

ALGORITMO DE INTELIGENCIA ARTIFICIAL PARA LA DETECCIÓN DE CULTIVOS DE CACAO (*TEOBROMA CACAO* L.), BANANO (*MUSA PARADISIACA* L.) Y PALMA AFRICANA (*ELAEIS GUINEENSIS* J.).



AUTORES:

Oliver Michael Zambrano Garcia

Lidia Vlassova

**Algoritmo de inteligencia artificial
para la detección de cultivos de
cacao (*Teobroma cacao* L.),
banano (*Musa paradisiaca* L.) y
palma africana (*Elaeis guineensis*
J.).**

Autor/es:

Zambrano Garcia, Oliver Michael

Vlassova, Lidia

© Publicaciones Editorial Grupo AEA Santo Domingo – Ecuador

Publicado en: <https://www.editorialgrupo-aea.com/>

Contacto: +593 983652447; +593 985244607 Email: info@editorialgrupo-aea.com

Título del libro:

Algoritmo de inteligencia artificial para la detección de cultivos de cacao (*Teobroma cacao* L.), banano (*Musa paradisiaca* L.) y palma africana (*Elaeis guineensis* J.).

© Zambrano Garcia Oliver Michael, Vlassova Lidia.

© Diciembre, 2023

Libro Digital, Primera Edición, 2023

Editado, Diseñado, Diagramado y Publicado por Comité Editorial del Grupo AEA, Santo Domingo de los Tsáchilas, Ecuador, 2023

ISBN: 978-9942-651-17-4



<https://doi.org/10.55813/egaea.l.2022.61>

Como citar: Zambrano Garcia, O. M., Vlassova, L. (2023). Algoritmo de inteligencia artificial para la detección de cultivos de cacao (*Teobroma cacao* L.), banano (*Musa paradisiaca* L.) y palma africana (*Elaeis guineensis* J.). Editorial Grupo AEA. Ecuador. <https://doi.org/10.55813/egaea.l.2022.61>

Palabras Clave: Maximum Likelihood, Random Forest, Cultivos, NDVI, RESI.

Cada uno de los textos de Editorial Grupo AEA han sido sometido a un proceso de evaluación por pares doble ciego externos (double-blindpaperreview) con base en la normativa del editorial.

Revisores:



Ing. Rojas Felipe Edwin, Ph.D.

Universidad Nacional
Huancavelica – Perú

de



Ing. Jorge Washington Rodríguez
Deza, Ph.D.

Universidad Nacional
Huancavelica – Perú

de



Los libros publicados por “**Editorial Grupo AEA**” cuentan con varias indexaciones y repositorios internacionales lo que respalda la calidad de las obras. Lo puede revisar en los siguientes apartados:



Editorial Grupo AEA



Aviso Legal:

La información presentada, así como el contenido, fotografías, gráficos, cuadros, tablas y referencias de este manuscrito es de exclusiva responsabilidad del/los autor/es y no necesariamente reflejan el pensamiento de la Editorial Grupo AEA.

Derechos de autor ©

Este documento se publica bajo los términos y condiciones de la licencia Creative Commons Reconocimiento-NoComercial-CompartirIgual 4.0 Internacional (CC BY-NC-SA 4.0).



El “copyright” y todos los derechos de propiedad intelectual y/o industrial sobre el contenido de esta edición son propiedad de la Editorial Grupo AEA y sus Autores. Se prohíbe rigurosamente, bajo las sanciones en las leyes, la producción o almacenamiento total y/o parcial de esta obra, ni su tratamiento informático de la presente publicación, incluyendo el diseño de la portada, así como la transmisión de la misma de ninguna forma o por cualquier medio, tanto si es electrónico, como químico, mecánico, óptico, de grabación o bien de fotocopia, sin la autorización de los titulares del copyright, salvo cuando se realice confines académicos o científicos y estrictamente no comerciales y gratuitos, debiendo citar en todo caso a la editorial. Las opiniones expresadas en los capítulos son responsabilidad de los autores.

RESEÑA DE AUTORES

**Zambrano Garcia, Oliver Michael**

Universidad Técnica Estatal de Quevedo

oliver.zambrano2017@uteq.edu.ec

El Ingeniero ambiental Oliver Zambrano se graduó de la Universidad Técnica Estatal de Quevedo en julio de 2023. La presente publicación se basa en los resultados del proyecto de investigación “Aplicación de algoritmo de inteligencia artificial para la detección de cultivos de cacao (*Theobroma cacao* L.), banano (*Musa paradisiaca* L.) y palma africana (*Elaeis guineensis* J.)” donde activamente colaboró. Dentro del proyecto se llevó a cabo una comparación para evaluar la eficacia de dos métodos en la identificación de cultivos de banano, cacao, palma africana y otras coberturas: Máxima Verosimilitud (ML) y Bosques Aleatorios (RF).

**Vlassova, Lidia**

Universidad Técnica Estatal de Quevedo

lvlassova@uteq.edu.ec<https://orcid.org/0000-0001-5025-5691>

Geógrafa Lidia Vlassova es una docente e investigadora de la Universidad Técnica de Quevedo (Ecuador) con amplia experiencia en aplicación de métodos de teledetección, herramientas de Sistemas de Información Geográfica y algoritmos de Inteligencia Artificial en ordenamiento territorial y gestión ambiental. Los resultados de sus investigaciones fueron presentados en los eventos científicos y publicados en las revistas de renombre internacional.

Índice

Reseña de Autores	VII
Índice	VIII
Índice de Tablas.....	XI
Índice de Figuras	XII
Índice de Anexos	XIII
Índice de Ecuaciones	XIV
Introducción	XV
Capítulo I: Contextualización y fundamentación teórica de la investigación	19
1.1. Problema de investigación	21
1.1.1. Planteamiento del problema	21
1.1.2. Diagnóstico.....	22
1.1.3. Pronóstico	23
1.1.4. Formulación del problema	23
1.1.5. Sistematización del problema	23
1.2. Objetivos.....	24
1.2.1. Objetivo general	24
1.2.2. Objetivos específicos.....	24
1.3. Justificación	24
1.4. Fundamentación teórica de la investigación.....	26
1.4.1. Marco conceptual	26
1.4.1.1. Producción agrícola sostenible.....	26
1.4.1.2. Ordenamiento territorial.....	26
1.4.1.3. Teledetección.....	26
1.4.1.4. Imágenes multiespectrales.....	27
1.4.1.5. Firmas espectrales	28
1.4.1.6. Sentinel-2.....	28

1.4.1.7. Bandas espectrales.....	28
1.4.1.7.1. Bandas del Sentinel-2.....	29
1.4.1.8. Clasificación de imágenes satelitales	29
1.4.1.8.1. Clasificación supervisada.....	30
1.4.1.8.2. Clasificación no supervisada.....	30
1.4.1.8.3. Método Máxima Probabilidad.....	30
1.4.1.8.4. Método Bosques Aleatorios	30
1.4.1.9. Índices espectrales	31
1.4.1.9.1. Índice NDVI.....	31
1.4.1.9.2. Índice RESI.....	31
1.5. Marco referencial	31
1.5.1. Estudio referencial 1	32
1.5.2. Estudio referencial 2.....	33
1.5.3. Estudio referencial 3.....	33
1.5.4. Estudio referencial 4	34
1.5.5. Estudio referencial 5	35
Capítulo II: Metodología de la investigación	37
2.1. Localización	39
2.2. Tipo de investigación	40
2.3. Métodos de investigación.....	40
2.4. Fuentes de recopilación de información.....	41
2.5. Diseño de la investigación	41
2.5.1. OE 1: Ejecutar el preprocesamiento de imágenes multiespectrales del satélite Sentinel-2 con las funciones del software especializado (ENVI y ArcGIS) para la posterior clasificación	42
2.5.1.1. Búsqueda de imágenes.....	42
2.5.1.2. Comprobación de imágenes.....	44

2.5.1.3. Corrección atmosférica de las imágenes satelitales Sentinel-2	45
2.5.1.4. Unión de las bandas espectrales	46
2.5.1.5. Extracción y recorte de polígono real para imagen.....	47
2.5.2. OE 2: Realizar la clasificación supervisada de las imágenes de Sentinel-2 con los métodos de Máxima Verosimilitud y Bosques Aleatorios e índices espectrales para detectar los cultivos principales en la zona de estudio	48
2.5.2.1. Cálculos de índices espectrales	48
2.5.2.2. Cálculo de índices espectrales en el software	49
2.5.2.3. Creación de áreas de entrenamiento	49
2.5.2.4. Regiones de interés (ROIs).....	50
2.5.2.5. Aplicación del clasificador Máxima Probabilidad (ML)	51
2.5.2.6. Aplicación del clasificador Bosques Aleatorios (RF).....	51
2.5.3. OE 3: Llevar a cabo el análisis estadístico de las clasificaciones para evaluar la precisión de los dos métodos en detección de los cultivos de cacao, banano y palma africana	52
2.5.3.1. Matriz de confusión	52
2.5.3.2. Coeficiente Kappa	53
2.5.3.3. Análisis estadístico.....	53
2.6. Recursos y materiales	54
2.6.1. Programas informáticos	54
2.7. Sitios web	55
Capítulo III: Resultados, discusión, conclusiones y recomendaciones.	57
3.1. Preprocesamiento de imágenes satelitales	59
3.1.1. Selección de imagen satélite Sentinel-2	59
3.1.2. Unión de bandas espectrales y recorte de imagen para área de estudio	62
3.2. Clasificación supervisada de las imágenes de Sentinel-2	64

3.2.1. Cálculo de índices espectrales NDVI y RESI	64
3.2.2. Cálculo de Stack de índices.....	65
3.2.3. Clasificación supervisada	67
3.2.3.1. Creación de áreas de entrenamiento	67
3.2.3.2. Creación de ROIs.....	69
3.2.3.3. Clasificación con Máxima Verosimilitud.....	70
3.2.3.4. Clasificación con Bosques Aleatorios	71
3.3. Análisis estadístico de los resultados.....	73
3.3.1. Validación de resultados.....	73
3.3.1.1. Áreas de entrenamiento al azar	74
3.3.1.2. Matriz de confusión	75
3.3.1.3. Coeficiente Kappa	77
3.3.2. Análisis estadístico	77
3.4. Discusión	79
3.5. Conclusiones	82
3.6. Recomendaciones	84
Referencias Bibliográficas.....	87
Anexos.....	97

Índice de Tablas

Tabla 1 <i>Bandas de Sentinel-2</i>	29
Tabla 2 Ecuaciones de Índices Espectrales.....	48
Tabla 3 <i>Formato para registro de coordenadas para definir áreas de entrenamiento</i>	50
Tabla 4 <i>Modelo de matriz de contingencia con aditamento de resultados del coeficiente Kappa</i>	52
Tabla 5 <i>Nivel de valorización del índice de Kappa</i>	53
Tabla 6 <i>Características de imagen 1 de Sentinel-2</i>	59
Tabla 7 <i>Características de imagen 2 de Sentinel-2</i>	61

Tabla 8 Línea de código aplicado para índices espectrales	64
Tabla 9 Áreas de entrenamiento.....	67
Tabla 10 ROIs de Stack de índices.....	69
Tabla 11 Extensión de coberturas clasificadas con método Máxima Verosimilitud (ML)	70
Tabla 12 Extensión de coberturas clasificadas con método Bosques Aleatorios (RF)	72
Tabla 13 Áreas de entrenamiento al azar	74
Tabla 14 Fuerza de Concordancia según el coeficiente de Kappa para las clasificaciones de Máxima Verosimilitud (ML) y Bosques Aleatorios (RF)	77
Tabla 15 Prueba estadística de normalidad para los métodos de clasificación Máxima Verosimilitud (ML) y Bosques Aleatorios (RF).....	77
Tabla 16 Prueba estadística de la Suma de Rangos de Wilcoxon para las comparaciones entre métodos clasificadores Máxima Verosimilitud (ML) y Bosques Aleatorios (RF)	78

Índice de Figuras

Figura 1 Opacidad atmosférica a la radiación electromagnética.....	28
Figura 2 <i>Mapa de localización del área de estudio: Cantones Buena Fe, Valencia y Empalme</i>	39
Figura 3 Flujograma del proceso metódico de la investigación.....	42
Figura 4 Servicio web Geológico "USGS Earth Explorer"	43
Figura 5 Comprobación de nubosidad en el área de estudio	44
Figura 6 Corrección atmosférica de bandas espectrales Sentinel-2 con el Semi Automatic Classification Plugin del año 2021	45
Figura 7 Configuración de combinación de bandas espectrales según polígono del área de estudio.....	46
Figura 8 Definición de polígono recortado del área de estudio en función de las coberturas reales	47
Figura 9 Ejemplo del cálculo para índice espectral en software ENVI - Herramienta Band Math	49
Figura 10 Herramienta algorítmica de clasificación "Dzetsaka" para Random Forest.....	51

Figura 11 Imagen 1: Combinación RGB de Sentinel-2 (15/08/2019) con área de estudio sobrepuesta.....	60
Figura 12 Imagen 2: Combinación RGB de Sentinel-2 (20/08/2019) con área de estudio sobrepuesta.....	62
Figura 13 Recorte de área de estudio para imagen preprocesada de Sentinel-2	63
Figura 14 Representación de Índices Espectrales: NDVI (A) & RESI (B)	65
Figura 15 Stack de Índices NDVI + RESI calculados.....	66
Figura 16 Clasificación supervisada con método Máxima Verosimilitud	71
Figura 17 Clasificación supervisada con método Bosques Aleatorios	73
Figura 18 Matriz de confusión de las clasificaciones de imágenes: A(ML) & B(RF)	75

Índice de Anexos

Anexo 1 Metadatos completos de la imagen 1 de Sentinel-2	99
Anexo 2 Metadatos completos de la imagen 2 de Sentinel-2	101
Anexo 3 Informe completo de la separabilidad espectral de ROIs (Superior e Inferior) de la imagen Sentinel-2	103
Anexo 4 Toma de coordenadas para áreas de entrenamiento y generación de ROIs para zonas de cultivo de interés y otros	104
Anexo 5 Toma de coordenadas para áreas de entrenamiento al azar y generación de ROIs para proceso de validación con zonas de cultivo de cacao y banano.....	104
Anexo 6 Toma de coordenadas para áreas de entrenamiento al azar y generación de ROIs para proceso de validación con zonas de otras coberturas y asentamientos humanos	105
Anexo 7 Determinación de áreas de entrenamiento para generación de ROIs en la zona de estudio	105
Anexo 8 Determinación de áreas de entrenamiento al azar para generación de ROIs y posterior análisis estadístico	106
Anexo 9 Generación de puntos al azar para método estadístico	106

Anexo 10 Obtención de datos aleatorios para examinar las diferencias estadísticas mediante la utilización de la función "Extract Multi Values to Point"	107
Anexo 11 Datos obtenidos a partir de herramienta SAGA de QGIS para el análisis de matriz de confusión y Coeficiente Kappa.....	107
Anexo 12 Gráfico Q-Q para normalidad de los datos de Maximum Likelihood	108
Anexo 13 Gráfico Q-Q para normalidad de los datos de Random Forest	108

Índice de Ecuaciones

Ecuación 1 <i>NDVI</i>	48
Ecuación 2 <i>RESI</i>	48
Ecuación 3 Coeficiente Kappa	53

Introducción

Actualmente la expansión e intensificación de la agricultura no regulada e insostenible afectan al medio natural mundial, continuando con las causas de los desperfectos ambientales (Graziano Ceddia, 2019). Asimismo, la creciente demanda de alimentos, piensos y energía está provocando una mayor intensidad de la agricultura y cambios en el paisaje. Dicha expansión agrícola insostenible conduce principalmente a la degradación de la tierra, especialmente en zonas fértiles (Leng et al., 2020).

En América Latina este proceso es más prominente, con tasas de incremento de áreas agrícolas sustancialmente superiores a la media mundial. Durante las últimas décadas, se han producido cambios significativos en el paisaje natural, en gran parte debido al aumento de la producción agrícola destinada a la exportación. Un ejemplo claro de este fenómeno es el aumento del 43% en la superficie cultivada en cinco países del continente, desde el inicio del ciclo agrícola en el año 2000 hasta su finalización en 2011 (Vega et al., 2015). Por consiguiente, es esencial establecer una planificación territorial exhaustiva que permita la evolución ordenada de los sistemas territoriales en la región.

A nivel nacional, la situación es similar, especialmente debido a que el país se destaca en la exportación de diversos productos agrícolas, como bananos, maíz amarillo, caña de azúcar, soja, cacao, entre otros (Miguel Ayala et al., 2016). Según el Ministerio de Agricultura, Ganadería, Acuacultura y Pesca (MAGAP), la degradación del suelo en el territorio ecuatoriano está estrechamente relacionada con el uso de prácticas agrícolas insostenibles (Ponce Cevallos, 2016). El problema también se presenta en la provincia de Los Ríos, la segunda región agrícola más apta del país (OECD & Food and Agriculture Organization of the United Nations, 2016). Los principales cultivos de la provincia son el cacao, el café, el arroz, el banano y la palma africana. Semejante a la provincia del Guayas que se destaca por cultivos de arroz, banano, cacao, café, al igual de la exportación y producción para consumo familiar de caña de azúcar, maíz y pastos.

Se ha comprobado que el crecimiento agrícola puede tener un mayor impacto en la pobreza que otros sectores de la economía, siendo una de las actividades más

afectadas por el cambio climático. Se estima que la disminución de la productividad y la producción agrícola provocada por este fenómeno retrasa el logro de las metas de reducción de la pobreza y representará un desafío para la seguridad alimentaria (NU. CEPAL, 2017).

Comprender los factores involucrados en la diversificación adecuada de la tierra agrícola es una base importante para que los gobiernos implementen políticas y programas para mejorar el sector agrícola a fin de reducir la pobreza, especialmente para aquellos que viven en el sector rural. En el marco del Objetivo 15 de la Agenda 21, centrado en el desarrollo rural y de la agricultura, la Comisión Brundtland destacó la importancia de que los gobiernos locales fomenten la adopción de modelos de producción agrícola diversificados que respalden una agricultura sostenible y amigable con el medio ambiente (NU. CEPAL, 2017). Asimismo, se ha observado un incremento en la práctica de extender la cubierta forestal mediante la transición a nuevos bosques secundarios, la cual se ha vuelto cada vez más común (Muñoz Murcillo et al., 2020).

La tecnología de teledetección ayuda en la solución de problemas de planificación territorial. Se utiliza para evaluar la degradación y conservación de recursos naturales, especialmente el uso del suelo y los cambios en la cobertura vegetal. La teledetección espacial contribuye a la investigación ambiental, permitiendo monitorear procesos dinámicos a partir de imágenes capturadas desde órbitas estables y repetitivas (Gonzaga Aguilar, 2014).

La cantidad y calidad de la información obtenida de imágenes satelitales depende de su procesamiento, y actualmente se están utilizando métodos basados en inteligencia artificial. Los algoritmos de aprendizaje automático más utilizados incluyen Random Forest (RF), bagging (embolsado), boosting (aumento), árbol de decisión, red neuronal artificial, máquina de vectores soportados, k-nearest neighbor (vecino más próximo), etc. Estos algoritmos aprenden la relación entre los datos de predicción y respuesta sin ninguna suposición sobre la distribución de los datos (Ok et al., 2012). Con un conjunto de datos y parámetros adecuados, pueden encontrar el mejor modelo para nuevos datos. El algoritmo de clasificación RF es superior a muchos algoritmos

basados en árboles debido a su robustez ante el ruido y la resistencia al sobreajuste (Ok et al., 2012). Ha sido demostrado que clasifica con éxito tanto grupos complejos como homogéneos de plantas con una precisión global del 88.37% (Jay et al., 2009). Cutler et al. (2007), clasificaron especies de plantas invasoras usando RF y datos ecológicos (Cutler et al., 2007). En este sentido, el procesamiento de imágenes satelitales basado en inteligencia artificial ofrece una gran oportunidad para el monitoreo de la vegetación y la detección de cambios en la cobertura vegetal.

La detección mediante procesamiento y clasificación de imágenes multiespectrales puede ayudar a compensar la falta de información actual y objetiva sobre la distribución de cultivos. Además de los métodos de clasificación tradicionales, se han propuesto alternativas basadas en inteligencia artificial y aprendizaje automático para mejorar la precisión de los resultados (Cánovas García et al., 2016). Estas estrategias innovadoras pueden ser útiles para mejorar la eficacia de la detección de cultivos y para desarrollar estrategias de gestión agrícola más sostenibles. En este proyecto se comparará los resultados obtenidos con el método tradicional de Máxima Verosimilitud y el método de inteligencia artificial Bosques Aleatorios, para la clasificación de imágenes multiespectrales de las provincias Guayas y Los Ríos, con enfoque en la detección de cultivos de cacao, banano y palma africana.

CAPITULO 01

CONTEXTUALIZACIÓN Y FUNDAMENTACIÓN TEÓRICA DE LA INVESTIGACIÓN

Contextualización y fundamentación teórica de la investigación

1.1. Problema de investigación

1.1.1. Planteamiento del problema

Para cubrir la creciente demanda de alimentos, originada por el aumento de la población y la variación en los hábitos alimenticios, es esencial mejorar la producción agrícola (Ngoma et al., 2021). Sin embargo, la expansión incontrolada de los cultivos y la aprobación del uso de tierras agrícolas sin un enfoque integral de la planificación de tierras ha generado problemas ambientales. Por ello, se han fijado objetivos en la Agenda 2030 para lograr un progreso sostenible (Muñoz Murcillo et al., 2020).

