MAPBIOMAS
[FUEGO]

COLECCIÓN 2 PERÚ

Documento de base teórica sobre algoritmos (ATBD)

MapBiomas Fuego

Colección 2.0

Versión 1

Coordinadora General

Ethel Rubin de Celis Llanos

Equipo

Johnny Huamaní

Vanesa Llaulli

Ana Lucía Lagos

Diego Perez

Dilmer Gonzales

Maricielo Loli

Beto Cuellar

Rossangela Calle

Melanie Regalado

Resumen Ejecutivo

El Instituto del Bien Común (IBC) es una asociación civil peruana sin fines de lucro, fundada en 1998, dedicada a la protección y gestión sostenible de los bienes comunes: ríos, lagos, bosques, recursos pesqueros, áreas naturales protegidas y territorios comunales. Estos recursos son esenciales para la vida de los pueblos amazónicos y, en el contexto del cambio climático, su conservación contribuye al bienestar tanto de las comunidades locales como de todo el país.

Desde 2009, como parte de la Red Amazónica de Información Socioambiental Georreferenciada (RAISG), el IBC participa en la elaboración de mapas de deforestación de la Amazonía peruana mediante el uso de tecnologías satelitales avanzadas. En 2017, RAISG y MapBiomas Brasil dieron origen a la iniciativa MapBiomas Amazonía, de la cual el IBC es miembro activo. Posteriormente, en 2022, el IBC asumió el liderazgo de MapBiomas Perú, consolidando una red colaborativa de organizaciones no gubernamentales, universidades y empresas tecnológicas orientada a producir datos abiertos y de calidad sobre los cambios en el territorio nacional.

Desde el 2022, en el marco de la red MapBiomas Perú, el IBC es el líder técnico de la iniciativa Cobertura y Uso del Suelo y de la iniciativa Agua. En ellas genera series históricas de mapas anuales de cobertura, uso del suelo, agua y glaciares, utilizando imágenes satelitales Landsat desde 1985 y procesamiento en la nube con inteligencia artificial.

En 2025, la red se fortaleció con una nueva iniciativa en alianza con el Laboratorio Sala de Observación Perú de la Facultad de Ciencias Forestales de la Universidad Nacional Agraria La Molina (antes Sala de Observación del Perú, creada por la OTCA en 2012). Esta colaboración permitió desarrollar por primera vez mapas de áreas quemadas a nivel nacional.

Este laboratorio tiene su origen en la Sala de Observación Perú (SdO-Perú), creada en 2012 por la Organización del Tratado de Cooperación Amazónica (OTCA) para implementar sistemas participativos de monitoreo de la cobertura forestal amazónica y fortalecer la coordinación regional para el manejo sostenible de los bosques. Tras seis años de ejecución, sus bienes fueron transferidos a la UNALM, que le brindó infraestructura y apoyo técnico, consolidando su desarrollo educativo, científico y tecnológico. Hoy, este espacio ofrece a estudiantes, egresados y profesionales acceso a imágenes satelitales, colecciones de datos y tecnologías de avanzada, fomentando la investigación aplicada, la formación técnica y la generación de evidencia para la gestión sostenible de los ecosistemas y la conservación de la biodiversidad.

De esta manera, el Laboratorio Sala de Observación Perú (SdO-Perú) se consolida como socio técnico fundamental de MapBiomas Perú en la co-creación de mapas de áreas quemadas. Fruto de esta alianza, y tras el exitoso lanzamiento de la Colección 1 en 2025, el presente documento introduce la Colección 2, consolidando un monitoreo espacial y temporal de largo plazo, aportando evidencia científica clave para la gestión del riesgo de incendios, la conservación de ecosistemas y la toma de decisiones informadas frente al cambio climático en el Perú.

Índice

Lista de Tablas **3**
Lista de Figuras **3**
1. Introducción **4**
1.1. Objetivo del documento 4
1.2. MapBiomas Fuego Perú 4
1.3. Google Earth Engine, Google Colab y Google Cloud 5
1.4. Otras iniciativas de mapeo 5
2. Metodología **8**
2.1 Delimitación de regiones de trabajo 9
2.2 Generación de mosaicos anuales 10
2.3. Colecta de muestras de entrenamiento 11
2.4 Clasificación 12
2.4.1. Fase de entrenamiento 12
2.4.2. Fase de predicción 12
2.5. Post clasificación 13
2.6. Evaluación de la Clasificación 14
3. Análisis de concordancia **16**
3.1. Cálculo de área para sorteo de puntos 16
3.2. Cálculo del tamaño de muestra 17
3.3. Generación de métricas 17
4. Productos derivados de los mapas **21**
4.1. Área quemada anual 21
4.2. Área quemada mensual 21
4.3. Área quemada anual por tamaño de cicatriz 21
4.4. Área quemada acumulada 21
4.5. Datos de frecuencia de incendios 21
4.6. Año de la última ocurrencia de incendio 22
5. Consideraciones prácticas y desafíos **22**
6. Referencias **22**

Lista de Tablas

Tabla 1. Proporción relativa de área para sorteo de puntos por clases

Tabla 2. Disponibilidad de datos multifuente para periodo 1990 al 2025

Tabla 3. Matriz de confusión

Lista de Figuras

Figura 1. Límite de biomas identificados para Perú

Figura 2. Síntesis metodológica de la Colección 2.0

Figura 3. Mapa del mosaico de calidad (QM) de 2025 para Perú (RGB SWIR-1, NIR, Red), con ejemplos de cicatrices: (1) Bosque seco ecuatorial, (2) Desierto costero, (2, 4, 5) Andes y (6, 7, 8, 9, 10) Amazonía.

Figura 4. Ejemplos de clasificación de áreas quemadas (2025) para diferentes tipos de cicatrices, con el mosaico de Landsat utilizado para la clasificación y el área clasificada como quemada, en rojo

Figura 5. Comparaciones de SERFOR (10 m), MODIS MCS64A1 (500M), FIRE CCI (250m), ASQ2 LATAM (10m), GABAM (30m), FIRMS (1000m), Landsat 7/8 Mosaic - 2019 y Colección 1 MapBiomas Fuego Perú (30m)

Figura 6. Área quemada anual multireferencia: FIRE CCI (resolución: 250 m), MODIS MCD64A1 (resolución: 500 m), SERFOR (resolución: 10 m), AQS2 LATAM (resolución: 10 m), GABAM (resolución: 30 m), Colección 1 MapBiomas Fuego (resolución: 30 m) y FIRMS (resolución: 1000 m)

1. Introducción

1.1. Objetivo del documento

Este documento describe la base teórica, la justificación y los métodos aplicados para producir mapas anuales de áreas quemadas desde el año 1990 hasta 2025 de la Colección 2 de mapas anuales de áreas quemadas del Perú, en el marco de la iniciativa MapBiomas Fuego.

Con este documento, el usuario conocerá todos los métodos de clasificación de imágenes Landsat (L5, L7, L8, L9), la arquitectura de procesamiento de imágenes y el enfoque para integrar los biomas y regiones presentes en el país. También se presenta un contexto histórico y antecedentes, así como una descripción general del conjunto de datos de imágenes satelitales.

1.2. MapBiomas Fuego Perú

La red MapBiomas Perú inicia en julio de 2021 con el propósito de apoyar el entendimiento de la dinámica de la cobertura y uso del suelo (LULC, por sus siglas en inglés) en todo el territorio nacional. Posteriormente, implementó su segunda iniciativa llamada MapBiomas Agua, con la finalidad de obtener datos anuales y mensuales de cuerpos de agua y glaciares en todo el país. En 2025, la red implementó su tercera iniciativa, MapBiomas Fuego, brindando a los usuarios información gratuita de mapas anuales de áreas quemadas a nivel nacional.

La red se desarrolla en el marco de: i) el uso de tecnología para el procesamiento en la nube de grandes cantidades de datos espaciales mediante algoritmos de **Google Earth Engine**; ii) la producción de datos y metodologías totalmente gratuitas, replicables y mejorables cada año (colecciones); iii) el equipo técnico multidisciplinario que con su experiencia lleva a cabo el mapeo del territorio; y iv) el apoyo de instituciones y financiadores que apoyan la visión de la red.

Tras el lanzamiento oficial de su primera colección en octubre de 2025 —la cual comprendió un periodo de 12 años (2013–2024)—, MapBiomas Fuego Perú presenta la **Colección 2**, logrando un avance significativo al extender la serie histórica a 36 años de monitoreo continuo (1990–2025). La metodología utilizada empleó algoritmos de Deep Learning mediante el modelo Deep Neural Network (DNN), integrando imágenes de 30 metros de resolución espacial de la constelación de satélites Landsat (L5, L7, L8 y L9). Los responsables de la co-creación de esta iniciativa son el Laboratorio de la Facultad de Ciencias Forestales de la Universidad Nacional Agraria La Molina (antes Sala de Observación del Perú, creada por la OTCA en 2012) y el Instituto del Bien Común.

