

RAISG

RED AMAZÓNICA DE INFORMACIÓN
SOCIOAMBIENTAL GEORREFERENCIADA



MAPBIOMAS
[AMAZONIA]

Apéndice - Minería

Tema transversal en Colección 5

Versión 1

Equipo responsable
Emanuel Valero
Rodrigo Lazo
José Sánchez

Noviembre 2023

Índice

1. Introducción	3
2. Construcción y definición de mosaicos Landsat	3
3. Variables de clasificación	5
4. Clasificación	5
4.1. Búsqueda y actualización de datos	5
4.2. Preparación de los datos	6
4.3. Muestras de entrenamiento (minería/no minería)	6
4.4. Clasificación con algoritmo Random Forests	7
5. Post-clasificación	8
5.1 Filtro espacial	8
5.2 Filtro de frecuencia	8
5.3 Integración	9
6. Resumen metodológico	11
7. Referencias	11

Lista de figuras

<i>Figura 1: Área de estudio</i>	4
<i>Figura 2: Tabla de coordenadas y mapa de puntos con la localización de las zonas mineras</i>	6
<i>Figura 3: Polígono de zonas mineras</i>	6
<i>Figura 4: Polígono de Mina y No-Mina</i>	7
<i>Figura 5: Der. Mina de bauxita. Izq. Clasificación de zona minera</i>	8
<i>Figura 6: Esquema de la funcionalidad del filtro espacial</i>	8
<i>Figura 7: Efecto del filtro de frecuencia</i>	9
<i>Figura 8: Clasificación de zona minera integrada</i>	9
<i>Figura 9: 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)</i>	10
<i>Figura 10: Pasos para la clasificación de minería</i>	11

Lista de Tablas

<i>Tabla 1: Variables para la clasificación de zonas mineras de la Colección 5</i>	5
--	---

1. Introducción

La cobertura minera, se clasifica como un tema transversal dentro de la metodología de MapBiomias, debido a que genera confusión en las clasificaciones que incluyen varias clases de cobertura, al no poder diferenciarse de manera adecuada de otras como herbazales, pastos o agricultura. En la colección 5 se incluyen los países: Bolivia, Colombia, Ecuador, Guyana, Guyana Francesa, Perú, Surinam y Venezuela.

Este documento describe la metodología utilizada en la colección 5 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. Construcción y definición de mosaicos Landsat

Para fines de procesamiento, la metodología de MapBiomias, divide el área del proyecto en una grilla de cartas regulares definidas con base a la grilla de las Cartas Internacionales del Mundo a la Millonésima, a escala de 1:250.000. Cada carta rectangular cubre un área de 1°30' de longitud por 1° de latitud. En la figura 1 se muestra el área de estudio de RAISG conformada por 8 países (Guayana Francesa, Surinam, Guyana, Venezuela, Colombia, Ecuador, Perú y Bolivia).

En cada una de estas cartas o celdas de la grilla, se construyen los mosaicos anuales de imágenes Landsat. Esto con el objetivo de generar imágenes periódicas sin nubes que permitan su clasificación e identificación de cambios. Para mayor detalle revisar la sección 3.1 del documento: [ATBD General MapBiomias Amazonia 5.0](#). En esta colección para la construcción de los mosaicos se utilizó la Landsat Collection 2, la cual presenta la versión más actualizada del reprocesamiento de los archivos de Landsat, con varias mejoras en la calidad de las imágenes.

La construcción de los mosaicos se inicia con la selección de las imágenes Landsat. Para ello se construye una tabla de parametrización con los siguientes parámetros, almacenados como metadatos en cada asset construido:

- ID: Identificador o código único de la unidad carta/región
- Año: Año de la serie (comprendiendo el período 1985 a 2022) al que corresponde el mosaico.
- Carta: Código identificador de la carta
- Fecha inicial/ Fecha final: Período del año (fecha de inicio y final) para la selección de imágenes del catálogo de datos de imágenes Landsat de Google Earth Engine.