En la zona de influencia de las provincias del Guayas y Los Ríos, existe un reto recurrente relacionado con el cultivo agrícola de cacao, que, a pesar de ser una fuente sólida de economía, puede verse afectado por el cambio climático y generar consecuencias negativas en el medio ambiente (Sporchia et al., 2021). Algunos de estos impactos incluyen la contaminación del suelo y de los cuerpos de agua, tanto subterráneos como superficiales, como resultado de la expansión de los cultivos, que puede ser insostenible y producir residuos peligrosos o no peligrosos a través de las actividades de producción.

Por otro lado, el cultivo del banano también enfrenta problemas similares, con un aumento en el uso del agua durante la temporada de sequía debido a los efectos del cambio climático. Además, este tipo de prácticas agrícolas insostenibles pueden llevar a una generación de emisiones de carbono, al igual que otros procesos de producción o transporte de los productos (Coltro & Karaski, 2019). En este sentido, es probable que la expansión del cultivo de banano en la zona de estudio, que es una fuente importante de PIB para las provincias de Los Ríos y Guayas, tenga un impacto significativo en el carbono.

En cuanto al crecimiento de la producción de palma africana, existe una fuerte conexión con la degradación ambiental y la emisión de gases de efecto

invernadero, especialmente dióxido de carbono. A pesar de las críticas en contra de su impacto en los agricultores pequeños, los expertos creen que las restricciones comunes en la gestión de la producción de aceite de palma limitan las iniciativas para prevenir y reducir estos efectos negativos (Castellanos-Navarrete, 2021).

Por lo tanto, es necesario implementar soluciones para abordar los efectos negativos de la expansión agrícola insostenible en el medio ambiente. Una opción es la planificación regional eficaz basada en información geográfica actual y objetiva. El uso de imágenes satelitales multiespectrales de acceso público podría ayudar a identificar medidas preventivas o de mitigación específicas relacionadas con la expansión agrícola. Sin embargo, estos métodos son relativamente nuevos y requieren ser evaluados en estudios de casos reales para garantizar su precisión.

1.1.2. Diagnóstico

El Plan de Desarrollo y Ordenamiento Territorial 2015-2019; alineado al Plan Nacional de Desarrollo 2017-2021 del Cantón San Jacinto de Buena Fe de la Provincia de los Ríos, incluye la clasificación de uso de suelo de 2013, donde la mayor parte del área rural se dedica a la agricultura, con 62.65% del área total, especialmente cacao, palma africana y maíz, plátano, ábaco, palma y otros menos importantes (GADM San Jacinto de Buena Fe, 2017).

Por otro lado, el Plan de Desarrollo y Ordenamiento Territorial 2016-2020 del Cantón Valencia, indica que el 93% del territorio (90374,2 ha) está dedicado a la agricultura intensiva y extensiva de cultivos como el banano, palma, café, cacao y agricultura tropical principalmente; existen zonas donde producción agrícola se destina al consumo local y es un pilar para la soberanía alimentaria (GADM de Valencia, 2016).

En caso del cantón El Empalme, en el Plan de Desarrollo y Ordenamiento Territorial - El Empalme 2015-2024, destacan las plantaciones forestales (bosque plantado): caña guadua, teca, balsa y caucho, con extensión de 1.502,5 hectáreas (2,31%). Además, es importante la presencia del cultivo de la palma africana que cubre un área de 990 ha (1,52 %) y de los misceláneos indiferenciados (mezcla de algunos cultivos) que tienen una extensión de 2.044

hectáreas representando el 3,15 %. Otros cultivos presentes en el cantón son el café, soya, mandarina, naranja, maracuyá y sandía con superficies relativamente pequeñas (GADM El Empalme, 2015).

1.1.3. Pronóstico

Mediante el uso de sensores remotos, aplicando métodos precisos para detectar cultivos y otras coberturas terrestres, es posible identificar características relevantes de la vegetación tales como: estado de salud, actividad fotosintética, estrés hídrico; además de un posible uso no sustentable de la tierra. Esto ayudará a los formuladores de políticas estatales o privados a responder de manera rápida y eficiente a cualquier problema identificado.

1.1.4. Formulación del problema

¿Tienen el algoritmo de aprendizaje automático Bosques Aleatorios una precisión mayor, menor o equivalente al de Máxima Verosimilitud para detectar cultivos de cacao, banano y palma africana?

1.1.5. Sistematización del problema

¿Es conveniente el uso de los índices espectrales calculados a partir de las imágenes multiespectrales de sensores remotos para la detección de los cultivos agrícolas?

¿Se pueden utilizar los métodos de Máxima Probabilidad y Bosques Aleatorios para representar la distribución y características de los cultivos de cacao, banano y palma en la zona de estudio?

¿Qué precisión tienen los métodos de clasificación de imágenes multiespectrales derivados del Sentinel-2?

1.2. Objetivos

1.2.1. Objetivo general

Evaluar los métodos de Bosques Aleatorios y Máxima Verosimilitud en la detección de los principales cultivos de la zona norte de las provincias del Guayas y de Los Ríos, a partir de las imágenes satélites Sentinel-2.

1.2.2. Objetivos específicos

- Ejecutar el preprocesamiento de imágenes multiespectrales del satélite Sentinel-2 con las funciones del software especializado (ENVI y ArcGIS) para la posterior clasificación.
- Realizar la clasificación supervisada de las imágenes de Sentinel-2 con los métodos de Máxima Verosimilitud y Bosques Aleatorios e índices espectrales para detectar los cultivos principales en la zona de estudio.
- Llevar a cabo el análisis estadístico de las clasificaciones para evaluar la precisión de los dos métodos en la detección de cultivos de cacao, banano y palma africana.

1.3. Justificación

Es importante conocer las características de los cultivos de cacao, banano y palma en la zona de estudio debido a su importancia económica y su amplio uso del área. La identificación de estas características a través de una clasificación individual es crucial para comprender su distribución en la región norte del Guayas y Los Ríos y mejorar la planificación territorial.

La elaboración de un análisis detallado de la distribución y las características de los cultivos, representados en un Sistema de Información Geográfica (SIG), puede permitir la adopción de medidas que tengan en cuenta la interacción entre el medio ambiente y los factores biológicos y abióticos. Una gestión integral de estos procesos agrícolas puede ayudar a orientar acciones hacia la reducción de emisiones de carbono, contribuyendo a cumplir uno de los Objetivos de Desarrollo Sostenible (ODS 15) que se encaminan en fomentar un uso sostenible

de los ecosistemas terrestres, cuyo principal enfoque es la prevención de la deforestación y la degradación del suelo.

Con base en lo anterior, se puede concluir que la evaluación de las tierras agrícolas es crucial para la planificación futura. Los planificadores estratégicos utilizan diversas técnicas como observaciones en el campo, muestreo, análisis de laboratorio, sensores remotos y SIG para concebir proyectos de planificación espacial (Elbeih, 2021). La combinación de análisis de imágenes multifocales o espectrales y el uso de SIG permite una gestión global de la cobertura del suelo en diferentes áreas y su representación.

Desde otro punto de vista, existe la necesidad de determinar la tasa de crecimiento de los límites agrícolas y caracterizar el problema. Este tema ha sido abordado por varios autores que han centrado sus investigaciones en una metodología basada en el uso de SIG (Viteri-Salazar & Toledo, 2020). Aunque todavía se han mostrado nuevos métodos tecnológicos que emplean los algoritmos de aprendizaje autónomo, involucrando el mapeo en específico de características que pueda poseer la cobertura del suelo (Aksoy et al., 2022).

Aunque las técnicas de teledetección pueden parecer modernas, el uso de imágenes multiespectrales para analizar la distribución agrícola y determinar su frontera ha estado vinculado a los SIG durante mucho tiempo (Viteri-Salazar & Toledo, 2020). Sin embargo, con el avance en la informática, han surgido nuevos métodos, muchos de los cuales utilizan algoritmos de aprendizaje automático, como el Bosques Aleatorios utilizado en el proyecto de investigación en cuestión (Aksoy et al., 2022).

A partir de la aparición de nuevos métodos de inteligencia artificial en sustitución de los métodos tradicionales, surgió la necesidad de comprobar la eficacia de estos diversos procesos metódicos. Dentro de la investigación, se utilizará metodologías para comparar los algoritmos de precisión tradicionales y nuevos (de inteligencia artificial) con relación a la detección y manipulación de imágenes satelitales (Sentinel-2) en zonas de cultivos de interés (cacao, banano y palma) en los diferentes cantones de las provincias del Guayas y Los Ríos, a través del cálculo de los índices espectrales NDVI (Natural Differential Vegetation Index) y RESI (Red-Edge Spectral Indices).

1.4. Fundamentación teórica de la investigación

1.4.1. Marco conceptual

1.4.1.1. Producción agrícola sostenible

Se crea un software para administrar la sostenibilidad en base a las condiciones únicas de un sitio, zona o región que posee características similares (Zepeda-Jazo, 2018). Dichas zonas agrícolas tienen en común factores como los tipos de cultivos y variedades de plantas, el clima, las prácticas agrícolas locales, las plagas y sus contrincantes naturales, y la situación socioeconómica del agricultor y el mercado (Zepeda-Jazo, 2018).

1.4.1.2. Ordenamiento territorial

El ordenamiento territorial debe entenderse como un procedimiento general encaminado a alterar el espacio geomorfológico para alcanzar el uso óptimo, racional y sostenible del territorio por parte de los habitantes en un proceso de participación y concertación (Márquez Poblete & Veloso Pérez, 2021).

1.4.1.3. Teledetección

La teledetección permite el análisis de cambios de variables biofísicas de la vegetación, siendo útiles como indicadores o métodos de detección para cada tipo de alteración que pueda encontrarse en la vegetación ante los procesos o factores diversos que existan. De igual forma, mantiene disposiciones actualizadas ante las series temporales de periodos largos para los diferentes datos satelitales nivel global brindados por servicios como “Google Earth Engine” (GEE), permitiendo la posibilidad de analizar procesos en la cobertura del suelo a través de la verificación de modelos teóricos y fórmulas de modelos de predicción (Hernandez-Clemente & Hornero, 2021).

Durante las últimas décadas, la teledetección ha evolucionado en una herramienta clave para diversos ámbitos de nuestra sociedad. Se han visto múltiples ejemplos de su aplicación como fundamento para la toma de decisiones en la administración eficiente de la agricultura y la silvicultura, los recursos naturales, la meteorología, la planificación y la cartografía del uso del suelo, entre otros (Labrador García et al., 2012). Además de aparecer como un instrumento

esencial para responder a los requisitos de un proceso en específico, ya que ofrece un medio no destructivo de proporcionar información recurrente desde la escala local a la global de forma sistemática, permitiendo así la caracterización de la variabilidad espaciotemporal dentro de un área determinada (Weiss et al., 2020).

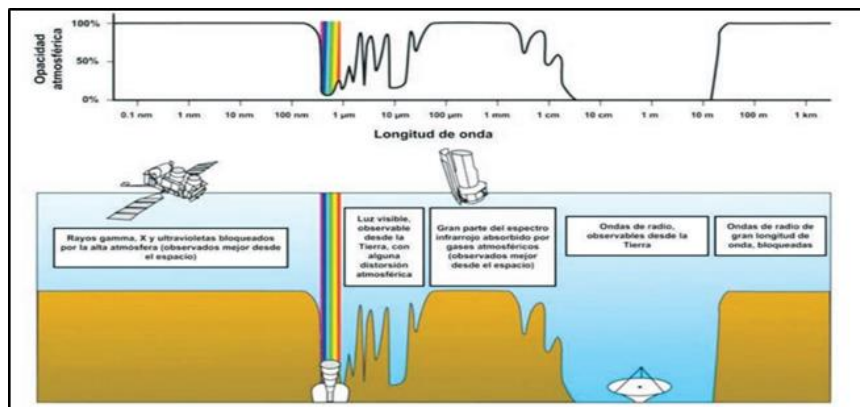
1.4.1.4. Imágenes multiespectrales

Una imagen multiespectral es un conjunto de columnas y filas que forman una matriz de datos numéricos que representan la intensidad de la energía electromagnética irradiada o pronunciada por los objetos en la superficie terrestre. Estas imágenes se pueden grabar en varias bandas del espectro electromagnético (individualmente), al igual que el satélite Landsat 8 tiene 11 bandas espectrales, es decir, una misma escena se captura en diferentes bandas (banda 1, banda 2, banda 3, etc.) (Sánchez-Méndez & Arguijo-Hernández, 2018).

Las imágenes multiespectrales pueden verse como una matriz 3D que tiene dos dimensiones espaciales y una espectral. Las imágenes multiespectrales tienen en general de 3 a 10 bandas espectrales, mientras que las hiperespectrales tienen cientos de bandas. Los escáneres funcionan en regiones concretas del espectro electromagnético. Como las imágenes multiespectrales representan la misma zona espacial de la tierra leyéndola varias veces, la información obtenida se presenta en forma de cubo (Vura et al., 2023). En particular, las características de propagación en la Tierra, y la posibilidad de comunicarse en el espacio a través de satélites a través de la atmósfera, hacen que esta banda de frecuencias sea ideal para una variedad de propósitos de transmisión de información, dando lugar a las radiocomunicaciones (Figura 1).

Figura 1

Opacidad atmosférica a la radiación electromagnética



Nota: Extraído de Ordoñez (2012)

1.4.1.5. Firmas espectrales

Las firmas espectrales permiten la identificación de diferentes tipos de cultivos, cuerpos de agua, suelos y otras tipologías de la corteza terrestre (Pech May et al., 2021).

Los satélites de observación de la Tierra son instrumentos que orbitan la misma y registran la radiación electromagnética emitida desde la zona terrestre. Su funcionamiento se basa en las llamadas firmas espectrales, la capacidad de ciertos objetos o sustancias para reflejar o emitir energía electromagnética (Pech May et al., 2021).

1.4.1.6. Sentinel-2

El proyecto Sentinel-2 es un cometido satelital europea para proporcionar datos de teledetección basados en tierra; el mismo satélite refleja de dos satélites (Sentinel-2A y Sentinel-2B) y provee observaciones de alta resolución espacial (10, 20 y 60 metros) (Yang et al., 2021).

1.4.1.7. Bandas espectrales

Se define una franja de frecuencias de onda en el espectro electromagnético como una banda espectral. Un ejemplo son las bandas de color rojo, verde y azul, que se extienden desde 0,4 hasta 0,7 μm de longitud de onda. Al combinar estas bandas en una sola imagen, se forma el espectro visible conocido como RGB (Rodríguez Pérez et al., 2015). Por ello, la información que obtienen los

detectores de satélites en la banda espectral se denomina banda, entendida como la "banda de imagen" de la teledetección.

1.4.1.7.1. Bandas del Sentinel-2

Los correspondientes requisitos de los usuarios han impulsado el diseño hacia un sistema multiespectral de observación de la Tierra fiable que cuenta con un Instrumento Multiespectral (MSI) con 13 bandas espectrales que van desde el visible y el infrarrojo cercano hasta el infrarrojo de onda corta. La resolución espacial varía de 10 m a 60 m (Tabla 1), dependiendo de la banda espectral, con un campo de visión de 290 km (European Space Agency & Fletcher, 2012). Esta combinación única de alta resolución espacial, amplio campo de visión y amplia cobertura espectral representa un gran paso adelante en comparación con otras misiones multiespectrales.

Tabla 1

Bandas de Sentinel-2

Número de banda	Nombre de banda	Longitud de onda central (nm)	Resolución espacial (m)
1	Aerosol costero	443	60
2	Azul	490	10
3	Verde	560	10
4	Rojo	665	10
5	Borde rojo de vegetación 1	705	20
6	Borde rojo de vegetación 2	740	20
7	Borde rojo de vegetación 3	783	20
8	Infrarroja cercana (NIR)	842	10
8A	Infrarroja cercana (NIR) estrecha	865	20
9	Vapor de agua	945	60
10	SWIR - Cirrus	1380	60
11	SWIR 1	1610	20
12	SWIR 2	2190	20

Nota: Extraído de European Space Agency (2012)

1.4.1.8. Clasificación de imágenes satelitales

La clasificación de imágenes hace referencia a la tarea de extraer capas de información de un mapa de bits de multidifusión (Arcgis Desktop, s.f.). A través

de esto, se pueden aplicar dos tipos de clasificación: supervisada y no supervisada.

1.4.1.8.1. Clasificación supervisada

La clasificación supervisada emplea las firmas espectrales obtenidas de las muestras de entrenamiento con el fin de categorizar las imágenes. Asimismo, es posible generar de manera sencilla un archivo de firma a partir de las muestras de capacitación, el cual será empleado por las herramientas de clasificación multivariante para clasificar la imagen (Arcgis Desktop, s.f.).

1.4.1.8.2. Clasificación no supervisada

El clasificador no supervisado encuentra clases espectrales (o grupos) en imágenes multibanda sin la intervención del analizador. La barra de herramientas de clasificación de imágenes va a favorecer a realizar una clasificación desatendida al proporcionar dirección a las herramientas de recopilación, la capacidad de comparar la calidad de la recopilación y el acceso a las herramientas de clasificación (Arcgis Desktop, s.f.).

1.4.1.8.3. Método Máxima Probabilidad

El Maximum Likelihood en inglés, también conocido como método de máxima verosimilitud, es un enfoque ampliamente utilizado en teledetección. En este método, se utilizan la matriz de covarianza y el vector de media de clase para reducir la distancia entre clases y aumentar la probabilidad de que una característica sea asignada a la clase seleccionada, empleando funciones discriminantes cuadráticas o lineales. Sin embargo, este método puede tener un ajuste excesivo a las probabilidades, lo que puede resultar en una clasificación errónea cuando se descartan los supuestos (Arai, 2020).

1.4.1.8.4. Método Bosques Aleatorios

El algoritmo de Random Forest en inglés, también conocido como método de Bosques Aleatorios, combina diferentes árboles de decisión para crear un modelo más robusto y confiable. Cada árbol se entrena con una muestra de datos y características elegidas al azar. Las predicciones se realizan haciendo en cada árbol y luego combinándolas a través de un proceso de votación. Este

método es ampliamente utilizado para identificar factores importantes, detectar datos atípicos y predecir variables categóricas y continuas (Waśniewski et al., 2020). Sin embargo, es importante llevar a cabo una validación adecuada para evitar resultados sesgados.

1.4.1.9. Índices espectrales

1.4.1.9.1. Índice NDVI

El Índice de Vegetación de Diferencia Normalizada (NDVI, en su idioma de inglés Normalized Difference Vegetation Index) proporciona un proceso metódico para la detección de los niveles de vegetación saludable, pudiendo aplicarse fácilmente en zonas de interés con diferentes escalas espaciales en toda superficie terrestre (Rhew et al., 2011). Los datos proporcionados por el indicador NDVI son recopilados empleando teledetección y son de acceso gratuito para el público en general. Siendo un beneficio para la replicación y comparación de estudios de casos en diferentes ubicaciones regionales (Rhew et al., 2011).

1.4.1.9.2. Índice RESI

El Índice de Bordes Rojos (RESI, en su idioma de inglés Red-Edge Spectral Indices) es un mejor indicador de la condición de la planta que el NDVI para cultivos de mediados a tardíos que han acumulado mucha clorofila (EOS Data Analytics, 2020). La razón es que la luz roja del borde (la que se usa en NDRE) puede penetrar más profundamente a través de las hojas que la luz roja (la que se usa en NDVI). Cuando se usa en grandes cantidades durante la temporada de crecimiento, es más útil que el NDVI, que a menudo se vuelve inexacto cuando las plantas acumulan la mayor cantidad de clorofila posible (EOS Data Analytics, 2020).

1.5. Marco referencial

En los últimos tiempos, se están haciendo cada vez más frecuentes los estudios de clasificación supervisada utilizando distintos algoritmos e imágenes de satélite. Estos estudios buscan encontrar la manera más efectiva de clasificar los

temas de investigación en función de las necesidades específicas de cada región. En este contexto, varios estudios han evaluado la precisión de los algoritmos empleados en este proyecto a partir de imágenes de satélite Sentinel-2 y con un objetivo de clasificación similar.

1.5.1. Estudio referencial 1

Santiago, Mallqui, y Ríos en el 2021 publicaron un estudio titulado "Mapeo de la cobertura vegetal en la subcuenca Quillcay (Ancash, Perú) con el clasificador de árbol de decisiones". En este trabajo, los autores se enfocaron en determinar la cobertura vegetal de la subcuenca en cuestión, utilizando el método de clasificación supervisada de árbol de decisiones. La investigación constó de varios pasos, que incluyeron una revisión bibliográfica y una validación de resultados finales (Santiago Bazan et al., 2021).

Inicialmente, identificaron los tipos de cobertura vegetal en la subcuenca a través de una revisión de fuentes secundarias. Luego, descargaron una imagen satelital de Landsat-8 desde el sitio web de USGS. La imagen fue corregida radiométrica y atmosféricamente, y se definieron las variables geográficas (altitud y pendientes). Después, se calcularon los índices de Humedad de Diferencia Normalizada (NDMI), se obtuvieron las firmas espectrales y se llevó a cabo la clasificación por medio del método Spectral Angle Mapper (con toma de puntos en campo). Finalmente, se completó la delimitación de la cobertura vegetal mediante árboles de decisión y se realizó una validación exhaustiva de los resultados (Santiago Bazan et al., 2021).

En el estudio realizado en la subcuenca en cuestión, se identificaron cuatro tipos dominantes de cobertura vegetal, incluyendo bofedal, matorral, pajonal y bosque, los cuales se encuentran en altitudes entre 3200 y 5600 metros sobre el nivel del mar, y en pendientes mayores a 8.5%. Un árbol de decisión fue creado con 14 nodos binarios, que incluyeron características tanto geográficas como espectrales, para discriminar píxeles de áreas sin o con poca vegetación. Los resultados de la validación fueron exitosos, con una precisión general del 90% en la matriz de confusión y un coeficiente de Kappa de 0.8998 (Santiago Bazan et al., 2021).

