

MAPBIOMAS
PERÚ

Apéndice Agricultura

Tema transversal en Colección 3

Versión 1

Índice

1 Introducción	3
2 Mosaicos de imágenes	4
2.1 Parametrización de mosaicos anuales	4
2.2 Variables de clasificación o feature space	4
2.3 Variables específicas para el transversal agricultura	
3. Clasificación	4
3.1. Definición de la clase agricultura (ID18) y clases adicionales obtenidas: pastos(ID 15), Plantaciones forestales(ID 9), palma aceitera(ID 35) y mosaico de agricultura y pastos(ID 21).	
3.2 Estrategias complementarias previas a la clasificación	
3.2.1 Regiones de clasificación	4
3.2.2 Colecta de puntos	5
3.2.3 Importancia de variables	6
3.2.4 Mapas de Referencia	7
3.3. Resumen metodológico	8
4 Post-clasificación	10
5.1 Llenado de vacíos de información	11
5.2 Filtro temporal	12
5.3 Filtro espacial	12
5.4 Filtro de frecuencia	13
4.5 Integración con las otras clases MapBiomias Amazonía	13
5 Resultados	
6 Referencias	

1 Introducción

En las colecciones 1 y 2 de MapBiomias Amazonía se trabajó sólo en el mapa general la clase de "mosaico de agricultura y pastos", la cual abarca cultivos perennes, semiperennes, anuales y plantaciones como la palma aceitera y árboles frutales, así también como los pastos cultivados para ganado. Para la colección 3, se buscó separar ambas clases "agricultura" y "pastos", y trabajarlas como temas transversales, sin embargo no se logró este objetivo en su totalidad, pero sí se pudo obtener una versión mejorada de la clase mosaico de agricultura y pastos, así como una metodología ya consolidada. Dentro de esta metodología se adaptaron procesos particulares que permitieron llegar a una mejor versión, como por ejemplo la aplicación de muestras por rango de año, definición poligonales de áreas o mapas base como área interés. En consecuencia, el tema transversal "agricultura" al final de la col.3 se llamó transversal de "mosaico de agricultura y pastos", los países que trabajaron esta clase como transversal fueron Perú y Bolivia. El resto de países clasificaron esta clase dentro del mapa general. En la colección 4, se realizaron ajustes en la metodología y la experiencia de la colección anterior, todos los países obtuvieron la clase de "agricultura", incluso en algunos países la metodología también permitió obtener la separación de otras clases vinculadas a agricultura como plantaciones forestales, palma aceitera, pasto y nuevamente una versión mejorada de la clase "mosaico de agricultura y pastos".

Para la Colección 5.0, se realizaron mejoras con respecto a la colección pasada, y se agregaron funcionalidades dentro de las clasificaciones que infirieron mejores resultados para este transversal.

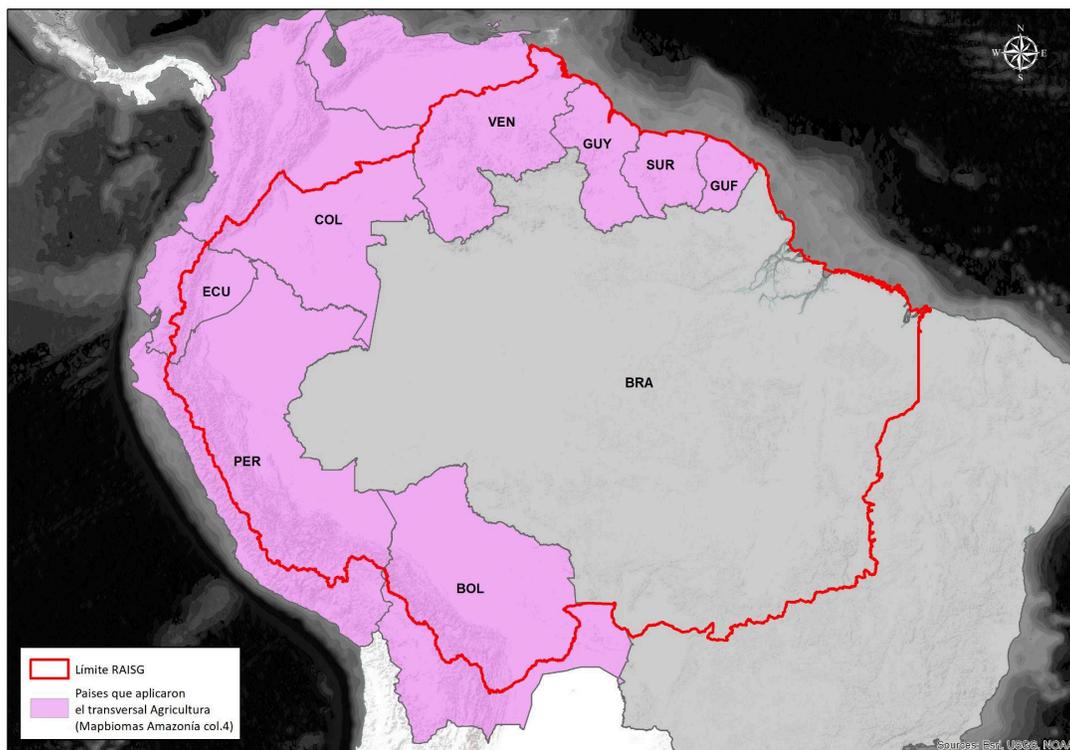


Figura 1. Países donde se aplicó la metodología del transversal Agricultura.

2 Mosaicos de imágenes

Para la clasificación del transversal "Agricultura" se utilizaron los mosaicos de imágenes generados previamente para la clasificación del mapa general de coberturas de la tierra de la Colección 5.0 de MapBiomás Amazonía. Se realizó la composición a través de imágenes satelitales de reflectancia de superficie con corrección atmosférica del sensor Landsat 5, 7, 8 y 9 que cubren la superficie asociada al límite RAISG que involucra los ocho países de la región amazónica (Bolivia, Colombia, Ecuador, Guyana, Guayana Francesa, Perú, Surinam y Venezuela) generados anualmente en una serie temporal de 38 años comprendidos desde 1985 hasta 2022¹.

2.1 Parametrización de mosaicos anuales

Un mosaico anual es la agregación de píxeles de varias imágenes Landsat a partir de los que se genera un mosaico representativo para un año, construido a partir de los siguientes parámetros²:

- ID: Identificador único de la unidad carta-región
- Año: Año de la serie (1985 a 2022) al que corresponde el mosaico.
- Carta: Código identificador de la carta
- Fecha inicial/ Fecha final: Periodo 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 OLI o combinaciones de Landsat 5 y Landsat 7 y Landsat 8 y Landsat 9.
- Nubosidad: 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.

Los parámetros de construcción de mosaicos anuales son definidos por el intérprete y representan los criterios de selección de imágenes disponibles en la Colección de datos Landsat a partir de los cuales se construye el mosaico anual. Las imágenes seleccionadas por año fueron *reducidas* a una imagen individual, o mosaico anual, empleando operadores llamados *reductores* existentes en Google Earth Engine, como se ilustra en la Figura 8.

