

Proyecto “Realizar el cálculo y análisis de la exactitud temática de los datos de actividad, estimadores ajustados e incertidumbre, para los mapas temáticos de Superficie de Bosque natural y cambio de la superficie de bosque (deforestación) a nivel nacional y regional para los reportes anuales de los años 2020, 2021 y 2022, de acuerdo con las especificaciones técnicas del Sistema de Monitoreo de Bosques y Carbono (SMBYC) de IDEAM”

Producto 1: Entrega del Informe final: i) Evaluación de exactitud temática de los productos cartográficos “Superficie de Bosque natural” del 2022 y “Cambio en la superficie de Bosque natural” a nivel nacional y regional del 2019-2020, ii) Capas geográficas y salidas gráficas de los puntos de muestreo evaluados.

Contrato No. VA-II-CCON-025-2024

Presentado por:

4D ELEMENTS CONSULTORES

Equipo:

Milton Romero Ruiz

Adriana Sarmiento Dueñas

Carlos Pedraza

Dallan Beltrán

Bogotá, 5 de septiembre de 2024

TABLA DE CONTENIDO

1. INTRODUCCIÓN	5
2. OBJETIVO.....	6
2.1 Objetivos específicos	6
3. METODOLOGIA	7
3.1 Recopilación de la información	7
3.2 Selección del área a evaluar	8
3.2.1. Seleccionar la capa temática a evaluar.....	8
3.2.2. Elaboración de la mascará del área de interés.....	8
3.3 Generación y diseño de la muestra.....	8
3.3.1 Definición del porcentaje de exactitud temática en general	8
3.3.2 Definición del tamaño de la muestra	8
3.3.3 Generación del proyecto.....	9
3.4 Interpretación de la muestra.....	10
3.4.1 Conformación del equipo de validación.....	10
3.4.2 Asignación de etiquetas de referencia	10
3.4.1 Consolidación y control de calidad.....	11
3.5 Generación de las estadísticas de validación.....	11
4. RECOPIACIÓN DE LA INFORMACIÓN	12
5. RESULTADOS: MAPA DE BOSQUE – NO BOSQUE, PERIODO 2022.....	17
5.1. Generación de la muestra y el proyecto	18
5.1.1. Definición del porcentaje de exactitud temática.....	18
5.1.2. Definición del tamaño de la muestra.....	18
5.1.3. Generación del proyecto	19
5.2. Interpretación de la muestra.....	20
5.2.1. Conformación del equipo de validación	20
5.2.2. Control de calidad de la interpretación	20
5.2.3. Generación de las estadísticas de validación (Matriz de error, cálculos de exactitud, estimadores de área ajustada, incertidumbre).....	22
6. RESULTADOS MAPA DE CAMBIO DE SUPERFICIE DE BOSQUE – NO BOSQUE, PERIODO 2019-2020	23
6.1. Selección de área a evaluar	23
6.1.1. Selección de la capa temática a evaluar	23
6.1.2. Generación del mapa de riesgo	24
6.2. Generación y diseño de la muestra.....	25

6.2.1.	Definición de porcentajes de exactitud temática:	25
6.2.2.	Definición de tamaño de la muestra:	26
6.2.3.	Generación del proyecto	27
6.3.	Interpretación y control de calidad de los puntos	28
6.3.1.	Conformación del equipo de validación	28
6.3.2.	Asignación de etiquetas de referencia.....	28
6.3.3.	Control de calidad de la interpretación	30
6.4.	Generación de las estadísticas de validación (Matriz de error, cálculos de exactitud, estimadores de área ajustada, incertidumbre)	31
6.	BIBLIOGRAFIA	33
7.	ANEXOS	35
	Anexo 1: Ejemplos de puntos de no coincidencia revisados para el mapa de bosque-no bosque 2022	35
	Anexo 2: Ejemplos de puntos de no coincidencia revisados para el mapa de cambio de superficie de bosque 2019-2020	39

INDICE DE FIGURAS

Figura 1. Diagrama metodológico para el desarrollo del trabajo.	7
Figura 2. Mosaicos de los productos Landsat last y median: a) Landsat 2019_22042019; b). Landsat_2019_Mediana_06042020; c). Landsat 2020_Last_pixel_07042021 y d). Landsat 2020_Mediana_07042021 e) Last_pixel_2022_160223 y f). Landsat_2022_Mediana_V2_09032023	14
Figura 3. Mosaicos de los productos Sentinel A). sentinel_2019_LastPixel_Col_z18; B) sentinel_media_2019_sininfor; C). Last_pixel_sentinel_2020_050321; D). Mediana_sentinel_2020_utm; E). Mediana_sentinel_2022_230220_utm_Z18N_09032023 y F) Last sentinel-xxxxx	16
Figura 4. Mapa de cobertura de bosque / no bosque del 2022.	17
Figura 5. Puntos de muestreo a validar para el mapa de bosque no bosque del 2022.	19
Figura 6. Proyecto QGIS, base para la validación del mapa de bosque no bosque 2022	20
Figura 7. Distribución de los puntos de no coincidencia entre los intérpretes (puntos negros: coincidencias, puntos rojos: sin coincidencias).	21
Figura 8. Mapa de cambio de superficie de bosque – no bosque del 2019-2020.	24
Figura 9. Mapa de riesgo a la deforestación	25
Figura 10. Puntos de muestreo a validar para el mapa de cambio de superficie de bosque 2019-2020.	27
Figura 11. Proyecto QGIS, base para la validación del mapa de cambio en la superficie de bosque 2019-2020	28
Figura 12. Vinculación de la ventana diseño de Acatama, con Google Earth	29
Figura 13. Vinculación de la ventana diseño de Acatama, con CCD	29
Figura 14. Distribución de los puntos de no coincidencia entre los intérpretes, puntos negros 2 coincidencias, puntos rojos sin coincidencias, , para el periodo 2019-2020	31

INDICE DE TABLAS

Tabla 1. Calculo tamaño de muestra, para el mapa de bosque no bosque del 2022	18
Tabla 2. Número de puntos de acuerdo entre los intérpretes.	21
Tabla 3. Matriz de error tradicional y medidas de fiabilidad	22
Tabla 4. Matriz de error de la proporción de área y medidas de fiabilidad	23
Tabla 5. Intervalos de confianza del área estimada por clase.	23
Tabla 6. Calculo tamaño de muestra, para el mapa de cambio superficie bosque no bosque del 2019-2020	26
Tabla 7. Número de puntos de acuerdo entre los intérpretes analizados para el periodo 2019-2020.	30
Tabla 8. Matriz de error tradicional y medidas de fiabilidad para el periodo 2019-2020	32
Tabla 9. Matriz de error de la proporción de área y medidas de fiabilidad para el periodo 2019-2020	32
Tabla 10. Matriz de error de la proporción de área y medidas de fiabilidad para el periodo 2019-2020	33
Tabla 11. Intervalos de confianza del área estimada por clase para el periodo 2190-2020	33

1. INTRODUCCIÓN

La teledetección se ha convertido en la herramienta más confiable que permiten monitorear e interpretar el cambio en la cobertura del suelo y el uso de la tierra (Toure, Storw, Shih, Weeks, & Lopez-Carr, 2018). La información proveniente de las imágenes satelitales captadas por los sensores remotos contiene valiosa información acerca de la cobertura terrestre, la cual refleja las características de la superficie terrestre, facilitando el estudio de muchas disciplinas y convirtiéndose en apoyo y complemento las técnicas convencionales de muchos estudios actuales y futuros. Hoy en día, existe una amplia disponibilidad de imágenes satelitales, de diferentes resoluciones tanto espaciales como espectrales que han permitido que se desarrolle técnicas y métodos de clasificación e interpretación basados principalmente en píxeles y en objetos (Liu, Du, & Wang, 2018).

Las ventajas y desventajas de esas técnicas se han discutido en muchos estudios, y actualmente los procesos de monitoreo de coberturas se basan en su gran mayoría, en la detección de cambios temáticos a nivel de píxeles provenientes de imágenes satelitales de resolución espacial media en especial de imágenes LANDSAT. Un análisis multitemporal de imágenes satelitales más allá del paradigma de interpretación basados en píxel o de objetos geográficos, debe permitir obtener, una mayor coherencia en los reportes generados sobre las dinámicas de cambio de las coberturas, lo que se traduce en un mejor análisis y predicción de los cambios en la extensión y delimitación de un determinado fenómeno continuo en el espacio (Geroffrey & Steiniger, 2018)

Bajo el contexto de avance en la teledetección, Colombia se acoge a los acuerdos de la Convención Marco de las Naciones Unidas sobre Cambio Climático– CMNUCC durante las Conferencias de las Partes de 2009 y 2010 (COP 15 y 16, respectivamente), y recientemente en Varsovia (COP 19), que instituye que los países en desarrollo establezcan un sistemas nacionales de monitoreo forestal que permitan cuantificar las emisiones/absorciones de GEI y los cambios en la superficie de los Bosques y las reservas forestales de Carbono. Para ello se define como cobertura de Bosque la: *Tierra ocupada principalmente por árboles que puede contener arbustos, palmas, guaduas, hierbas y lianas, en la que predomina la cobertura arbórea con una densidad mínima del dosel de 30%, una altura mínima del dosel (in situ) de 5 m al momento de su identificación, y un área mínima de 1,0 ha. Se excluyen las coberturas arbóreas de plantaciones forestales comerciales, cultivos de palma, y árboles sembrados para la producción agropecuaria.*

Esta definición es consecuente con los criterios definidos por la CMNUCC en su decisión 11/COP.7, con la definición adoptada por Colombia ante el Protocolo de Kyoto (MAVDT, 2002), así como con la definición de la cobertura de Bosque natural incluida en la adaptación para Colombia de la leyenda de la metodología CORINE Land Cover -CLC Colombia-. Para el presente protocolo la deforestación está definida como: *La conversión directa y/o inducida de la cobertura de Bosque a otro tipo de cobertura de la Tierra en un periodo de tiempo determinado* (Defries, y otros, 2006). Por su parte, la regeneración se define como: *La recuperación de la cobertura de Bosque de zonas donde no estaba presente en fechas anteriores.*

De esta manera, en el 2011 los investigadores del IDEAM, evalúa diferentes técnicas de procesamiento de datos de sensores remotos para la detección de cobertura del Bosque y sus respectivos cambios en el tiempo, generando como resultado el “Protocolo de procesamiento digital de imágenes para la cuantificación de la deforestación en Colombia a Nivel Nacional - Escala gruesa y fina” (Cabrera, y otros, 2011), la cual hoy ha sido la apuesta metodológica para la cuantificación histórica de la deforestación a nivel nacional para los periodos 1990-2000, 2000-2005 y 2005-2010.

Este protocolo, además de dar las directrices necesarias para el procesamiento de imágenes, proponen seguir una serie de recomendaciones y consideraciones, que han sido aceptadas por la comunidad de sensores remotos como buenas prácticas, para al evaluar la exactitud de los mapas finales de cambio, la cual en gran medida acoge los lineamientos determinados por investigadores de la Universidad de Boston (Olofsson, Foody, Stehman, & Woodcock, 2012), (Olofsson, Foody, & Herold, 2014). De acuerdo con estas recomendaciones se debe validar la

exactitud temática de los productos que se generen en la elaboración de mapas de bosque/no bosque y del mapa de Cambio de bosque. Para ello se define el término validación como un *conjunto de técnicas para determinar la calidad de un mapa en particular*. Las técnicas incluyen la evaluación de la exactitud de un mapa dado, con base a observaciones tales como la precisión general y los errores de omisión y comisión por clase de cobertura de la tierra, y para ello se aborda principalmente tres componentes: (i) un diseño de muestreo probabilístico; (ii) un diseño de la respuesta que sea coherente con la información de referencia a usar; y (iii) un análisis para estimar la exactitud, a partir una matriz de error construida mediante estimadores insesgados. Acogiendo a estos lineamientos, este proyecto muestra el marco metodológico para el diseño de muestreo y su implementación, evaluación de la exactitud temática, cálculos de estimadores de área e incertidumbre, así como los procedimientos de control de calidad de la interpretación y seguimiento de las actividades.

Bajo este esquema, en el 2021, la empresa 4D Elements, realizó el cálculo y análisis de la exactitud temática de los datos de actividad, estimadores ajustados e incertidumbre, para los mapas temáticos de superficie de bosque natural y cambio de la superficie de bosque (deforestación) a nivel nacional y regional para el año 2018, de acuerdo a las especificaciones técnicas del Sistema de Monitoreo de Bosques y Carbono (SMBYC) de IDEAM, siendo este el piloto de implementación bajo los protocolos definidos por el IDEAM y los diferentes autores de referencia a la manera de generar este tipo de análisis. A partir de este análisis, en esta etapa se realiza el procedimiento de validación, retomando las lecciones aprendidas en este primer proceso y utilizando las ventajas en la evolución que se ha tenido en la información satelital, los avances de la herramienta ACATAMA y la experiencia que tuvo el equipo interpretador el cual está conformado por los mismos profesionales que elaboraron el primer análisis.

Este primer informe recopila la primera fase del proyecto en la cual se describe la conformación del equipo de trabajo, el diseño de muestreo, así como su implementación y los primeros resultados de avance de la evaluación de la exactitud del mapa de bosque no bosque del 2022.

2. OBJETIVO

Realizar el cálculo y análisis de la exactitud temática de los datos de actividad, estimadores ajustados e incertidumbre, para los mapas temáticos de Superficie de Bosque natural para el 2022 y cambio de la superficie de bosque (deforestación) a nivel nacional y regional para el reporte anual del año 2020, de acuerdo con las especificaciones técnicas del Sistema de Monitoreo de Bosques y Carbono (SMBYC) de IDEAM.

2.1 Objetivos específicos

Evaluación temática de los productos cartográficos “Superficie de Bosque natural” del 2022 y “Cambio en la superficie de Bosque Natural” a nivel nacional y regional de Colombia para el 2020, a partir de imágenes de satélite.

- Tipo y cantidad de imágenes e insumos utilizados
- Descripción del pre-procesamiento realizado a las imágenes en caso de ser necesario
- Tipificación de los errores de omisión y comisión detectados, socializados con el equipo técnico del SMBYC del IDEAM
- Matriz de error, cálculos de exactitud, estimadores de área ajustada, incertidumbre y demás cálculos y análisis requeridos
- Análisis de los datos recolectados.
- Reporte y análisis de la evaluación de exactitud temática de los productos cartográficos “Superficie de Bosque natural” del 2022 y “Cambio en la superficie de Bosque natural” a nivel nacional y regional para el año 2020 a partir de imágenes de satélite producidos por el IDEAM.

- Productos cartográficos generados en el proceso de clasificación y validación, incluido el shape con los puntos de muestreo clasificados, evaluados tipificados entre otros.
- Tabla con los cálculos de exactitud temática, estimaciones de área ajustada, incertidumbre y demás cálculos generados

3. METODOLOGIA

La Figura 1 muestra las metodológica implementada para validación de los mapas de cambio de bosque / no bosque para los reportes anuales de los años 2020, 2021 y 2022.



Figura 1. Diagrama metodológico para el desarrollo del trabajo.

3.1 Recopilación de la información

La primera fase consistió en la obtención de la información necesaria para realizar la evaluación temática del producto. Para ello se recopiló la siguiente información:

- i) Mapa de bosque/no bosque del periodo 2022 el cual contiene la información de coberturas distribuidas en tres categorías: bosque estable, no bosque estable y sin información.
- ii) Mapa de cambio de superficie de bosque del periodo 2019-2020, el cual contiene la información por las clases de bosque estable, no bosque estable, deforestación, regeneración y sin información.
- iii) Información satelital necesaria para la validación, la cual incluye imágenes de Landsat consolidadas e individuales del periodo de análisis, imágenes Sentinel a 10 metros de resolución, con 4 bandas

espectrales, link gratuito de observación de imágenes de alta resolución (Planet Scope) y aquellos insumos satelitales disponibles en páginas web (Google, ESRI, Bing) para los periodos analizados.