1.5.2. Estudio referencial 2

En el artículo "An evaluation of Landsat, Sentinel-2, Sentinel-1 and MODIS data for crop type mapping" por Peng Song, Huang, Hansen y Potapov (2021), se evaluó la capacidad de diferentes satélites para representar los cultivos de maíz y soya en Estados Unidos. Se comparó la utilidad y precisión de las imágenes de Landsat-8, Landsat-7, Sentinel-2, Sentinel-1 y MODIS. La investigación consistió en llevar a cabo experimentos de clasificación utilizando estas imágenes satelitales y datos del Departamento de Agricultura de los Estados Unidos (Song et al., 2021).

El estudio se llevó a cabo en bloques de 20 km², y se clasificaron de acuerdo con la intensidad promedio de cultivos. Se realizó un preprocesamiento y se diseñaron experimentos en cuatro categorías para evaluar la precisión de clasificación y la importancia de cada sensor, banda espectral y fecha de generación de imágenes satelitales. Para la clasificación se utilizó el algoritmo "árboles de decisión", que brindó dos resultados importantes: una estimación de la precisión de la clasificación si se tienen áreas de entrenamiento suficientes y correctas, y la selección y clasificación interpretando la estructura del árbol resultante (Song et al., 2021).

La eficacia y precisión potencial la evaluaron a través de varios sensores satelitales para la identificación de cultivos de maíz y soja en Estados Unidos. La investigación mostró una precisión general de 90% para todos los sensores, con Landsat y Sentinel-2 logrando una precisión de 96,8% y 96,6% respectivamente. Los investigadores concluyen que los datos de teledetección con resolución moderada pueden lograr una precisión nominal del 95% para cultivos en grandes extensiones (Song et al., 2021).

1.5.3. Estudio referencial 3

En el artículo "Comparison of Classification Algorithms for Detecting Typical Coastal Reclamation in Guangdong Province with Landsat 8 and Sentinel 2 Images", publicado por Bin Ai, Ke Huang, Shaojie Sun, Zhuokai Jian y Xiaoding Liu en el año 2022, se buscó determinar la técnica adecuada para representar la recuperación costera común en la provincia de Guangdong, China, a gran

escala, mediante el uso de métodos de clasificación, incluyendo el árbol de decisiones (Ai et al., 2022).

Los datos satelitales de Landsat-8 y Sentinel-2 se obtuvieron sin nubes y se realizaron correcciones radiométricas, atmosféricas y geométricas. Se evaluaron diferentes algoritmos de clasificación, incluyendo bosque aleatorio, máquina de vectores de soporte, árbol de decisión y métodos orientados a objetos basados en reglas. El mejor algoritmo para visualizar geográficamente la recuperación costera se determinó mediante análisis estadístico usando matrices de confusión y la prueba de McNemar (Ai et al., 2022).

Los autores de este artículo encontraron que la precisión en la clasificación de la zona costera osciló entre un 83.05% y un 88.57% cuando se aplicaron los algoritmos a las imágenes de Sentinel-2 en el oeste de la zona costera. En el este, la precisión varió entre un 86.47% y un 88.50%. El mejor algoritmo para Sentinel-2 fue el orientado a objetos con un coeficiente Kappa de 0.827. Para las imágenes de Landsat-8, los resultados fueron similares, siendo el mejor clasificador orientado a objetos en el oeste de la zona costera (86.39%) y el método de Árbol de Decisiones con una precisión del 82.62% en el oeste y del 81.36% en el Delta del río y el este (Ai et al., 2022).

1.5.4. Estudio referencial 4

En un artículo publicado por Campos-Taberner et al. (2020) titulado "Deep Learning para la clasificación de usos de suelo agrícola con Sentinel-2", se comparó el desempeño de un algoritmo de inteligencia artificial llamado 2-BiLSTM con otros métodos de clasificación, incluyendo árboles de decisión, k-vecinos más cercanos, redes neuronales, máquinas de soporte vectorial y bosques aleatorios, para clasificar usos de suelo agrícola en imágenes de Sentinel-2. El estudio se llevó a cabo en la Comunidad Valenciana, España, utilizando dos imágenes de Sentinel-2 que abarcan un área total de 220 x 220 km². Se eligieron 16 usos de suelo no urbanos para comparar los diferentes métodos de clasificación y se valoraron con una matriz de confusión y un índice de precisión.

Entre los resultados, se encontró que el algoritmo 2-BiLSTM fue el más preciso con 98.6% de identificación, seguido de los algoritmos 1-BiLSTM, 3-BiLSTM, 4-

BiLSTM y bosques aleatorios. El algoritmo de máquinas de soporte vectorial, redes neuronales, k-vecinos más cercanos y árboles de decisión tuvieron una precisión menor. El algoritmo 2-BiLSTM obtuvo la mayor precisión en la clase tierra con un 99.9% de identificación precisa, mientras que la clase frutales tuvo los errores más grandes. Una vez seleccionado el mejor algoritmo, se utilizó para generar un mapa de usos de suelo y un mapa de confianza de clasificación (Campos-Taberner et al., 2020).

1.5.5. Estudio referencial 5

En el estudio publicado por Xiao et al. (2020) titulado “Sentinel-2 red-edge spectral indices (RESI) suitability for mapping rubber boom in Luang Namtha Province, northern Lao PDR”, un nuevo algoritmo llamado índices espectrales de borde rojo (RESI) fue diseñado y utilizado para mapear las plantaciones de caucho en la provincia de Luang Namtha, en el norte de Laos. La precisión global y el coeficiente kappa de los mapas alcanzaron un 92,50% y 0,91, respectivamente. La estimación del área de plantaciones maduras de caucho en 2018 fue de 771,2 km² en la provincia. Según los autores, la creación del RESI es un nuevo enfoque para mapear las plantaciones de caucho utilizando las bandas de borde rojo de Sentinel-2 A/B a nivel regional. La investigación demuestra que los algoritmos normalizados para mapear plantaciones de caucho son viables con imágenes de Landsat y Sentinel-2.

CAPITULO 02

METODOLOGÍA DE LA INVESTIGACIÓN

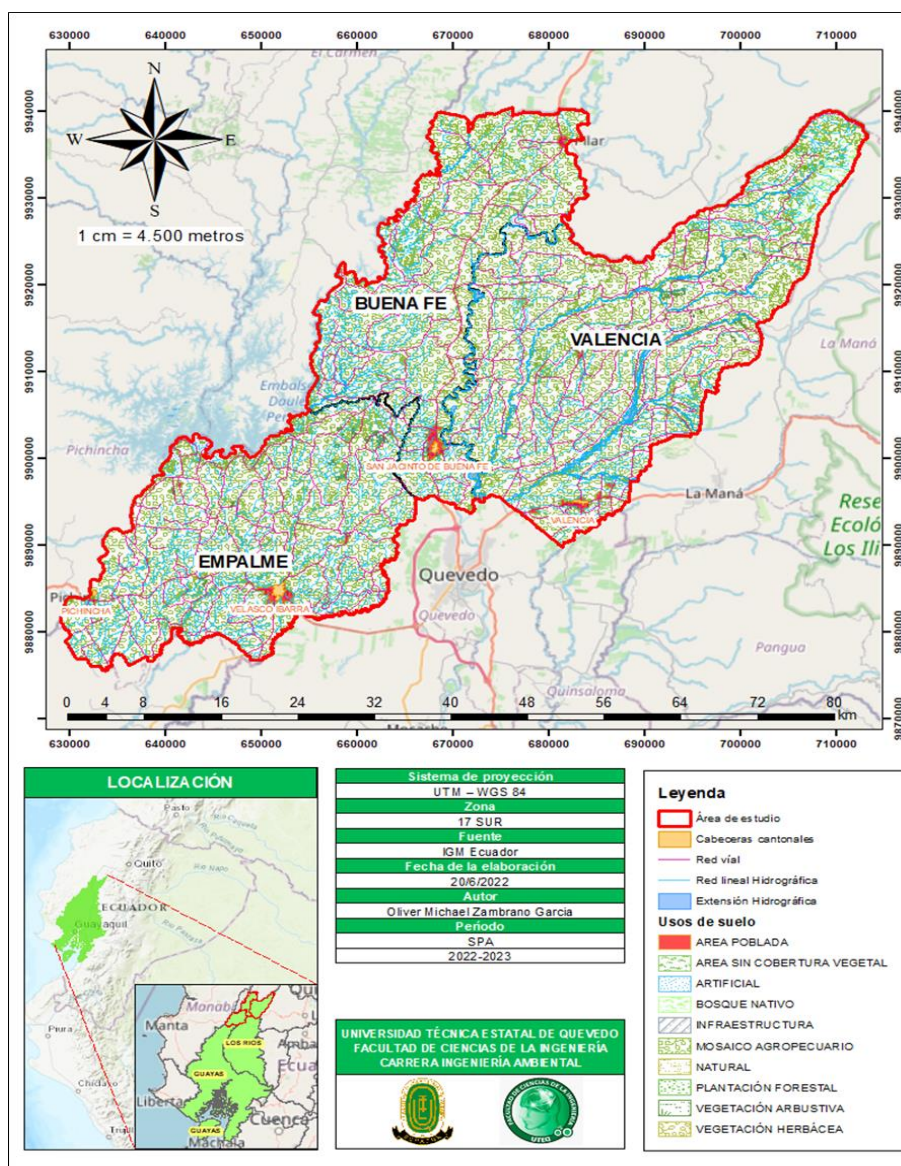
Metodología de la investigación

2.1. Localización

El área de estudio se sitúa en la zona norte de las provincias del Guayas y de Los Ríos incluye el cantón de El Empalme perteneciente a la provincia del Guayas, y a los cantones Valencia y Buena Fe pertenecientes a la provincia de Los Ríos (Figura 2).

Figura 2

Mapa de localización del área de estudio: Cantones Buena Fe, Valencia y Empalme



Nota: Autor (2023)

En estos cantones las actividades agropecuarias juegan un papel fundamental, siendo pilares de la economía de las provincias correspondientes. El proyecto de investigación se concentrará en el procesamiento de imágenes satelitales para la detección de cultivos de interés (banano, cacao y palma), comparando la efectividad de las metodologías de clasificación, así como el nivel de utilidad que poseen las imágenes del satélite Sentinel-2.

2.2. Tipo de investigación

La investigación desarrollada es de tipo exploratoria, cuantitativa y transversal.

- **Experimental:** En el proyecto de investigación se evaluó la precisión del método de inteligencia artificial Bosques Aleatorios a partir de imágenes satelitales de Sentinel-2 en el área de estudio, y se realizó una comparación de métodos para la clasificación supervisada de imágenes multiespectrales.
- **Cuantitativa:** En el proyecto de investigación, la validación de resultados se realizó a partir de una matriz de confusión y se obtuvo el coeficiente Kappa para cada clasificación supervisada ejecutada. La selección del mejor método clasificador permitió obtener la información objetiva sobre la extensión de los cultivos de cacao, banano y palma africana en el área de estudio.
- **Transversal:** En el proyecto de investigación, se compararon la precisión del método de inteligencia artificial Bosques Aleatorios y el método tradicional Máxima Verosimilitud, en la detección de cultivos de cacao, banano y palma africana en el área de estudio a partir de las imágenes satelitales Sentinel-2.

2.3. Métodos de investigación

El proyecto de investigación utilizó el siguiente método:

- **Método inductivo:** En el proyecto de investigación, para determinar el mejor clasificador para las imágenes multiespectrales, se ejecutó un

proceso desde lo particular (procesamiento de las imágenes, cálculo de índices espectrales y validación de resultados) a lo general (determinación del mejor método clasificador a partir de una imagen multiespectral).

2.4. Fuentes de recopilación de información

La información para el proyecto de investigación se obtuvo de las siguientes fuentes:

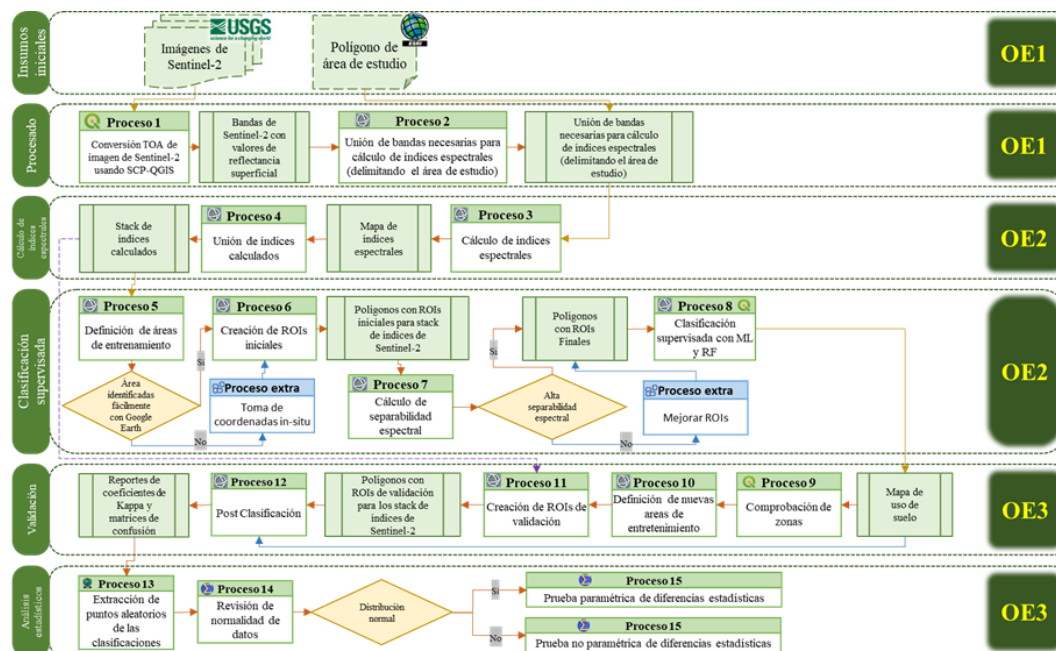
- **Fuentes primarias:** Para la definición de áreas de entrenamiento, previo a la ejecución de la clasificación supervisada con los métodos ya mencionados, se tomaron coordenadas in-situ (puntos referenciales) para el cultivo de cacao; de la misma forma, se tomaron coordenadas de los puntos de localización de las demás coberturas de interés usando la información disponible a través del programa Google Earth Pro. En el mismo sentido, se tomaron coordenadas in-situ y desde Google Earth Pro para la validación de resultados con la matriz de confusión y el coeficiente Kappa.
- **Fuentes secundarias:** Se descargaron imágenes multiespectrales de Sentinel-2 desde el sitio web de la USGS (USGS-U.S. Geological Survey, s.f.), adicionalmente se recurrió al sitio web SciHub – Copernicus (Copernicus Open Access Hub, s.f.) para obtener la información sobre las características de las imágenes. Estas mismas que sirvieron para el cálculo de los índices espectrales, la clasificación supervisada con los métodos posteriormente descritos y la evaluación de su precisión para la validación de resultados.

2.5. Diseño de la investigación

En el proyecto de investigación se empleó un proceso metódico para lograr los objetivos específicos (OE) propuestos (Figura 3).

Figura 3

Flujograma del proceso metodológico de la investigación



Nota: Autor (2023)

A continuación, se muestra los métodos y técnicas manejados en el desarrollo de la investigación:

2.5.1.OE 1: Ejecutar el preprocesamiento de imágenes multiespectrales del satélite Sentinel-2 con las funciones del software especializado (ENVI y ArcGIS) para la posterior clasificación

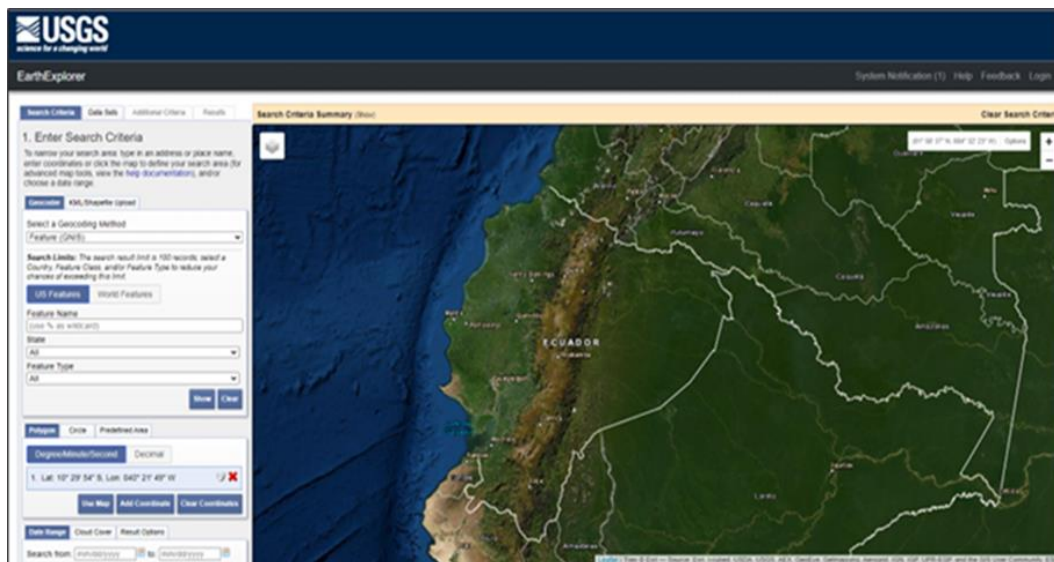
Para el cumplimiento del objetivo propuesto sobre el preprocesamiento de imágenes multiespectrales del satélite Sentinel-2, se plantearon los siguientes puntos:

2.5.1.1. Búsqueda de imágenes

Se empleó el método de búsqueda para las imágenes multiespectrales que brinda el satélite Sentinel-2, a través de un sitio web que presta el servicio geológico estadounidense nombrando como “USGS Earth Explorer”. Dicho portal digital permitió el acceso, visualización y manipulación de datos geoespaciales de múltiples satélites orbitantes a nuestro planeta Tierra (Figura 4).

Figura 4

Servicio web Geológico "USGS Earth Explorer"



Nota: USGS Science for a changing world (USGS-U.S. Geological Survey, s.f.)

Posterior a lo mencionado se implementó el siguiente proceso para la adquisición de las imágenes satelitales:

- La búsqueda de dichas imágenes se ejecutó en una de las secciones de configuración del sitio web "Geocorder", la cual se ubica en la opción de "Search Criteria".
- En el apartado mencionado se estableció ciertos parámetros para buscar una localización en específico, de manera que se preseleccionó opciones de "World Features"
- Seguido se estableció un país ("Country"), la clase características ("Feature Class") y el tipo de función ("Feature Type"), esto según lo que se deseó aplicar en la investigación, habiendo características de cultivos con funciones agrícolas en la zona más cercana al área de estudio el cual son las provincias de Los Ríos y Guayas.
- Una vez aplicado, se seleccionó "Data sets" y se especificó el conjunto de datos de los diferentes proveedores de imágenes satelitales, tomando como referencia el de Sentinel, seguido de la categoría "Sentinel-2".
- Posteriormente se estableció un rango temporal para disponer de diferentes imágenes a lo largo de ese periodo; esto se realizó en la

sección inferior de “Search from ... establecer fecha from a otra fecha más actual”.

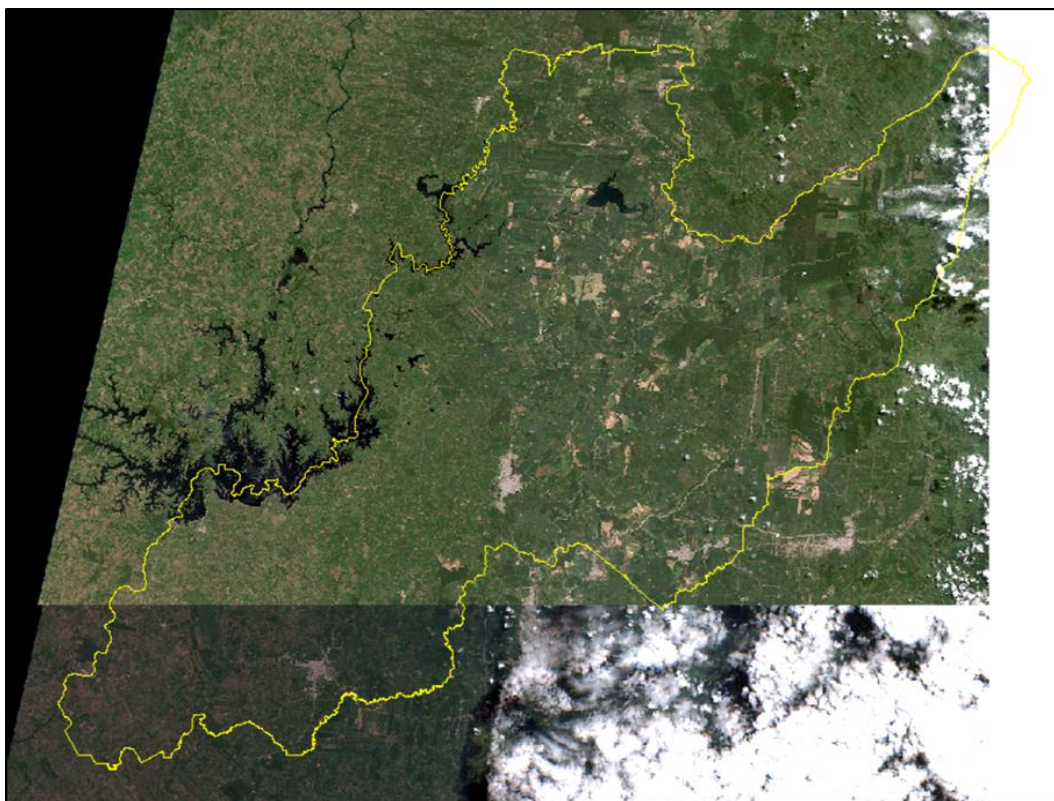
- f) Una vez aplicado el proceso de búsqueda, se identificó la mejor imagen satelital (con la menor presencia de nubosidad sobre el área de estudio) y se realizó la respectiva descarga en formato multibanda.