Las colecciones MapBiomas pretenden contribuir al desarrollo de un método rápido, fiable colaborativo y de bajo coste para procesar conjuntos de datos a gran escala y generar series temporales históricas de mapas anuales. Todos los datos, procesos, códigos, estadísticas y otros análisis son de libre acceso a través de la plataforma MapBiomas Perú (https://peru.mapbiomas.org/)

Los datos generados para la **Colección 2.0** son los siguientes:

- Áreas quemadas anuales a nivel nacional del periodo 1990 al 2025 (36 años).
- Áreas quemadas según las clases de Uso y Cobertura del Suelo de MapBiomas.
- Frecuencia anual de áreas quemadas en Perú.
- Área quemada mensual en Perú de 1990 a 2025.
- Áreas quemadas acumuladas en Perú.
- Área quemada anual por intervalo de tamaño de cicatriz de incendio.

- Año de la última ocurrencia del incendio.;

Los productos pueden ser consultados de manera gratuita en la pagína web de la red peru.mapbiomas.org El presente documento se encuentra alojado en la sección (https://peru.mapbiomas.org/iniciativas-y-productos/fuego/mapeo-anual/) y los algoritmos de clasificación están disponibles en el Github de MapBiomas Perú (https://github.com/mapbiomas-peru).

1.3. Google Earth Engine, Google Colab y Google Cloud

Los mapas de áreas quemadas de MapBiomas Fuego Perú son procesados gracias a las herramientas de Google, tales como Google Earth Engine, Google Colab y Google Cloud.

Google Earth Engine (GEE) es una plataforma en la nube para el análisis científico de conjuntos de datos geoespaciales a escala global. Los principales componentes de GEE son: 1) Catálogos de datos de sensoramiento remoto públicamente disponibles, 2) Infraestructura de procesamiento computacional en la nube a escala petabyte, 3) APIs para JavaScript y Python para interactuar con los servidores del GEE, y 4) y un ambiente de desarrollo integrado (IDE) en línea que permite el desenvolvimiento de aplicaciones mediante scripts[1] denominado el Code Editor[2].

Los componentes mencionados hacen de GEE la herramienta idónea para el procesamiento de grandes cantidades de datos que requiere la clasificación de imágenes satelitales con fines de mapeo de áreas quemadas. Adicionalmente, GEE permite la aplicación de metodologías compartidas a un mismo set de datos, y la replicación de procesos por parte de terceros interesados en evaluar conceptos o metodologías. Para complementar esta capacidad, **Google Colab** es una herramienta potente y flexible que nos permite realizar tareas complejas de clasificación, aprovechando la infraestructura de computación en la nube para procesar grandes volúmenes de información.

Todo este flujo de trabajo se beneficia de la infraestructura de almacenamiento de **Google Cloud** que posee la capacidad de almacenar y organizar los datos de manera estructurada en Google Cloud Storage mediante buckets. Este proceso es fundamental para garantizar la replicabilidad y mejorar la gestión del proyecto. El uso de buckets para organizar los datos permite un control de versiones y un acceso seguro a los insumos y resultados.

A todo este flujo de trabajo se le suma la visualización y análisis de datos a través de **Looker Studio**, una potente herramienta que permite transformar los datos procesados en la nube en informes y paneles interactivos. Looker Studio facilita la creación de gráficos dinámicos y mapas interactivos, permitiendo al equipo técnico explorar, analizar e interpretar la información sobre las áreas quemadas de forma intuitiva, así como comparar los resultados con datos de referencia.

1.4. Otras iniciativas de mapeo

En los últimos años se han desarrollado diversas herramientas para el mapeo de áreas quemadas, algunas de ellas de alcance global, con enfoques cada vez más robustos. Todos tienen en común su interés en aportar en uno de los temas más urgentes actualmente: conocer el estado actual de la afectación de los incendios en el contexto del cambio climático. Estas iniciativas han aportado a la detección de la deforestación, el monitoreo de los ecosistemas terrestres y acuáticos, y a fortalecer la

[1] https://developers.google.com/earth-engine/
[2] code.earthengine.google.com

detección de la degradación forestal, la conservación, la detección de focos de calor, etc. A continuación enlistamos las iniciativas más relevantes.

- **MODIS Fire and Thermal Anomalies (NASA/FIRMS)**, es una iniciativa que usa datos satelitales MODIS (Terra/Aqua) y VIIRS (Suomi NPP, NOAA-20) para identificar focos de calor casi en tiempo real. El portal FIRMS (Fire Information for Resource Management System) permite la descarga y visualización global de los focos de calor. La detección de focos de calor se basa en algoritmos de temperatura y brillo espectral, y se actualiza cada 3 horas dependiendo de la órbita donde esté el sensor. Su cobertura incluye toda la superficie terrestre.
- **Fire Information for Resource Management System (FIRMS) – Global**, es un servicio de la NASA para detectar fuegos activos a través de imágenes térmicas. Incluye API y acceso mediante Google Earth Engine y plataformas GIS. Es ampliamente utilizada por gobiernos, ONGs, investigadores y profesionales en gestión de recursos naturales, monitoreo ambiental, respuesta a emergencias, investigación y planificación territorial.
- **Global Forest Watch Fires (GFW Fires)**, es una plataforma del World Resources Institute (WRI) que combina datos satelitales (VIIRS, MODIS) con capas de cobertura boscosa y alertas de incendios, especialmente útil en regiones tropicales. Está diseñada para detectar, visualizar y analizar incendios activos, quemas agrícolas y deforestación en tiempo cuasi real, especialmente en regiones tropicales como la Amazonía (Perú, Brasil, Colombia, Bolivia, etc.)
- **Copernicus Emergency Management Service (CEMS) – European Forest Fire Information System (EFFIS)**, es un sistema europeo que monitorea incendios forestales en Europa, el Mediterráneo y América Latina, usando datos de Sentinel-2, Sentinel-3, MODIS, VIIRS y Meteosat/MSG. El objetivo de EFFIS es proporcionar información precisa, actualizada y armonizada sobre incendios forestales para la prevención, detección, monitoreo y evaluación de impacto en Europa y regiones vecinas. Cuenta con una base de datos desde el año 2000 sobre eventos de incendios forestales.
- **Amazon Conservation – MAAP (Monitoring of the Andean Amazon Project)**, es una iniciativa regional liderada por la ONG Amazon Conservation (ACCA / Conservación Amazónica) con foco en la Amazonía (Perú, Brasil, Bolivia) que usa imágenes satelitales para mapear quemas, deforestación e incendios ilegales. Fue lanzado en 2015, desde entonces ha producido más de 200 informes que combinan ciencia, datos satelitales y visualización geográfica para facilitar la toma de decisiones basada en evidencia.
- **GEOGloWS / GEO Global Agricultural Monitoring (GEOGLAM-Fire)**, es una iniciativa dentro del Group on Earth Observations (GEO) que integra datos de incendios para evaluar impacto en cultivos y ecosistemas agrícolas. Algunas de sus aplicaciones incluyen detectar y cuantificar incendios sobre tierras de cultivo, cruce entre focos de calor VIIRS y capas agrícolas (ej. arroz, soya, maíz) para evaluar pérdidas, y estimar emisiones por quema de biomasa agrícola.
- **GOFC-GOLD Fire Mapping and Monitoring**, es un grupo internacional respaldado por NASA y FAO para desarrollar estándares y redes de datos sobre incendios a escala global. Apoya redes regionales de observación satelital. Su objetivo es mejorar la capacidad global y regional para observar, mapear, modelar y comprender los regímenes de fuego en ecosistemas naturales y productivos, a través de datos satelitales, metodologías estandarizadas y redes colaborativas.
- **Global Wildfire Information System (GWIS)**, es una iniciativa conjunta entre Copernicus y FAO que proporciona datos integrados sobre incendios forestales en una sola plataforma global. El objetivo de GWIS es armonizar la información mundial sobre incendios forestales y mejorar la capacidad de los países para monitorear, prevenir y gestionar eventos relacionados con el fuego. Su cobertura geográfica incluye todos los continentes, con un foco especial en áreas críticas (Amazonía, África subsahariana, Europa mediterránea, Australia y Asia Central.)
- **AFIS (Advanced Fire Information System – África)**, es un sistema de monitoreo de incendios en África, desarrollado por CSIR y NASA, pero replicable en otras regiones. Está enfocado en

alertas tempranas y gestión forestal. Es ampliamente utilizada en gestión ambiental, prevención de desastres, protección de áreas naturales y planificación agrícola y forestal en contextos vulnerables al fuego. El sistema cuenta con alertas personalizadas de incendio por mensaje de texto, correo electrónico o notificaciones según el área de interés.

