada de incendios”, estos eventos son los que alimentaron al modelo. Por otro lado, a los meses restantes se los denominó “Temporada antes de los incendios”, estos eventos no se usaron en el modelo porque el número de incidentes fue muy pequeño o nulo. También, el norte registra un pico de incendios en el mes de septiembre, mientras que para la zona sur es el mes de diciembre (*Figura 23a*). Por otro lado, se puede observar en la *Figura 23b* que para la zona norte existe un pico de incendio cada tres años, durante los primeros 9 años pero a partir del 2010 este valle entre picos se va reduciendo. En cambio, en la zona sur se puede notar que, en general, sigue un patrón donde después de un año existe un pico en los eventos. Además, se observa que el 2012 es el año que registra mayor cantidad de incendios para la zona norte; mientras que la zona sur lo registra para el año 2005.

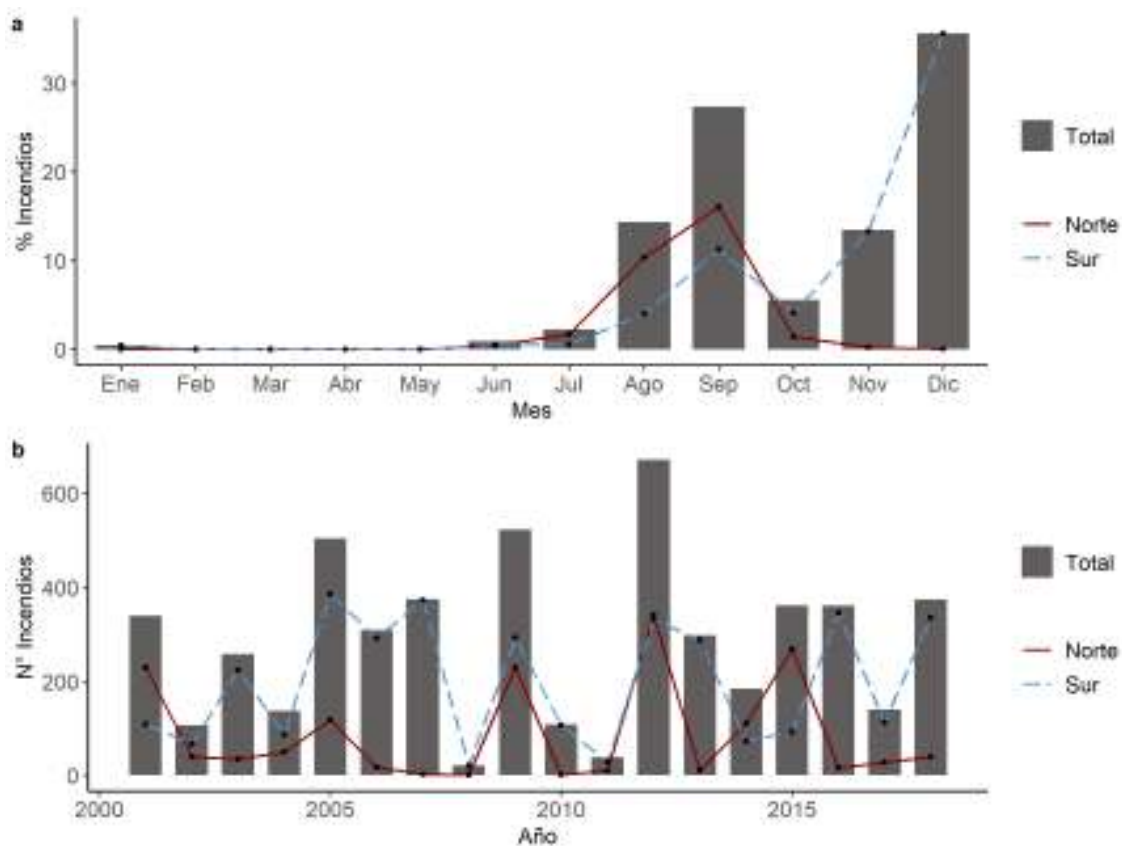


Figura 23. Análisis temporal del comportamiento de incendios. La gráfica **a** es un análisis mensual y la gráfica **b** es un análisis anual. El valor total mencionado en las gráficas es la suma de los incendios del norte y del sur.

Elaborado: Abigail Navarrete

De acuerdo con los datos de Uso de Suelo para el nivel 2 (*Figura 24*) y de áreas protegidas, se obtuvo información, más específica, de donde se producen los incendios. En la zona norte casi el 30% de su total de eventos ocurren en el páramo y en el sur aproximadamente el 35% se producen en el Mosaico Agrícola

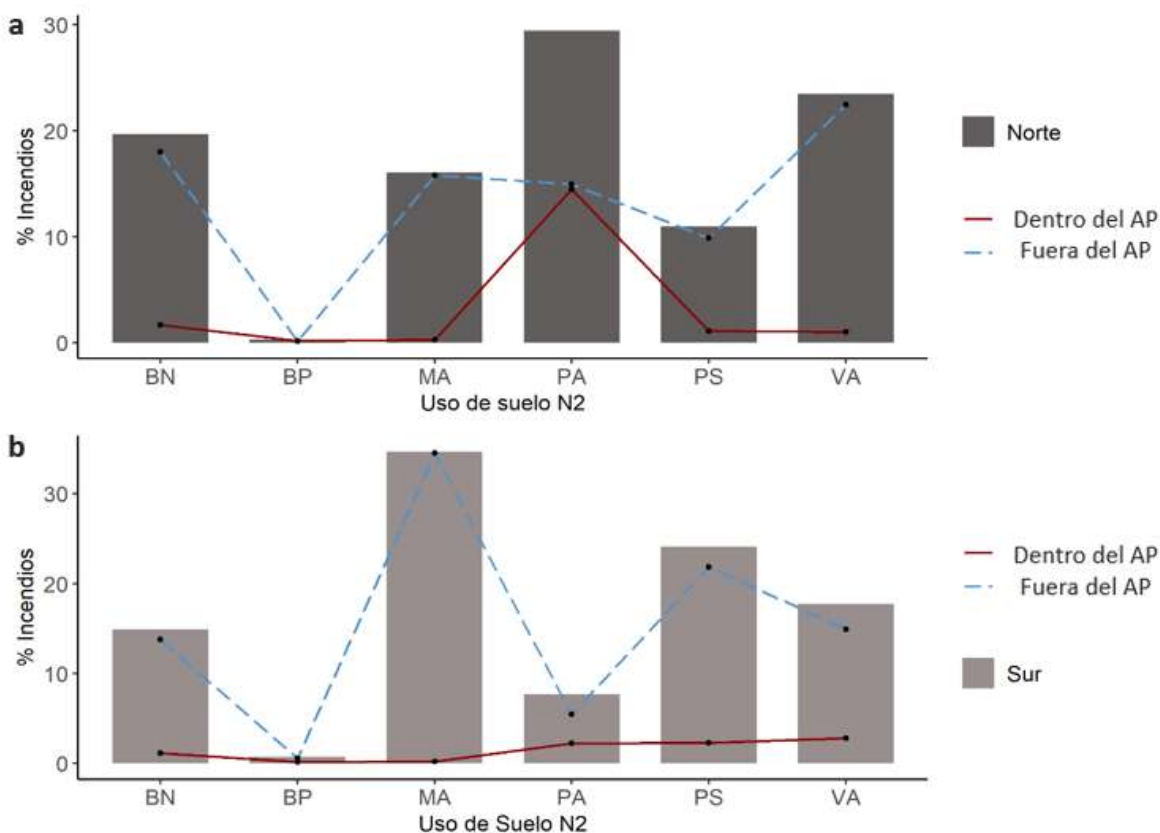


Figura 24. Análisis de la relación del uso de suelo nivel 2 con los incendios. Donde las barras muestran el total de incendios para cada zona; BN es Bosque Nativo, BP es Bosque de Plantación, MA es Mosaico Agrícola, PA es Páramo, PS es Pasto, VA es Vegetación. La gráfica **a** es el análisis para la zona norte y la gráfica **b** es el análisis para la zona sur. **AP** área protegida.

Elaborado: Abigail Navarrete

La **Figura 25** registra, para la zona sur, que la mayor cantidad de eventos se dieron en lugares con poca elevación y poca pendiente, y mientras estas aumentan los incendios disminuyen, esto se puede atribuir a la accesibilidad al terreno en relación a que un poco más de la tercera parte de sucesos se produjeron en Mosaico agropecuario. Por otro lado, en la zona norte no se registra gran cantidad de acontecimientos en zonas bajas, más bien se acumulan en elevación entre 2000 y 4000m.s.n.m; esto podría deberse a que las ciudades capitales de esta provincia se encuentran, aproximadamente, entre los 2200 a 3000m.s.n.m; también hay que recordar que casi la tercera parte de los incendios se produjeron en los páramos. En la **Figura 25c** se puede observar que los incendios se producen, especialmente, en direcciones como Suroeste, Oeste y Noroeste. Mientras que, en el norte se producen en direcciones como Noreste, Oeste y Noroeste.

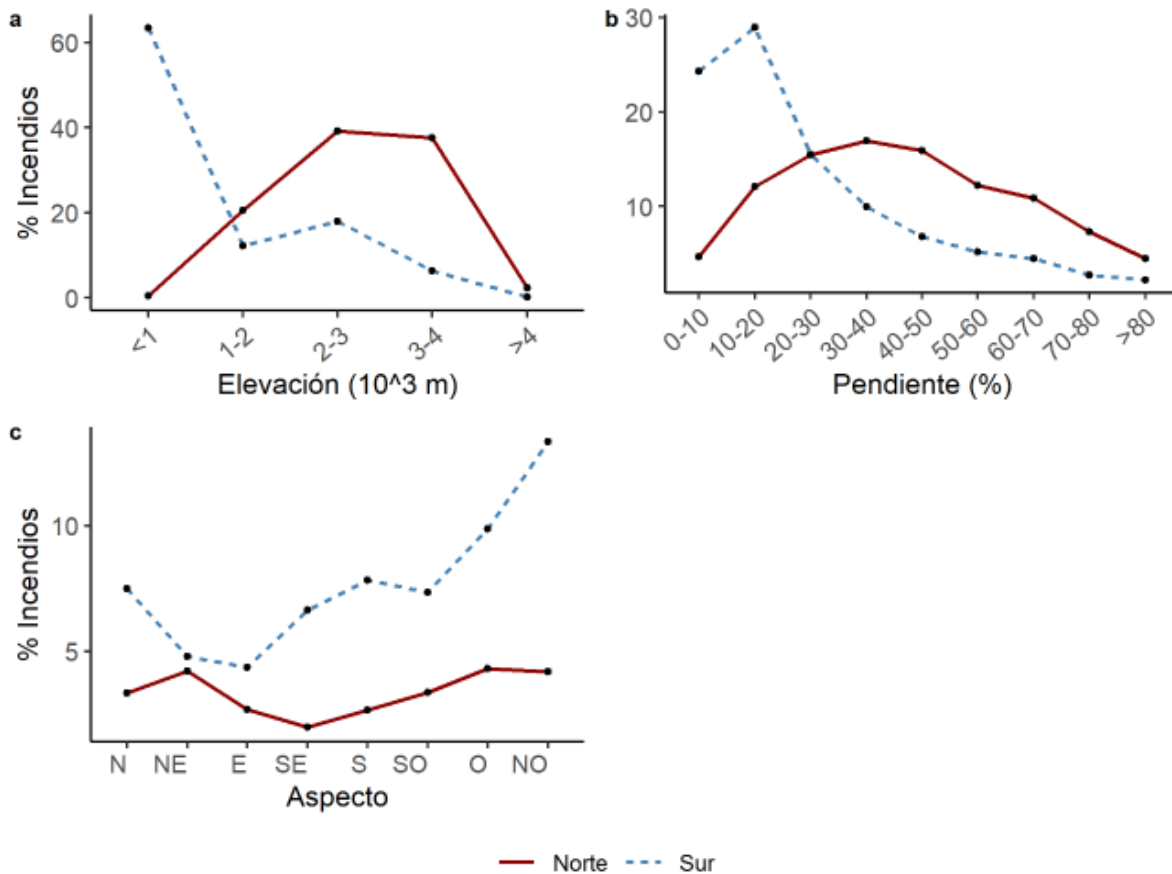


Figura 25. Relación entre variables topográficas e incendios. La figura (a) es la relación con elevación, (b) relación con la pendiente, (c) relación con Aspecto
Elaborado: Abigail Navarrete

En la *Figura 26* parte a y b se representa la relación de los incendios en función de las variables de infraestructura. En ambas figuras el comportamiento es similar, donde los incendios, en general, disminuyen conforme aumenta la distancia. La *Figura 26a* muestra que más del 20% de los eventos se dan en distancias entre 5 a 10 Km de a los centros poblados. Además, se observa que para la *Figura 26b* el sur poseen una pendiente más pronunciada, es decir los cambios no son suaves y de la misma forma que en la *Figura a*, más del 50% los incendios se dan en distancias de 500m, es decir muy cerca de las vías.