2.5.1.2. Comprobación de imágenes

Se empleó un método de comprobación digital con ayuda del software ArcGIS, donde se insertaron y sobrepusieron las imágenes previamente obtenidas (en una banda por cada imagen satelital) y de la capa shapefile de los cantones de las zonas de interés en el territorio ecuatoriano (en la tabla de contenido del software). Se llevó a cabo una comparación de los colores de las bandas de las imágenes con la capa shapefile de los cantones para resaltar las diferencias y mejorar la presentación visual (Figura 5).

Figura 5

Comprobación de nubosidad en el área de estudio



Nota: Autor (2023)

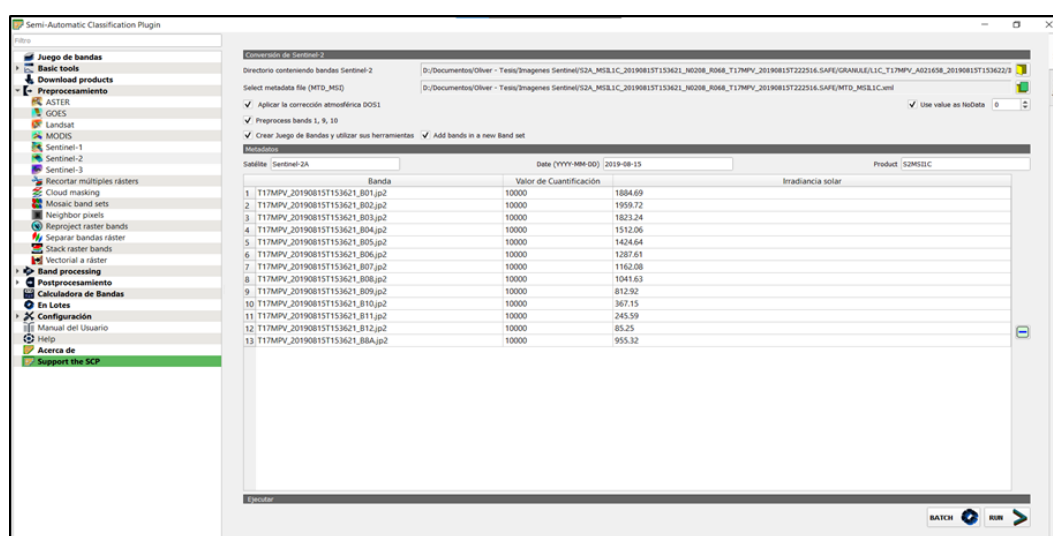
2.5.1.3. Corrección atmosférica de las imágenes satelitales Sentinel-2

Se empleó el método de corrección atmosférica para mejorar la imagen satelital que brinda el satélite Sentinel-2. Dicho proceso se utilizó debido a que la imagen disponible para descarga no estaba corregida y presenta valores afectados por la influencia de factores atmosféricos (nivel “1C”).

Este tipo de corrección se ejecutó en el software QGIS, donde se instaló previamente el paquete complementario llamado “Semi-Automatic Classification Plugin” (SCP) (Congedo, 2021). Una vez instalado el paquete, se dispuso a aplicar los siguientes pasos: se seleccionó “SCP” → “Preprocesamiento” → “Sentinel-2” (Figura 6). Una vez en la sección del apartado de metadatos, se escogió la ruta de la carpeta que posee información de todas las bandas espectrales, al igual que el archivo “MTD_MSI” (el cual contiene información de todas las bandas espectrales de la imagen satelital). Para finalizar, se seleccionó la opción “Aplicar la corrección atmosférica” con un visto y se corrió el programa en la opción “RUN”.

Figura 6

Corrección atmosférica de bandas espectrales Sentinel-2 con el Semi Automatic Classification Plugin del año 2021



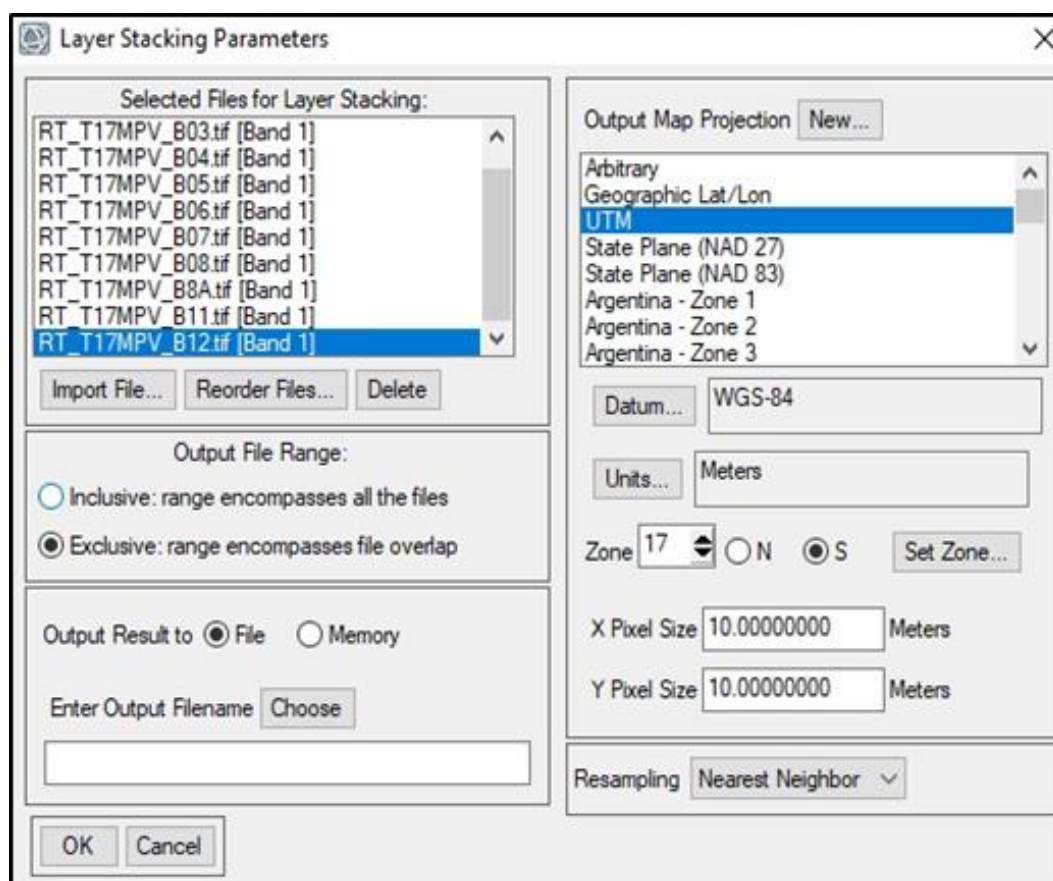
Nota: Autor (2023)

2.5.1.4. Unión de las bandas espectrales

Se empleó un método de unión de las imágenes del satélite Sentinel-2 previamente procesadas (comprobadas y corregidas), a través de una opción “Layer Stacking” localizada en la caja de herramientas digital del software “ENVI” (Figura 7). Las bandas espectrales agrupadas en una sola imagen multispectral sirvieron para el posterior cálculo de los índices espectrales y visualización aplicando la combinación RGB (referido al espectro visible). Se utilizó la caja de herramientas de unión de imágenes para combinar las bandas seleccionadas, logrando así una extensión espacial completa de la imagen satelital.

Figura 7

Configuración de combinación de bandas espectrales según polígono del área de estudio



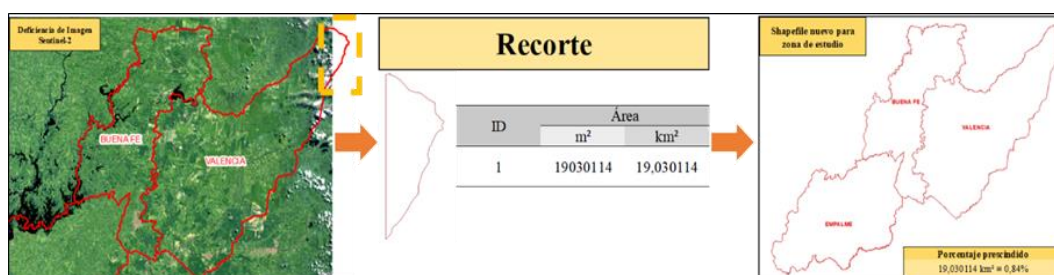
Nota: Autor (2023)

2.5.1.5. Extracción y recorte de polígono real para imagen

Inicialmente, se delimitó un área de estudio a través de la creación de un polígono que incluía los cantones El Empalme, Buena Fe y Valencia. Esta extracción se realizó utilizando el software ArcMap y la herramienta de "Export Data". Sin embargo, la imagen de Sentinel-2 descargada no cubría toda el área prevista, por lo que se tuvo que ajustar el polígono en función de la cobertura real de la imagen satelital. Se recortó manualmente en ArcMap con la herramienta "Cut Polygons" y se excluyó una pequeña porción del área total (1%) para evitar errores en la clasificación supervisada (Figura 8).

Figura 8

Definición de polígono recortado del área de estudio en función de las coberturas reales



Nota: Autor (2023)

Como siguiente paso, se empleó un proceso de recorte de la imagen multibanda (resultado de la unión de bandas espectrales), y se realizaron recortes con los límites de las zonas del estudio. Se inició con la extracción del polígono que concentre a los cantones de Valencia, Buena Fe y El Empalme (áreas de estudio) a través del software ArcGIS, mediante una selección de los cantones de interés en la tabla de atributos de la capa del shapefile de cantones del terreno ecuatoriano y siendo exportado con la herramienta "Export Data" de la opción "Data".

Una vez extraído el polígono de los cantones (formato shapefile), se ingresó dicho shapefile al software "ENVI", con el fin de recortar las imágenes multiespectrales de Sentinel-2 con la herramienta "Resize Data" para la zona de estudio. Continuamente, se escogió las imágenes satelitales y se seleccionó la opción "Spatial Subset ROI/EVF" para elegir el archivo shapefile de la zona de

interés de estudio (previamente procesada), y finalizando con la selección de la carpeta destinada al recorte.

2.5.2.OE 2: Realizar la clasificación supervisada de las imágenes de Sentinel-2 con los métodos de Máxima Verosimilitud y Bosques Aleatorios e índices espectrales para detectar los cultivos principales en la zona de estudio

Para el cumplimiento del objetivo propuesto sobre el cálculo de los índices espectrales (NDVI y RESI) a partir de las imágenes multiespectrales preprocesadas, se planteó los siguientes puntos:

2.5.2.1. Cálculos de índices espectrales

Para la aplicación de los índices espectrales se utilizaron 2 tipos de ecuaciones que representan las bandas espectrales (variables), según la siguiente tabla:

Tabla 2

Ecuaciones de Índices Espectrales

Índice	Ecuación	Detalle	Referencia
NDVI	Ecuación 1 NDVI $\frac{NIR - RED}{NIR + RED}$	Dónde, NIR significa banda de infrarrojo cercano, mientras que RED significa banda roja. Por definición, los valores de NDVI van de -1 a 1 con valores más altos para vegetación densa y valores negativos para áreas con aguas superficiales, rocas y nubes.	(Pech May et al., 2021)
RESI	Ecuación 2 RESI $\frac{\rho_{RE3} + \rho_{RE2} - \rho_{RE1}}{\rho_{RE3} + \rho_{RE2} + \rho_{RE1}}$	Donde, dónde ρ_{Rojo} , ρ_{RE1} , ρ_{RE2} , ρ_{RE3} , ρ_{NIR} y ρ_{SWIR1} son los valores de superficiales de las bandas roja, RE1, RE2, RE3, RE3, NIR y SWIR1 en las imágenes de las imágenes Sentinel-2, respectivamente.	(Xiao et al., 2020)

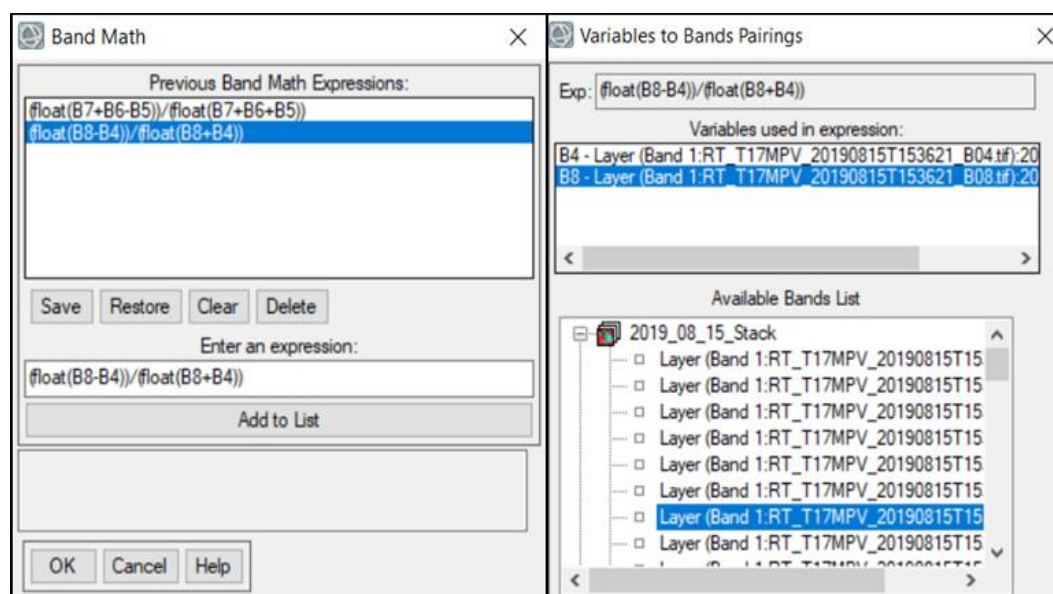
Nota: Autor (2023)

2.5.2.2. Cálculo de índices espectrales en el software

Para cálculo de los índices espectrales NDVI y RESI se utilizó el software ENVI, y se aplicó la herramienta Band Math. A través de esto se asignó la **¡Error! No se encuentra el origen de la referencia.** y **¡Error! No se encuentra el origen de la referencia.** en la sección de expresión de variables (manteniendo una escritura lineal en la aplicación de esta), donde se seleccionó las bandas correspondientes a la formulación de cada ecuación (Figura 9).

Figura 9

Ejemplo del cálculo para índice espectral en software ENVI - Herramienta Band Math



Nota: Autor (2023)

2.5.2.3. Creación de áreas de entrenamiento

Se implementó un proceso de clasificación supervisada utilizando una selección de elementos de la imagen llamada "área de entrenamiento", cuya clasificación perteneciente se conoce con una alta precisión. Este tipo de clasificación de grupos de imágenes es el proceso mediante el cual a cada elemento de la imagen se le establece una categoría, en función de los atributos presentes en las áreas de entrenamiento (Borrás et al., 2017).

Para definir las respectivas categorías de interés, se analizó la información disponible en el software Google Earth y datos en campo. Consecuentemente,

se realizó la toma y registro de varias coordenadas en relación con cada categoría de interés en el área de estudio. Para ello fue implementado el formato mostrado en la tabla 3.

Tabla 3

Formato para registro de coordenadas para definir áreas de entrenamiento

Categoría		Coordenada UTM-WGS84		Origen	
Detalle	Codificación	X	Y	Google Earth	In-situ
Zona de cultivos de cacao	ZC_C				
Zona de cultivos de banano	ZC_B				
Zona de cultivos de palma africana	ZC_PA				
Zona de otras coberturas	ZC_O				
Zonas sin vegetación aparente	ZS_VA				
Zona de asentamientos humanos	ZA_H				

Nota: Autor (2023)

Para la toma de coordenadas se utilizó la aplicación móvil “Russian Topo Maps”, donde se determinó varias zonas tales como ZC_C, ZC_O, ZS_VA y ZA_H, las cuales se consideraron más accesibles para dicha toma in-situ y mantener registros fotográficos de la misma. Consecuentemente, para las zonas ZC_B y ZC_PA se utilizó el software Google Earth, con el fin de identificar las zonas de interés a través de la interfaz de navegación satelital que brinda éste.

2.5.2.4. Regiones de interés (ROIs)

Para definir las regiones de interés de las imágenes satelitales de Sentinel-2, se utilizaron coordenadas en ENVI 5.3 para cada stack de índices calculados. Se crearon 60 regiones de entrenamiento para Sentinel-2, divididas en categorías predefinidas, con un total de 1974 píxeles de 100 m² (19,74 km²). Además, se aplicó el principio de separación espectral en cada área relevante, lo que permitió confirmar la existencia de una relación entre las categorías o clases generadas (teniendo en cuenta que cada píxel tiene características únicas y distintas) (Vera Rojas, 2020).

2.5.2.5. Aplicación del clasificador Máxima Probabilidad (ML)

Después de crear las regiones de interés, se aplicó un clasificador de Máxima Probabilidad (ML) en ENVI 5.3 en ambos stacks de índices. Este método utiliza una distribución normal de las estadísticas para cada clase y determina la probabilidad de que un píxel pertenezca a una clase específica. Finalmente, se asigna a cada píxel a la clase con la probabilidad más alta, aunque si la probabilidad es menor que un límite específico, el píxel permanece sin clasificar. Este proceso se utilizó para clasificar el stack de índices obtenido de la imagen de Sentinel-2.

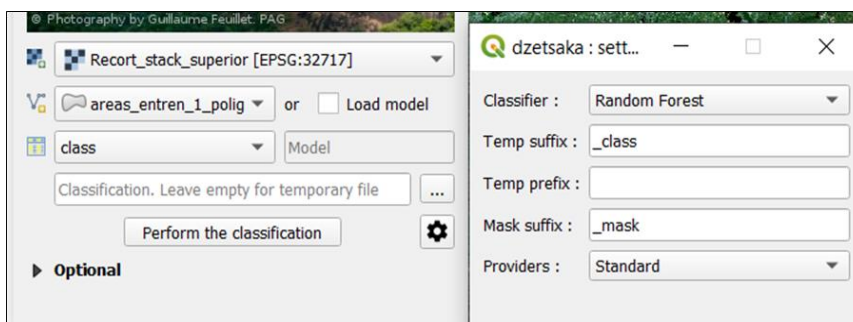
Para clasificar, se importó el stack de índices de Sentinel-2 al software ENVI y se utilizó la herramienta "Maximum Likelihood Classification" dentro de la sección "Classification" del Toolbox. Se eligió el stack de índices y las categorías previamente definidas en las regiones de interés (Ha et al., 2020). Se estableció un límite de probabilidad (0,3 a 1), por lo que para cada píxel se asignó a la clase con la probabilidad más alta.

2.5.2.6. Aplicación del clasificador Bosques Aleatorios (RF)

Se utilizó un clasificador no paramétrico de inteligencia artificial en QGIS para clasificar la imagen de Sentinel-2 en el stack de índices calculados. Primero, se importaron las regiones de interés y los stacks de índices al software y luego se instaló el complemento "Dzetsaka" (Nogueira de Brito & Silva Brito, 2020), (Titti et al., 2022). Una vez abierto el complemento, se verificó que el stack de índices y las regiones de interés tengan la misma proyección cartográfica y se eligió el clasificador RF y la columna "class" para ejecutar la clasificación (Figura10).

Figura 10

Herramienta algorítmica de clasificación "Dzetsaka" para Random Forest



Nota: Autor (2023)

2.5.3.OE 3: Llevar a cabo el análisis estadístico de las clasificaciones para evaluar la precisión de los dos métodos en detección de los cultivos de cacao, banano y palma africana

Para el cumplimiento del objetivo propuesto sobre el análisis estadístico de las clasificaciones fue una técnica valiosa para evaluar la precisión de los métodos de detección de cultivos y pudo ser aplicado de manera efectiva en la detección de los cultivos de cacao, banano y palma africana. Así, se planteó los siguientes puntos:

2.5.3.1. Matriz de confusión

Dentro este proceso se aplicó una matriz de confusión o tabla de contingencia, constando con dos tipos de entradas, siendo para cada categoría (de cultivos) representado en un fila y columna dentro de la tabla. Estas dos entradas compararon los valores reales con los resultados de la clasificación obtenida, y el grado de la precisión se identificó mediante la diagonal de la matriz (destacando la cantidad de píxeles reales y la clasificación concordante por categoría). Por consiguiente, las que se muestran verticalmente representaron el porcentaje de píxeles reales que se confundirá con el porcentaje horizontal de píxeles que se desordenaron después de la clasificación (Borrás et al., 2017).

Esta tabla de contingencia mostrada en la tabla 4 fue elaborada para cada clasificador en función de las imágenes satelitales obtenidas, y de la ayuda del software ENVI 5.3.