- Sensor: El satélite y su respectivo sensor: Landsat 5 TM, Landsat 7 ETM+, Landsat 8 OLI, Landsat 9 o una combinación de Landsat 5 y Landsat 7.
- Porcentaje de nubes: Porcentaje máximo de cobertura de nubes aceptado de cada imagen Landsat que será usado para construir el mosaico de imágenes. Este dato proviene de la metadata de las imágenes Landsat.
- Blacklist: Imágenes que por su calidad son excluidas de la construcción del mosaico.

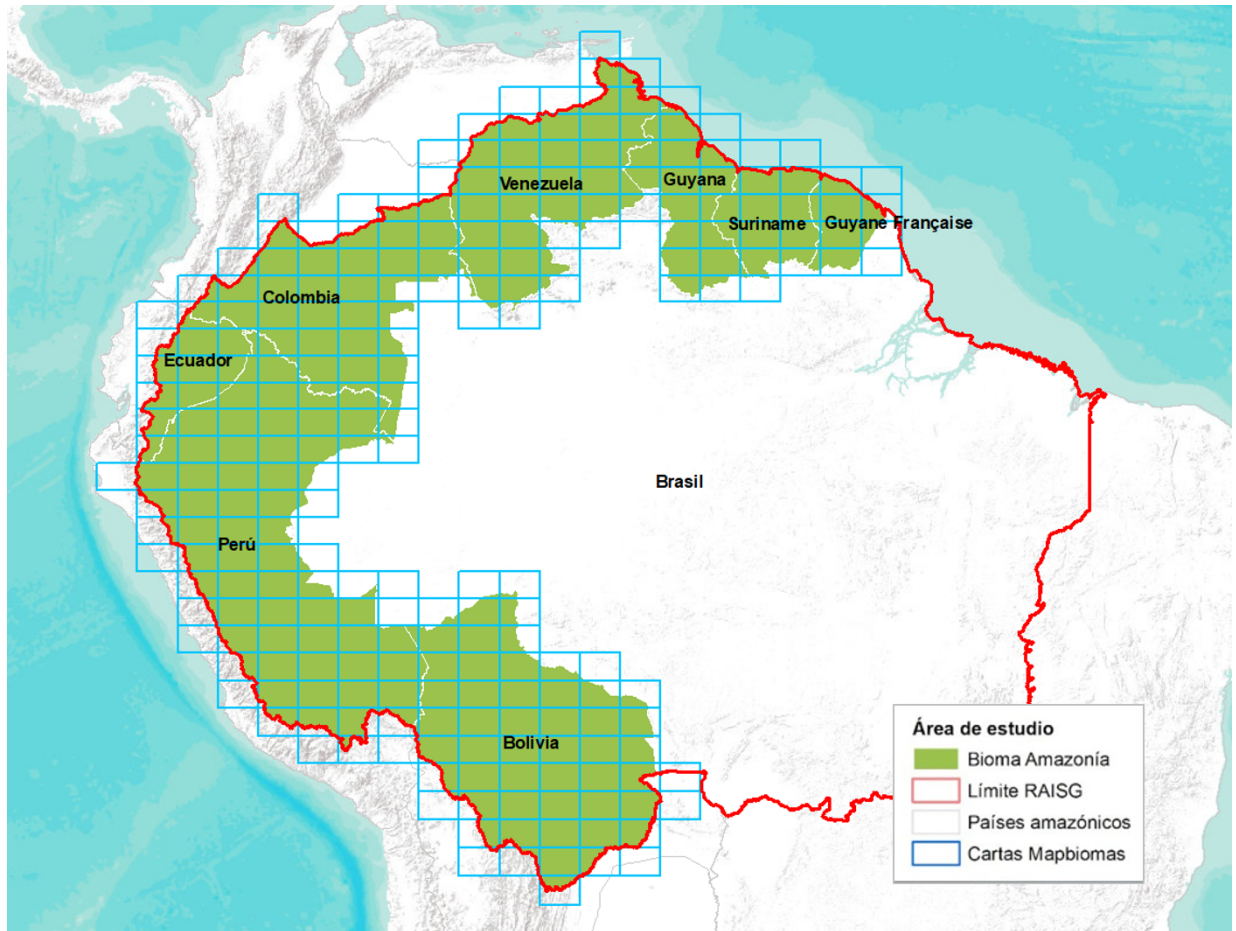


Figura 1: Área de estudio

Con esta información se construyen los mosaicos anuales de imágenes, con la herramienta Google Earth Engine. Todos los mosaicos generados se almacenan como una colección de imágenes de Google Earth Engine y están disponibles en la ruta: [projects/mapbiomas-raisg/MOSAICOS/mosaics-2](https://earthengine.google.com/projects/mapbiomas-raisg/MOSAICOS/mosaics-2). Los mosaicos presentan inicialmente los valores de mediana de cada una de las bandas de los satélites Landsat. Con base en estos datos, se agregaron a cada mosaico bandas adicionales de índices espectrales, fracciones espectrales, variables de textura e índices topográfico-geomorfológicos derivados del modelo de elevación digital. Se aplica un cálculo estadístico a cada una de las bandas de los mosaicos, para generar información específica sobre: mediana, mediana estación seca (corresponde al percentil 25), mediana estación lluviosa (correspondiente al percentil 75), amplitud (rango de valores), la desviación estándar y el valor mínimo. Todos estos datos componen nuestras variables, que se utilizan como fuente para el clasificador de bosque aleatorio (Random Forest).

