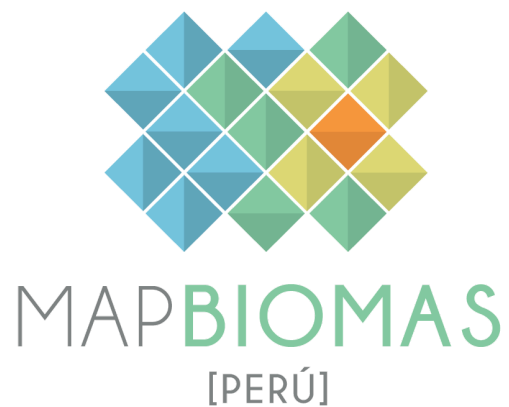




Apéndice - Minería

Tema transversal

Colección 4.0

Equipo responsable

Sergio Campos
Lucy Flores
Katherin Ivala
Zuley Cáceres
Andrea Bravo
Valeria Llactayo
Emanuel Valero

ÍNDICE

1. Introducción	3
2. Marco teórico	3
3. Metodología	3
3.1. Área de estudio	3
3.2. Generación de mosaicos anuales	6
3.3. Variables de clasificación o feature space	8
3.4. Clasificación	8
3.4.1. Búsqueda y actualización de datos	9
3.4.2. Preparación de los datos	9
3.4.3. Muestras de entrenamiento	10
3.4.4. Algoritmo de clasificación	10
3.5. Posclasificación	11
3.5.1. Llenado de vacíos de información - Filtro Gap fill	11
3.5.2. Filtro espacial	12
3.5.3. Filtro temporal	12
3.6. Integración	14
3.7. Resumen metodológico	15
4. Referencias	15

Lista de figuras

Figura 1: Mapa de acumulado de la clasificación de la clase Minería	4
Figura 2: Regiones de clasificación	5
Figura 3: Método de generación de mosaicos anuales Colección 4.0 Mapbiomas Perú	7
Figura 4: Tabla de coordenadas y mapa de puntos con la localización de las zonas mineras	9
Figura 5: Polígono de zonas mineras	9
Figura 6: Polígono de Mina y No-Mina	10
Figura 7: Mina de bauxita y clasificación de zona minera	11
Figura 8: Esquema de funcionalidad de filtro Gap Fill	12
Figura 9: Esquema de la funcionalidad del filtro espacial	12
Figura 10: Esquema de la funcionalidad del filtro temporal	13
Figura 11: Clasificación de zona minera integrada	14
Figura 12: Imágenes de referencia (resultado de la integración del tema transversal de Minería en el mapa general, imagen de alta resolución de Google Earth y paisaje)	15
Figura 13: Diagrama de flujo metodológico	15

Lista de Tablas

Tabla 1: Regiones de clasificación	5
Tabla 2: Colecciones de datos Landsat utilizadas en la Colección 4.0	7
Tabla 3: Variable de clasificación o feature space	8

1. Introducción

Este documento describe la metodología utilizada en la Colección 4 de MapBiomias Perú para la clasificación de las zonas mineras; definidas como todas las áreas de extracción de materiales pétreos o minerales con clara exposición del suelo, sin hacer diferenciación del tipo de actividad (industrial, artesanal, legal o ilegal).

Es importante señalar que la metodología original del trabajo se basó en la experiencia del equipo de MapBiomias Brasil, y en especial de Cesar Diniz (Solutions in Geoinformation, Solved, Brasil), quien ofreció una presentación al equipo MapBiomias para la identificación de la minería industrial de ese país. Esta fue implementada en la herramienta Google Earth Engine con ajustes menores para mejorar el rendimiento y la calidad.

2. Marco teórico

En el contexto de MapBiomias Perú, se considera como transversal de Minería a las áreas destinadas a la extracción de minerales y rocas, incluyendo canteras y otras superficies asociadas directamente con actividades extractivas, caracterizadas principalmente por la remoción de la cobertura natural y la exposición del suelo o del material geológico superficial. La clasificación comprende tanto actividades de minería metálica como no metálica, sin establecer diferenciaciones en función de su escala, modalidad de operación o condición legal, por lo que se incluyen actividades industriales, artesanales, formales e ilegales.

Desde el punto de vista de la teledetección, las áreas mineras presentan características espectrales y espaciales que permiten diferenciarlas de otras coberturas de la superficie terrestre. Entre estas características destacan la pérdida o ausencia de vegetación, la exposición de suelo y roca, la modificación de la superficie del terreno y la presencia de patrones geométricos o espaciales asociados a las actividades de extracción. Sin embargo, estas características pueden variar en función del tipo de mineral extraído, las condiciones ambientales, la escala de la operación y el entorno en el que se desarrolla la actividad minera.

La identificación mediante imágenes satelitales permite realizar un seguimiento espacial y temporal de estas áreas, facilitando la detección de superficies que presentan cambios asociados a la actividad extractiva. En este sentido, la clasificación busca representar la ocupación de la superficie por actividades mineras, sin pretender determinar aspectos relacionados con la legalidad de la operación, el método de extracción o su nivel de formalización.

3. Metodología

3.1. Área de estudio

El ámbito de estudio comprende todo el territorio peruano y considera los cuatro biomas definidos para la clasificación de MapBiomias Perú: Desierto Costero, Bosque Seco Ecuatorial, Andes y Amazonía.