Tabla 4

Modelo de matriz de contingencia con aditamento de resultados del coeficiente Kappa

	Clasificación	Categoría vertical			Total ↓
		C1	C2	C3	
Categoría Horizontal	C1				
	C2				
	C3				
Total →					
Tipo de Error	Omisión				
	Comisión				
Precisión General					

Nota: Análisis de uso de suelo a partir de imágenes satelitales Sentinel-2 en el cantón Buena Fe, provincia de Los Ríos (Vera Rojas, 2020)

2.5.3.2. Coeficiente Kappa

Adicionalmente se aplicó el coeficiente Kappa para evaluar la precisión. Siguiendo las indicaciones de la investigación de Velásquez sobre el “Análisis de uso de suelo a partir de imágenes satelitales Sentinel-2 en el cantón Buena Fe, Provincia de Los Ríos” (Vera Rojas, 2020), donde destaca la ecuación para dicho cálculo:

Ecuación 3

Coeficiente Kappa

$$k = \frac{N \sum_{i=1}^r X_{ii} - \sum_{i=1}^r (X_{i+} * X_{+i})}{N^2 - \sum_{i=1}^r (X_{i+} * X_{+i})}$$

Siendo **N** simboliza el número total de observaciones, **r** el número de clases en la matriz, **X_{i+}** simboliza el total de observaciones en la fila **i**, y **X_{+i}** el total de observaciones en la columna **i**, y **X_{ii}** el número de observaciones en el elemento tanto como en la fila y columna **i** (Vera Rojas, 2020).

A través del cálculo se pudo indicar el nivel de ajuste o precisión, la interpretación de los rangos de valores del coeficiente es como se muestra en la Tabla 5:

Tabla 5

Nivel de valorización del índice de Kappa

Valor del Coeficiente k	Fuerza de Concordancia
0,00	Pobre
0,01 – 0,20	Leve
0,21 – 0,40	Aceptable
0,41 – 0,60	Moderada
0,61 – 0,80	Considerable
0,81 – 1,00	Casi perfecta

Nota: Ladis y Koch (1997). *The measurement of observer agreement for categorical data. Biometrics*. Referenciado por Cerda y Villaroel (2008).

2.5.3.3. Análisis estadístico

Se llevó a cabo un análisis estadístico para determinar la presencia de diferencias significativas entre las clasificaciones de imágenes Sentinel-2. Se seleccionaron puntos al azar de las cuatro clasificaciones, siendo entre 350 a

400 puntos proporcional al tamaño del área de estudio (2256,10274 km²). Se eligió un punto por cada clasificación por kilómetro cuadrado. Para seleccionar los puntos al azar, se utilizaron las herramientas "Create Random Points" y "Extract Multi Values to Point" del software ArcGIS 10.8.

Para evaluar las diferencias significativas entre las clasificaciones de las imágenes de Sentinel-2, se llevó a cabo un análisis estadístico con el software SPSS. Debido a que los datos no cumplieron con los criterios de normalidad, se aplicó la prueba no paramétrica de Wilcoxon-Mann-Whitney para muestras independientes, también conocida como "Suma de Rangos de Wilcoxon" (Sánchez Turcios, 2015), (Marrón Ramos et al., 2022).

La prueba de Wilcoxon-Mann-Whitney combina los datos de dos muestras en una sola. No se calcula la diferencia entre las medianas, pero se utilizó para comparación estadística debido a su capacidad para determinar la diferencia de dispersión de datos entre dos grupos. Esta diferencia se establece a partir de una selección de datos calculados para los dos grupos comparados (Smida et al., 2022).

2.6. Recursos y materiales

2.6.1. Programas informáticos

- **ArcMap - ArcGIS [versión 10.8.12790]:** Es un componente de la plataforma ArcGIS, un software de SIG (sistemas de información geográfica) desarrollado por Esri. Se utilizó para el preprocesamiento y visualización de imágenes satélite Sentinel-2.
- **ENVI [versión 5.3]:** Es un software de procesamiento de imágenes de sensores remotos, que permite el análisis de imágenes de satélite, incluyendo imágenes Sentinel-2. Se utilizó para el preprocesamiento y visualización de imágenes, permitiendo clasificar cultivos a través de métodos de máxima verosimilitud.
- **Google Earth Pro [versión 7.3.6.9345]:** Es un software de visualización geográfica que permitió explorar imágenes satelitales y otra información geográfica en una interfaz gráfica de usuario intuitiva. En el contexto de la

detección de cultivos en la zona, Google Earth Pro fue utilizado para visualizar imágenes Sentinel-2 y verificar los resultados de la clasificación.

- **QGIS [versión 3.26.0]:** Es un software de SIG de código abierto que permite la visualización, análisis y edición de datos geográficos. En el contexto de la detección de cultivos, QGIS fue utilizado para visualizar imágenes preprocesadas y verificar los resultados de la clasificación utilizando métodos de Bosques Aleatorios.
- **Russian Topo Maps [versión 6.9.1]:** Es un software que permite visualizar mapas topográficos. En el contexto de la detección de cultivos, Russian Topo Maps fue utilizado como una herramienta complementaria para la visualización de información geográfica y determinación de coordenadas geográficas.

2.7. Sitios web

- **USGS Earth Explorer.** Es un sitio web mantenido por el Servicio Geológico de los Estados Unidos (USGS) que permite a los usuarios buscar y descargar imágenes satelitales de una amplia variedad de satélites, incluyendo Sentinel-2. Con el sitio web, los usuarios pueden definir áreas geográficas de interés y establecer criterios de fecha y resolución para limitar los resultados de la búsqueda. Una vez que se han seleccionado las imágenes adecuadas, los usuarios pueden descargarlas en formatos que sean compatibles con una amplia variedad de programas informáticos, incluyendo ArcGIS, QGIS y ENVI.

CAPITULO 03

RESULTADOS Y DISCUSIÓN

Resultados y discusión

3.1. Preprocesamiento de imágenes satelitales

3.1.1. Selección de imagen satélite Sentinel-2

De acuerdo con el área de estudio, se seleccionaron 2 imágenes multiespectrales disponibles en la base de datos web USGS. La imagen multiespectral 1 denota una fecha de registro del 15 de agosto del 2019 (Figura 11), y corresponde a un 85% del área total de la zona de estudio. Mientras que la imagen multiespectral 2 denota una fecha posterior del 20 de agosto del 2019 (Figura 12), cubriendo un 15% del área total de la zona de interés.

Tabla 6

Características de imagen 1 de Sentinel-2

Imagen 1	
S2A_MSIL1C_20190815T153621_ N0208_R068_T17MPV_20190815T222516	
Resolución espacial (m)	10, 20 y 60
Número de bandas espectrales	13
Fecha	2019-08-15T15:36:21.024Z
Porcentaje de cobertura de nubes	1,3742
Nivel de procesamiento	1C
DATUM	WGS84
Proyección de mapa	Universal Transversal de Mercator
Zona UTM	17 Sur

Nota: USGS Science for a changing world (USGS-U.S. Geological Survey, s.f.)

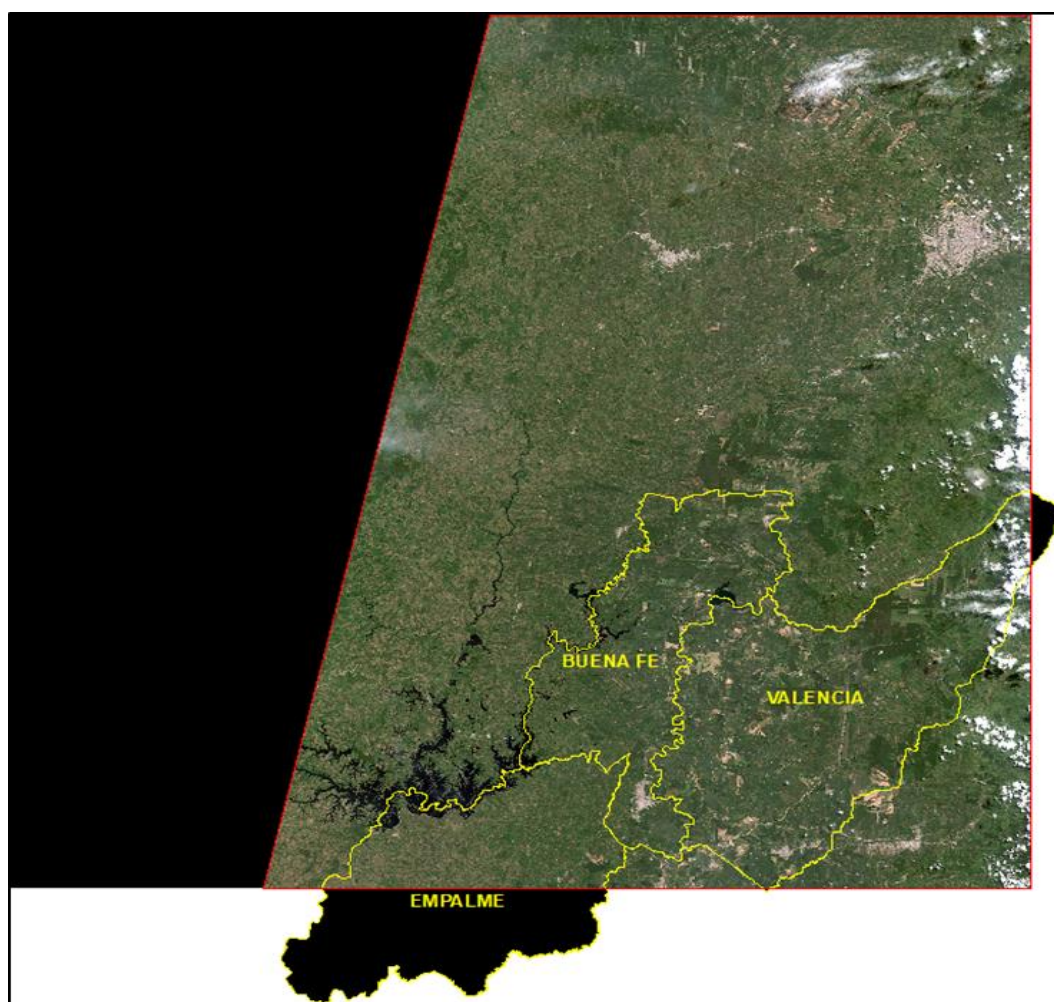
Según la Tabla 6, la primera imagen adquirida posee características de 3 tipos de resoluciones espaciales (10, 20, 60 m), contando con 13 bandas espectrales. Estas resoluciones también permiten caracterizar las diferentes bandas con una mayor precisión. Para la resolución de 10 metros, se utilizaron las bandas 2, 3, 4 y 8; para la resolución de 20 metros, se utilizaron las bandas 8-A, 11 y 12; y para la resolución de 60 metros, se utilizaron las bandas 1, 9 y 10. Por otro lado, la imagen presentó un porcentaje de cobertura de nubes muy bajo (1,3742%), siendo perfecta para la aplicación de los posteriores procesos. Asimismo, mostró

un nivel de procesamiento de 1C, identificando la necesidad de aplicar una corrección atmosférica.

Según la investigación con las imágenes satelitales Sentinel-2, la imagen número 1 fue elegida debido a su alcance para implementar los procedimientos posteriores. Según la Figura 11, se sobrepuso el área de estudio en la imagen 1 para identificar el contenido a abarcar en la zona, demostrando vacíos de resolución espacial en el área de estudio. A través de esto, la mayor proyección de nubosidad en la imagen se presenta en la zona este, donde la importancia de niveles bajos de nubosidad es precedente para la selección de la imagen, a través de la visualización se catalogó a la imagen como óptima para el proceso del estudio.

Figura 11

Imagen 1: Combinación RGB de Sentinel-2 (15/08/2019) con área de estudio sobrepuesta



Nota: USGS Science for a changing world (USGS-U.S. Geological Survey, s.f.)

La tabla 7 presenta las características de la imagen 2 que son similares a las de la imagen 1 (resolución espacial, bandas espectrales). La cobertura de nubes en la imagen 2 constituye 29,4774% a diferencia de la imagen 1 (1,3742%). Ambas imágenes son de nivel 1C, siendo necesario aplicar una corrección atmosférica.

Tabla 7

Características de imagen 2 de Sentinel-2

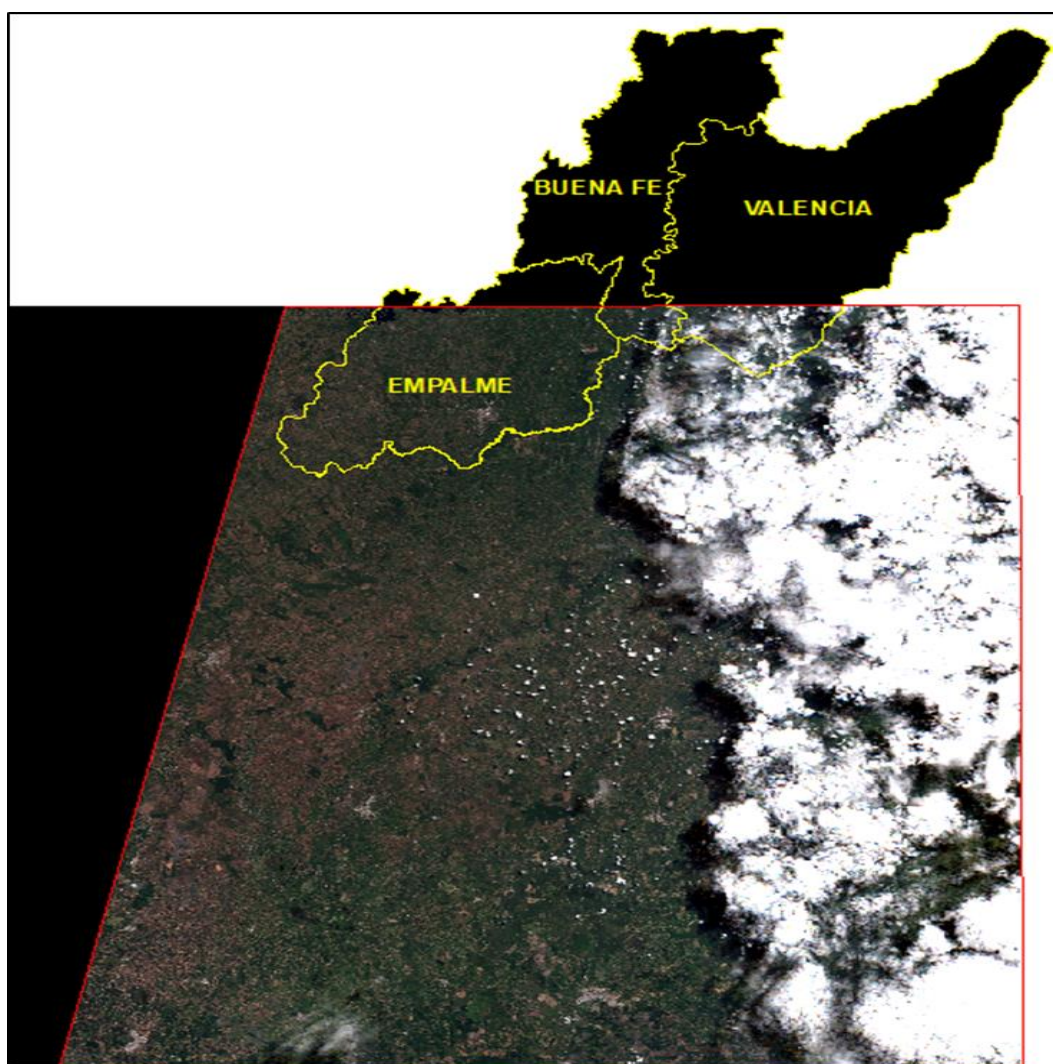
Imagen 2	
S2B_MSIL1C_20190820T153629_	
N0208_R068_T17MPU_20190820T202350	
Resolución espacial (m)	10, 20 y 60
Número de bandas espectrales	13
Fecha	2019-08-20T15:36:29.024Z
Porcentaje de cobertura de nubes	29,477399999999996
Nivel de procesamiento	1C
DATUM	WGS84
Proyección de mapa	Universal Transversal de Mercator
Zona UTM	17 Sur

Nota: USGS Science for a changing world (USGS-U.S. Geological Survey, s.f.)

Debido a la insuficiencia de la imagen 1, que no cubre toda la zona en estudio, se buscó una segunda imagen que pudiera cubrir la parte inferior faltante. La figura 12 muestra la presencia de nubes en el sector este de la imagen, aunque esta cubre una amplia zona en el área de estudio, fue considerada adecuada y aplicable para los siguientes procesos.

Figura 12

Imagen 2: Combinación RGB de Sentinel-2 (20/08/2019) con área de estudio sobrepuesta



Nota: USGS Science for a changing world (USGS-U.S. Geological Survey, s.f.)

3.1.2. Unión de bandas espectrales y recorte de imagen para área de estudio

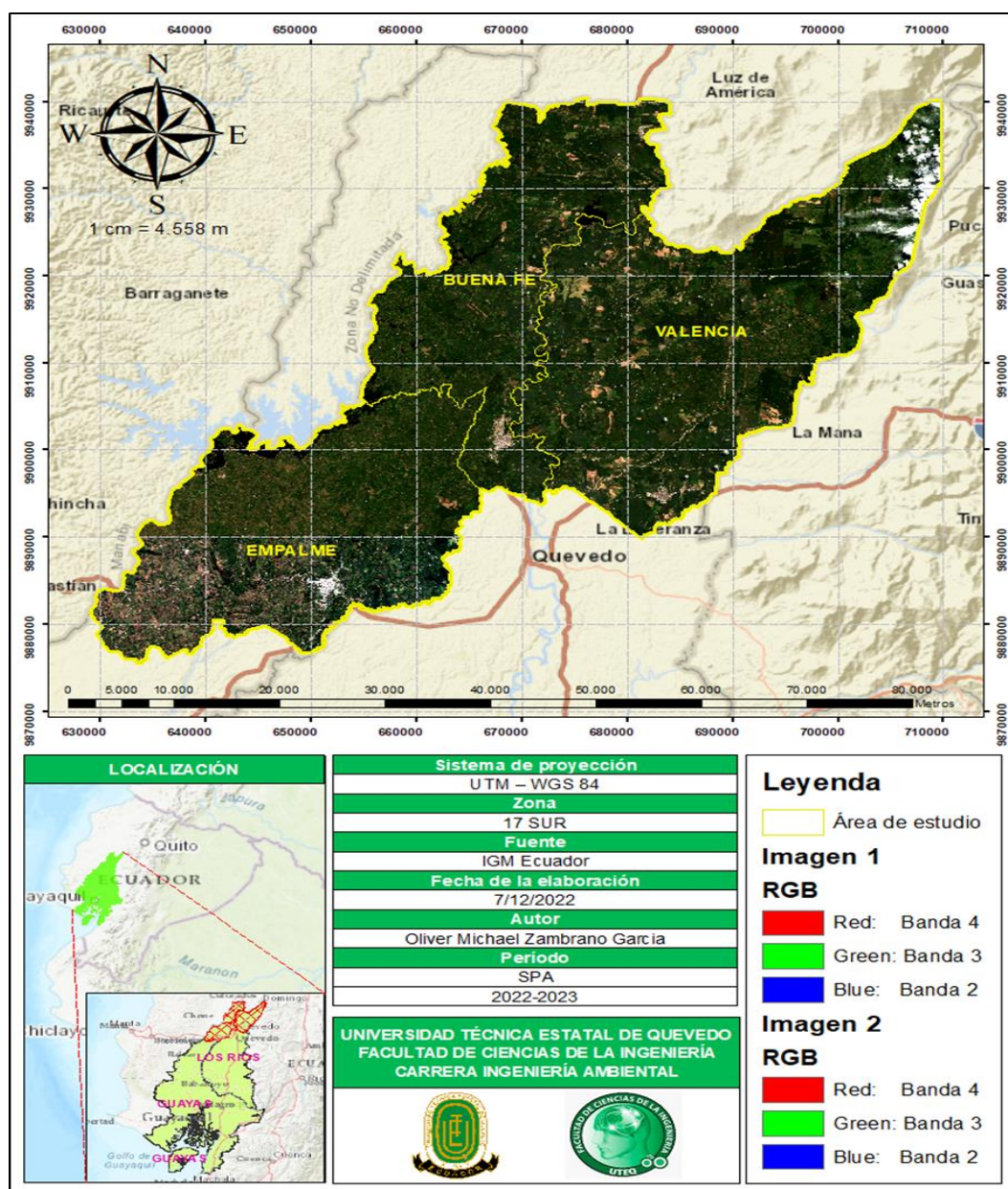
Tras la aplicación de la corrección atmosférica para las bandas espectrales de ambas imágenes Sentinel-2, se continuó con la unión de estas bandas espectrales procesadas para generar la combinación RGB, calcular los índices espectrales y realizar la clasificación supervisada.

Para la imagen multiespectral, se combinaron las bandas espectrales 2 (azul), 3 (verde), 4 (rojo), 5 (Borde rojo de vegetación 1), 6 (Borde rojo de vegetación 2), 7 (Borde rojo de vegetación 3), 8 (Infrarroja cercana), 8A (Infrarroja cercana

estrecha), 11 (SWIR 1) y 12 (SWIR 2). Este proceso se aplicó a ambas imágenes (1 y 2) y luego se recortó para ajustarse al área de estudio antes de continuar con los cálculos de índices espectrales y la clasificación supervisada. Además, para superponer las imágenes con el área de estudio, se utilizó la imagen 1 para cubrir la parte inferior de los cantones Buena Fe y Valencia, ya que en la imagen 2 dicha zona presenta una alta cantidad de nubes que dificultan la aplicación de los siguientes cálculos. Este resultado se puede ver en la Figura 13, donde se representaron las bandas espectrales 4, 3 y 2 respectivamente.