3.2 Selección del área a evaluar

Las técnicas de validación dependen en gran medida del diseño de muestreo de probabilidad para la obtención de los datos. Por ello, para la validación del mapa de bosque -no bosque – deforestación para los periodos de análisis se siguió la metodología planteada por (Olofsson P. , Foody, Stehman, & Woodcock, 2012), que establece un diseño metodológico a través de una **selección de tipo de muestreo** considerando la proporción de cada clase que ocupa en el territorio, siendo necesario identificarlas y asignarles un tamaño de muestra lo suficientemente grande como para obtener un error estándar pequeño en la estimación de la exactitud del usuario y la obtención de muestras a través de un **muestreo aleatorio estratificado** para conseguir los mejores resultados en el proceso de validación. Para la obtención de las muestras a validar se utilizó la herramienta ACATAMA (Accuracy Assessment of Thematic Maps) plugin que trabaja bajo el software de QGIS versión 3.10 en adelante y el cual fue desarrollada por el equipo del SMByC donde se incorpora los lineamientos de (Olofsson, Foody, & Herold, 2014). Los parámetros que se emplean para la evaluación de los mapas se describen a continuación:

3.2.1. Seleccionar la capa temática a evaluar

La evaluación se realizó para el mapa de bosque – no bosque del periodo 2022 y el mapa de cambio (bosque – no bosque – deforestación) para el periodo de 2019-2020, los cuales se encuentran en el sistema de coordenadas geodésicas WGS84, EPSG 32618, con una resolución espacial de 30 x 30 metros, generado por el Sistema de Monitoreo de Bosques y Carbono – SMByC.

3.2.2. Elaboración de la mascaró del área de interés.

Para evaluar el mapa de cambio de superficie de bosque, se elaboró un mapa de riesgo a partir de la creación de una mascaró de área de interés, a partir de los mapas históricos de bosque – no bosque – deforestación de los periodos comprendidos entre el 2012 – 2022. La finalidad de esta mascaró es el de incrementar los esfuerzos para la validación en las superficies donde pueda existir mayor probabilidad de deforestación y de esta forma realizar un balance de áreas entre las coberturas de bosque estable y no bosque estable respecto a las de cambio. De esta manera, el Sistema de Monitoreo de Bosques y Carbono – SMByC del IDEAM, generó el mapa de riesgo el cual se incorporó a los mapas de cambio y así se identificó los puntos de validación para cada uno de los mapas de cambio de superficie en evaluación, dividiendo cada cobertura en las áreas de alto riesgo y debajo riesgo para cada categoría de análisis.

3.3 Generación y diseño de la muestra

La generación de la muestra se basó en el método aleatorio estratificado por clases, acorde a lo propuesto por Pontius- Oloson y el IDEAM a lo largo de los años. Este método se basa en la proporción que define el número de puntos calculados acorde a la proporción del área y del error. En otras palabras, a aquellas clases que ocupan una pequeña porción del territorio, se les asignó un tamaño de muestra lo suficientemente grande para obtener un error estándar pequeño en la exactitud del usuario. Este mismo procedimiento se realizó para el mapa de bosque no bosque del periodo 2022. Los parámetros iniciales del modelo fueron:

3.3.1 Definición del porcentaje de exactitud temática en general

El valor inicial preliminar, seleccionado para la definición del porcentaje de exactitud temática de la muestra es el planteado por ACATAMA, acorde a la información proveniente de la capa temática. Este valor sirve de parámetro inicial el cual se ajusta al definir los parámetros del tamaño de la muestra.

3.3.2 Definición del tamaño de la muestra

El tamaño total de la muestra es determinado a partir de tres componentes acorde a la ecuación propuesta por (Cochran, 1977) definida como

$$n = \left[\frac{\sum W_i S_i}{S(\hat{O})} \right]^2 \leftrightarrow$$

Donde:

n = Tamaño de la muestra,

El primer componente corresponde a la proporción de área mapeada de la clase i $W_i = \frac{A_{m,i}}{A_{tot}}$, donde $A_{m,i}$ es el área mapeada de la clase i y A_{tot} es área total de las clases evaluadas en el mapa y donde

Área

$$A = \frac{T_P^2 * (\text{Pixel numbers})}{10000}$$

El segundo es el producto entre $U_i * U_i - 1$, donde U_i es la exactitud esperada de cada clase clasificada por el usuario, $U_i - 1$ es el complemento de U_i , y se considera el error máximo esperado o la probabilidad de que el área mínima de mapeo sea asignada a otra clase distinta.

$$S_i = \sqrt{U_i * (1 - U_i)}$$

Para las clases de Bosque y No Bosque, se les asignó un valor a del 90% ya que las firmas espectrales tienen un intervalo de confianza de 90%, se espera que la exactitud de cada cobertura este por encima o igual del valor de los intervalos de confianza. La clase no estudio es incluida debido a que se necesita puntos de validación para verificar la clasificación de dicha clase.

El tercer componente es el error estándar general y $S\hat{O}$, que se determina de acuerdo a la exactitud esperada de la validación, así el tamaño total de la muestra es la sumatoria de la multiplicación entre la proporción del área de cada clase (W_i) y el error estándar de cada clase (S_i), dividido entre el error estándar general (S_0). Lo anterior elevado al cuadrado.

Para ajustar la desviación estándar U_i se le asignó un valor teórico de exactitud de usuario sugeridos por el SMyC, los cuales corresponden a 0,9 para las coberturas de bosque estable y no bosque estable, por considerarse que esta clase suele tener una alta exactitud, mientras que, para las clases de deforestación, sin información y regeneración se les asignó 0,8. Estos valores se analizaron y ajustaron en conjunto con los especialistas del SMyC Tabla 1.

3.3.3 Generación del proyecto

Para la generación del proyecto metodológicamente se definió la distribución espacial de los sitios de verificación aleatoria acorde a los siguientes parámetros:

- **Distancia entre las muestras:** 500 m, este valor se definió con el fin de que no se traslapen los pixeles a seleccionar con respecto a la mínima unidad cartográfica que es 1 ha, (9 pixeles dentro de una rejilla de 3 x 3 = 180 para que no se traslapen las dos clases). Acorde al protocolo del IDEAM se designó una separación de 500 metros.

- **Número de vecinos:** 8 puntos alrededor de la clase que va ser evaluada. La evaluación se realizó únicamente para el punto central, no para los vecinos.
- **Mínimo número de vecinos de la misma clase:** 8 puntos alrededor de nuestra clase sea el mismo. Es decir, si se tomó un punto de deforestación, los 4 vecinos del pixel central también tienen que corresponder a puntos de deforestación. Si se elige más de 8 vecinos, aquellas áreas que tienen solamente una hectárea; las cuales tendrían solamente 8 vecinos de la misma clase no serían candidatos a ser seleccionados como muestra perdiéndose gran proporción de la información. Igualmente, esta selección obedeció a que se quiso evitar en la medida de lo posible, que algunos puntos cayeran al borde de cambio entre clases y debido a que ese límite muchas veces no es claro en las imágenes satelitales a evaluar se debe evitar estos límites.
- **Número máximo de interacciones:** La definición de las interacciones depende de la cantidad de memoria RAM y en nuestro caso se dejó de 6000. Se generó una rejilla a todo el país de puntos acorde a los parámetros determinados anteriormente. A definir el número de interacciones, se permite ubicar de la mejor manera los puntos, con el fin de cubrir mejor el área a evaluar.
- **Semilla aleatoria:** El valor corresponde al año de análisis que para bosque/no bosque fue 2022 y para el mapa de cambio de superficie fue de 2019. Este valor se refiere a la generación de números pseudoaleatorios, que tratan de imitar las propiedades de generaciones independientes de una distribución uniforme en intervalo (0,1 bite). De esta manera al hacer esta semilla el generador, puedes predecir (y reproducir) la secuencia.

3.4 Interpretación de la muestra

Una vez obtenidos los puntos a validar se inicia el proceso de interpretación de los mismos los cuales se realizan en los siguientes pasos:

3.4.1 Conformación del equipo de validación

La primera tarea una vez generado los puntos a validar, es la conformación del equipo de validación, acorde a lo establecido por el documento metodológico de operación estadística, monitoreo de la superficie de Bosque Natural en Colombia (IDEAM, 2019). Este equipo está constituido por:

- El líder de la evaluación, encargado de coordinar el trabajo de los intérpretes, realizar el diseño e implementación de una muestra probabilística; realizar la consolidación y verificación de la interpretación; y efectuar el análisis de exactitud.
- Tres (3) profesionales con amplia experiencia en interpretación visual y digital, y entrenados para aplicar el concepto de Cobertura boscosa en imágenes ópticas de media resolución.

3.4.2 Asignación de etiquetas de referencia

Inicialmente, se hizo una reunión del coordinador con su grupo de intérpretes, con el fin de considerar los mismos parámetros para la clasificación de los puntos. Posteriormente, cada uno de los intérpretes de manera individual inicia la comprobación de cada uno de los puntos de muestreo obtenidos a partir de la interpretación visual de las imágenes de alta resolución de Planet Scope (<https://www.planet.com/>). Estas imágenes están disponibles para consulta de usuarios mensualmente; los mosaicos de los satélites Landsat7 ETM+, Landsat 8 OLI, y Sentinel generados por el Sistema de monitoreo de Bosques y Carbono SMBYC del IDEAM; la información de las imágenes LANDSAT individuales y las herramientas de apoyo de imágenes disponibles por internet (ESRI, BING, GOOGLE).

Para ello, cada uno de los intérpretes realiza la interpretación con ayuda de la herramienta de QGIS Acatama, definiendo una escala de referencia de 200, la cual se va modificando acorde a la dificultad de la interpretación y siguiendo los siguientes parámetros:

- Para la definición de bosque se optó el concepto de la FAO como aquella superficie de árboles con altura mínima del dosel de 5 metros que cubren una hectárea en más del 30%. Para ello, se excluyen aquellas coberturas arbóreas de plantaciones forestales comerciales, cultivos de palma y mosaicos de cultivos arbóreos.

- Se evalúa cada una de las clases que se encuentran inmersas dentro del mapa de bosque no bosque (Bosque, sin información, no bosque) y superficie de cambio (bosque estable, bosque no estable, deforestación, regeneración y sin información).
- Las imágenes Planet del 2019 al 2022 son de soporte para revisar y corroborar la información de los puntos. Se tiene precaución del desplazamiento y revisar el último trimestre de cada año (octubre a diciembre) para tener certeza de que clase se evalúa.
- La información de mediana de otros años se utilizaras para corroborar las clases (esta información es útil para evaluar regeneración).
- Es importante considerar la clase sin información y evaluarla correctamente, si en los compuestos de Mediana y Last no se tiene información, pero en las imágenes Planet Scope si, el punto se evalúa dependiendo el tipo de cobertura presente en dichas imágenes. Las áreas que no se pueda demostrar información con los insumos otorgados en uno de los años evaluados, se categorizará en la clase sin información
- En los casos que los mosaicos LANDSAT media y last no tuvieran información ya sea por efectos como las nubes o sombras o simplemente porque el satélite no capturo información, se realizará la verificación de las imágenes individuales empezando por el último mes del año (diciembre). De no tener información en el mes correspondiente, se verificará el punto con el mes anterior, es decir, que si se estaba verificando un punto en el mes de diciembre y este no tenía información se verificará con el mes de noviembre.

3.4.1 Consolidación y control de calidad

Una vez cada interprete termine la fase de validación visual, se realiza la integración de los resultados individuales con el fin de identificar las coincidencias y desacuerdos que existan entre los puntos evaluados. Para ello se genera la matriz de confusión preliminar que permite determinar las clases con mayor confusión, y así realizar el control de calidad y los ajustes necesarios de la interpretación del muestreo. Para ello se tomaron en cuenta los siguientes parámetros:

- Si los tres interpretes coinciden en su interpretación, se deja el valor asignado de referencia por los mismos.
- En caso de existir coincidencia en al menos dos de las tres interpretaciones, se evalúa junto con el líder de la evaluación la diferencia y se llega a un consenso.
- En caso de existir diferencias en la interpretación de los tres intérpretes, se realizó una revisión de los puntos junto con el coordinador y se llegó a un consenso.
- Finalmente, con el equipo del SMByC se realizó la revisión final de los puntos de omisión y comisión y se llegó a un consenso sobre la validación de los mismos.

3.5 Generación de las estadísticas de validación

Una vez llegado al consenso final con todos los puntos, se elaboró nuevamente una matriz de confusión final. Esta es una matriz cuadrada en la que se compara la clasificación de la capa evaluada obtenida por el grupo evaluador con la elaborada por el productor, en este caso el SMCyB. A través de la matriz de confusión se evalúa la exactitud de la clasificación, situando en las filas las clases o categorías del mapa y en las columnas las mismas clases para la verdad de terreno. Esta matriz presenta una visión general de las asignaciones, tanto de las clasificaciones correctas (elementos de la diagonal) como de las migraciones o fugas (elementos fuera de la diagonal), así mismo, recoge los errores de omisión y de comisión (*Errores de comisión: elementos que no perteneciendo a una clase aparecen en ella y Errores de omisión: elementos que perteneciendo a esa clase no aparecen en ella por estar erróneamente incluidos en otra*) (Boca & Rodríguez, 2012).

Respecto a las medidas utilizadas para expresar la exactitud del mapa se emplea:

- i) Exactitud del productor (\hat{P}_j) calculada como la razón entre el número de unidades muestreadas que fueron correctamente clasificadas (diagonal) y el número de unidades que pertenecen a esa categoría (campo) o sea que no fueron asignadas en el mapa a la clase correspondiente. Se utiliza para calcular los errores de omisión.
- ii) Exactitud del usuario (\hat{U}_j) se calcula como razón entre el número de unidades muestreadas que fueron correctamente clasificadas (diagonal) y el número total de unidades que forman esa clase en el mapa. O sea, el número de clases asignadas a una categoría en el mapa pero que no pertenecen a ella. Se utiliza para calcular los errores de comisión (Boca & Rodríguez, 2012).
- iii) Exactitud general del mapa $\hat{O} = \sum_{j=1}^q \hat{p}_{jj}$ donde \hat{p} es un estimador insesgado de la proporción de superficie en la celda i, j , de la matriz de error y se calcula como $\hat{p}_{jj} = W_i \frac{n_{ij}}{n_i}$

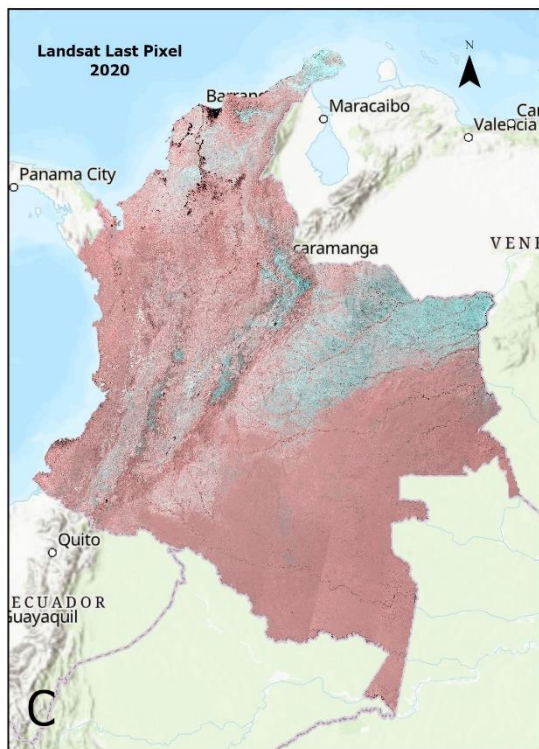
Estas medidas se utilizan teniendo en cuenta la proporción de la superficie de cada clase para la estimación de la matriz de error. Sin embargo, igualmente y como establece (Ramírez, Ramírez, Galindo, & Espejo, 2014), se realizó la estimación de las medidas de fiabilidad de manera directa a partir de matriz tradicional de recuentos.