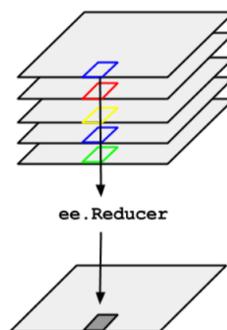


Figura 2. Esquema de la aplicación de un reductor a una colección de imágenes (Google, 2020³).

¹ La composición de mosaicos se generó a partir del reductor de píxel de mediana con una resolución espacial de 30 metros.

² Cada país evaluó y definió un periodo óptimo del año tomando como criterios: 1) la disponibilidad o cobertura de datos, 2) el contraste espectral entre clases y 3) las características fenológicas de la cobertura vegetal en diferentes momentos del año.

³ Tomado de: https://developers.google.com/earth-engine/guides/reducers_image_collection

Al momento de parametrizar los mosaicos se consideró que se puede alcanzar valores de precisión mayores al utilizar mosaicos de imágenes satelitales cuyos datos cuenten con la menor cantidad de ruidos. Es por ello que se buscó que cada mosaico tenga la menor presencia de nubes e interferencias posible y la mayor extensión de cobertura de datos Landsat disponible dentro del periodo definido. En casos excepcionales se dio el caso de que no hubo imágenes disponibles para el período seleccionado, donde se optó por extender el periodo de búsqueda de imágenes.

Nubes y sombras de nubes son enmascaradas previamente para que únicamente píxeles libres de nubes y sombra de nubes sean seleccionados de las imágenes disponibles. Los métodos de enmascaramiento de nube y sombra de nube fueron el Cfmask y CloudScore.

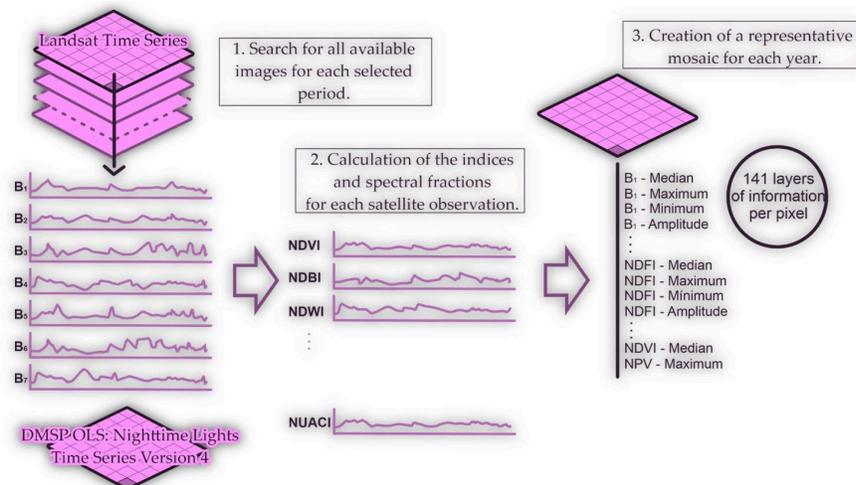


Figura 3. Proceso de generación de mosaicos anuales de imágenes Landsat

2.2 Variables de clasificación o feature space

Se calcularon variables (*feature space*) a partir del mosaico anual que representan los insumos del proceso de clasificación. Las bandas Landsat, junto con las variables de clasificación se encuentran consolidados en archivos ráster compuestos por 141 bandas en total⁴ que incluyen: las bandas landsat espectrales, índices espectrales, información fraccional y de textura derivada de las mismas e índices de las fracciones espectrales. Adicionalmente se usaron 7 variables estáticas: HAND, shademask2, slppost, altitud, pendiente, latitud y longitud; que ayudaron a la clasificación de clases que espectralmente son muy similares pero se logran diferenciar por estos aspectos topográficos.

A las imágenes disponibles en cada año, se aplicó el cálculo de **reductores estadísticos** para generar los valores de cada píxel. Estos reductores son:

- Mediana: Mediana⁵ de todos los valores disponibles en el mosaico anual para esa ubicación (píxel).
- Mediana época seca: Cálculo de la mediana estadística aplicada a los píxeles del cuartil 25 (con los menores valores) de NDVI (proxy de época seca).
- Mediana época húmeda: Cálculo de mediana estadística aplicada a los píxeles del cuartil 75 (con los mayores valores) de NDVI (proxy de época lluviosa).
- Amplitud: Extensión de la variación entre todos los píxeles disponibles en el mosaico anual.

⁴ Disponibles para descarga en la plataforma MapBiomias Amazonía.

⁵ Mediana es el valor que separa la mitad superior de la mitad inferior de una muestra de datos o una población. [Documentación](#) de la herramienta en Google Earth Engine.

- Desviación estándar: Desviación estándar de los valores de todos los píxeles disponibles en el mosaico anual para una ubicación determinada.
- Mínimo: Menor valor de todos los píxeles disponibles en el mosaico anual en una ubicación determinada.
- Máximo: Mayor valor de todos los píxeles disponibles en el mosaico anual en una ubicación determinada.
- Mínimo del periodo seco: Cálculo del menor valor de todos los píxeles disponibles de las imágenes del cuartil con los menores valores de NDVI (proxy de época seca).
- Mínimo del periodo húmedo: Cálculo del menor valor de todos los píxeles disponibles de las imágenes del cuartil con los mayores valores de NDVI (proxy de época lluviosa).
- Máximo del periodo seco: Cálculo del mayor valor de todos los píxeles disponibles de las imágenes del cuartil con los menores valores de NDVI (proxy de época seca).
- Máximo del periodo húmedo: Cálculo del mayor valor de todos los píxeles disponibles de las imágenes del cuartil con los mayores valores de NDVI (proxy de época lluviosa).
- QMO del periodo seco: El valor más alto que tiene la banda en el índice evi2 en la estación seca.
- QMO del periodo lluvioso: El valor más alto que tiene la banda en el índice evi2 en la estación húmeda.

2.3 Variables específicas para el transversal agricultura

En la colección anterior (col.3) previo a un análisis, se obtuvieron 33 variables específicas o de mayor importancia para la clasificación de la clase agricultura, estas 33 variables con sus reductores fueron incorporadas a las 141 variables del mapa general.