La clasificación de la clase Minería se realizó para toda la serie temporal de la colección actual, permitiendo identificar su distribución espacial y evolución a escala nacional. Como referencia espacial del área de estudio, la Figura 1 presenta el mapa de acumulado de la clasificación de Minería para el conjunto de años analizados, donde se representa la totalidad de píxeles

identificados como pertenecientes a esta clase al menos en uno de los años de la serie temporal. Esta representación permite visualizar la extensión y distribución de las áreas de minas identificadas a nivel nacional y proporciona un contexto espacial para el análisis desarrollado posteriormente.



Figura 1. Mapa de acumulado de la clasificación de la clase Minería (ID 30) correspondiente a toda la serie temporal (1985–2025).

Para la clasificación de Minería se definieron treinta y seis (36) regiones de mapeo o clasificación, usando las mismas que en el ATBD General MapBiomás Perú. Cada región fue trabajada independientemente para cada año de la serie temporal.

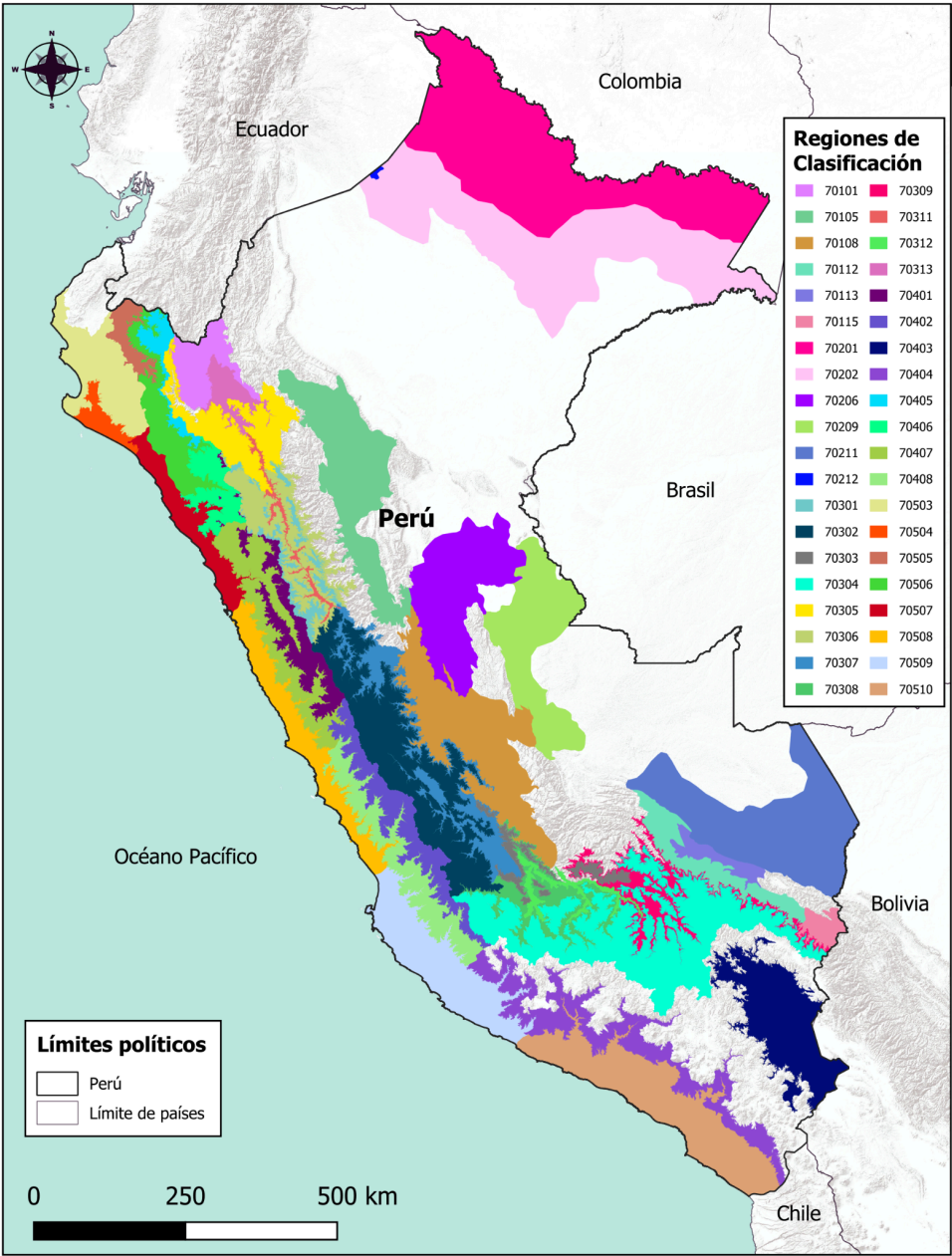


Figura 2. Regiones de clasificación.

Tabla 1. Regiones de clasificación.

Región de mosaico	Nombre de región de mosaico	Identificador	Códigos
Región 701	Amazonía Alta	5	70105
		8	70108
		9	70109
		12	70112
		13	70113
		15	70115

Región 702	Amazonía Baja	2	70202
		6	70206
		9	70209
		11	70211
Región 703	Andes Alto	1	70301
		2	70302
		3	70303
		4	70304
	Andes Bajo	1	70305
		2	70306
		3	70307
		4	70308
		5	70309
	Andes Bosque Seco Interandino	2	70311
Región 704	Andes Pacífico Alto	1	70401
		2	70402
		3	70403
	Andes Pacífico Bajo	1	70404
		2	70405
		3	70406
		4	70407
		5	70408
Región 705	Costa Norte	3	70503
		4	70504
		5	70505
		6	70506
	Costa Centro	1	70507
		2	70508
	Costa Sur	1	70509
		2	70510

Fuente: Elaboración propia

3.2. Generación de mosaicos anuales

Para la clasificación transversal de la clase Minería se utilizaron los mosaicos anuales generados previamente para la clasificación general de Cobertura y Uso del Suelo de la Colección 4.0 de MapBiomás Perú. Estos mosaicos fueron producidos a partir de imágenes satelitales de reflectancia de superficie con corrección atmosférica de los sensores Landsat 4, 5, 7, 8 y 9, y comprenden una serie temporal de 41 años, desde 1985 hasta 2025. La Tabla 2 presenta las colecciones de imágenes Landsat utilizadas y disponibles en Google Earth Engine.