3. Variables de clasificación

La selección de variables (de los mosaicos) para la identificación de las zonas mineras en la Pan-Amazonía es el paso inicial para la clasificación. Se realizaron pruebas que permitieron identificar un conjunto de variables específicas para la detección de la minería. En total se usaron 18 variables, las cuales se presentan en la tabla 1.

Tabla 1 - Variables para la clasificación de zonas mineras de la Colección 5, MapBiomias Amazonia

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_mcfeeters	$(\text{Green} - \text{NIR}) / (\text{Green} + \text{NIR})$	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

4. Clasificación

Para generar la clasificación se desarrolló un algoritmo de procesamiento en seis etapas, cuatro de las cuales corresponden al proceso de clasificación y las restantes 2 a la etapa de post-clasificación. 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 5) de MapBiomias Amazonía. A continuación se describe cada paso asociado a la clasificación.

4.1. Búsqueda y actualización de datos

Se realizó una búsqueda de fuentes de datos auxiliares (mapas, reportes, informes), 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 2)

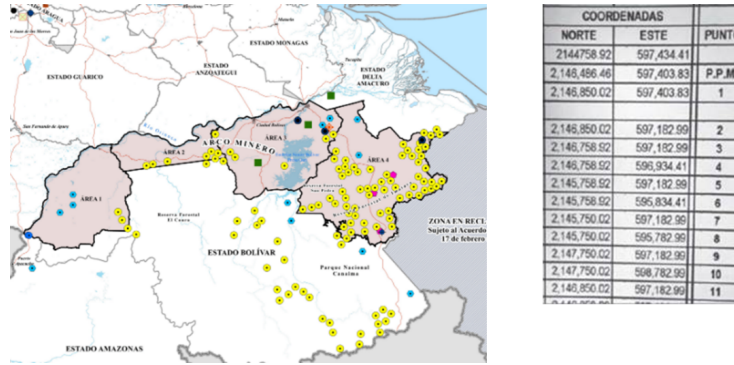


Figura 2: Tabla de coordenadas y mapa de puntos con la localización de las zonas mineras

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 2). Usando estos datos se identificaron y generaron polígonos envolventes alrededor de las zonas mineras, pudiendo abarcar una o varias minas (figura 3). 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.