Figura 13

Recorte de área de estudio para imagen preprocesada de Sentinel-2



Nota: Autor (2023)

3.2. Clasificación supervisada de las imágenes de Sentinel-2

3.2.1. Cálculo de índices espectrales NDVI y RESI

El proceso de cálculo de los índices espectrales se llevó a cabo utilizando el software ENVI y se visualizaron en el software ArcMap (Figura 12). El objetivo principal de la investigación era clasificar de manera supervisada a través de dos métodos, y para ello, se hizo uso del cálculo y agrupación de los índices espectrales (stack de índices) (Figura 13). La herramienta "Band Math" en ENVI se utilizó para calcular los índices espectrales a partir de las ecuaciones 1 y 2. Para ello según las ecuaciones de los índices espectrales, se plantearon las líneas de código para insertar en la herramienta de ENVI (tabla 8).

Tabla 8

Línea de código aplicado para índices espectrales

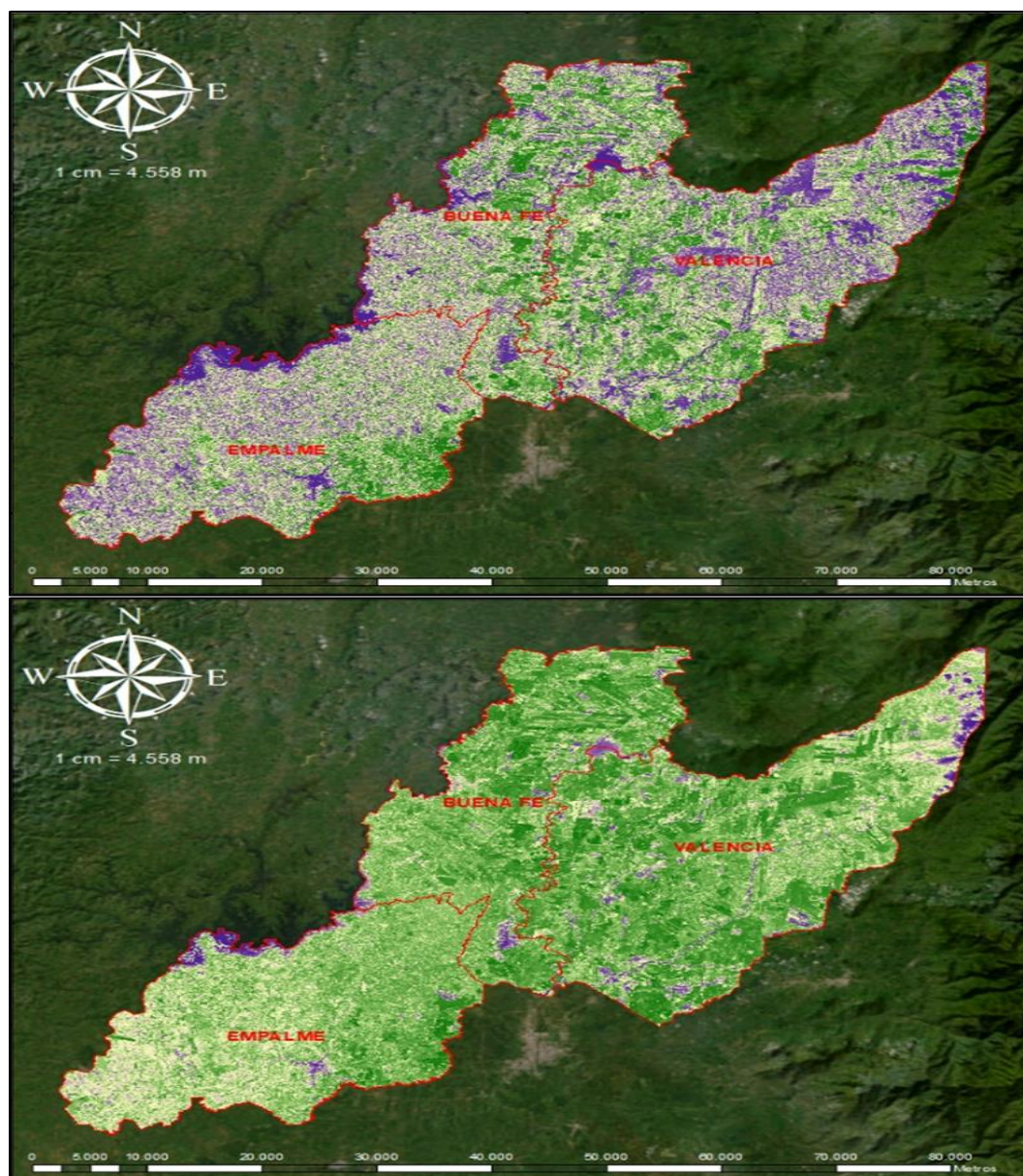
Índice espectral	Ecuación	Línea de código
NDVI	$NDVI = \frac{B8 - B4}{B8 + B4}$	((float((b8)-(b4)))/(float((b8)+(b4))))
RESI	$RESI = \frac{B7 + B6 - B5}{B7 + B6 + B5}$	((float((b7)+(b6)-(b5)))/(float((b7)+(b6)+(b5))))

Nota: Autor (2023)

De acuerdo con la Figura 14, el NDVI presentó un rango de valores que varía desde -0,314202 hasta 0,829794; los valores más altos indican la presencia de densa vegetación verde en el área de estudio. Por otro lado, el RESI tuvo valores que oscilaron desde 0,183321 hasta 0,82621, y los valores más elevados en la representación muestran una mayor actividad fotosintética de la vegetación de la zona.

Figura 14

Representación de Índices Espectrales: NDVI (A) & RESI (B)



Nota: Autor (2023)

3.2.2. Cálculo de Stack de índices

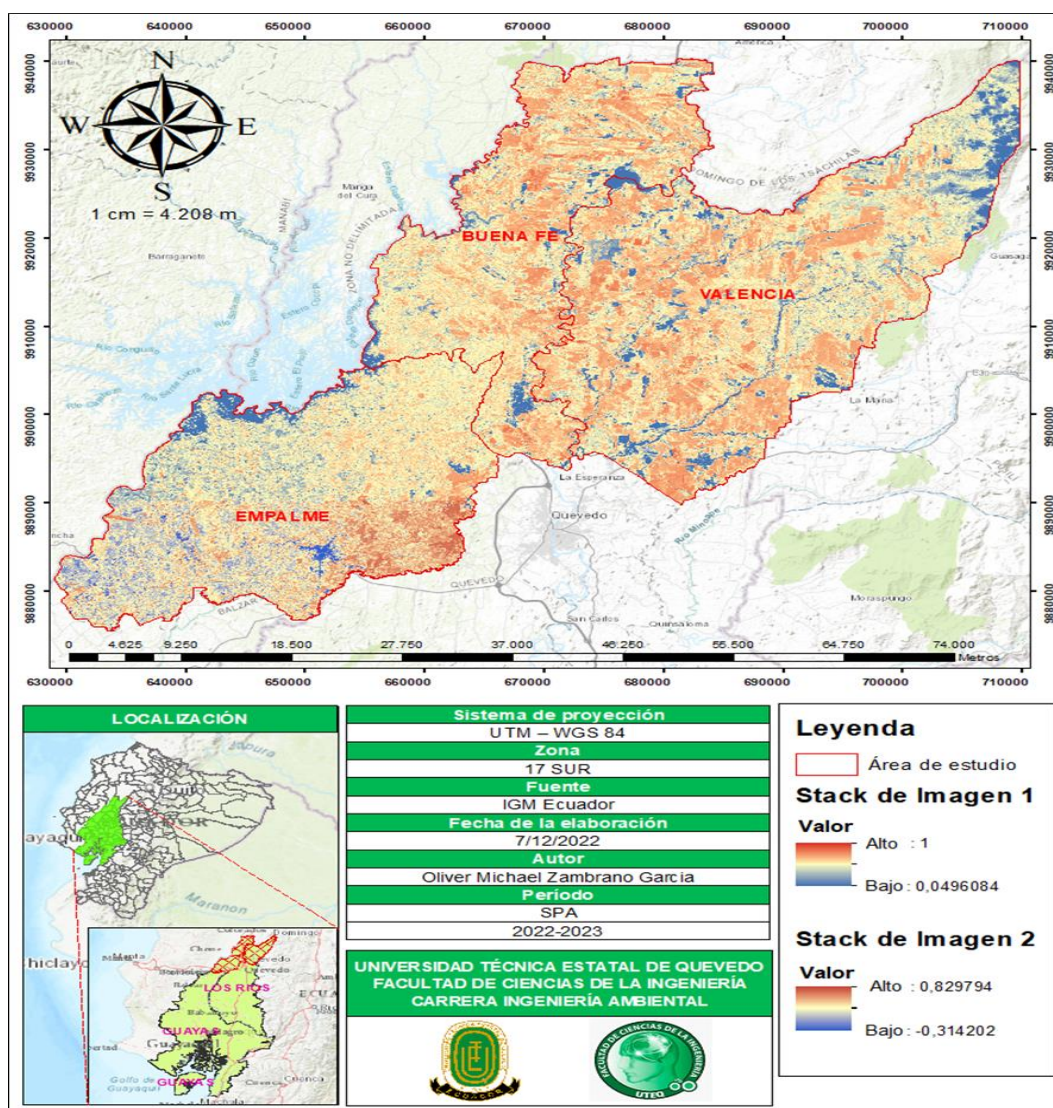
Se creó una pila de los índices obtenidos para la imagen Sentinel-2, y a partir de ella se designaron las áreas de entrenamiento para los clasificadores de Máxima Verosimilitud y Bosques Aleatorios. La combinación de esta pila se convirtió en un archivo ráster para la imagen, que contenía tres bandas para el índice NDVI y cuatro bandas para el índice RESI (que habían sido calculados previamente).

A pesar de que cada índice espectral utiliza diferentes bandas, la visualización gráfica de la pila combina estas bandas (Figura 14).

La figura 15 muestra la combinación de bandas de índices en cada Stack para la imagen 1 y 2. La banda roja (R) se asignó para representar el índice RESI, mientras que para las bandas verde (G) y azul (B) se asignaron para representar el índice NDVI. Se utilizaron dos índices para crear esta combinación de bandas. Para su visualización, se aplicó una escala de contenido máximo y mínimo, presentando un rango de valores que varía desde 0,0496084 hasta 1 para la parte superior (Stack imagen 1). A diferencia de la parte inferior (Stack imagen 2) se presentaron rango de valores de 0,314202 a 0,829794.

Figura 15

Stack de Índices NDVI + RESI calculados



Nota: Autor (2023)

3.2.3. Clasificación supervisada

3.2.3.1. Creación de áreas de entrenamiento

Para los clasificadores Máxima Verosimilitud y Bosques Aleatorios, se identificaron áreas de entrenamiento en el área de estudio y se clasificaron en diferentes coberturas de suelo. Estas incluyen zonas de cultivos de cacao, banano, palma africana, otras coberturas, asentamientos humanos y sin vegetación aparente. Las ubicaciones de estas diferentes coberturas se presentan en la tabla 9.

Tabla 9

Áreas de entrenamiento

Categorías		Coordenada UTM-WGS84		Origen	
Detalle	Codificación	X	Y	Google Earth	In-situ
Zonas de cultivos de cacao	ZC_C	654288,60	9886947,70		X
		654261,40	9886925,30		X
		649439,90	9884861,10		X
		649352,70	9884830,80	X	
		666896,50	9899372,70	X	
Zonas de cultivos de banano	ZC_B	654471,30	9886788,90	X	
		654761,80	9886625,10	X	
		655071,50	9885786,50	X	
		655689,10	9886453,90	X	
		649155,40	9885350,50	X	
		667743,20	9897717,50	X	
		667613,60	9898317,10	X	
		685687,00	9895342,50	X	
		685792,90	9895711,10	X	
Zonas de cultivos de palma africana	ZC_PA	668750,60	9894734,50	X	
		653700,80	9882438,00	X	
		668367,80	9900122,50	X	
		666891,80	9900680,50	X	
		666722,60	9900640,40	X	
		667821,10	9903646,10	X	
		668348,20	9904372,30	X	
		667950,00	9904457,20	X	
		668109,50	9897464,30	X	

		667673,50	9898069,90	X
		667522,60	9898093,00	X
		668594,90	9896024,50	X
		681850,60	9894530,40	X
		681820,40	9894703,60	X
		679307,80	9894241,00	X
		683093,40	9893744,70	X
		649712,60	9884342,00	X
		649735,40	9884930,10	X
		667883,40	9897283,20	X
		668164,90	9896996,50	X
Zonas de otras coberturas	ZC_O	667883,40	9897283,20	X
		668464,30	9895783,80	X
		668746,80	9895616,00	X
		681090,80	9896240,40	X
		684842,50	9895095,70	X
		684257,20	9895122,70	X
		651390,30	9884290,00	X
		651215,90	9884928,20	X
		652106,30	9884515,10	X
		651149,20	9882282,80	X
		649613,60	9884856,10	X
		668026,00	9901161,30	X
Zonas de asentamientos humanos	ZA_H	666960,40	9900704,40	X
		668343,70	9900937,30	X
		668061,70	9902322,30	X
		667628,60	9900149,80	X
		683449,40	9894523,80	X
		682883,10	9894294,30	X
		683560,70	9894018,50	X
		683310,00	9895402,90	X
		685234,10	9895341,80	X
		649531,20	9883826,70	X
Zonas sin vegetación aparente	ZS_VA	650082,70	9883871,90	X
		666800,20	9900515,90	X
		668078,50	9895903,40	X
		683971,00	9896550,20	X

Nota: Autor (2023)










Se definieron coordenadas de 60 puntos de apoyo para definir las áreas de entrenamiento, las cuales se utilizaron para crear las regiones de interés (Tabla 9). Se basó principalmente en Google Earth para tomar las coordenadas, identificando zonas extensas a partir de imágenes reales captadas en el terreno y que cumplieran con las coberturas necesarias según las categorías establecidas. Además, se tomaron 3 coordenadas in-situ, las cuales fueron para la categoría "Zonas con cultivos de cacao".

3.2.3.2. Creación de ROIs

Una vez que se definieron las áreas de entrenamiento, se crearon las regiones de interés basadas en las categorías de la tabla 9, tanto para el stack de índices de la imagen 1 e imagen 2 (representado en la Figura 14). La Tabla 10 muestra la cantidad de píxeles incluidos en cada categoría y el número total de píxeles, ya que ambas imágenes provienen del mismo satélite (Sentinel-2).

Tabla 10

ROIs de Stack de Índices

Categoría		Número de píxeles		
Detalle	Color	Imagen 1	Imagen 2	Total
Zona de cultivos de cacao		16	81	97
Zona de cultivos de banano		36	59	95
Zona de cultivos de palma africana		26	30	56
Zona de otras coberturas		85	33	118
Zonas sin vegetación aparente		388	61	449
Zona de asentamientos humanos		118	103	221
Cuerpos de agua		446	203	649
Nubes		201	88	289
Sin categorización		No considerado		

Nota: Autor (2023)

De acuerdo con la información presentada en la tabla 10, las asignaciones de píxeles variaron entre las imágenes 1 y 2. Se reflejaron asignaciones mayores en las ROIs de "Zonas sin vegetación aparente", "Zonas de asentamiento humano", "Cuerpos de agua" y "Cobertura de nubes". Debido a la presencia de nubosidad en ambas imágenes, se presentaron categorías donde se reflejó una mayor cantidad de píxeles en la imagen 1 debido al contenido de cobertura de










nubes en la zona este de la imagen y en diversos espacios distribuidos por encima de otras áreas de la cobertura vegetal. La diferencia en las asignaciones de píxeles entre la imagen 2 y la imagen 1 se debe a que las zonas de nubes son más amplias en la imagen 2.

3.2.3.3. Clasificación con Máxima Verosimilitud

En la tabla 11 se muestra que el clasificador de Máxima Verosimilitud para la imagen Sentinel-2 identificó una mayor área "sin categorización", que totalizó 3153,8997 km². Al contrario, la cobertura con menor extensión fue la de "Zonas de asentamientos humanos", con 24,6776 km². Para las principales áreas de zonas de cultivos de cacao, banano y palma africana, se identificaron 67,2412 km², 73,3114 km² y 123,3667 km², respectivamente. En total, se encontró una superficie de 4189,3570 km², lo que equivale a la identificación de 41893570 píxeles con una resolución espacial de 10 m x 10 m (100 m² por píxel).

Tabla 11

Extensión de coberturas clasificadas con método Máxima Verosimilitud (ML)

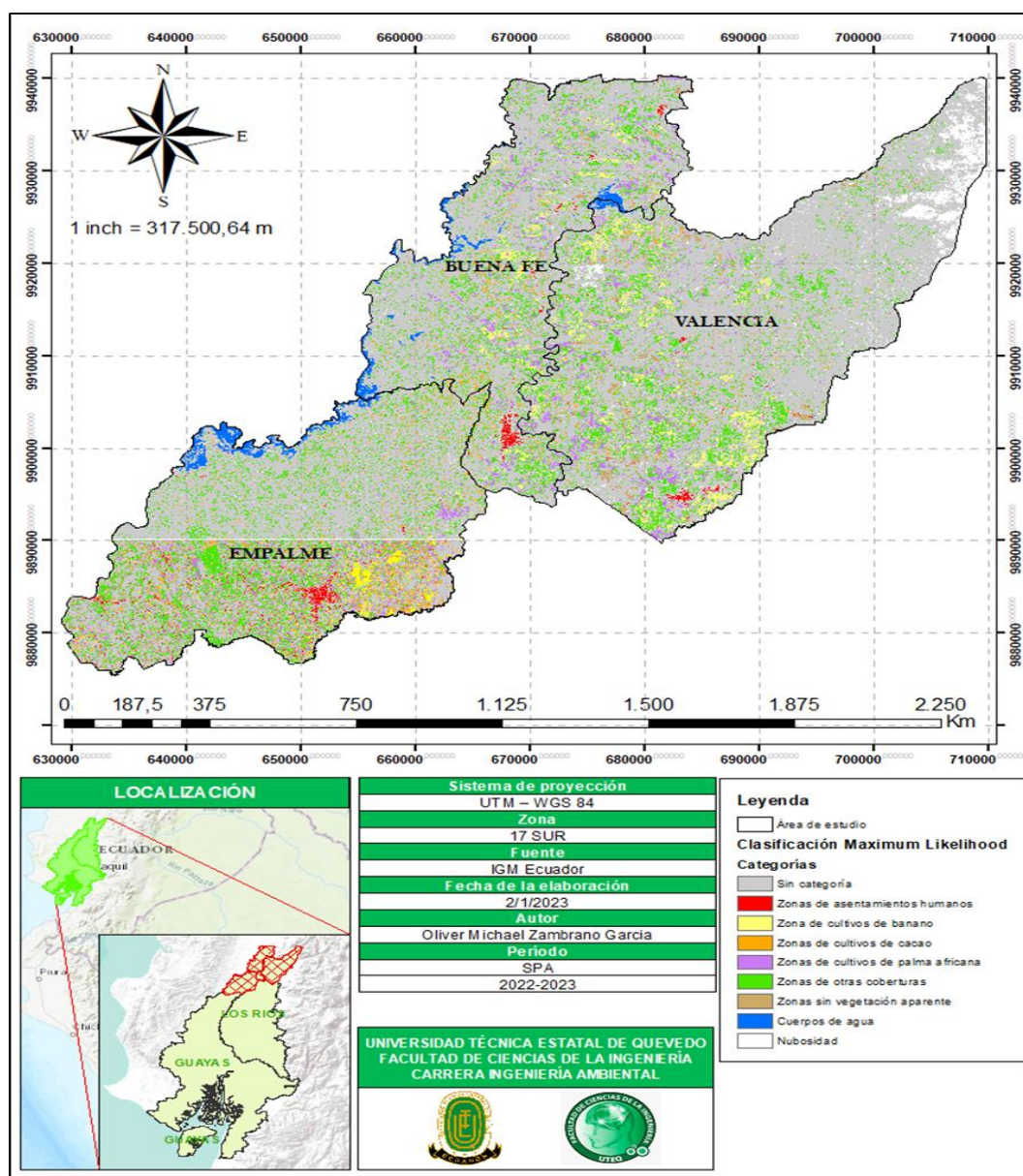
Categoría		Conteo pixeles	Área	
Detalle	Color		m ²	km ²
Zona de cultivos de cacao		672412	67241200	67,2412
Zona de cultivos de banano		733114	73311400	73,3114
Zona de cultivos de palma africana		1233667	123366700	123,3667
Zona de otras coberturas		4836076	483607600	483,6076
Zonas sin vegetación aparente		547735	54773500	54,7735
Zona de asentamientos humanos		246776	24677600	24,6776
Cuerpos de agua		943841	94384100	94,3841
Nubes		1140952	114095200	114,0952
Sin categorización		31538997	3153899700	3153,8997
Total	→	41893570	4189357000	4189,3570

Nota: Autor (2023)

La figura 16 presenta los resultados obtenidos con el clasificador Máxima Verosimilitud. Se puede ver que el método tiene tendencia a cometer errores en la categorización, ya que una gran parte del área quedó sin categorizar. Precisamente, la clasificación de los cultivos de interés mostró una falta de precisión, debido a la escasa distribución de los cultivos, así como la interpretación limitada del algoritmo convencional en la zona de estudio.