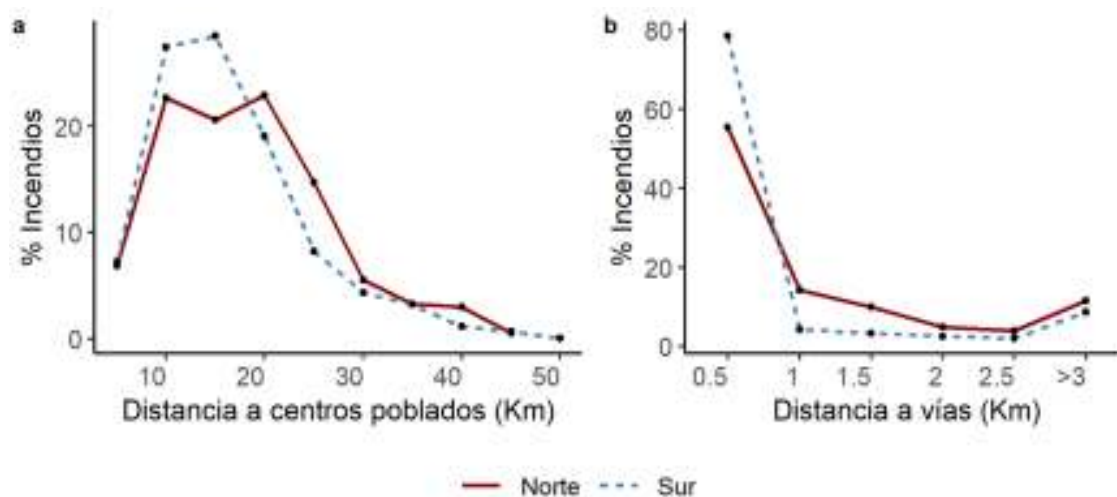


Figura 26. Relación entre variables de infraestructura e incendios. (a) Relación con Distancia a centros poblados, (b) Relación con distancia a vías

Elaborado: Abigail Navarrete

Con respecto a la relación entre incendios y variables climáticas como Radiación, Precipitación media anual, Temperatura media anual, y relación con la variable biológica VCI, se obtuvo la **Figura 27** a, b, c y d. En la radiación media anual (irradiación horizontal global) la **Figura a** muestra distribuciones similares pero opuestas para las dos áreas de estudio, donde el 56% de los incendios se acumula en un rango de 1500-1750 kWh/m² para la zona norte; en tanto que para la zona sur el 75% de los incendios se presentan en un rango de 1750 a 2000 kWh/m². También, se observa que el 49% de los puntos de incendios de la zona norte se ubicaron entre 1000 y 1500 mm de precipitación media anual; mientras que, para la zona sur, aproximadamente el 77% de los eventos ocurrieron en un rango de 500 a 1000 mm. Además, la gráfica de Temperatura media anual revela, que al igual que la radiación, las distribuciones son similares pero opuestas entre las dos zonas. Los incendios para la zona norte concentran el mayor número de incendios entre los 15 y 20°C acumulando el 49% de eventos en este rango; mientras que para la zona sur el 52% de incendios se produjeron entre los 25 y 30°C. La **Figura d** describe la relación entre la variable biológica VCI y los incendios, en la zona norte cerca del 30% de los eventos se producen entre el rango de 70 a 80% VCI, rango que está en un nivel de sequía ligera; mientras que para el sur el 31% de los incendios se produjeron en un VCI entre 50-60%, representando por un nivel de sequía moderado.

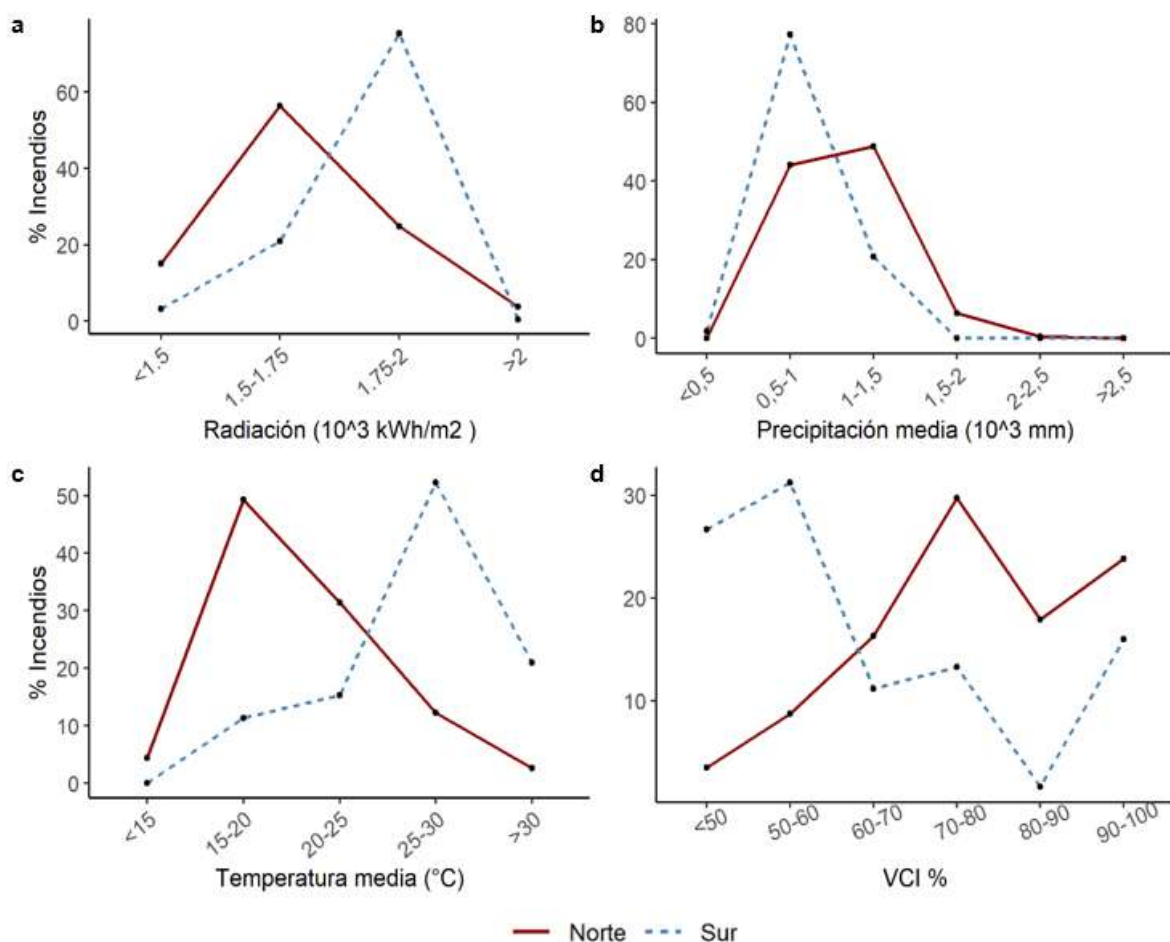


Figura 27. Relación entre variables climáticas-biológicas e incendios, (a) Relación con la radiación (Irradiación Horizontal Global), (b) Relación con precipitación media anual, (c) Relación con la temperatura media, (d) Relación con el Índice de condición vegetal.

Elaborado: Abigail Navarrete

5.3. ANÁLISIS ESTADÍSTICO

En esta sección se realizó una prueba de hipótesis paramétrica a las variables cualitativas y una prueba de hipótesis no paramétrica a las variables cuantitativas, la significancia de las variables se obtuvo a través de los índices de los Árboles de Decisión MDG y MDA, y la prueba de multicolinealidad. Los resultados de estas pruebas, se presentan abajo.

Las pruebas de hipótesis no paramétrica y paramétrica se aplicaron a todas las variables según como correspondía, por lo que aquí solo se presenta un ejemplo de las mismas, el resto de los resultados se muestran el **Anexo (1, 2,3)**. La **Figura 28** muestra la distribución de los incendios (verde) y los no incendios (rojo) en función de la distancia a vías para la zona norte, y se observa que estas

distribuciones no son iguales, por lo tanto esta información si podría contribuir al modelo porque el comportamiento de la variable dicotómica es diferente y por ello se rechaza la hipótesis nula de igualdad (celeste, *Figura 28*).

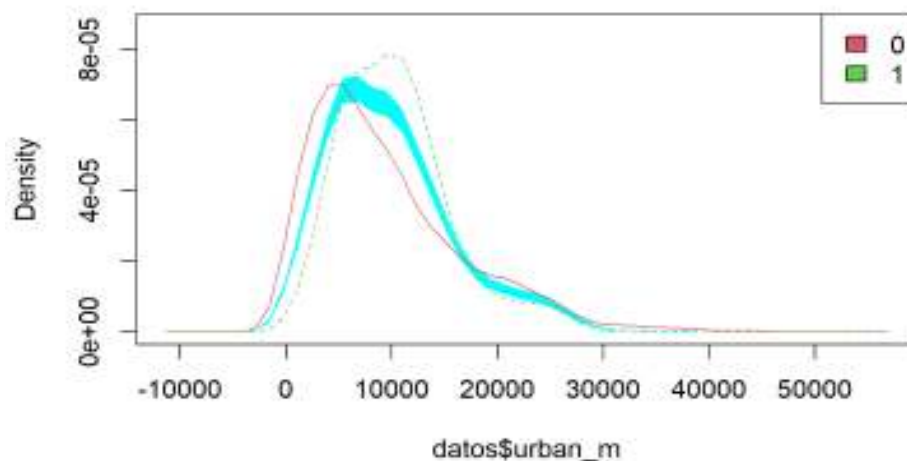


Figura 28. Resultados de la prueba No Paramétrica para las variables Distancia a centros poblados Norte. Los incendios representados por la línea punteada verde, No Incendios representados por la línea roja y la banda celeste es una banda de referencia de la hipótesis nula.

Elaborado: Abigail Navarrete

La prueba paramétrica muestran resultados en función del p_value para un 95% de confianza por lo que los valores menores a 0.05 en todas las categorías de la variable significará que se puede incluir a la variable en el modelo. En la **Tabla 6** se observa la prueba de hipótesis paramétrica para la variable Aspecto de la zona norte, y los valores de p_value presentan que solo la categoría Noreste y Sureste entran en el parámetro de aceptación, por lo tanto está variable no aporta con información al modelo porque se está aceptando la hipótesis nula para el resto de categorías.

Tabla 6.

Resultados de la prueba paramétrica para la variable Aspecto de la zona Norte

	Aspect_1 <chr>	p_value <dbl>	prop_incendios <dbl>
1	NE	0.0111	0.455
2	SE	0.0155	0.334
3	S	0.237	0.370
4	W	0.326	0.421
5	NW	0.341	0.381
6	SW	0.630	0.412
7	N	0.932	0.403
8	E	1	0.401

Estas pruebas nos permitieron descartar las variables Distancias a cuerpos de agua y Aspecto para el sur, y para el norte las mismas variables más Distancia a vías.

Las pruebas de significancia revelaron los siguientes resultados. En la *Figura 29* se muestra en la parte superior las variables con mayor importancia y conforme desciende, la importancia también disminuye. También se puede observar que las variables con mayor importancia son las variables climáticas, precipitación y radiación, la variable de infraestructura más importante es Distancia a centros poblados, entre las variables topográficas con mayor importancia están la elevación (dem_m) y la pendiente (slope), por otro lado las variables biológicas tienen la menor importancia entre el resto de variables.

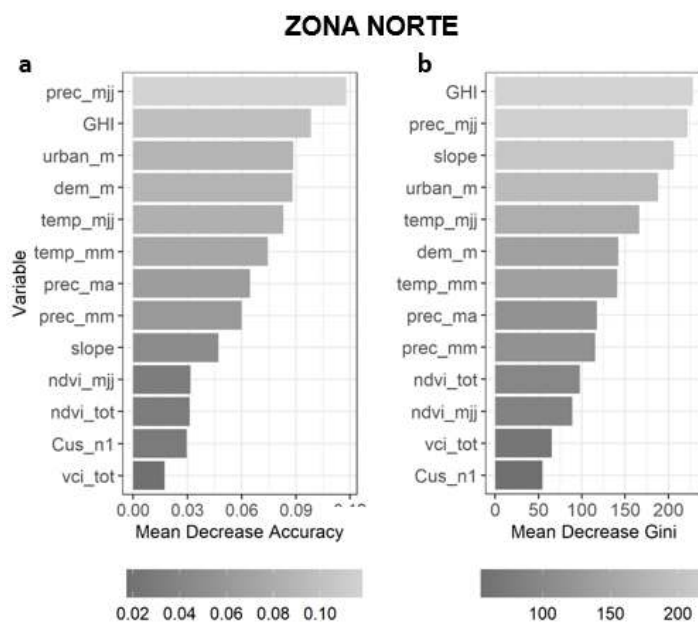


Figura 29. Resultados de las pruebas MDA (a) y MDG (b) para el Norte. Donde prec_mjj es la precipitación media de noviembre a mayo (temporada antes de incendios), GHI es la Irradiación horizontal global, urban_m es la distancia a centros poblados, temp_mjj es la temperatura media de la temporada antes de incendios, temp_mm temperatura media mensual, prec_ma precipitación media anual, prec_mm precipitación media mensual, slope es pendiente, ndvi_mjj es el Índice de Vegetación de Diferencia Normalizada de la temporada antes de incendios, ndvi_tot Índice de Vegetación de Diferencia Normalizada del periodo de estudio, Cus_n1 es la cobertura de uso de suelo nivel 1 representada en números, vci_tot es el Índice de la Condición de la Vegetación del periodo de estudio.

Elaborado: Abigail Navarrete

En la *Figura 30* se puede observar que la variable con mayor importancia es la de topografía - elevación, luego están las variables climáticas temperatura, precipitación y radiación, entre las variables biológicas con mayor importancia está la cobertura de uso

de suelo (Cus_n1), por otro lado, las variables de infraestructura tienen la menor importancia entre el resto de las variables.