- **INPE – Queimadas (Brasil)**, es el sistema brasileño de monitoreo de focos de calor y ocurrencias de incendios forestales lanzado en 1998. Utiliza datos de sensores como MODIS, VIIRS, GOES, para detectar y reportar en casi tiempo real la localización y extensión de los focos de calor en todo el territorio brasileño y América del Sur. Los datos generados son públicos y se actualizan diariamente, permitiendo el monitoreo continuo de la dinámica del fuego.
- **Cicatrices y ocurrencias de incendios** (2017-2024), elaborado por el **Servicio Nacional Forestal y de Fauna Silvestre del Perú** (SERFOR). SERFOR es la autoridad nacional en materia forestal y de fauna silvestre en Perú. Está adscrito al Ministerio de Desarrollo Agrario y Riego (MIDAGRI), y dirige el Sistema Nacional de Gestión Forestal y de Fauna Silvestre (SINAFOR). Ellos monitorean anualmente las ocurrencias de incendios forestales a nivel nacional.

2. Metodología

La cadena de procesamiento de MapBiomas está basada enteramente en tecnología de Google y, como tal, la iniciativa MapBiomas Perú se ha apoyado en GEE para 1) procesar imágenes satelitales en infraestructura computacional en la nube, 2) desarrollar códigos (scripts) en Javascript y Python, 3) almacenar los datos generados usando almacenamiento en la nube, y 4) desplegar los resultados (mosaicos, mapas anuales de áreas quemadas, estadísticas, información metodológica, entre otros) a través de una plataforma web de consulta pública accesible en http://plataforma.peru.mapbiomas.org.

El flujo de trabajo de la Colección 1 MapBiomas Perú se encuentra sintetizado en la **Figura 1** y es detallado en las siguientes secciones: Delimitación de regiones de trabajo (sección 3.1), Generación de mosaicos anuales (sección 3.2), Colecta de muestras de entrenamiento (Sección 3.3), Clasificación (sección 3.4), Post clasificación (sección 3.5) y Evaluación de la clasificación (sección 3.6).

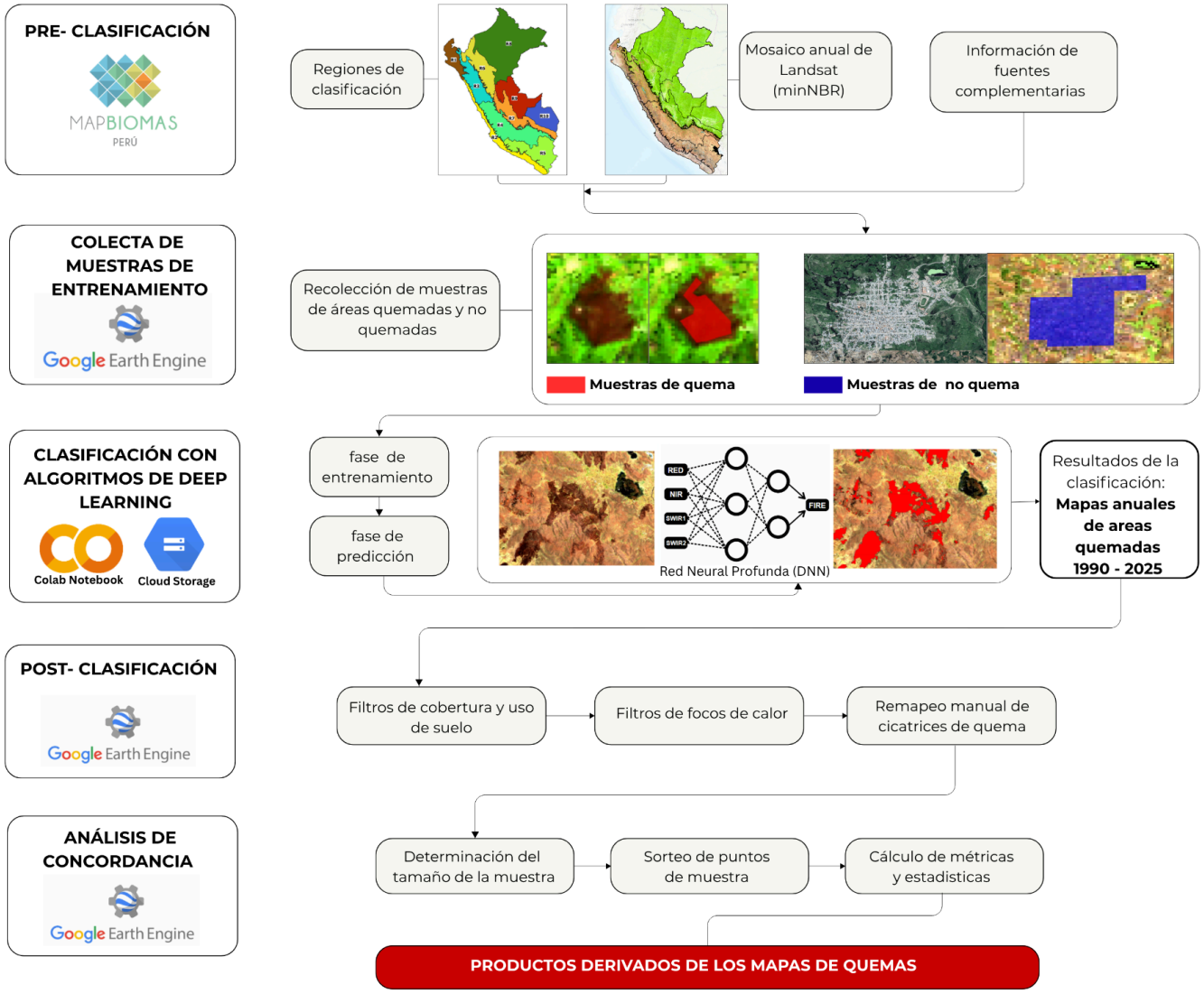


Figura 1. Síntesis metodológica de la Colección 2.0

2.1 Delimitación de regiones de trabajo

El área de trabajo abarca el territorio continental peruano, que tiene una extensión aproximada de 129 430 615 hectáreas y excluye la superficie insular (0.01%).

Debido a la diversidad del territorio peruano en términos de coberturas y usos de suelo, así como la distribución heterogénea de las áreas quemadas, se decidió dividir el país en 10 regiones de clasificación para facilitar el análisis (Figura 2). Esta segmentación es necesaria para manejar la complejidad inherente a los algoritmos de aprendizaje automático y la heterogeneidad de los biomas previamente definidos por MapBiomas Perú: Bosque seco ecuatorial, Desierto costero, Andes y Amazonía.

