

**ANÁLISIS COMPARATIVO DEL RENDIMIENTO DE CLASIFICACIÓN SUPERVISADA EN
IMÁGENES COMPUESTAS Y MODIFICADAS POR REDUCCIÓN EN PNN LOS
NEVADOS Y ALREDEDORES**

¹Manuela Rentería Cárdenas

Informe final de trabajo de grado presentado como requisito parcial para optar al título de
Especialista en Sistemas de Información Geográfica

Director (a):

MSc César Anibal Olmos Severiche

Universidad de Manizales

Facultad de Ciencias e Ingeniería

Especialización en Sistemas de Información Geográfica

Manizales, 2024

Resumen

El Parque Nacional Natural Los Nevados y sus zonas de amortiguación, en la región centro-occidental de Colombia, presentan condiciones meteorológicas caracterizadas por la persistencia de nubosidad y periodos de fumarola activa. Esto supone un reto importante para el estudio de las coberturas terrestres a partir de imágenes satelitales y algoritmos de clasificación supervisada, en términos de análisis y captura de la información por el ruido presente

Ante este panorama, resulta de gran interés explorar los resultados de clasificación supervisada de coberturas asociados a dos tipos de imágenes, la primera correspondiente a imágenes satelitales con bandas compuestas sin modificación y la segunda a imágenes satelitales con bandas modificadas por reducción, donde a partir de múltiples imágenes para un rango de tiempo, se genera un único producto cuyos píxeles responden a la elección estadística del usuario; esto para efectos de un análisis comparativo sobre su rendimiento e intentar esclarecer cuál insumo resulta pertinente implementar en aquellas zonas donde la nubosidad sea problemática, siendo esto muy común en el contexto colombiano. Para ello se implementan herramientas SIG y plataformas para análisis geoespacial como Google Earth Engine, donde se llevan a cabo diferentes procesamientos y algoritmos, desde la visualización de imágenes, selección, limpieza, entrenamiento del modelo, visualización de clases, entre otras. Estas actividades se realizan con el fin de constatar la validez de este método como primer paso para estudios de coberturas terrestres, y abrir un cuestionamiento sobre la clasificación de este tipo de insumos.

Palabras clave: Imágenes satelitales, reducción, clasificación supervisada, coberturas terrestres, Google Earth Engine, SIG.

Abstract

An important challenge for the study of land covers from satellite imagery lies in the application of supervised classification algorithms, particularly in regions with unique meteorological conditions such as Los Nevados National Natural Park in Colombia's central-western region. Here, persistent cloudiness and active fumaroles pose significant limitations on remote sensing analysis.

To address this challenge, there is growing interest in comparing the results of supervised classification using two types of images. The first type consists of satellite images with composite bands without modification, while the second type involves images with bands modified by reduction, through the synthesis of multiple images over a time range, a single product is generated, enabling pixels to respond to user-defined statistical parameters. This comparative analysis aims to discern the most relevant input for implementation in areas plagued by cloud cover, a common issue in the Colombian context. Leveraging GIS tools and platforms like Google Earth Engine, various processing steps and algorithms were employed, involving tasks from image visualization and selection, to model training and class visualization. The efficacy of this method was evaluated through these processes and served as an initial approach for land cover studies while also prompting inquiry into the classification of such inputs.

Keywords: Satellite images, reduction, supervised classification, land covers, Google Earth Engine, GIS.

TABLA DE CONTENIDO

1. PLANTEAMIENTO DEL PROBLEMA.....	8
2. JUSTIFICACIÓN	10
3. OBJETIVOS	11
3.1 Objetivo General	11
3.2 Objetivos específicos	11
4. ANTECEDENTES	12
5. MARCO NORMATIVO.....	19
6. MARCO TEÓRICO.....	20
7. METODOLOGÍA.....	24
7.1 Enfoque metodológico.....	24
7.2 Tipo de estudio.....	24
7.3 Área de estudio	24
7.4 Procedimiento	25
8. RESULTADOS	59
9. CONCLUSIONES.....	64
10. RECOMENDACIONES	66
11. REFERENCIAS	68

LISTADO DE FIGURAS

Figura 1: Localización del área de estudio.	25
Figura 2: Fase 1 del proyecto, con objetivos 1 y 2 avanzando de forma paralela.	26
Figura 3: Imagen Sentinel 2 seleccionada, con nubosidad del 0 – 10%.	28
Figura 4: Código asociado al proceso “de sorting” o sorteo en la plataforma GEE.	28
Figura 5: Código asociado al proceso “de sorting” o sorteo en la plataforma GEE.	29
Figura 6: Biblioteca Sentinel-2 MSI armonizada.	30
Figura 7: Código desarrollado en lenguaje JAVA.	31
Figura 8: Filtros de fecha y nubosidad empleados en la consola de GEE.	31
Figura 9: Escenas que cumplen con los requisitos de fecha y nubosidad especificados.	32
Figura 10: Código relacionado con el recorte de la información y configuración del estadístico.	33
Figura 11: Código relacionado con el enmascaramiento de las nubes en la consola de GEE.	33
Figura 12: Código relacionado con la exportación del producto obtenido.	34
Figura 13: Fase 2 del proyecto, con objetivos 3 y 4 avanzando de forma paralela en la primera parte y de forma conjunta al final.	34
Figura 14: Imagen Sentinel 2A obtenida a partir de extraer del ráster original la información correspondiente al área de estudio.	36
Figura 15: Mosaico de Sentinel 2 para un periodo de 2018 hasta 2023 generado mediante algoritmos en la plataforma GEE.	37
Figura 16: Semillas o muestras de entrenamiento dispuestas para fines de clasificación supervisada.	39
Figura 17: Semillas implementadas en la identificación de la clase “cuerpos de agua”. Imagen Mosaico (Izquierda) e Imagen S2A 2020 (Derecha), combinación en color real (B4+B3+B2).	39
Figura 18: Semillas implementadas en la identificación de la clase “tejido urbano”. Imagen Mosaico (Izquierda) e Imagen S2A 2020 (Derecha), combinación en color real (B4+B3+B2).	40
Figura 19: Municipio de Salento, Quindío representando la clase de tejido urbano.	40
Figura 20: Semillas implementadas en la identificación de la clase “vegetación densa”. Imagen Mosaico (Izquierda) e Imagen S2A 2020 (Derecha), combinación en color real (B4+B3+B2).	41
Figura 21: Semillas implementadas en la identificación de la clase “vegetación baja”. Imagen Mosaico (Izquierda) e Imagen S2A 2020 (Derecha), combinación en color real (B4+B3+B2).	41
Figura 22: Semillas implementadas en la identificación de la clase “periglacial”. Imagen Mosaico (Izquierda) e Imagen S2A 2020 (Derecha), combinación en color real (B4+B3+B2).	42
Figura 23: Semillas implementadas en la identificación de la clase “roca”. Imagen Mosaico (Izquierda) e Imagen S2A 2020 (Derecha), combinación en color real (B4+B3+B2).	42
Figura 24: Semillas implementadas en la identificación de la clase “nieve”. Imagen Mosaico (Izquierda) e Imagen S2A 2020 (Derecha), combinación en color real (B4+B3+B2).	43
Figura 25: Semillas implementadas en la identificación de la clase “nieve”. Imagen Mosaico (Izquierda) e Imagen S2A 2020 (Derecha), combinación en falso color (B11+B8A+B4).	43

Figura 26: Semillas implementadas en la identificación de la clase “suelo desnudo”. Imagen Mosaico (Izquierda) e Imagen S2A 2020 (Derecha), combinación en color real (B4+B3+B2).	44
Figura 27: Semillas implementadas en la identificación de la clase “nube”. Imagen Mosaico (Izquierda) e Imagen S2A 2020 (Derecha), combinación en color real (B4+B3+B2).	44
Figura 28: Mapa de sombras de la zona de estudio.	45
Figura 29: Semillas implementadas en la identificación de la clase “sombras”. Imagen Mosaico (Izquierda) e Imagen S2A 2020 (Derecha), combinación en color real (B4+B3+B2).	45
Figura 30: Relación entre las semillas de la clase sombra y el mapa de sombras de la zona.	46
Figura 31: Clasificación supervisada obtenida de la imagen S2A del año 2020.	47
Figura 32: Clasificación supervisada obtenida mosaico S2A.....	48
Figura 33: Ejercicio de clasificación supervisada realizado por estudiantes del Politécnico Colombiano Jaime Isaza Cadavid.	49
Figura 34: Coberturas de la tierra Corine Land Cover nivel 3, año 2018 100 K.....	50
Figura 35: Muestras aleatorias dispuestas para el mosaico Sentinel 2A.	52
Figura 36: Muestras aleatorias dispuestas para la imagen Sentinel 2A año 2020.	53

LISTADO DE TABLAS

Tabla 1: Matriz de confusión imagen mosaico Sentinel 2A.....	54
Tabla 2: Matriz de confusión imagen Sentinel 2A año 2020	54
Tabla 3: Índices calculados para la imagen mosaico Sentinel 2A.	57
Tabla 4: Valores calculados para la imagen mosaico Sentinel 2A.....	57
Tabla 5: Índices calculados para la imagen Sentinel 2A año 2020.....	57
Tabla 6: Valores calculados para la imagen S2A año 2020.....	57
Tabla 7: Valoración del coeficiente kappa.	58

1. PLANTEAMIENTO DEL PROBLEMA

En el estudio y análisis de las coberturas terrestres mediante el uso de imágenes satelitales de acceso libre, es común la discretización y cuantificación de dichas coberturas mediante la asignación de clases aleatorias o entrenadas (clasificación no supervisada y supervisada) en software GIS (ArcGIS, QGIS). En este proceso de análisis, la presencia de nubes y sombras suponen un gran reto en términos de disponibilidad espacial y temporal de la información, puesto que el ruido que ellas generan no permite estandarizar la información en los píxeles asociados a cada una de las clases existentes, generando vacíos en las zonas donde no es posible observar la superficie dadas las condiciones meteorológicas. En países tropicales como Colombia, donde la topografía juega un papel clave en la existencia y persistencia de nubosidad, se encuentran zonas con grandes dificultades para ser estudiadas de cuenta de esta situación.

Es por lo anterior que dentro de las alternativas de solución se adopta la construcción de imágenes compuestas, donde no solo se contempla un momento en específico sino la recopilación de datos para un rango de tiempo determinado y se elige el píxel que se adapte a las condiciones requeridas. En este proceso denominado reducción, se toman múltiples imágenes de un mismo sitio, captadas en momentos diferentes, y se genera un único producto, cuyos píxeles provienen del grupo de imágenes y que representan un estadístico particular (media, mediana, desviación estándar, máximo, mínimo, etc) de los datos. Este proceso, genera entonces un producto con características particulares de presencia o ausencia de información sensible al proceso, pero que no supone una imagen satelital en sí sino un compuesto artificial basado en ellas y que por ello no es un continuo de información en lo referente a lo temporal, más sí en lo espacial.

Para realizar este análisis se consideró el Parque Nacional Natural Los Nevados y sus zonas de amortiguación, ubicado entre los departamentos de Caldas, Quindío,

Risaralda y Tolima. Dentro de este parque se encuentran diferentes nevados de interés y gran importancia, entre los que figura el Nevado del Ruiz, el cual se eleva hasta los 5321 m.s.n.m. y se ubica entre dos valles interandinos con circulación de vientos particular, que genera ascenso de masas de aire cargadas de humedad que se condensan y generan persistencia de nubosidad. Adicionalmente, el volcán posee en ciertos períodos una fumarola activa, que se aprecia como una forma de nubosidad persistente con su respectiva sombra. La combinación de factores anteriores configura este macizo volcánico como un sitio de monitoreo satelital con retos particulares, especialmente para el estudio de coberturas a través de imágenes satelitales de un momento particular o imágenes con reducción para limpieza de ruido ambiental.

Ante este panorama es necesario conocer si al implementar imágenes satelitales para la clasificación supervisada de coberturas, el rendimiento y sus resultados asociados son diferentes a los que se obtienen con estos revolucionarios productos que se han venido adoptando con cada vez más frecuencia como solución viable a los problemas de ruido ambiental. Esta necesidad se relaciona con la validez de este método como primer paso para estudios de coberturas terrestres, y abriría un cuestionamiento sobre la clasificación de este tipo de insumos.

2. JUSTIFICACIÓN

Las imágenes satelitales son un insumo fundamental en la ejecución de estudios y análisis del territorio, esto dadas sus múltiples bondades en términos de extensión, puesto que permiten analizar grandes porciones del terreno en forma simultánea e instantánea (Chuvienco, 2008). Sumado a la cantidad y diversidad de datos obtenidos a partir de ellas y aplicables a muchas áreas del conocimiento, constituyendo la base para análisis geoespaciales de aspectos estratégicos e indispensables para el correcto aprovechamiento y distribución de los recursos particulares de cada zona (Chuvienco, 2008). Sin embargo, existen limitaciones en su calidad por la presencia de ruido en ellas sea nubosidad, sombras u otros, esto generalmente se encuentra condicionado por factores climáticos.

El Parque Nacional Natural Los Nevados en particular el volcán Nevado del Ruiz y sus zonas de amortiguación, suponen un área de gran interés geológico y de condiciones meteorológicas complejas dada su configuración geomorfológica, caracterizado por alta presencia de nubosidad y periodos de fumarola activa. Conocer los cambios y modificaciones que se han presentado en la cobertura terrestre de la zona, es de gran importancia para entender las dinámicas ambientales y geológicas.

Dentro de las novedades del presente trabajo se encuentra el análisis comparativo del rendimiento de dos tipos de imágenes, para la obtención de coberturas terrestres mediante los algoritmos involucrados en la clasificación supervisada. La primera imagen corresponde a un momento en específico y la segunda a una construcción a partir de píxeles obtenidos de la recopilación de datos para un rango de tiempo determinado, cuya continuidad espacial permite obtener imágenes con reducción de ruido y mejor visualización. Estas actividades se realizan con el fin de consolidar conocimiento sobre las diferentes alternativas existentes para el estudio de coberturas, pudiendo establecer qué herramientas son óptimas y cuáles permiten la toma de mejores decisiones para su conservación y cuidado, en la medida que los resultados se aproximen a las condiciones del sitio.

3. OBJETIVOS

3.1 Objetivo General

Determinar el rendimiento de imágenes satelitales con bandas compuestas sin modificación versus imágenes satelitales con bandas modificadas por reducción, para fines de clasificación supervisada en el área del Parque Nacional Natural Los Nevados y sus zonas de amortiguación, mediante un análisis comparativo que involucre la implementación de herramientas SIG y plataformas de análisis espacial.

3.2 Objetivos específicos

3.2.1 Revisar a partir de criterios de calidad las imágenes satelitales de la zona de estudio disponibles en las plataformas de acceso libre.

3.2.2 Generar mediante algoritmos de reducción en la herramienta Google Earth Engine imágenes satelitales con variaciones de los filtros de metadatos y calidad relacionada con fechas y porcentajes de nubosidad.

3.2.3 Elaborar una clasificación supervisada de cobertura vegetal a partir de imágenes satelitales con bandas compuestas sin modificación e imágenes satelitales con bandas modificadas por reducción para obtener mapas de coberturas de la zona de interés.

3.2.4 Comparar los resultados obtenidos y su confiabilidad a partir de ambas fuentes, para efectos de la elaboración de cartografía temática que permita la correcta toma de decisiones.

4. ANTECEDENTES

A continuación, se enuncian diferentes trabajos relacionados con el estudio de coberturas terrestres, uso de herramientas SIG y Google Earth Engine para el procesamiento y análisis de imágenes satelitales.

Arango Gutiérrez et al., (2005) abordan la clasificación no supervisada de coberturas en imágenes digitales de los sensores remotos Landsat ETM, como una alternativa para responder en primera instancia al déficit de inventarios de las especies vegetales en Colombia. Se realizó una clasificación supervisada distinguiendo un total de 8 clases de coberturas, esto con el fin de obtener un referente respecto al cual se pueda comparar el desempeño de los resultados obtenidos a partir de los clasificadores no supervisados ISODATA y SOM. Para lo anterior, los autores tuvieron en cuenta aspectos estadísticas de las imágenes, características del sensor, definición de las mejores bandas y sus combinaciones, depuración de datos, entre otros. La información resultante es comparada con la clasificación supervisada mediante matrices de confusión, sugiriendo que no existe una buena convergencia entre los métodos de clasificación supervisada para las coberturas vegetales.

Peña Araya, (2007) implementan procedimientos de corrección en imagen satelital ASTER para la estimación de la fracción de cobertura vegetal (VCF) y el índice de vegetación de diferencia normalizada (NDVI) en la cuenca del río Mirta ubicada en la región Aisén, Chile. Entre los procedimientos están la corrección atmosférica, topográfica y de subpíxel. Se pudo encontrar que para el VCF la corrección de sombras a nivel de subpíxel aumenta notablemente su promedio, lo cual incide en valores más homogéneos, así como la modificación de la distribución espacial de valores, por lo cual ayudan al cálculo de parámetros de vegetación en bosques donde incide el ángulo solar oblicuo, altas pendientes y sombras de vegetación superior. Para el caso del NDVI, es la corrección atmosférica la

que mejora la toma de valores, teniendo en cuenta que se debe validar la precisión de los parámetros calculados, lo cuales deberán ser contrastados con medidas equivalentes *in situ*, con el uso de matrices de error y relaciones de asociatividad.

Paz Pellat et al., (2015) se enfocan en analizar las bondades y restricciones de los índices de vegetación implementando parámetros teórico experimentales, siendo estos índices una herramienta cada vez más utilizada en el análisis de coberturas, pero con cuestionamientos en relación con su diseño y aplicación. El análisis se realiza mediante modelos radiativos de interacción de orden uno y dos, para un total de sesenta índices los cuales se construyen en un formato asociado a patrones de curvas espectrales. Los resultados obtenidos sugieren que solo un índice de vegetación cumple con requerimientos tanto teóricos como experimentales, esto para la etapa vegetativa reproductiva.

Shelestov et al., (2017) buscan determinar el grado de eficiencia de los procesos de clasificación de imágenes multitemporales en la plataforma Google Earth Engine, esto comparando el rendimiento de los últimos clasificadores disponibles para la obtención de mapas de cultivos, con resolución de 30 metros en áreas extensas de Ucrania, particularmente la región de Kiev. El comparativo de las fortalezas y debilidades de los clasificadores en cuestión, se realizan respecto al clasificador de referencia en el que se implementa un enfoque de red neuronal de estudios anteriores. De acuerdo con los resultados obtenidos, se establece que GEE presenta ventajas en términos de acceso a los productos de detección remota a través de la nube, no obstante, en lo que compete a la precisión en la clasificación, se tiene que el enfoque basado en redes neuronales es más preciso que el generado por las máquinas de vectores de soporte.

Gorelick et al., (2017) compilan las diferentes posibilidades que ofrece la plataforma Google Earth Engine para el análisis y estudios de problemáticas de carácter climático, ambiental, hídrico, entre otras. Al ser una plataforma basada en la nube permite utilizar

herramientas informáticas competentes para el procesamiento de grandes volúmenes de datos. Dentro de sus grandes ventajas, se encuentra el catálogo de datos donde se almacenan imágenes aéreas y satelitales como lo son Landsat, Sentinel-1 y Sentinel-2, también se pueden encontrar datos de cobertura terrestre, pronóstico del clima, entre otras. El ritmo con el que se actualiza este catálogo es constante siendo una herramienta útil y pertinente para diferentes estudios, por otra parte, la plataforma GEE permite compartir hallazgos y resultados fácilmente entre usuarios de la comunidad científica, facilitando al usuario no solo utilizar los insumos del catálogo público sino también de su entorno local.

Las condiciones de nubosidad y topografía de la Reserva Natural de Fanjingshan en China, dificultan el mapeo y monitoreo de la vegetación. Bajo esta premisa los autores Tsai et al., (2018) abordan técnicas avanzadas de clasificación de imágenes a través de la plataforma Google Earth Engine, involucrando procesos como la normalización de imágenes, construcción de compuestos de imágenes multi-estacionales, adecuación de parámetros para los métodos de clasificación de aprendizaje automático, generación de mapas de clasificación y finalmente definición de la precisión de estos productos respecto a los datos recopilados a través de jornadas de campo con parcelas de muestreo. De acuerdo con los resultados obtenidos, se establece que es un método con altos valores de precisión para el mapeo remoto de coberturas y usos del suelo, dada la eficacia de la herramienta Google Earth Engine.