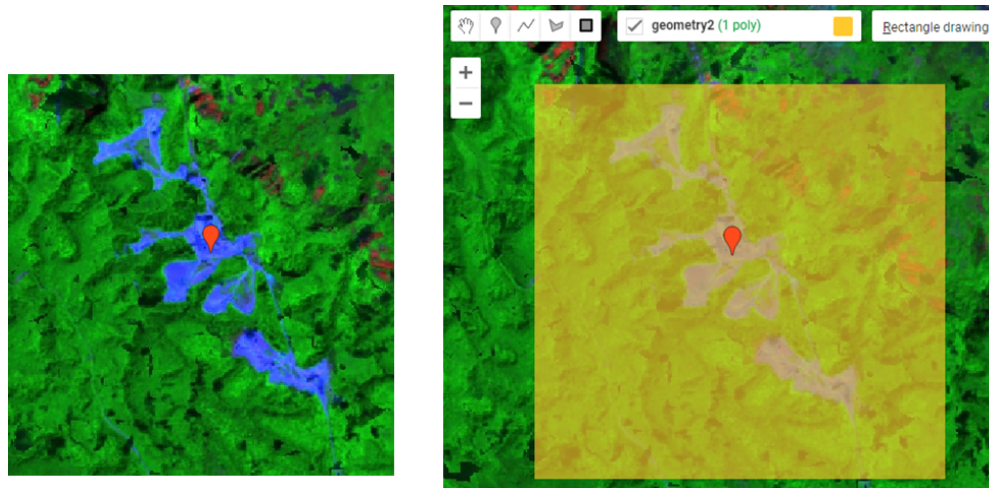


Figura 3: Polígono de zonas mineras

4.3. Muestras de entrenamiento (minería/no minería)

Dentro de las envolventes (polígonos) que delimitan las áreas mineras de interés (sección 4.2, ver figura 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 el segundo definido como 'No_Mina' (ver figura 4)

que abarca la clase 'no observado' o aquellas clases no correspondientes a minería (ID 27).

Las muestras de entrenamiento tomadas dentro del área de interés (envolventes cuadradas/geométricas) son las que se utilizan para generar la clasificación de la zona de minería. Todo lo que está fuera de la envolvente delimitada no es clasificado ya que se considera como zona no minera y/o asignada a otra clase dentro de la leyenda MapBiomás.

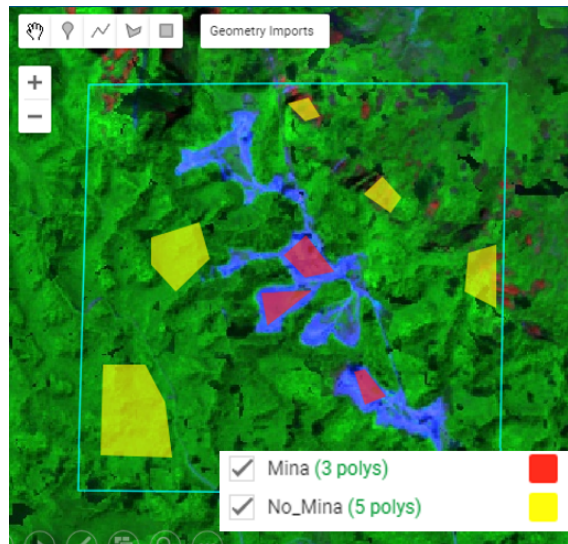


Figura 4: Polígono de Mina y No-Mina

4.4. Clasificación con algoritmo Random Forests

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), 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 ejecutó dentro de las zonas delimitadas en los polígonos dibujados según lo explicado en la sección 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 5), dentro de cada cuadrícula definida previamente.



Figura 5: Der. Mina de bauxita. Izq. Clasificación de zona minera

5. Post-clasificación

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. Para ello se utilizó un filtro espacial y uno de frecuencia.

5.1 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 (figura 6). Para ambas clases (ID: 27, ID: 30), la conexión mínima establecida fue de 5 píxeles (aprox. 0.5 ha).

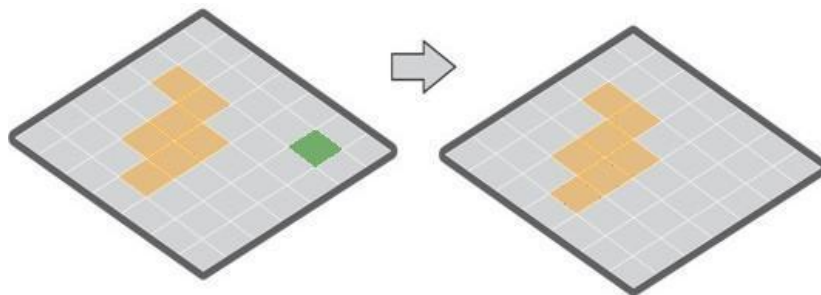


Figura 6: Esquema de la funcionalidad del filtro espacial

5.2 Filtro de frecuencia

Este filtro toma en consideración la frecuencia de ocurrencia de la clase minería (ID: 30) y no observado (ID: 27) en toda la serie temporal, para rellenar o reemplazar por el valor de la clase correspondiente el año que cumple con el valor mínimo de frecuencia.