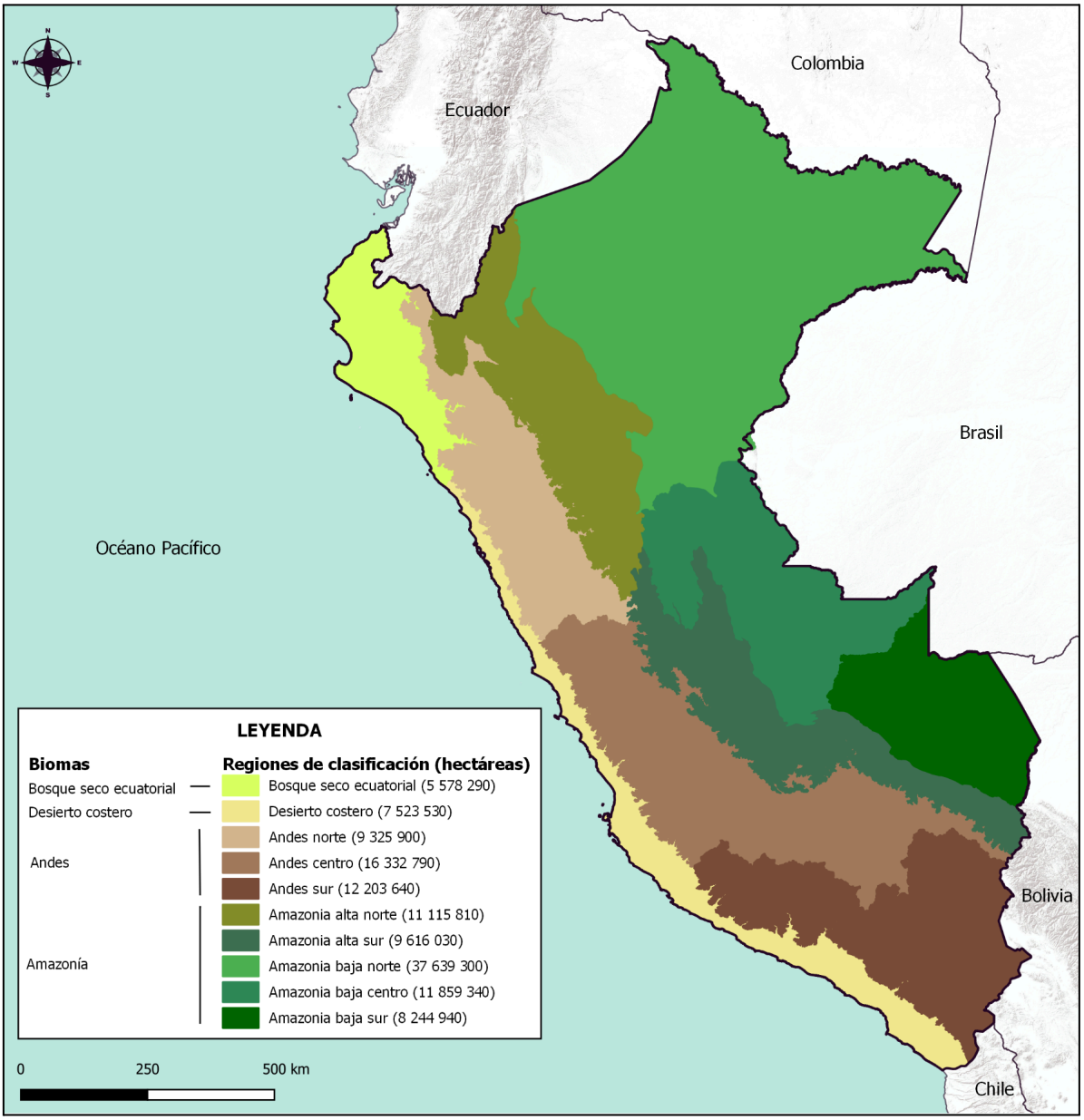


Figura 2. Límite de regiones de clasificación por bioma identificados para Perú

Para el proceso se utilizaron criterios técnicos sobre la distribución de ecosistemas a partir de mapas nacionales de referencia, distribución de las áreas quemadas a nivel de territorio y a la experiencia de las colecciones de cobertura y uso del suelo. Las coberturas utilizadas fueron:

- Mapa de cobertura vegetal (MINAM, 2015) y
- Mapa de cobertura y uso del suelo Colección 3 (MapBiomas Perú, 2025).

2.2 Generación de mosaicos anuales

A través de la plataforma de Google Earth Engine, se utilizan datos de reflectancia de superficie (SR) del catálogo de Landsat[3] C2 con nivel de corrección Tier 1 (30m × 30m), para el periodo de 1990 a 2025 a bordo de los satélites Landsat 5 (1990–1998), Landsat 7 (L7, 1999–2021), Landsat 8 (L8, 2013–2025) y Landsat 9 (L9, 2022 en adelante), con un intervalo de retorno de 16 días.

La reflectancia superficial de Landsat viene acompañada de dos bandas de evaluación de calidad bit a bit (**QA_PIXEL** y **QA_RADSAT**) que indican los píxeles con problemas radiométricos e instrumentales, incluyendo un indicador de probabilidad. Utilizamos la banda **QA_PIXEL** para seleccionar y enmascarar los píxeles con altos niveles de confianza (67–100%) de 'nube' y 'sombra'. Luego, empleamos **QA_RADSAT** para evitar píxeles con saturación radiométrica en cualquier banda de reflectancia superficial. Finalmente, descartamos los píxeles con valores negativos en la reflectancia superficial para eliminar anomalías y ruido en la composición del mosaico de calidad anual.

Utilizamos un **enfoque estadístico anual** para resumir esta cantidad de datos y optimizar la clasificación sin descartar información espectral a nivel de píxel. Este enfoque nos permitió crear mosaicos anuales mediante la composición de todas las imágenes de 16 días en un único mosaico de calidad (QM), usando el **Índice espectral NBR mínimo (Normalized Burn Ratio)** (ecuación 1 — Key y Benson, 2006) como función de ordenamiento por píxel. Se seleccionó el píxel con el valor NBR más bajo, y todas sus características de reflectancia espectral, incluyendo los metadatos de la escena con la fecha de ese píxel seleccionado, se utilizaron para crear el mosaico de calidad anual.

$$\lambda QM=[Azul,Verde,Rojo,NIR,SWIR1,SWIR2]=\text{fecha en la que se obtuvo el min}(\lambda NIR+\lambda SWIR1\lambda NIR-\lambda SWIR1)[xi...j]$$

Donde λ representa los valores de reflectancia de las bandas de calidad que componen el mosaico de calidad (QM), recuperados de la fecha en que cada píxel alcanzó su valor NBR mínimo (mín) en un año determinado (x), considerando el conjunto de todas las escenas disponibles, desde la primera (i) hasta la última (j). λNIR es la reflectancia superficial del infrarrojo cercano y λSWIR1 es la reflectancia superficial del infrarrojo de onda corta utilizada para calcular el índice espectral NBR.

En otras palabras, calculamos el NBR para cada píxel con una observación válida dentro de un año específico y los apilamos en una imagen multibanda. Se seleccionaron los píxeles con el NBR más bajo dentro de la imagen multibanda, y su información espectral (Tabla 1) se utilizó para componer el mosaico de calidad anual (QM). Además de la información espectral, conservamos la información de metadatos de la escena, incluyendo la fecha en la que cada píxel mostró su valor NBR más bajo. El mosaico de calidad NBR creado con la información espectral del NBR mínimo tuvo un buen desempeño en la diferenciación de usos y coberturas del suelo quemados y no quemados en las regiones de clasificación utilizadas para el territorio peruano (Figura 3).

[3] https://developers.google.com/earth-engine/datasets/catalog/landsat/

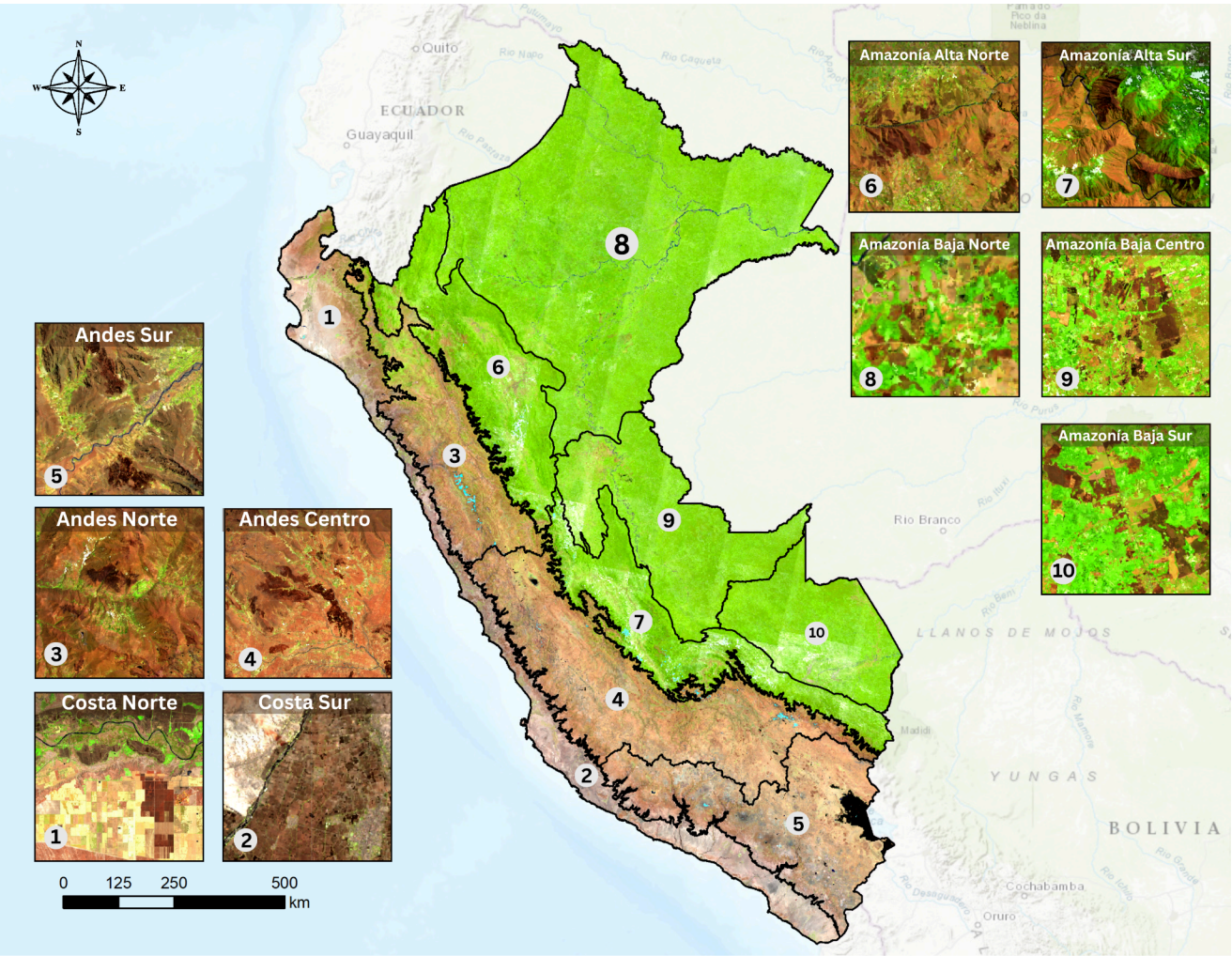


Figura 3. Mapa del mosaico de calidad (QM) de 2025 para Perú (RGB SWIR-1, NIR, Red), con ejemplos de cicatrices: (1) Bosque seco ecuatorial, (2) Desierto costero, (3, 4, 5) Andes y (6, 7, 8, 9, 10) Amazonía.