Avogadro & Padró Garcia, (2019) buscan comparar y analizar las ventajas entre el uso de la clasificación automática (CA) de imágenes de sensores MSI (MultiSpectral Imager) de Sentinel-2 (S2) y OLI (Operational Land Imager) de Landsat-8 (L8) en cuanto a la fiabilidad y precisión de los datos, para identificar las plantaciones de Eucalyptus y Pinus en el departamento de Concordia, Argentina. Mediante el uso de los SIG se desarrolló el inventario forestal con el cual se obtuvo el mapa de plantaciones forestales en Argentina, permitiendo cuantificar su superficie y disponer de cobertura digital de las plantaciones para

consolidar su inventario. La evaluación de la calidad de los datos obtenidos se realizó mediante matrices de confusión de clasificación supervisada y no supervisada, tanto en S2 como en L8. De este análisis se determinó que ambas herramientas son precisas y confiables en la cuantificación de las plantaciones forestales, cuya diferencia radica en mejoras de resolución espacial y espectral si se opta por el S2, pero se cuenta con mejoras en resolución radiométrica en L8.

Otro estudio para tener en cuenta fue el realizado por Arenas Herrera et al., (2019), donde se analiza la utilidad de Google Earth Engine en herramientas fundamentales de procesamiento y preprocesamiento de imágenes Sentinel 2, basados en el proyecto de sensores remotos presentado por el IGAC en 2019. GEE es una plataforma que provee catálogos de imágenes satelitales y datasets geográficos, con información en diferentes escalas para conocer patrones de cambio, cuantificar tendencias y obtener información de superficies. Con los resultados dados por GEE para el preprocesamiento de imágenes con S2 (corrección atmosférica, reducción de colecciones, generación de mosaicos), así como en procesamiento de imágenes con S2 (cálculo del índice normalizado de vegetación NDVI, clasificación supervisada) se obtuvo información con alto nivel de precisión, la cual permite ahorrar costos y ser eficientes en los procesos de investigación.

Zhang & Zhang, (2020) implementaron la herramienta Google Earth Engine para estudiar los cambios de cobertura entre 1987 y 2017, para la zona baja del río Yangtze en Nanjing. Este proceso incluyó la adquisición y procesamiento de seis compuestos de datos (periodos de tres años cada uno), así como la selección de muestras de cobertura para el entrenamiento del modelo y posteriormente la clasificación mediante el algoritmo de Random Forest basada en las imágenes satelitales Landsat y datos auxiliares con las muestras de entrenamiento seleccionadas, utilizando cuatro categorías de cobertura terrestre (agua, vegetación, área construida y áreas áridas). La precisión del modelo se estableció a partir de los puntos de muestreos seleccionados aleatoriamente. De acuerdo

con los resultados obtenidos se estableció que las categorías de vegetación y agua superficial tienen una precisión alta, mientras que las áreas áridas y edificadas una baja, debido principalmente a sus características espectrales similares.

LUO et al., (2021) mediante el uso de Google Earth Engine (GEE), llevaron a cabo el mapeo de la distribución de cultivos en la provincia de Heilongjiang, China. Lo anterior, involucra insumos como imágenes Sentinel 1 y 2 de todas las zonas cubiertas dentro del área para el periodo de crecimiento crítico en el año 2018, imágenes mensuales compuestas, índices de vegetación y bandas de polarización, además de procesos como la clasificación de los cultivos a través de Random Forest. Los resultados obtenidos sugieren una precisión del 89,75% para el mapa de distribución, al combinar imágenes Sentinel con Random Forest, en esta misma línea el uso de imágenes de series temporales, la incorporación de bandas de infrarrojos de onda corta, y el uso de índices de vegetación se consideran elementos clave para mejorar progresivamente los resultados.

El presente estudio se centra en mapear áreas urbanas en tres ciudades de China, para ello los autores Zhang et al., (2021) se interesan en determinar la efectividad de la composición de imágenes de series temporales anuales puntualmente las de Landsat 8, esto con apoyo de Google Earth Engine. Su metodología se basa en llevar a cabo algoritmos de reducción en la herramienta Google Earth Engine, específicamente por valores mínimos, máximos y la mediana, esto para elaborar cuatro esquemas que permitieran sintetizar series temporales anuales, una vez construidos las áreas urbanas se mapean con el algoritmo de Random Forest. En lo que respecta a la precisión de la información, se tiene que el mapeo de áreas urbanas obtenido al combinar los métodos de reducción mencionados supera el 90%, a su vez se destaca que los compuestos de imágenes anuales mejoran significativamente la presencia de nubosidad y sombras asociadas. Así mismo, al implementar varias funciones de reducción los resultados son de calidad superior a diferencia de cuando se implementa sólo una.

Por otra parte, con el fin de evaluar la aplicabilidad en la clasificación orientada a objetos en dos áreas de cultivos en Keshan y Tongnan, Luo et al., (2021) estudian las imágenes satelitales Sentinel-1 de Google Earth Engine (GEE), pudiendo establecer que esta herramienta ofrece una mejor resolución gracias a la información recopilada por el Big Data y Machine Learning, para diferentes intervalos de tiempo mediante el método de segmentación de imágenes de agrupamiento no iterativo simple (SNIC). Por los resultados obtenidos, se definió que la respuesta dada en un intervalo de tiempo menor proporciona información más detallada para la clasificación de cultivos. Ahora bien, si se combinan la clasificación de los datos Sentinel-1 en conjunto con GEE se mejora aún más la precisión de la clasificación de cultivos.

Hay Chung et al., (2021) destacan que en los estudios ambientales urbanos son de gran importancia los mapas de zonas climáticas locales (LCZ) de alta precisión y a gran escala, para ello plantean el mejoramiento del flujo de trabajo mediante la plataforma Google Earth Engine (GEE), basados en datos óptimos y selección de algoritmos de aprendizaje automático (ML), para las zonas de Guandong, Macao y Hong Kong al sureste de China. Se encontró que RF y SVM son los algoritmos de ML más apropiados para la clasificación LCZ, obteniendo precisiones con mejoras del 6.47% y 8.37% respectivamente, en comparación con el enfoque tradicional. Si bien es cierto que RF brinda resultados más precisos, son más homogéneos los datos obtenidos con SVM, al tener una variabilidad mucho menor con respecto a la precisión.

Lindsay et al., (2022) a través de Google Earth Engine buscan mejorar la visibilidad de deslizamientos de tierra y facilitar su detección, utilizando imágenes Sentinel 1 y 2 para la zona del antiguo condado de Song of Fjordane (Noruega) y tomando como referencia el índice de vegetación de diferencia normalizada (NDVI). Con los resultados obtenidos, se confirma la utilidad de las imágenes S1 y S2 en escalas temporales y espaciales diversas, para comprender los deslizamientos. Si se toma el NDVI como referencia de estudio, se

recomienda la combinación S2-MT, en caso de vegetación verde, y enfoque combinado S1-MT y S2-MT cuando no se tenga espectro libre de nubosidad. Finalmente, se presenta una serie de hallazgos en esta región de Noruega, tales como cantidad de deslizamientos de tierra y su resolución.

Eraso Terán et al., (2022) abordan el uso de Google Earth Engine (GEE) para el estudio de incendios forestales, en lo que concierne a cambios de usos del suelo y coberturas y evaluación de superficies quemadas, particularmente en el Parque Nacional Natural El Tuparro ubicado en el departamento de Vichada, Colombia. Allí se presentó un incendio forestal que afectó a más de 3000 hectáreas para el año 2021. Allí se pueden elaborar mapas de áreas quemadas a partir de imágenes Landsat o Sentinel 2 de las que también pueden obtenerse índices espectrales. Mediante GEE eligen la imagen más apropiada, complementan con información de mapas de vientos, índices de humedad del suelo, niveles freáticos, entre otros. Permitiendo exaltar la funcionalidad de GEE para el monitoreo de terrenos de difícil acceso.

5. MARCO NORMATIVO

A continuación, se hace una breve reseña de las principales normas que sustentan los estudios y análisis de coberturas terrestres apoyado en el uso de imágenes satelitales, sistemas de información geográfica y herramientas de análisis geoespacial. Siendo esta una actividad de gran importancia para conocer las dinámicas que allí confluyen y establecer acciones para su conservación en el marco del desarrollo sostenible.

Decreto 2811 de 1974: Por el cual se dicta el Código Nacional de Recursos Naturales Renovables y de Protección al Medio Ambiente. (Parques Nacionales Naturales de Colombia, 1974)

Resolución 0148 de 1974: Expedida por el entonces Ministerio de Agricultura, se aprobó el Acuerdo de la Junta Directiva del Instituto de Desarrollo de Recursos Naturales Renovables - INDERENA -, por medio del cual se delimitó y reservó, como área protegida el Parque Nacional Natural Los Nevados con “el objeto de conservar la flora, la fauna y las bellezas escénicas naturales.” (Parques Nacionales Naturales de Colombia, 2017).

Decreto 3570 de 2011: El cual establece que son funciones de la Dirección de Asuntos Ambientales Sectorial y Urbana, entre otras: “Apoyar la innovación, desarrollo y adaptación de las tecnologías para aprovechar sosteniblemente la oferta ambiental y prevenir, mitigar o corregir los impactos y efectos ambientales de las actividades socioeconómicas y gestionar, ante las autoridades competentes y el sector privado, la inversión de capital requerido”. (DNP, 2022)

Resolución 0393 de 2017: Por la cual se adopta el Plan de Manejo del Parque Nacional Natural Los Nevados. (Parques Nacionales Naturales de Colombia, 2017)

6. MARCO TEÓRICO

Teledetección: Corresponde a la técnica que posibilita la obtención o registro de información a distancia, dado que no se tiene contacto directo con los objetos localizados en la superficie terrestre. Lo anterior, requiere que se genere interacción de alguna índole entre el sensor y los objetos.

En cualquier sistema de teledetección es imperante la existencia de tres elementos, entre los que figuran el sensor, el objeto observado y el flujo energético, donde además las tres formas de obtener información a partir de un sensor remoto son la reflexión, emisión y emisión – reflexión (Chuvienco, 2008).

Espectro electromagnético: Se refiere a la disposición u organización de bandas de longitudes de onda o frecuencia, en las que la radiación electro – magnética exhibe similitud. Abarcando diversas longitudes de onda, que van desde las ondas cortas (como son los rayos X y rayos gamma) cuya unidad de medida más frecuente son micrómetros, hasta las ondas largas (como las tele-comunicaciones) con unidades de centímetros o metros (Chuvienco, 2008).

En el ámbito de la teledetección, existen una serie de bandas espectrales con alta implementación que merecen la pena destacarse, entre las que figuran:

Espectro visible: Su nombre hace alusión a la única banda perceptible por el ojo humano abarca desde los 0,4 hasta los 0,7 μm , en ella se identifican la banda azul (0,4 a 0,5 μm), verde (0,5 a 0,6 μm) y roja (0,6 a 0,7 μm).

Infrarrojo cercano, reflejado o fotográfico: Su identificación se da a partir de películas con emulsiones especiales, abarca desde los 0,7 a 1,3 μm , permitiendo identificar las masas vegetales.

Infrarrojo medio: Allí tienen lugar los procesos de reflexión de luz solar y emisión de la superficie terrestre, abarca desde los 1,3 a 2,5 μm para el infrarrojo de onda corta y de 3

a 5 μm para el infrarrojo medio, siendo la primera utilizada para fines de estimación de humedad y la segunda para la identificación de zonas de alta temperatura.

Infrarrojo lejano o térmico: Abarca desde los 8 a 14 μm abarcando la porción del espectro donde es posible registrar el calor originario de las cubiertas terrestres.

Micro-ondas: Son aquellas bandas con una longitud superior a 1 mm, cuyo interés radica en considerarse una energía transparente a la nubosidad.

Tipos de sensores: La clasificación de los sensores está determinada en función de la energía que implementan para adquirir la información, entre los que figuran los sensores pasivos como aquellos que requieren y dependen de la energía ambiental proveniente del sol, permitiendo la medición de la radiación electromagnética que refleja la superficie terrestre y atmosférica. Por su parte, los sensores activos son aquellos que, para la obtención de información, generan su propia interacción energética con el objetivo de estudio, recibiendo así la señal que rebota de la superficie terrestre. (Veneros et al., 2020)

Programa de Observación de la Tierra de la Unión Europea (Copérnico):
Corresponde a el programa de observación terrestre liderado por la Comisión Europea (CE) y la Agencia Espacial Europea (ESA), con miras a suministrar y facilitar información actualizada y precisa que permita la comprensión de las dinámicas existentes en la superficie.

Dentro del programa existen una serie de misiones Sentinel, dentro de las cuales se destaca Sentinel-2 para fines del presente trabajo, este genera imágenes de alta resolución facilitando información de la vegetación, la cobertura de agua, el uso, zonas costeras, entre otras. Su constelación consta de dos satélites gemelos o idénticos, denominados Sentinel 2A y Sentinel 2B dispuestos en la misma órbita y con distancia entre sí de 180° , con años de lanzamiento de 2015 y 2017 respectivamente (Agencia Espacial Europea (ESA), 2024).

Google Earth Engine: Plataforma computacional online del ambiente Google que desde 2010 brinda soporte en la gestión de grandes volúmenes de datos de forma eficiente, sin incurrir en gastos de máquina del usuario, dado que se apoya en los procesadores de Google Cloud (Gorelick et al., 2017; Perilla & Mas, 2020). Esta plataforma cuenta además con un extenso catálogo de archivos digitales, que se actualiza cada día y que permite la obtención, edición, cómputo y exportación de datos a partir de una interfaz de explorador y otra de editor de código mediante lenguaje de programación (Perilla & Mas, 2020).

Reducción: Este método de Google Earth Engine permite compilar grupos de datos y reducirlos a un único dato, que responde a un estadístico particular (mínimo, máximo, media, mediana, desviación estándar); y que permite así tratar los ruidos de las imágenes satelitales captadas en un momento particular (producidos por nubes, sombra, humo, niebla, reflejos, saltos de escaneo, etc) (Olmos S et al., 2023); (Teluguntla et al., 2018); (Vélez-Castaño et al., 2020).

Píxel: Es la unidad homogénea mínima que conforma un RASTER (imagen digital) y que contiene un único valor de la variable representada en esta matriz digital. Dichos píxeles poseen profundidades, que se relacionan con la cantidad de valores que pueden representar (ESRI, s.f.).

Nubosidad: Condición meteorológica relacionada con la presencia de nubes sobre la superficie terrestre. Tiene gran importancia en el campo del sensoramiento remoto satelital, dado que supone una restricción para la captura de información de la superficie para los sensores remotos pasivos (que funcionan con la reflectancia de la luz solar) y que por ende puede causar pérdida de información. (IDEAM, 2019).

Clasificación supervisada: Método mediante el cual se procesan RASTER con información visual de la superficie, y que mediante el entrenamiento de los algoritmos

clasificadores puede generar un nuevo RASTER con las clases ingresadas en el entrenamiento. Esta herramienta permite discretizar en clases zonales que facilitan la comprensión de la distribución de estas y su cuantificación (Corso, 2009).

Corine Land Cover: Metodología europea de clasificación que busca generar un inventario homogéneo de coberturas de la superficie terrestre, por medio de la interpretación visual de imágenes satelitales y la generación de bases de datos geográficas. Dicha metodología fue adaptada para Colombia por el IDEAM en 2010, mediante criterios jerárquicos y análisis de estados fisiológicos de altura y densidad de las coberturas vegetales (Ministerio del Medio Ambiente, 2010).

Composición de bandas: Proceso mediante el cual se toman diferentes zonas del espectro electromagnético empaquetadas en forma de bandas satelitales, y se combinan para generar colores que pueden corresponder con el color verdadero o con una amplia gama de "falsos colores" que facilitan la identificación de elementos en la superficie terrestre y es un apoyo para fotointerpretación y clasificación supervisada (Chuvieco, 2008).

7. METODOLOGÍA

7.1 Enfoque metodológico

De acuerdo con los objetivos planteados en el presente trabajo, el enfoque metodológico corresponde a la categoría de investigación aplicada, dado que no persigue el descubrimiento de leyes o principios científicos y se utilizará el conocimiento básico en pro de su aplicabilidad para el campo de las coberturas vegetales, particularmente en el procedimiento de clasificación supervisada y composición de imágenes mediante herramientas SIG entre las que figuran el software ArcMap de la suite ArcGIS, Google Earth, plataformas para visualización y descarga de imágenes satelitales como ESA Copernicus y Google Earth Engine para la descarga, análisis y gestión de información espacial. Bajo esta misma línea, este trabajo también presenta un enfoque cuantitativo, debido a que no se aceptan juicios de valor, sino que las apreciaciones se rigen por los resultados obtenidos a partir de algoritmos y su verificación se realiza mediante matrices que involucran porcentajes de aproximación.

7.2 Tipo de estudio

Teniendo en cuenta los diferentes procedimientos contemplados, se define como una investigación de tipo correlacional, puesto que a partir de los insumos de entrada, sea imágenes satelitales con bandas compuestas o imágenes con bandas modificadas por reducción, se realizará un análisis comparativo mediante el cual se determinará la incidencia en el rendimiento de resultados al ingresar al modelo píxeles de un solo momento vs píxeles asociados a varios momentos para fines de la clasificación supervisada de coberturas en la plataforma ArcGIS.

7.3 Área de estudio

El área de estudio seleccionada está delimitada por una geometría cuadrada, cuya área total es de 176.551,13 hectáreas (MAGNA Colombia CTM12), con jurisdicción en los

departamentos de Caldas, Quindío, Risaralda y Tolima (Figura 1). En su interior abarca el polígono correspondiente al Parque Nacional Natural Los Nevados (PNN Los Nevados), su declaratoria de área protegida se realizó en el marco del Decreto 2811 del 18 de diciembre de 1974 (Ministerio de Ambiente y Desarrollo Sostenible, 1974), con miras a la protección de la fauna y flora presente en los ecosistemas de alta montaña, donde además de su importancia en términos ambientales también lo posee en términos turísticos por la presencia de páramos y nevados.

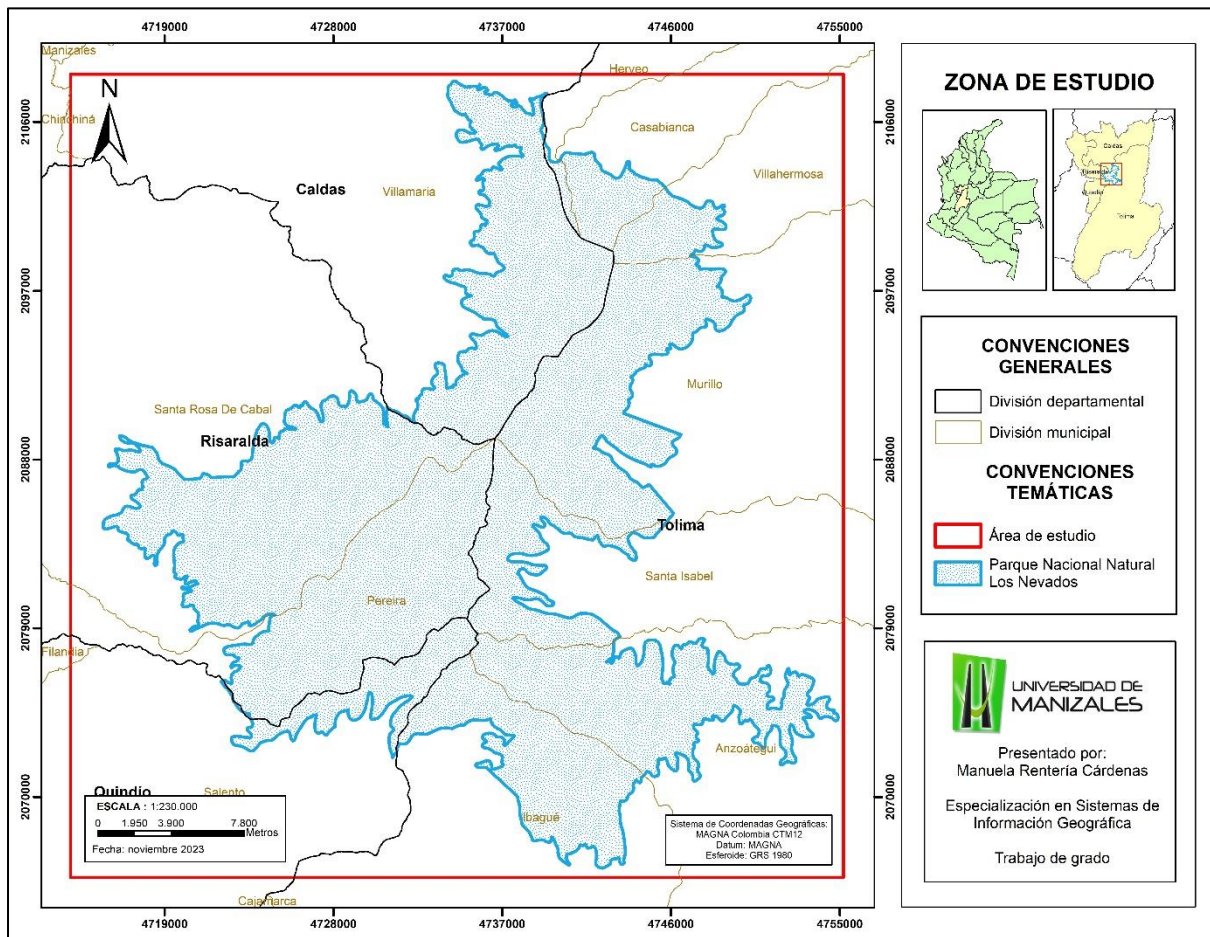


Figura 1: Localización del área de estudio.
Fuente: Elaboración propia a partir de productos ArcGIS.