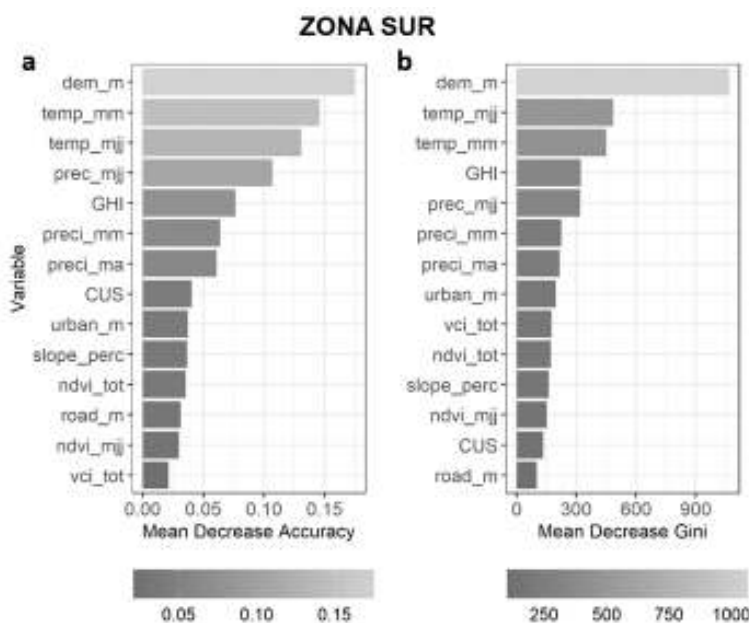


Figura 30. Resultados de las pruebas MDA (a) y MDG (b) para el Sur. Donde prec_mjj es la precipitación media de enero a julio (temporada antes de incendios), GHI es la Irradiación horizontal global, urban_m es la distancia a centros poblados, temp_mjj es la temperatura media de la temporada antes de incendios, temp_mm temperatura media mensual, prec_ma precipitación media anual, prec_mm precipitación media mensual, slope es pendiente, ndvi_mjj es el Índice de Vegetación de Diferencia Normalizada de la temporada antes de incendios, ndvi_tot Índice de Vegetación de Diferencia Normalizada del periodo de estudio, Cus_n1 es la cobertura de uso de suelo nivel 1 representada en números, vci_tot es el Índice de la Condición de la Vegetación del periodo de estudio.

Elaborado: Abigail Navarrete

La correlación entre las variables independientes se muestra en la **Figura 31**. Mientras los colores rojo y azul disminuyen en su intensidad la correlación entre dos variables es menos fuerte; el color azul muestra una relación positiva, el color rojo es una relación negativa y cero representa falta de relación entre variables. Se puede observar que, para la zona norte, existen casillas que superan el 90% de correlación, solamente en la dirección positiva (azul). Sin embargo, esta relación se da entre variables con la misma información, por ejemplo: prec_mm (precipitación media mensual) vs prec_ma (precipitación media anual). Por otro lado, el resto de las variables tienen correlaciones inferiores a 65% en ambas direcciones.

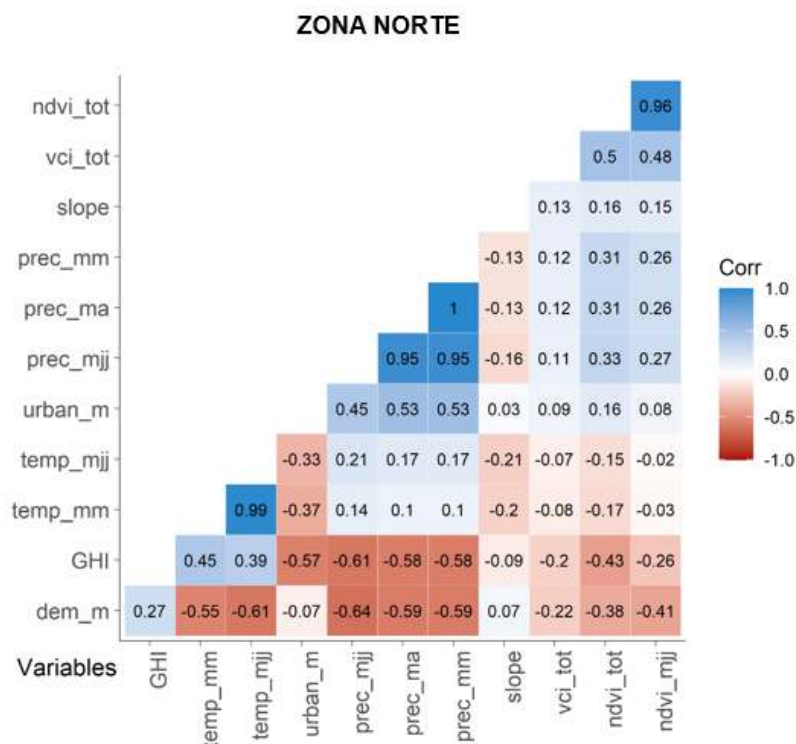


Figura 31. Resultados de la prueba de Multicolinealidad para el Norte. Donde prec_mjj es la precipitación media de la temporada antes de incendios (nov-may), GHI es la Irradiación horizontal global, urban_m es la distancia a centros poblados, temp_mjj es la temperatura media de la temporada antes de incendios, temp_mm temperatura media mensual, prec_ma precipitación media anual, prec_mm precipitación media mensual, slope es pendiente, ndvi_mjj es el Índice de Vegetación de Diferencia Normalizada de la temporada antes de incendios, ndvi_tot Índice de Vegetación de Diferencia Normalizada del periodo de estudio, vci_tot es el Índice de la Condición de la Vegetación del periodo de estudio.

Elaborado: Abigail Navarrete

En la zona sur (*Figura 32*), se identificaron casillas que superan el 80% de correlación en ambas direcciones (rojo y azul). Según lo mencionado anteriormente, encontramos que hay correlación entre variables que tienen la misma información. También se encontró correlación negativa entre temperatura y elevación (temp_mm, temp_mjj y dem_m), y correlación positiva entre temperatura y radiación (temp_mm, temp_mjj y GHI). Sin embargo, en función de la *Figura 32* las variables elevación, temperatura y radiación son significativas dentro del modelo.

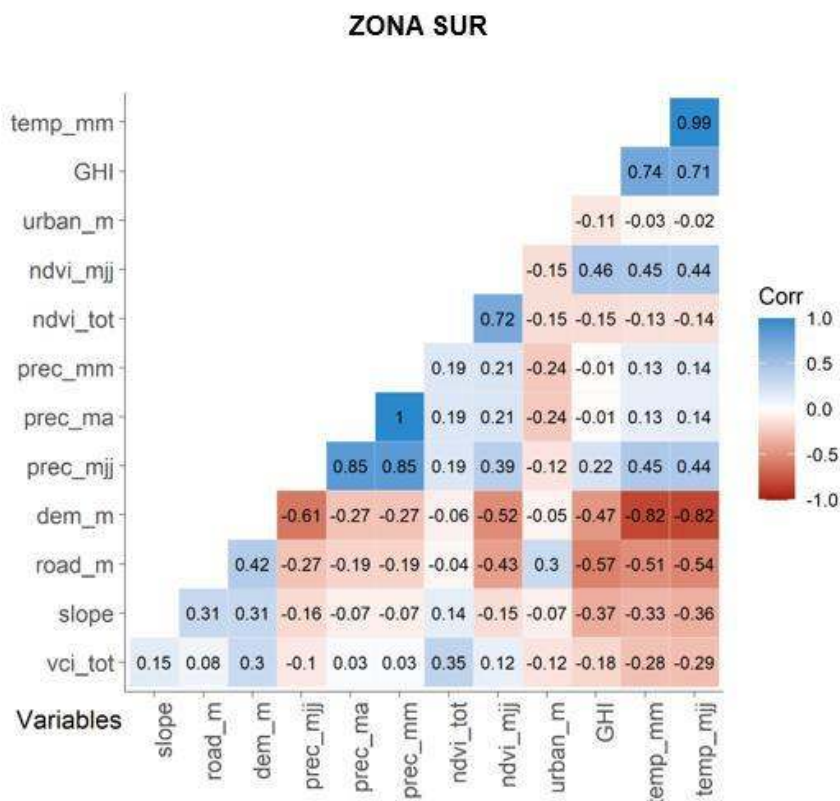


Figura 32. Resultados de la prueba de Multicolinealidad para el Sur. Donde prec_mjj es la precipitación media de la temporada antes de incendios (ener-jul), GHI es la Irradiación horizontal global, urban_m es la distancia a centros poblados, temp_mjj es la temperatura media de la temporada antes de incendios, temp_mm temperatura media mensual, prec_ma precipitación media anual, prec_mm precipitación media mensual, slope es pendiente, ndvi_mjj es el Índice de Vegetación de Diferencia Normalizada de la temporada antes de incendios, ndvi_tot Índice de Vegetación de Diferencia Normalizada del periodo de estudio, vci_tot es el Índice de la Condición de la Vegetación del periodo de estudio.

Elaborado: Abigail Navarrete

5.4.VARIABLES QUE SE USARON EN EL MODELO

Luego de analizar la significancia de las variables. Las combinaciones (receta) para la creación de los modelos que obtuvieron mejores resultados, para cada zona, fueron:

- **Receta Norte:** $CUS + prec_mjj + ndvi_mjj + slope + urban_m + GHI + vci_tot$.
- **Receta Sur:** $CUS + dem_m + GHI + vci_tot + urban_m + ndvi_tot + prec_mm$

5.5.RESULTADOS DEL MODELO LOGÍSTICO

En las **Tabla 7 y 8** se presenta un resumen de los resultados del modelo logístico para cada zona. El coeficiente representa la relación que hay entre la presencia o no de incendios, que sería la variable dependiente y las variables independientes. El signo representa si está relación es inversa o directamente proporcional. El error estándar señala la desviación estándar que existe de la estimación del coeficiente. El VIF evalúa la multicolinealidad. La prueba de Wald indica si el coeficiente de la variable es o no significativo para una significancia del 95%.

5.5.1. ZONA NORTE

La **Tabla 7** presenta los resultados del modelo para las ocho variables predictivas de la zona norte. El VIF revela que todas las variables están dentro del rango aceptable de multicolinealidad, es decir entre 1 a 4. Además, la prueba de Wald muestra que todas las variables están por debajo del 5%, por lo tanto los coeficientes son significativos en el modelo. Se puede concluir que todas las variables son significativas y no tienen multicolinealidad. La variable Uso de Suelo (Bosque, Tierra Agropecuaria) y la precipitación (prec_mjj) son las 3 variables con mayor influencia dentro del modelo y además tienen una relación inversamente proporcional. Las siguiente 5 variables tienen una relación directamente proporcional con la variable respuesta; y en función de la importancia dentro del modelo tenemos lo siguiente, índice de vegetación de diferencia normalizada antes de los incendios (ndvi_mjj), pendiente (slope), distancia a centros poblados (urban_m), irradiación horizontal global (GHI), índice de condición vegetal (vci_tot). De acuerdo con la regresión logística la importancia de las variables que arroja el modelo difiere totalmente a la importancia de los árboles de decisión.

Tabla 7.

Variables resultantes del modelo logit en la zona Norte. Cobertura de uso de suelo (Bosque, Tierra Agropecuaria), Precipitación media antes de los incendios nov-may (prec_mjj), Índice de vegetación de diferencia normalizada media antes de los incendios nov-may (ndvi_mjj), Pendiente (slope), Distancia a centros poblados(urban_m), Irradiación horizontal global (GHI), Índice de condición vegetal del periodo de estudio(vci_tot)

Variable	Coeficiente	Signo	Error Estándar	VIF	Wald test	Importancia		
						GLM	Accu	Gini
Bosque	1.649	-	0.167	1.989	<0.001	1	5	7
T. Agrope	1.356	-	0.143	1.857	<0.001	2	6	8
prec_mjj	0.803	-	0.084	1.962	<0.001	3	1	1
ndvi_mjj	0.777	+	0.074	2.004	<0.001	4	7	5
slope	0.731	+	0.059	1.161	<0.001	5	4	4
urban_m	0.613	+	0.071	1.761	<0.001	6	2	3
GHI	0.386	+	0.081	2.963	<0.001	7	3	2
vci_tot	0.230	+	0.059	1.256	<0.001	8	8	6

5.5.2. ZONA SUR

La **Tabla 8** muestra los resultados del modelo para las ocho variables predictivas de la zona sur. El VIF revela que todas las variables están dentro del rango aceptable de multicolinealidad (1-4). Además, la prueba de Wald muestra que los coeficientes están por debajo del 5%, por lo tanto estos son significativos en el modelo. Entonces, podemos concluir que todas las variables son significativas y que no tienen multicolinealidad fuerte. La cobertura de uso de suelo ocupa el primer y tercer lugar con mayor influencia en el modelo, el segundo lugar es de la variable topográfica (dem_m), el cuarto y octavo lugar lo ocupan las variables climáticas radiación (GHI) y precipitación media mensual (prec_mm), luego están las variables biológicas que son el índice de condición vegetal (vci_tot) y el índice de vegetación de diferencia normalizada (ndvi_tot), que ocupan el quinto y séptimo puesto respectivamente, y la distancia a centros poblados (urban_m) está el sexto lugar

dentro de la escala de influencia del modelo. Se observa que la importancia del modelo y la importancia de los árboles de decisión también difieren para esta zona.

Tabla 8.

Variables resultantes del modelo logit en la zona Sur. Elevación (dem_m), Temperatura media mensual (temp_mm), Irradiación global horizontal (GHI), Índice de condición vegetal dentro del período de estudio (vci_tot), Distancia a centros poblados (urban_m), Precipitación media mensual (prec_mm), Índice de vegetación de diferencia normalizada de todo el período (ndvi_tot).