4. RECOPIACIÓN DE LA INFORMACIÓN

Para la validación de los mapas temáticos, se utilizó la información satelital generada por el SMBYC y los productos individuales de Landsat sin necesidad de generar ningún tipo de pre-procesamiento de las mismas. En el momento de la validación a los mosaicos de imágenes Landsat y Sentinel se les realizó un estiramiento de corte acumulativo utilizando para ello una extensión y límites predeterminados. Esta técnica permitió mejorar la calidad de la imagen resaltando las características locales que la hicieron más visible al contrastar localmente detalles.

La información satelital empleada para la validación de los productos fue:

- **Landsat:** Los cuatro mosaicos de los productos Last y mediana que corresponden a Last_pixel2019_22042019; Landsat_2019_Mediana_06042020; Landsat_2020_last_pixel_07042021 y Landsat_2020_Mediana_07042021 que permitieron evaluar el mapa de cambio de superficie del periodo 2019 – 2020 (**¡Error! No se encuentra el origen de la referencia.** A,B,C,D) y los dos mosaicos de las imágenes Last_pixel_2022_160223 y Landsat_2022_Mediana_V2_09032023 para evaluar el mapa de bosque no bosque del 2022 (**¡Error! No se encuentra el origen de la referencia.** E,F). Estos productos fueron generados por el Sistema de Monitoreo de Bosques y Carbono SMBYC los cuales fueron construidos a partir de compuestos multianuales de imágenes Landsat a las que previamente se les ha aplicado procesos de calibración y corrección radiométrica, eliminando los píxeles contaminados atmosféricamente. Los productos Landsat Mediana, corresponde al valor promedio de los valores de reflectancia que se encuentran entre el percentil 25 y 75 del compuesto, generado para cada píxel de cada banda espectral. Por su parte, los productos Landsat Last pixel, corresponde a la mediana de los tres últimos valores sin nubes de cada banda obtenida del compuesto de imágenes multianual. Los productos incluyen cuatro bandas espectrales de las imágenes Landsat, a saber: Rojo (0.626-0.693 μ m), Infrarrojo cercano- NIR (0.776- 0.904 μ m), Infrarrojo de onda corta SWIR1 (1.567-1.784 μ m) e infrarrojo de onda media SWIR2 (2.097-2.349 μ m).



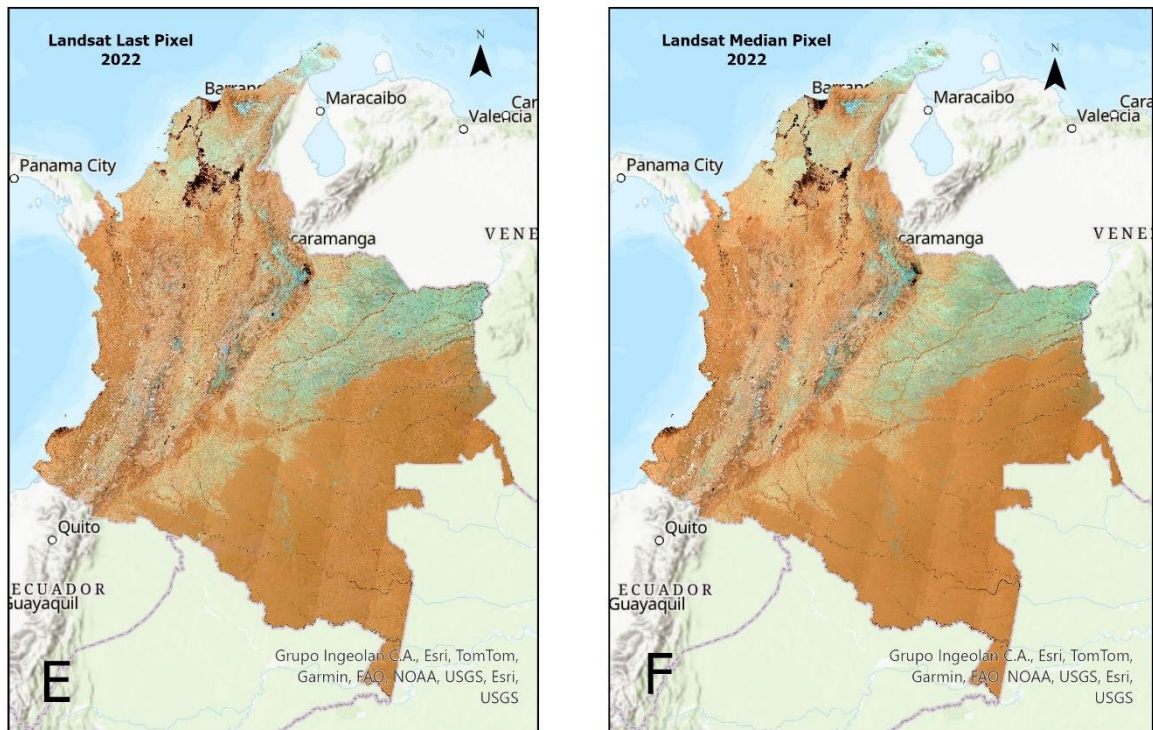
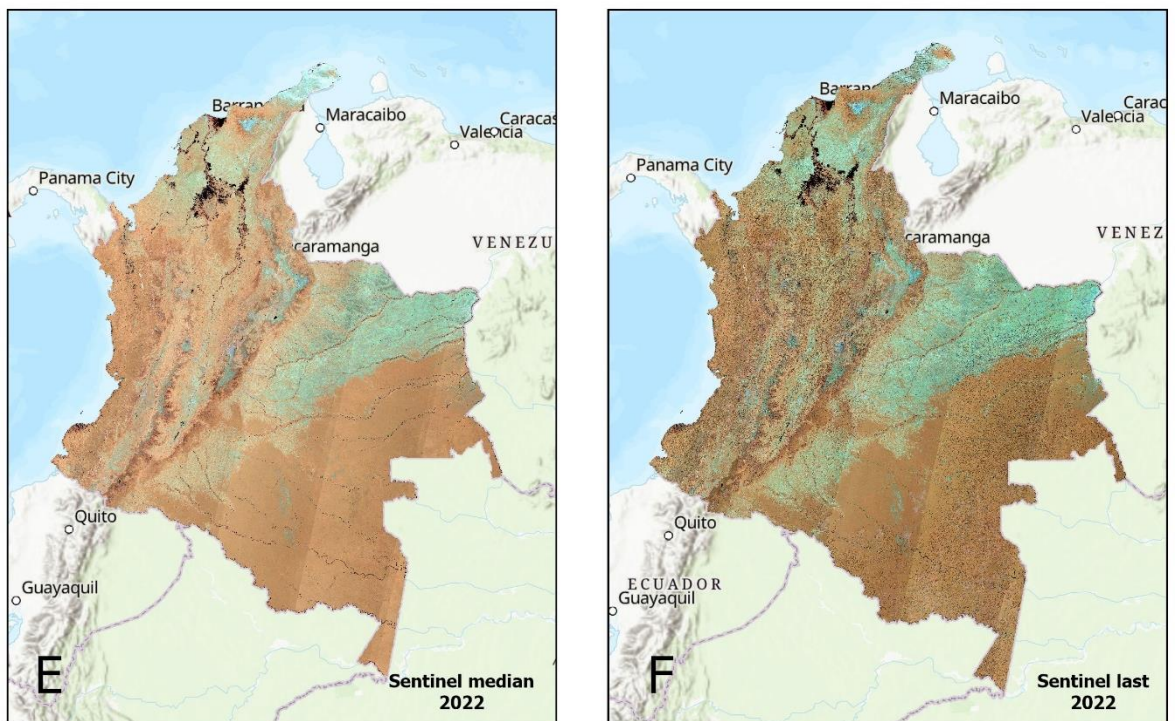


Figura 2. Mosaicos de los productos Landsat last y median: a) Landsat 2019_22042019; b). Landsat_2019_Mediana_06042020; c). Landsat 2020_Last_pixel_07042021 y d). Landsat 2020_Mediana_07042021 e) Last_pixel_2022_160223 y f). Landsat_2022_Mediana_V2_09032023

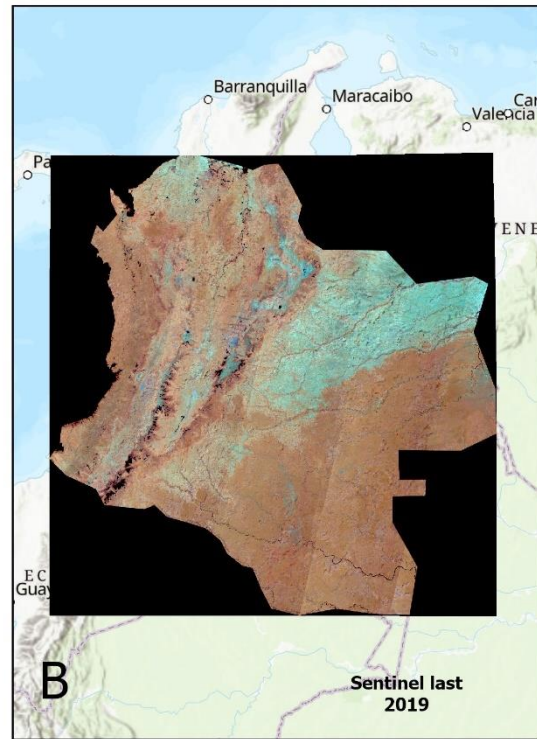
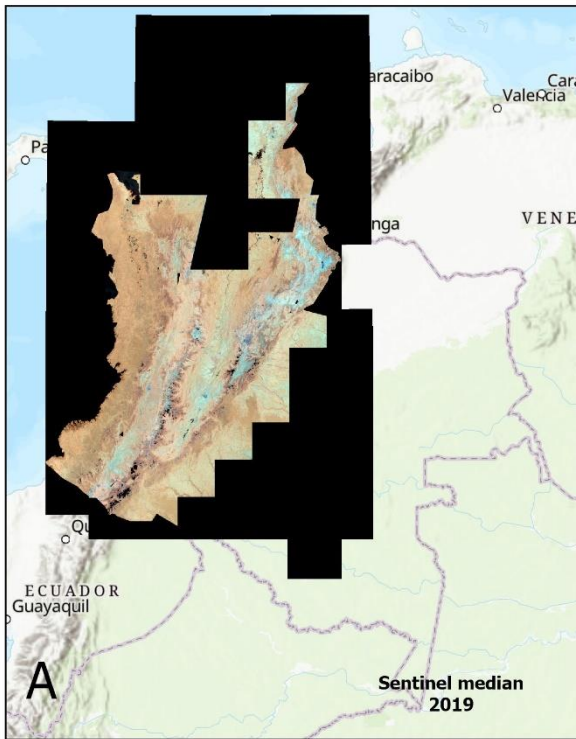
Sentinel: Los cuatro mosaicos last y mediana de sentinel_2019_LastPixel_Col_z18; sentinel_media_2019_sininfor_Landsat; Last_piexel_sentinel_2020_050321 y Mediana_sentinel_2020_utm; que permitieron evaluar el mapa de cambio de superficie del periodo 2019 – 2020 (



A,B,C,D(Error! No se encuentra el origen de la referencia.) **y los dos compuestos de last_sentinel_2022_230222_utm_z18 y Mediana_sentinel_2022_230220_utm_Z18N 09032023 para evaluar el mapa de bosque no bosque del 2020 (**



➤ , E,F) Al igual que las imágenes Landsat, estos mosaicos fueron generados por el Sistema de Monitoreo de Bosques y Carbono SMBYC los cuales fueron construidos a partir de compuestos multianuales de imágenes a las que previamente se les ha aplicado procesos de calibración y corrección radiométrica, eliminando los pixeles contaminados atmosféricamente.



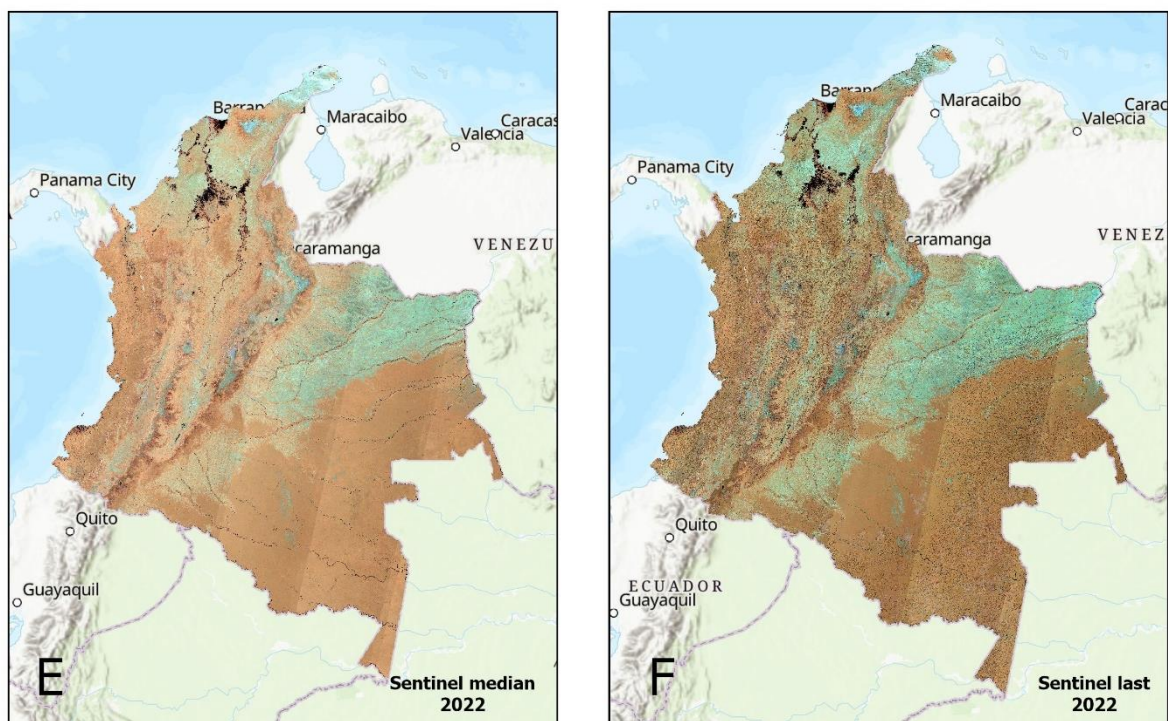


Figura 3. Mosaicos de los productos Sentinel A). `sentinel_2019_LastPixel_Col_z18`; B) `sentinel_media_2019_sininfor`; C). `Last_pixel_sentinel_2020_050321`; D). `Mediana_sentinel_2020_utm`; E). `Mediana_sentinel_2022_230220_utm_Z18N_09032023` y F) `Last sentinel-xxxx`

- **Planet:** El servicio WMTS de las que permite filtrar, localizar y visualizar imágenes de planet, https://api.planet.com/basemaps/v1/mosaics/wmts?api_key=06da105ff5cd493ab05068a0e8a99810, los cuales fueron desplegados para los periodos de análisis.
- **ESRI, BING, GOOGLE:** El plugin de QGIS QuickMapServices: que permite el acceso a información de imágenes satelitales que provienen de servicios Web, mediante protocolos WMS y WMTS generados dentro de Bing, ESRI, Google, OpenStreetMap entre otros.
- **Landsat Individuales:** Las imágenes Landsat individuales de toda el área continental de Colombia, que corresponden a los últimos tres meses del año 2019 a 2022. Las imágenes tienen seis bandas espectrales de las imágenes Landsat, a saber: Azul (0.45-0.52 μm), Verde (0.52-0.60 μm), y Rojo (0.626-0.693 μm), Infrarrojo cercano- NIR (0.776- 0.904 μm), Infrarrojo de onda corta SWIR1 (1.567-1.784 μm) e infrarrojo de onda media SWIR2 (2.097-2.349 μm).
- **Sentinel y Landsat 2019 a 2021:** Los compuestos de las imágenes Landsat `2021_Last_pixel_v3_11052022` y `Landsat_2021_Media_V3_11052022`, `Sentinel_last_pixel_2021_270422_UTMZ18` y `Sentinel_mediana_2021_270422_UTMZ18`. Las imágenes del 2019 a 2020 fueron apoyo en casos que no se tenía certeza de la evolución de la cobertura.