Tabla 2. Colecciones de datos Landsat utilizadas en la Colección 4.0.

Sensor	ID Colección en GEE
Landsat 4 TM	LANDSAT/LT04/C02/T1_SR
Landsat 5 TM	LANDSAT/LT05/C02/T1_SR
Landsat 7 ETM+	LANDSAT/LE07/C02/T1_SR
Landsat 8 OLI /TIRS	LANDSAT/LC08/C02/T1_SR
Landsat 9	LANDSAT/LC09/C02/T1_L2

Fuente: Elaboración propia.

La Figura 3 resume la metodología general empleada para la generación de los mosaicos anuales de la Colección 4.0 de MapBiomás Perú. El proceso considera la disponibilidad anual de imágenes Landsat, la grilla de cartas millonésimas utilizada para la organización espacial de los mosaicos y la definición de parámetros asociados a la temporalidad de las imágenes, los sensores disponibles, la cobertura de nubes y la exclusión de escenas que pudieran afectar la calidad de las composiciones anuales.

Adicionalmente, la metodología general incorpora información topográfica derivada del modelo digital de elevación SRTM (Shuttle Radar Topography Mission) disponible en Google Earth Engine.

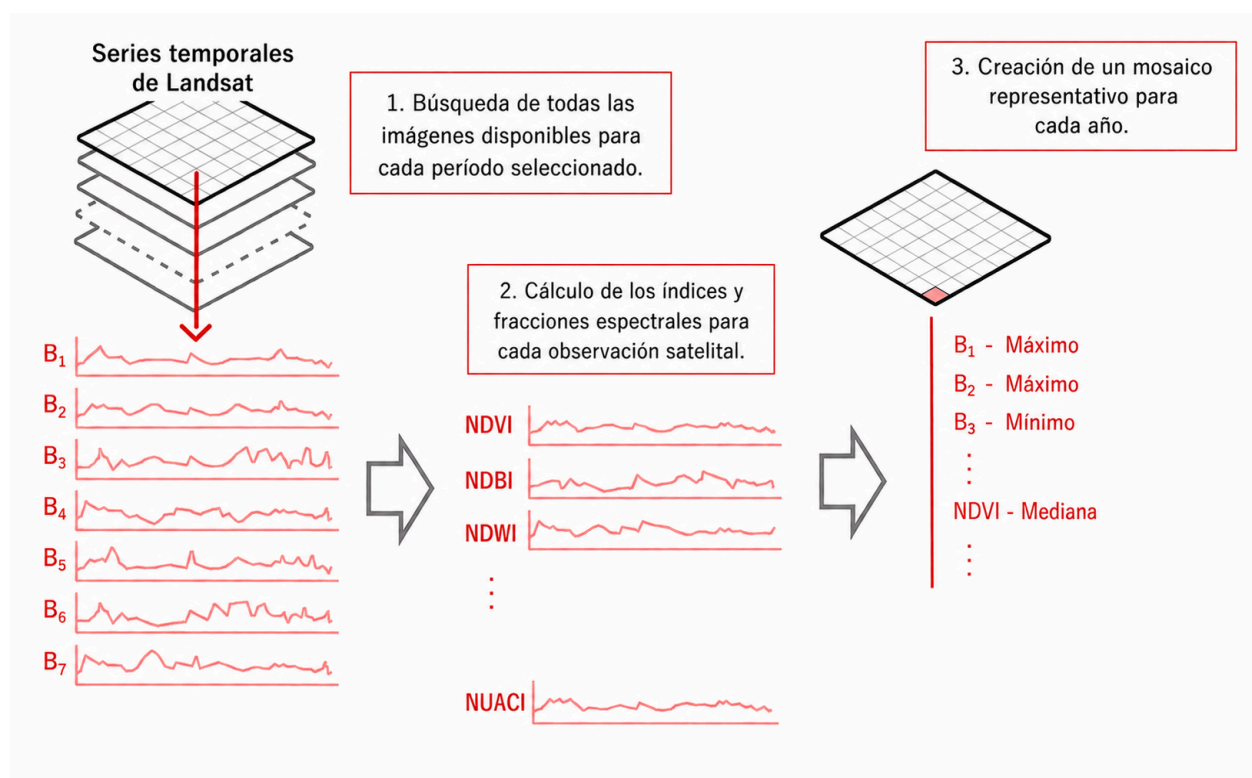


Figura 3. Método de generación de mosaicos anuales Colección 4.0 Mapbiomas Perú.

Como parte del preprocesamiento general de las imágenes Landsat, se aplicaron procedimientos para la identificación y enmascaramiento de píxeles afectados por nubes, sombras y otras condiciones que pudieran comprometer la calidad de las observaciones. Una vez eliminados estos píxeles, se generaron las composiciones anuales y los reductores estadísticos utilizados para construir las variables de clasificación. La descripción detallada de los parámetros, algoritmos de enmascaramiento y procedimientos empleados para la generación de los mosaicos se encuentra en el ATBD General de MapBiomás Perú Colección 4.0.