Variable	Coeficiente	Signo	Error Estándar	VIF	Wald test	Importancia		
						GLM	Accu	Gini
Bosque	1.461	-	0.114	1.749	<0.001	1	4	7
dem_m	1.011	-	0.050	1.755	<0.001	2	1	1
T.agrope	0.802	-	0.090	1.802	<0.001	3	5	8
GHI	0.504	+	0.040	1.262	<0.001	4	2	2
vci_tot	0.319	-	0.036	1.209	<0.001	5	8	5
urban_m	0.227	+	0.038	1.230	<0.001	6	6	4
ndvi_tot	0.170	+	0.041	1.369	<0.001	7	7	6
prec_mm	0.128		0.039	1.334	<0.001	8	3	3

5.6.RESULTADOS DEL MODELO

La *Figura 33* presenta la relación que existe entre Verdaderos Positivos y Falsos Positivos de la matriz de confusión, donde la sensibilidad está relacionada con la fracción de VP y la especificidad con la fracción de FP. Entonces, se puede observar que ambas curvas se encuentran sobre y alejándose de la línea de referencia, esto implica que el número de verdaderos positivos no son iguales al número de Falsos Positivos.

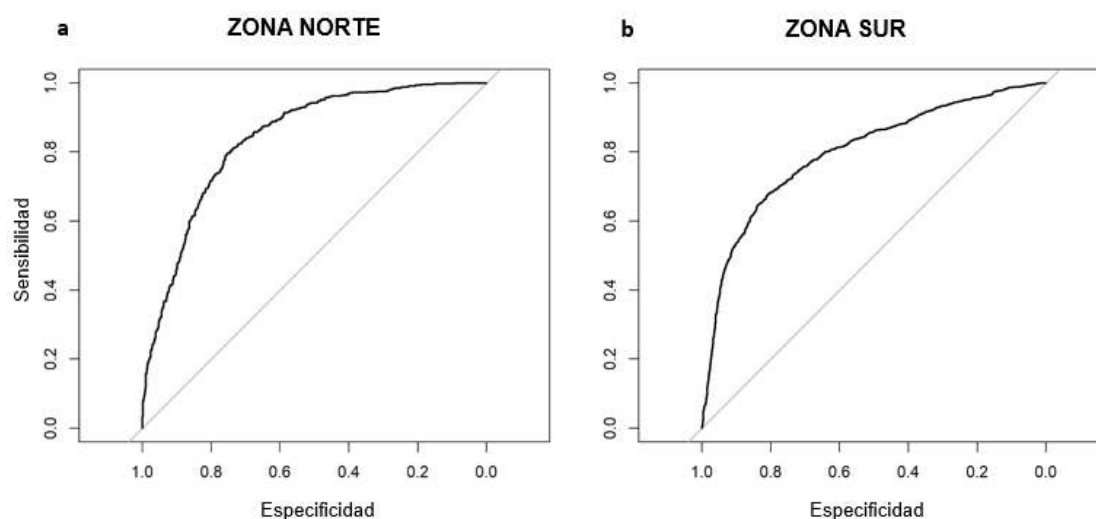


Figura 33. Curva ROC
Elaborado: Abigail Navarrete

La **Tabla 9** revela los resultados globales del modelo. La clasificación se da para las etapas de entrenamiento (Entrena) y prueba, donde se obtuvo la siguiente precisión, para el norte tiene un valor de 0.75 y el sur tiene 0.76, esto para ambas etapas. Los resultados de la matriz de confusión revelan que, para ambas zonas, existen más aciertos en la clasificación de los no Incendios que en la de los incendios. El Pseudo R^2 indica qué tan bien explica el modelo a la variable respuesta, es así como para el norte se obtuvo un 77% con un error del 23% y para el sur un 76% con un error del 24%. Finalmente, al final de la tabla encontramos los resultados del área bajo la curva ROC (AUC) demuestra que la exactitud global de la prueba está en el rango admitido por intervalo de confianza. Esto demuestra que este modelo es robusto y tiene un buen ajuste.

Tabla 9.
Resultados del modelo

Zona	Precisión		Validación Cruzada	Matriz de confusión		Error	Valor AUC	
	Entrena	Prueba	Precisión	0	1		Media 95% CI	
Norte	0.745	0.745	0.745	0.822	0.685	23.3%	0.830	0.819-0.859
Sur	0.764	0.750	0.757	0.810	0.675	24.4%	0.799	0.783-0.846

5.7.RESULTADOS DE LA INTERPOLACIÓN

En la *Figura 34* se observa la transformación de la distribución de las probabilidades que se obtuvieron del análisis estadístico. Los gráficos a y c muestran la distribución inicial de los datos de probabilidad Norte y Sur, los gráficos b y d es la transformación Box-Cox con parámetro 0.64 y 0.35, respectivamente, estas fueron las transformaciones que más se aproximaron a una distribución normal.

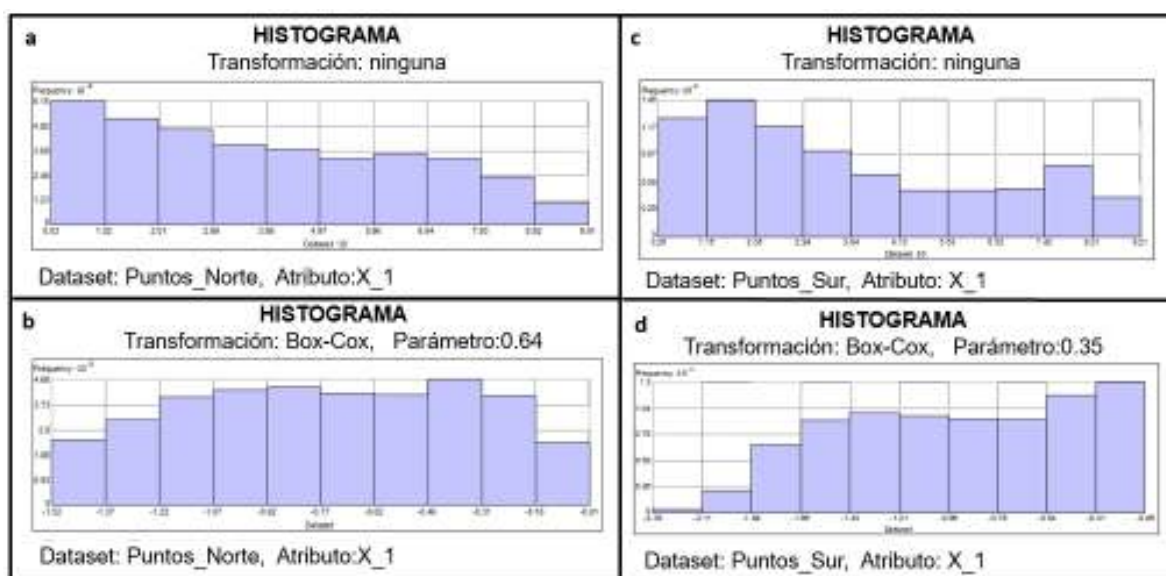


Figura 34. Histograma para transformación de distribución.
Elaborado: Abigail Navarrete

Dentro de la **Tabla 10** se puede encontrar los resultados de los modelos que se escogió para interpolar la probabilidad de incendio de cada zona. El resultado de la interpolación es el cálculo de un valor promedio ponderado de las muestras circundantes. Entonces, el error va a ser igual a lo modelado menos lo observado (muestras). Por lo tanto, el modelo Circular para el norte ha acertado en el 86.2%, mientras que el modelo exponencial para el sur ha acertado en 93%.

Resultados del método Kriging de ArcGIS

Tabla 10.
Resultados del Kriging

	Error de predicción (root-mean-square)	Modelo
Norte	0.138	Circular
Sur	0.070	Esférico

5.8.MAPA DE PROBABILIDAD DE OCURRENCIA DE INCENDIOS

Se presenta un mapa de probabilidades de ocurrencia de incendios forestales para el periodo 2001-2018, es el resultado final de la interpolación Kriging de los resultados del modelo lineal generalizado. Las probabilidades fueron clasificadas en niveles de riesgo, el primero y el que tiene menor impacto es el riesgo Muy bajo, le sigue el Bajo, y desde aquí el riesgo comienza a tener mayor impacto en el área de estudio (Medio, Alto y Muy alto). A continuación, se presenta un mapa para cada zona.

La *Figura 35* es el mapa de probabilidad de riesgo de incendio para la zona norte. En este mapa la clasificación dominante en el territorio es la Muy baja (0-0.2) y Baja (0.2-0.4) ocupando el 36 y 33% del territorio respectivamente, el riesgo Medio (0.4-0.6) ocupa el 20%, el 11% corresponde a Alta (0.6-0.8) probabilidad y solamente 1% del área de estudio tiene Muy alta (0.8-0.95) probabilidad de ocurrencia de incendio. También, se puede observar que la provincia que tiene mayor frecuencia de Media – Alta - Muy Alta probabilidad de ocurrencia de incendios es Imbabura, sobre todo se presentan en cantones como Cotacachi, San Miguel de Urququí e Ibarra. Además, en la Provincia del Carchi las probabilidades entre Media a Muy alta se presentan en los cantones de El Ángel, Mira y en la zona rural de Tulcán. Mientras que, para Pichincha estas probabilidades se concentran en el norte y sureste de Quito, centro de Cayambe y en el cantón Mejía.

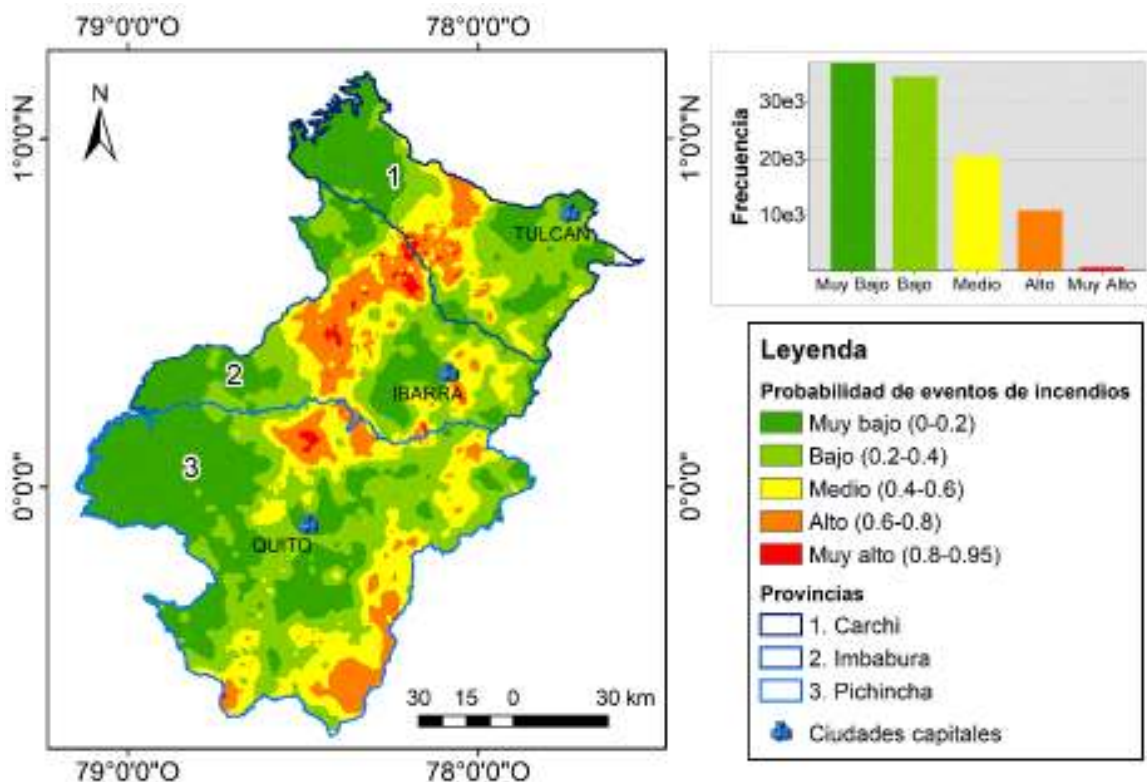


Figura 35. Mapa de probabilidad para la Zona Norte
Elaborado: Abigail Navarrete

En la **Tabla 11**, se analizó la relación entre las tres probabilidades más influyentes (Medio, Alto, Muy Alto) y la cobertura de uso de suelo, y se encontró que la probabilidad Muy Alta, tiene las siguientes proporciones, 66% se encuentra en Vegetación Arbustiva y Herbácea, 18% en Bosque y 16% en Tierra Agropecuaria. Para la probabilidad Alta se encontró que 53% es para Vegetación Arbustiva y Herbácea, Bosque 25%, Tierra Agropecuaria 20% y Otros 2%. Mientras que, para el riesgo Medio la Vegetación Arbustiva y Herbácea tiene 43%, Tierra Agropecuaria 37%, Bosque 18% y Otros 2%. Esto demuestra que la Vegetación Arbustiva y Herbácea está más expuesta a los incendios.

Tabla 11.

Resumen de la probabilidad de incendio, Zona Norte.

Probabilidad	% territorio ocupado	Bosque	Veg. Arb	Tierra Agrop	Otros
Muy baja	36	-	-	-	-
Baja	32	-	-	-	-
Media	20	18	43	37	2
Alta	11	25	53	20	2
Muy Alta	1	18	66	16	0

La *Figura 36* es el mapa de probabilidad de riesgo de incendio para la zona sur. En este mapa la clasificación dominante es la Muy baja (0-0.2) ocupando el 41% del territorio, le sigue la categoría Baja (0.2-0.4) con el 28%, el riesgo Medio (0.4-0.6) representa el 19%, mientras el 10% corresponde a Alta (0.6-0.8) probabilidad y solamente 1% del área de estudio tiene Muy alta (0.8-0.95) probabilidad de ocurrencia de incendio. Se observa que Loja es la única provincia que tiene un riesgo Muy alto, mientras que para Azuay y El Oro los riesgos Medio y Alto, se concentran en los límites junto a Loja.