5. RESULTADOS: MAPA DE BOSQUE – NO BOSQUE, PERIODO 2022

El mapa objeto de evaluación fue elaborado por el Sistema de Monitoreo de Carbono y Bosque SMCyB del IDEAM para el periodo 2022 (Figura 4), el cual se encuentra en el sistema de coordenadas geodésicas WGS84, UTM, zona 18N EPSG 32618. Se observa que el área Continental de Colombia, tiene una extensión de 114.062.089 ha, y a 2022 el país contaba con un área total de 59.290.764 ha de bosque natural, equivalente a 51,98%. Por su parte, un total de 54.694.6765 ha de superficie continental presenta coberturas no boscosas, equivalente al 47,95% del territorio. Para este periodo se registró un total de 76.649 ha en áreas sin información las cuales no pudieron ser evaluadas.

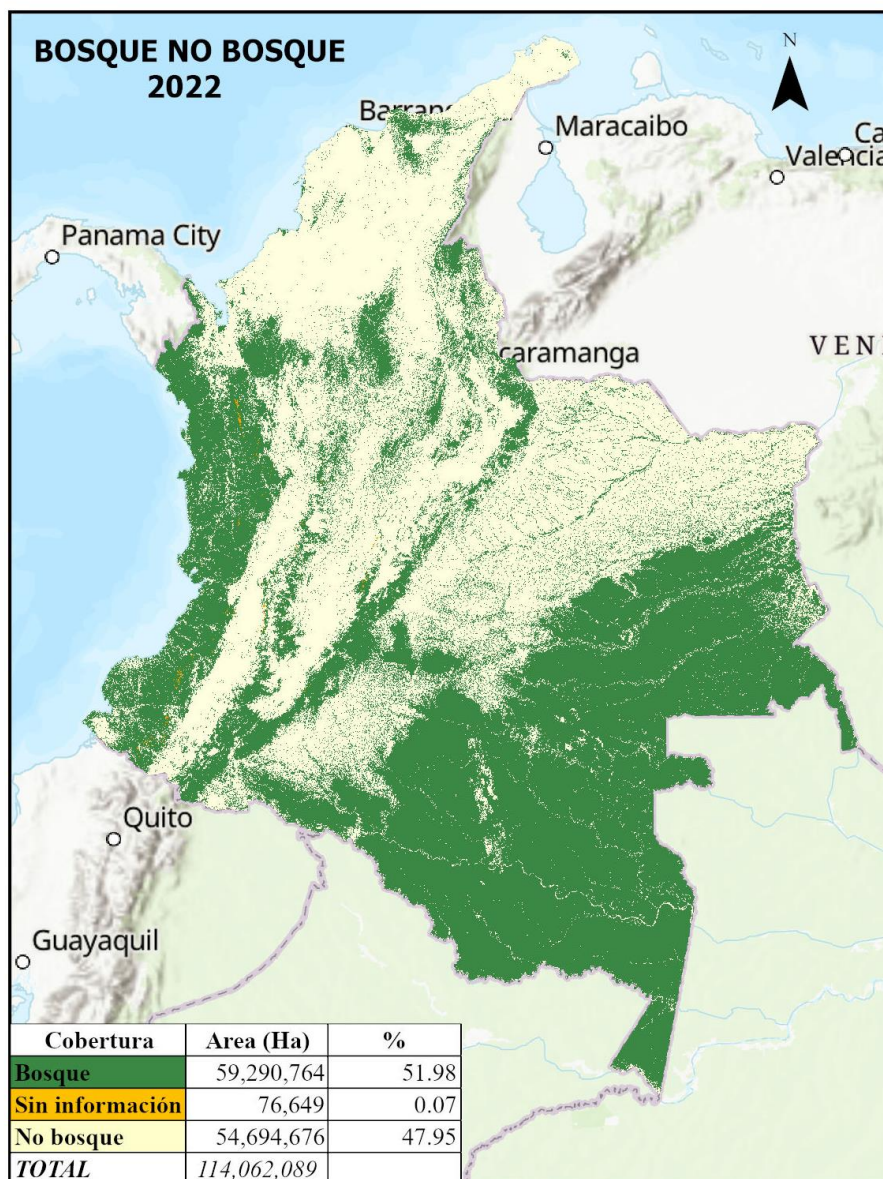


Figura 4. Mapa de cobertura de bosque / no bosque del 2022.

5.1. Generación de la muestra y el proyecto

La generación de la muestra del mapa temático del 2022 se realizó en 3 fases, cuyos resultados se describen a continuación:

5.1.1. Definición del porcentaje de exactitud temática

Para la estimación del tamaño de muestra, se siguió los parámetros establecidos en la metodología, asumiendo un error estándar esperado para todo el mapa de 0,003. Igualmente, y acorde a (Olofsson, Foody, & Herold, 2014), que han mostrado que los errores de comisión para las clases de cambio son más altos que los cometidos en las clases estables, se asumió una exactitud de usuario para la clase de cambio correspondiente a sin información de 0,8, en las categorías de riesgo alto y bajo; mientras que para las clases de bosque y no bosque estable correspondió a 0,90 (Tabla 1).

Tabla 1. Calculo tamaño de muestra, para el mapa de bosque no bosque del 2022

CLASES	Área (Píxeles)	Área (ha)	Wi	Ui	SI
1 Bosque	658.786.272	59.342.870	0.51981131	0.9	0.3
2. Sin Información	851.657	76.649	0.00067199	0.8	0.4
3. No Bosque Estable	607.718.619	54.696.676	0.47951669	0.9	0.3
Total	1.267.356.548	114.062.089			

5.1.2. Definición del tamaño de la muestra

Una vez discutida y concertada con el equipo técnico del SMByC y, con ayuda de la herramienta desarrollada en el Sistema de Monitoreo de Bosques y Carbono SMCYC ACATAMA (Accuracy Assessment of Thematic Maps), se inició la generación de muestras acorde a los parámetros estadísticos establecidos dentro de la metodología. De esta manera, se realizó el proceso de evaluación de la exactitud temática de las coberturas bosque estable, no bosque estable, y sin información. De esta manera, se generó la distribución aleatoria de los sitios de verificación, una vez se efectuó la corrida bajo los parámetros establecidos en el marco metodológico (Figura 5).

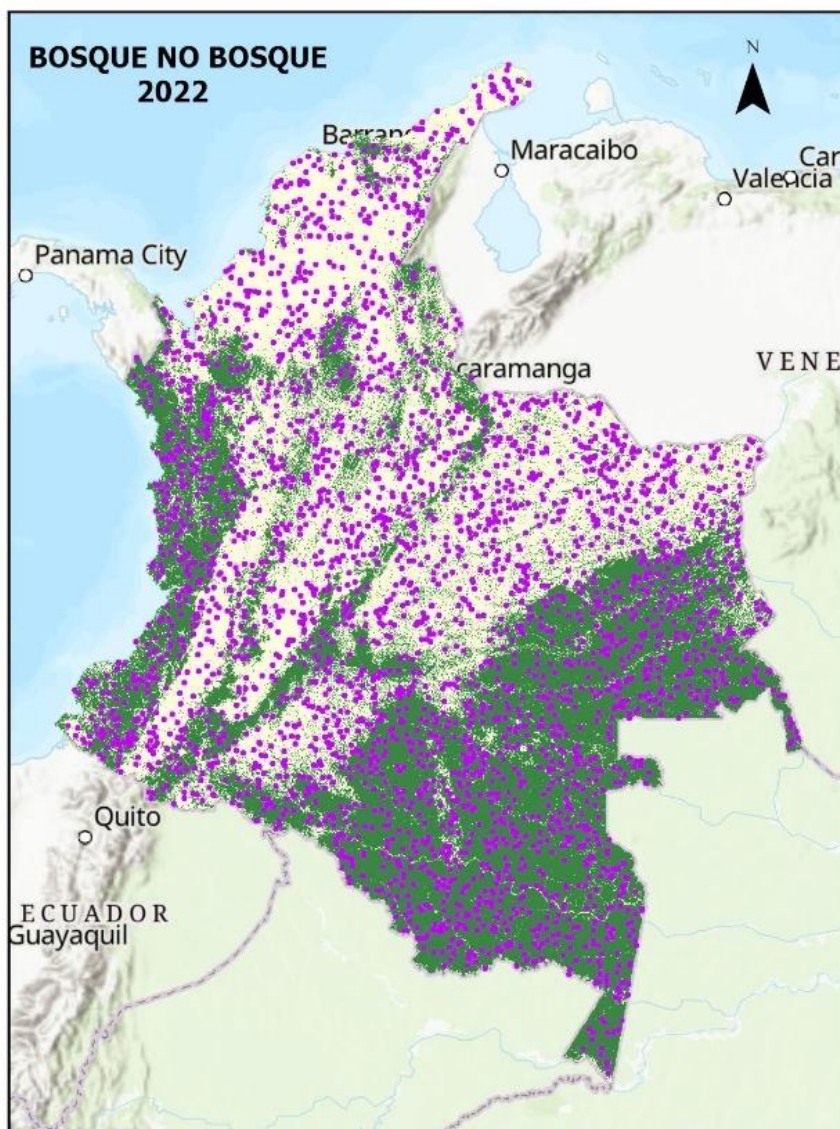


Figura 5. Puntos de muestreo a validar para el mapa de bosque no bosque del 2022.

De esta manera se generó un total de 3.085 puntos los cuales se distribuyen en: 1.650 en bosque, 1.385 en bosque no estable y 50 en sin información.

5.1.3. Generación del proyecto

Para realizar la validación se construyó el proyecto dentro de software QGIS el cual fue almacenado dentro de la carpeta ..\Puntos_Bnb_2022\Proyecto_bnb.qgz. Dentro del proyecto se cargaron las imágenes de referencia de los mosaicos de Landsat y sentinel del 2021 y 2022, así como las imágenes disponibles en la plataforma de Planet, las imágenes Google satélite, Bing satélite, ESRI y algunas capas de información temática disponible como apoyo para la interpretación (Figura 6).

El proyecto fue compartido a cada uno de los intérpretes con la finalidad de cada uno de los mismos partiera de la misma información y con ello realizar la validación visual individual de los puntos seleccionados.

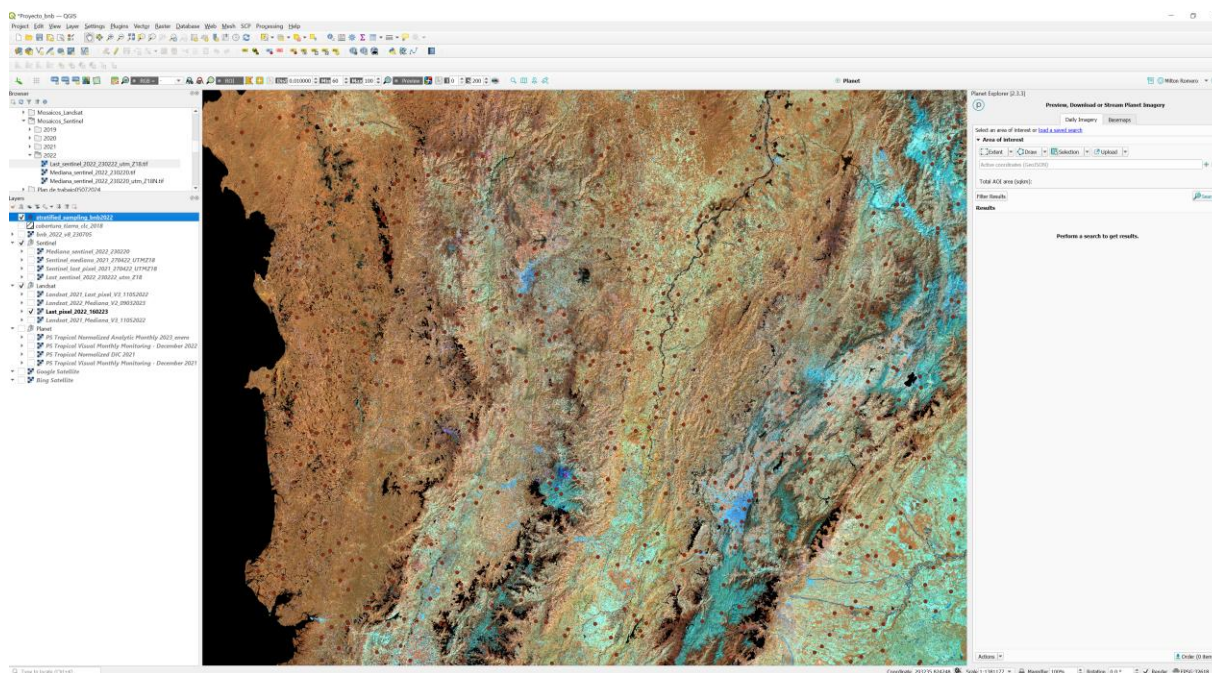


Figura 6. Proyecto QGIS, base para la validación del mapa de bosque no bosque 2022

5.2. Interpretación de la muestra

Una vez obtenidos los puntos a validar se inició el proceso de interpretación de los mismos los cuales se realizan en los siguientes pasos:

5.2.1. Conformación del equipo de validación

El equipo encargado de la validación de los 3.085 puntos realizó la interpretación visual individualmente. Posteriormente, el líder del grupo quien permanentemente estuvo evaluando los avances de la interpretación, brindó los lineamientos para llegar a la evaluación final del producto, para obtener un resultado acorde a la metodología planteada anteriormente.

5.2.2. Control de calidad de la interpretación

La capa de los 3.085 puntos de muestreo, junto con el proyecto inicial fue entregada a cada uno de los intérpretes de manera independiente de esta manera, cada uno de los intérpretes con la utilización de la herramienta de QGIS ACATAMA, y definiendo una escala de referencia de 200, la cual se fue modificando acorde a la dificultad de la interpretación.

Una vez finalizada la interpretación, se procedió a generar el control de calidad de la interpretación realizada por los tres intérpretes para: identificar inconsistencias sistemáticas en los criterios de interpretación, identificar áreas con dificultad de interpretación y en caso de ser necesario redefinir clases. Las diferencias en la interpretación de las clases con mayor confusión según la matriz se dieron por: no acuerdo entre los intérpretes y áreas sin información en cualquiera de los dos periodos de evaluación. La Tabla 2 muestra el número de puntos y el grado de acuerdo a los que se llegó en la primera interpretación.

Tabla 2. Número de puntos de acuerdo entre los intérpretes.

No. Puntos	Grado de acuerdo
2.979	Totalmente de acuerdo (3 intérpretes)
106	Parcialmente de acuerdo (2 intérpretes)

Para este mapa, los 106 puntos de no coincidencia, se evidencio tan solo dos intérpretes estuvieron de acuerdo en la asignación de la clase de referencia. Por otra parte, las 2.979 muestras restantes, existió acuerdo total de los intérpretes en la asignación de las clases de referencia.

La Figura 7 muestra la distribución de los puntos no coincidentes, los cuales son representados en color negro para aquellos donde la coincidencia fue para dos intérpretes y en color rojo los que no existió ninguna coincidencia entre los tres intérpretes.