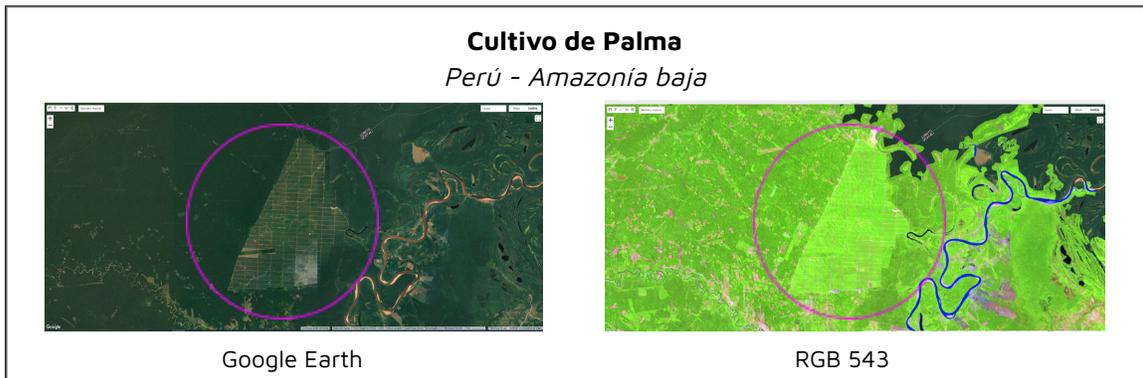
Tabla 1. Variables y sus reductores de mayor importancia para el transversal agricultura

Bandas	min	median	max	qmo
green_dry		x		x
red_dry	x	x	x	x
nir_dry		x		x
swir1_dry		x	x	x
swir2_dry	x	x		x
green_wet	x	x		x
red_wet			x	
nir_wet		x		x
swir1_wet	x	x	x	x
swir2_wet		x	x	x
NDWI_wet	x	x	x	x
CEI_wet	x			
NDWI_dry	x			

3 Clasificación

3.1 Definición de la clase agricultura (ID: 18) y otras obtenidas en la metodología de este transversal.

- **Agricultura (ID 18):** Son áreas donde la cobertura original ha sido modificada o reemplazada por cultivos anuales, temporales y perennes. Pueden estar activos o ser tierras en reposo. Comprende cultivos de arroz, árboles frutales, palma (excepto Colombia, ya que logró diferenciarlo y extraerlo como clase a parte) y otros tipos de cultivos. Todos los países mapearon la clase agricultura.

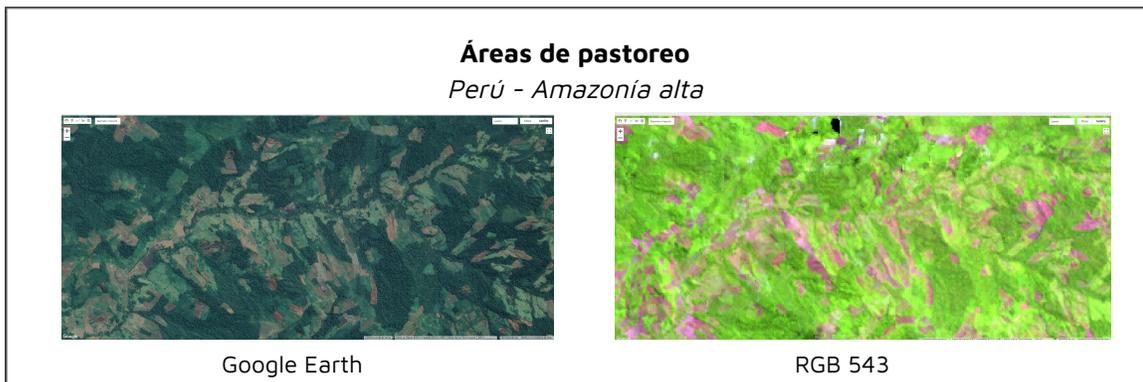


Complementario a la clase "Agricultura", la metodología también permitió clasificar la clase "pastos", la mayoría de países optó por los resultados de "pastos" obtenidos en el transversal de agricultura que los obtenidos en el propio transversal de "pastos". Esto más que todo por la diferencia de variables en los mosaicos.

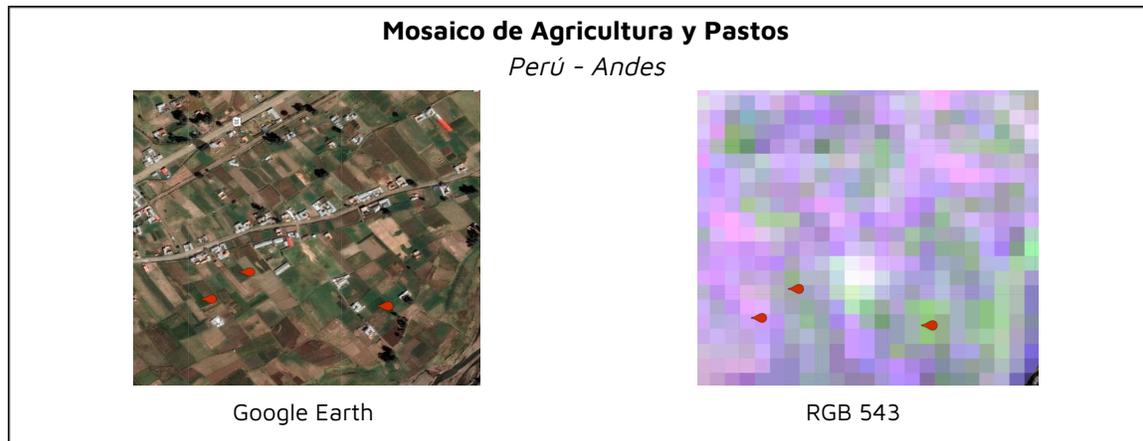
Adicional a la clase pastos, también se obtuvo el mapeo de plantaciones forestales, palma aceitera y una versión mejorada de la clase "mosaico de agricultura y pastos"; cabe indicar que estas clases "adicionales" sólo se dieron en algunos países.

A continuación la definición de las clases adicionales conseguidas en la metodología del transversal agricultura:

- **Pasto (ID 15):** Áreas de pasto, plantado o natural, vinculadas a la actividad pecuaria. Los países que mapearon pasto con la metodología del transversal agricultura fueron Perú, Bolivia y Colombia.



- **Mosaico de Agricultura y Pasto (ID 21):** Áreas de uso agropecuario en las cuales no fue posible separar las clases agricultura y pasto; ó un uso agropecuario conformadas por arreglos en sistemas agroforestales o silvopastoriles.
Los países que que utilizaron este transversal para obtener la clase mejorada de "mosaico de agricultura y pastos" fueron Perú y Colombia.



3.2 Estrategias complementarias previas a la clasificación

Las estrategias previas a la clasificación de agricultura, permiten obtener un mejor resultado de mapeo de la clase y reducir los posibles errores de sobreestimación o subestimación. Entre ellas está la generación de regiones de clasificación específicas para esta clase, máscaras y muestras de entrenamiento para obtener la importancia de variables según biomas o regiones.

3.2.1 Regiones de clasificación.

Uno de los primeros pasos fue la creación de regiones de clasificación específicas para mapear esta clase, basándonos en características comunes; ya sea por similitud en tipos de cultivo, zonas de mayor o menor actividad agrícola, etc. Esto se realizó basándonos que al agrupar por similitudes, la respuesta espectral en cada región iba ser similar y por ende reducir posibles confusiones del clasificador Random Forest.

Cabe indicar, que la metodología te permitía usar regiones de clasificación específicas de agricultura y a la vez, regiones de clasificación del mapa general.