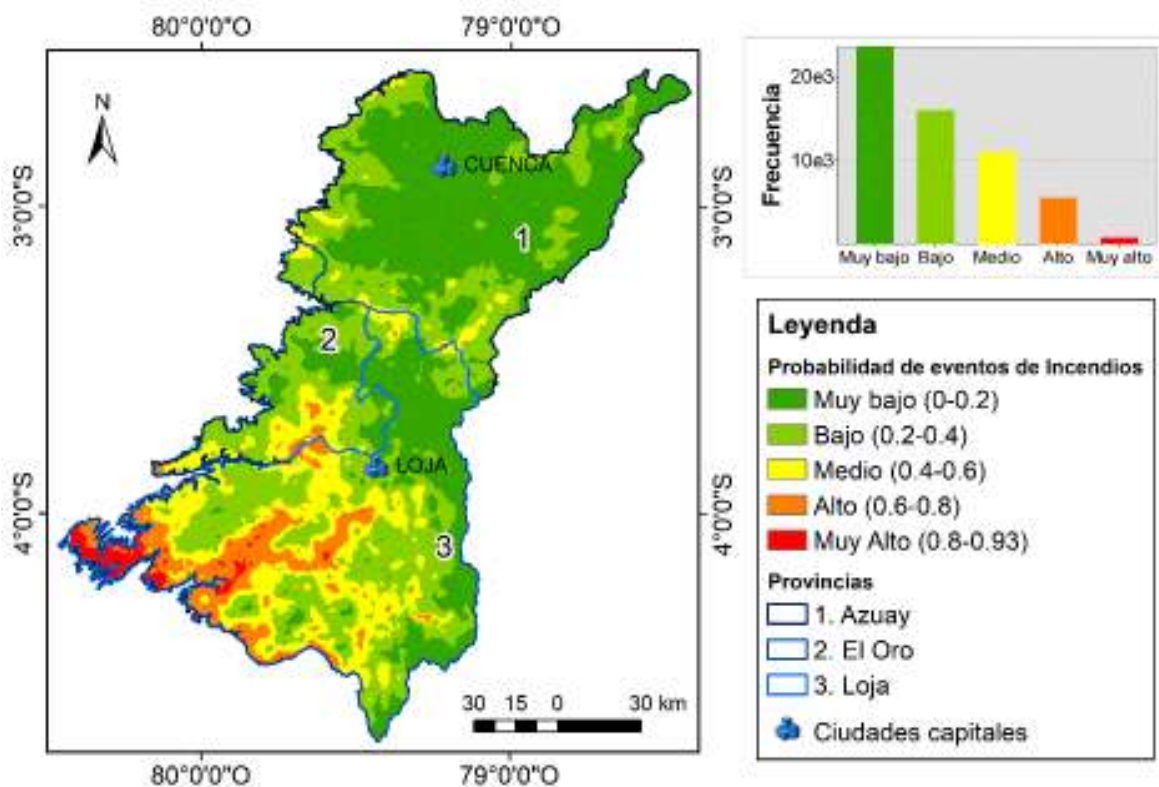


Figura 36. Mapa de probabilidad para la Zona Sur
Elaborado: Abigail Navarrete

Dentro del AE, el nivel Muy Alto se presentó en las siguientes categorías de uso de suelo, Tierra Agropecuaria un 42%, Bosque 33%, Vegetación Arbustiva y Herbácea 24% y Otros 1%. Mientras que, para la probabilidad Alta los porcentajes se distribuyeron así, Vegetación Arbustiva y Herbácea 40%, Tierra Agropecuaria 35%, Bosque 20%, Otros 5%. Finalmente, para la probabilidad Media, la Tierra agropecuaria concentra el 44% de esta probabilidad, sigue la Vegetación Arbustiva y Herbácea con 29%, Bosque 25% y Otros 2%. Esto demuestra que la Tierra Agropecuaria tiene mayor probabilidad, Media y Muy Alta, de sufrir incendios forestales (**Tabla 12**).

Tabla 12.*Resumen de la probabilidad de incendio, Zona Sur.*

Probabilidad	% territorio ocupado	Bosque	Veg. Arb	Tierra Agrop.	Otros
Muy baja	42	-	-	-	-
Baja	28	-	-	-	-
Media	19	25	29	44	2
Alta	10	20	40	35	5
Muy Alta	1	33	24	42	1

5.9.DISCUSIÓN

No existen otros estudios que analicen los patrones espaciales y temporales para esta clase de eventos (incendios forestales), específicamente en estas dos zonas. Sin embargo, hay estudios sobre incendios forestales sobre todo en el cantón Quito (Pichincha) y en algunos otros lugares del país. En este estudio se registró una temporada de incendios, para la zona norte, que abarca de junio a octubre, y el mayor número de eventos se produjeron en agosto y septiembre; Pazmiño (2019) mencionan que los incendios forestales en Quito se registran principalmente en julio, agosto y septiembre dentro del periodo 1997-2012. También en este trabajo, para la zona sur, se observó que la temporada de incendios se da entre agosto - diciembre; Rodas (2015) describe en su trabajo que, los meses con mayor número de eventos en la cuenca de Paute (Azuay) fueron agosto, septiembre y octubre.

En este estudio también se pudo observar una diferencia en el comportamiento de los incendios en relación con la Cobertura de Uso de Suelo (nivel1). En el norte el 53% del total de incendios se produjeron en la Vegetación Arbustiva (matorrales) y Herbácea (páramos). Mientras que, en la zona sur el 58 % de eventos se produjeron en la Tierra Agropecuaria. Esto también es mencionado por Estacio y Narvárez (2012) y Rodas (2015), respectivamente, aunque no se menciona la misma proporción, si se lo hace en el riesgo de incendio. Los pocos estudios mencionados en esta discusión hacen evidente la escasa información sobre los eventos de incendios en estas zonas de estudios y la necesidad de adquirir mayor información sobre estos fenómenos.

CAPITULO 6 CONCLUSIONES Y RECOMENDACIONES

6.1.CONCLUSIONES

- El rendimiento de los equipos computacionales puede afectar el procesamiento de la base de datos en ArcGIS por lo que, en ocasiones, se limitó a la computadora a realizar una operación a la vez y se requirió de una combinación de softwares de geoestadística para procesar la base de datos. Es así que, la rasterización de vectores como Cobertura de Uso de Suelo, toma mucho tiempo e intentos en ser procesada, debido a la cantidad de información que tiene esta capa. Además, el cálculo del VCI (Índice de Condición Vegetal) se pudo automatizar en MATLAB y así optimizar el tiempo. Es decir se necesitó más herramientas que las de ArcGIS para procesar los datos cualitativos y cuantitativos, y así armar la base de datos.
- En la valoración de la significancia de las variables predictivas, se obtuvo diferentes resultados entre la evaluación de la importancia reflejada por los Árboles de Decisión y el resultado del modelo en relación de la significancia de los coeficientes de las variables. Aunque los Árboles de Decisión ayudaron dentro del proceso, también sus resultados dificultaron las combinaciones que se debía hacer para los modelos, pues variables que tenían poca significancia no iban a ser incluidas dentro del modelo. Por ejemplo, en los Árboles de Decisión la variable Uso de Suelo fue catalogada con baja importancia dentro de la prueba, para la zona norte, pero esta misma variable en el modelo tuvo el mayor efecto para describir a la variable dependiente.
- La variable que se repite con una importancia significativa, en ambas zonas, es el Uso de Suelo, donde el nivel Tierra agropecuaria ocupa el segundo y tercer lugar en la zona norte y sur, respectivamente. Esto revela la importante influencia que existe del factor humano dentro de los patrones. Por otro lado, las pruebas de significancia apuntaban a que la temperatura jugaría un papel importante, dentro del modelo bajó la precisión global que se observó en relación a otras combinaciones de variables predictivas.

- El modelo probabilístico mostró mejor precisión para la zona sur que para la zona norte. El área bajo la curva (AUC) de ROC reveló que la zona norte tiene un valor igual a 0.83 y la zona sur tiene un valor aproximado a 0.79, esto demuestra que el norte obtuvo mayor exactitud global en la prueba porque está más cerca de 1. Finalmente, en los resultados que arroja la matriz de confusión se observa que los No incendios se identifican mejor que los Incendios.
- En el modelo logístico se pudo evidenciar que, los patrones para la producción de incendios forestales difieren entre la zona norte y la sur, esto hace que se acepte la hipótesis planteada en este estudio. Se pudo observar que existe una variable topográfica que describe mejor una zona que otra; además la temporada en la que se promedió una variable también influye en la descripción de los incendios y no incendios en cada zona. Por ejemplo, para el norte la pendiente, el NDVI y la precipitación de la temporada antes de los incendios, describen mejor los patrones de incendio. En cambio para el sur, las variables que describen mejor estos eventos fueron la elevación, la precipitación media mensual, y el NDVI total. Por otro lado, se pudo observar que el uso de suelo es la variable que mejor explica los incendios en ambas zonas de estudio.
- Hubo dificultades en el desarrollo del mapa de probabilidad de incendios, pues se tuvo que realizar transformaciones a la información para que pueda ser interpolada. Debido a que los datos para la interpolación con Kriging deben tener una distribución normal y los datos de probabilidad de incendio de este estudio no cumplían con este requisito, se procedió a realizar transformaciones Box-Cox de las probabilidades, pero no se pudo obtener una distribución completamente normal, especialmente para los datos de la zona sur. Sin embargo, se prefirió esta herramienta porque optimizó el tiempo de trabajo y además los resultados tienen mayor respaldo que otros tipos de interpolación por trabajar con geoestadística.
- Del mapa de probabilidades se puede concluir que los lugares con mayor afectación por la probabilidad Media, Alta y Muy Alta son las zonas rurales y este riesgo incrementa conforme aumenta la distancia a las capitales provinciales, exceptuando Loja. Por otro lado, Loja es la provincia en la cual

se produjo más del 50% de los incendios de la zona sur y el mapa refleja que su territorio se encuentra en gran riesgo de producción incendios.

6.2.RECOMENDACIONES

- Se recomienda automatizar el cálculo de índices, pues si hace este tipo de procesamiento en ArcGIS, al ser ingresados de forma manual, tomará tiempo en obtener un resultado.
- Usar otras pruebas estadísticas de significancia que permitan escoger con mayor efectividad las variables predictivas. Esto podría agilizar el proceso en la creación del modelo predictivo y evitar la múltiple combinación de variables para escoger el modelo con mejor precisión.
- En este estudio se usó un método de interpolación para predecir el valor de la probabilidad en cada área de estudio. Debido a los inconvenientes con la distribución de los datos, se recomienda buscar otro método de interpolación de la probabilidad o usar la ecuación del modelo logístico para obtener el valor del número total de celdas.
- Como se observó en el mapa, la zona rural de las provincias son los lugares con mayor riesgo de incendios forestales. Además, se sabe que más del 90% de los incendios son provocados por actividad humana, debido a que en este trabajo no se pudo analizar más variables que expliquen la relación de estas actividades con los incendios, se podría incluir un análisis social más profundo como actividad económica o ingreso per cápita, y verificar si esta información describe con mayor precisión estos eventos.
- Se pudo evidenciar que no existen muchos estudios para identificar los patrones que determinan los incendios forestales. Por ello se sugiere agregar más información para ampliar la base de datos y así entender mejor los patrones de incendios. Algunos autores han usado otras variables que influyen en los incendios, como datos de viento, humedad relativa, incluso se podría analizar si el fenómeno del niño, que afecta a nuestro país, ejerce algún tipo de impacto en la producción de estos eventos. En tanto exista información disponible para esta u otra área de estudio.
- Otro aspecto importante que se puede tratar es el comportamiento de los incendios en función del aumento de temperatura global como consecuencia

del cambio climático. Sin duda este escenario implica una gran presión sobre los ecosistemas mundiales, por lo que muchos patrones podrían cambiar a futuro. Poder comprender estos cambios, ayudaría a las autoridades a tomar acciones que prevengan o mitiguen el impacto del cambio climático en este fenómeno.