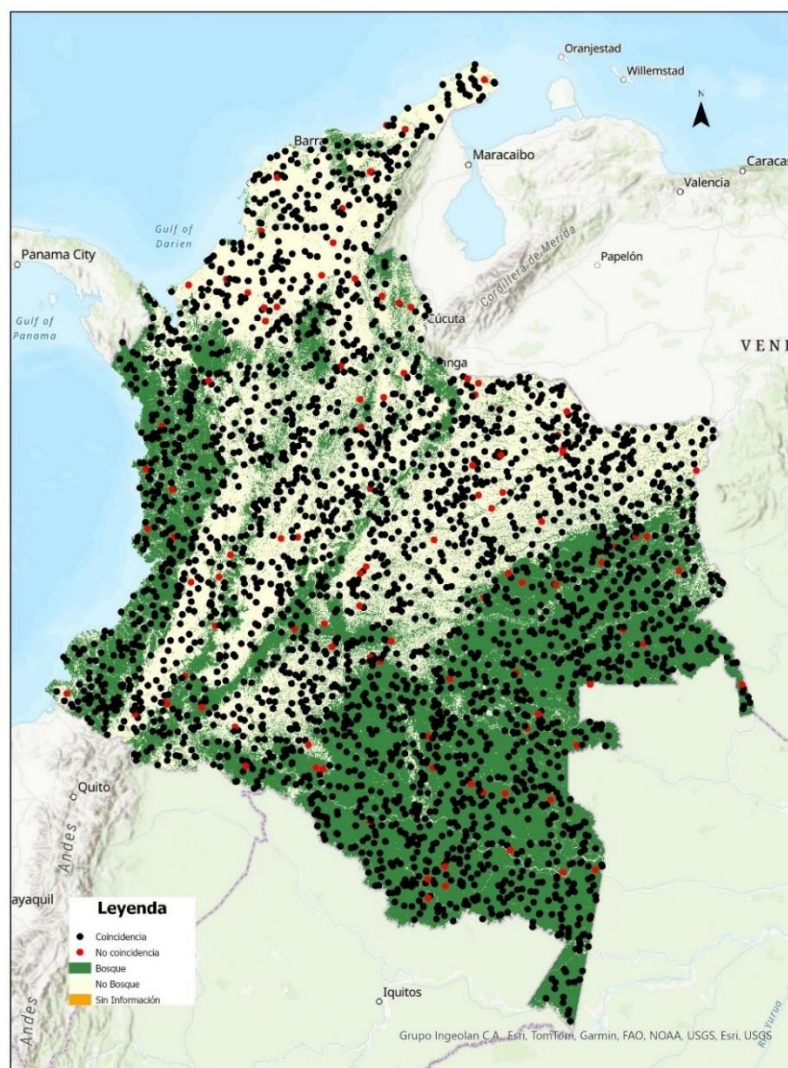


Figura 7. Distribución de los puntos de no coincidencia entre los intérpretes (puntos negros: coincidencias, puntos rojos: sin coincidencias).

Los 106 puntos en que no existió acuerdo, fueron revisados entre el equipo y el coordinador del proyecto, llegando a un consenso de asignación para la clase. Aquellos puntos donde no se llegó a un acuerdo fueron igualmente revisados y discutidos con el equipo del SMyC, llegando un acuerdo sobre la clase a la que pertenecía.

5.2.3. Generación de las estadísticas de validación (Matriz de error, cálculos de exactitud, estimadores de área ajustada, incertidumbre)

Una vez realizado el control de calidad, se generó la **matriz de error** o confusión convencional y la de proporción por área, así como las medidas de fiabilidad acorde al protocolo establecido por (Galindo, Espejo, Rubiano, Vergara, & Cabrera, 2014), tomando para la evaluación de la exactitud temática el 100% de las unidades de la muestra (3.085). A partir de los datos que se observan en la Tabla 3, del total de los 3.085 puntos analizados un total de 3.037 fueron clasificados acorde al valor asignado de referencia. Estos corresponden a 1.620 puntos en la categoría de bosque, 1.367 a la categoría no bosque y 50 a la categoría sin información, la cual corresponde a la diagonal principal de la Tabla 3, la cual se resalta en verde.

En total las no coincidencias corresponden a 48 puntos, estos fueron revisados en QGIS ACATAMA, analizando el contexto general y detallando el punto para tener una visión amplia del territorio, para esto se realizó mejoramiento del histograma y tener una mejor visualización.

Tabla 3. Matriz de error tradicional y medidas de fiabilidad

CLASES	Bosque (1)	No bosque (2)	Sin información (3)	Total	Exactitud Usuario	Wi
Bosque (1)	1.620	29	1	1.650	0.9818	0.51981
No bosque (2)	18	1.367	0	1.385	0.987	0.47952
Sin información (3)	0	0	50	50	1	0.00067
total	1.638	1.396	51	3.085		
Exactitud productor	0.9879	0.9811	0.6808		0.9843	

Comisión: Del total de 1.650 puntos clasificados en bosque dentro del mapa temático evaluado, se encontró que 1.620 puntos estaban bien clasificados, 29 correspondían a no bosque y 1 a sin información. Para los 1.385 puntos clasificados en no bosque dentro del mapa temático evaluado, se encontró que 1.367 estaban bien clasificados y 18 correspondían a bosque. Finalmente, los 50 puntos clasificados en sin información fueron bien clasificados.

Omisión: De lo anterior inferimos, que en la clase bosque faltaron 18 puntos que fueron clasificados en no bosque dentro del mapa temático. Respecto a no bosque se identificó 29 puntos en omisión, que fueron clasificados dentro del mapa temático como bosque; y finalmente, se detectó que 1 punto que le pertenece a la clase sin información fue clasificado como bosque.

La **exactitud** global del mapa es de **98.43%**, la cual está dada como el cociente entre la sumatoria de puntos de la diagonal principal sobre el total de los puntos evaluados.

En cuanto a la exactitud total del usuario fue del **98.01%** en donde para la clase bosque fue de 98.2%; no bosque 98.7% y sin información 100%.

En cuanto a la exactitud total del productor esta fue de **88.3%** en donde para la clase bosque fue de 98.8%, no bosque 98.1% y sin información 68%.

Al contrastar los resultados entre las exactitudes del productor versus las del usuario asociada a los resultados obtenidos a partir de la proporción de área Tabla 4, se observa que la exactitud del productor para la clase, sin información fue sobreestimada.

Tabla 4. Matriz de error de la proporción de área y medidas de fiabilidad

CLASES	Bosque (1)	No bosque (2)	Sin información (3)	Usuario	Productor
Bosque (1)	0.51036	0.0914	0.00032	0.982	0.988
No Bosque (2)	0.00623	0.47328	-	0.987	0.981
Sin información (3)	-	-	0.00067	1.000	0.681
total	0.51659	0.48242	0.00099	0.984	0.883

Junto con esto, para cada clase se determinó el **área estimada ajustada** observando un error aproximado de 256.481 ha para bosque 254.041 para no bosque estable y 35.934 para sin información (Tabla 5).

Tabla 5. Intervalos de confianza del área estimada por clase.

CLASES	Área (ha)	Error	Límite inferior	Límite superior	Incertidumbre
1 (Bosque)	58,923.583,910	256.481	58,420.881	59,426,286	0.00853
2 (No bosque)	55.025.922.482	254.041	54.528.003	55.523,841	0.00905
3 (Sin Información)	112.582.926	35.934	42.153	183.013	0.62559
Total	114,062,089.32				

Por otra parte, se encontró que la clase bosque tiene una incertidumbre del 0.8%; el no bosque 0.9% y sin información del 62.3%. Finalmente, se obtuvo una mejor exactitud general del mapa de cambio, en función de la proporción de área de 98.4% este valor no presenta una diferencia significativa con respecto al obtenido mediante la matriz de error tradicional (96,9%) que se estima a partir del índice de Kappa.

6. RESULTADOS MAPA DE CAMBIO DE SUPERFICIE DE BOSQUE – NO BOSQUE, PERIODO 2019-2020

El segundo mapa objeto de evaluación fue el mapa de cambio de superficie de bosque -no bosque del periodo 2019-2020 elaborado por el Sistema de Monitoreo de Carbono y Bosque SMCyB del IDEAM (Figura 8), el cual se encuentra en el sistema de coordenadas geodésicas WGS84, UTM, zona 18N EPSG 32618.

6.1. Selección de área a evaluar

Para la obtención del área a evaluar y posterior generación de puntos de muestreo del mapa temático de cambio de superficie de bosque-no bosque 2019-2020 se realizaron 2 actividades, cuyos resultados se describen a continuación:

6.1.1. Selección de la capa temática a evaluar

Se observa que el área Continental de Colombia, tiene una extensión de 114.062.089 ha, y a 2022 el país contaba con un área total de 59.290.764 ha de bosque natural, equivalentes a 51,98%. Por su parte, un total de 54.694.6765 ha de superficie continental presenta coberturas no boscosas, equivalente al 47,95% del territorio. Para este periodo se registró un total de 76.649 ha en áreas sin información las cuales no pudieron ser evaluadas.

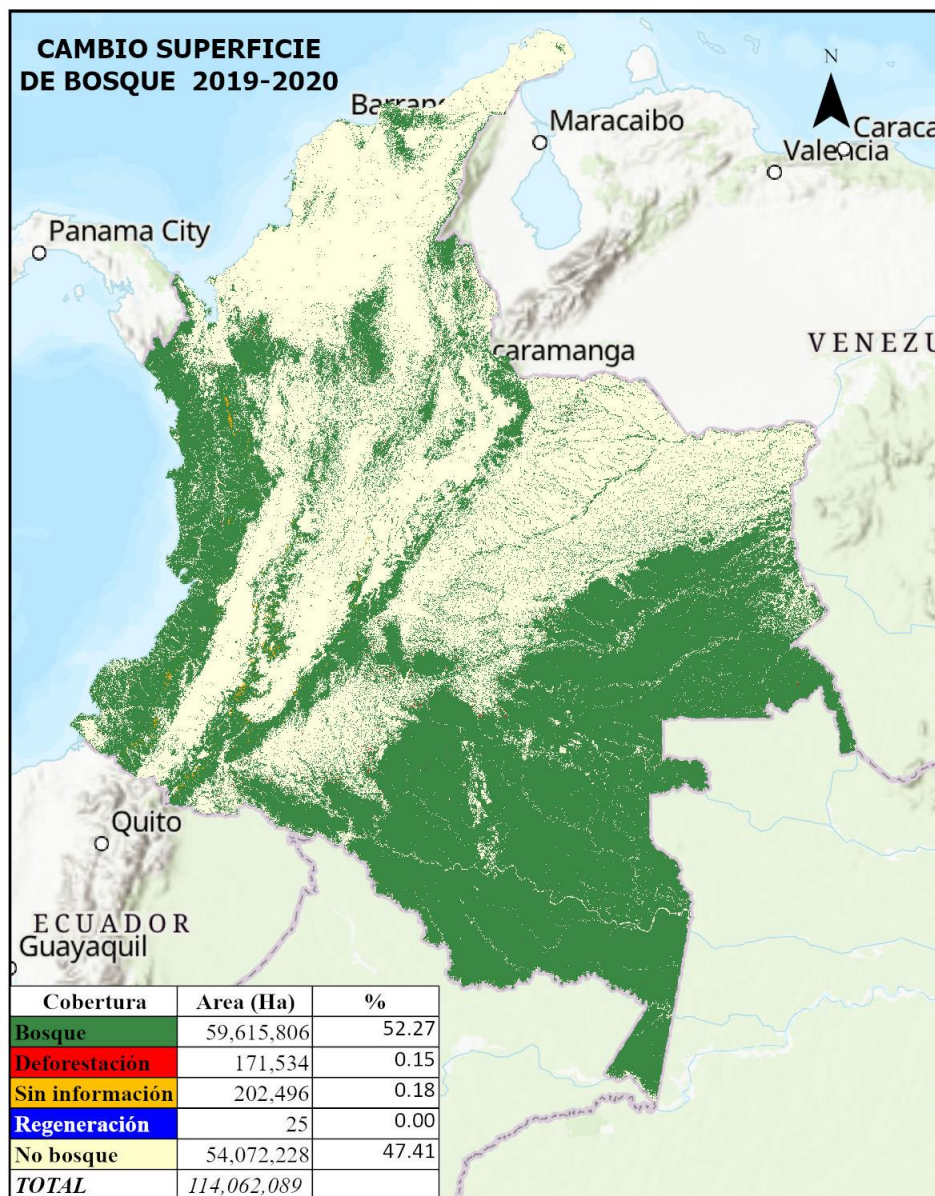


Figura 8. Mapa de cambio de superficie de bosque – no bosque del 2019-2020.

6.1.2. Generación del mapa de riesgo

Antes de iniciar el proceso de generación de muestras del mapa de cambio de superficie de bosque – no bosque, se elaboró el mapa de riesgo a la deforestación partiendo con la identificación del área de deforestación donde se concentra la deforestación en el país, con ayuda de la información histórica de los mapas de bosque – no bosque generando un total de 10 estratos para el muestro y la validación del mapa de cambio de la superficie que corresponde a: (1) Bosque estable en zonas de alto riesgo de deforestación BE-RA, (2) No bosque estable en zonas de alto riesgo de deforestación NB-RA, (3) Deforestación en zonas de alto riesgo de deforestación DEF-RA, (4) Bosque estable en áreas de bajo riesgo de deforestación BE-RB, (5) No Bosque estable en áreas de bajo riesgo de deforestación NB-RB y (6) Deforestación en áreas de bajo riesgo de deforestación DEF-RB (Figura 9).



Figura 9. Mapa de riesgo a la deforestación

6.2. Generación y diseño de la muestra

Los resultados de la generación y diseño de la muestra que involucra la definición del porcentaje de exactitud temática, el tamaño de la muestra y la generación del proyecto se describen a continuación:

6.2.1. Definición de porcentajes de exactitud temática:

Para la estimación del tamaño de muestra, se siguió los parámetros establecidos en la metodología, se asumió un error estándar esperado para todo el mapa de 0,003. Igualmente, y acorde a (Olofsson, Foody, & Herold, 2014), que han mostrado que los errores de comisión para las clases de cambio son más altos que los cometidos en las clases estables, se asumió una exactitud de usuario para la clase de cambio correspondiente a sin información de

0,8, en las categorías de riesgo alto y bajo; mientras que para las clases de bosque y no bosque estable correspondió a 0,90. Finalizada esta etapa, se discutió, concertó y aceptó con el equipo técnico del SMByC del IDEAM (Tabla 6Tabla 1).

Tabla 6. Calculo tamaño de muestra, para el mapa de cambio superficie bosque no bosque del 2019-2020

CLASES	Área (Píxeles)	Área (ha)	Wi	Ui	SI
1 Bosque riesgo bajo	412.230.093	37.100.708	0.32526766	0.9	0.3
2 Bosque riesgo alto	250.167.750	22.515.098	0.19739335	0.9	0.3
2. Deforestación riesgo bajo	91.710	8.254	0.00007236	0.8	0.4
2. Deforestación riesgo alto	1.814.225	163.280	0.00143150	0.8	0.4
3. Sin información riesgo bajo	1.801.516	162.136	0.00142148	0.8	0.4
3. Sin información riesgo alto	448.445	40.360	0.00035384	0.8	0.4
4. Regeneración riesgo bajo	15	1	0.00000001	0.8	0.4
4. Regeneración riesgo alto	262	24	0.00000021	0.8	0.4
5. No bosque estable riesgo bajo	334.540.083	30.108.607	0.26396682	0.9	0.3
5. No bosque estable riesgo alto	266.262.449	23.939.620	0.21009277	0.9	0.3
Total	1.267.356.548	114.062.089			

6.2.2. Definición de tamaño de la muestra:

Una vez discutida y concertada con el equipo técnico del SMByC y, con ayuda de la herramienta desarrollada en el Sistema de Monitoreo de Bosques y Carbono SMCYC ACATAMA (Accuracy Assessment of Thematic Maps), se inició la generación de muestras acorde a los parámetros estadísticos establecidos dentro de la metodología. De esta manera, se realizó el proceso de evaluación de la exactitud temática de las coberturas bosque estable, no bosque estable, deforestación y sin información en las diferentes categorías de riesgo. De esta manera, se generó un total de 10.011 puntos con distribución aleatoria de los sitios de verificación, una vez se efectuó la corrida bajo los parámetros establecidos en el marco metodológico (Figura 10).