2.3. Colecta de muestras de entrenamiento

Creamos una librería espectral basada en la delineación manual de áreas quemadas y no quemadas para ser utilizada como muestras de entrenamiento. Estas muestras fueron estratificadas por sensores Landsat (recopiladas en diferentes años) y por cada bioma/región. La recolección de muestras de entrenamiento se realizó en las 10 regiones de clasificación, asegurando la representación de las distintas características espectrales presentes en cada región. Finalmente, dividimos nuestra librería espectral en 10 paquetes (uno para cada región de clasificación) y la usamos como entrada en el paso de clasificación.

Algunos insumos que se utilizaron durante esta etapa fueron:

- Cartas nacional de 55x55 km, que nos permitieron obtener muestras distribuidas en cada región de forma homogénea;
- Google Earth Pro, que nos permitió visualizar el color verdadero de muchas zonas que podían generar confusión o con la finalidad de comprobar que la muestra a colectar pertenecía a la categoría correspondiente;
- Colección 4 de MapBiomas Cobertura y uso del suelo (1990 - 2025), para identificar los diferentes tipos de cobertura y uso;

- Cicatrices de incendios del SERFOR (2016-2024);
- Cicatrices de Áreas quemadas de Sentinel-2 (mensuales y anuales entre el 2019 y el 2023).
- Mapa Global Anual de Áreas Quemadas (GABAM) (30m de resolución espacial) (1992, 1995, 1996, 1998-2021);
- Registro de Datos Climáticos de Área Quemada Fire CCI Versión 5.1 (2000-2020);
- Áreas quemadas de MODIS (MCD64A1) (500m de resolución espacial) (2000-2025);
- Cicatrices de incendios del Sistema de Información sobre Incendios para la Gestión de Recursos (FIRMS) (1km de resolución espacial) (2000-2025) y
- Focos de calor del satélite AQUA M-T del INPE (2000-2025).

2.4 Clasificación

El modelo de clasificación aplicado fue la **Deep Neural Network (DNN)**, un marco computacional capaz de realizar aprendizaje profundo y reconocimiento de patrones visuales. Específicamente, se implementó una estructura de **Multi-Layer Perceptron Network (MLPN)**, que incorpora múltiples capas de unidades computacionales, o "neuronas", interconectadas. Cada nodo (neurona) en una capa está conectado a un nodo en la siguiente, formando tres tipos principales de capas: de entrada, oculta y de salida (Hu & Weng, 2009).

Para este modelo DNN, las capas de entrada correspondieron a las bandas espectrales **RED, NIR, SWIR1 y SWIR2**, mientras que las capas de salida consistieron en dos clases: **quemado y no quemado**. El algoritmo de mapeo de áreas quemadas involucra dos pasos principales: entrenamiento y predicción.

2.4.1. Fase de entrenamiento

En la fase de entrenamiento, los siguientes parámetros fueron utilizados en base a pruebas realizadas previamente por MapBiomas Fuego Brasil: tasa de aprendizaje (0.001), tamaño de lote (1000), número de iteraciones (7000) e entradas para clasificación (Arruda et al. 2021). Las entradas de clasificación utilizadas en este modelo fueron los datos espectrales SR obtenidos de los mosaicos anuales de calidad, utilizando las muestras de entrenamiento de áreas quemadas y no quemadas.

Basándose en la biblioteca espectral de las muestras de entrenamiento de áreas quemadas y no quemadas, se utilizaron las siguientes bandas espectrales como entradas para el modelo de clasificación de áreas quemadas: rojo (RED—0,65 µm), infrarrojo cercano (NIR—0,86 µm) e infrarrojo de onda corta (SWIR1—1,6 µm y SWIR2—2,2 µm). Estas bandas espectrales de Landsat fueron seleccionadas por su sensibilidad a los eventos de incendio en distintos usos y coberturas del suelo.

Los datos de entrada para el entrenamiento se dividieron en dos conjuntos de datos: el 70% de las muestras se utilizaron para el entrenamiento y el 30% para la prueba, con el fin de estimar la capacidad del algoritmo DNN para mapear áreas quemadas con precisión.

2.4.2. Fase de predicción

La clasificación se realizó utilizando los mosaicos anuales de calidad de Landsat para cada una de las regiones y para cada sensor (Landsat 5, Landsat 7, Landsat 8 y Landsat 9). Esto resultó en 10 mapas de áreas quemadas para todo el Perú (Figura 4). Este enfoque nos permitió aprovechar las potentes capacidades de las DNN (redes neuronales profundas) para un mapeo preciso y eficiente de las áreas quemadas en los diversos biomas/regiones de Perú.

Se aplicó un filtro espacial para eliminar el ruido y definir un área mínima de mapeo. Se eliminaron las áreas quemadas más pequeñas o iguales aproximadamente a 0.36 hectáreas (4 píxeles). Este método aseguró la eliminación de píxeles de ruido aislados.

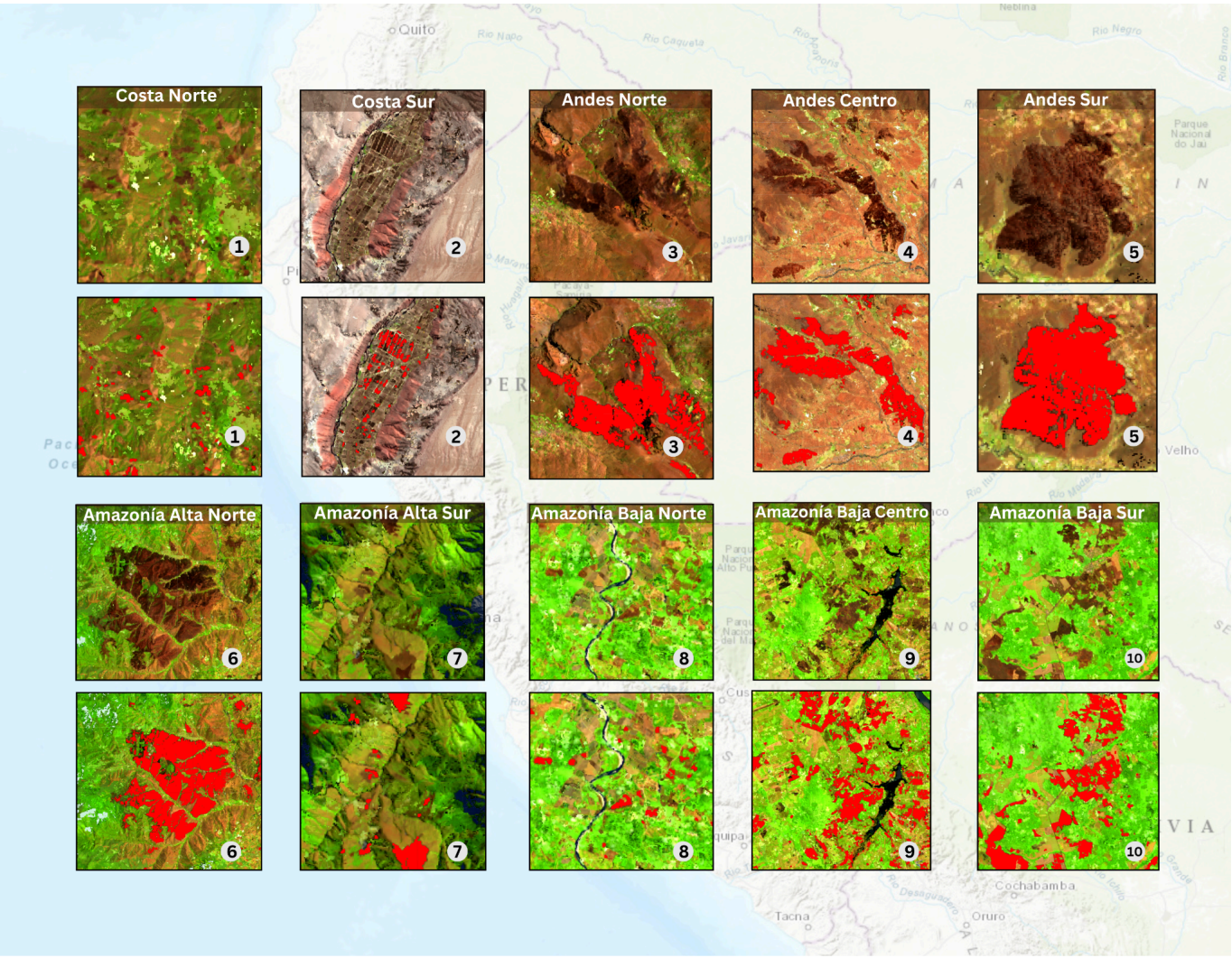


Figura 4. Ejemplos de clasificación de áreas quemadas (2025) para diferentes tipos de cicatrices, con el mosaico de Landsat utilizado para la clasificación y el área clasificada como quemada, en rojo.