3.3. Variables de clasificación o feature space

La selección de variables (de los mosaicos) para la identificación de las zonas mineras es el paso inicial para la clasificación. Se realizaron pruebas que permitieron identificar un conjunto de variables específicas para la detección de Minería. En total se usaron 18 variables¹, las cuales se presentan en la Tabla 3 a continuación.

Tabla 3. Variables de clasificación o feature space

Banda	Expresión / Descripción	Reductor
blue	Valor de la banda azul	Mediana
green	Valor de la banda verde	Mediana
red	Valor de la banda roja	Mediana
nir	Valor de la banda infrarrojo cercano (NIR)	Mediana
swir1	Valor de la banda SWIR1	Mediana
swir2	Valor de la banda SWIR2	Mediana
ndfi		Mediana
ndfib		Mediana
npv		Mediana
soil		Mediana, Amplitud
evi2	$2.5 * ((\text{NIR} - \text{Red}) / (\text{NIR} + 2.4 * \text{Red} + 1))$	Mediana
ndvi	$(\text{NIR} - \text{Red}) / (\text{NIR} + \text{Red})$	Mediana
ndwi_gao	$(\text{Green} - \text{SWIR1}) / (\text{Green} + \text{SWIR1})$	Mediana
ndwi_mcfeters	$(\text{Green} - \text{NIR}) / (\text{Green} + \text{NIR})$	Mediana
mndwi	$(\text{Green} - \text{NIR}) / (\text{Green} + \text{NIR})$	Mediana
gcvi	$(\text{nir} / \text{green}) - 1$	Mediana
nuaci	$UNTL * (1 - \sqrt{(\text{NDWI} - a\text{NDWI})^2 + (\text{NDVI} - a\text{NDVI})^2} + (\text{NDBI} - a\text{NDBI})^2)$	Mediana
pri	$(\text{blue} - \text{green}) / (\text{blue} + \text{green})$	Mediana

Fuente: Elaboración propia

3.4. Clasificación

Para generar la clasificación se desarrolló un algoritmo de procesamiento organizado en seis etapas principales: cuatro asociadas a la preparación y clasificación, una etapa de posclasificación y una etapa final de integración. Se generaron los resultados y se almacenaron en Google Earth Engine. El proceso involucró la definición de las áreas mineras, toma de muestras y clasificación, para luego integrar todas las minas identificadas en una capa única y, finalmente, con el mapa general (colección 4) de MapBiomás Perú. A continuación se describe cada paso asociado a la clasificación.

¹ Las variables derivadas y fracciones espectrales corresponden a las variables disponibles en los mosaicos generales de la Colección 4.0 de MapBiomás Perú. Su descripción detallada y procedimiento de cálculo se encuentran en el ATBD General.

3.4.1. Búsqueda y actualización de datos

Se realizó una búsqueda de fuentes de datos auxiliares (mapas, reportes, informes, geovisores), con el objetivo de localizar las principales áreas mineras. Con ello se generó una tabla con las coordenadas de cada zona minera, que luego se exportó a una herramienta de sistemas de información geográfica (SIG) para su visualización espacial como un mapa de puntos (ver Figura 4).

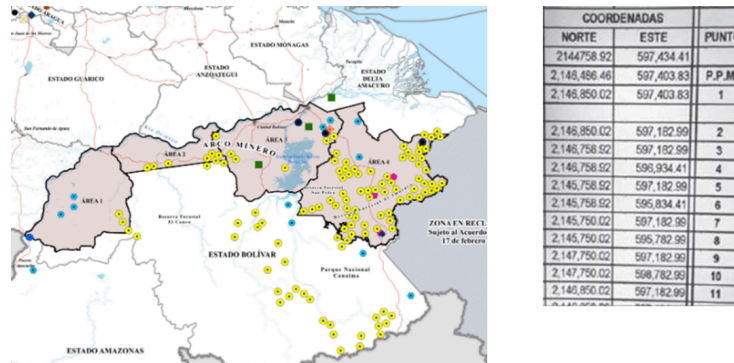


Figura 4. Tabla de coordenadas y mapa de puntos con la localización de las zonas mineras.

3.4.2. Preparación de los datos

El mapa de puntos de las zonas mineras generado se importó a Google Earth Engine, como referencia espacial para delimitar las áreas mineras y la toma de muestras de entrenamiento. Como mapa base se utilizaron las imágenes de alta resolución espacial de Google Maps y los mosaicos de Landsat (ver sección 3.2). Usando estos datos se identificaron y generaron polígonos envolventes alrededor de las zonas mineras, pudiendo abarcar una o varias minas (ver Figura 5). Estas envolventes se delimitaron manualmente utilizando las herramientas de dibujo de Google Earth Engine, cuidando que el área de las zonas mineras y no mineras fuese proporcional, para facilitar la correcta identificación de coberturas por parte del algoritmo de clasificación. Para cumplir ese requerimiento, en ciertas regiones se dibujó una gran cantidad de polígonos.