CAPITULO 7 BIBLIOGRAFÍA

- Achaerandio, L. (2001). Las variables de la investigación. In *Iniciación a la práctica de la investigación* (pp. 81–96). Retrieved from http://biblio3.url.edu.gt/publiclg/biblio_sin_paredes/fac_politicas/2018/tecnico_trab/inici_pracinvves/cont/
- Altamirano, A., Salas, C., Yaitul, V., Smith-Ramirez, C., & Ávila, A. (2013). Influencia de la heterogeneidad del paisaje en la ocurrencia de incendios forestales en Chile Central. *Revista de Geografía Norte Grande*, 170(55), 157–170.
- Amat, J. (2016). *Validación de modelos de regresión: Crss-validation, OneleaveOut, Bootstrap*. 17. Retrieved from <https://github.com/JoaquinAmatRodrigo/Estadistica-con-R>
- ArcGIS for Desktop. (2016). Cómo funciona Kriging—Ayuda | ArcGIS for Desktop. Retrieved August 29, 2020, from Conceptos del conjunto de herramientas Interpolación de ráster website: <https://desktop.arcgis.com/es/arcmap/10.3/tools/3d-analyst-toolbox/how-kriging-works.htm>
- Beltrán, D. D., & Ponte, E. B. (2015). *Análisis clasificadorio de las gestantes según vía de culminación del parto aplicando regresión logística binaria*. Universidad Nacional Agraria La Molina.
- Borowik, T., Pettorelli, N., Sönrichsen, L., & Jędrzejewska, B. (2013). Normalized difference vegetation index (NDVI) as a predictor of forage availability for ungulates in forest and field habitats. *European Journal of Wildlife Research*, 59(5), 675–682. <https://doi.org/10.1007/s10344-013-0720-0>
- Breiman, L., Culter, A., Liaw, A., & Wiener, M. (2018). *Breiman and Cutler's Random Forests for Classification and Regression. Package "randomForest". Version 4.6-14*. Retrieved from <https://cran.r-project.org/web/packages/randomForest/randomForest.pdf>
- Brownson, J. (2014). Meteorology: the Many Facets of the Sky. In *Solar Energy Conversion Systems* (1st ed., pp. 99–133). <https://doi.org/10.1016/b978-0-12-397021-3.00005-3>
- Buendia, L., Colás, P., & Hernadez, F. (2001). Tipos de variables en un experimento. In *Metodos de Investigacion en Psicopedagogia* (pp. 1–7). Madrid: McGraw-Hill.
- Burrough, P. A., & McDonnell, R. A. (1986). Spatial Analysis Using Continuous Fields. *Principles of Geographical Information Systems*, pp. 183–219. <https://doi.org/10.1519/r-13104.1>
- Bustos, A., Segura, D., Coronel, J., Onofa, Á., Ministerio del Ambiente del Ecuador, Subsecretaría de Patrimonio Natural, & Programa Amazonia Sin Fuego. (2019, March 25). Los incendios forestales en Ecuador. Retrieved August 12, 2020, from Osbodigital. Todo sobre gestión forestal. website: <https://osbodigital.es/2019/03/25/los-incendios-forestales-en-ecuador/>

- Catry, F. X., Rego, F. C., Bação, F. L., & Moreira, F. (2009). Modeling and mapping wildfire ignition risk in Portugal. *International Journal of Wildland Fire*, 18(8), 921–931. <https://doi.org/10.1071/WF07123>
- CAZALAC. (2018). Índice de la Condición de la Vegetación (VCI). Retrieved April 22, 2020, from [http://www.climatedatalibrary.cl/CAZALAC/maproom/Canicula/Monitoring/Agricultural/VCI.html?Set-Language=es®ion=irids%3ASOURCES%3AFeatures%3APolitical%3ACentralAmerica%3Adistricts%3Agid%40811%3Ads&T=21-30 Sep 2018#tabs-1](http://www.climatedatalibrary.cl/CAZALAC/maproom/Canicula/Monitoring/Agricultural/VCI.html?Set-Language=es®ion=irids%3ASOURCES%3AFeatures%3APolitical%3ACentralAmerica%3Adistricts%3Agid%40811%3Ads&T=21-30%20Sep%202018#tabs-1)
- Chang, Y., Zhu, Z., Bu, R., Chen, H., Feng, Y., Li, Y., ... Wang, Z. (2013). Predicting fire occurrence patterns with logistic regression in Heilongjiang Province, China. *Landscape Ecology*, 28(10), 1989–2004. <https://doi.org/10.1007/s10980-013-9935-4>
- Chuvienco, E. (2008). *Teledetección Ambiental: La observación de la Tierra desde el Espacio*. (3era edici; S. A. Ariel, Ed.). Retrieved from <https://siglibreuruguay.wordpress.com/2016/09/01/libro-gratuito-teledeteccion-ambiental-la-observacion-de-la-tierra-desde-el-espacio-de-emilio-chuvienco/>
- Climate Hazards Center. (2020). CHIRPS: Rainfall Estimates from Rain Gauge and Satellite Observations. Retrieved April 23, 2020, from <https://www.chc.ucsb.edu/data/chirps>
- Corporación Nacional Forestal de Chile. (2019). Detección de incendios forestales. Retrieved from <http://www.conaf.cl/incendios-forestales/combate-de-incendios-forestales/deteccion-de-incendios-forestales/>
- Cruz, I., & Rainer, R. (2012). Detección y monitoreo de incendios forestales mediante imágenes de satélites. *Biodiversitas*, 100, 12–13.
- De Groot, R. S. (1992). Functions of nature : evaluation of nature in environmental planning, management and decision making. In *Functions of nature : evaluation of nature in environmental planning, management and decision making*. Retrieved from <https://www.cabdirect.org/cabdirect/abstract/19931980587>
- Department of Statistics Online Program. (2018). 10.7 - Detecting Multicollinearity Using Variance Inflation Factors | STAT 462. Retrieved August 12, 2020, from Leson 10: Regression Pitfalls website: <https://online.stat.psu.edu/stat462/node/180/>
- Didan, K. (2015). MOD13A3 MODIS/Terra vegetation Indices Monthly L3 Global 1km SIN Grid V006 [Data set]. NASA EOSDIS Land Processes DAAC. Retrieved September 13, 2019, from <https://lpdaac.usgs.gov/products/mod13a3v006/#citation>
- Dietterich, T. (1995). Overfitting and Undercomputing in Machine Learning. *ACM Computing Surveys (CSUR)*, 27(3), 326–327. <https://doi.org/10.1145/212094.212114>
- Dirección de Monitoreo de Eventos Adversos. (2016). *Informe de Situación 90-Incendios Forestales (28/10/2016)*. Quito.

- Dirección de Monitoreo de Eventos Adversos. (2017). *Informe de Situación- Incendios Forestales 26/10/2017*. Retrieved from <http://www.serviciometeorologico.gob.ec/mapas-climaticos/>
- Dirección de Monitoreo de Eventos Adversos. (2018). *Informe de Situación - Incendios Forestales a nivel nacional 2018*. Retrieved from <https://www.gestionderiesgos.gob.ec/wp-content/uploads/2018/10/SITREP-5- Informe-de-Situación-SGR-2018-031008.pdf>
- Dirección de Monitoreo de Eventos Adversos. (2019). *Informe de Situación No.39 - Incendios Forestales a nivel nacional 2019*. Retrieved from <https://www.gestionderiesgos.gob.ec/wp-content/uploads/2018/10/SITREP-5- Informe-de-Situación-SGR-2018-031008.pdf>
- Dobson, A., & Barnett, A. (2018). *An Introduction to Generalized Linear Models* (4th ed.; Taylor & Francis Group, Ed.).
- Ecuador Forestal. (2013). INCENDIOS FORESTALES: ¿Cómo se producen, cómo evitarlos y qué efectos genera en el ecosistema? Retrieved from <https://ecuadorforestal.org/actualidad-forestal/incendios-forestales-como-se-producen-como-evitarlos-y-que-efectos-genera-en-el-ecosistema/>
- Escobar, C. (2017). Facultad de Ciencias Sociales - Universidad de Chile. Retrieved from <http://www.facso.uchile.cl/noticias/130905/los-problemas-sociopoliticos-detras-de-los-incendios-forestales>
- Estacio, J., & Narváez, N. (2012). Incendios forestales en el Distrito Metropolitano de Quito (DMQ): Conocimiento e intervención pública del riesgo. *Letras Verdes*, 11, 27–52. <https://doi.org/10.17141/letrasverdes.11.2012.914>
- Eva, H., & Lambin, E. F. (2000). Fires and land-cover change in the tropics: A remote sensing analysis at the landscape scale. *Journal of Biogeography*, 27(3), 765–776. <https://doi.org/10.1046/j.1365-2699.2000.00441.x>
- Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), 861–874. <https://doi.org/10.1016/j.patrec.2005.10.010>
- Flández, J. (2011). *Análisis de cobertura visual de puntos de observación terrestre fija de detección de incendios forestales de CONAF Región de Los Ríos* (Universidad Australa de Chile). Retrieved from <http://cybertesis.uach.cl/tesis/uach/2011/fiff584a/doc/fiff584a.pdf>
- Food and Agriculture Organization. (2009). El problema de los incendios forestales. Retrieved from <http://www.fao.org/news/story/es/item/29145/icode/>
- Fox, J. (2015). Generalized Linear Models. In *Applied Regression Analysis and Generalized Linear Models* (3rd ed., pp. 379–424). Retrieved from https://www.sagepub.com/sites/default/files/upm-binaries/21121_Chapter_15.pdf
- Frost, J. (2019). *Regression Analysis* (1st ed.).
- Garcia, C. V., Woodard, P. M., Titus, S. J., Adamowicz, W. L., & Lee, B. S. (1995). A logit model for predicting the daily occurrence of human caused forest fires. *International Journal of Wildland Fire*, 5(2), 101–111.

<https://doi.org/10.1071/WF9950101>

- Giglio, L., Justice, C., Boschetti, L., & Roy, D. (2015). MCD64A1 MODIS/Terra+Aqua Burned Area Monthly L3 Global 500m SIN Grid V006 [Data set]. NASA EOSDIS Land Processes DAAC. <https://doi.org/https://doi.org/10.5067/MODIS/MCD64A1.006>
- Giglio, Louis, Boschetti, L., Roy, D., Hoffman, A. A., & Humber, M. (2016). Collection 5 MODIS Burned Area product User Guide Version 1.0. In Nasa. Retrieved from https://modis-land.gsfc.nasa.gov/pdf/MODIS_C6_BA_User_Guide_1.0.pdf
- Good, P. (2005). Permutation, Parametric, and Bootstrap Tests of Hypotheses. In *Journal of the American Statistical Association* (third, Vol. 100). <https://doi.org/10.1198/jasa.2005.s48>
- Guo, F., Su, Z., Wang, G., Sun, L., Lin, F., & Liu, A. (2016). Wildfire ignition in the forests of southeast China: Identifying drivers and spatial distribution to predict wildfire likelihood. *Applied Geography*, 66, 12–21. <https://doi.org/10.1016/j.apgeog.2015.11.014>
- Guo, F., Su, Z., Wang, G., Sun, L., Tigabu, M., Yang, X., & Hu, H. (2017). Understanding fire drivers and relative impacts in different Chinese forest ecosystems. *Science of the Total Environment*, 605–606, 411–425. <https://doi.org/10.1016/j.scitotenv.2017.06.219>
- Han, H., Guo, X., & Yu, H. (2016). Variable selection using Mean Decrease Accuracy and Mean Decrease Gini based on Random Forest. *Proceedings of the IEEE International Conference on Software Engineering and Service Sciences, ICSESS*, 0, 219–224. <https://doi.org/10.1109/ICSESS.2016.7883053>
- Hashim, H., Abd, Z., & Adnan, N. A. (2019). Urban Vegetation Classification With NDVI Threshold Value Method With Very High Resolution (VHR) Pleiades Imagery. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-4/W16*(October), 237–240. <https://doi.org/10.5194/isprs-archives-xxii-4-w16-237-2019>
- Hauck, W., & Donner, A. (1977). Wald's Test as Applied to Hypotheses in to Logit Analysis. *Journal of the American Statistical Association*, 72(360), 851–853. Retrieved from <http://www.jstor.org/stable/2286473> .
- Hirschberger, P. (2017). *FORESTS ABLAZE. Causes and effects of global forest fires*. Retrieved from <https://www.worldwildlife.org/stories/forest-fires-the-good-and-the-bad>
- Instituto Nacional de Estadística y Censos. (2010). Base de Datos-Censo de Población y Vivienda 2010. Retrieved September 9, 2020, from Población y Demografía website: <https://www.ecuadorencifras.gob.ec/base-de-datos-censo-de-poblacion-y-vivienda-2010-a-nivel-de-manzana/>
- Jiménez Moreno, A. (1991). Modelización cartográfica de densidades mediante estimadores Kernel'. *Treballs de La Societat Catalana de Geografia*, 155–170.
- Jones, M. W., Santín, C., van der Werf, G. R., & Doerr, S. H. (2019). Global fire

- emissions buffered by the production of pyrogenic carbon. *Nature Geoscience*, 12(9), 742–747. <https://doi.org/10.1038/s41561-019-0403-x>
- Kapetsky, J. M., & Aguilar, J. (2009). Sistemas de información geográfica , sensores remotos y mapeo para el desarrollo y la gestión de la acuicultura marina. FAO Documento Técnico de Pesca No. 458. In FAO, *Documento Tecnico*. Retrieved from <http://agris.fao.org/agris-search/search/display.do?f=2011/XF/XF1102.xml;XF2006444791>
- Kitzberger, T., Brown, P. M., Heyerdahl, E. K., Swetnam, T. W., & Veblen, T. T. (2006). *Contingent Pacific-Atlantic Ocean influence on multicentury wildfire synchrony over western North America*. Retrieved from www.pnas.org/cgi/content/full/
- Kogan, F. N. (1997). Global Drought Watch from Space. *Bulletin of the American Meteorological Society*, 78, 621–636. <https://doi.org/https://doi.org/10.1080/01431169608949106>
- Laguna, C. (2014). *INFERENCIA PARÁMETRICA: RELACIÓN ENTRE DOS VARIABLES CUALITATIVAS*. Retrieved from <http://www.ics-aragon.com/cursos/salud-publica/2014/pdf/M2T07.pdf>
- Lavorel, S., Flannigan, M. D., Lambin, E. F., & Scholes, M. C. (2007). Vulnerability of land systems to fire: Interactions among humans, climate, the atmosphere, and ecosystems. *Mitigation and Adaptation Strategies for Global Change*, 12(1), 33–53. <https://doi.org/10.1007/s11027-006-9046-5>
- Liaw, A., & Wiener, M. (2002). Classification and Regression by randomForest. *R News*, 2(3), 18–22.
- Long, J. S., & Freese, J. (2001). Regression Models for Categorical Dependent Variables Using STATA. In *Sociology The Journal Of The British Sociological Association*. <https://doi.org/10.1186/2051-3933-2-4>
- López-González, E., & Ruiz-Soler, M. (2011). Análisis de datos con el Modelo Lineal Generalizado. Una aplicación con R. *Revista Española de Pedagogía*, 248, 59–80. Retrieved from <https://reunir.unir.net/handle/123456789/3839>
- López, I., & Pita, S. (1998). Curvas ROC. *Cad Aten Primaria*, 5(4), 229–235.
- MAE. (2012). *Metodología para la representación Cartográfica de los Ecosistemas del Ecuador Continental*. Retrieved from http://www.ambiente.gob.ec/wp-content/uploads/downloads/2013/03/Documento_Metodología_28_05_2012_v2_1.pdf
- Maingi, J. K., & Henry, M. C. (2007). Factors influencing wildfire occurrence and distribution in eastern Kentucky, USA. *International Journal of Wildland Fire*, 16(1), 23–33. <https://doi.org/10.1071/WF06007>
- Martínez, J., Chuvieco, E., & Martín, P. (2004). Estimación de factores de riesgo humano de ignición en España mediante regresión logística. *II Simposio Internacional Sobre Políticas, Planificación y Economía de Los Programas Contra Incendios Forestales*, 281–294.
- Martinez, R., Ruiz, D., Andrade, M., Blacutt, L., Pabón, D., Jaimes, E., ...