2.5. Post clasificación

Después de evaluar los resultados de la clasificación, también se aplicaron máscaras de post-clasificación para reducir los errores de comisión de tipos de uso y cobertura del suelo con firmas espectrales similares a las de áreas recientemente quemadas, como cuerpos de agua, áreas urbanas y algunos tipos de cobertura específicos por región/bioma. Definimos reglas específicas por región/bioma para eliminar los píxeles que fueron clasificados erróneamente como quemados dentro de las distintas clases de cobertura y uso del suelo de la Colección 3 de MapBiomas. Las reglas para cada bioma se detallan a continuación.

Región	Nombre de región	Clases
r1	Costa Norte	(24), (33), (68) y (25)
r2	Costa Sur	(24), (25), (30), (11), (13), (70), (68), (4) y (33)

r3	Andes Norte	(24), (25), (29) y (30)
r4	Andes Centro	(23), (24), (25), (29), (31), (33), (34), (61), (68) y (11)
r5	Andes Sur	(24), (25), (29), (68) y (33)
r6	Amazonía Alta Norte	(23), (24), (25), (68) y (33)
r7	Amazonía Alta Sur	(23), (24), (25), (26), (68),(30) y (33)
r8	Amazonía Baja Norte	(24) y (33)
r9	Amazonía Baja Centro	(24) y (33)
r10	Amazonía Baja Sur	(23), (24), (25), (31), y (33)

Un paso adicional del proceso de post-clasificación incluye un enmascaramiento de píxeles quemados (comisión) que no coincidan con focos de calor aplicado para los tipos de cobertura como matorral (66), pastizal/herbazal (12) y otra formación no boscosa (13), para las regiones de Andes (norte, centro y sur), que son donde priman estos tipos de vegetación. Para aquellos pixeles que coincidían con los focos de calor, se procedió a aplicar un buffer con radio que varía entre 2 y 5 km (según las necesidades de cada región) para no eliminar las cicatrices. Cabe resaltar que solo para la región Costa norte, fue necesario aplicar una máscara de polígonos debido a que la estacionalidad del bosque seco requiere un trabajo detallado del análisis sobre los errores por comisión que puedan surgir.

Finalmente se realizó el último paso del **procesamiento post-clasificación** que incluía recuperar la información de la fecha del píxel quemado del mosaico anual construido a partir del NBR mínimo, identificando el mes en que se mapea la cicatriz del incendio.

Adicionalmente se incorporó el uso de la herramienta ReMap dentro del Toolkit, para realizar una reclasificación de los píxeles que luego de todo el proceso aún sean considerados por error como quemados para que no sean considerados al final del proceso. Esta reasignación de píxeles se consideró como de valor 13 para todas las regiones, sin perder el valor del píxel quemado del mosaico anual identificado por mes.

2.6. Evaluación de la Clasificación

Las clasificaciones de cicatrices de incendios forestales y quemas fueron evaluadas tanto mediante inspección visual como a través de análisis estadísticos. La inspección visual trató de comparar detalladamente las cicatrices clasificadas con los mosaicos originales de Landsat para verificar la precisión. Las discrepancias identificadas durante estas inspecciones fueron registradas y utilizadas para refinar las muestras de los algoritmos. Se realizaron análisis estadísticos para validar los resultados de la clasificación, comparándolos con productos de áreas quemadas.

Estas inspecciones visuales se complementaron con el conocimiento de expertos locales y con conjuntos de datos adicionales para validar las clasificaciones de cicatrices de incendios y quemas. Las cicatrices clasificadas se compararon con múltiples conjuntos de datos de referencia, que sirvieron como línea base para la validación. Entre los productos de área quemada se encuentran MODIS MCD64A1 (500 m), los datos de cicatrices de SERFOR (10 m), los mapas globales de áreas quemadas de alta resolución de GABAM (30 m) y los productos globales de áreas quemadas como FireCCI (250 m) y FIRMS (1000 m). Esta evaluación, que integró métodos automatizados con evaluaciones guiadas por expertos, garantizó la solidez, calidad y fiabilidad de las clasificaciones de áreas quemadas (Figuras 5 y 6).

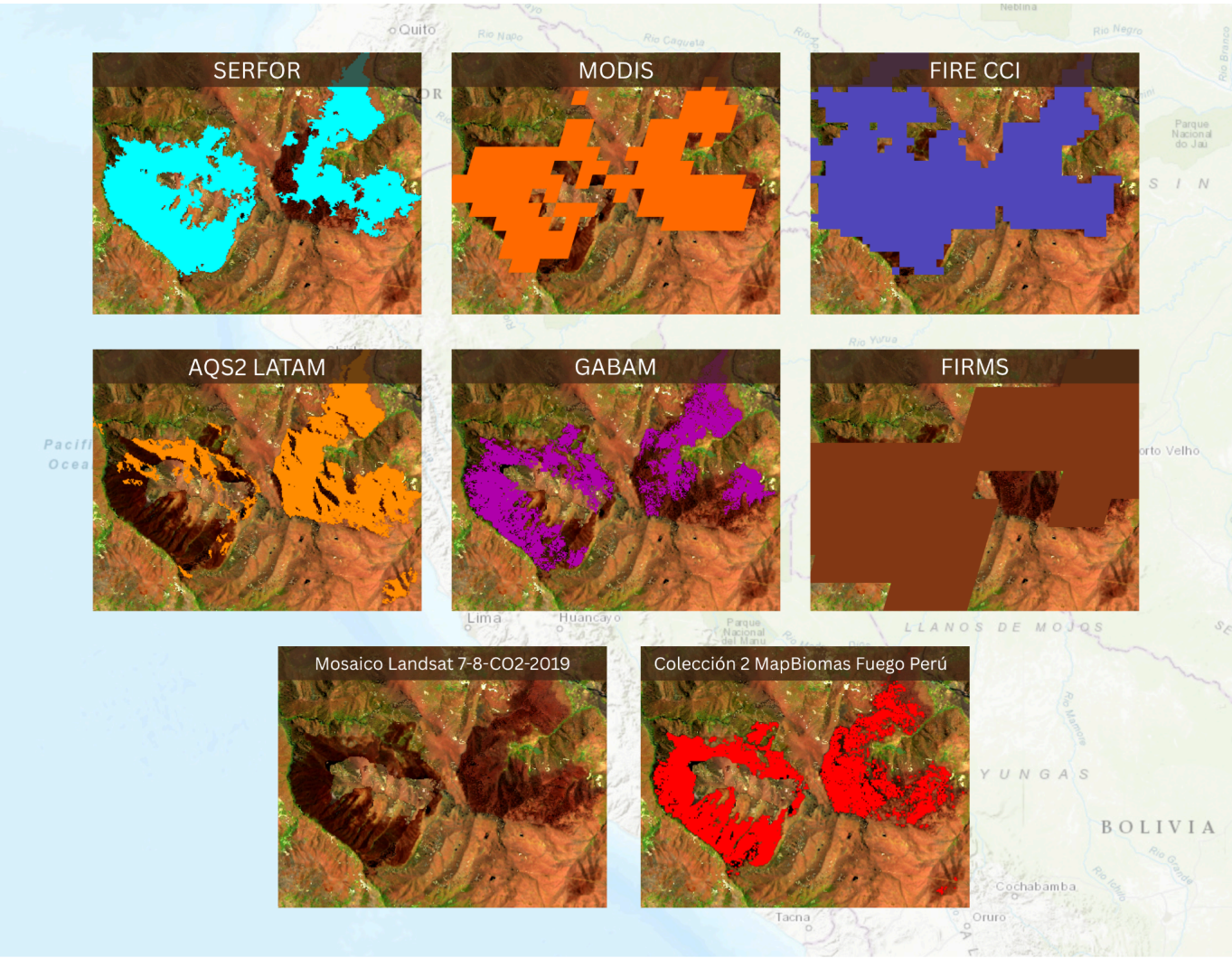


Figura 5. Comparaciones de SERFOR (10 m), MODIS MCS64A1 (500M), FIRE CCI (250m), ASQ2 LATAM (10m), GABAM (30m), FIRMS (1000m), Landsat 7/8 Mosaic - 2019 y Colección 1 MapBiomas Fuego Perú (30m).