7.4 Procedimiento

La metodología de este proyecto se concibió como un grupo de procesos concatenados y causales, que daban lugar a otros mediante la adición de decisiones y datos procedentes de diferentes herramientas. Adicionalmente, se tendrían procedimientos

que avanzaban de forma paralela a las fases del proyecto como la consulta bibliográfica y la consulta a expertos.

Se contemplaron dos fases del proyecto, siendo la primera de ellas relacionada a los objetivos 1 y 2, y la segunda relacionada con los objetivos 3 y 4.

Como se observa en la figura 2, para la primera fase se tiene como punto de partida una misma fuente de datos, pero se tomaron caminos distintos para llegar a un mismo objetivo, correspondiente a la consolidación de una base de datos procesados. Estos datos a su vez tienen diferentes grados de corrección, que va desde los datos RAW sin corregir, pasando por los datos corregidos en lo alto de la atmósfera (TOA) y los corregidos en superficie (BOA).

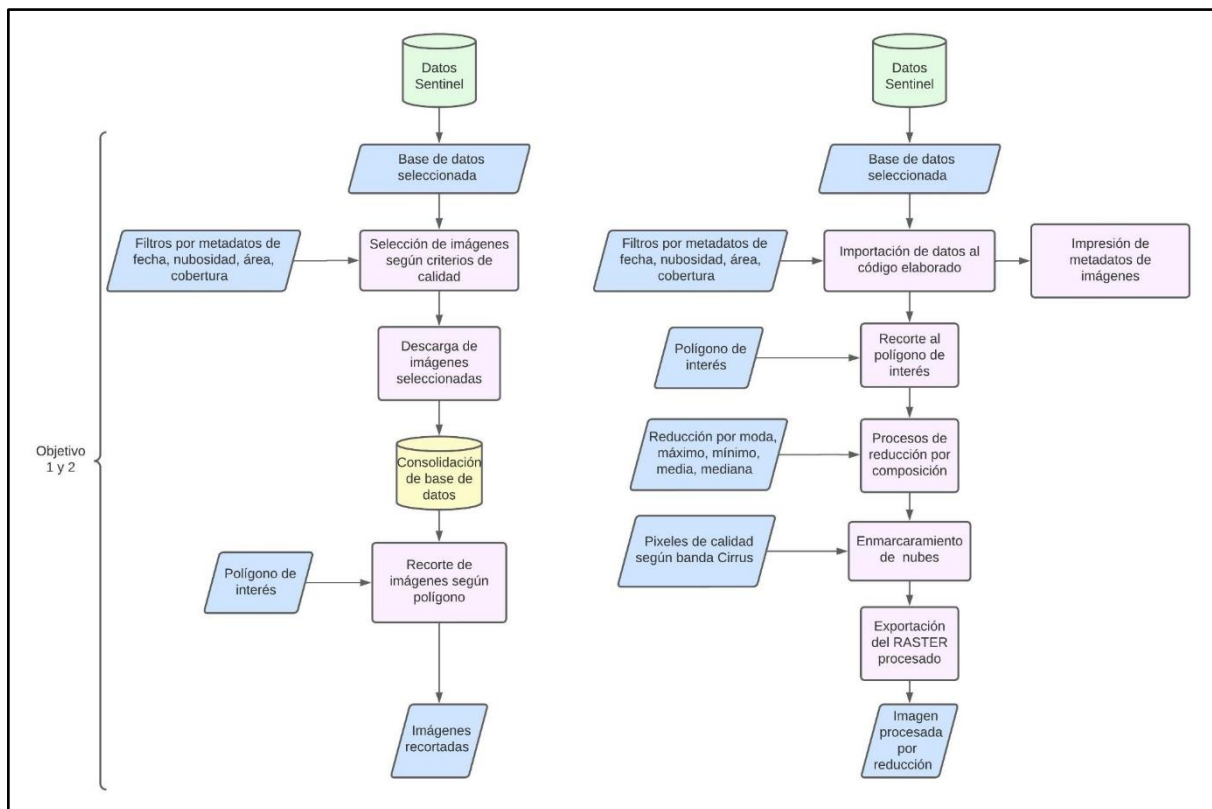


Figura 2: Fase 1 del proyecto, con objetivos 1 y 2 avanzando de forma paralela.
Fuente: Elaboración propia.

En la primera fase de esta investigación se avanzó de forma paralela en los primeros dos objetivos, cada uno comprendía actividades específicas, que se detallan a continuación:

Objetivo 1: Revisar a partir de criterios de calidad las imágenes satelitales de la zona de estudio disponible en las plataformas de acceso libre.

Actividades

1.1 Selección de la base de datos a trabajar

Para la búsqueda de imágenes Sentinel existen diferentes bases de datos, es importante mencionar que para la fecha en la cual se realizó dicha revisión (octubre de 2023), estas imágenes se encontraban disponibles en la plataforma de la ESA Copernicus, donde están migrados todos los datos que generan los sensores gemelos S2A y S2B, que son equivalentes entre sí y muestrean de forma alternada estos se sitúan en la misma órbita y están separados 180° entre sí. Actualmente dicha información se encuentra alojada en Copernicus Data Space.

1.2 Selección de las imágenes según criterios de calidad y disponibilidad

Para ambas plataformas de satélite, Sentinel A y Sentinel B, el principal criterio de calidad designado fue el relacionado a la nubosidad y a partir de este se escogió la mejor imagen disponible para la zona de trabajo.

Se evaluó la información obtenida ingresando múltiples rangos variables de nubosidad, el máximo valor permitido para Sentinel 2A fue del 45% y para Sentinel 2B fue del 55%. Las imágenes que resultaron llamativas porque se percibieron relativamente “despejadas” para el contexto geográfico y climático, fueron de los años 2020 y 2023. Finalmente realizando un sondeo por toda el área y priorizando la de interés, se eligió la imagen de Sentinel 2A del mes de febrero del año 2020, cuyo rango de nubosidad en la escena es máximo del 10% y se identificó con el código “S2A_MSIL2A_20200218T15

2631_N0214_R025_T18NVL_20200218T193733” (Figura 3).

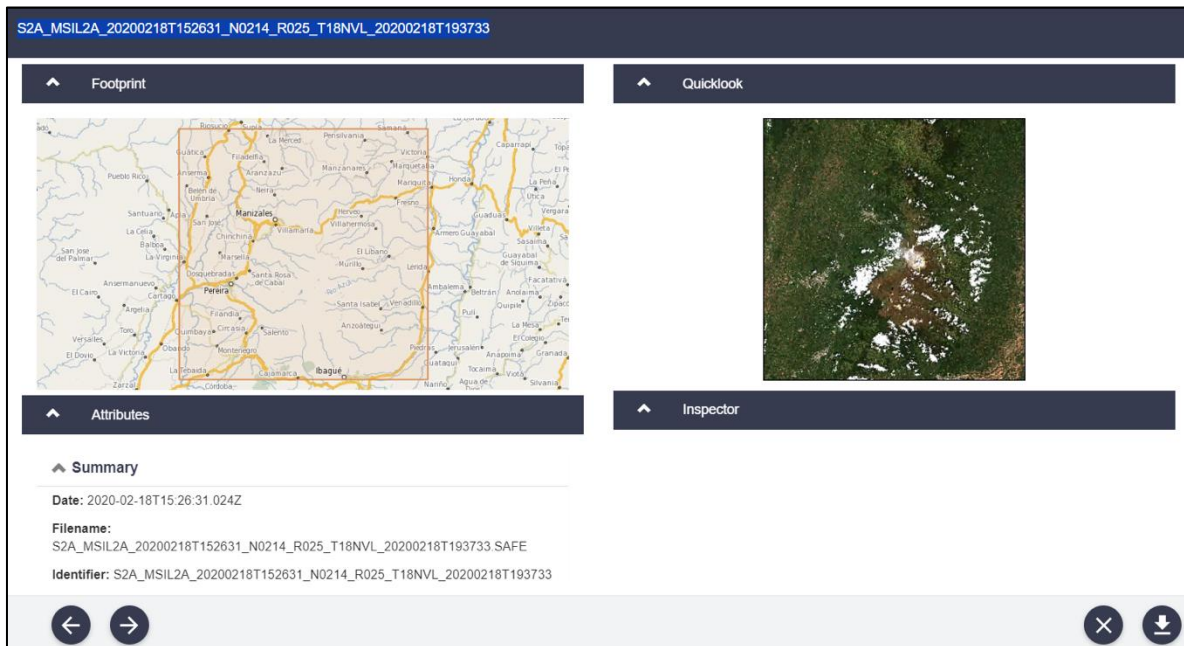


Figura 3: Imagen Sentinel 2 seleccionada, con nubosidad del 0 – 10%.
Fuente: Obtenido de la plataforma ESA Copernicus.

Sumado al criterio del experto, la selección de esta imagen se confirmó mediante el algoritmo de “Sorting” (sorteo) implementado en GEE, donde se especificó un rango de nubosidad y el programa escogió de la biblioteca de datos la que mejor cumplía, coincidiendo con la imagen escogida anteriormente y dejando en claro que esta era la imagen que mejor se adaptaba a las condiciones requeridas (Figura 4 y figura 5).

```
New Script *
16 var outline = ee.Image().byte().paint({featureCollection: geometry,
17 Map.addLayer(outline, {palette: ['red']}, 'PNNLN');
18
19
20
21 //Llamamos a la colección de Sentinel 2
22 var S2 = ee.ImageCollection('COPERNICUS/S2')
23 .filterDate('2016-01-01', '2023-10-28')
24 .filter(ee.Filter.lt('CLOUDY_PIXEL_PERCENTAGE', 100))
25 .filterBounds(geometry)
26
27 //Mostramos sus metadatos
28 print (S2);
29
30 Map.centerObject(S2, 10);
31
```

Figura 4: Código asociado al proceso “de sorting” o sorteo en la plataforma GEE.
Fuente: Obtenido del editor de código de la plataforma Google Earth Engine (GEE).

```
24 .filter(ee.Filter.lt('CLOUDY_PIXEL_PERCENTAGE',100))
25 .filterBounds(geometry)
26
27 //Mostramos sus metadatos
28 print(S2);
29
30 Map.centerObject(S2, 10);
31 var sorteo = S2.sort('CLOUDY_PIXEL_PERCENTAGE');
32 var sentinel2 = ee.Image(sorteo.first());
33 Map.addLayer (sentinel2, {bands: ['B8']}, 'S2');
34
35
```

The Inspector panel shows the following objects:

- PNLNL: Image (1 band)
 - type: Image
 - bands: List (1 element)
 - properties: Object (1 property)
- S2: Image COPERNICUS/S2/20200218T1526...
 - type: Image
 - id: COPERNICUS/S2/20200218T152631_20...
 - version: 1582128758637604
 - bands: List (16 elements)
 - properties: Object (66 properties)

Figura 5: Código asociado al proceso “de sorting” o sorteo en la plataforma GEE. Fuente: Obtenido del editor de código de la plataforma Google Earth Engine (GEE).

1.3. Descarga de imágenes seleccionadas

Se realizaron las solicitudes de descarga y estas se llevaron a la terminal, pues este objetivo se trabajó desde el software licenciado (ArcGIS), que implicaba descargar y visualizar las imágenes en el computador. En algunos casos dichas solicitudes debían hacerse desde el archivo de datos, pues algunas escenas captadas por los satélites no cumplieron con los criterios de calidad mínimos establecidos para toda el área, pero eventualmente para el área de interés sí la cumplían, por lo que había que solicitar su extracción del archivo, lo que podía tardar un poco más de tiempo en obtener las imágenes.

1.4 Consolidación de base de datos

Se consolidó el lugar donde se alojarán los archivos obtenidos a través de la plataforma de la ESA y se organizaron para su posterior uso en el software escogido.

1.5 Recorte de imágenes según el polígono.

Este paso, de carácter repetitivo, consistía en tomar todas las bandas multiespectrales que se involucraron en los procesos posteriores de esta investigación y en recortarlas (con la herramienta clip) a la forma predefinida del área de estudio, según el polígono de interés que se escogió. En este punto se tuvieron las imágenes listas para ser

empleadas en los pasos subsiguientes que conducían a alcanzar los demás objetivos específicos y posteriormente el objetivo general de este proyecto.

Objetivo 2: Generar mediante algoritmos de reducción en la herramienta Google Earth Engine, imágenes satelitales con variaciones de los filtros de metadatos y calidad relacionada con fechas y porcentajes de nubosidad.

Actividades

2.1 Selección de la base de datos a trabajar

Una vez más fue necesario escoger la base de datos de Sentinel con la que se trabajaría, y dicha escogencia ya no se realizó en la plataforma de la ESA Copernicus del objetivo 1, sino en la biblioteca de repositorios (Data Catalog) de Google Earth Engine donde se desarrolló este objetivo. La biblioteca elegida corresponde a la de Sentinel-2 MSI armonizada (Harmonized Sentinel-2 MSI) donde se estandarizan los valores espectrales y se llevan al rango común de las imágenes Sentinel (al dividirlos entre 1000). Los datos tienen la misma resolución espectral, espacial y temporal del resto de datos capturados en las misiones Sentinel-2 (Figura 6).

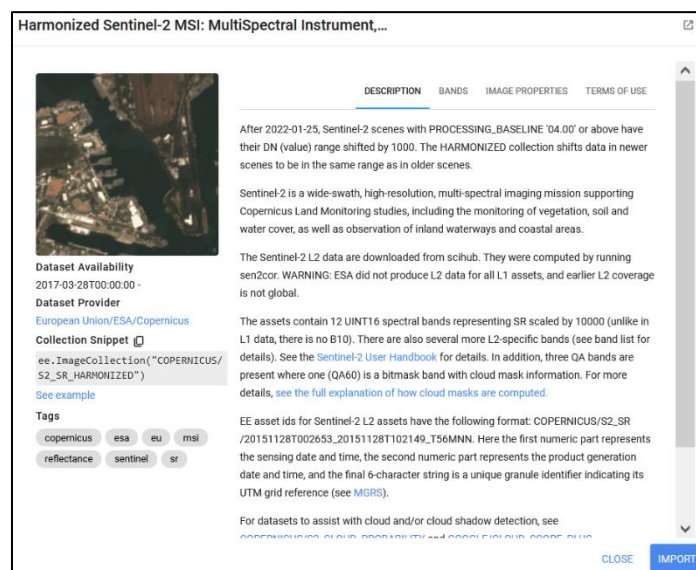


Figura 6: Biblioteca Sentinel-2 MSI armonizada.
Fuente: Obtenido de la biblioteca de imágenes de la plataforma GEE.

2.2 Importación de imágenes al código elaborado

Una vez seleccionada la fuente de datos de Sentinel, se generó el código en lenguaje JAVA desarrollado para la gestión de datos. Allí se aplicaron los filtros de nubosidad, área y fecha, y se obtuvieron los metadatos de las imágenes a emplear, que cumplieran con los filtros especificados (Figura 7).

```
Composicion *
Imports (1 entry)
1
2 //Se crea la máscara de nubes
3 - function maskS2clouds(image) {
4   var qa = image.select('QA60');
5
6   // Bits 10 and 11 are clouds and cirrus, respectively.
7   var cloudBitMask = 1 << 10;
8   var cirrusBitMask = 1 << 11;
9
10  // Both flags should be set to zero, indicating clear conditions.
11  var mask = qa.bitwiseAnd(cloudBitMask).eq(0)
12    .and(qa.bitwiseAnd(cirrusBitMask).eq(0));
13
14  return image.updateMask(mask).divide(10000);
15 }
16
17 //Llamamos a la colección de Sentinel 2
18 var S2 = ee.ImageCollection('COPERNICUS/S2_SR_HARMONIZED')
19   .filterDate('2016-01-01','2023-10-28')
20   .filterBounds(geometry)
21   .filter(ee.Filter.lt('CLOUDY_PIXEL_PERCENTAGE',20))
22   .map(maskS2clouds);
23 //Mostramos sus metadatos
24 print(S2);
25
26 //Obtenemos el valor de su estadístico
27 var S2median = S2.median();
28 Map.addLayer(S2median, {min:0.0, max: 0.3, bands: ['B4','B3','B2']}, 'S2_clip_2');
29
30 var S2clip=S2median.clip(geometry);
31 Map.centerObject(S2, 10);
32 Map.addLayer(S2clip, {min:0.0, max: 0.3, bands: ['B11','B8A','B4']}, 'S2_clip');
33 Map.addLayer(S2clip, {min:0.0, max: 0.3, bands: ['B4','B3','B2']}, 'S2_clip_2');
34 Map.addLayer(S2clip, {min:0.0, max: 0.3, bands: ['B8A','B4','B3']}, 'S2_clip_3');
35
36 //Índices tierra agua
37 //NDWI Mc Feeters
38 var NDWIMcF = S2clip.normalizedDifference(['B3','B6']);
39 Map.addLayer(NDWIMcF,{min:-1,max:1,'palette':'FF0000,0000FF'},'NDWIMcFeeters_2021_S2');
40
41 //Exportamos el mosaico
42 Export.image.toDrive({image:S2median.select(['B1','B2','B3','B4','B5','B6','B7','B8','B8A','B9','B11','B12']), description: 'Mosaico_PNNLN_S2', region: geometry, scale: 10});
43 Export.image.toDrive({image:S2clip.select(['B11','B8A','B4']), description: 'Mosaico2_PNNLN_S2', region: geometry});
44 Export.image.toDrive({image:S2clip.select(['B8A','B4','B3']), description: 'Mosaico3_PNNLN_S2', region: geometry});
```

Figura 7: Código desarrollado en lenguaje JAVA.

Fuente: Obtenido del editor de código de la plataforma Google Earth Engine (GEE).

Como se aprecia en el código, los filtros de fecha y nubosidad empleados tienen los siguientes factores: el producto resultante de la implementación del algoritmo de reducción, se construyó a partir de imágenes obtenidas entre el 1 de enero de 2016 y el 28 de octubre de 2023 (fecha en la que se elaboró el código); escenas con menos del 20% de cobertura nubosa para toda la región de captura. Se agregó un filtro de límites para enfocar el código a la zona de interés, predefinida en una geometría cargada al procesador de código (Code editor) de Google Earth Engine (Figura 8).

```
//Llamamos a la colección de Sentinel 2
var S2 = ee.ImageCollection('COPERNICUS/S2_SR_HARMONIZED')
  .filterDate('2016-01-01','2023-10-28')
  .filterBounds(geometry)
  .filter(ee.Filter.lt('CLOUDY_PIXEL_PERCENTAGE',20))
  .map(maskS2clouds);
//Mostramos sus metadatos
```

Figura 8: Filtros de fecha y nubosidad empleados en la consola de GEE.

Fuente: Obtenido del editor de código de la plataforma Google Earth Engine (GEE).