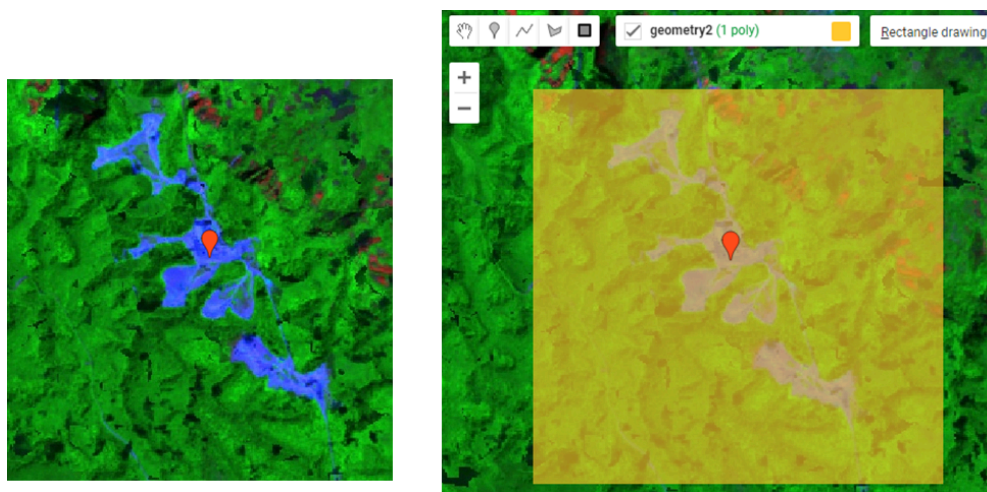


Figura 5. Polígono de zonas mineras

3.4.3. Muestras de entrenamiento

Dentro de las envolventes (polígonos) que delimitan las áreas mineras de interés (sección 3.4.2, se definieron manualmente dos tipos de polígonos de entrenamiento: uno definido como 'Mina' (ID 30), que contiene estrictamente el área de explotación minera o la mina y un segundo conjunto denominado 'No_Mina' (ID 27) (ver Figura 6), utilizado como clase complementaria durante la clasificación binaria y conformado por áreas que no corresponden a Minería (ID 30).

Las muestras de entrenamiento tomadas dentro del área de interés (envolventes cuadradas/geométricas) son las que se utilizan para generar la clasificación de Minería. Las áreas ubicadas fuera de las envolventes delimitadas no fueron incluidas en el proceso de clasificación de esta clase y corresponden a otras coberturas dentro de la leyenda de MapBiomás Perú.

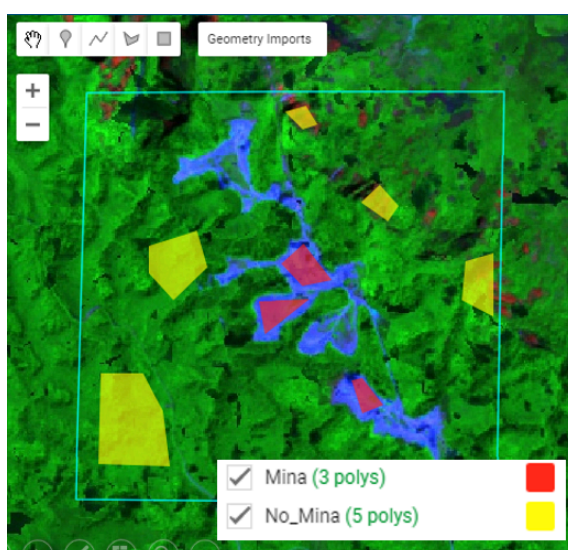


Figura 6. Polígono de Mina y No-Mina.

3.4.4. Algoritmo de clasificación

La clasificación de las zonas mineras se llevó a cabo en la plataforma de Google Earth Engine usando el algoritmo Random Forests (bosques aleatorios, de su traducción del inglés) (Breiman, 2001; Hartshorn, 2016). con base en la metodología implementada por el equipo de MapBiomás Brasil (Solutions in Geoinformation, Solved, Brasil) para la identificación de la minería industrial de ese país.

La clasificación se configuró con 50 árboles para todas las regiones de clasificación y se ejecutó dentro de las zonas delimitadas en los polígonos dibujados según lo explicado en la sección 3.4.2 de este documento. El resultado fue una imagen formada por píxeles con valor binario indicando la presencia o ausencia de zonas mineras (área definida en color rojo en la Figura 7), dentro de cada cuadrícula definida previamente.



Figura 7. Mina de bauxita y clasificación de zona minera.

3.5. Posclasificación

Los resultados de la clasificación se sometieron a la aplicación de una secuencia de filtros con el propósito de reducir las inconsistencias temporales, vacíos de información y otros problemas de clasificación.

3.5.1. Llenado de vacíos de información - Filtro Gap fill

El filtro de relleno de vacíos temporales (*Gap Fill*) se aplicó con el objetivo de completar los píxeles sin información presentes en la serie anual de clasificación. Previamente, se aseguró la existencia de una banda correspondiente a cada año del periodo 1985–2025; cuando alguna banda anual no se encontraba disponible, esta se incorporó como una banda enmascarada. Asimismo, para los años definidos en la parametrización del proceso, se identificaron como vacíos aquellos píxeles donde los mosaicos Landsat no presentaban información válida.

El procedimiento se ejecutó mediante dos barridos consecutivos de la serie temporal. En el primer barrido, realizado desde el inicio hacia el final de la serie (1985–2025), los píxeles sin información fueron completados utilizando el último valor válido disponible en los años precedentes. De esta manera, cuando se presentaron vacíos consecutivos, el último valor válido observado se propagó hacia los años siguientes. Posteriormente, se realizó un segundo barrido en sentido inverso, desde el final hacia el inicio de la serie (2025–1985), mediante el cual los vacíos que permanecieron sin completar fueron rellenados utilizando el valor válido posterior más próximo.