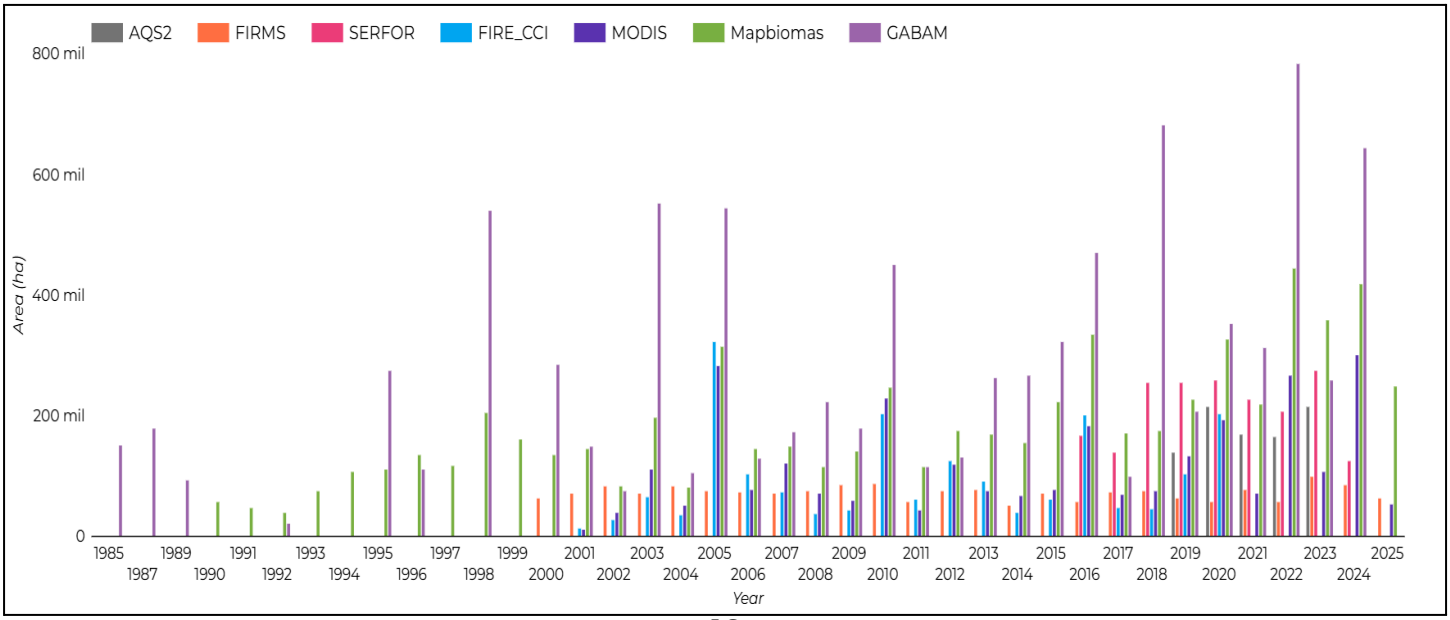


Figura 6. Área quemada anual multireferencia: FIRE CCI (resolución: 250 m), MODIS MCD64A1 (resolución: 500 m), SERFOR (resolución: 10 m), AQS2 LATAM (resolución: 10 m), GABAM (resolución: 30 m), Colección 1 MapBiomas Fuego (resolución: 30 m) y FIRMS (resolución: 1000 m)

3. Análisis de concordancia

Para evaluar la concordancia entre los resultados del mapa de áreas quemadas y la información de referencia (SERFOR, GABAM, Fire CCI, MODIS, FIRMS y Áreas quemadas de Sentinel-2 Anual), se llevó a cabo un análisis de concordancia diseñado para comparar los datos según las particularidades de cada región de clasificación, considerando sus tipos de cobertura, uso de suelo y topografía. Con este fin, se siguieron tres pasos esenciales: determinación del tamaño de la muestra, sorteo de puntos y cálculo de métricas para medir el grado de acuerdo entre los resultados obtenidos y la información de referencia. A continuación, se detalla el proceso utilizado para este análisis:

3.1. Cálculo de área para sorteo de puntos

Se delimitó el área de estudio para el sorteo de puntos, tomando en cuenta las áreas quemadas anualmente entre 1990 y 2025. Este enfoque nos permitió establecer un marco espacial y temporal preciso para el análisis.

Adicionalmente, se generó un estrato de "no fuego" para crear un área de control. Esta se definió con un buffer diferenciado alrededor de las cicatrices de fuego:

- Cicatrices pequeñas (≤ 8 píxeles): se aplicó un buffer de 30 metros.
- Cicatrices grandes (≥ 9 píxeles): se aplicó un buffer de 60 metros.

El estrato de "no fuego" no representa una zona sin incendios, sino más bien el área inmediata y externa a las cicatrices, lo que asegura que la comparación se realice en zonas con características similares y expuestas al riesgo de incendio. Este buffer diferenciado permite compensar la dispersión y el tamaño de las cicatrices, logrando una proporción equilibrada entre las clases "fuego" y "no fuego".

Finalmente, se calculó la proporción relativa de cada clase ("fuego" y "no fuego") dentro del buffer. Esto nos proporcionó una visión cuantitativa del equilibrio entre ambas clases, con un promedio general del 60% para la clase "fuego" y 40% para la clase "no fuego" en todas las regiones de clasificación para el periodo 1990 al 2025. En la Tabla 1, se muestran estos datos a nivel de región.

Tabla 1. Proporción relativa de área para sorteo de puntos por clases

Regiones	Clase Fuego (%)	Clase No Fuego (%)
Bosque seco ecuatorial	56,8	43,2
Desierto costero	53,7	46,3
Andes norte	59,6	40,4

Andes centro	60,2	39,8
Andes sur	62,3	37,7
Amazonía alta norte	58,7	41,3
Amazonía alta sur	60,7	39,3
Amazonía alta norte	54,8	45,2
Amazonía alta centro	58,8	41,2
Amazonía alta sur	60,2	39,8

3.2. Cálculo del tamaño de muestra

El cálculo de muestras debe ser estadísticamente significativo, por lo cuál se decidió establecer que la precisión fuera del 95% con un error del 2%. Para ello se utilizó la fórmula donde se calcula el **tamaño de muestra preliminar o inicial (n_0)**, asumiendo una población infinita o muy grande:

Donde:

- n_0: es el **tamaño de muestra inicial** (sin ajuste por población finita).
- Z: es el **valor crítico** de la distribución normal estándar para el nivel de confianza deseado.
- p: es la **proporción esperada** de la característica en la población.
- q: es la **proporción complementaria** (1−p).
- E: es el **margen de error permitido**.

Una vez que se tiene n_0, se aplica el **factor de corrección por población finita** para obtener el tamaño de muestra ajustado (n):

Donde:

- n: es el **tamaño de la muestra ajustado** para una población finita.
- n_0: es el **tamaño de muestra inicial** calculado previamente.
- N: es el **tamaño total de la población**.

Finalmente, la muestra total se reparte en dos estratos:

- 80% de los puntos en áreas de fuego.
- 20% de los puntos en áreas de no-fuego (anillo).

Este balance busca dar mayor peso a la clase positiva (cicatrices), sin dejar de considerar el contexto.

3.3. Generación de métricas

A continuación, se muestra la disponibilidad de las fuentes utilizadas para la generación de métricas de la matriz de confusión por año (Tabla 2):

Tabla 2. Disponibilidad de datos multifuente para periodo 1990 al 2025

Año	Áreas quemadas de Sentinel 2	Fire CCI	FIRMS	GABAM	MODIS	SERFOR
1990						
1991						
1992				X		
1993						
1994						
1995				X		
1996				X		
1997						
1998				X		
1999				X		
2000		X	X	X	X	
2001		X	X	X	X	
2002		X	X	X	X	
2003		X	X	X	X	
2004		X	X	X	X	
2005		X	X	X	X	
2006		X	X	X	X	
2007		X	X	X	X	
2008		X	X	X	X	
2009		X	X	X	X	
2010		X	X	X	X	
2011		X	X	X	X	
2012		X	X	X	X	
2013		X	X	X	X	

2014		X	X	X	X	
2015		X	X	X	X	
2016		X	X	X	X	X
2017		X	X	X	X	X
2018		X	X	X	X	X
2019	X	X	X	X	X	X
2020	X	X	X	X	X	X
2021	X		X	X	X	X
2021	X		X		X	X
2022	X		X		X	X
2023	X		X		X	X
2024			X		X	X
2025			X		X	

Todas estas fuentes entregan información sobre áreas quemadas, aunque con distintas metodologías (cicatrices anuales, anomalías térmicas, bases nacionales, etc.).

Para evaluar la concordancia entre los puntos de validación (verdad de terreno) y cada mapa de referencia (fuente satelital o institucional), se utiliza una matriz de confusión (Tabla 3).

- Lo que se observa en la realidad (clase Real: fuego o no fuego).
- Lo que indica el mapa (Predicho: fuego o no fuego).