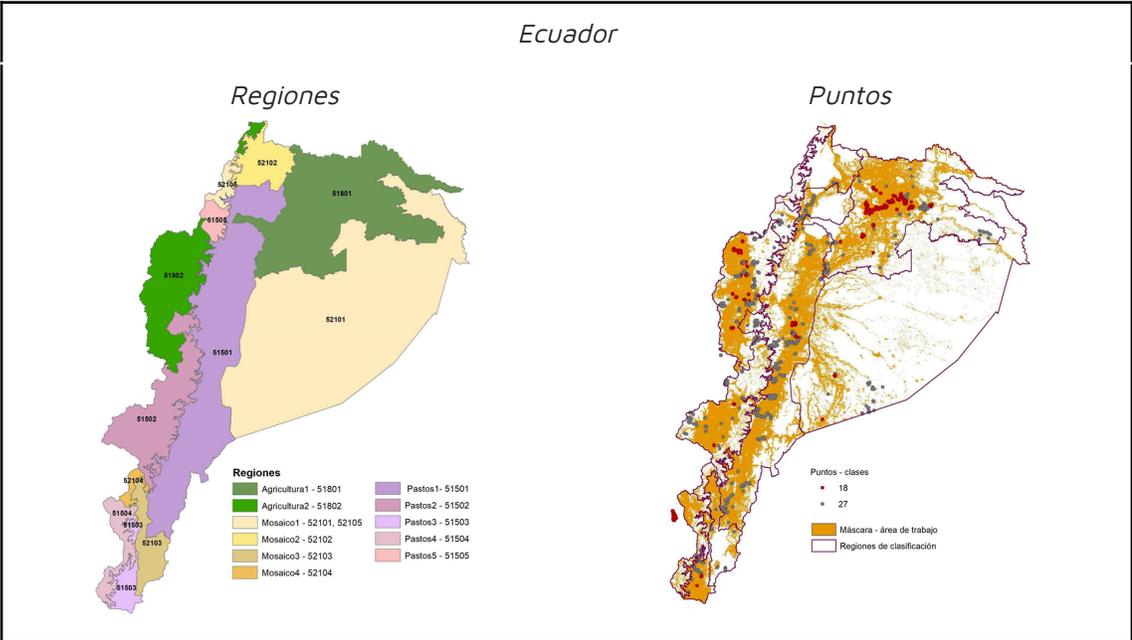
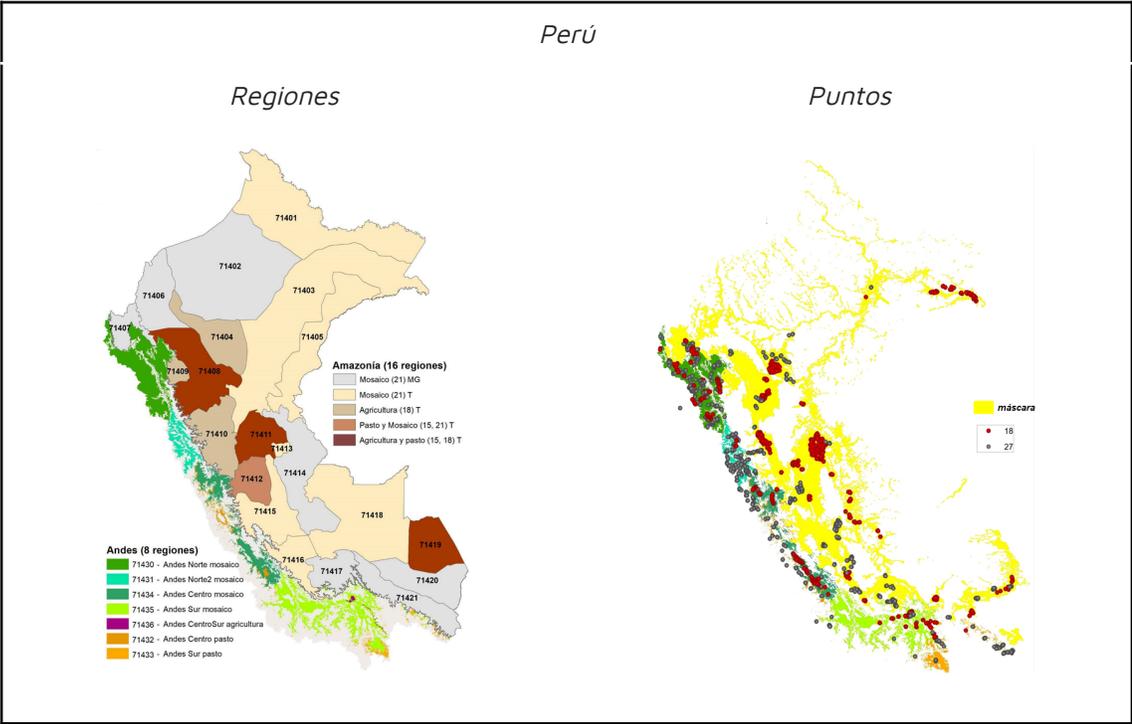
Resaltar que estas son estrategias para mejorar la detección de la clase, los resultados son integrados posteriormente.

Algunos países también generaron máscaras, es decir, áreas más reducidas que las regiones de clasificación, donde el mapeo de agricultura se daría sólo dentro de estas máscaras, reduciendo de esta manera el ruido o sobreestimación de la clase. Estas máscaras se obtuvieron a partir del acumulado de la clase "mosaico de agricultura y pastos" de la colección anterior.

3.2.2 Colecta de puntos para importancia de variables.

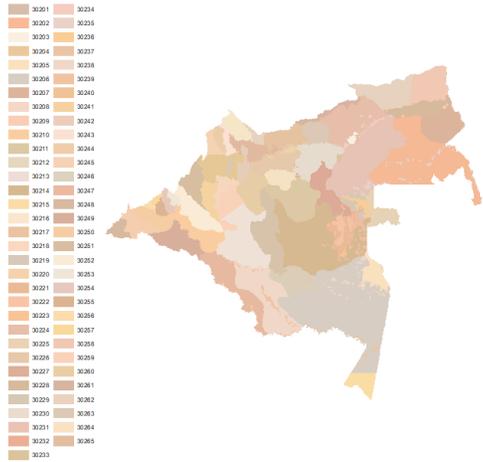
La colecta de puntos para la importancia de variables, se realizó por un año específico o un periodo de tiempo determinado, se obtuvieron a través de mapas de referencia por países o con la ayuda de imágenes de satélite de mejor resolución.

A continuación se muestran las regiones de clasificación específicas de agricultura, máscaras y puntos de muestreo generados por los países.