- Euscátegui, C. (2011). Synthesis of the Climate of the Tropical Andes. In *Climate change and biodiversity in the tropical Andes* (pp. 97–109). <https://doi.org/10.13140/2.1.3718.4969>
- Mataveli, G. A. V., Silva, M. E. S., Pereira, G., Da Silva Cardozo, F., Shinji Kawakubo, F., Bertani, G., ... Da Silva, V. V. (2018). Satellite observations for describing fire patterns and climate-related fire drivers in the Brazilian savannas. *Natural Hazards and Earth System Sciences*, 18(1), 125–144. <https://doi.org/10.5194/nhess-18-125-2018>
- McBratney, A. B., & Webster, R. (1986). Choosing functions for semi-variograms of soil properties and fitting them to sampling estimates. *Journal of Soil Science*, 37(4), 617–639. <https://doi.org/10.1111/j.1365-2389.1986.tb00392.x>
- McCullagh, P., & Nelder, J. A. (2019). *Generalized Linear Models* (Segunda). <https://doi.org/10.1201/9780203753736>
- Medina, R. F., & Ñique, C. I. (2017). Bosques aleatorios como extensión de los árboles de clasificación con los programas R y Python. *Interfases*, 0(010), 165. <https://doi.org/10.26439/interfases2017.n10.1775>
- Ministerio del Ambiente Ecuador. (2015). *Análisis de vulnerabilidad ante Riesgos Naturales Hidrometeorológicos en las Áreas Protegidas del Ecuador Continental*.
- Ministerio del Ambiente Ecuador. Reglamento al Código Orgánico del Ambiente. , Pub. L. No. Registro Oficial Suplemento 507, 752 Fielweb Evolución Jurídica 1 (2019).
- Ministerio del Ambiente y Agua. (2014). Plan de Contingencia de Incendios Forestales buscar prevenir, mitigar y controlar este fenómeno natural. Retrieved August 12, 2020, from Noticias website: <https://www.ambiente.gob.ec/plan-de-contingencia-de-incendios-forestales-buscar-prevenir-mitigar-y-controlar-este-fenomeno-natural/>
- Mohammadi, F., Bavaghar, M. P., & Shabanian, N. (2014). Forest Fire Risk Zone Modeling Using Logistic Regression and GIS: An Iranian Case Study. *Small-Scale Forestry*, 13(1), 117–125. <https://doi.org/10.1007/s11842-013-9244-4>
- Montiel, C., Herrero, G., & Solano, J. (2009). *Detección de incendios forestales: Sistemas de detección*. Retrieved from <http://www.conaf.cl/incendios-forestales/combate-de-incendios-forestales/deteccion-de-incendios-forestales/>
- Moreno Ruiz, J. A., García Lázaro, J. R., Del Águila Cano, I., & Leal, P. H. (2013). Burned area mapping in the North American boreal forest using terra-MODIS LTDR (2001-2011): A comparison with the MCD45A1, MCD64A1 and BA GEOLAND-2 products. *Remote Sensing*, 6(1), 815–840. <https://doi.org/10.3390/rs6010815>
- Müller, H.-G. (2006). Density Estimation—II. In *Encyclopedia of Statistical Sciences*. <https://doi.org/10.1002/0471667196.ess0285.pub2>
- Oliveira, S., Oehler, F., San-Miguel-Ayanz, J., Camia, A., & 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>
- Oliver, M. A., & Webster, R. (1990). Kriging: a method of interpolation for geographical information systems. *International Journal of Geographical Information Systems*, (4:3), 313–332. <https://doi.org/http://dx.doi.org/10.1080/02693799008941549>
- Ospina, E. (2017). *Mapeo de áreas quemadas por incendios forestales en el Ecuador utilizando información satelital de la NASA, período 2000-2016* (Escuela Politécnica Nacional). Retrieved from <http://bibdigital.epn.edu.ec/bitstream/15000/8103/4/CD-2254.pdf>
- Pazmiño, D. (2019). Peligro de incendios forestales asociado a factores climáticos en Ecuador. *FIGEMPA: Investigación y Desarrollo*, 1(1), 10–18. <https://doi.org/10.29166/revfig.v1i1.1800>
- Pérez, C., & Muñoz, Á. L. (2002). *Teledetección: nociones y aplicaciones* (C. Pérez Gutiérrez, Ed.). Retrieved from <https://dialnet.unirioja.es/servlet/extart?codigo=1011761>
- Privitera, G. (2014). Introduction to Hypothesis Testing. In *Statistics for the Behavioral Sciences* (Second). Retrieved from https://www.sagepub.com/sites/default/files/upm-binaries/40007_Chapter8.pdf
- Refaeilzadeh, P., Tang, L., & Huan, L. (2009). Cross-Validation. In *Encyclopedia of database systems* (Vol. 5, pp. 532–538). <https://doi.org/https://doi.org/10.1007/978-0-387-39940-9>
- Rodas, M. (2015). *Análisis del comportamiento de incendios forestales en la cuenca del río Paute mediante sensores remotos* (Universidad de Cuenca). Retrieved from <http://dspace.ucuenca.edu.ec/bitstream/123456789/22325/1/Tesis.pdf>
- Rodríguez, G. (2007). Linear Models for Continuous Data. In *Lecture Notes on Generalized Linear Models*. Retrieved from http://chianti.ucsd.edu/~rsaito/ENTRY1/WEB_RS3/PDF/Learn/Statistics/generalized_linear_model_RodriguezG_lecnote21.pdf
- Rosales, W. (2019). 62.000 hectáreas de bosque afectadas por incendios desde 2012 en Ecuador. Retrieved from <https://www.eluniverso.com/noticias/2019/09/08/nota/7506112/62000-hectareas-bosque-afectadas-incendios-2012-ecuador>
- Rowell, A., & Moore, P. (2000). *Global Review of Forest Fires*. Retrieved from <https://portals.iucn.org/library/efiles/documents/2000-047.pdf>
- Sabuco, P. (2013). The problem of forest fires and the basis for its teledetection in Perú. *Apuntes de Ciencia & Sociedad*, 03(02), 5–8.
- Salazar, C., & Del Castillo, S. (2018). *Fundamentos Básicos De Estadística* (Primera). Retrieved from [http://www.dspace.uce.edu.ec/bitstream/25000/13720/3/Fundamentos Básicos de Estadística-Libro.pdf](http://www.dspace.uce.edu.ec/bitstream/25000/13720/3/Fundamentos_Básicos_de_Estadística-Libro.pdf)

- Secretaría de Gestión de Riesgos. (2018). Plan Nacional. In *Plan Nacional de adaptación al cambio climático*. Retrieved from http://www.planificacion.gob.ec/wp-content/uploads/downloads/2017/10/PNBV-26-OCT-FINAL_0K.compressed1.pdf
- Sifakis, N. I., Iossifidis, C., Kontoes, C., & Keramitsoglou, I. (2011). Wildfire Detection and Tracking over Greece Using MSG-SEVIRI Satellite Data. *Remote Sensing*, 3(3), 524–538. <https://doi.org/10.3390/rs3030524>
- SOLARGIS. (2019). *User's Guide GHI*. 1–3. Retrieved from [https://msdn.microsoft.com/en-us/library/aa291809\(v=vs.71\).aspx](https://msdn.microsoft.com/en-us/library/aa291809(v=vs.71).aspx)
- Stambaugh, M. C., & Guyette, R. P. (2008). Predicting spatio-temporal variability in fire return intervals using a topographic roughness index. *Forest Ecology and Management*, 254(3), 463–473. <https://doi.org/10.1016/j.foreco.2007.08.029>
- Statistical Consulting Group. (2020). FAQ: What are pseudo R-squareds? Retrieved August 11, 2020, from UCLA website: <https://stats.idre.ucla.edu/other/mult-pkg/faq/general/faq-what-are-pseudo-r-squareds/>
- Swetnam, T. W. (1993). Fire history and climate change in giant sequoia groves. *Science*, 262(5135), 885–889. <https://doi.org/10.1126/science.262.5135.885>
- Tan, Y. K., & Panda, S. K. (2011). Self-autonomous wireless sensor nodes with wind energy harvesting for remote sensing of wind-driven wildfire spread. *IEEE Transactions on Instrumentation and Measurement*, 60(4), 1367–1377. <https://doi.org/10.1109/TIM.2010.2101311>
- Turner, M. G., & Romme, W. H. (1994). Landscape dynamics in crown fire ecosystems. *Landscape Ecology*, 9(1), 59–77. <https://doi.org/10.1007/BF00135079>
- Universidad de Maryland. (2017). MODIS Active Fire and Burned Area Products - Burned Area. Retrieved December 17, 2020, from <https://modis-fire.umd.edu/ba.html>
- Vallejo-Villalta, I., Rodríguez-Navas, E., & Márquez-Pérez, J. (2019). Mapping forest fire risk at a local scale—A case study in Andalusia (Spain). *Environments - MDPI*, 6(3). <https://doi.org/10.3390/environments6030030>
- van Der Werf, G. R., Randerson, J. T., Giglio, L., van Leeuwen, T. T., Chen, Y., Rogers, B. M., ... Kasibhatla, P. S. (2017). Global fire emissions estimates during 1997-2016. *Earth System Science Data*, 9(2), 697–720. <https://doi.org/10.5194/essd-9-697-2017>
- Vilar del Hoyo, L., Martín Isabel, M. P., & Martínez Vega, F. J. (2011). Logistic regression models for human-caused wildfire risk estimation: analysing the effect of the spatial accuracy in fire occurrence data. *European Journal of Forest Research*, 130(6), 983–996. <https://doi.org/10.1007/s10342-011-0488-2>
- Wan, Z., Hook, S., & Hulley, G. (2015). MOD11B3 MODIS/Terra Land Surface Temperature/Emissivity Monthly L3 Global 6km SIN Grid V006.