El resultado se organiza en cuatro posibles casos:

- **TP (Verdaderos Positivos):** el mapa identifica fuego donde realmente hubo fuego.
- **TN (Verdaderos Negativos):** el mapa identifica no fuego donde realmente no hubo fuego.
- **FP (Falsos Positivos):** el mapa marca fuego donde en realidad no lo hubo.
- **FN (Falsos Negativos):** el mapa no detecta fuego donde sí ocurrió.

La matriz se organiza así:

Tabla 3. Matriz de confusión

	Predicho:No fuego	Predicho:Fuego
Real: No fuego	TN (aciertos en no fuego)	FP (falsos positivos)
Real: Fuego	FN (falsos negativos)	TP (aciertos en fuego)

A partir de esta tabla se calculan las métricas de exactitud que permiten cuantificar qué tan bien funciona cada mapa.

Para evaluar qué tan bien coinciden los mapas de áreas quemadas con la "verdad de terreno" representada en los puntos de validación, se utilizan indicadores estadísticos clásicos derivados de la matriz de confusión. :

- **Exactitud Global (OA)**
  Indica la proporción de aciertos totales del mapa, es decir, cuántas veces clasificó correctamente un píxel como fuego o no-fuego respecto al total evaluado.proporción de aciertos sobre el total.

$$OA = \frac{TP + TN}{TP + TN + FP + FN}$$

  Donde:
  TP = verdaderos positivos
  TN = verdaderos negativos
  FP = falsos positivos y
  FN = falsos negativos

- **Kappa de Cohen (κ)**
  Mide el nivel de concordancia entre el mapa y la referencia, pero corrigiendo el efecto del acuerdo esperado por azar (Cohen, 1960).

$$k = \frac{p_0 - p_e}{1 - p_e}$$

  Donde $p_0 = OA$ y p_e es la probabilidad de acuerdo esperado por azar.

- **Exactitud del Productor (PA)**
  Evalúa la capacidad del mapa para detectar correctamente cada clase, también conocida como *recall*.

$$PA_{no\ fuego} = \frac{TN}{TN + FP}, \quad PA_{fuego} = \frac{TP}{TP + FN}$$

  Por ejemplo, PA_{fuego} muestra qué proporción de los píxeles que realmente fueron quemados fueron identificados como tales en el mapa.

- **Exactitud del Usuario (UA)**
  Indica la confiabilidad de cada clase, es decir, qué tan seguros podemos estar de que un píxel clasificado en el mapa corresponde realmente a esa clase (*precisión*)

$$UA_{no\ fuego} = \frac{TN}{TN + FN}, \quad UA_{fuego} = \frac{TP}{TP + FP}$$

Estas métricas se calculan año por año y para cada fuente de referencia (MODIS, FireCCI, FIRMS, MapBiomas, etc.), permitiendo comparar el desempeño relativo de cada conjunto de datos.

4. Productos derivados de los mapas

Con base a los mapas anuales de áreas quemadas, se calculan las subproductos de superficie en hectáreas (ha) tanto anuales como mensuales de las áreas quemadas

4.1. Área quemada anual

Datos de **área quemada anual** desde 1990 hasta 2025, que indican las áreas mapeadas como quemadas para cada año. También incluye la **Cobertura Quemada Anual**, que representa el área quemada anual para cada clase de uso y cobertura del suelo, donde cada píxel contiene el valor del código de uso y cobertura del suelo (Colección 3) de la clase que se quemó.

4.2. Área quemada mensual

Datos de **área quemada mensual** que cubren el período de 1990 hasta 2025, con los datos de cada píxel recuperados del mosaico anual basado en la fecha de la imagen satelital. Los datos de **Área Quemada Mensual** indican el mes (1 a 12) en que ocurrió el incendio para cada píxel dentro del período de 1990 hasta 2025.

4.3. Área quemada anual por tamaño de cicatriz

Datos anuales sobre los **tamaños de cicatrices de áreas quemadas** desde 1990 hasta 2025, categorizados en 8 intervalos de tamaño diferentes. Una cicatriz única se define como todos los píxeles que estuvieron conectados continuamente durante el año actual.

4.4. Área quemada acumulada

Datos de **área quemada acumulada** construidos incrementando el área quemada cada año, contando cada píxel solo una vez, independientemente de múltiples ocurrencias de incendios. Estos datos representan áreas con al menos un evento de incendio durante diferentes períodos. El **Área Quemada Acumulada por Uso y Cobertura del Suelo** representa el área total con al menos un evento de incendio para diferentes períodos, clasificada según las clases de Uso y Cobertura del Suelo de la Colección 3 de MapBiomas del último año.

4.5. Datos de frecuencia de incendios

Los mapas de **frecuencia de área quemada** representan cuántas veces el mismo píxel fue mapeado como quemado durante un período. Los datos de frecuencia de incendios se agregan en un solo mapa con 11 clases para el período de 1990 hasta 2025. La clase 1 representa píxeles que se quemaron una vez, la clase 2 representa píxeles que se quemaron dos veces, y así sucesivamente. Estos datos

también incluyen los respectivos usos y coberturas del suelo de la Colección 3 de MapBiomas para el último año.

4.6. Año de la última ocurrencia de incendio

Los datos de **Año de la Última Ocurrencia** de Incendio representan el año en que cada píxel se quemó por última vez, cubriendo el período de 1990 hasta 2025. Cada píxel contiene el valor del año en que fue mapeado por última vez como quemado, hasta el año correspondiente.

5. Consideraciones prácticas y desafíos

La colección 2.0 de mapas de cicatrices de áreas quemadas en el Perú es un instrumento de monitoreo estratégico que refleja la historia del país, en 36 años. La producción de este volumen de información multianual ha dado lugar a aplicaciones para estimar tendencias en el monitoreo de las cicatrices, así como para comprender los factores que modifican la dinámica del fuego.

Para el desarrollo de este proyecto, con un alcance espacial y temporal inédito, se utilizó una metodología estandarizada que es factible de ser replicada en otras áreas del planeta. El uso de las plataformas de trabajo en la nube de Google Earth Engine y la tecnología de código abierto se ha mostrado prometedor para la accesibilidad y el procesamiento de datos a gran escala.

A través del aprendizaje y la experiencia adquirida en la producción de esta colección, junto con el intercambio de ideas con los equipos de las diversas iniciativas de MapBiomas, fue posible lograr una mayor eficiencia en términos de tiempo y procesos. A partir del trabajo colaborativo y en red de un equipo multidisciplinar, fue posible llegar a una metodología ajustada a las necesidades particulares de cada territorio.

El uso de algoritmos de *Deep learning* utilizando el modelo Deep Neural Network (DNN) como clasificador, combinado con un protocolo de mapeo flexible, permite a cada país de las diversas iniciativas definir su espacio de características y muestras.

La aplicación de filtros de post clasificación permitió reducir los efectos asociados a la baja calidad y baja disponibilidad de imágenes de satélite que ocurren principalmente al inicio de la serie temporal.

El siguiente paso de este proyecto es ampliar la serie histórica anual de cicatrices de fuego hasta el año 1985.

6. Referencias

- Arruda, V. L. S.; Piontekowski, V. J.; Alencar, A.; Pereira, R. S.; Matricardi, E. A. T. An alternative approach for mapping burn scars using Landsat imagery, Google Earth Engine, and Deep Learning in the Brazilian Savanna. Remote Sensing Applications: Society and Environment, [s. l.], v. 22, p. 100472, 2021.
- European Space Agency (ESA). (2021). *Fire_cci Burned Area Climate Data Record v5.1* [Data set]. European Space Agency. Recuperado de https://esa-fire-cci.org/

- 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 Land Processes Distributed Active Archive Center. https://doi.org/10.5067/MODIS/MCD64A1.006 Date Accessed: 2025-09-15

- Instituto Nacional de Pesquisas Espaciais (INPE). (2024). Focos de calor del satélite AQUA M-T (2013-2014). [Conjunto de datos]. Recuperado de http://www.inpe.br/queimadas/bdqueimadas/

- Long Tengfei; Zhang Zhaoming; He Guojin, 2021, "30 m Resolution Global Annual Burned Area Product", https://doi.org/10.7910/DVN/3CTMKP, Harvard Dataverse, V1

- MapBiomas. (2025). *MapBiomas Colección 3 de cobertura y uso del suelo*. [Conjunto de datos].

- MapBiomas. (2026). *Colección 4.0 de la Cobertura y Uso del Suelo de Perú* [Conjunto de datos].

- Ministerio del Ambiente. 2015. Mapa de Cobertura Vegetal del Perú. Memoria Descriptiva - Perú. 108p. https://sinia.minam.gob.pe/repositorio-de-mapas?page=11

- NASA. (2024). *Fire Information for Resource Management System (FIRMS)*. NASA Earth Science Data and Information System (ESDIS). Recuperado de https://earthdata.nasa.gov/earth-observation-data/near-real-time/firms