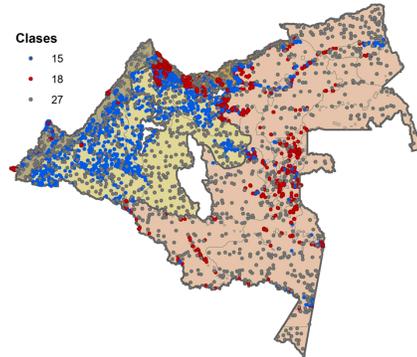


Colombia

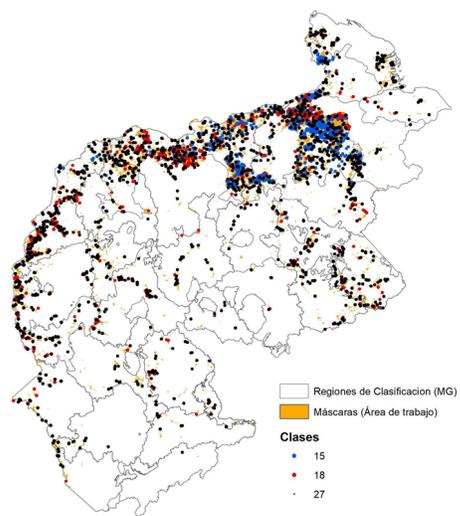
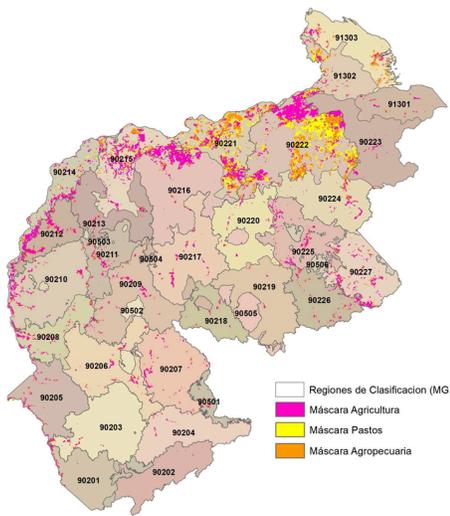
Regiones

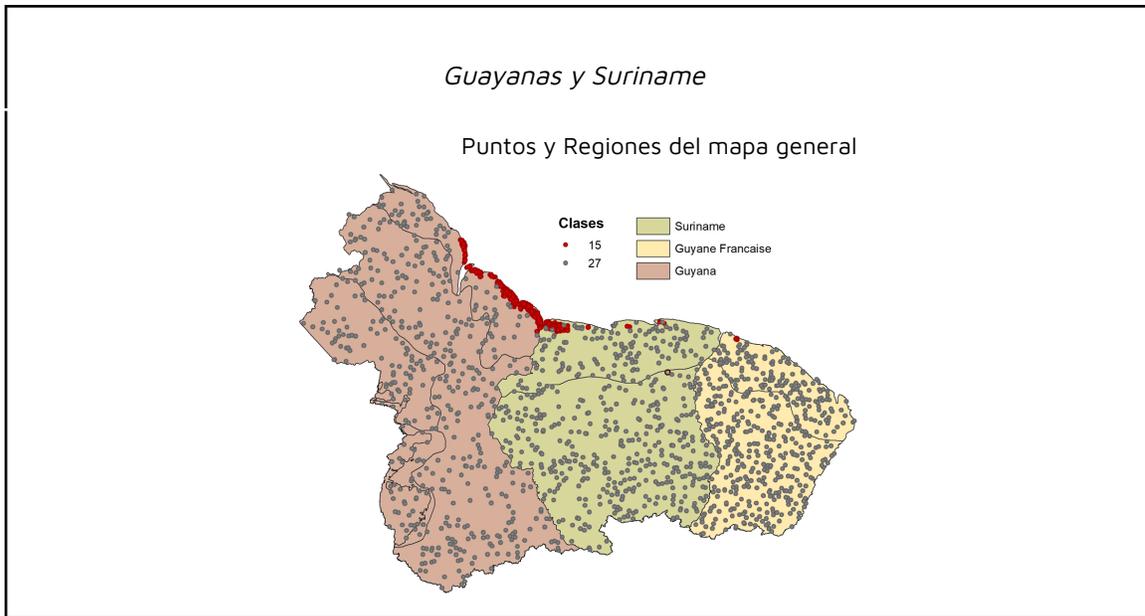


Puntos



Venezuela

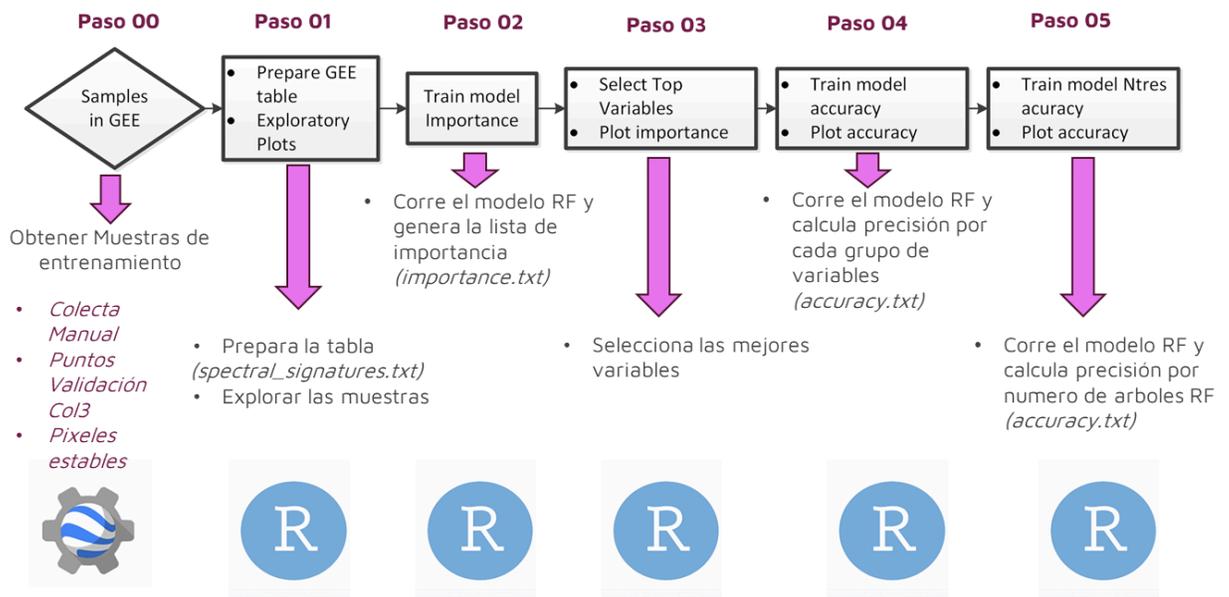




3.2.3 Importancia de variables.

La importancia de variables se generó por países, a su vez, cada país pudo obtener las variables más importantes por biomas o regiones de clasificación, debido a diferencias por tipos de cultivos o diferencias climáticas de latitudes y altitudes.

Figura 4. Flujograma metodológico para obtener la importancia de variables



3.2.4 Mapas de Referencia

La siguiente lista de mapas de referencia por países se utilizó como información base para la generación de colecta de puntos (generación de importancia de variables), máscaras o regiones de clasificación de agricultura.