<https://doi.org/https://doi.org/10.5067/MODIS/MOD11B3.006>

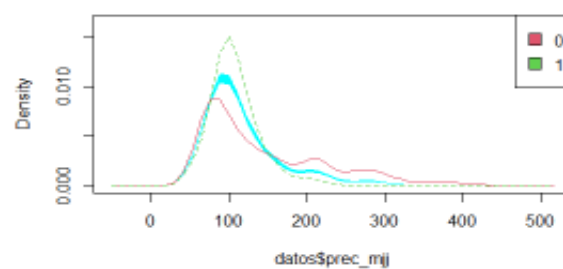
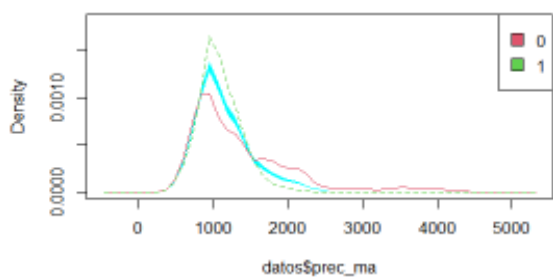
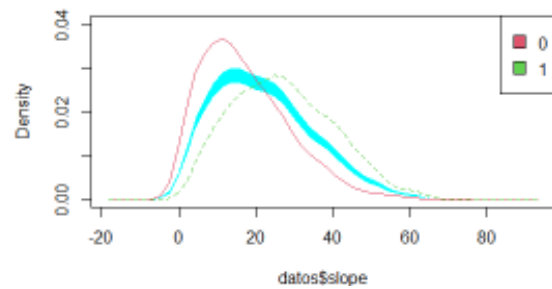
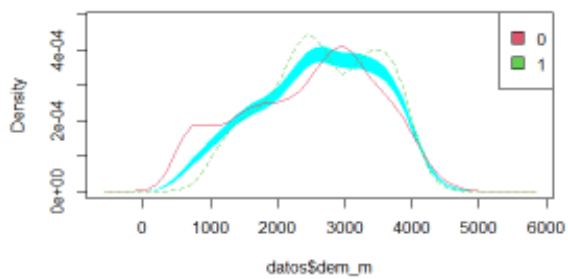
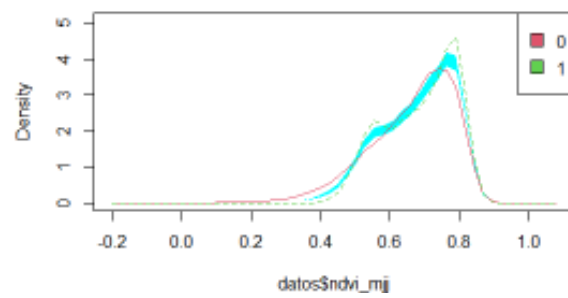
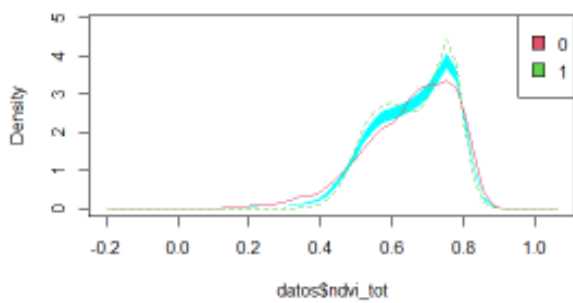
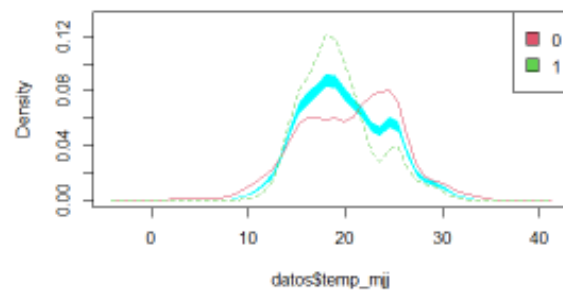
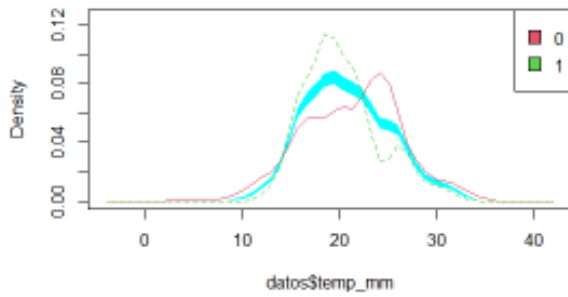
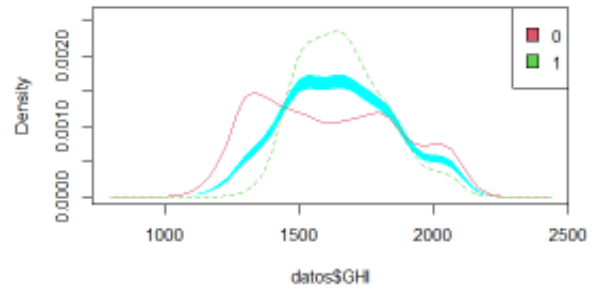
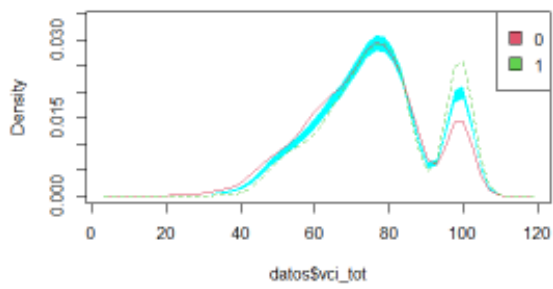
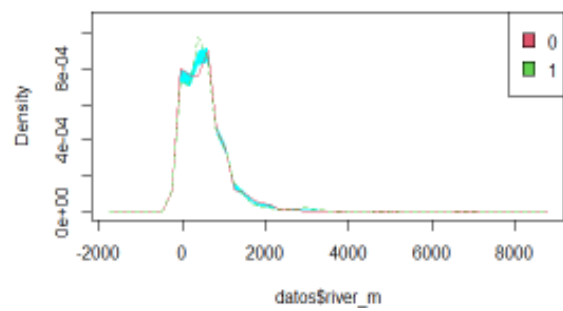
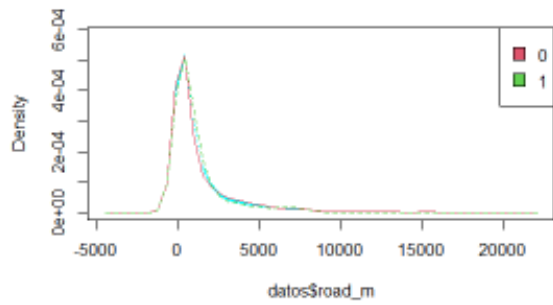
Williams, R. (2019). *Scalar Measures of Fit: Pseudo R 2 and Information Measures (AIC & BIC)*. Retrieved from <https://www3.nd.edu/~rwilliam/stats3/L05.pdf>

ANEXOS

ANEXO 1

Pruebas No Paramétricas para la Zona Norte

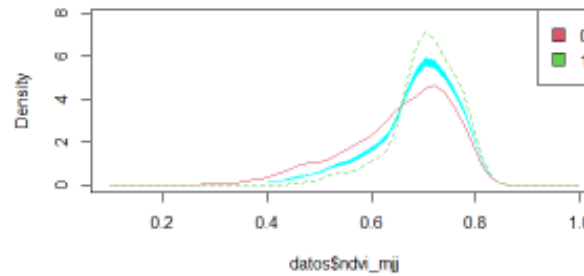
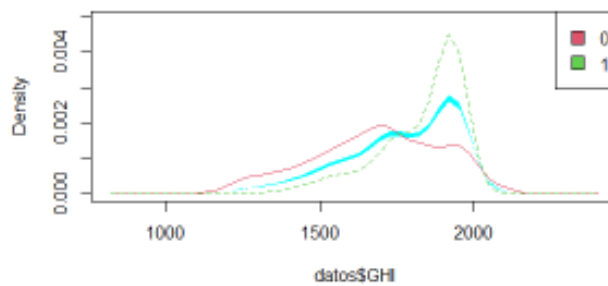
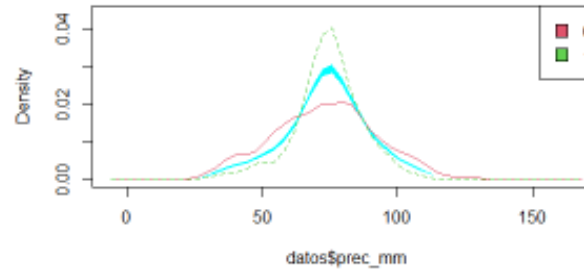
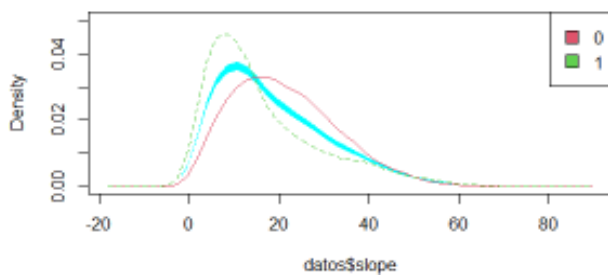
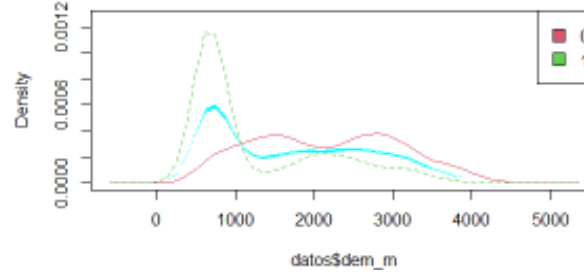
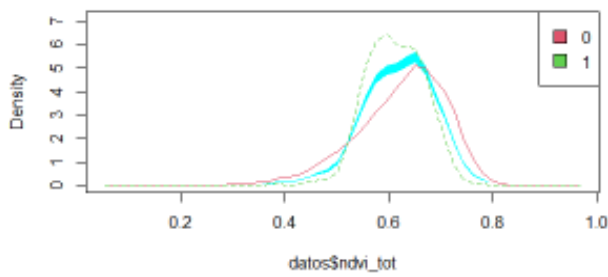
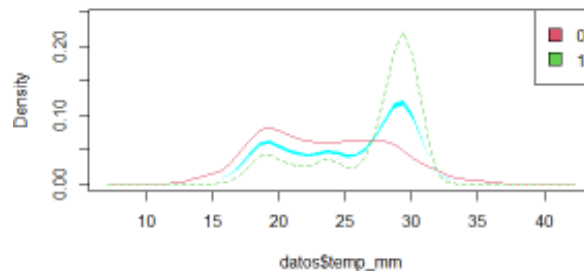
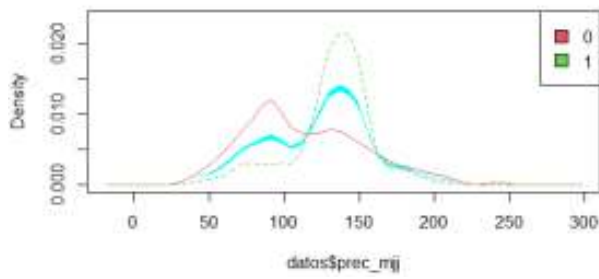
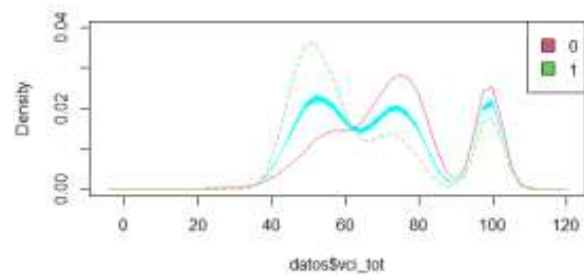
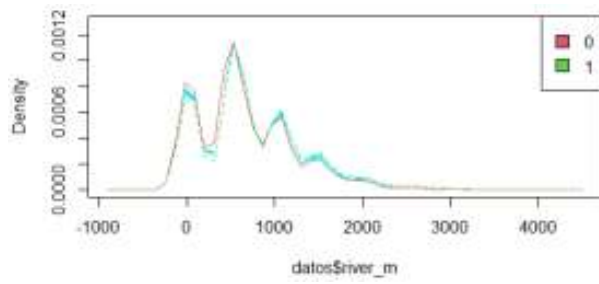
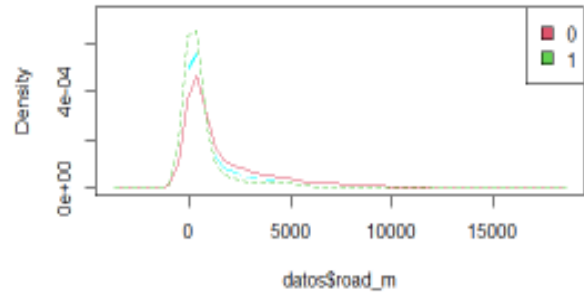
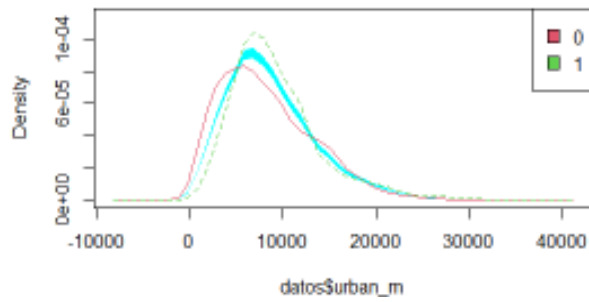
Donde urban_m es la distancia a centros poblados, river_m es la distancia a ríos en metros, vci_tot es el Índice de la Condición de la Vegetación del periodo de estudio, GHI es la Irradiación horizontal global, temp_mm temperatura media mensual, temp_mjj es la temperatura media de la temporada antes de incendios (noviembre-mayo), ndvi_tot Índice de Vegetación de Diferencia Normalizada del periodo de estudio, ndvi_mjj es el Índice de Vegetación de Diferencia Normalizada de la temporada antes de incendios (nov-may), dem_m modelo de elevación digital en metros, slope pendiente en porcentaje, prec_ma precipitación media anual, prec_mjj es la precipitación media de la temporada antes de incendios (nov-may).



ANEXO 2

Pruebas No Paramétricas para la Zona Sur

Donde urban_m es la distancia a centros poblados, river_m es la distancia a ríos en metros, vci_tot es el Índice de la Condición de la Vegetación del periodo de estudio, GHI es la Irradiación horizontal global, temp_mm temperatura media mensual, temp_mjj es la temperatura media de la temporada antes de incendios (enero-julio), ndvi_tot Índice de Vegetación de Diferencia Normalizada del periodo de estudio, ndvi_mjj es el Índice de Vegetación de Diferencia Normalizada de la temporada antes de incendios (ener-jul), dem_m modelo de elevación digital en metros, slope pendiente en porcentaje, prec_ma precipitación media anual, prec_mjj es la precipitación media de la temporada antes de incendios (ener-jul).



ANEXO 3

Pruebas Paramétricas para la Zona Norte y Sur

Donde Aspect_1 es el aspecto y CUS es la cobertura de uso de Suelo.

ZONA NORTE

	Aspect_1	p_value	prop_incendios
	<chr>	<dbl>	<dbl>
1	NE	0.0111	0.455
2	SE	0.0155	0.334
3	S	0.237	0.370
4	W	0.326	0.421
5	NW	0.341	0.381
6	SW	0.630	0.412
7	N	0.932	0.403
8	E	1	0.401

	CUS	p_value	prop_incendios
	<fct>	<dbl>	<dbl>
	VEGETACION ARBUSTIVA	1.69e-43	0.581
	TIERRA AGROPECUARIA	1.12e-18	0.285
	BOSQUE	1.27e- 8	0.311

ZONA SUR

	Aspect_1	p_value	prop_incendios
	<chr>	<dbl>	<dbl>
1	E	0.0000000362	0.303
2	NW	0.0000000552	0.466
3	F	0.0172	0.875
4	NE	0.0262	0.359
5	SW	0.0871	0.375
6	SE	0.558	0.410
7	S	0.652	0.407
8	W	0.705	0.405
9	N	0.925	0.398

	CUS	p_value	prop_incendios
	<fct>	<dbl>	<dbl>
1	TIERRA AGROPECUARIA	1.20e-23	0.474
2	BOSQUE	3.39e-20	0.296
3	VEGETACION ARBUSTIVA	1.17e- 7	0.347