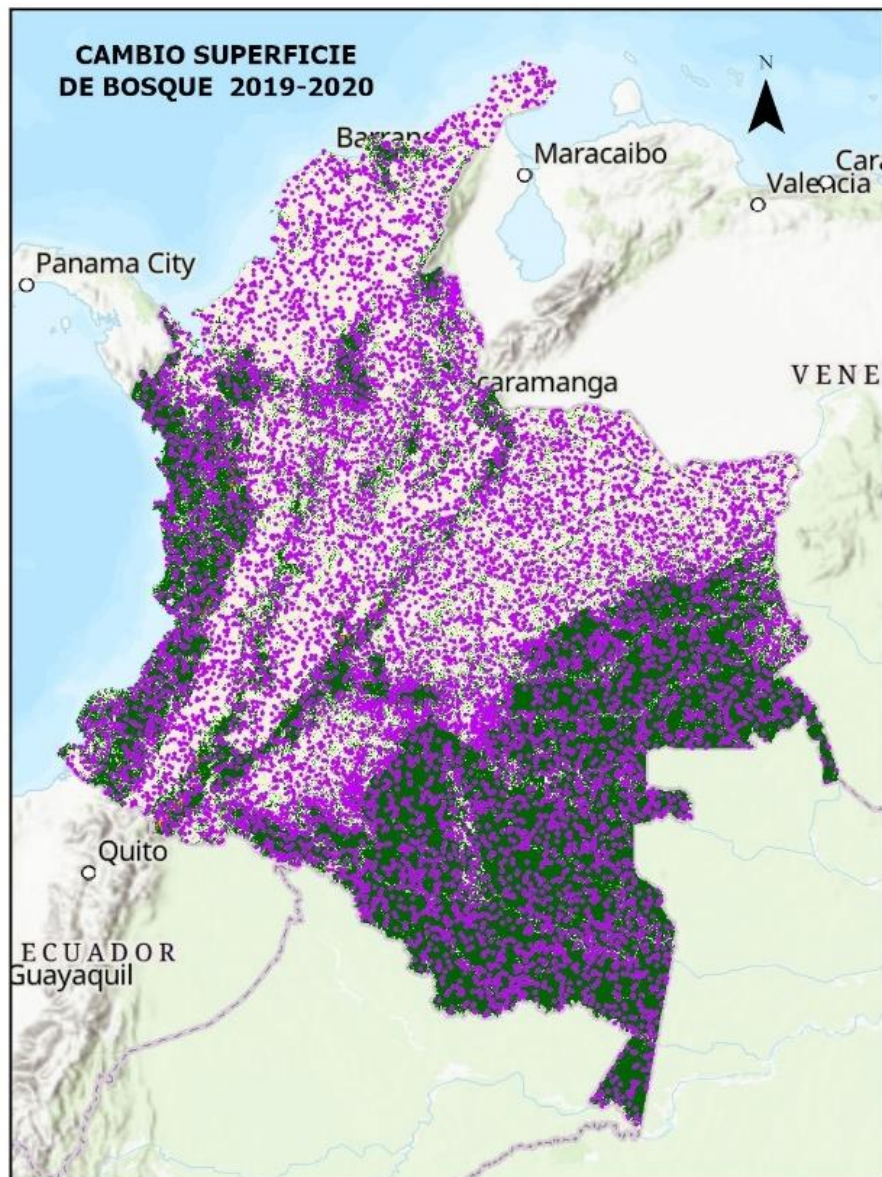


Figura 10. Puntos de muestreo a validar para el mapa de cambio de superficie de bosque 2019-2020.

6.2.3. Generación del proyecto

Para realizar la validación se construyó el proyecto dentro de software QGIS el cual fue almacenado dentro de la carpeta ..\Puntos_2019_2020\cambio_2019_2020.qgz. Dentro del proyecto se cargaron las imágenes de referencia de los mosaicos de Landsat y sentinel del 2019 y 2020, así como las imágenes disponibles en la plataforma de Planet, las imágenes Google satélite, Bing satélite, ESRI y algunas capas de información temática disponible como apoyo para la interpretación (Figura 11).

El proyecto fue compartido a cada uno de los intérpretes con la finalidad de cada uno de los mismos partiera de la misma información y con ello realizar la validación visual individual de cada uno de los puntos seleccionados. Cada

uno de los intérpretes igualmente tenía las imágenes individuales las cuales a su vez sirvieron de apoyo en el momento de tener dudas sobre alguna de las muestras a clasificar.

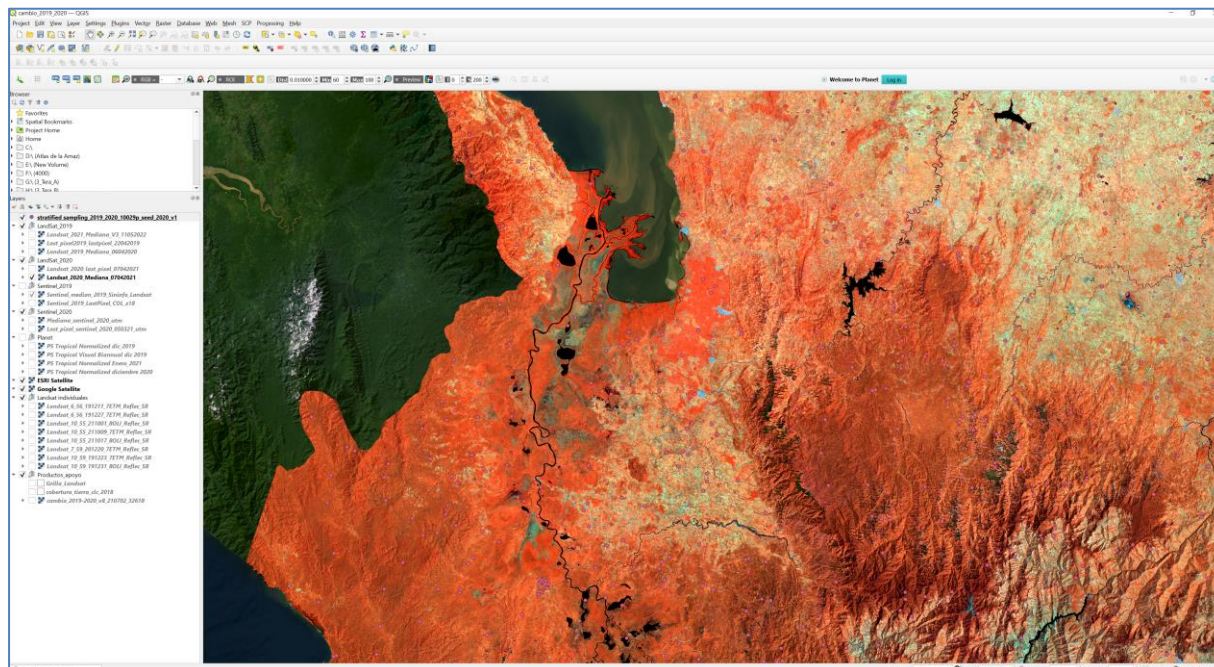


Figura 11. Proyecto QGIS, base para la validación del mapa de cambio en la superficie de bosque 2019-2020

6.3. Interpretación y control de calidad de los puntos

6.3.1. Conformación del equipo de validación

El equipo encargado de la validación del mapa de cambio de superficie de bosque 2019-2020 realizó la interpretación visual individualmente, junto con el líder del grupo quien permanentemente estuvo evaluando los avances de la interpretación y dando los lineamientos para llegar a la evaluación final del producto, para obtener un resultado acorde a la metodología planteada anteriormente. Este fue el mismo equipo que trabajo el mapa de bosque-no bosque del 2022.

6.3.2. Asignación de etiquetas de referencia

La capa de los 10.011 puntos de muestreo para este periodo, fue entregada a cada uno de los intérpretes de manera independiente para su interpretación visual, acorde a lo establecido en la metodología propuesta. Para ello, al igual que el mapa de bosque – no bosque, cada uno de los intérpretes con la utilización de la herramienta de QGIS ACATAMA, definiendo una escala de referencia de 200, la cual se fue modificando acorde a la dificultad de la interpretación fue realizando la interpretación de manera independiente.

Para la validación de la misma igualmente se utilizaron las ayudas de las herramientas disponibles en la ventana de diseño de respuesta de Acatama, Plugin de QGIS como son:

- **Visualizador Google Earth** integrado dentro de la ventana de diseño de respuesta en Acatama, el cual vincula a Google Earth y permite ver el histórico de imágenes de Google que ayudan a analizar las tendencias de cada uno de los puntos a lo largo de varios años (de ser disponibles estas imágenes, utilizando Google Imagen (Figura 12)

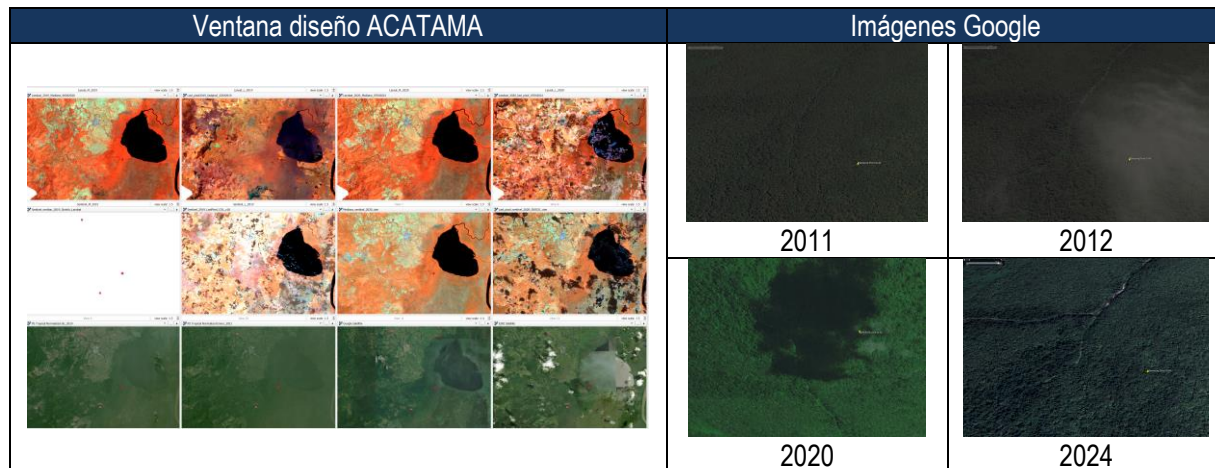


Figura 12. Vinculación de la ventana diseño de Acatama, con Google Earth

- **Integración del complemento de detección de cambios continuos (CCD)** ubicado dentro de la ventana de diseño de respuesta en Acatama. CCD ayuda a analizar las tendencias y los puntos de quiebre del cambio de las muestras a lo largo de varios años, utilizando Landsat y Sentinel a través de Google Earth Engine (Figura 13).

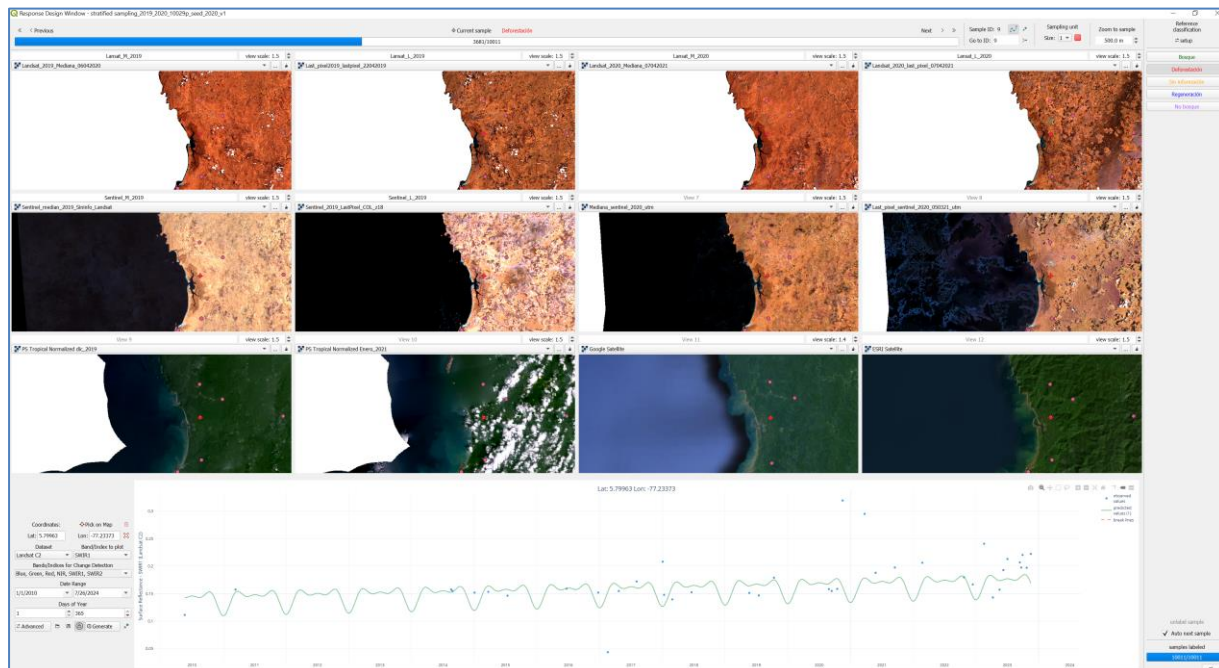


Figura 13. Vinculación de la ventana diseño de Acatama, con CCD

Igualmente, y como se mencionó anteriormente, se realizó un estiramiento de corte acumulativo utilizando para ello una extensión y límites predeterminados, así como la utilización de imágenes individuales de Landsat.

Una vez finalizada la interpretación individual se integró la información y se identificó las coincidencias y desacuerdos que existieron en esta validación. La Tabla 7 muestra el número de puntos y el grado de acuerdo y desacuerdo a los que se llegó en la primera interpretación.

Tabla 7. Número de puntos de acuerdo entre los intérpretes analizados para el periodo 2019-2020.

No. Puntos	Grado de acuerdo
4	Totalmente desacuerdo
9.552	Totalmente de acuerdo (3 interpretes)
455	Parcialmente de acuerdo (2 interpretes)

6.3.3. Control de calidad de la interpretación

En esta etapa se realizó una verificación de las asignaciones realizada por los tres intérpretes para: identificar inconsistencias sistemáticas en los criterios de interpretación, identificar áreas con dificultad de interpretación y en caso de ser necesario redefinir clases. Las diferencias en la interpretación de las clases con mayor confusión según la matriz se dieron por no acuerdo entre los intérpretes y áreas sin información en cualquiera de los dos periodos de evaluación.

Para este mapa, los 459 puntos de no coincidencia, se evidencio que tan solo 4 puntos estaban totalmente en desacuerdo los tres intérpretes y 455 dos intérpretes estuvieron de acuerdo en la asignación de la clase de referencia. Por otra parte, las 9.552 muestras restantes, existió acuerdo total de los intérpretes en la asignación de las clases de referencia.

La Figura 14Figura 7 muestra la distribución de los puntos no coincidentes, los cuales son representados en color negro para aquellos donde la coincidencia fue para dos intérpretes y en color rojo los que no existió ninguna coincidencia entre los tres intérpretes.