Tabla 2. Fuentes de Referencia

Pais	Nombre	Fuente	Año
Perú	Mapa Nacional de Superficie Agrícola, según departamento	MINAGRI	2020
Bolivia	Pérdida de cobertura vegetal	Global Forest Change	2020
Ecuador	Mapa de Cobertura y uso de la tierra y Sistemas productivos agropecuarios del Ecuador continental (versión editada por el Ministerio de Agricultura y Ganadería en 2020), escala 1:25.000	MAG	2009-2015
Colombia	Territorios artificializados (Cultivos) - Coberturas de la Tierra 100K Mapa de bosque no bosque de Colombia	IDEAM	2018
		SMBYC	2021
Venezuela	Guayana Venezolana: Mapa de Vegetación Capa de referencia interpretada y digitalizada a partir de imágenes de alta resolución espacial	Otto Huber	1995
		Provita	2018
Guyanas	ESA World Cover 2020 v100 ESRI Global LULC 10m Global cropland expansion in the 21st century	ESA	2020
		ESRI	2021
		GLAD	2021

3.3 Resumen metodológico.

Los siguientes diagramas muestran el resumen por pasos de la metodología de agricultura.

Figura 5. Flujograma metodológico resumido del transversal agricultura

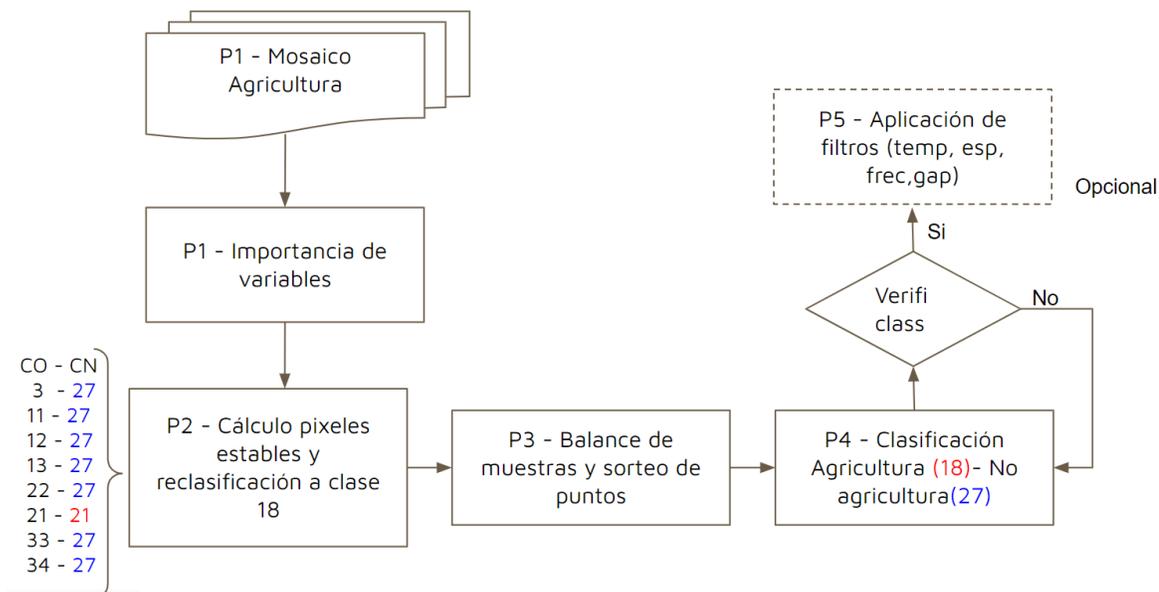
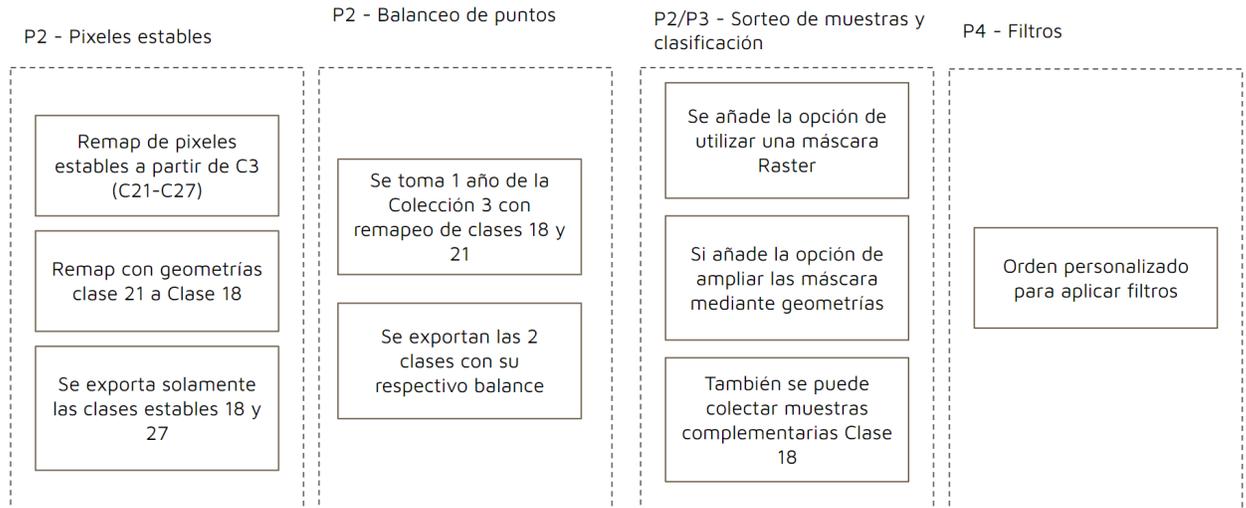


Figura 6. Flujograma metodológico descriptivo del transversal agricultura



4 Post-clasificación

Debido a la naturaleza basada en píxeles del método de clasificación y la serie temporal muy larga, se aplicó una cadena de filtros de clasificación posterior. El proceso de post clasificación incluye la aplicación de un filtro de relleno (gapfill), temporal, espacial y de frecuencia.

No existe una secuencia definida para la aplicación de los filtros, esto varía según el resultado de cada región de clasificación, sin embargo se recomienda empezar por el llenado de vacíos (gapfill).

4.1 Llenado de vacíos de información (gapfill)

En una serie de tiempo larga de regiones severamente afectadas por nubes, se espera que los vacíos puedan ocurrir en el resultado de la mediana. En este filtro, los vacíos son reemplazados por la clasificación temporalmente más cercana. En este procedimiento, si no hay disponibilidad de un valor en el año siguiente, entonces el espacio se reemplaza con el valor del año precedente. Se pueden utilizar hasta tres años anteriores para cubrir vacíos. Por lo tanto, las brechas solo deberían existir si un píxel dado se ha clasificado permanentemente como sin información en toda la serie temporal.

Para realizar un seguimiento de los orígenes temporales de los píxeles, se construyó una máscara de años, como se muestra en la Figura 7.