De manera complementaria, se generó una imagen que registra el año de procedencia del valor utilizado en el relleno de cada píxel, permitiendo identificar el origen temporal de la información incorporada. De este modo, el filtro permite reducir los vacíos de información de la serie antes de la aplicación de los filtros posteriores de posclasificación (ver Figura 8).

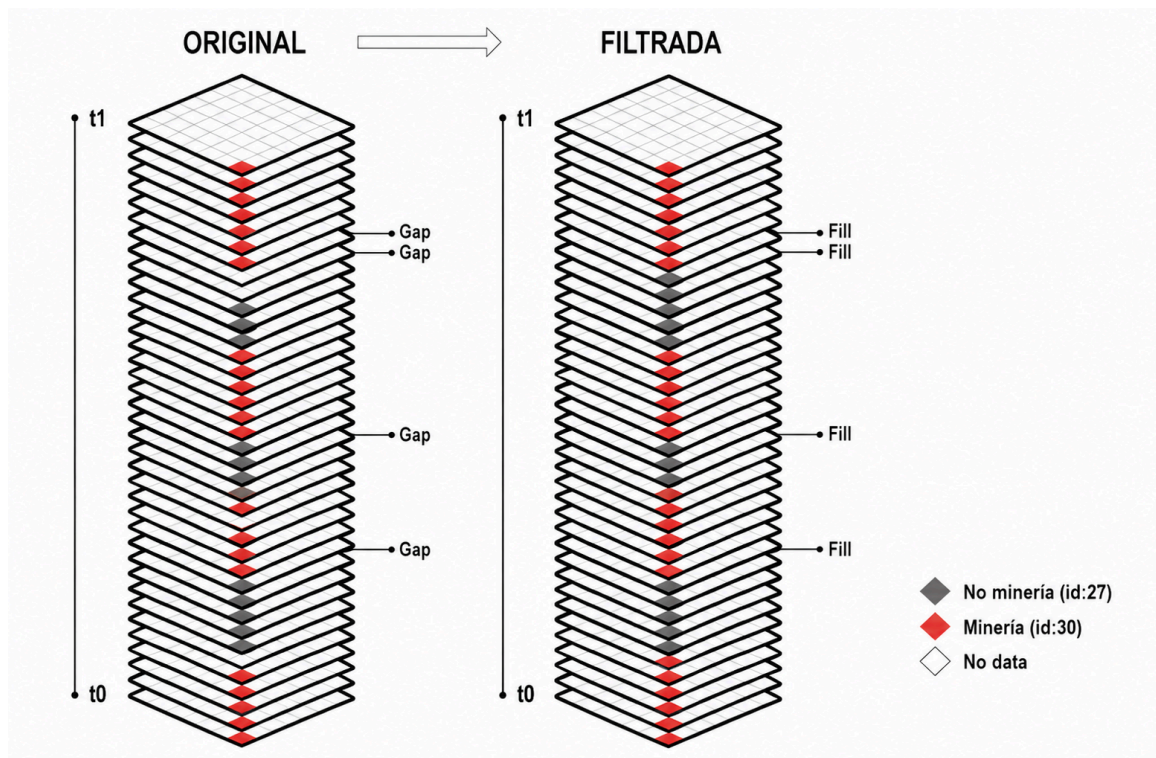


Figura 8. Esquema de funcionalidad de filtro Gap Fill, en donde t0 representa el primer año y t1 el último año.

3.5.2. Filtro espacial

El filtro espacial está basado en la función "connectedPixelCount" de Google Earth Engine, que identifica al conjunto de píxeles vecinos que comparten un mismo valor. Utilizando esta función y considerando una cantidad mínima preestablecida de píxeles idénticos, aquellos que no presentan conexión (valores diferentes) se consideran píxeles aislados y fueron reclasificados. Para ambas clases (ID: 27, ID: 30), la conexión mínima establecida fue de 5 píxeles (aprox. 0.5 ha) (ver Figura 9).

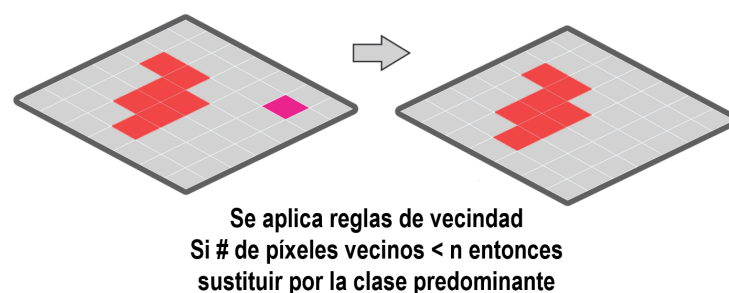


Figura 9. Esquema de la funcionalidad del filtro espacial.

3.5.3. Filtro temporal

Se ejecutó un filtro temporal utilizando ventanas de 3, 4 y 5 años en toda la serie de tiempo, tomando en cuenta el contexto de cada región de clasificación. El filtro inspecciona el valor de cada píxel clasificado en relación al valor de ese píxel en clasificaciones temporalmente

consecutivas. Para ello emplea una ventana móvil unidireccional que toma en consideración secuencias de clasificaciones de 3 a 5 años e identifica transiciones temporales no permitidas. El filtro temporal se aplica a cada píxel de todos los años de la colección. Dependiendo del año que la regla modificará, existen tres tipos de reglas:

Reglas generales (RG): Aplicadas a píxeles de años en posiciones intermedias en secuencias de 3 a 5 años. Esta regla se aplica únicamente en casos donde hay una inconsistencia temporal; por ejemplo, cuando secuencias de años consecutivos tienen valores idénticos a excepción del píxel en posición central. En estos casos, el filtro modificará el valor del píxel central para que guarde consistencia con los píxeles que le anteceden y suceden. En el caso de secuencias de 3 años, solo existe una opción de posición central o año intermedio. En el caso de secuencias de 4 o 5 años, existen dos o tres alternativas de posiciones centrales. Esta regla modifica los valores de las clasificaciones de los años 1986 al 2024.