Las escenas que cumplieron con los filtros indicados (Figura 9) se obtuvieron a través de los metadatos de la misma. El código de la escena aparece en el item `system.index`:

```
* ImageCollection COPERNICUS/S2_SR_HARMONIZED (9 elements)
type: ImageCollection
id: COPERNICUS/S2_SR_HARMONIZED
version: 1705274735583590
bands: []
*features: List (9 elements)
  *0: Image (23 bands)
    type: Image
    bands: List (23 elements)
    properties: Object (1 property)
    system:index: 20181225T152631_20181225T153106_T18NVL
  *1: Image (23 bands)
    type: Image
    bands: List (23 elements)
    properties: Object (1 property)
    system:index: 20190104T152631_20190104T153031_T18NVL
  *2: Image (23 bands)
    type: Image
    bands: List (23 elements)
    properties: Object (1 property)
    system:index: 20190822T152641_20190822T152641_T18NVL
  *3: Image (23 bands)
    type: Image
    bands: List (23 elements)
    properties: Object (1 property)
    system:index: 20190827T152649_20190827T152643_T18NVL
  *4: Image (23 bands)
    type: Image
    bands: List (23 elements)
    properties: Object (1 property)
    system:index: 20200104T152639_20200104T153054_T18NVL
  *5: Image (23 bands)
    type: Image
    bands: List (23 elements)
    properties: Object (1 property)
    system:index: 20200109T152631_20200109T152633_T18NVL
  *6: Image (23 bands)
    type: Image
    bands: List (23 elements)
    properties: Object (1 property)
    system:index: 20200218T152631_20200218T152633_T18NVL
  *7: Image (23 bands)
    type: Image
    bands: List (23 elements)
    properties: Object (1 property)
    system:index: 20230128T152639_20230128T152835_T18NVL
  *8: Image (23 bands)
    type: Image
    bands: List (23 elements)
    properties: Object (1 property)
    system:index: 20230826T152639_20230826T153029_T18NVL
```

Figura 9: Escenas que cumplen con los requisitos de fecha y nubosidad especificados. Fuente: Obtenido del editor de código de la plataforma Google Earth Engine (GEE)

2.3 Recorte al polígono de interés

Se hizo además un recorte de toda la información al polígono de interés, con el objetivo de no procesar zonas que no corresponden con el proyecto, ahorrando volúmenes importantes de trabajo en la nube de Google Cloud (GCP).

2.4 Proceso de reducción

Este paso es uno de los más importantes en este objetivo, consistía en recopilar toda la información de diferentes fechas, recortada previamente, superponerla, y escoger con base en el rendimiento cuál de los estadísticos disponibles (media, mediana, mínimo, máximo) genera el mejor producto libre de ruido. El algoritmo con base en el estadístico

elegido tomó todos los píxeles de un mismo lugar captados en diferentes fechas y generó un único producto que es una composición reducida de toda la información, y con altas posibilidades de no captar elementos generadores de ruido. El estadístico seleccionado fue la mediana, dado que esta medida de tendencia central que corresponde a la mitad de los datos y que es menos susceptible a los valores extremos en la parte inferior o superior de los datos, que pueden llegar a ser outliers como nubes muy densas o sombras de nubes muy oscuras (Figura 10).

```
var S2clip=S2median.clip(geometry);
Map.centerObject(S2, 10);
Map.addLayer (S2clip, {min:0.0, max: 0.3, bands: ['B11','B8A','B4']}, 'S2_clip');
Map.addLayer (S2clip, {min:0.0, max: 0.3, bands: ['B4','B3','B2']}, 'S2_clip_2');
Map.addLayer (S2clip, {min:0.0, max: 0.3, bands: ['B8A','B4','B3']}, 'S2_clip_3');
```

Figura 10: Código relacionado con el recorte de la información y configuración del estadístico.
Fuente: Obtenido del editor de código de la plataforma Google Earth Engine (GEE).

2.5 Enmascaramiento de nubes

Dado que aún con la reducción del paso anterior eventualmente algunas nubes queden en el RASTER, era necesario hacer un enmascarado de nubes (Figura 11), que consiste en elegir píxeles de la imagen resultante del paso anterior dentro de un rango de valores dado para la banda correspondiente a calidad por limpieza de nubes (QA60), y escogiendo solo píxeles que estén libres de nubes, y eliminando aquellos píxeles que no lo estén. Este algoritmo de limpieza es estándar para las imágenes de Sentinel 2 y viene incluido en el repositorio empleado. Esto garantiza un producto final con las mejores condiciones posibles de limpieza de ruido.

```
//Se crea la máscara de nubes
function maskS2clouds(image) {
  var qa = image.select('QA60');

  // Bits 10 and 11 are clouds and cirrus, respectively.
  var cloudBitMask = 1 << 10;
  var cirrusBitMask = 1 << 11;

  // Both flags should be set to zero, indicating clear conditions.
  var mask = qa.bitwiseAnd(cloudBitMask).eq(0)
    .and(qa.bitwiseAnd(cirrusBitMask).eq(0));

  return image.updateMask(mask).divide(10000);
}
```

Figura 11: Código relacionado con el enmascaramiento de las nubes en la consola de GEE.
Fuente: Obtenido del editor de código de la plataforma Google Earth Engine (GEE).

2.6 Exportación de RASTER procesado.

El producto obtenido en el paso anterior finalmente se exporta a Google Drive para poder ser descargado (Figura 12) y continuar el resto de la gestión de los datos en la terminal mediante el software licenciado.

```
//Exportamos el mosaico  
Export.image.toDrive({image:S2median.select("B1", "B2", "B3", "B4", "B5", "B6", "B7", "B8", "B8A", "B9", "B11", "B12"), description: 'Mosaico_PNNLN_S2', region: geometry, scale: 10});  
Export.image.toDrive({image:S2clip.select("B11", "B8A", "B4"), description: 'Mosaico2_PNNLN_S2', region: geometry});  
Export.image.toDrive({image:S2clip.select("B8A", "B4", "B3"), description: 'Mosaico3_PNNLN_S2', region: geometry});
```

Figura 12: Código relacionado con la exportación del producto obtenido.
Fuente: Obtenido del editor de código de la plataforma Google Earth Engine (GEE).

En este punto, finalizados los objetivos 1 y 2, se tienen dos bases de datos con datos listos para proceder a realizar los objetivos 3 y 4. Como se observa en la figura 13, se partía de los datos obtenidos en la fase 1, y replicando el mismo proceso para ambos consistente en la clasificación supervisada de los mismos. Luego, se validaron los resultados obtenidos por ambas vías y se compararon los resultados mediante métodos adecuados para ello.

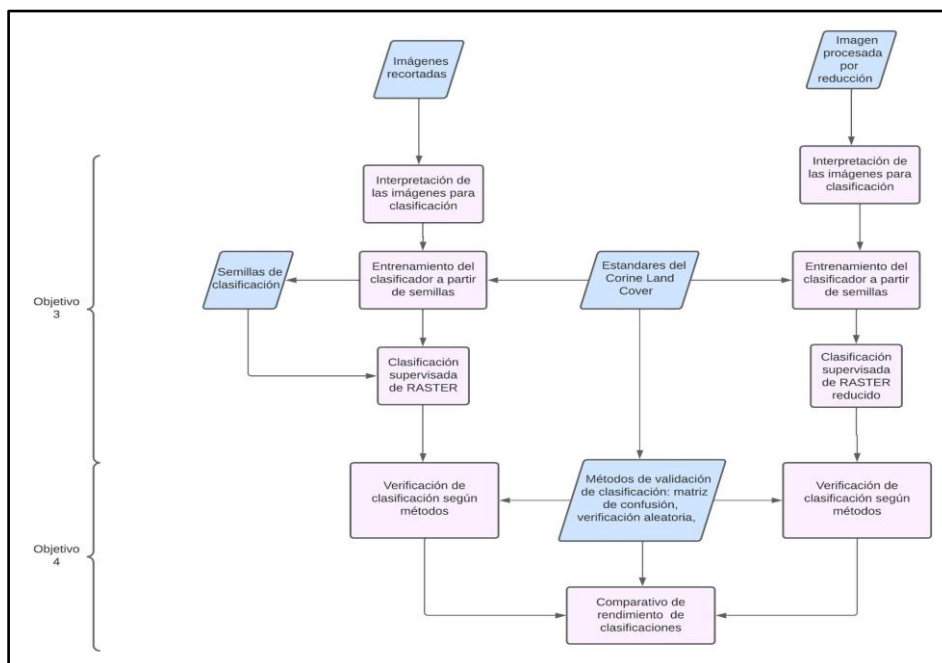


Figura 13: Fase 2 del proyecto, con objetivos 3 y 4 avanzando de forma paralela en la primera parte y de forma conjunta al final.
Fuente: Elaboración propia.

Como en la fase anterior, cada objetivo tiene sus respectivas actividades, que se detallan a continuación:

Objetivo 3: Elaborar una clasificación supervisada de cobertura vegetal a partir de imágenes satelitales con bandas compuestas sin modificación e imágenes satelitales con bandas modificadas por reducción para obtener mapas de coberturas de la zona de interés.

Actividades

3.1 Interpretación de las imágenes para clasificación

Se desplegaron en ArcMap de ArcGIS los productos obtenidos por las diferentes vías en la fase anterior (objetivos 1 y 2) y se hizo una interpretación de lo que se observa, para definir clases de clasificación.

La imagen en color real del área de estudio, obtenida de combinar las bandas roja, verde y azul recortadas al tamaño del polígono (Figura 14), da cuenta de un área de tonalidades verdes de diferente intensidad (verdes oscuros asociados a vegetaciones densas y verdes claros a vegetaciones bajas), con amplias zonas en color café de diferente tonalidad (tonalidades oscuras en las zonas de transición entre la alta montaña y los picos nevados, donde la tonalidad cambia a clara), una zona en color grisáceo, cercana a la fumarola activa del volcán Nevado del Ruíz y que refleja ya sea ceniza caída y acumulada, o una suerte de esclarecimiento del fondo café por la refracción de los gases de la fumarola. Se destaca el color blanco asociado a la presencia de nubes en los bordes del área protegida, siendo especialmente densa en la parte noroccidental; y a los casquetes de nieve de los picos más elevados. Debido a la posición del sol al momento de paso del satélite, con fuente de luz en aproximadamente a 130° de azimuth y 65° de altitud, lo que proyecta sombras en dirección noroccidental. El color de estas sombras es muy parecido al de los cuerpos de agua de alta profundidad, con tonalidades negras; teniéndose también algunos de tonalidades azul clara que denotan alta transparencia y baja profundidad. Finalmente, se

aprecia el trazado urbano de una población con techos color café y ocre en la zona suroccidental del polígono de interés.

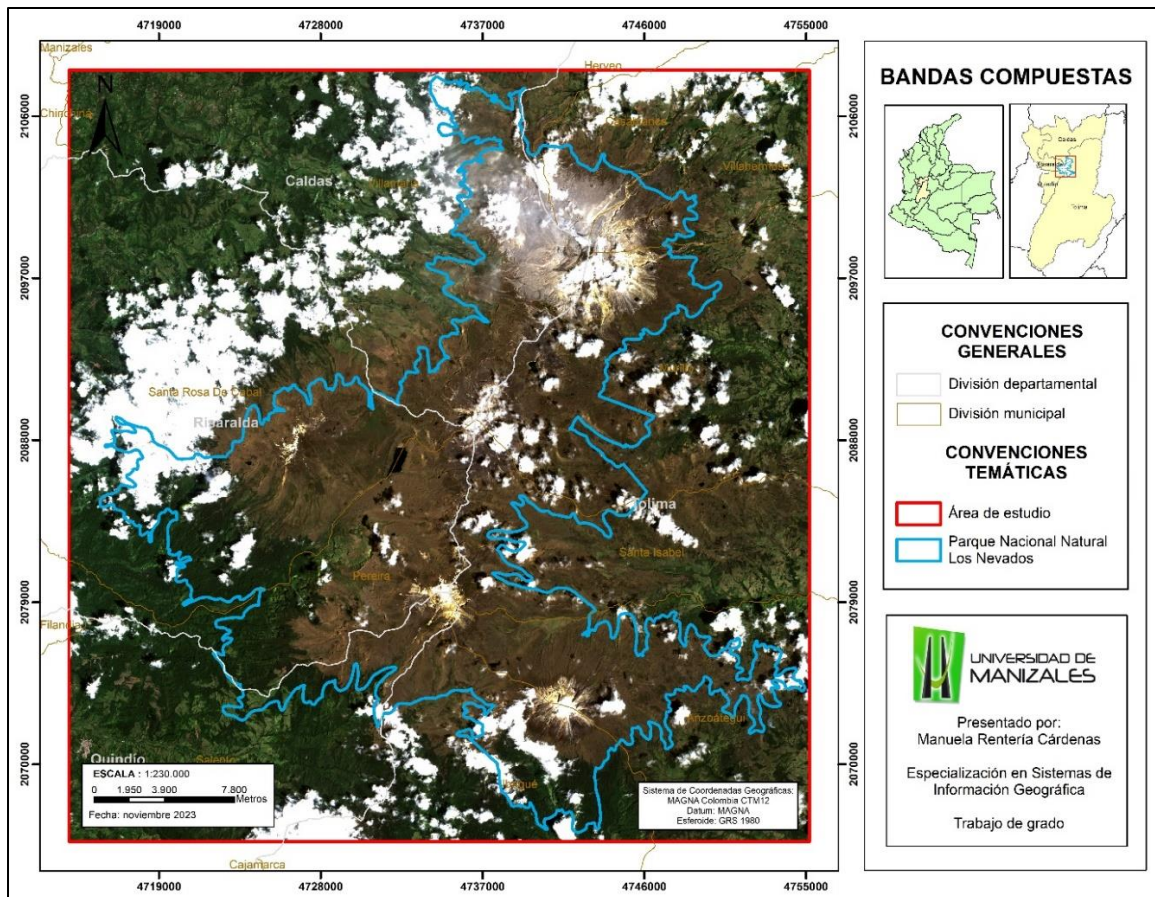


Figura 14: Imagen Sentinel 2A obtenida a partir de extraer del ráster original la información correspondiente al área de estudio.

Fuente: Elaboración propia a partir de productos ArcGIS.

Se aprecia en la figura 15 una escena compuesta con pixeles de múltiples momentos donde aún se encuentran algunas nubes que demuestran la persistencia de nubosidad para esas zonas específicas, pues el valor de la mediana correspondió a esa condición nubosa. Además, aparecen muchas menos sombras de nubes que en la imagen de un momento específico, fruto de la composición realizada. Por demás, las clases corresponden en alta medida a la imagen de un momento específico denotando baja variabilidad espacial en la zona al menos para el rango temporal considerado.

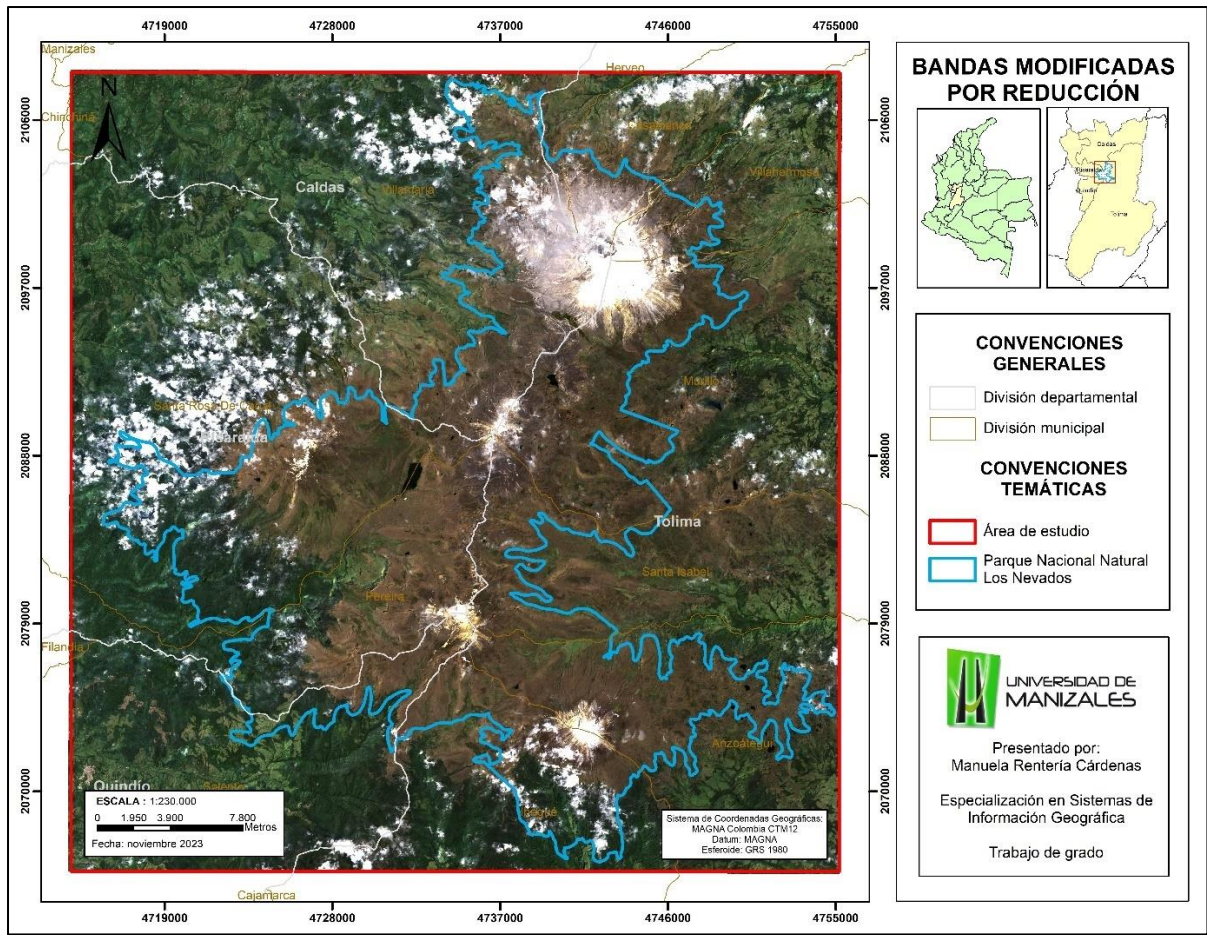


Figura 15: Mosaico de Sentinel 2 para un periodo de 2018 hasta 2023 generado mediante algoritmos en la plataforma GEE.
Fuente: Elaboración propia a partir de productos ArcGIS.

A partir de un sondeo por ambas imágenes, se lograron identificar un total de 10 clases de coberturas, entre las que figuran: cuerpo de agua; nieve; nube; periglaciario; roca; sombra; suelo desnudo; tejido urbano; vegetación densa; vegetación baja. En un comparativo entre ambas imágenes se aprecia que no hay variaciones importantes en la distribución de las clases, siendo la de nubes, sombras y la de nieve; las que manifiestan mayor variabilidad, debido al estadístico empleado en la reducción. Para estas clases, especialmente la de sombras, por la notable reducción en el mosaico obtenido por el algoritmo, habría que tomar ciertas medidas para garantizar posibilidad de comparación entre los dos productos.

Las herramientas implementadas para su identificación y características son descritas en la actividad 3.2.

3.2 Entrenamiento del clasificador a partir de semillas

En este punto se realizó para cada los RASTER el entrenamiento del clasificador en el software ArcGIS, mediante la colocación ordenada de polígonos (denominados muestras de capacitación o semillas) que luego servirían para obtener una clasificación supervisada lo más acorde posible con lo que se observa en las imágenes empleadas. Para la obtención de resultados precisos, pues requiere disponer gran cantidad de pequeñas semillas para lograr una correcta discriminación de las clases escogidas en el punto anterior.

Es importante mencionar, que para este ejercicio se implementaron semillas cuyo principal requisito es presentar correspondencia para ambas imágenes, es decir, que tanto en la imagen de Sentinel del año 2020 como en la imagen de tipo mosaico generada en GEE, la clase a la cual se hace alusión estuviera presente. Lo anterior, con el fin de mantener parámetros y condiciones iniciales que permitan la comparación entre ambos resultados y/o productos. Con esta estrategia se pretendía eliminar la variabilidad que pudiese existir por el entrenamiento del modelo para cada imagen, que acarrea errores propios de la experticia y el grado de dedicación del usuario que la ejecuta. De esta forma, con una sola distribución de polígonos, se garantizaba que la variabilidad de la clasificación supervisada se asociaba directamente al agrupamiento en clústeres que hace directamente el software para los valores espectrales de los productos sobre los que se colocaron las semillas.

Después de haber realizado el barrido o sondeo por toda la imagen y tener claridad de las clases presentes, se dispusieron diferentes semillas, cuyos nombres y número de píxeles se relacionan en la figura 16.

ID	Class Name	Value	Color
1	Cuerpo de agua	1	Blue
2	Nieve	2	Light Blue
3	Nube	3	White
4	Periglaciár	4	Yellow
5	Roca	5	Brown
6	Sombra	6	Red
7	Suelo desnudo	7	Orange
8	Tejido urbano	8	Magenta
9	Vegetación baja	9	Light Green
10	Vegetación densa	10	Dark Green

Figura 16: Semillas o muestras de entrenamiento dispuestas para fines de clasificación supervisada.
Fuente: Elaboración propia a partir de productos ArcGIS.