Figura 16

Clasificación supervisada con método Máxima Verosimilitud

**Nota:** Autor (2023)










3.2.3.4. Clasificación con Bosques Aleatorios

En la tabla 12 se muestra que el clasificador de Bosques Aleatorios para la imagen Sentinel-2 identificó una mayor área en las "Zonas de otras coberturas", que totalizó 1600,2380 km². Al contrario, la cobertura con menor extensión fue la de "Zonas de asentamientos humanos", con 18,6295 km². Para las principales áreas de zonas de cultivos de cacao, banano y palma africana, se identificaron 148,7444 km², 8,4236 km² y 93,5421 km², respectivamente. En total, se encontró

una superficie de 2255,7797 km², lo que equivale a la identificación de 22557797 píxeles con una resolución espacial de 10 m x 10 m (100 m² por píxel).

Tabla 12

Extensión de coberturas clasificadas con método Bosques Aleatorios (RF)

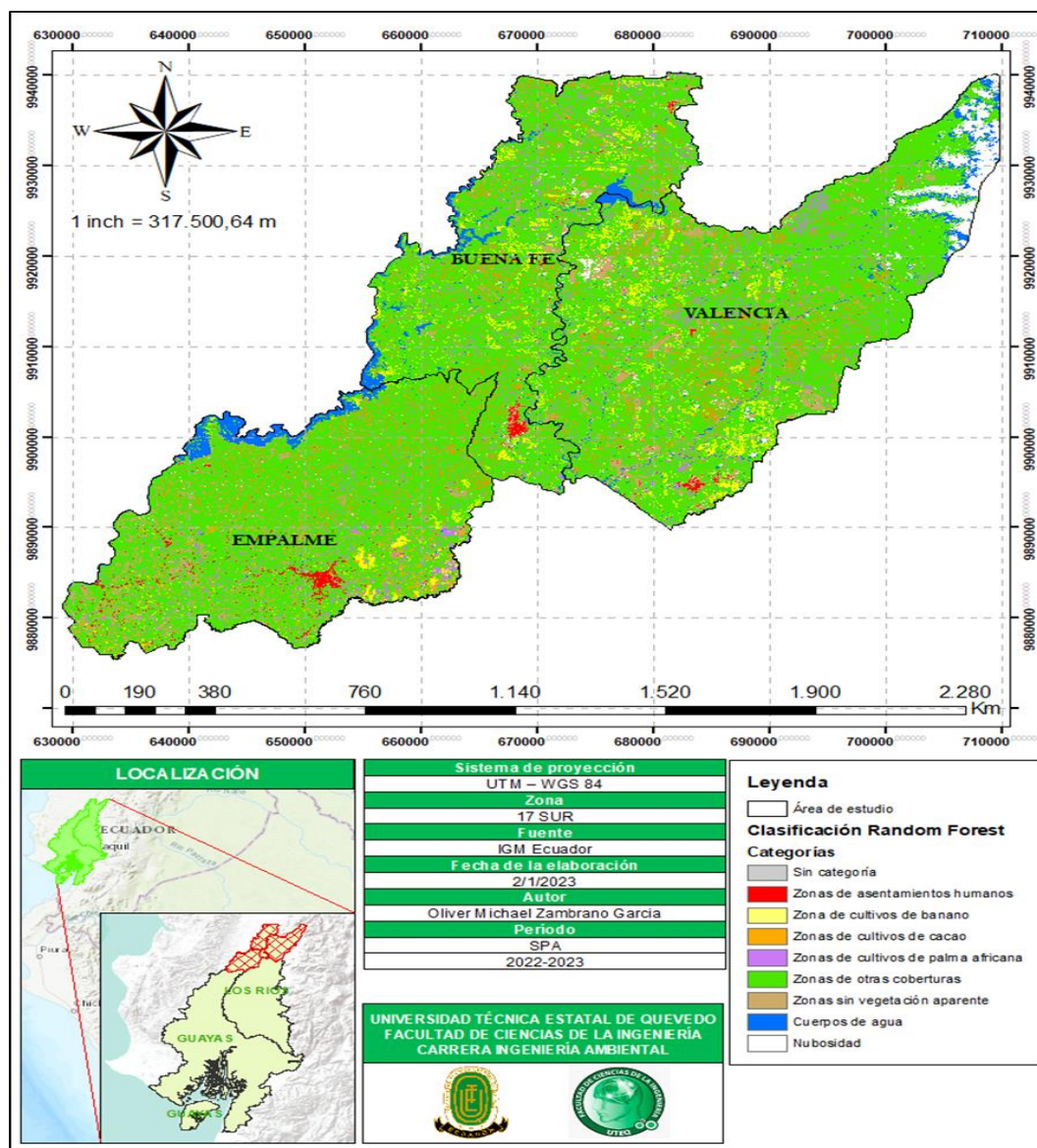
Categoría		Número de píxeles	Área	
Detalle	Color		m ²	km ²
Zona de cultivos de cacao		1487444	148744400	148,7444
Zona de cultivos de banano		842360	84236000	8,4236
Zona de cultivos de palma africana		935421	93542100	93,5421
Zona de otras coberturas		16002380	1600238000	1600,2380
Zonas sin vegetación aparente		1656619	165661900	165,6619
Zona de asentamientos humanos		186295	18629500	18,6295
Cuerpos de agua		868224	86822400	86,8224
Nubes		579054	57905400	57,9054
Sin categorización		0	0	0
Total	→	22557797	2255779700	2255,7797

Nota: Autor (2023)

La figura 17 ilustra las coberturas categorizadas por el clasificador Bosques Aleatorios. Este método demostró ser más efectivo en la categorización de las coberturas definidas debido a su capacidad de utilizar cientos de combinaciones de árboles para alcanzar una clasificación precisa. Para el área de estudio, se identificó una mayor categorización de coberturas de cultivos de interés, visualizando en un promedio de 7%, 4% y 4% respectivamente para las zonas de cultivos de cacao, banano y palma africana.

Figura 17

Clasificación supervisada con método Bosques Aleatorios



Nota: Autor (2023)

3.3. Análisis estadístico de los resultados

3.3.1. Validación de resultados

Para la validación de los resultados obtenidos, se generaron áreas de entrenamiento a partir de coordenadas elegidas al azar (Tabla 13). Con estas áreas se crearon nuevos ROIs y se elaboraron las matrices de confusión (Figura 18) y se calculó el coeficiente Kappa (Tabla 14) basándose en las clasificaciones realizadas previamente (secciones 4.2.3.3 y 4.2.3.4).

3.3.1.1. Áreas de entrenamiento al azar

La tabla 13 presenta las áreas de entrenamiento que fueron generadas al azar. Para su creación se recurrió principalmente a la utilización de Google Earth, pero para las categorías "Zonas con cultivos de cacao", "Zonas de otras coberturas" y "Zonas de asentamientos humanos", se obtuvieron coordenadas de manera in-situ.

Tabla 13

Áreas de entrenamiento al azar

Categorías		Coordenada UTM-WGS84		Origen	
Detalle	Codificación	X	Y	Google Earth	In-situ
Zonas de cultivos de cacao	ZC_C	649423,50	9884854,70		X
		667864,20	9897904,80	X	
		668009,00	9897591,50	X	
		666903,00	9899360,10	X	
Zonas de cultivos de banano	ZC_B	653668,50	9882828,00	X	
		653438,60	9882878,20	X	
		667839,40	9897884,30	X	
		685733,50	9895504,10	X	
		680933,80	9894466,60	X	
Zonas de cultivos de palma africana	ZC_PA	651138,80	9880406,90	X	
		667654,70	9898259,70	X	
		666856,10	9898855,80	X	
		681829,90	9894606,30	X	
Zonas de otras coberturas	ZC_O	653439,50	9882802,30		X
		652299,60	9886163,00		X
		652034,80	9882845,90	X	
		668311,40	9896941,20	X	
		668209,70	9904286,40	X	
		669176,10	9903522,10	X	
		684693,80	9895148,60	X	
		685408,40	9895291,40	X	
Zonas de asentamientos humanos	ZA_H	651551,10	9885131,10		X
		651586,50	9884772,10		X
		650474,50	9884870,80		X
		667870,90	9900737,90	X	
		668348,00	9901531,90	X	

		683315,50	9894656,90	X
		683617,10	9894935,90	X
		685172,00	9895450,00	X
		653775,50	9882828,60	X
Zonas sin		651920,50	9882971,90	X
vegetación	ZS_VA	649532,50	9883803,40	X
aparente		668590,40	9899306,90	X
		681598,70	9893504,30	X

Nota: Autor (2023)


En la toma de coordenadas se utilizó la aplicación móvil “Russian Topo Maps”, donde se determinó varias zonas tales como ZC_C, ZC_O, ZS_VA y ZA_H, las cuales se consideraron más accesibles para dicha toma in-situ. Consecuentemente, para las zonas ZC_B y ZC_PA se utilizó el software Google Earth, a través de la interfaz de navegación satelital que brinda este.

3.3.1.2. Matriz de confusión

Después de establecer los ROIs basados en la tabla 13, se produjeron las matrices de confusión. La figura 18 muestra ambas matrices, con los porcentajes de precisión general y específicos para cada categoría de clasificación.

Figura 18


Matriz de confusión de las clasificaciones de imágenes: A(ML) & B(RF)

A Máxima Verosimilitud (ML) - Imagen 1-2 								
	ZC_C V	ZC_B V	ZC_PA V	ZC_O V	ZC_H V	ZS_VA V	S_C V	Total V
ZC_C >	545	0	0	0	0	0	0	545
ZC_B >	0	564	0	0	0	0	0	564
ZC_PA >	0	0	994	0	0	0	0	994
ZC_O >	0	0	0	3016	0	0	0	3016
ZA_H >	0	0	0	0	184	0	0	184
ZS_VA >	0	0	0	0	0	460	0	460
S_C >	0	0	0	0	0	0	45837	45837
Total >	545	564	994	3016	184	460	45837	
Acuerdo/Exactitud [%]	100,00	100,00	100,00	100,00	100,00	100,00	100,00	
Precisión general	52,96%							

Porcentaje
de Precisión

0-25

25-50

B Bosques Aleatorios (RF) - Imagen 1-2 								
	ZC_O V	ZC_PA V	S_C V	ZC_C V	ZS_VA V	ZC_B V	ZA_H V	Total V
ZC_O >	260	6	0	3	2	2	0	273
ZC_PA >	1	10	0	0	0	0	0	11
S_C >	2	0	12	0	0	0	0	14
ZC_C >	5	0	0	23	0	0	0	28
ZS_VA >	1	0	0	0	26	0	0	27
ZC_B >	1	0	0	0	0	13	0	14
ZA_H >	0	0	0	0	0	0	3	3
Total >	270	16	12	26	29	15	3	
Acuerdo/Exactitud [%]	96,30	62,50	100,00	88,46	92,86	86,67	100,00	
Precisión general	93,78%							

Nota: Autor (2023)

De acuerdo con la figura 18, en la clasificación de Máxima Verosimilitud (ML) para los cultivos de cacao (ZC_C), banano (ZC_B), palma africana (ZC_PA), zonas de otros tipos de cultivos (ZC_O), zonas de asentamientos humanos (ZA_H) y zonas sin vegetación visible (ZS_VA) mostraron niveles de precisión-exactitud del 100%. Como resultado, estas zonas clasificadas fueron detectadas de manera eficaz debido a la asignación de puntos al alzar supuestos en las diferentes categorizaciones. De igual forma, en la categoría sin clasificación (S_C), que incluyó datos sobre nubosidad, cuerpos de agua y categorías no definidas o sin clasificación, se obtuvieron niveles de acuerdo-exactitud del 100% debido a la amplia cobertura de nubosidad y la extensa proyección sin clasificación representada en la figura 16.

Para la clasificación de Bosques Aleatorios (RF), en los cultivos de cacao (ZC_C), banano (ZC_B) y palma africana (ZC_PA) se obtuvieron niveles de precisión-exactitud del 88,46%, 86,67% y 66,50%, respectivamente. Por su parte, en las zonas con otros tipos de cultivos (ZC_O), zonas de asentamientos humanos (ZA_H) y zonas sin vegetación visible (ZS_VA) se alcanzaron niveles de precisión-exactitud del 96,30%, 100% y 92,86%, respectivamente. En la clasificación sin categorización (S_C), que incluyó datos sobre nubosidad, cuerpo de agua y categorías no definidas, la categoría no definida no presentó resultados variables en la matriz. La categoría S_C demostró niveles de precisión-exactitud del 100%, lo que refleja el excelente proceso realizado con el método RF.

Siendo así que la clasificación RF demostró ser mejor con un nivel de precisión general del 93,78%. Por otro lado, la clasificación ML tuvo una precisión general de solo 52,96%. Al evaluar la detección de los cultivos de interés, se encontraron diferencias significativas en términos de niveles de precisión-exactitud individual, con un 96,30% para el método RF y un 100% para el método ML.

3.3.1.3. Coeficiente Kappa

Asimismo, se llevaron a cabo cálculos de los coeficientes de Kappa para las cuatro clasificaciones realizadas. La tabla 14 presenta los resultados obtenidos.

Tabla 14

Fuerza de Concordancia según el coeficiente de Kappa para las clasificaciones de Máxima Verosimilitud (ML) y Bosques Aleatorios (RF)

Clasificación	Coeficiente de Kappa	Precisión
ML	0,1543	Leve
RF	0,860818	Casi perfecta

Nota: Autor (2023)

Los resultados obtenidos de las clasificaciones realizadas por ambos métodos muestran una marcada diferencia según los datos presentados en la tabla 16. El método RF logró una clasificación sobresaliente, con un coeficiente Kappa de 0,860818, el cual representa una fuerza de concordancia o precisión casi perfecta. Por otro lado, el método ML no logró un rendimiento satisfactorio, con un coeficiente Kappa de solo 0,1543, lo cual refleja una precisión leve.

3.3.2. Análisis estadístico

Después de obtener los valores de la medición, se realizó un análisis estadístico para comprobar la normalidad de los datos (Tabla 15).

Tabla 15

Prueba estadística de normalidad para los métodos de clasificación Máxima Verosimilitud (ML) y Bosques Aleatorios (RF)

Métodos de Clasificación		Kolmogórov-Smirnov			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Categorías	ML	0,424429141	395	1,8966E-206	0,637742537	395	4,71875E-28
	RF	0,421167729	395	3,2441E-203	0,6474668	395	9,6003E-28

Nota: Autor (2023)

Según lo indicado en la Tabla 15, para evaluar la normalidad de los datos se compararon variables aleatorias entre los métodos de aprendizaje automático (ML) y selección de características (RF), y en ambos casos se obtuvo un nivel de significancia inferior a 0,05. Por lo tanto, se concluye que los datos no provienen de una distribución normal y se utilizó la prueba de suma de rangos de Wilcoxon.

Para complementar los resultados, se realizó un análisis estadístico no paramétrico utilizando la prueba Wilcoxon-Mann-Whitney para muestras independientes, y los resultados se presentan en la Tabla 16.

Tabla 16

Prueba estadística de la Suma de Rangos de Wilcoxon para las comparaciones entre métodos clasificadores Máxima Verosimilitud (ML) y Bosques Aleatorios (RF)

Estadísticos de prueba	Valores por puntos al azar
U de Mann-Whitney	40083,500
W de Wilcoxon	118293,500
Z	-12,458
Sig. asintótica(bilateral)	1,26033E-35

Nota: Autor (2023)

Según la tabla 16, el valor p (Sig. asintótica (bilateral)) de la comparación entre los métodos de clasificación ML y RF es inferior a 0.05. En consecuencia, se rechaza la hipótesis nula y se puede afirmar que existen diferencias estadísticamente significativas entre los resultados obtenidos con ML y RF para la clasificación de la imagen de Sentinel-2.

Los resultados del análisis estadístico son coherentes con los reportados por las matrices de confusión y los coeficientes de Kappa. La diferencia en las clasificaciones ML y RF fue del 40,82% en las matrices de confusión y 0.71 en los coeficientes de Kappa, lo que demuestra que existen diferencias significativas entre los resultados obtenidos con estos métodos.

3.4. Discusión

Durante las últimas décadas de procesos de investigaciones para la clasificación de imágenes multiespectrales, el método de Máxima Verosimilitud ha sido ampliamente utilizado. Sin embargo, han surgido nuevos enfoques basados en inteligencia artificial, como Random Forest (Athey et al., 2019), (Genuer et al., 2017), aun así, estos nuevos métodos de clasificación supervisada requieren estudios que comprueben los niveles de precisión. Estos tipos de métodos que apoyan a los procesos de teledetección favorecen a la actualización de base de datos nacionales en carácter de distribución de uso de suelo más significativos como pueden ser en áreas de cultivos agrícolas (Borrás et al., 2017), siendo así un requisito indispensable para mejorar la toma de decisiones en procesos de sostenibilidad y ordenamiento del territorio. Particularmente, las provincias Guayas y Los Ríos dentro de Ecuador tienen una producción agrícola sorprendente, siendo uno de sus principales sectores con beneficios socioeconómicos (MAG, 2020). A su vez el manejo de los métodos encaminado al área de estudio y sus adyacentes, valdrán para identificar posibles inestabilidades al uso del suelo enfocado en actividades agrícolas.

Los índices espectrales de vegetación son comúnmente empleados en procesos de detección cultivos, debido a la inmensa posibilidad de obtener categorizaciones precisas (Pech May et al., 2021). En diversos estudios se han analizado algunos índices de vegetación, como NDVI y una variación del índice de bordes rojos (RESI), entre otros (Xiao et al., 2020), (Campos-Taberner et al., 2020). Precisamente, el índice NDVI es el más común y ampliamente utilizado para estimar la variación de los cultivos en una determinada zona durante periodos de tiempo. En la mayoría de los valores que NDVI puede mostrar, datos resultantes como 0,2 y 0,4 conciernen a una vegetación escasa, mientras que una vegetación moderada tiende a variar entre 0,4 y 0,6 (Pech May et al., 2021); cualquier valor por encima de 0,6 indicaría una mayor densidad de vegetación en la zona (Pech May et al., 2021). Por otro lado, el índice de RESI basado en Sentinel muestra una ventaja obvia en la detección de cobertura vegetal de la transición defoliación-foliación (Xiao et al., 2020). Aludiendo unos ejemplos, se emplearon en una investigación para evaluar variabilidad espacio-temporal en cuatro manglares periurbanos (Denis Ávila et al., 2020), o también en cambios

en la vegetación forestal durante 4 semestres en 2 parques nacionales (Lastovicka et al., 2020), entre otros diversos estudios de investigación realizados alrededor del planeta.

Para el estudio, se aprovecharon los índices NDVI y RESI para crear una imagen multibanda (Stack). Los resultados mostraron una diferencia significativa en ambos índices, con RESI mostrando una mejor capacidad para identificar las características de clorofila en las coberturas vegetales a comparación con NDVI. Aunque ambos índices asumen un mismo objetivo, RESI utilizó con mayor eficacia la detección de los niveles infrarrojos de la cobertura de la vegetación y los cultivos de interés. Semejante a lo mencionado, (Xiao et al., 2020) en una investigación desarrollaron un método de índices espectrales de bordes rojos (RESI) mediante la normalización de tres bandas de bordes rojos y lo aplicaron para identificar y cartografiar plantaciones de caucho reflejando que tienen un potencial significativo para el seguimiento de las plantaciones de caucho, sino que también mejora los métodos de teledetección de la cartografía de otros tipos coberturas vegetales mediante la introducción de la red-EDGE.

Tanto el método Máxima Verosimilitud (ML) como el método Bosques Aleatorios (RF) son utilizados para detectar áreas de cultivos agrícolas con imágenes de Sentinel-2 (Athey et al., 2019), (Denis Ávila et al., 2020), (Borrás et al., 2017). El estudio demostró que el mejor rendimiento en la clasificación de banano, cacao y palma africana se logró con el uso de RF. Por otro lado, ML tuvo una precisión más baja, como se puede ver en los porcentajes de precisión y el índice Kappa (Athey et al., 2019). La razón principal por la que ML tuvo una baja precisión en la detección de estos cultivos fue debido a la carencia de información incorporada para la clasificación de la imagen de Sentinel-2 (Valero Medina & Alzate Atehortúa, 2019), afectando la precisión en la detección de cultivos.

Al revisar las estadísticas que verifican las categorizaciones, el rendimiento sobresaliente del algoritmo Bosques Aleatorios en la cartografía de coberturas agrícolas destaca en primer lugar, lo que coincide con los hallazgos obtenidos en investigaciones previas que utilizaron la misma técnica de clasificación. En este tipo de investigaciones, las amplias dimensiones de los terrenos agrícolas hacen posible obtener una amplia cantidad de muestras representativas de cada

categoría, necesarias para una correcta formación del algoritmo. Recientemente, en otras partes del mundo, (Ramírez et al., 2020) utilizaron la muestra de la cobertura terrestre como referencia para llevar a cabo la clasificación, la cual fue capturada por un experto en el tema de manera manual. Donde obtuvieron una distribución del 70% para el entrenamiento del algoritmo Bosques Aleatorios y un 30% para validar la clasificación. En otro caso, (Morell-Monzó et al., 2022) analizaron el uso de datos multitemporales Sentinel-2 para clasificar el estado de pequeñas parcelas de cítricos demostrando el potencial de los datos Sentinel-2 para identificar el estado de estas parcelas, donde lograron obtener una precisión global del 90,30% mediante la clasificación supervisada basada en Bosques Aleatorios.

A pesar de que la técnica de Máxima Verosimilitud (ML) y Bosques Aleatorios (RF) han mostrado resultados novedosos en la clasificación de imágenes Sentinel-2 con otros índices de vegetación y preprocesamientos, para el caso de estudio la precisión obtenida en la clasificación supervisada con el método RF fue casi perfecta, con una precisión del 93,78% y coeficiente Kappa de 0,860818. Este impresionante rendimiento puede tener sus raíces en la investigación de campo previa, en la cual se recopilaban datos sobre los diferentes cultivos de interés, permitiendo la elaboración de categorías para los grupos de entrenamiento y validación. Sin embargo, con una mayor cantidad y calidad de datos de entrenamiento, se esperaría un aumento en la precisión, llegando a un rango entre el 94% y 98% (Borrás et al., 2017).