Reglas de primer año (RP): Aplicadas únicamente al primer año de la serie temporal. Esta regla modifica los valores de la clasificación del año 1985.

Reglas de último año (RU): Aplicada al último año de la clasificación. Esta regla modifica los valores de la clasificación del año 2025.

De este modo, los filtros temporales reducen vacíos de información e inconsistencias temporales o cambios que no son posibles o no están permitidos (ver Figura 10).

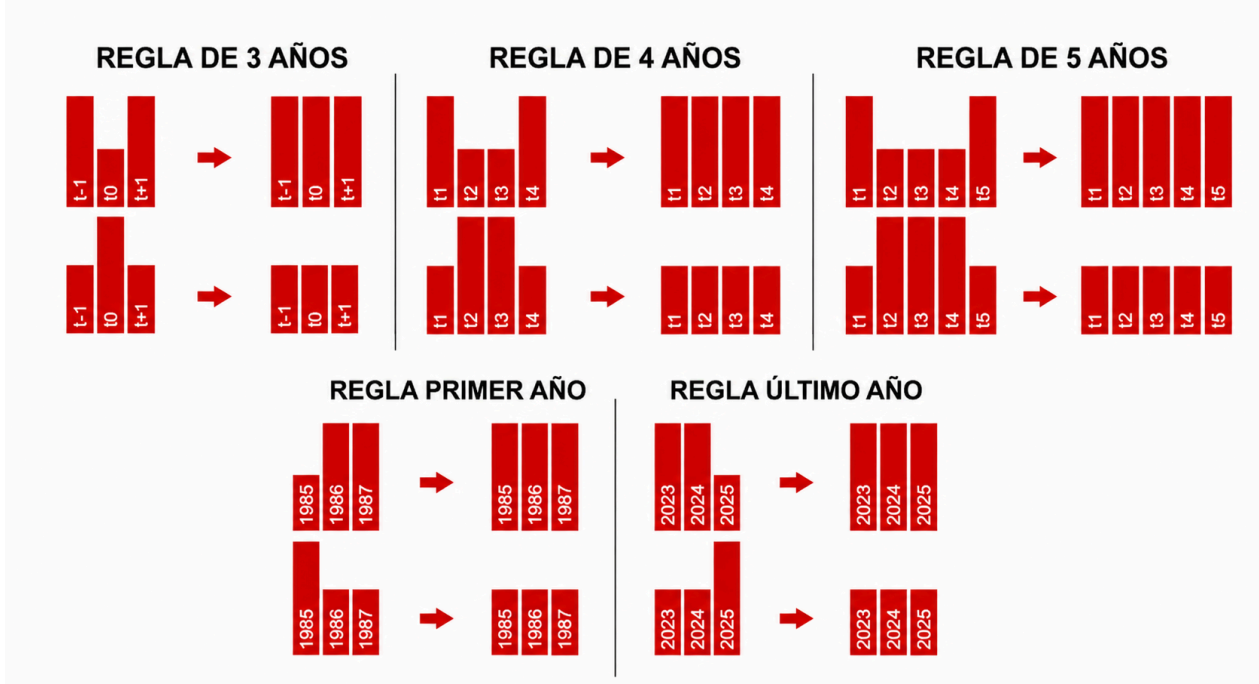


Figura 10. Esquema de la funcionalidad del filtro temporal.

De manera complementaria, se emplearon herramientas y procedimientos adicionales de posclasificación, entre ellos procesos de unión (join) con la colección anterior (Colección 3.0) en el caso de Minería para todas las regiones de clasificación, con el propósito de aprovechar y conservar información previamente validada y con buen desempeño en la clasificación. Asimismo, se aplicaron ajustes específicos de carácter temporal y espacial en función de las características y necesidades identificadas en cada región o periodo analizado, a partir de procesos de revisión y ajuste sustentados en un control de calidad visual. Estas herramientas

permitieron resolver inconsistencias puntuales, preservar patrones de clasificación consolidados y mejorar la continuidad y coherencia de las trayectorias de cambio a lo largo de la serie temporal. En este sentido, la secuencia de posclasificación no se estableció como un procedimiento rígido y uniforme para todo el territorio, sino como un conjunto de herramientas aplicadas de manera contextual, cuya selección y configuración dependen de las condiciones espaciales y temporales de cada área, contribuyendo a optimizar la consistencia final de la clasificación.

3.6. Integración

Al obtener la clasificación de las zonas mineras se realizó la integración de todas las zonas identificadas con minas para todo el país (ver Figura 11), y este resultado se integró al mapa general de clasificación nacional (ver Figura 12). Esta integración se generó con base a una serie de reglas de prevalencia u orden de integración por jerarquía, para cada una de las clases dentro de la leyenda MapBiomas.