Cuerpos de agua: Se distribuyen en diferentes zonas del área de estudio, algunos ejemplares importantes están en cercanías al nevado de Santa Isabel, como lo son la Laguna del Otú, Laguna Verde, entre otras (Figura 17).

Para su reconocimiento se requirió de herramientas como el cálculo de índices, particularmente el NDWI el cual indica las zonas con alto contenido de humedad. Dicho índice se obtuvo por la vía de código en Google Earth Engine (para el mosaico) y por álgebra de mapas en ArcGIS (para la imagen de un solo momento). Así mismo fue útil implementar combinaciones como el infrarrojo, donde el agua se observa en tonalidades oscuras.

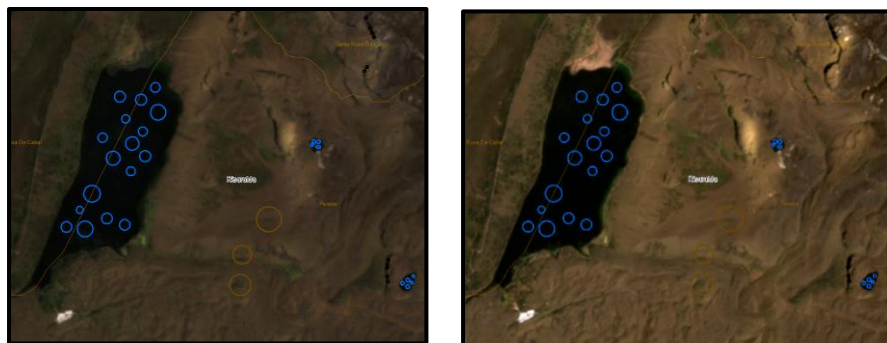


Figura 17: Semillas implementadas en la identificación de la clase “cuerpos de agua”. Imagen Mosaico (Izquierda) e Imagen S2A 2020 (Derecha), combinación en color real (B4+B3+B2).
Fuente: Elaboración propia a partir de productos ArcGIS.

Tejido urbano: Esta clase se puede encontrar por fuera del polígono del PNN Los Nevados, el centro poblado más importante corresponde al municipio de Salento y pueden

observarse otros caseríos en el sector noroccidental del polígono de estudio (Figura 18 y figura 19). El tejido urbano resulta particularmente complejo de identificar para el modelo especialmente en Colombia, dado que la gama de colores que involucra esta clase es amplia, traduciendo así la diversidad de materiales con las que se construyen los tejados.



Figura 18: Semillas implementadas en la identificación de la clase “tejido urbano”. Imagen Mosaico (Izquierda) e Imagen S2A 2020 (Derecha), combinación en color real (B4+B3+B2).
Fuente: Elaboración propia a partir de productos ArcGIS.



Figura 19: Municipio de Salento, Quindío representando la clase de tejido urbano.
Fuente: Elaboración propia a partir de productos de Google Earth.

Vegetación densa y vegetación baja: Estas clases se encuentran distribuidas ampliamente por el polígono de estudio (Figura 20 y figura 21). Para su identificación además de hacer uso de combinaciones como el infrarrojo, donde la vegetación densa se observaba de color marrón oscuro y la vegetación baja de color rosado de color. También se utilizó el índice NDVI el cual está asociado al verdor o salud de la vegetación, por lo que para la vegetación densa presenta valores más elevados que para la vegetación baja. Dicho

índice se obtuvo por la vía de código en Google Earth Engine (para el mosaico) y por álgebra de mapas en ArcGIS (para la imagen de un solo momento).

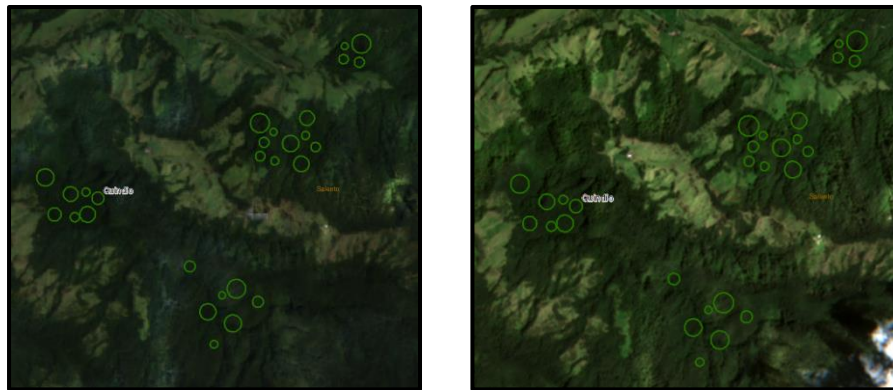


Figura 20: Semillas implementadas en la identificación de la clase “vegetación densa”. Imagen Mosaico (Izquierda) e Imagen S2A 2020 (Derecha), combinación en color real (B4+B3+B2).
Fuente: Elaboración propia a partir de productos ArcGIS.

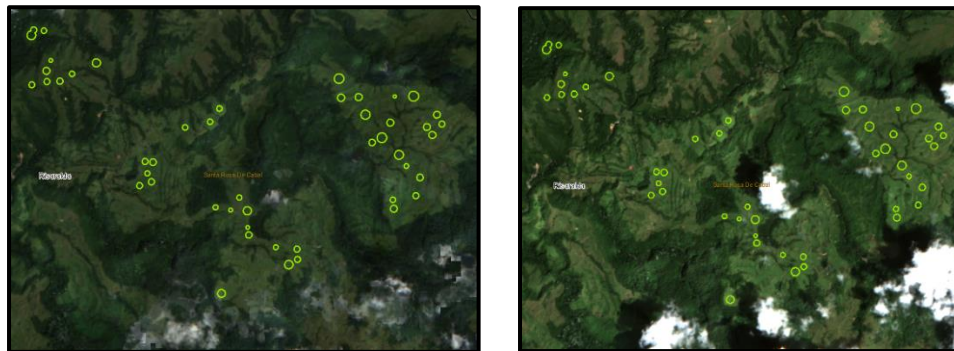


Figura 21: Semillas implementadas en la identificación de la clase “vegetación baja”. Imagen Mosaico (Izquierda) e Imagen S2A 2020 (Derecha), combinación en color real (B4+B3+B2).
Fuente: Elaboración propia a partir de productos ArcGIS.

Periglaciár: Esta clase también conocida como desierto periglaciár, se encuentra distribuida en las zonas circundantes a los glaciares donde eventualmente ocurren periodos de nevadas (Figura 22). Se caracteriza principalmente, por sus tonalidades amarillentas en la combinación de color real y tonalidades rosadas en la combinación de falso color, un importante ejemplar de esta clase de cobertura corresponde al Paramillo del Quindío.

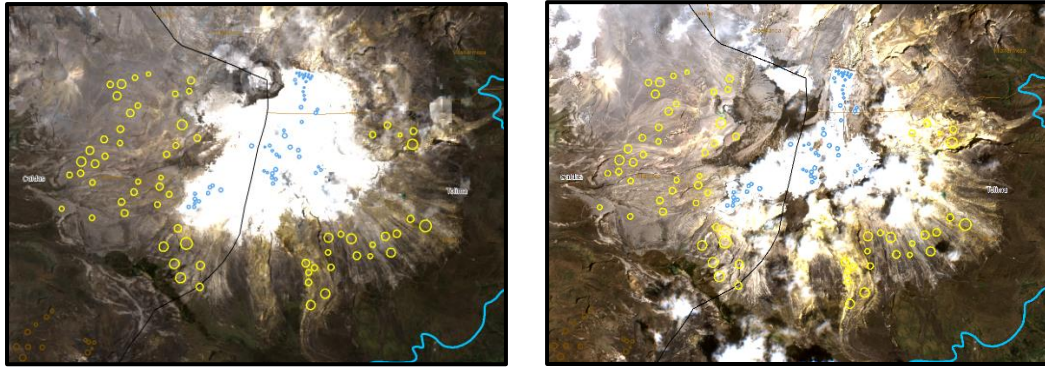


Figura 22: Semillas implementadas en la identificación de la clase “periglacial”. Imagen Mosaico (Izquierda) e Imagen S2A 2020 (Derecha), combinación en color real (B4+B3+B2).
Fuente: Elaboración propia a partir de productos ArcGIS.

Roca: Esta clase de cobertura se distribuye en las proximidades a las clases de desierto periglacial y nieve (Figura 23), donde no hay desarrollo de vegetación, por el contexto geológico de la zona, las características litológicas obedecen a un origen volcánico. Eventualmente en las zonas de roca pueden existir nevadas o caída de material proveniente de la actividad volcánica, por lo que las características del suelo allí presentes no permiten el desarrollo de una capa de vegetación significativa.

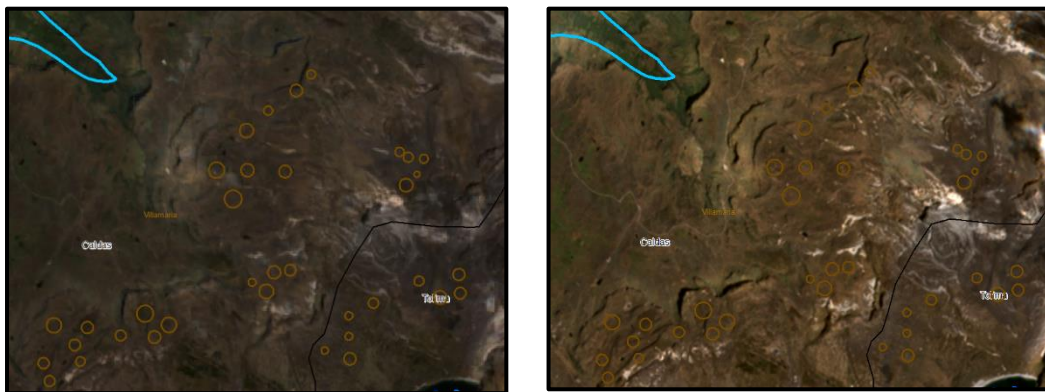


Figura 23: Semillas implementadas en la identificación de la clase “roca”. Imagen Mosaico (Izquierda) e Imagen S2A 2020 (Derecha), combinación en color real (B4+B3+B2).
Fuente: Elaboración propia a partir de productos ArcGIS.

Nieve: Dentro de los atractivos de este parque natural está la presencia de ecosistemas de alta montaña, entre los que figuran los nevados. Esta clase de cobertura se encuentra principalmente en los picos que se catalogan bajo los nombres de Nevado del Ruiz, Nevado de Santa Isabel y Nevado del Tolima (Figura 24). Para su identificación fue necesaria la combinación de bandas en falso color, permitiendo diferenciar entre las clases

de nubes y nieve, observándose de tonalidades rosadas y azules respectivamente (Figura 25). Por su parte, para la combinación en color real se ajustaron los valores máximos y mínimos de la función extensión (stretch) de las imágenes, que permite ajustar brillo, luminosidad y contraste de la imagen, permitiendo destacar elementos de las escenas.

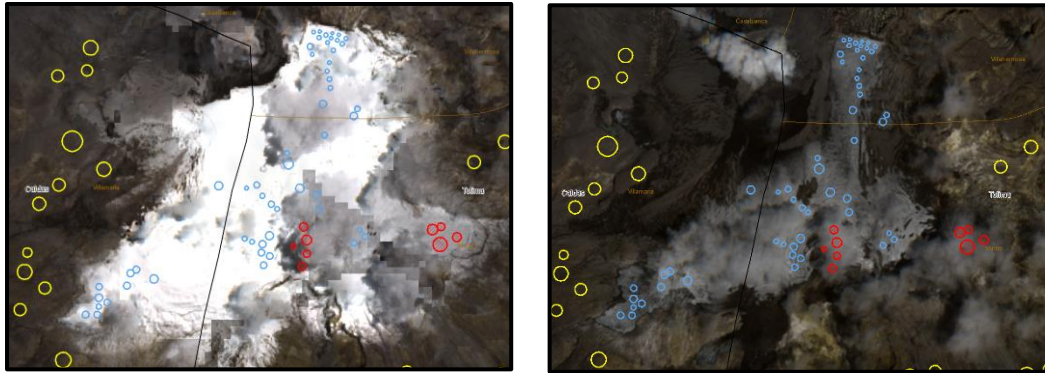


Figura 24: Semillas implementadas en la identificación de la clase “nieve”. Imagen Mosaico (Izquierda) e Imagen S2A 2020 (Derecha), combinación en color real (B4+B3+B2).
Fuente: Elaboración propia a partir de productos ArcGIS.

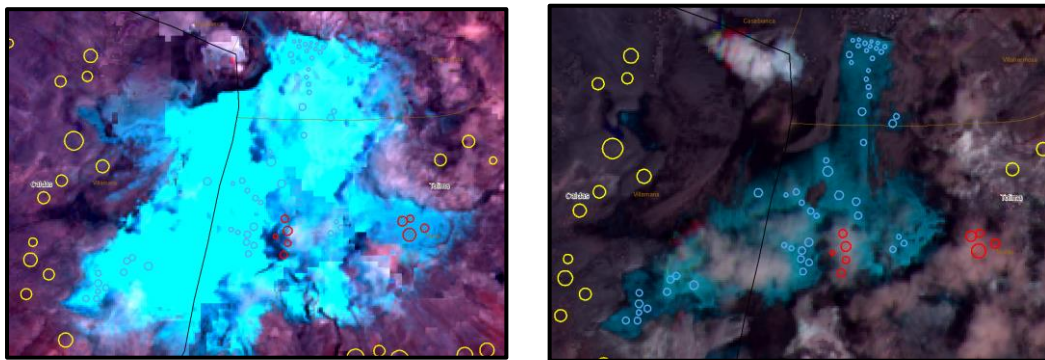


Figura 25: Semillas implementadas en la identificación de la clase “nieve”. Imagen Mosaico (Izquierda) e Imagen S2A 2020 (Derecha), combinación en falso color (B11+B8A+B4).
Fuente: Elaboración propia a partir de productos ArcGIS.

Suelo desnudo: Esta clase de cobertura se puede observar de manera frecuente en el polígono de estudio, su aparición obedece a procesos naturales o intervenciones antrópicas (Figura 26). Su identificación es relativamente sencilla en la combinación de bandas de color real, se caracteriza por sus tonalidades naranja – marrones. Sin embargo, lo anterior supone dificultades para el modelo dado que puede confundirse con la clase de tejido urbano cuyas razones fueron expuestas anteriormente.

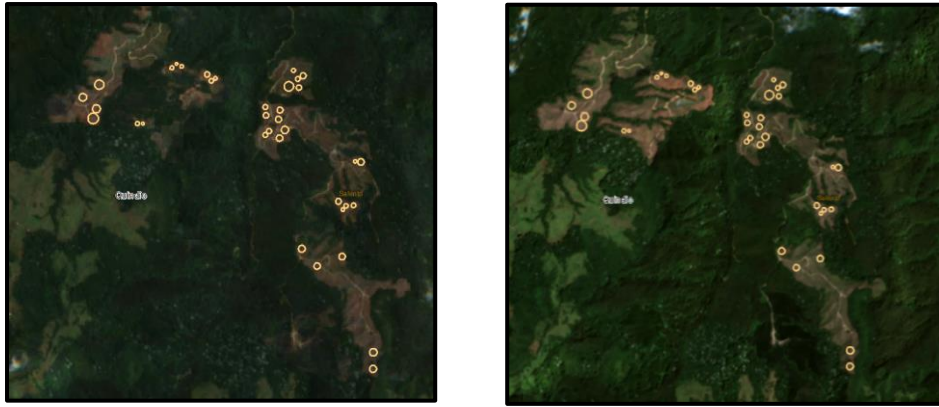


Figura 26: Semillas implementadas en la identificación de la clase “suelo desnudo”. Imagen Mosaico (Izquierda) e Imagen S2A 2020 (Derecha), combinación en color real (B4+B3+B2).
Fuente: Elaboración propia a partir de productos ArcGIS.

Nubes: La persistencia de nubosidad y condiciones meteorológicas suponen un gran reto en la zona de estudio, donde además los periodos de fumarola activa incrementan la presencia de esta clase de cobertura (Figura 27). A partir de los diferentes algoritmos implementados en la plataforma de GEE como el enmascaramiento y la composición por reducción, se pudo remover parte de la nubosidad, sin embargo, existen zonas donde estas aún permanecen por lo que fue requerido tener en cuenta esta clase en la clasificación supervisada. Para su identificación y distinción de la clase nieve en aquellas zonas de cráter, fue necesario utilizar la combinación de falso color (Color azul nieve y color rosado nubes), mientras que para la combinación en color real se ajustaron los valores máximos y mínimos de la extensión (stretch).

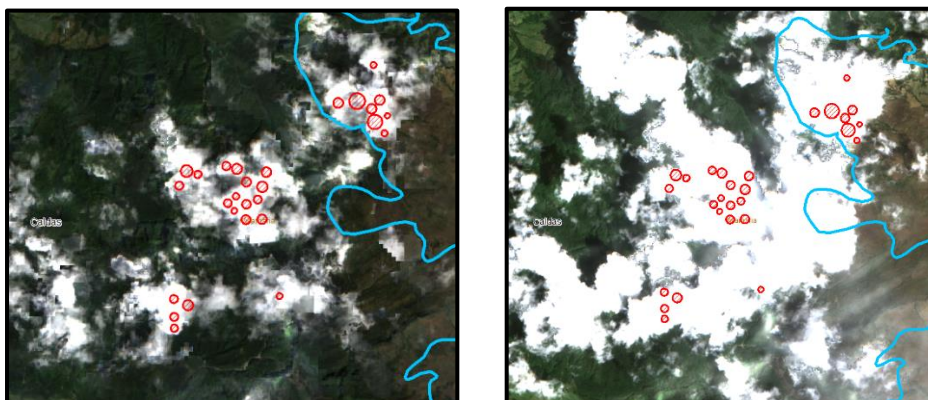


Figura 27: Semillas implementadas en la identificación de la clase “nube”. Imagen Mosaico (Izquierda) e Imagen S2A 2020 (Derecha), combinación en color real (B4+B3+B2).
Fuente: Elaboración propia a partir de productos ArcGIS

Sombras: Dados los diferentes algoritmos implementados sobre las imágenes, la clase de sombras para la imagen de mosaico no es claramente apreciable, contrario a lo observado en la imagen Sentinel del año 2020. Es por lo anterior que las semillas fueron dispuestas en las sombras asociadas al terreno (Figura 28 y figura 29), donde por el trayecto o Path Row seguido por el satélite y la posición del sol, a aproximadamente 130° de azimuth y 60° de altitud, quedan registradas en la imagen de un mismo momento y en la imagen Mosaico (Figura 30), permitiendo tener concordancia para esta clase.

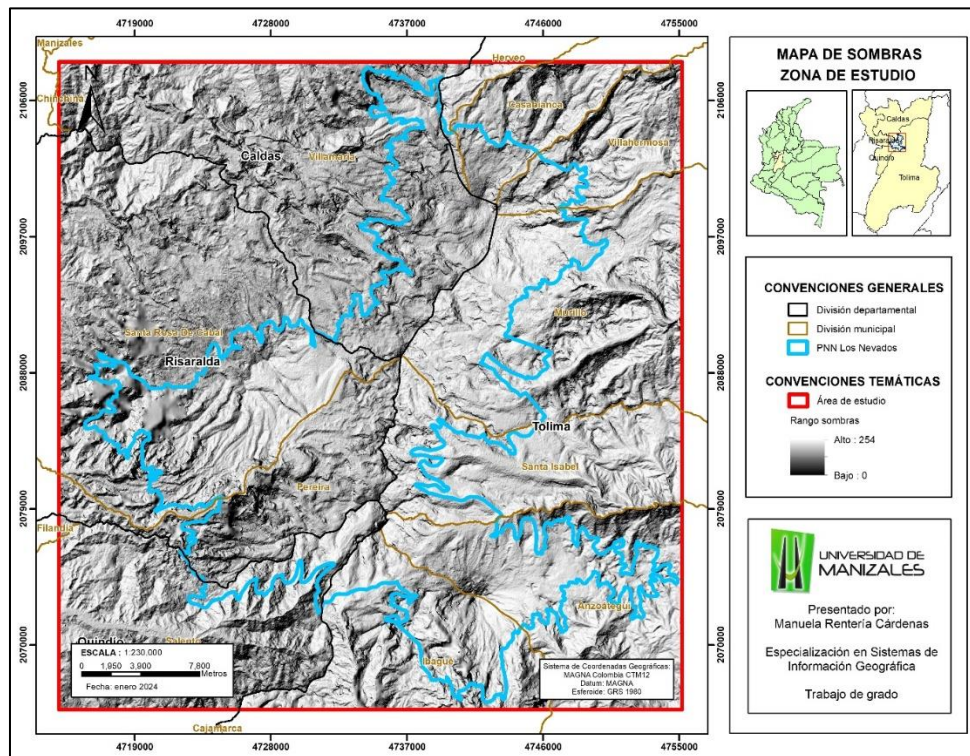


Figura 28: Mapa de sombras de la zona de estudio.
Fuente: Elaboración propia a partir de productos ArcGIS.