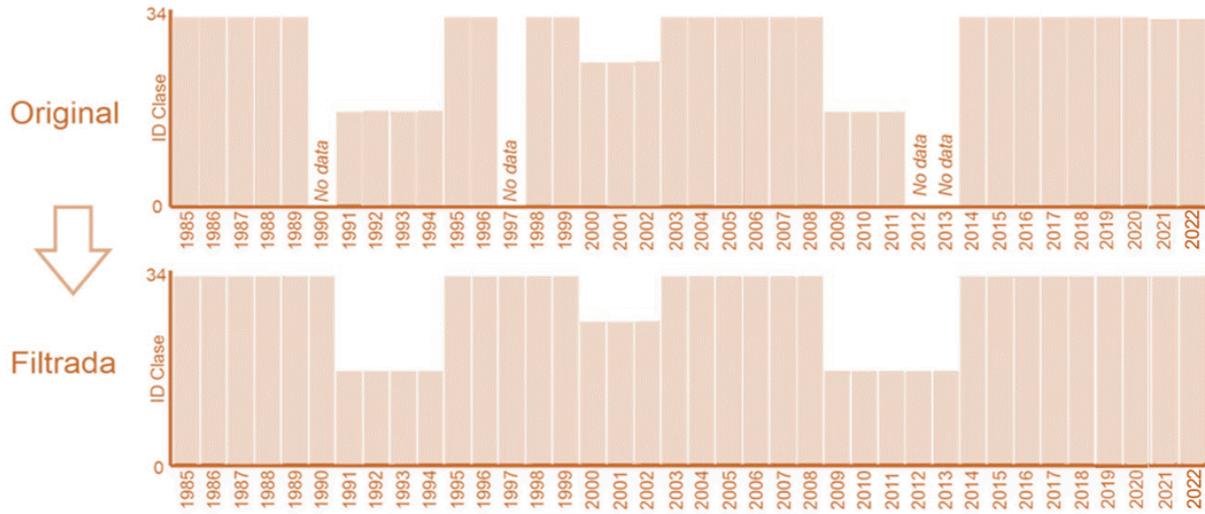


Figura 7 – Funcionalidad del filtro de Gap Fill

4.2 Filtro temporal

El filtro temporal utiliza clasificaciones secuenciales en una ventana móvil unidireccional de 3, 4 o 5 años para identificar transiciones temporalmente incoherentes. Basado en una sola regla genérica (RG), el filtro temporal inspecciona la posición central de los años consecutivos, y si los extremos del análisis son idénticos pero la posición central no lo es, entonces el píxel central se reclasifica hacia la clase del vecino temporal.

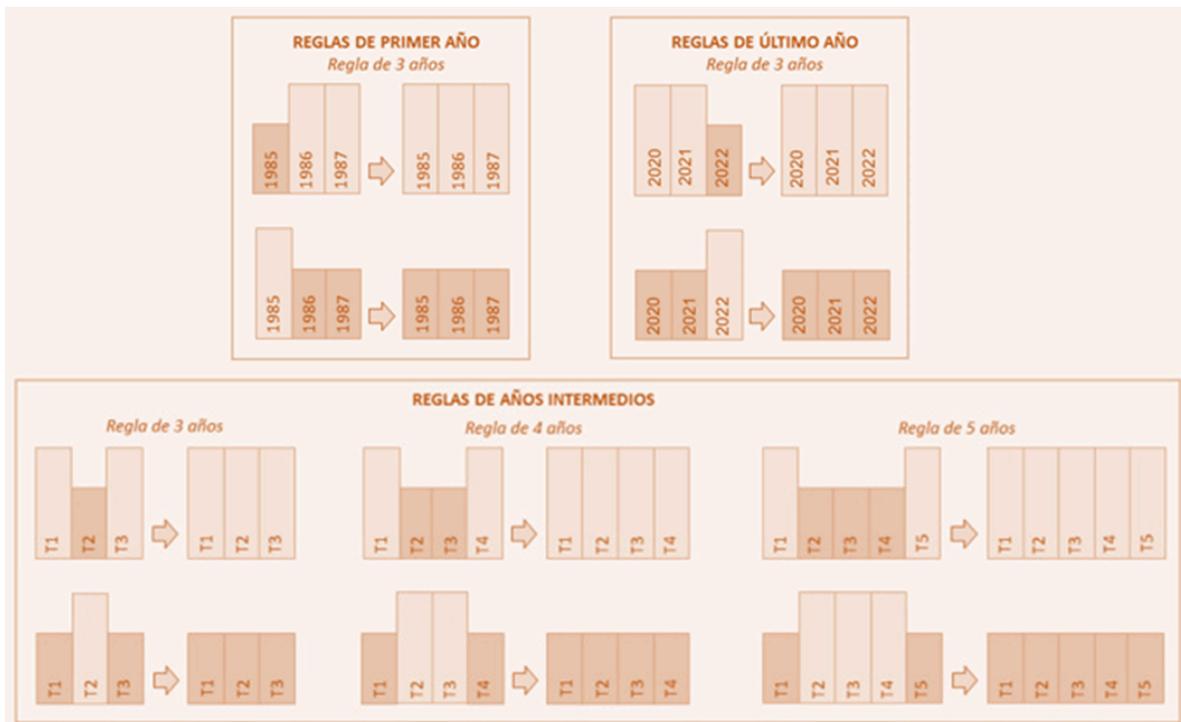


Figura 8 – Filtro Temporal

4.3 Filtro espacial

El filtro espacial basado en la función "connectedPixelCount", que identifica al conjunto de píxeles (vecindario) que comparten un mismo valor de píxel. En consecuencia, sólo los píxeles que no presentan conexión considerando una cantidad mínima predefinida de píxeles idénticos son definidos como píxeles aislados y reclasificados (Figura 9).

Para la clase de Agricultura, la conexión mínima establecida fue de 3 píxeles (aprox. 0.3 ha).



Figura 9 – Efecto de la aplicación del filtro espacial

4.4 Filtro de frecuencia

Este filtro considera la ocurrencia de la clase a lo largo de la serie temporal para normalizar la trayectoria histórica de la clase. Teniendo en cuenta que el comportamiento de agricultura manifiesta una tendencia incremental y de expansión hacia los últimos años, se aplicó el filtro con el fin de regular el aumento consistente de los píxeles de agricultura y evitar fluctuaciones anómalas de un año a otro.