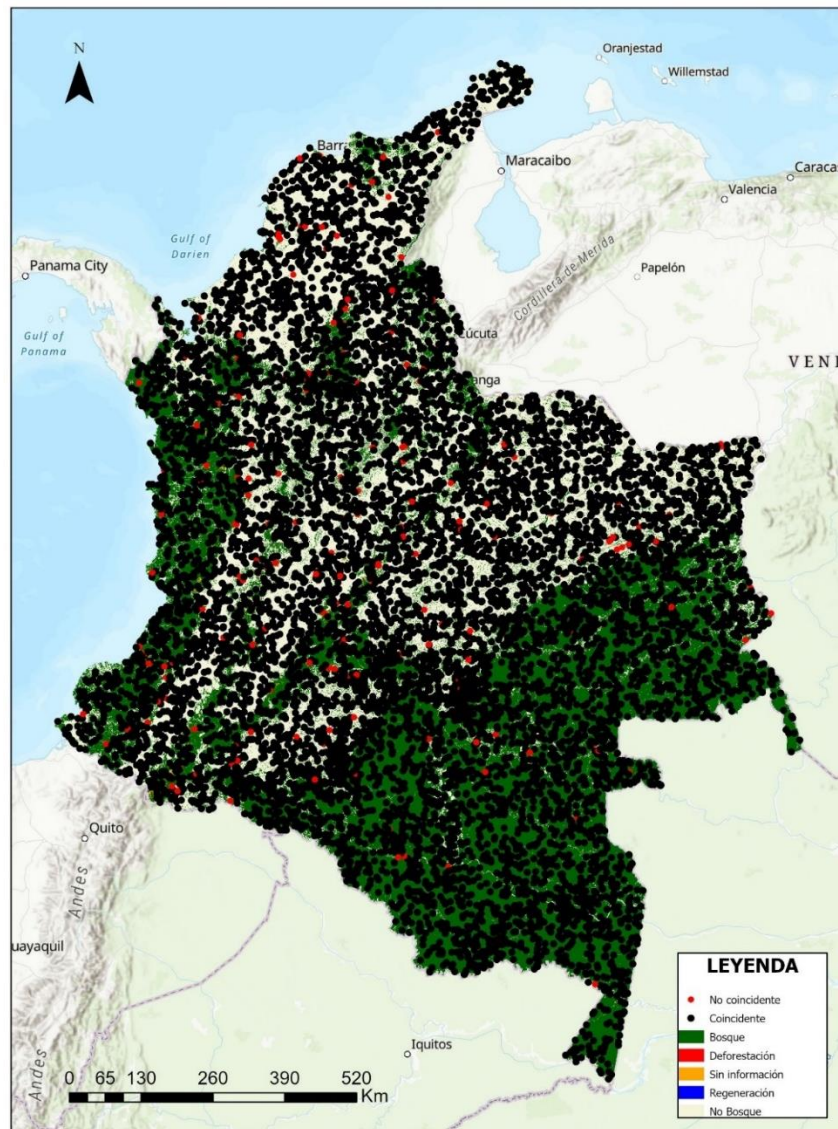


Figura 14. Distribución de los puntos de no coincidencia entre los intérpretes, puntos negros 2 coincidencias, puntos rojos sin coincidencias, , para el periodo 2019-2020

Estos 459 puntos fueron evaluados en conjunto con el líder del equipo, llegando a un consenso sobre el mismo y 14 de los mismos fueron discutidos con el grupo del SMByC, llegando igualmente a un consenso sobre la designación de los mismos.

6.4. Generación de las estadísticas de validación (Matriz de error, cálculos de exactitud, estimadores de área ajustada, incertidumbre)

Una vez realizado el control de calidad, se generó la matriz de error o confusión convencional y la de proporción por área, así como las medidas de fiabilidad acorde al protocolo establecido por (Galindo, Espejo, Rubiano, Vergara, & Cabrera, 2014), tomando para la evaluación de la exactitud temática el 100% de las unidades de la muestra (10.011). A partir de los datos que se observan en la Tabla 8, del total de los 10.011 puntos analizados un total de 9.817 fueron clasificados acorde al valor asignado de referencia. Estos corresponden a 4.466 puntos en la

categoría de bosque, 953 en deforestación, 37 en la categoría sin información y 4.361 a la categoría no bosque, la cual corresponde a la diagonal principal de la Tabla 8 la cual se resalta en verde. En total las no coincidencias corresponden a 194 puntos, estos fueron revisados en QGIS ACATAMA, analizando el contexto general y detallando el punto para tener una visión amplia del territorio, para esto se realizó mejoramiento del histograma y tener una mejor visualización.

Tabla 8. Matriz de error tradicional y medidas de fiabilidad para el periodo 2019-2020

CLASES	Bosque (1)	Deforestación (2)	Sin información (3)	No bosque (5)	Total	Exactitud Usuario	Wi
Bosque (1)	4466	3	2	100	4571	0.97703	0.52266
Deforestación (2)	12	953	0	35	1000	0.953	0.00150
Sin información (3)	2	0	37	1	40	0.92500	0.00178
No bosque (5)	39	0	0	4361	4400	0.99114	0.47406
total	4519	956	39	4497	10011		
Exactitud productor	0.99163	0.80688	0.87776	0.97605		0.98359	

Los elementos ubicados fuera de la diagonal principal, indica los errores de asignación (marginales), ya sea porque perteneciendo a una misma clase han sido asignados a categorías distintas lo cual se denomina omisión (al cual se les asigna a los productos)

Comisión: Del total de 4.571 puntos clasificados en bosque dentro del mapa temático evaluado, se encontró que 4.466 puntos estaban bien clasificados, 3 correspondían a deforestación, 2 a sin información y 100 a no bosque. Para los 1.000 puntos clasificados en deforestación se encontró que 953 estaban bien clasificados, 12 correspondían a bosque y 35 a no bosque. Por su parte de los 40 evaluados en la clase sin información 37 estaban bien clasificados, 2 corresponden a bosque y 1 a no bosque y finalmente de los 4.400 puntos en no bosque, se encontró que 4.361 estaban bien clasificados y 39 pertenecen a bosque.

Omisión: De lo anterior se infiere que en la clase bosque faltaron 12 puntos que fueron clasificados en deforestación, 2 en sin información y 39 en no bosque dentro del mapa temático. Respecto a deforestación se encontraron 3 clasificados en bosque. Para la clase sin información se encontraron 2 clasificados en bosque y finalmente para la clase no bosque se encontró 100 en bosque, 35 en deforestación y 1 en sin información.

La **exactitud** global del mapa es de **98.36%**, la cual está dada como el cociente entre la sumatoria de puntos de la diagonal principal sobre el total de los puntos evaluados.

En cuanto a la exactitud total del usuario fue del **96.2%** en donde para la clase bosque fue de 97.7%; deforestación de 95.3%, sin información de 92.5% y no bosque 99.1%.

Por su parte matriz de error del usuario (Tabla 9), que mide los errores de la comisión, manifiesta que el mayor error en comisión se presentó en la clase clasificada en bosque que pertenecen a no bosque (0.01143).

Tabla 9. Matriz de error de la proporción de área y medidas de fiabilidad para el periodo 2019-2020

CLASES	Bosque (1)	Deforestación (2)	Sin información (3)	No bosque (5)	Usuario	Productor
Bosque (1)	0.51066	0.00034	0.00023	0.01143	0.97703	0.99163
Deforestación (2)	2.00E-05	0.00143	0	5.00E-05	0.953	0.80688
Sin información (3)	9.00E-05	0	0.00164	4.00E-05	0.925	0.87776
No Bosque (5)	0.0042	0.00011	0	0.46986	0.99114	0.97605
total	0.51496	0.00178	0.00187	0.48139		

Por su parte la matriz de error del productor (Tabla 10), mide los errores de la omisión, mostrando que, los mayores errores se presentaron en: i) clase clasificada como deforestación que corresponde a bosque (0.19312); ii) clase clasificada en sin información que corresponde a bosque (0.12224); iii) clase clasificada en bosque que corresponde a no bosque (0.0081) y iv) clase clasificada como bosque que corresponde a sin información (0.00017). Cabe resaltar que los errores más grandes se encuentran en donde las clases tienen un área proporcionalmente inferior a las demás clases. Esto hace que una omisión de una clase de poca área con respecto a una clase de alta área aumente de forma exponencial.

Tabla 10. Matriz de error de la proporción de área y medidas de fiabilidad para el periodo 2019-2020

CLASES	Bosque (1)	Deforestación (2)	Sin información (3)	No bosque (5)	Usuario	Productor
Bosque (1)	0.99163	0.19312	0.12224	0.12224	0.97703	0.99163
Deforestación (2)	4.00E-05	0.80688	0	0.00011	0.953	0.80688
Sin información (3)	0.00017	0	0.87776	9.00E-05	0.925	0.87776
No Bosque (5)	0.00816	0	0	0.97605	0.99114	0.97605

Junto con esto, para cada clase se determinó el **área estimada observando** un error aproximado de 152,784 ha para bosque, 22,614 para deforestación, 20,324 para sin información y 150,023 para no bosque estable (Tabla 11Tabla 5).

Tabla 11. Intervalos de confianza del área estimada por clase para el periodo 2190-2020

CLASES	Área (ha)	Error	Límite inferior	Límite superior	Incertidumbre
1 (Bosque)	58,737,836	152,784	58,438,379	59,037,295	0.0051
2 (Deforestación)	202,598	22,614	158,275	246,921	0.21877
3 (Sin Información)	213,393	20,324	173,559	253,229	0.18667
5 (no Bosque)	54,908,235	150,023	54,614,190	55,202,281	0.00536
Total	114,062,064				

Finalmente, la incertidumbre reportada por Acatama es de 0.5% en bosque, 21.9% en deforestación 19% en sin información y 0.5% en no bosque.

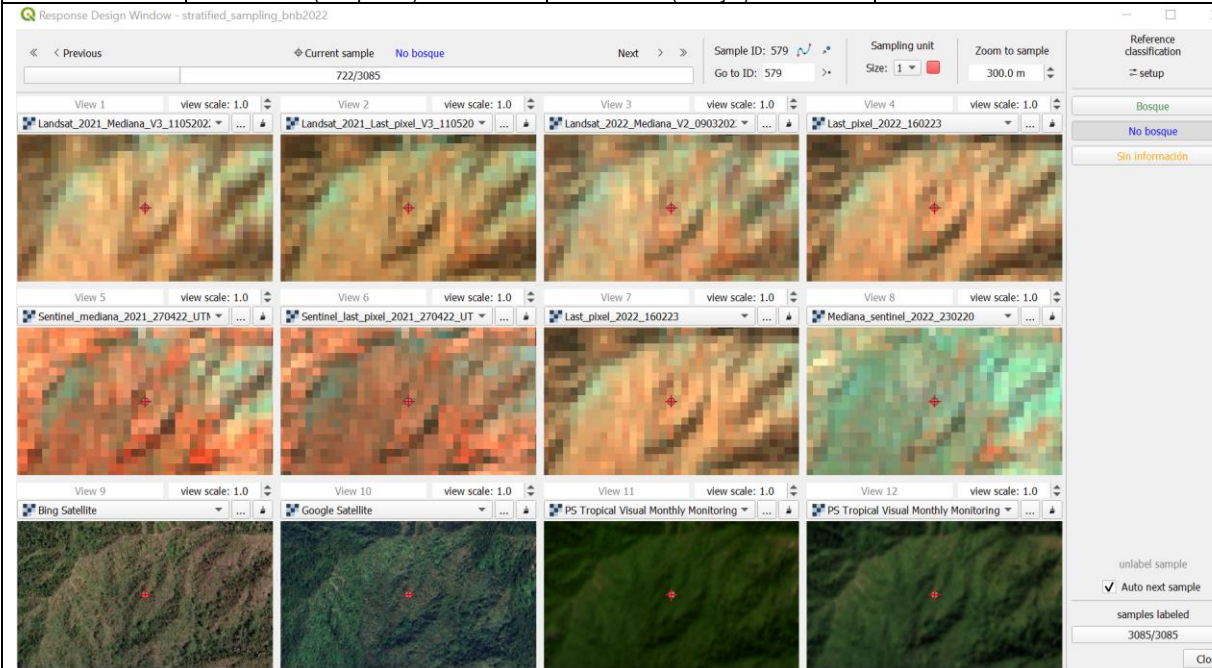
6. BIBLIOGRAFIA

- Boca, T., & Rodríguez, G. (2012). *Métodos estadísticos de la evaluación de la exactitud de productos derivados de sensores remotos*. Buenos Aires.
- Cabrera, E., Vargas, D., Galindo, G., García, M., Ordoñez, M., Vergara, L., . . . Galindo, P. (2011). *Memoria técnica de la cuantificación de la deforestación histórica nacional escalas gruesas y finas*. Bogotá.
- Cochran, W. (1977). *Sampling techniques*. New York: Joh Wiley & Sons.
- Defries, R., Achard, F., Brown, S., Herold, M., Murdiyarto, D., Schalamadinger, B., & Souza, D. (2006). Reducing greenhouse gas in temperate forest. *Remote Sensing Reviews*, 13(207-852).
- Galindo, G., Espejo, O., Rubiano, J., Vergara, L., & Cabrera, E. (2014). *Protocolo de procesamiento digital de imágenes para la cuantificación de la deforestación en Colombia V 2.0*. Bogotá - Colombia: Instituto de Hidrología, Meteorología y Estudios Ambientales - IDEAM.
- Geroffrey, J., & Steiniger, S. (2018). Free and open-source geographic information tools for landscape ecology. *Ecological Informatics*, 4(183-195).
- Liu, P., Du, Q., & Wang, L. (2018). Remote sensing big data: theory, methods and applications. *Remote Sensing*.

- Olofsson, P., Foody, G., & Herold, M. (2014). Good practices for estimating area and assessing accuracy of land change . *Remote Sensing of Environment*, 148 (42-57).
- Olofsson, P., Foody, G., Stehman, S., & Woodcock, C. (2012). Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. *Remote Sensing of Environment*.
- Ramirez, S., Ramirez, M., Galindo, G., & Espejo, O. (2014). Evaluación de la exactitud temática del mapa de cambio de bosque para la región amazónica colombiana en el periodo 2010-2012. *Memorias XVI Simposio Internacional Selper 2014*. Medellin.
- Toure, D., Storw, D., Shih, H., Weeks, J., & Lopez-Carr, D. (2018). Land cover and land use changes analysis using multi-spectral resolution data and object-based image analysis. *Remote sensing of environment*.