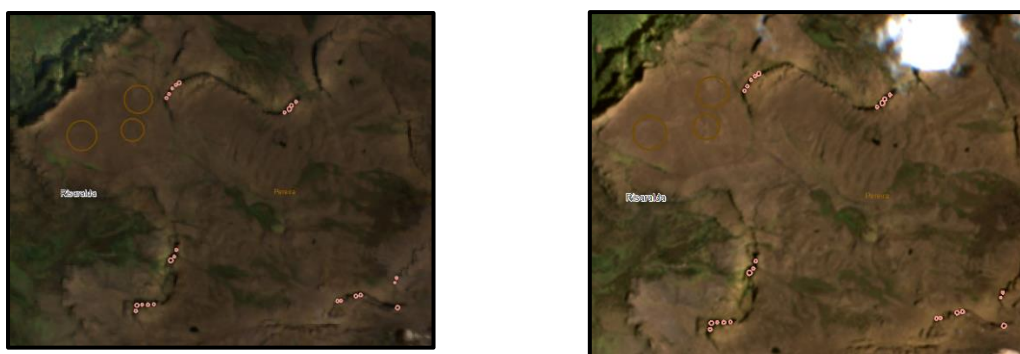


Figura 29: Semillas implementadas en la identificación de la clase "sombras". Imagen Mosaico (Izquierda) e Imagen S2A 2020 (Derecha), combinación en color real (B4+B3+B2).
Fuente: Elaboración propia a partir de productos ArcGIS.

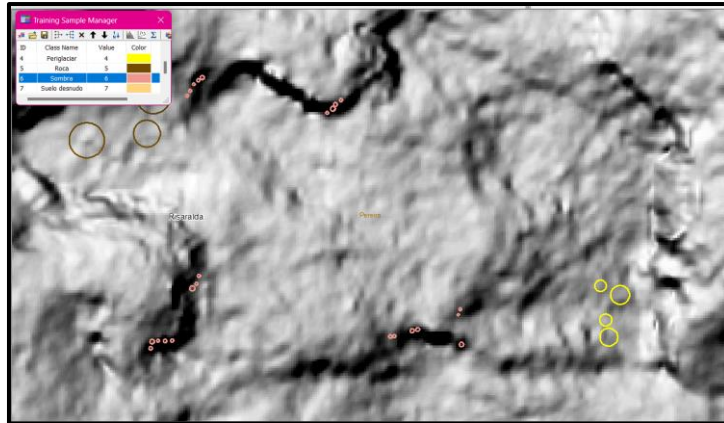


Figura 30: Relación entre las semillas de la clase sombra y el mapa de sombras de la zona.
Fuente: Elaboración propia a partir de productos ArcGIS.

Una vez consolidadas las semillas o muestras de entrenamiento, se guardaron como la firma espectral para cada RASTER, siendo estos insumos requeridos por el clasificador en el paso siguiente.

3.4 Clasificación supervisada de los RASTER

En este paso se ejecutó la clasificación de los RASTER a partir de las firmas espectrales creadas, organizando por clases y por colores adecuados los resultados obtenidos. Las semillas empleadas en cuanto a disposición espacial fueron iguales, pero con valores espectrales diferenciales según fuese el mosaico o la imagen de un mismo momento. En este punto, los clasificadores realizaron trabajo sobre imágenes de diferentes proveniencias, siendo unas obtenidas en un mismo momento y las otras obtenidas en diversos momentos y procesadas en las actividades del objetivo 2.

Objetivo 4: Comparar los resultados obtenidos y su confiabilidad a partir de ambas fuentes.

Actividades

4.1 Verificación de clasificación según diversos métodos

A partir de las firmas espectrales obtenidas para cada uno de los RASTER, se procedió a ejecutar el proceso de clasificación supervisada mediante el algoritmo Maximum

Likelihood Classification en ArcGIS, donde los datos de entrada son el RASTER que se desea clasificar y su correspondiente firma espectral. Los resultados obtenidos se observan en la figura 31 y figura 32.

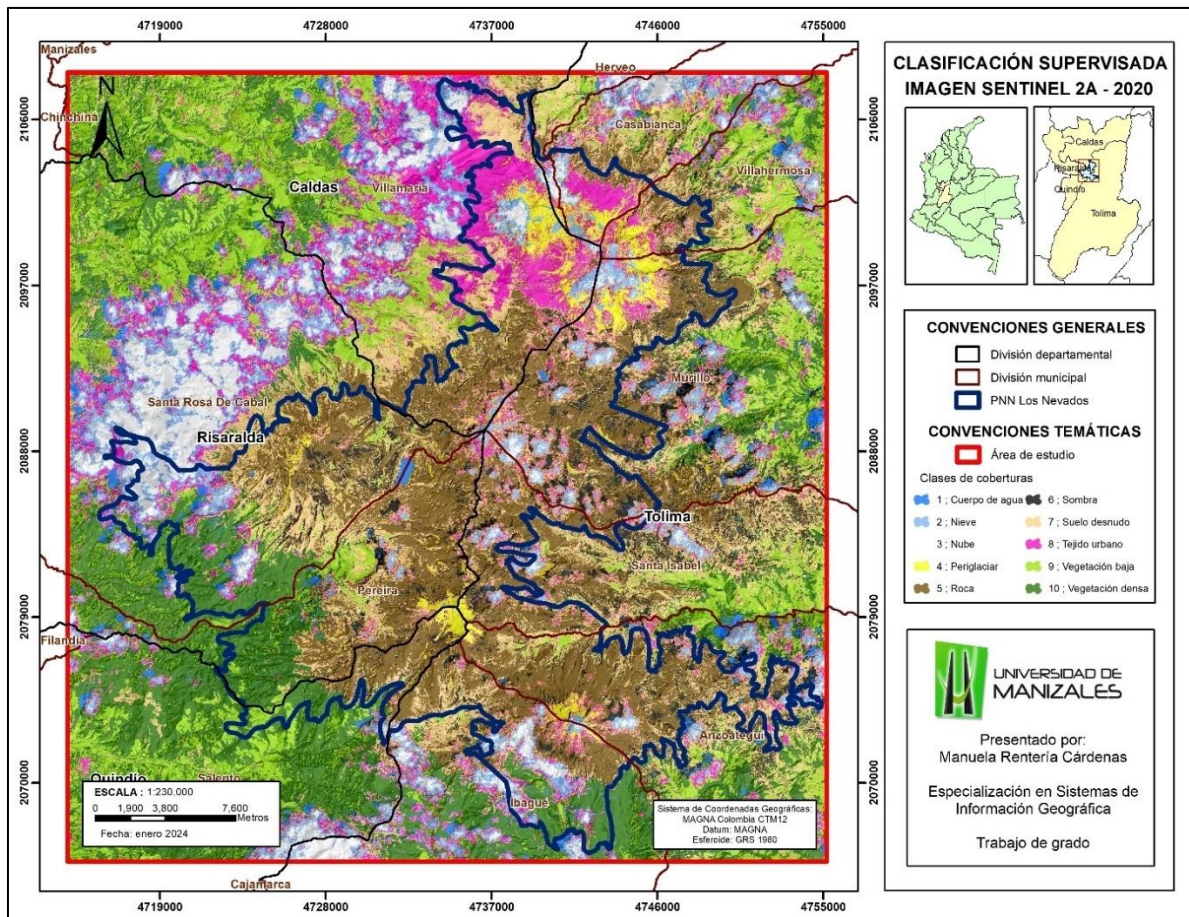


Figura 31: Clasificación supervisada obtenida de la imagen S2A del año 2020.
Fuente: Elaboración propia a partir de productos ArcGIS.

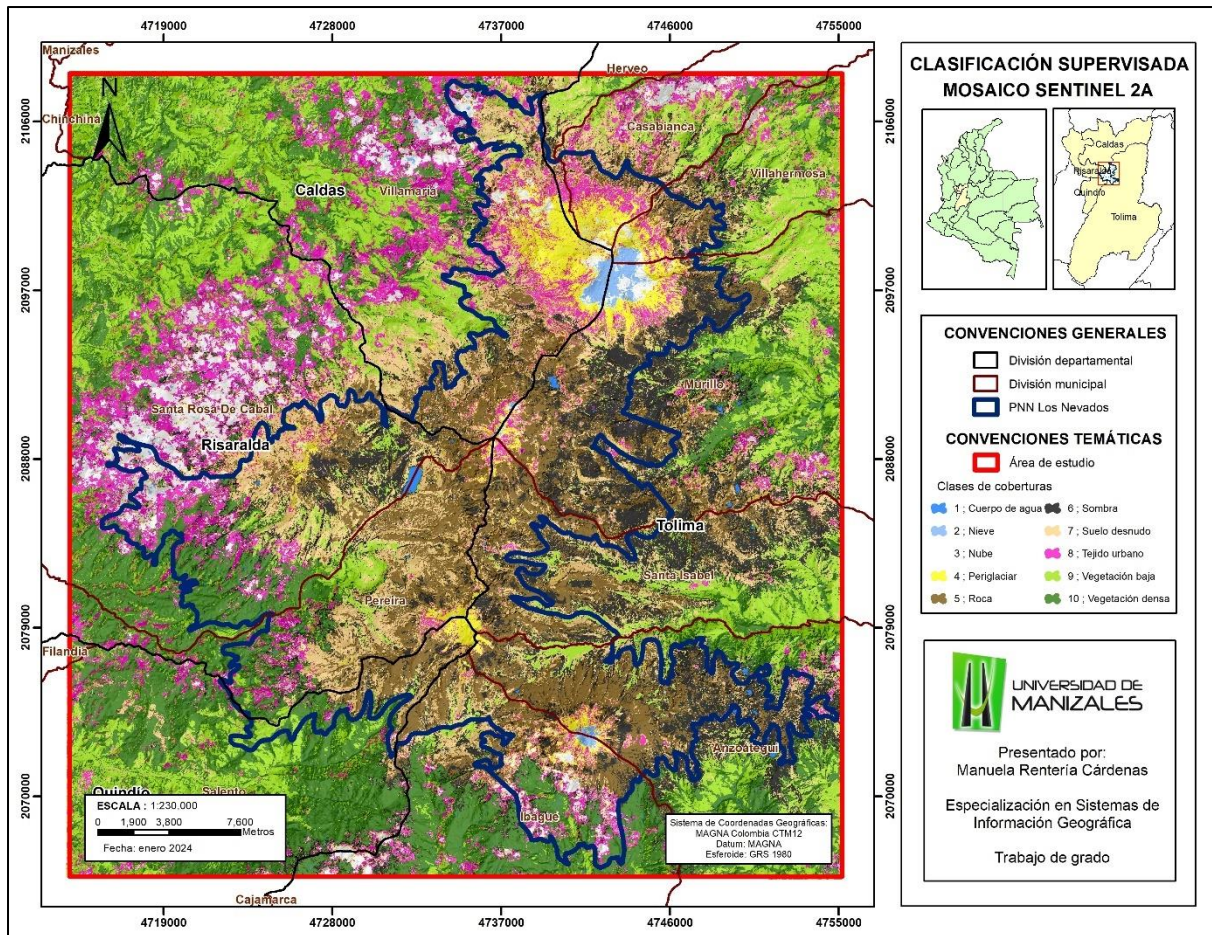


Figura 32: Clasificación supervisada obtenida mosaico S2A.
Fuente: Elaboración propia a partir de productos ArcGIS.

De las clasificaciones supervisadas anteriores se puede apreciar una distribución relativamente simétrica de las clases, exceptuando la relacionada a la nubosidad. Para dicha clase, es bastante notoria la disminución de sombras asociadas a nubes dado que la reducción por composición no toma usualmente en cuenta esos valores, que son extremos para el rango de valores considerado, y a que además se aplica un filtro de nubosidad sobre las imágenes. Ante ello, el algoritmo de clasificación para el mosaico de imágenes distribuye sombras en toda la escena, especialmente asociadas a elementos del terreno, cosa que no ocurre en la imagen de un solo momento. De manera preliminar, se puede encontrar menor uniformidad y más saltos de clase en el mosaico que en la imagen de un mismo momento, dada la forma de obtención del primero.

Se aprecia una alta uniformidad en ambos RASTER para la clase de vegetación densa baja, obteniéndose un rendimiento muy uniforme para ambos productos. Se destaca además para ambas imágenes que la clase asociada al tejido urbano presenta una dispersión errática en buena parte de la escena para ambos RASTER, especialmente en la zona de borde de la nube y en la zona de transición de la clase periglaciaria a roca. Lo anterior se asociaría a lo explicado anteriormente con relación al tipo de material que se emplea en nuestro país para la construcción de viviendas.

El anterior ejercicio se contrastó con el trabajo realizado por dos grupos de estudiantes del curso de geomática del semestre 2023-2 del Politécnico Colombiano Jaime Isaza Cadavid del programa de ingeniería civil (Figura 33), impartido por el asesor de esta investigación. Los estudiantes trabajaron sobre el mismo polígono de estudio, pero para clases diferentes, determinadas por el profesor, como trabajo final del curso. Algunos resultados obtenidos en la clasificación supervisada se presentan a continuación:

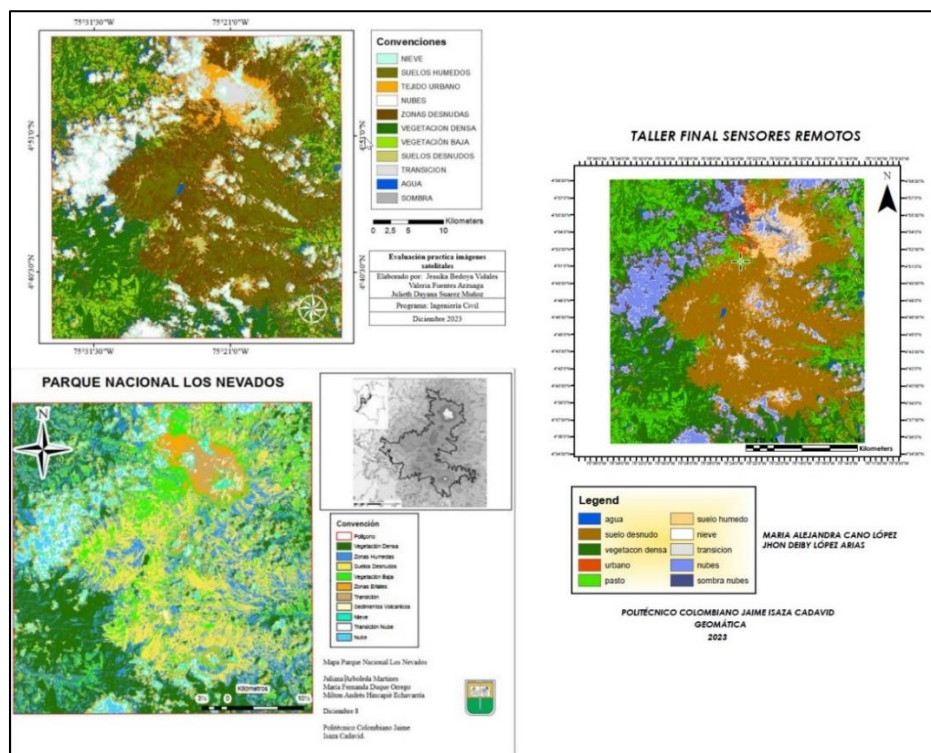


Figura 33: Ejercicio de clasificación supervisada realizado por estudiantes del Politécnico Colombiano Jaime Isaza Cadavid.

Fuente: Obtenido de ejercicio académico de estudiantes del Politécnico Jaime Isaza Cadavid.

En las clasificaciones anteriores para la misma zona y con clases diferentes, se aprecian problemas comunes manifestados también en el desarrollo de la clasificación supervisada de esta investigación y asociados con confusión del modelo para dividir clases, especialmente aquellas que tienen cierta similitud óptica, como son las sombras de las nubes y los cuerpos de agua; las zonas urbanas y los suelos desnudos; o las nubes y la nieve. Por su parte, las clases asociadas con vegetación densa y baja son mucho mejor separadas al momento de ejecutar la clasificación supervisada.

Adicionalmente se compararon las clasificaciones supervisadas obtenidas con la que entrega el proyecto Corine Land Cover (CLC) para Colombia (Figura 34), adaptado por el IDEAM con información hasta 2018, y se encontró una distribución de clases que, aunque es mucho más rica en este compendio de coberturas oficial del CLC, tiene distribuciones espaciales muy similares a las obtenidas en el ejercicio clasificatorio realizado.

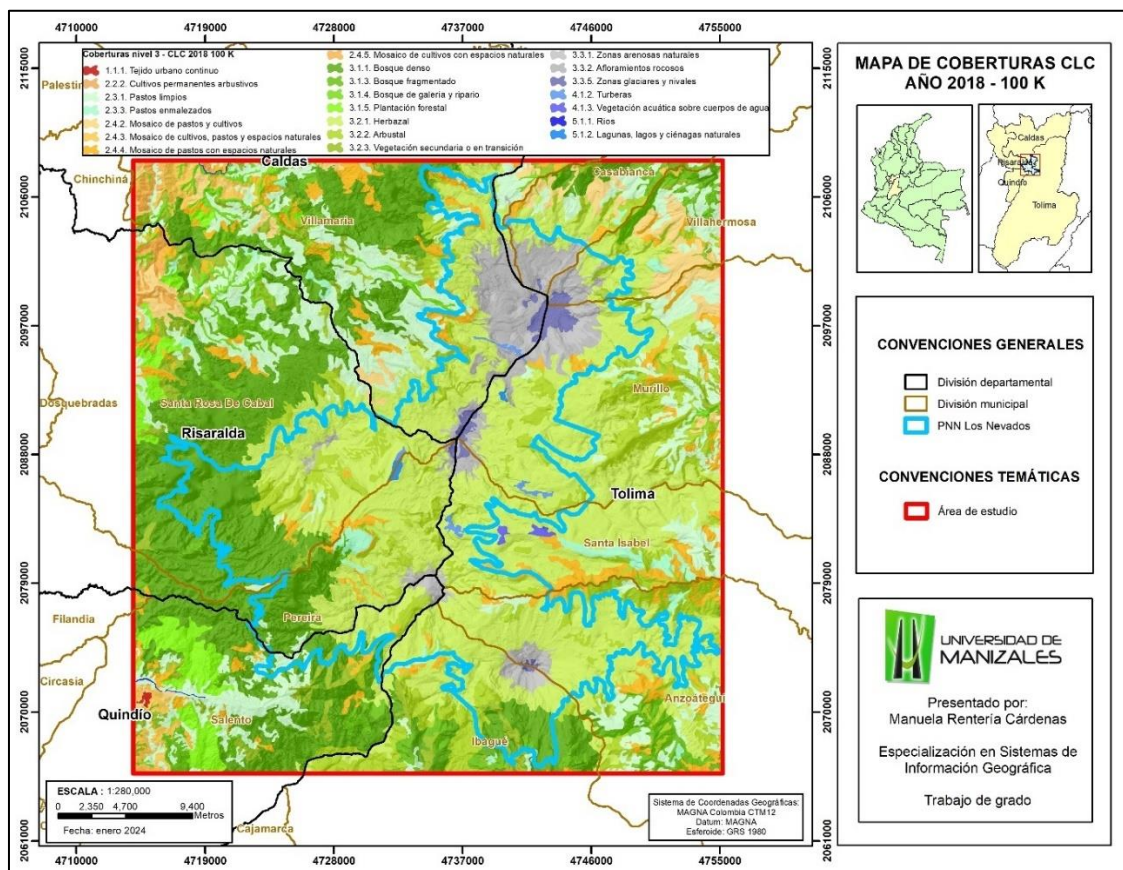


Figura 34: Coberturas de la tierra Corine Land Cover nivel 3, año 2018 100 K
Fuente: Elaboración propia a partir de productos ArcGIS.