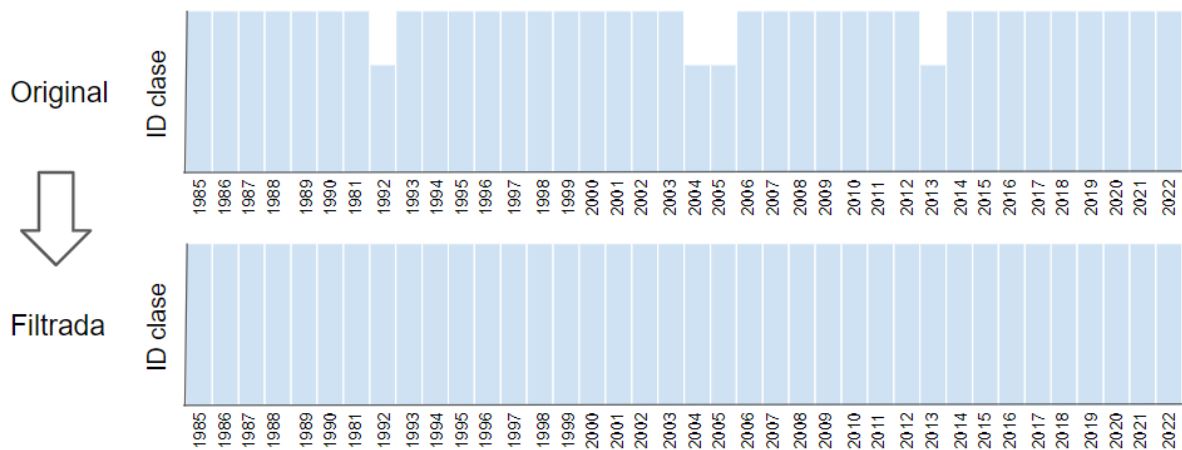


Figura 7: Efecto del filtro de frecuencia

5.3 Integración

Al obtener la clasificación de las zonas mineras se hizo una primera integración de todas las zonas identificadas con minas para cada país (figura 8), y este resultado se integró al mapa general de clasificación nacional (figura 9). 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 MapBiomás.