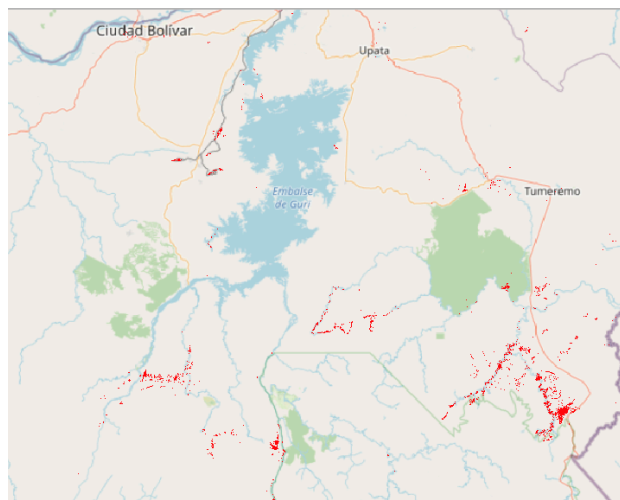


Figura 11. Clasificación de zona minera integrada.

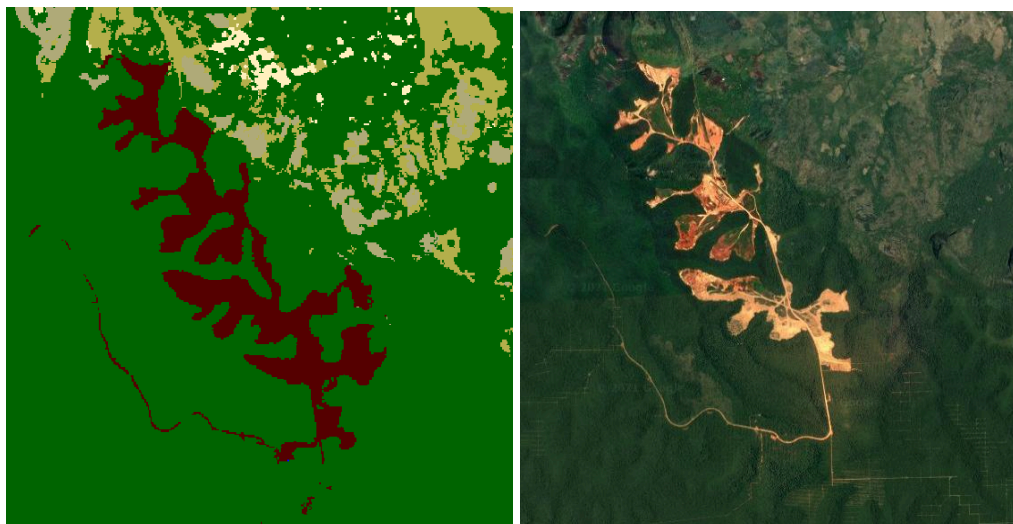




Figura 12. Imágenes de referencia (resultado de la integración del tema transversal de Minería en el mapa general, imagen de alta resolución de Google Earth y paisaje).

3.7. Resumen metodológico

Los 6 pasos de clasificación del tema transversal de Minería se representan esquemáticamente en el siguiente gráfico. (ver Figura 13).

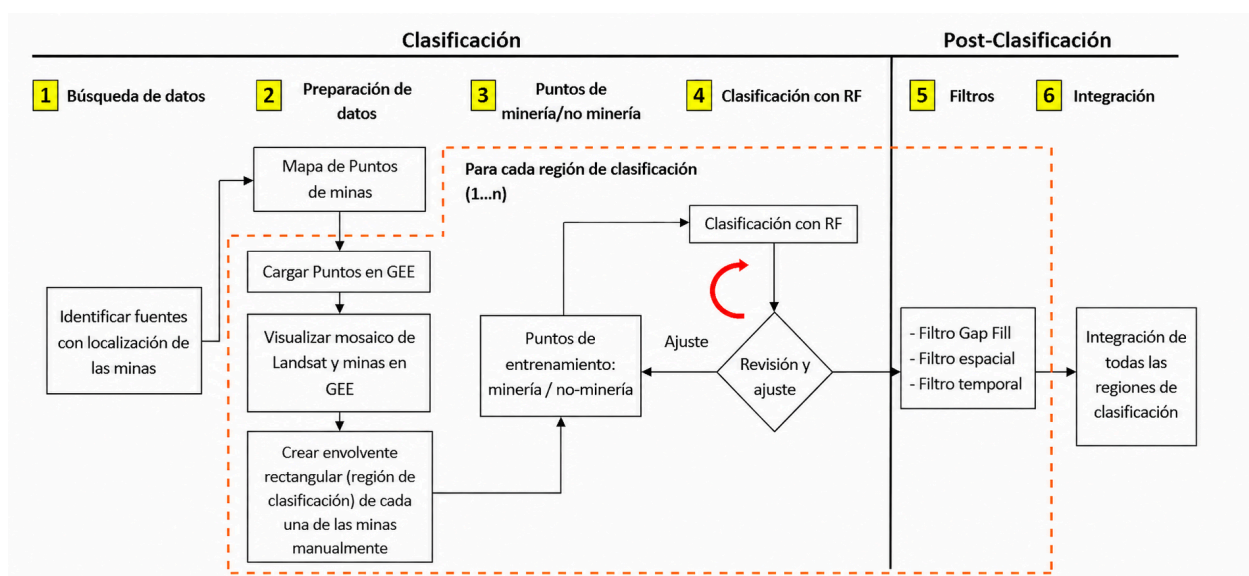


Figura 13. Diagrama de flujo metodológico.

4. Referencias

Breiman, Leo. "Random Forests." Machine Learning vol 45 (2001): 5–32.

Hartshorn, Scott. Machine Learning With Random Forests And Decision Trees A Visual Guide For Beginners. Amazon.com Services LLC, 2016.

MapBiomas. (s. f.). *MapBiomas Brasil*. <https://brasil.mapbiomas.org/>