Cabe destacar que un estudio como el CLC posee verificaciones espaciales en campo, una distribución de coberturas estandarizada, además de un desarrollo metodológico detallado a partir de recursos de misiones Landsat con resolución espacial de 30 metros. En el caso de este estudio se emplearon datos con resolución espacial con más nivel de detalle, pero con las limitaciones propias de realización remota del trabajo, que puede limitar el desempeño de este tipo de ejercicios. Llama la atención que la clase asignada por el CLC para las zonas aledañas a los picos nevados del PNN Los Nevados corresponde a herbazales, identificado como roca y suelo desnudo en el presente estudio, a partir de las ortofotografías existentes para la zona de alta montaña en cuestión, donde por limitaciones ecosistémicas no hay crecimiento de vegetación. Se aprecia además una alta coincidencia espacial en la distribución de las vegetaciones más densas y de la clase asociada con el agua del CLC con las obtenidas en este estudio. Las demás clases del CLC presentan diferentes niveles de coincidencia espacial con las clasificaciones supervisadas desarrolladas para las dos imágenes.

4.2 Comparación de rendimiento de clasificación

Con los resultados obtenidos en el punto anterior, se realizó el comparativo sobre cuál de las dos fuentes de datos (Imagen Sentinel 2A del año 2020 vs Imagen Mosaico Sentinel) tuvo un mejor rendimiento al momento de entrar en el clasificador, y si la no continuidad espacio-temporal del RASTER procesado obtuvo resultados muy diferentes al del RASTER sin procesar.

Lo anterior, se determinó mediante la construcción de una matriz de confusión, cuya función principal consiste en la identificación de errores o inconsistencias entre las características que obedecen a la realidad del terreno y los resultados obtenidos en las diferentes clasificaciones supervisadas, por lo que permite obtener un panorama o contexto sobre las clases asignadas de forma correcta o incorrecta, además de dar lugar a análisis

de naturaleza descriptiva y analítica en diferentes escalas, acorde a las necesidades de los estudios (Sánchez Muñoz, 2016).

En primera instancia, para la obtención de la matriz de confusión fue necesaria la ejecución de diferentes algoritmos de manera secuencial en el programa ArcMap del software ArcGIS, entre los que figuran la creación de un shape con geometría tipo punto, donde se dispusieron a mano las diferentes muestras aleatorias para cada una de las clases presentes acorde a lo observado o foto interpretado en cada una de las imágenes, es decir, tanto para el mosaico como para la imagen del año 2020. Estas muestras para efectos del presente trabajo se adoptaron como la información que correspondía a la realidad del terreno. Para obtener resultados significativos se escogieron 100 muestras o puntos aleatorios por cada clase que de acuerdo con su disposición se consideran como un muestreo de tipo aleatorio estratificado, obteniendo así un total de 1000 puntos aleatorios dada la existencia de 10 clases (Figura 35 y figura 36).

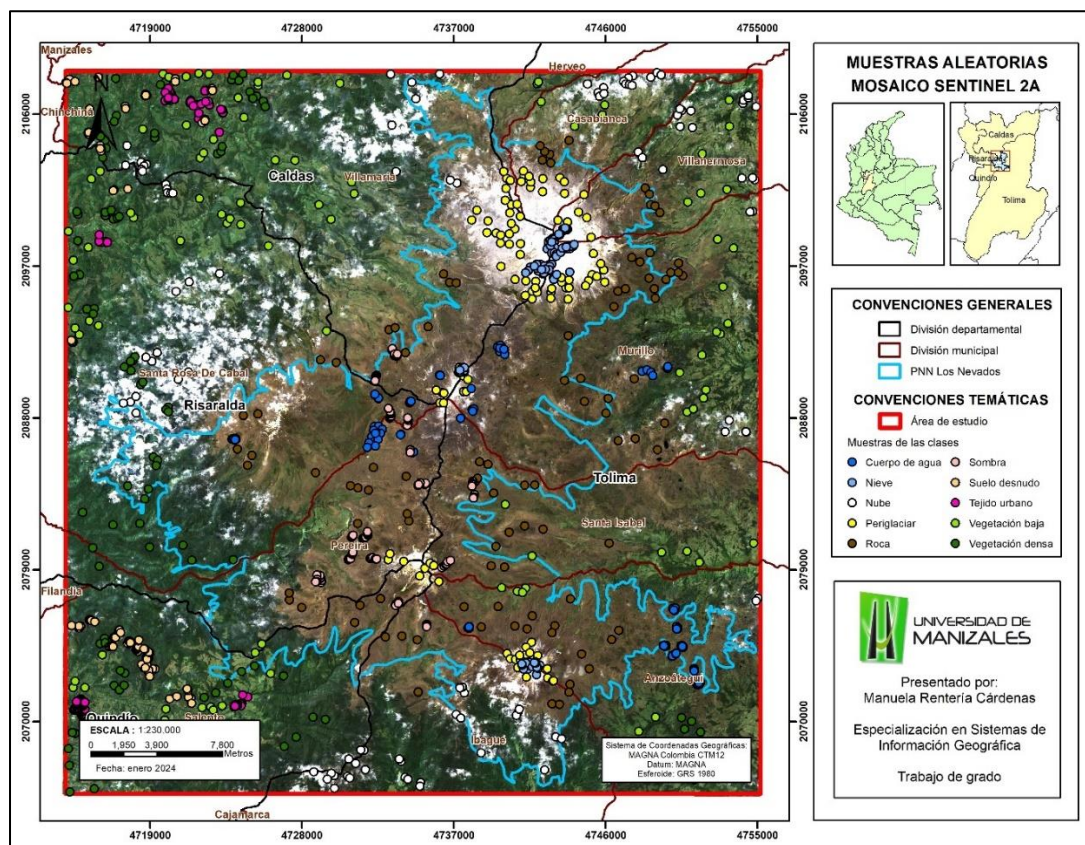


Figura 35: Muestras aleatorias dispuestas para el mosaico Sentinel 2A.
Fuente: Elaboración propia a partir de productos ArcGIS.

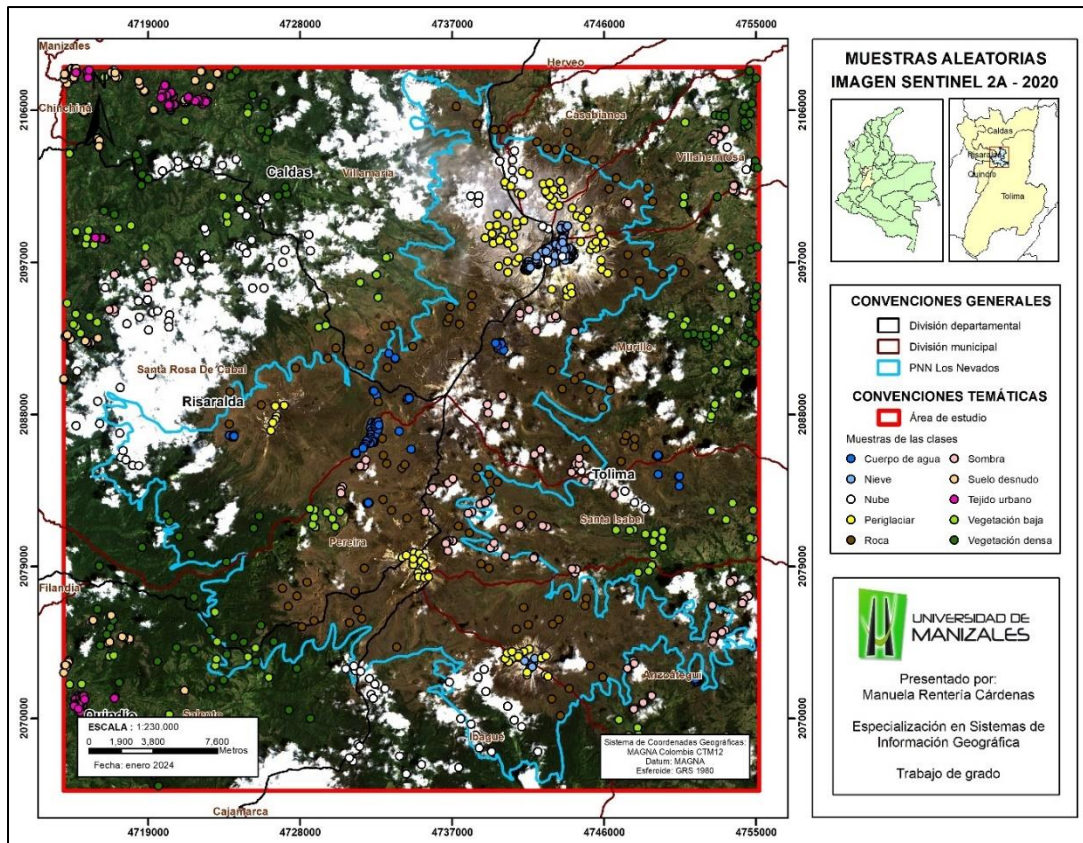


Figura 36: Muestras aleatorias dispuestas para la imagen Sentinel 2A año 2020.
Fuente: Elaboración propia a partir de productos ArcGIS.

Posterior a ello se realizó la conversión de estas muestras aleatorias de shape a RASTER, teniendo como valor de interés el ID con el cual se identificaban cada una de las clases. Posteriormente se ejecutó la herramienta “Combine” o combinar, disponible en ArcToolbox, la cual permite la combinación de varios archivos tipo RASTER con el propósito de obtener un valor de salida único para cada una de las combinaciones existentes, esto a partir de los valores de entrada correspondientes en este caso el ráster de las muestras y el ráster de clasificación supervisada. La tabla de combinaciones obtenida anteriormente se utilizó como dato de entrada en la herramienta “Pivot Table” o tabla pivote, mediante la cual se disminuye la redundancia de valores o registros, convirtiendo las relaciones de uno a muchos.

Los resultados generados con esta tabla pivote corresponden a los insumos base para la construcción de la matriz de confusión (Tabla 1 y Tabla 2), las clases consideradas

como verdaderas se incorporan en las columnas (acorde a lo interpretado en el terreno por el autor) y en las filas se disponen las clases asociadas a la clasificación supervisada realizada previamente.

Tabla 1: Matriz de confusión imagen mosaico Sentinel 2A.
Fuente: Elaboración propia a partir de productos ArcGIS.

MATRIZ DE CONFUSIÓN - IMAGEN MOSAICO S2A													
CLASES	ID	Cuerpo de agua	Nieve	Nube	Periglaciár	Roca	Sombra	Suelo desnudo	Tejido urbano	Vegetación baja	Vegetación densa	TOTAL	Error de comisión
Cuerpo de agua	1	84	0	0	0	0	3	0	0	0	1	88	4
Nieve	2	0	93	4	0	0	0	0	0	0	0	97	4
Nube	3	1	7	76	0	0	0	0	3	0	0	87	3
Periglaciár	4	0	0	6	76	0	3	6	12	0	0	103	21
Roca	5	0	0	0	0	63	19	11	0	0	0	93	30
Sombra	6	10	0	0	0	15	72	1	0	0	4	102	5
Suelo desnudo	7	0	0	0	2	21	2	75	13	0	0	113	13
Tejido urbano	8	3	0	13	22	0	0	7	69	1	3	118	4
Vegetación baja	9	0	0	1	0	1	0	0	3	99	3	107	3
Vegetación densa	10	2	0	0	0	0	1	0	0	0	89	92	0
TOTAL		100	100	100	100	100	100	100	100	100	100	1000	
Error de omisión		16	7	20	24	37	3	7	3	0	0		

Tabla 2: Matriz de confusión imagen Sentinel 2A año 2020
Fuente: Elaboración propia a partir de productos ArcGIS

MATRIZ DE CONFUSIÓN - IMAGEN S2A 2020													
CLASES	ID	Cuerpo de agua	Nieve	Nube	Periglaciár	Roca	Sombra	Suelo desnudo	Tejido urbano	Vegetación baja	Vegetación densa	TOTAL	Error de comisión
Cuerpo de agua	1	92	0	0	0	0	37	0	0	0	0	129	37
Nieve	2	0	73	28	4	0	0	0	3	0	0	108	35
Nube	3	0	27	70	0	0	0	0	0	0	0	97	0
Periglaciár	4	0	0	1	66	0	0	10	7	0	0	84	17
Roca	5	0	0	0	1	79	0	21	0	0	0	101	21
Sombra	6	3	0	0	0	3	62	0	0	0	0	68	0
Suelo desnudo	7	0	0	0	2	18	0	51	7	0	2	80	9
Tejido urbano	8	0	0	1	27	0	0	18	79	0	0	125	0
Vegetación baja	9	0	0	0	0	0	0	0	4	100	1	105	1
Vegetación densa	10	5	0	0	0	0	1	0	0	0	97	103	0
TOTAL		100	100	100	100	100	100	100	100	100	100	1000	
Error de omisión		8	27	2	30	21	1	18	4	0	0		

Los valores o números existentes en la diagonal principal se interpretan como las clasificaciones que se obtuvieron de manera correcta y los valores que se encuentran por fuera de esta, se interpretan como migraciones asociadas a errores, estos se dividen principalmente en dos categorías, errores de omisión y errores de comisión (Cerda & Villaroel, 2008).

Los errores de omisión son aquellos valores que se sitúan por debajo de la diagonal principal y corresponden a aquellos elementos que aun perteneciendo a una clase no se registran en ella, dado que están incluidos de forma errónea en otras clases. Por su parte, los errores de comisión son los valores que se disponen por encima de la diagonal principal y hacen referencia a aquellos elementos que sin pertenecer a una clase aparecen en ella (Sánchez Muñoz, 2016).

Las fórmulas utilizadas para el cálculo de ambos errores se relacionan a continuación:

$$\text{Error de omisión} = \frac{\text{Suma de valores de la columna clase por debajo de la diagonal principal}}{\text{Suma total de los valores en la columna clase}}$$

$$\text{Error de comisión} = \frac{\text{Suma de valores de la fila clase por encima de la diagonal principal}}{\text{Suma total de los valores en la fila clase}}$$

Posterior a ello se procede con el cálculo de dos índices globales (Tabla 3 y Tabla 5), el primero de ellos llamado por diferentes autores como proporción de acuerdo observado - P_o o también denominado porcentaje de acuerdo - P_a , este se define como la probabilidad de que una clase sea clasificada o no de forma correcta (Sánchez Muñoz, 2016). La fórmula implementada para su obtención se relaciona a continuación:

$$P_a = \frac{1}{N} \sum_{i=1}^M n_{i,i}$$

Representando:

M: número de clases

N: número total de muestras

n_{i,i}: número de casos en la diagonal

De igual manera se implementó el índice denominado coeficiente de acuerdo aleatorio (a posteriori) Ca_{ps} , el cual arroja el porcentaje de acuerdo esperado al azar considerando que ciertas clases están en mayor número de celdas que otras, incrementando así la posibilidad de su correcta clasificación (Sánchez Muñoz, 2016). La fórmula implementada para su obtención se relaciona a continuación:

$$Ca_{ps} = \sum_{i=1}^M P_{i+} \times P_{+i} = \frac{1}{N^2} \sum_{i=1}^M n_{i+} \times n_{+i}$$

Los dos índices descritos anteriormente, son requeridos para la obtención del denominado índice Kappa de ajuste k (Tabla 4 y Tabla 6), cuyo uso e interpretación es amplio y extendido, este representa la mejora de la clasificación en relación con la asignación aleatoria de cierto número de elementos en un determinado número de grupos, logrando establecer el porcentaje de acuerdo dada la eliminación del azar (Sánchez Muñoz, 2016). La fórmula implementada para su obtención se relaciona a continuación:

$$k = \frac{P_a - Ca_{ps}}{1 - Ca_{ps}}$$

Representando:

k: Índice Kappa de ajuste

P_a: Porcentaje de acuerdo

Ca_{ps}: Coeficiente de acuerdo aleatorio

Tabla 3: Índices calculados para la imagen mosaico Sentinel 2A.
Fuente: Elaboración propia a partir de productos ArcGIS.

IMAGEN MOSAICO S2A	ITEM	INDICE	%
Porcentaje de acuerdo	Pa	0.7960	79.60%
Coefficiente de acuerdo aleatorio	Ca _{ps}	0.0019	0.19%
Índice Kappa de ajuste	K	0.7956	79.56%

Tabla 4: Valores calculados para la imagen mosaico Sentinel 2A.
Fuente: Elaboración propia a partir de productos ArcGIS.

IMAGEN MOSAICO S2A						
Clases	Error com (%)	Error omi (%)	Porcentaje de acuerdo	Porcentaje de acuerdo (%)	Índice Kappa de ajuste	Índice Kappa de ajuste (%)
Cuerpo de agua	0.4	16				
Nieve	0.4	7				
Nube	0.3	20				
Periglacial	2.1	24				
Roca	3	37				
Sombra	0.5	3	0.7960	79.60%	0.7956	79.56%
Suelo desnudo	1.3	7				
Tejido urbano	0.4	3				
Vegetación baja	0.3	0				
Vegetación densa	0	0				

Tabla 5: Índices calculados para la imagen Sentinel 2A año 2020.
Fuente: Elaboración propia a partir de productos ArcGIS.

IMAGEN S2A 2020	ITEM	INDICE	%
Porcentaje de acuerdo	Pa	0.7690	76.90%
Coefficiente de acuerdo aleatorio	Ca _{ps}	0.0024	0.24%
Índice Kappa de ajuste	K	0.7685	76.85%

Tabla 6: Valores calculados para la imagen S2A año 2020.
Fuente: Elaboración propia a partir de productos ArcGIS.

IMAGEN S2A 2020						
Clases	Error com (%)	Error omi (%)	Porcentaje de acuerdo	Porcentaje de acuerdo (%)	Índice Kappa de ajuste	Índice Kappa de ajuste (%)
Cuerpo de agua	3.7	8				
Nieve	3.5	27	0.7690	76.90%	0.7685	76.85%
Nube	0	2				
Periglacial	1.7	30				

Roca	2.1	21
Sombra	0	1
Suelo desnudo	0.9	18
Tejido urbano	0	4
Vegetación baja	0.1	0
Vegetación densa	0	0

De acuerdo con el valor del índice Kappa de ajuste k , puede establecerse el nivel de concordancia entre la clasificación obtenida y la realidad del terreno (Tabla 7) (Cerde & Villaroel, 2008).

Tabla 7: Valoración del coeficiente kappa.
Fuente: Obtenido de Landis y Koch, 1977.

Índice Kappa	Fuerza de la concordancia
0.00	Pobre (Poor)
0.1 - 0.20	Leve (Slight)
0.21 - 0.40	Aceptable (Fair)
0.41 - 0.60	Moderada (Moderate)
0.61 - 0.80	Considerable (Substantial)
0.81 - 1.00	Casi perfecta (Almost perfect)

En este punto, con las clasificaciones obtenidas, se verificó según el método más viable la exactitud de la clasificación. Se empleó el método de las matrices de confusión a partir de puntos aleatorios distribuidos en la zona de interés. Además, se calcularon índices de exactitud de clasificación que permiten aproximarse a la determinación de la precisión de la clasificación obtenida para el RASTER reducido y el de un solo momento.

8. RESULTADOS

Una vez calculadas las matrices de confusión y el índice Kappa asociado a cada una de ellas se obtiene que, para la clasificación supervisada obtenida para la imagen satelital de tipo mosaico con bandas modificadas por reducción, el índice Kappa es de 0.7956 (79,56%) determinando así que la fuerza de concordancia entre la clasificación supervisada y las condiciones reales del terreno es considerable, dado que el valor se ubica entre el rango de 0.61 y 0.80.

En lo que respecta a la clasificación supervisada obtenida para la imagen satelital Sentinel 2A del año 2020 con bandas compuestas sin modificación, se tiene que el índice Kappa es de 0.7685 (76.85%) por lo que se establece que la fuerza de concordancia entre la clasificación supervisada y las condiciones reales del terreno es considerable, tal como ocurrió en el caso anterior.

Si bien la diferencia existente entre los valores de concordancia entre ambas clasificaciones supervisadas no es muy marcada o evidente, existen ciertas observaciones y apreciaciones que al desarrollarse dieron lugar a análisis y conclusiones más profundas sobre ambos productos, estas se relacionan a continuación:

El algoritmo de reducción permitió obtener una visibilidad adecuada en ciertas zonas donde sin su implementación, la presencia de nubes y sombras representaban una pérdida de información asociada al territorio. Esto representa una gran ventaja para una zona donde su contexto geográfico, climático, geológico y demás configuran una serie de condiciones permanentes que dificultan la obtención de información para ciertas regiones. Debe aclararse que, por el estadístico implementado en el algoritmo de reducción correspondiente a la mediana sumado a las condiciones predominantes del territorio, en la escena de interés no es posible una remoción total de la nubosidad más si una disminución de esta. Bajo la

misma línea, se menciona que la implementación de este estadístico da lugar a la aparición de nubosidad en zonas donde para la imagen del momento en específico no se registran, esto se debe a que dentro del compilado de imágenes a partir de las cuales obtendrá los píxeles para construir el ráster final, la condición mediana es de presencia de nubes.