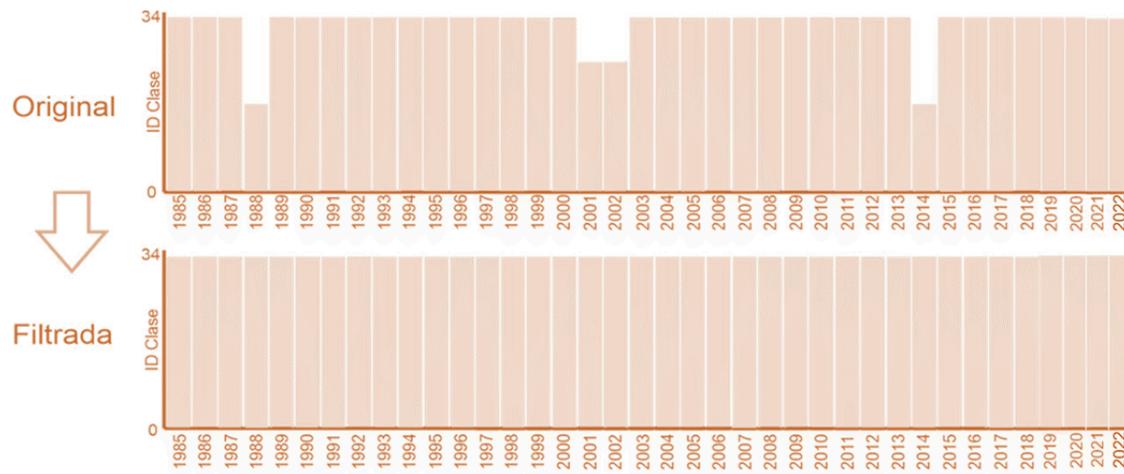


Figura 10 – Efecto de la aplicación del filtro de frecuencia

4.5 Filtro de incidencia

El filtro de incidencia estabiliza el valor de píxeles de clase agricultura que cambiaron demasiadas veces a lo largo de la serie temporal de 37 años. Todos los píxeles que cambian más de una cantidad de veces definida por el usuario y que están conectados con por lo menos n píxeles, fueron reemplazados por el valor de la clase de moda de ese píxel. Esto reduce cambios a lo largo de los bordes de las clases y ayuda a estabilizar las transiciones erradas por ruidos.

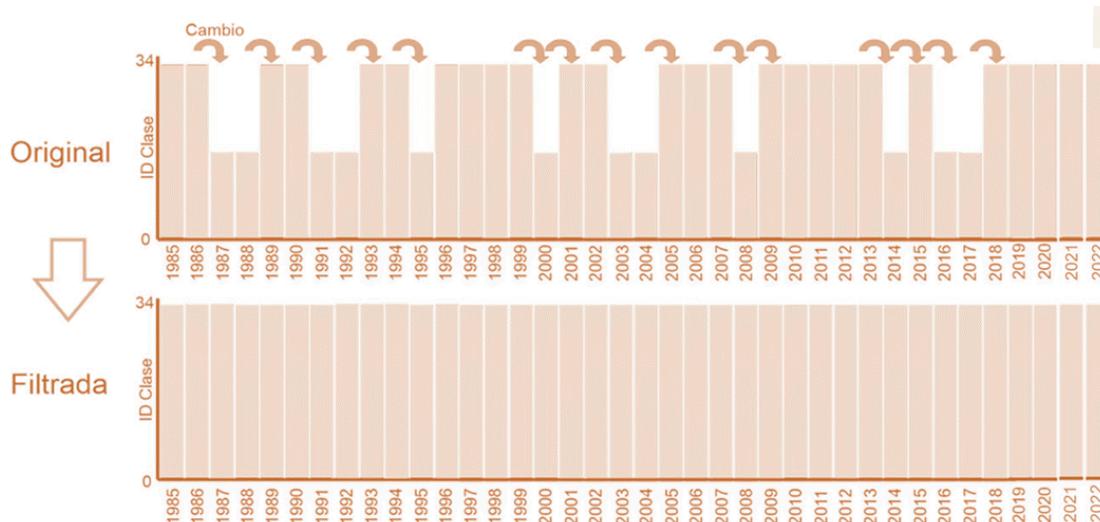


Figura 11 – Efecto de la aplicación del filtro de incidencia

5 Resultados

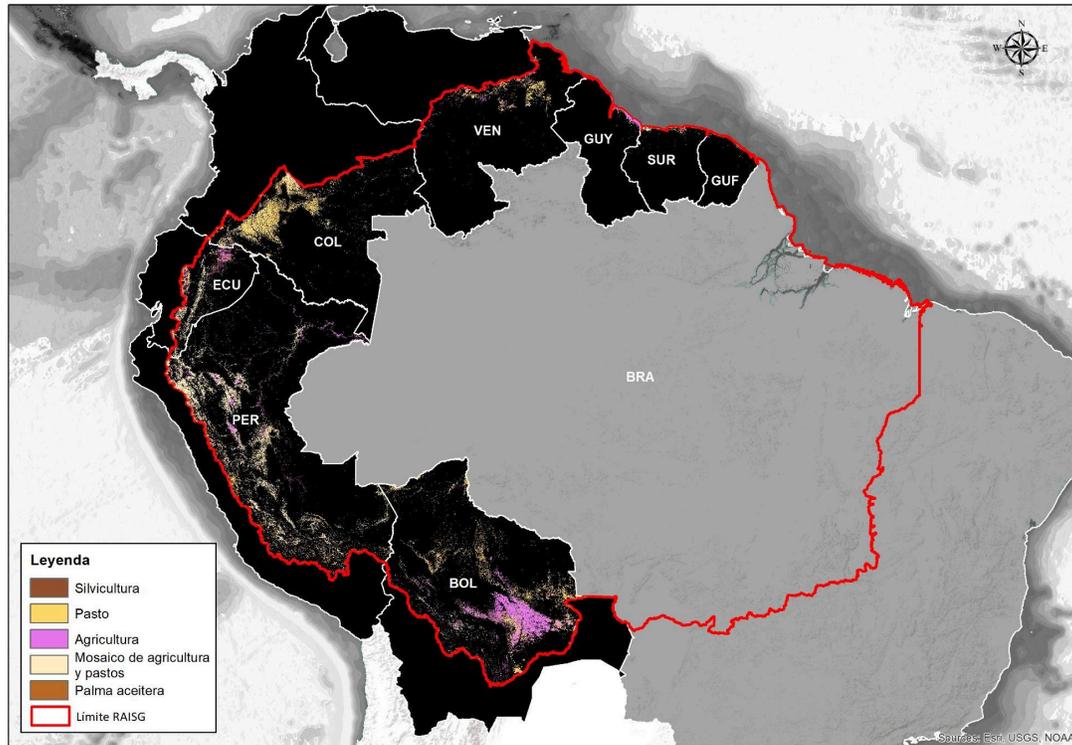
Por las distintas estrategias y características de la metodología se logró el objetivo de separar agricultura de pastos en todos los países; sin embargo, en muchas áreas aún resulta dificultoso esa separación, debido a la falta de fuentes de referencia, al amplio periodo de mapeo en la serie temporal y principalmente porque las extensiones de estas clases son pequeñas y mezcladas entre sí, por lo que resulta muy difícil la separación con un pixel de 30 metros de resolución. Este tipo de usos permanecieron en la clase de “mosaico de agricultura y pastos”.

A continuación se muestra los resultados o clases por países obtenidas por la metodología del transversal agricultura:

Tabla 3. Resultados del transversal agricultura por países

País	Clases mapeadas en transversal Agricultura
Bolivia	Agricultura (18), Pastos (15)
Perú	Agricultura (18), Pastos (15), Silvicultura (9), Mosaico de agricultura y pastos (21), Palma aceitera (35).
Ecuador	Agricultura (18) y Palma Aceitera(35).
Colombia	Agricultura (18), Pastos (15) , Mosaico de agricultura y pastos (21).
Venezuela	Agricultura (18)
Guyanas y Suriname	Agricultura (18)

Figura 12 – Mapa del resultado por países



6 Referencias

1. Breiman, Leo. "Random Forests." *Machine Learning* 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.
3. Mapbiomas. Agriculture and Forest Plantation - Appendix, Brazil 2020.