7. ANEXOS

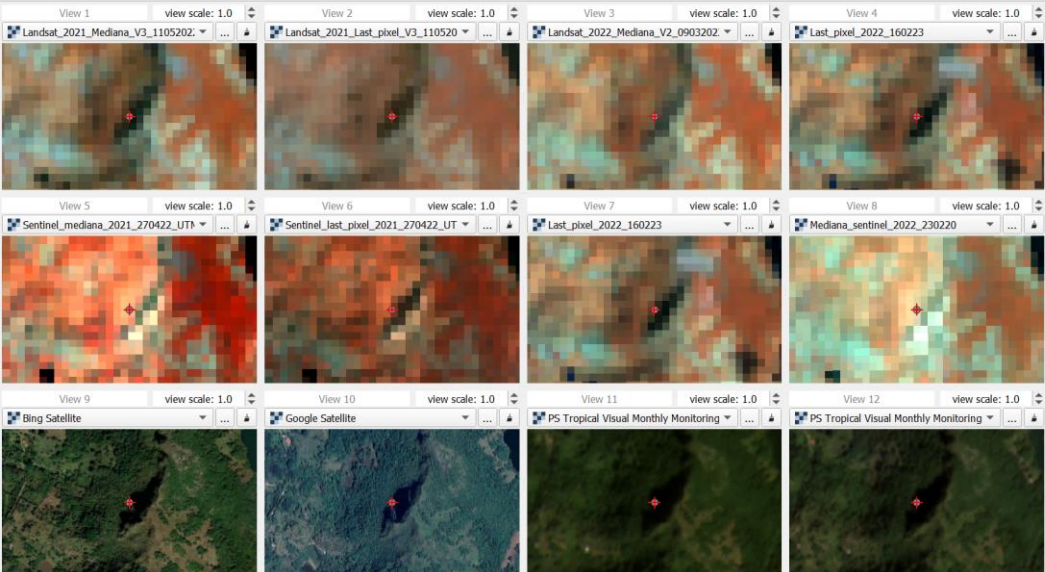
Anexo 1: Ejemplos de puntos de no coincidencia revisados para el mapa de bosque-no bosque 2022

579	No bosque estable (interpretes)	Bosque (SMBYC)	Todas las imágenes muestran cobertura arbustiva.
			

2100	No bosque estable (interpretes)	Bosque (SMBYC)	Todas las imágenes muestran cobertura arbustiva, sombra de montaña
------	---------------------------------	----------------	--

Response Design Window - stratified_sampling_bnb2022

Current sample: No bosque (986/3085) | Sample ID: 2100 | Sampling unit: 1 | Zoom to sample: 300.0 m



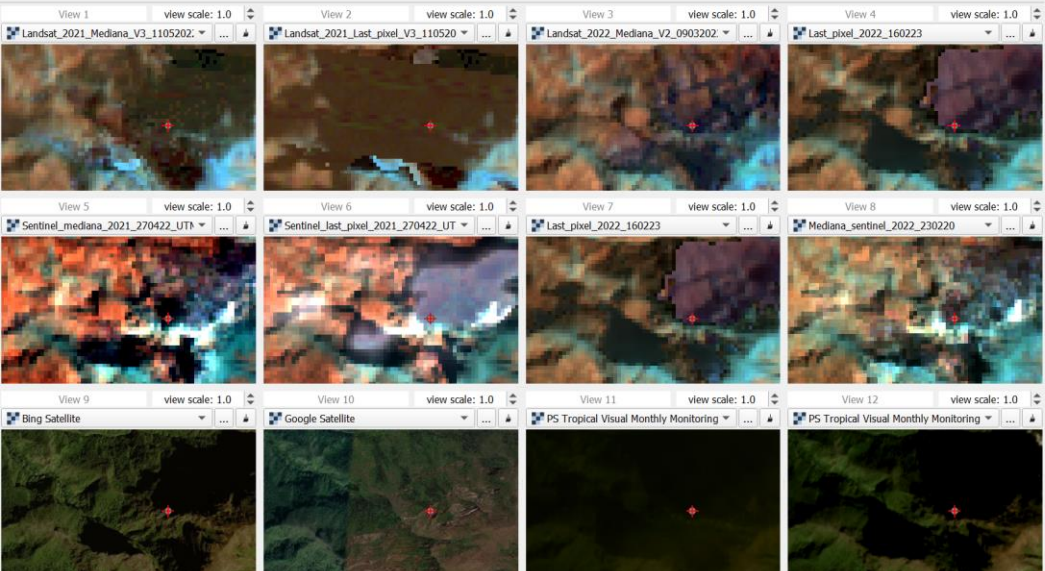
Reference classification: Bosque, No bosque, Sin información

Auto next sample: | samples labeled: 3085/3085

1156	No bosque estable (interpretes)	Bosque (SMBYC)	A 2008 se observaba no bosque y en la Google satélite se observa que es no bosque
------	---------------------------------	----------------	---

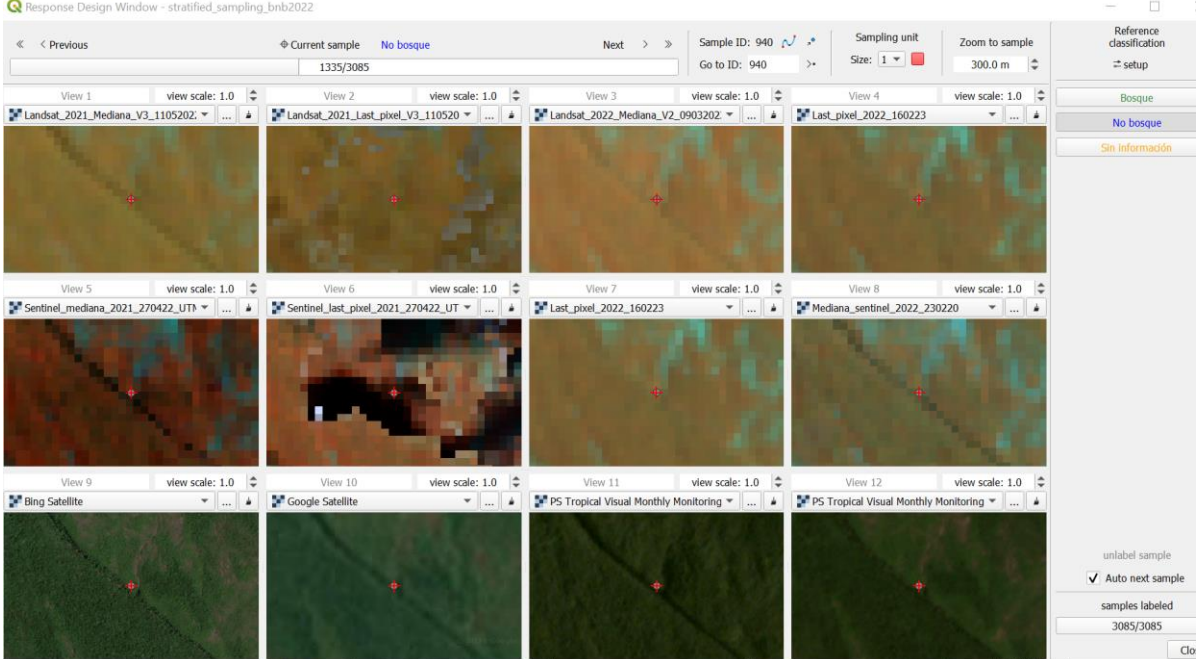
Response Design Window - stratified_sampling_bnb2022


Current sample: No bosque (1235/3085) | Sample ID: 1156 | Sampling unit: 1 | Zoom to sample: 300.0 m

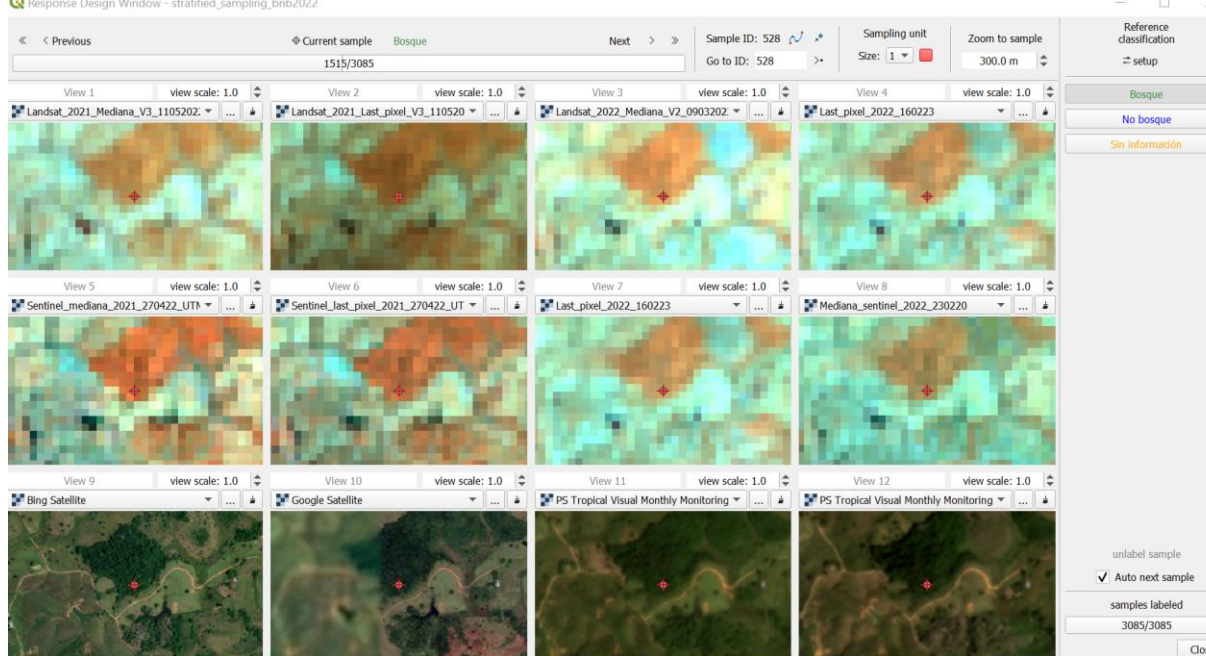


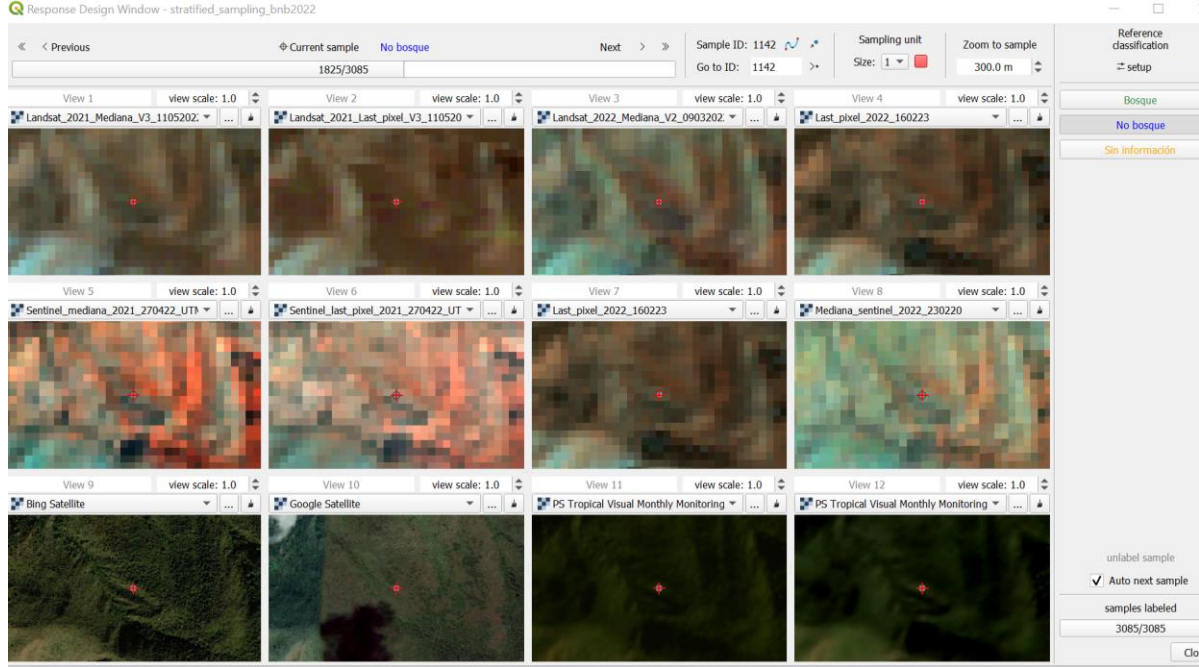
Reference classification: Bosque, No bosque, Sin información

Auto next sample: | samples labeled: 3085/3085

940	No bosque estable (interpretés)	Bosque (SMBByC)	Área límite de boque que no cumple con los requerimientos mínimos de áreas con los vecinos cercanos
			

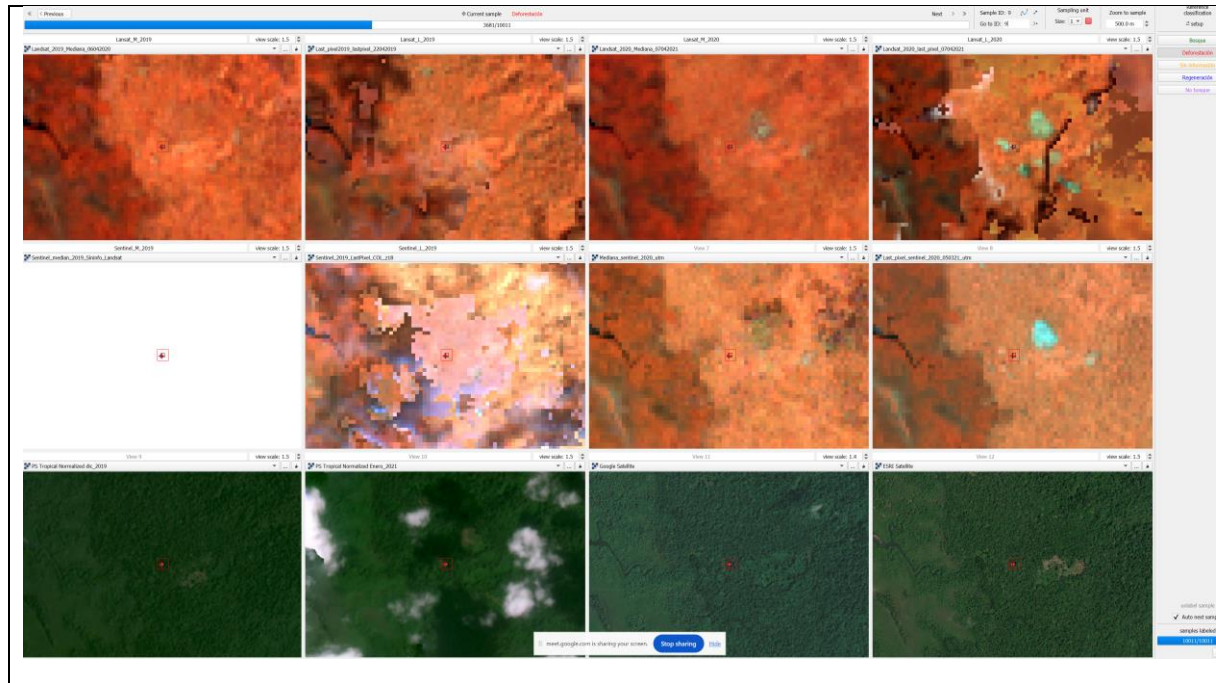
2114	No bosque estable (interpretés)	Bosque (SMBByC)	Las imágenes muestran para los dos periodos cobertura no boscosa
			

528	Bosque (interpretes)	No Bosque Estable (SMBYC)	Las imágenes muestran para los dos periodos cobertura boscosa, cumple con la hectárea y los requerimientos de fao para ser bosque.
 <p>The screenshot displays a software interface for monitoring a forest sample. At the top, it shows the sample ID '528' and the current classification 'Bosque'. Below this, there are 12 individual view windows arranged in a 3x4 grid. The first two rows show various satellite imagery (Landsat and Sentinel) and monitoring data (Last_pixel and Mediana) for the years 2021 and 2022. The third row shows Bing Satellite and Google Satellite imagery. Each view window includes a 'view scale: 1.0' dropdown and a red crosshair indicating the sample location. On the right side, there is a sidebar with a 'Reference classification' section containing 'Bosque' and 'No bosque' options, and a 'samples labeled' section with the ID '3085/3085'. The interface also includes navigation buttons like 'Previous' and 'Next', and a 'Zoom to sample' control.</p>			

1142	No Bosque Estable (interpretes)	Bosque (SMBYC)	Las imágenes muestran para los dos periodos cobertura no boscosa, y las imágenes Google históricas muestran no bosque
			

Anexo 2: Ejemplos de puntos de no coincidencia revisados para el mapa de cambio de superficie de bosque 2019-2020

9	Deforestación (interpretes)	Bosque (SMBYC)	Las imágenes muestran para los dos periodos cobertura no boscosa, y las imágenes Google históricas muestran no bosque
---	--------------------------------	-------------------	---



1192	Desforestación (interpretes)	NO Bosque (SMBYC)	Las imágenes muestran para los dos periodos cobertura no boscosa, y las imágenes Google históricas muestran no bosque