Por otra parte, dentro de las observaciones asociadas a este producto se tiene que en algunas zonas la presencia de píxeles con saltos o con cambios abruptos alrededor puede dar lugar a imprecisiones en la clasificación o entendimiento del terreno, es allí donde cobra importancia el criterio del experto y su conocimiento espacial de la zona para brindarle adecuado manejo e interpretación a estas zonas. La pérdida de continuidad temporal mas no espacial supone una realidad condicionada por el modelo que el usuario que emplee este tipo de productos no puede perder de vista y no debe ser tomada con ligereza, pues los productos asociados a este tipo de elementos procesados se condicionarán a tal situación.

Ahora bien, en lo que concierne a las dificultades encontradas al momento de clasificar se pueden destacar las siguientes:

Si bien en la disposición de las muestras de entrenamiento para las clases “nieve” y “nube”, se utilizaron herramientas como las combinaciones de bandas, en este caso la de falso color, para su distinción, el modelo obtenido no logra diferenciar de forma adecuada entre estas dos clases, lo que se refleja en que principalmente para la clasificación de la imagen de bandas sin modificar, la clase nieve aparece en zonas que corresponden a nubes y viceversa, esto puede deberse a que la firma espectral de ambas clases así como su apariencia son muy similares y por el tamaño de pixel de la imagen no hay la precisión suficiente para su diferenciación. Esto se manifiesta en los errores de comisión que para la clase nieve es de 35, haciendo referencia a que de los 100 puntos aleatorios dispuestos para la clase nieve 35 de ellos se incluyeron incorrectamente. También se registra el caso de elementos que aun perteneciendo a la clase nieve han sido excluidos o no se encuentran

relacionados en ella, ya que están incorporados incorrectamente en otras clases, reflejado en su error de omisión con valor de 27. El primer error puede corresponder a los casos en los que las zonas de nube fueron clasificadas como nieve y el segundo error a los casos en los que las zonas de nieve fueron clasificadas como nube.

Otro caso importante es el asociado a la clase de “tejido urbano”, pese a que la disposición de semillas se procuró realizar de manera detallada, la variedad de materiales implementados para la construcción de tejados induce a la existencia de ruido ya que muchos de los colores presentes se asemejan a otras clases, tal es el caso de suelo desnudo (naranjas), roca (marrón oscuro), nubes (blancos), amarillos (periglaciario), entre otras. Esto da lugar a que ciertas zonas del terreno sean clasificadas erróneamente, lo que se refleja por ejemplo en los errores de comisión de las clases roca y periglaciario, tanto para la clasificación de la imagen con bandas modificadas por reducción (30 y 21 respectivamente) como para la imagen con bandas sin modificar (17 y 21 respectivamente) explicando así los casos en los que el tejido urbano es incluido incorrectamente en las clases de roca o periglaciario. Esto explica el por qué se tienen altos valores de omisión para las clases roca y periglaciario, traduciendo que al incluir erróneamente la clase tejido urbano se termina por excluir elementos que realmente sí pertenecerían a estas clases.

Siguiendo esta misma línea, se menciona la clase cuerpo de agua, la cual presenta algunas inconsistencias en la clasificación supervisada obtenida para la imagen de bandas combinadas sin modificar, lo anterior se atribuye a que en este ráster dada la alta presencia de nubosidad y por ende de sombras asociadas, el modelo exhibe tendencia a confundir la clase sombra con cuerpo de agua. Lo anterior se refleja en el error de comisión, siendo además el error más alto de este tipo para esta matriz, cuya interpretación corresponde a que estos 37 puntos aleatorios están incluidos incorrectamente en esta clase, siendo probablemente sombras de nube más no cuerpo de agua como lo clasifica el modelo.

Finalmente se menciona la clase sombra, cuya dificultad radica en los cambios abruptos de la aparición de esta clase entre un ráster y otro, donde la implementación del algoritmo de reducción ocasiona que la presencia de sombras en la imagen de tipo mosaico sea muy reducida por lo que se optó por determinar sombras no solo de nubes sino también las del terreno, esto pudo ocasionar que para la clasificación supervisada obtenida del ráster modificado por reducción, el modelo se viera forzado y por ende sobreestimara la aparición de esta clase dada la existencia de un número bajo de píxeles que cumplieran con la condición, clasificando entonces erróneamente muchas zonas que corresponden a roca como sombra. Lo anterior, se ve reflejado en el error de omisión de la clase roca, siendo de hecho el valor de omisión más alto para esta matriz, lo que se traduce en que muchas zonas que son roca fueron excluidas y clasificadas erróneamente en otra clase, por ejemplo, sombra.

Teniendo en cuenta las apreciaciones descritas anteriormente, se evidencia que si bien la implementación de estos algoritmos tiene sus bondades en términos de recuperación de información por la remoción parcial de nubes, los resultados asociados a la misma no están exentos de las múltiples imprecisiones o dificultades que pueden encontrarse en el ejercicio de clasificación supervisada, que están asociados no solo a la precisión del usuario (errores de comisión) sino además a la precisión del productor (errores de omisión), por lo que no perder de vista las causales de los resultados puede permitir obtener un mayor provecho y dar lugar a interpretaciones que se aproximen a la realidad.

En relación con los resultados obtenidos por el índice Kappa deben contemplarse las diferentes limitaciones que este presenta, pues si bien sus resultados brindan una buena aproximación a la concordancia existente entre las clasificaciones supervisadas y la realidad del terreno, es importante mencionarlas y no perderlas de vista. En primera instancia los valores del índice están condicionados por los datos de entrada, donde no solo corresponden a las clasificaciones sino además a las muestras aleatorias cuyo tipo de

muestreo y cantidad de muestras incidirán sobre el valor del índice Kappa, por lo cual entre más muestras se implementen y mayor aleatoriedad se garantice, el valor de este índice podrá obedecer a resultados más verídicos. Adicionalmente el valor del índice de Kappa también se ve afectado por el número de clases (Epidat 4.0, 2012) dado que entre mayor sea el número de clases, más complejo será clasificarlas correctamente y por ende el valor del índice Kappa tiende a ser más bajo. Por lo que es no solo adecuado sino además complementario, el cálculo u obtención de otros índices que refuercen y amplíen el espectro de interpretaciones obtenidas, toda vez se tenga claridad de que todas son aproximaciones para establecer la concordancia entre el modelo y las condiciones reales.

Una vez alcanzados los objetivos 1 al 4, fue posible alcanzar el objetivo general del proyecto que consiste en: Determinar el rendimiento de imágenes satelitales con bandas compuestas sin modificación versus imágenes satelitales con bandas modificadas por reducción para fines de clasificación supervisada en el Parque Nacional Natural Los Nevados y sus zonas de amortiguación, mediante un análisis comparativo que involucre la implementación de herramientas SIG y plataformas de análisis espacial.

9. CONCLUSIONES

- Para el estudio de las coberturas de la tierra la implementación del algoritmo de reducción por composición en la plataforma Google Earth Engine (GEE), representa una herramienta estratégica para la disminución de ruido en la escena y la recuperación de información en aquellas zonas donde las condiciones de nubosidad y/o fumarolas activas son persistentes. No obstante, se destaca la alta importancia de la óptica del experto para la definición de los diferentes criterios de calidad, entre los que figuran el rango de nubosidad, el estadístico a implementar entre otros, siendo estos determinantes para la obtención o no del producto deseado.
- Tanto para la imagen con bandas modificadas por reducción como para la imagen de bandas compuestas sin modificar, la concordancia entre las clasificaciones supervisadas obtenidas y las condiciones reales del terreno es considerable, siendo ligeramente superior el valor de la concordancia para la clasificación supervisada asociada a la imagen obtenida por reducción. Lo anterior, permite constatar no solo la pertinencia de implementar estos productos dada su comparabilidad y coherencia, sino además el buen posicionamiento de estos dentro del espectro de posibles alternativas para brindar solución a la problemática de nubosidad distribuida ampliamente en el territorio nacional.
- La plataforma Google Earth Engine ofrece múltiples bondades en el ámbito de análisis geoespacial, destacándose las bibliotecas de imágenes obtenidas de diferentes programas satelitales, la ejecución de procesamientos en la nube otorgando mayor rendimiento en los procesos y la comunidad científica como principales usuarios que promueven la constante carga y retroalimentación de datos de interés. Dado su alto potencial es importante que se imparta y promueva su

implementación en ejercicios académicos y profesionales, obteniendo el máximo provecho de las herramientas que están a la vanguardia.

- En la obtención de productos basados en aprendizaje automático no se puede perder de vista la interpretación del usuario, dado que si bien éstas representan herramientas innovadoras no pueden adoptarse como productos infalibles o verdades absolutas sobre la representación del terreno. Es allí donde cobra relevancia la implementación de herramientas para evaluar la eficacia y el rendimiento de los modelos de clasificación como lo es la matriz de confusión, esta permite no solo comparar los modelos sino además identificar aquellas áreas que requieren mejora, acorde a lo observado en los errores de omisión y comisión para cada clase o elemento.
- Los ejercicios de clasificación supervisada representan un modelo aproximado de las condiciones del terreno, la calidad de sus resultados no solo está condicionado por las semillas de entrenamiento, sino además por ciertas particularidades del territorio. En este caso la similitud entre las clases, la diversidad de características que abarcan las clases, entre otras. Tener claridad sobre los causales asociados a los resultados, permite obtener un panorama mucho más amplio sobre las limitaciones o no del producto obtenido.
- Dada la importancia ecosistémica, geológica y cultural del Parque Nacional Natural Los Nevados, es importante que sea objeto constante de análisis y estudios en pro del entendimiento de las dinámicas del paisaje, cambios en el ecosistema y demás factores, que generalmente se asocian con cambios en las coberturas terrestres.

10. RECOMENDACIONES

- Se sugiere implementar algoritmos de sorteo, para cerciorarse que la imagen escogida para el procesamiento de un solo momento sea la más adecuada acorde a los criterios de calidad requeridos por el usuario.
- Para futuros proyectos que requieran manejar y gestionar altos volúmenes de información, se sugiere acudir e implementar plataformas como Google Earth Engine para los análisis o procesamientos asociados, teniendo en cuenta las limitaciones existentes y las consideraciones especiales de los productos que allí se generan.
- Se recomienda ser cuidadosos con la interpretación de los mapas temáticos de coberturas obtenidos a partir de procesos de reducción, dado que no son una representación continua del terreno en términos temporales, puesto que sus píxeles pertenecen a múltiples momentos de acuerdo con el rango de tiempo definido por el autor.
- Con miras a obtener resultados cada vez más acordes con las condiciones reales del terreno, se sugiere incorporar otros métodos adicionales a la matriz de confusión, además de calcular índices que otorguen mayor peso o veracidad a los resultados.
- Se recomienda para próximos ejercicios académicos relacionados con esta temática, realizar la clasificación supervisada incorporando clases existentes en el Corine Land Cover en cualquiera de sus niveles acorde al nivel de discernimiento entre clases. Lo anterior, permitirá no solo la comparación con fuentes oficiales sino además una mayor implementación en diferentes estudios.
- Se sugiere explorar la ejecución de la clasificación supervisada con métodos diferentes al Maximum Likelihood Classification, para obtener y eventualmente comparar otras metodologías estandarizadas capaces de discretizar la realidad en

clases, quizá de forma más óptima que este método que se considera el más común de todos.

- Se recomienda la búsqueda constante de herramientas y algoritmos que permitan la representación de la realidad del terreno y permitan sortear las limitantes geográficas y climáticas que dificultan los análisis, tales como problemas de nubosidad, inexistencia de información con continuidad temporal, ruidos por asuntos relacionados a la captura de información, entre otros. Este tipo de productos fortalecerá la capacidad creciente de monitorear de forma remota la superficie terrestre.
- Se recomienda ahondar en el análisis y validación de los productos obtenidos a partir de información que provenga de algoritmos de reducción por composición, dado que pueden ser tomados como soluciones altamente recomendables y con creciente auge de uso, dadas sus ventajas evidentes, pero con limitaciones no siempre tenidas en cuenta que pueden conllevar a subestimar los errores asociados a la representación de la realidad propios de este tipo de elementos.

11. REFERENCIAS

- Agencia Espacial Europea (ESA). (2024, 02 21). *The European Space Agency*. From https://www.esa.int/Space_in_Member_States/Spain/EI_programa_Copernico
- Arango Gutiérrez et al. (2005). CLASIFICACIÓN NO SUPERVISADA DE COBERTURAS VEGETALES SOBRE IMÁGENES DIGITALES DE SENSORES REMOTOS: "LANDSAT - ETM+". *Revista Facultad Nacional de Agronomía - Medellín*, 58(1), 2611-2634.
- Arenas Herrera et al. (2019). *ResearchGate*. doi:10.13140/RG.2.2.35252.37766
- Avogadro, E. G., & Padró Garcia, J. C. (2019). Diferenciación de plantaciones forestales en EntreRíos (Argentina): Comparación de métodos de clasificación aplicados a imágenes sentinel-2 y landsat-8",. *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica*, 117-139.
doi:<http://dx.doi.org/10.21138/GF.652>
- Cerda, J., & Villaroel, L. (2008). Evaluación de la concordancia inter-observador en investigación pediátrica: Coeficiente de Kappa. *Revista Chilena de Pediatría*, 79(1), 54-58. doi:<http://dx.doi.org/10.4067/S0370-41062008000100008>
- Chuvieco, E. (2008). *Teledetección ambiental* (Vol. 3ra edición). (G. P. (GBS), Ed.) Editorial Ariel.
- Corso, C. L. (2009). *Aplicación de algoritmos de clasificación supervisada usando Weka*. Córdoba: Universidad Tecnológica Nacional, Facultad Regional Córdoba.
- DNP, M. A. (2022). *Recuperación de Cobertura Vegetal en Áreas Disturbadas*. Bogotá: Departamento Nacional de Planeación.
- Epidat 4.0. (2012). www.sergas.es. Retrieved Enero, 2024 from Sergas: Servicio Gallego de Salud: <https://www.sergas.es/Saude-publica/Epidat-4-1-6--Concordancia-e-consistencia>
- Eraso Terán et al. (2022). Google Earth Engine aplicado al monitoreo de superficies quemadas y cambios de coberturas y usos del suelo. *GeoFocus, Revista*

Internacional de Ciencia y Tecnología de la Información Geográfica, 29, 89–113.

doi:<https://dx.doi.org/10.21138/GF.764>

ESRI. (n.d.). *Imágenes y teledetección / Fundamentos de imágenes y rasters*. Retrieved 2023 from <https://pro.arcgis.com/es/pro-app/latest/help/data/imagery/what-s-in-a-pixel.htm>

Gorelick et al. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202, 18 - 27.

doi:10.1016/j.rse.2017.06.031

Hay Chung et al. (2021). Improved machine-learning mapping of local climate zones in metropolitan areas using composite Earth observation data in Google Earth Engine.

Building and Environment, 199. doi:10.1016/j.buildenv.2021.107879

IDEAM. (2019). *Glosario Meteorológico*. Bogotá: Subdirección de Meteorología.

Lindsay et al. (2022). Multi-Temporal Satellite Image Composites in Google Earth Engine for Improved Landslide Visibility: A Case Study of a Glacial Landscape. *Remote Sensing*, 14(10). doi:10.3390/rs14102301

doi:10.3390/rs14102301

LUO et al. (2021). Monthly composites from Sentinel-1 and Sentinel-2 images for regional major crop mapping with Google Earth Engine. *Journal of Integrative Agriculture*,

20(7), 1944 - 1957. doi:10.1016/S2095-3119(20)63329-9

Luo et al. (2021). Using time series sentinel-1 images for object-oriented crop classification in google earth engine. *Remote Sensing*, 13(4), 1-19. doi:10.3390/rs13040561

Ministerio de Ambiente y Desarrollo Sostenible. (1974). *Portal web Ministerio de Ambiente y Desarrollo Sostenible*. From [https://www.minambiente.gov.co/wp-](https://www.minambiente.gov.co/wp-content/uploads/2021/10/Decreto-2811-de-1974.pdf)

[content/uploads/2021/10/Decreto-2811-de-1974.pdf](https://www.minambiente.gov.co/wp-content/uploads/2021/10/Decreto-2811-de-1974.pdf)

Ministerio del Medio Ambiente. (2010). *Leyenda nacional de coberturas de la tierra*.

Metodología CORINE Land Cover Adaptada para Colombia Escala 1:100.000. From

http://siatac.co/c/document_library/get_file?uuid=a64629ad-2dbe-4e1e-a561-

[fc16b8037522&groupId=762](http://siatac.co/c/document_library/get_file?uuid=a64629ad-2dbe-4e1e-a561-fc16b8037522&groupId=762)

- Olmos S et al. (2023). Obtención de índices multitemporales tierra-agua en cuencas del Oriente Antioqueño Colombiano; empleando Earth Engine e imágenes Landsat y Sentinel con enfoque a la planificación del recurso hídrico y del territorio. In P. Cristiano, J. G. Dalfré Filho, A. *ANALES del XXX Congreso Latinoamericano de Hidráulica*, 348-359.
- Parques Nacionales Naturales de Colombia. (1974). *Parques Nacionales Naturales de Colombia / Políticas y Marco Normativo*. From <https://www.parquesnacionales.gov.co/portal/es/sistema-nacional-deareas-protegidas-sinap/politicas-y-marco-normativo/>
- Parques Nacionales Naturales de Colombia. (2017). *Parques Nacionales Naturales de Colombia / Resoluciones*. From <https://www.parquesnacionales.gov.co/portal/es/normatividad/marco-normativo-del-sistema-de-parques-nacionales-naturales/resoluciones-3/>
- Paz Pellat et al. (2015). ALCANCES Y LIMITACIONES DE LOS ÍNDICES ESPECTRALES DE LA VEGETACIÓN: ANÁLISIS DE ÍNDICES DE BANDA ANCHA. *Terra Latinoamericana*, 33(1), 27-49.
- Peña Araya, M. A. (2007). Correcciones de una imagen satelital ASTER para estimar parámetros. *Bosque*, 28(2), 162-172.
- Perilla, G. A., & Mas, J.-F. (2020). Google Earth Engine (GEE): una poderosa herramienta que vincula el potencial de los datos masivos y la eficacia del procesamiento en la nube. *Investigaciones Geográficas*, 101, 0–2. doi:<https://doi.org/10.14350/rig.59929>
- Sánchez Muñoz, J. M. (2016). Análisis de Calidad Cartográfica mediante el estudio de la Matriz de Confusión. *Pensamiento Matemático*, 6(2), 9-18.
- Shelestov et al. (2017). Exploring Google earth engine platform for big data processing: Classification of multi-temporal satellite imagery for crop mapping. *Frontiers in Earth Science*, 5, 1-10. doi:10.3389/feart.2017.00017
- Teluguntla et al. (2018). A 30-m landsat-derived cropland extent product of Australia and China using random forest machine learning algorithm on Google Earth Engine cloud

computing platform. *ISPRS Journal of Photogrammetry and Remote Sensing*, 144, 325 - 340. doi:10.1016/j.isprsjprs.2018.07.017

Tsai et al. (2018). Mapping vegetation and land use types in Fanjingshan National Nature Reserve using google earth engine. *Remote Sensing*, 10(6), 927.
doi:10.3390/rs10060927

Vélez-Castaño et al. (2020). Erosion and progradation in the Atrato River delta: A spatiotemporal analysis with Google Earth Engine. *Revista Facultad de Ingeniería Universidad de Antioquia*, 99, 83–98. doi:10.17533/udea.redin.20200688

Veneros et al. (2020). Aplicación de sensores remotos para el análisis de cobertura vegetal y cuerpos de agua. *Idesia [online]*, 38(4), 99 - 107.
doi:http://dx.doi.org/10.4067/S0718-34292020000400099.

Zhang et al. (2021). Assessment of annual composite images obtained by google earth engine for urban areas mapping using random forest. *Remote Sensing*, 13(4), 1 - 21.
doi:10.3390/rs13040748

Zhang, D.-D., & Zhang, L. (2020). Land cover change in the central region of the lower yangtze river based on landsat imagery and the google earth engine: A case study in Nanjing, China. *Sensors (Switzerland)*, 20(7), 2091. doi:10.3390/s20072091