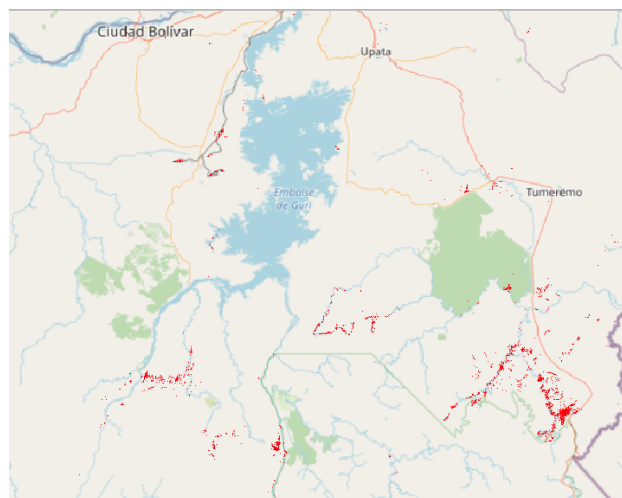


Figura 8: Clasificación de zona minera integrada

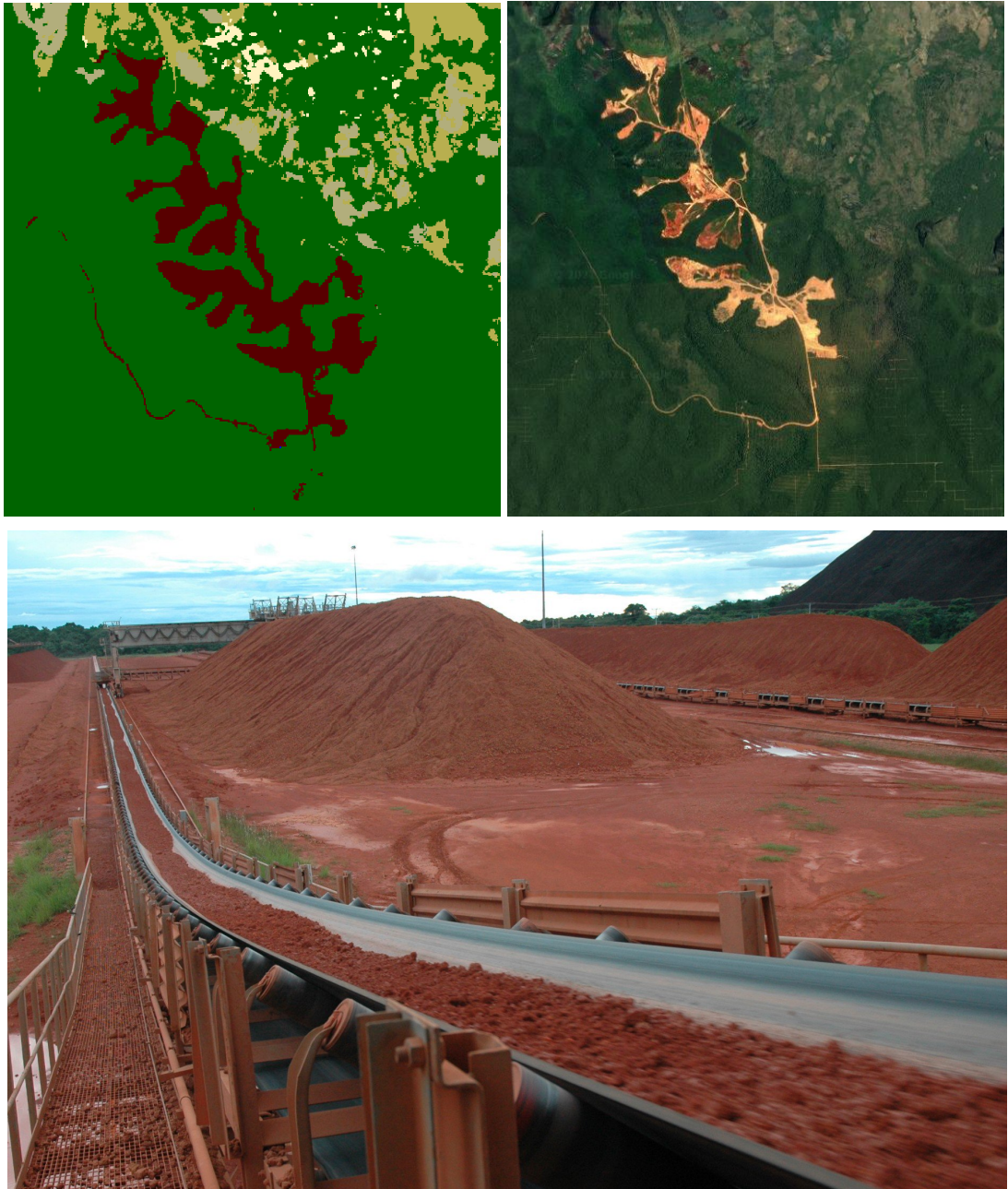


Figura 9: Imágenes de referencia (resultado de la integración del tema transversal de minería en el mapa general, imagen de alta resolución de Google Earth y paisaje)

6. Resumen metodológico

Los 6 pasos de clasificación del tema transversal de minería se representan esquemáticamente en el siguiente gráfico (figura 10).



Figura 10: Pasos para la clasificación de minería

7. Referencias

- 1) Breiman, Leo. "Random Forests." Machine Learning vol 45 (2001): 5–32.
- 2) Hartshorn, Scott. Machine Learning With Random Forests And Decision Trees A Visual Guide For Beginners. Amazon.com Services LLC, 2016.