En algunos casos el método de Bosques Aleatorios no es propiamente preciso para la determinación de clasificación y validación. En este sentido se resalta el uso de Máxima Verosimilitud, que distintivamente presentó resultados de precisión leves o intermedios (52,96%) para el estudio. En un caso similar, (Valero Medina & Alzate Atehortúa, 2019) compararon 3 tipos de técnicas de algoritmos para una imagen satelital en la identificación de varios campos de cultivo de algodón, obteniendo una clasificación con la máxima probabilidad del 68,95%, SVM del 81,325% y de bosques aleatorios del 78,925%. Esto demostró que, al realizar la prueba de confianza para las clasificaciones con intervalos no solapados, se obtuvieron valores más altos para el SVM. Sin embargo, en otro caso de estudio (Nolasco et al., 2017) compararon el rendimiento de los

algoritmos de clasificación Máxima Verosimilitud, Support Vector Machine y Bosques Aleatorios, y analizaron la influencia de la naturaleza y cantidad de categorías en los modelos. Al final, lograron determinar que el algoritmo de Máxima Verosimilitud fue el más conveniente, alcanzando una precisión del 99% y un índice k de 0,99.

Las clasificaciones de esta investigación tuvieron una precisión general aproximada de 53% (ML) y 94% (RF). Sin embargo, la precisión individual de detección para dos de los tres objetivos principales, cacao y banano, fue excelente con valores por encima del 88% (para el método RF). En cambio, la detección de la palma africana no fue precisa, ya que su separabilidad espectral con otras áreas resultó ser baja y obtuvo valores de acuerdo/exactitud por debajo del 63% (Weiss et al., 2020). Esto podría deberse a que la palma africana no se cultiva en grandes extensiones o la clasificación de este no se diferencia con la de banano u otros cultivos parecidos en la zona de estudio, lo que limitó la cantidad de muestras de campo para entrenar a los clasificadores.

La posible mejora en la precisión que presenta Bosques Aleatorios en comparación con Máxima Verosimilitud es un aspecto importante por considerar, especialmente en la clasificación de imágenes de Sentinel-2, donde se ha demostrado una diferencia estadísticamente significativa. Por tanto, se encomienda que en futuras investigaciones se tomen en cuenta más y mejores coordenadas in situ en zonas con cultivos de palma africana, ya que se cree que la baja precisión en la detección de estos cultivos con Sentinel-2 se debe a falta de incorporación de muestras para las clasificaciones. De manera que permita mejorar las condiciones para los clasificadores y evaluar verdaderamente la capacidad de mejora de Bosques Aleatorios en comparación con Máxima Verosimilitud en la clasificación de imágenes de Sentinel-2.

3.5. Conclusiones

El preprocesamiento de imágenes multiespectrales del satélite Sentinel-2 es un paso crítico para la toma de decisiones. El uso de software especializado como ENVI y ArcGIS permitió ejecutar este proceso de manera eficiente y precisa, preparando las imágenes para su posterior clasificación y análisis. Es importante

destacar que la calidad y precisión del preprocesamiento sosegaron directamente con los resultados finales.

Los valores obtenidos en los índices calculados mostraron diferencias notables debido a la aplicabilidad específica de cada uno, donde la cantidad de vegetación, actividad fotosintética y clorofila jugaron un papel fundamental para las posteriores clasificaciones. Tomando en cuenta que existieron diferencias en los datos y visualización de los índices, se puede afirmar que en conjunto NDVI y RESI ayudaron a la clasificación más eficiente de las zonas de cultivos de interés.

Los métodos de Máxima Verosimilitud (ML) y Bosques Aleatorios (RF), junto con los índices espectrales, han demostrado ser efectivos en la clasificación supervisada de imágenes Sentinel-2 para detectar los cultivos principales en la zona de estudio. Estos métodos combinaron la información espectral y la estructura de la imagen para mejorar la precisión de la clasificación, para el caso de RF. Sin embargo, es importante tener en cuenta que la selección adecuada de índices espectrales y la correcta validación de los modelos fueron cruciales para lograr resultados precisos.

Los resultados de las clasificaciones de Sentinel-2 con ML y RF fueron distintos entre sí. En las clasificaciones con ML, la categoría que ocupó la mayor extensión espacial fue "Zonas sin clasificación", mientras que la categoría con la menor extensión fue "Zonas de asentamientos humanos". Esto resultó en un aumento en la extensión espacial de la categoría "Zonas de asentamientos humanos" y en la clasificación sin categoría. Por otro lado, en la clasificación con RF, se redujo la extensión de la categoría "Zonas de cultivos de palma africana" y se aumentaron las coberturas de banano, cacao y zonas de asentamientos humanos, así como también se disminuyó la extensión de las coberturas de nubes, cuerpos de agua y las sin clasificación.

No fue posible detectar de forma precisa las zonas con cultivos de palma africana en imágenes de Sentinel-2 mediante el uso de ML y RF. Sin embargo, se detectaron con alta precisión las zonas con cultivos de cacao (con una precisión individual de más del 88.46% mediante RF) y banano (con una precisión individual de más del 86.67%) mediante RF. En general, para clasificar todas las

coberturas de cultivos de interés, la mejor opción es utilizar Bosques Aleatorios, mientras que la combinación Máxima Verosimilitud no es la más recomendada para la zona de estudio o localizaciones aledañas.

Los resultados obtenidos en las matrices de confusión y los coeficientes de Kappa fueron respaldados por un análisis estadístico riguroso. Las diferencias notables en las precisiones de clasificación entre los métodos ML y RF fueron confirmadas por la evaluación estadística, que mostraron diferencias significativas. Además, el análisis con la prueba de Wilcoxon-Mann-Whitney indicó que las variaciones en la precisión de clasificación según las matrices de confusión y los coeficientes de Kappa son estadísticamente significativas.

3.6. Recomendaciones

Se debe continuar investigando y mejorando el preprocesamiento de imágenes multiespectrales para la detección de cultivos. Sería útil explorar nuevos métodos y software para mejorar la precisión y eficiencia del preprocesamiento y la clasificación de imágenes satelitales, además de facilitar el acceso remoto y directo a los metadatos e imágenes satelitales. Igualmente, sería recomendable considerar la incorporación de otros datos relevantes para mejorar la detección de cultivos, como datos climáticos y de suelo. Esto permitirá obtener una visión más completa y precisa de la región y ayudará a tomar decisiones más informadas sobre la gestión y conservación de los recursos naturales.

Para mejorar la identificación de coberturas difíciles en Google Earth, es importante llevar a cabo una recopilación de muestras en el terreno con el objetivo de mejorar las áreas de interés y aumentar la separabilidad espectral para lograr clasificaciones más precisas, especialmente en áreas relativamente pequeñas y que a menudo están mezcladas.

Se sugiere utilizar métodos de Bosques Aleatorios en futuras investigaciones dentro o fuera de la zona de estudio para mejorar continuamente la toma de decisiones. Es aconsejable experimentar con otros métodos, como las Máquinas de Vectores de Soporte (SVMs), para identificar áreas más extensas y probar diferentes funciones de núcleo en la función de decisión con el fin de evaluar su grado de precisión.



REFERENCIAS BIBLIOGRÁFICAS

Referencias Bibliográficas

- Ai, B., Huang, K., Zhao, J., Sun, S., Jian, Z., & Liu, X. (2022). Comparison of classification algorithms for detecting typical coastal reclamation in Guangdong Province with Landsat 8 and Sentinel 2 images. *Remote Sensing*, 14(2), 385. <https://doi.org/10.3390/rs14020385>
- Aksoy, S., Yildirim, A., Gorji, T., Hamzehpour, N., Tanik, A., & Sertel, E. (2022). Assessing the performance of machine learning algorithms for soil salinity mapping in Google Earth Engine platform using Sentinel-2A and Landsat-8 OLI data. *Advances in Space Research: The Official Journal of the Committee on Space Research (COSPAR)*, 69(2), 1072–1086. <https://doi.org/10.1016/j.asr.2021.10.024>
- Arai, K. (2020). Maximum likelihood classification based on classified result of boundary mixed pixels for high spatial resolution of satellite images. *International Journal of Advanced Computer Science and Applications* : IJACSA, 11(9). <https://doi.org/10.14569/ijacsa.2020.0110904>
- Arcgis Desktop. (s/f). *¿Qué es la clasificación de imagen?* Arcgis Desktop. <https://desktop.arcgis.com/es/arcmap/10.4/extensions/spatial-analyst/image-classification/what-is-image-classification-.htm>
- Athey, S., Tibshirani, J., & Wager, S. (2019). Generalized random forests. *Annals of statistics*, 47(2). <https://doi.org/10.1214/18-aos1709>
- Borràs, J., Delegido, J., Pezzola, A., Pereira, M., Morassi, G., & Camps-Valls, G. (2017). Clasificación de usos del suelo a partir de imágenes Sentinel-2. *Revista de teledetección*, 48, 55. <https://doi.org/10.4995/raet.2017.7133>
- Campos-Taberner, M., García-Haro, F. J., Martínez, B., & Gilabert, M. A. (2020). Deep learning para la clasificación de usos de suelo agrícola con Sentinel-2. *Revista de teledetección*, 56, 35. <https://doi.org/10.4995/raet.2020.13337>

- Cánovas García, F., Alonso Sarría, F., & Gomariz Castillo, F. (2016). Modificación del algoritmo Random Forest para su empleo en clasificación de imágenes de teledetección. En *Aplicaciones geotecnológicas para el desarrollo económico sostenible. XVII Congreso Nacional de Tecnologías de la Información Geográfica* (pp. 359–368). Grupo de Tecnologías de la Información Geográfica.
- Castellanos-Navarrete, A. (2021). Oil palm dispersal into protected wetlands: Human–environment dichotomies and the limits to governance in southern Mexico. *Land Use Policy*, 103, 1–13. <https://doi.org/10.1016/j.landusepol.2021.105304>
- Cerda L, J., & Villarroel del P, L. (2008). Evaluación de la concordancia inter-observador en investigación pediátrica: Coeficiente de Kappa. *Revista Chilena de Pediatría*, 79(1). <https://doi.org/10.4067/s0370-41062008000100008>
- Coltro, L., & Karaski, T. U. (2019). Environmental indicators of banana production in Brazil: Cavendish and Prata varieties. *Journal of Cleaner Production*, 207, 363–378. <https://doi.org/10.1016/j.jclepro.2018.09.258>
- Congedo, L. (2021). Semi-Automatic Classification Plugin: A Python tool for the download and processing of remote sensing images in QGIS. *Journal of open source software*, 6(64), 3172. <https://doi.org/10.21105/joss.03172>
- Copernicus Open Access Hub. (s/f). *Europe's Eyes on Earth*. Copernicus Open Access Hub. <https://scihub.copernicus.eu/dhus/#/home>
- Cutler, D. R., Edwards, T. C., Jr, Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., & Lawler, J. J. (2007). Random forests for classification in ecology. *Ecology*, 88(11), 2783–2792. <https://doi.org/10.1890/07-0539.1>
- Denis Ávila, D., Curbelo, E. A., Madrigal-Roca, L. J., & Pérez-Lanyau, R. D. (2020). Variación espacio-temporal de la respuesta espectral en manglares de La Habana, Cuba, a través de sensores remotos. *Revista de biología tropical*, 68(1). <https://doi.org/10.15517/rbt.v68i1.39134>
- Elbeih, S. F. (2021). Evaluation of agricultural expansion areas in the Egyptian deserts: A review using remote sensing and GIS. *Egyptian Journal of*

- Remote Sensing and Space Sciences*, 24(3), 889–906.
<https://doi.org/10.1016/j.ejrs.2021.10.004>
- EOS Data Analytics. (2020). *¿Qué es la Diferencia Normalizada de Borde Rojo (NDRE)?* EOS Data Analytics.
<https://eos.com/es/industries/agriculture/ndre/>
- European Space Agency, & Fletcher, K. (2012). *SENTINEL 2: ESA's optical high-resolution mission for GMES operational services*.
- GADM de Valencia. (2016). *Plan de Desarrollo y Ordenamiento Territorial 2016-2020 del Cantón Valencia*.
<https://multimedia.planificacion.gob.ec/PDOT/descargas.html>
- GADM El Empalme. (2015). *Plan de Desarrollo y Ordenamiento Territorial - El Empalme 2015-2024*.
<https://multimedia.planificacion.gob.ec/PDOT/descargas.html>
- GADM San Jacinto de Buena Fe. (2017). *Plan de Desarrollo y Ordenamiento Territorial 2015-2019; Alineado al Plan Nacional de Desarrollo 2017-2021 del Cantón San Jacinto de Buena Fe, Provincia de los Ríos*.
<https://multimedia.planificacion.gob.ec/PDOT/descargas.html>
- Genuer, R., Poggi, J.-M., Tuleau-Malot, C., & Villa-Vialaneix, N. (2017). Random forests for big data. *Big Data Research*, 9, 28–46.
<https://doi.org/10.1016/j.bdr.2017.07.003>
- Gonzaga Aguilar, C. (2014). *Aplicación de índices de vegetación derivados de imágenes satelitales Landsat 7 ETM+ y ASTER para la caracterización de la cobertura vegetal en la zona centro de la provincia de Loja, Ecuador* [Universidad Nacional de La Plata].
<https://doi.org/10.35537/10915/34487>
- Graziano Ceddia, M. (2019). The impact of income, land, and wealth inequality on agricultural expansion in Latin America. *Proceedings of the National Academy of Sciences of the United States of America*, 116(7), 2527–2532.
<https://doi.org/10.1073/pnas.1814894116>

- Ha, N. T., Manley-Harris, M., Pham, T. D., & Hawes, I. (2020). A comparative assessment of ensemble-based machine learning and maximum likelihood methods for mapping seagrass using Sentinel-2 imagery in Tauranga Harbor, New Zealand. *Remote Sensing*, 12(3), 355. <https://doi.org/10.3390/rs12030355>
- Hernandez-Clemente, R., & Hornero, A. (2021). Monitoring and assessment of desertification using remote sensing. *Ecosistemas: revista científica y tecnica de ecologia y medio ambiente*, 30(3), 2240. <https://doi.org/10.7818/ecos.2240>
- Jay, S. C., Lawrence, R., Repasky, K., & Keith, C. (2009). Invasive species mapping using low cost hyper spectral imagery. *ASPRS 2009 Annual Conference*, 1–9.
- Labrador García, M., Évora Brondo, J. A., & Arbelo, M. (2012). *Satélites de teledetección para la gestión del territorio*. https://www.researchgate.net/publication/259230060_Satelites_de_teledeteccion_para_la_gestion_del_territorio
- Lastovicka, J., Svec, P., Paluba, D., Kobliuk, N., Svoboda, J., Hladky, R., & Stych, P. (2020). Sentinel-2 data in an evaluation of the impact of the disturbances on forest vegetation. *Remote Sensing*, 12(12), 1–26. <https://doi.org/10.3390/rs12121914>
- Leng, X., Feng, X., & Fu, B. (2020). Driving forces of agricultural expansion and land degradation indicated by Vegetation Continuous Fields (VCF) data in drylands from 2000 to 2015. *Global Ecology and Conservation*, 23, 1–16. <https://doi.org/10.1016/j.gecco.2020.e01087>
- MAG. (2020). *Resumen Ejecutivo de los Diagnósticos Territoriales del Sector Agrario*. Ministerio de Agricultura y Ganadería – Coordinación General de Planificación y Gestión Estratégica. https://www.agricultura.gob.ec/wp-content/uploads/2020/08/Resumen-Ejecutivo-Diagn%C3%B3sticos-Territoriales-del-Sector-Agrario_14-08-2020-1_compressed.pdf

- Márquez Poblete, M. A., & Veloso Pérez, E. (2021). El ordenamiento territorial en Chile: estado del arte. *Estado, gobierno, gestión pública*, 18(35). <https://doi.org/10.5354/0717-8980.2020.61424>
- Marrón Ramos, D. N., Reyes Valenzuela, R., González Torres, A., Juárez Rodríguez, R., & Mendoza Montero, F. Y. (2022). Evaluación de la deserción a nivel superior: dimensiones que inciden en carreras universitarias. *RIDE revista iberoamericana para la investigación y el desarrollo educativo*, 13(25). <https://doi.org/10.23913/ride.v13i25.1269>
- Miguel Ayala, L., van Eupen, M., Zhang, G., Pérez-Soba, M., Martorano, L. G., Lisboa, L. S., & Beltrao, N. E. (2016). Impact of agricultural expansion on water footprint in the Amazon under climate change scenarios. *The Science of the Total Environment*, 569–570, 1159–1173. <https://doi.org/10.1016/j.scitotenv.2016.06.191>
- Morell-Monzó, S., Estornell, J., & Sebastiá-Frasquet, M. T. (2022). Clasificación del estado de parcelas de cítricos utilizando datos multitemporales Sentinel-2. *Teledetección para una Agricultura Sostenible en la era del Big Data, Actas del XIX Congreso de la Asociación Española de Teledetección*. https://www.researchgate.net/publication/363349545_Clasificacion_del_estado_de_parcelas_de_citricos_utilizando_datos_multitemporales_Sentinel-2
- Muñoz Murcillo, J. L., Gentili, J., & Bustos Cara, R. (2020). Uso agrícola del suelo y demanda de agua para riego en la cuenca del río Vices (Ecuador) durante el período 1990 – 2014. *Investigaciones geográficas*, 1(59), 91. <https://doi.org/10.5354/0719-5370.2020.56958>
- Ngoma, H., Pelletier, J., Mulenga, B. P., & Subakanya, M. (2021). Climate-smart agriculture, cropland expansion and deforestation in Zambia: Linkages, processes and drivers. *Land Use Policy*, 107, 1–15. <https://doi.org/10.1016/j.landusepol.2021.105482>

- Nogueira de Brito, B., & Silva Brito, J. L. (2020). Mapeamento de qualidade em pastagens do Cerrado por meio de imagens Sentinel 2. *Geografia Ensino & Pesquisa*, 24(44). <https://doi.org/10.5902/2236499443352>
- Nolasco, M., Bocco, M., & Scavuzzo, C. (2017). Mapeo de cultivos agrícolas en Córdoba: una comparación del desempeño de tres algoritmos de clasificación. *IX Congreso Argentino de AgroInformática (CAI 2017) - JAIIO 46-CLEI 43* (Córdoba, 2017). <http://sedici.unlp.edu.ar/handle/10915/62830>
- NU. CEPAL. (2017). El cambio climático, la agricultura y la pobreza en América Latina. *Naciones Unidas*, 1, 1–17. <https://hdl.handle.net/11362/45696>
- OECD, & Food and Agriculture Organization of the United Nations. (2016). *OCDE-FAO Perspectivas Agrícolas 2016-2025*. OECD. https://doi.org/10.1787/agr_outlook-2016-es
- Ok, A. O., Akar, O., & Gungor, O. (2012). Evaluation of random forest method for agricultural crop classification. *European Journal of Remote Sensing*, 45(1), 421–432. <https://doi.org/10.5721/eujrs20124535>
- Pech May, F., Sánchez Hernández, J. V., & Sánchez Jacinto, H. (2021). Análisis de zonas de cultivo y cuerpos de agua mediante el cálculo de índices radiométricos con imágenes Sentinel-2. *Lámpsakos*, 24, 48. <https://doi.org/10.21501/21454086.3601>
- Ponce Cevallos, J. (2016). *La Política Agropecuaria Ecuatoriana hacia el desarrollo territorial rural sostenible: 2015-2025 Versión 1*. Quito, Ecuador: Sector Público Gubernamental. <http://www2.competencias.gob.ec/wp-content/uploads/2021/03/02-06PPP2015-POLITICA02-1.pdf>
- Ramírez, M., Martínez, L., Montilla, M., Sarmiento, O., Lasso, J., & Díaz, S. (2020). Obtención de coberturas del suelo agropecuarias en imágenes satelitales Sentinel-2 con la inyección de imágenes de dron usando Random Forest en Google Earth Engine. *Revista de teledetección*, 56, 49. <https://doi.org/10.4995/raet.2020.14102>
- Rhew, I. C., Vander Stoep, A., Kearney, A., Smith, N. L., & Dunbar, M. D. (2011). Validation of the normalized difference vegetation index as a measure of

- neighborhood greenness. *Annals of Epidemiology*, 21(12), 946–952.
<https://doi.org/10.1016/j.annepidem.2011.09.001>
- Rodríguez Pérez, D., Sánchez-Carnero, N., Domínguez Gómez, J. A., & Marta Pastrana, C. S. (2015). *Cuestiones de teledetección*. UNED - Universidad Nacional de Educación a Distancia.
<https://dialnet.unirioja.es/servlet/libro?codigo=574612>
- Sánchez Turcios, R. A. (2015). Prueba de Wilcoxon-Mann-Whitney: mitos y realidades. *Revista mexicana de endocrinología metabolismo & nutrición*, 2(1), 18–21.
- Sánchez-Méndez, A. G., & Arguijo-Hernández, S. P. (2018). Análisis de imágenes multiespectrales para la detección de cultivos y detección de plagas y enfermedades en la producción de café. *Research in Computing Science*, 147(7), 309–317. <https://doi.org/10.13053/racs-147-7-24>
- Santiago Bazan, F., Mallqui Meza, H., & Rios Recra, R. (2021). Mapeo de la cobertura vegetal en la subcuenca Quillcay (Ancash – Perú) con el clasificador de Árbol de decisiones. *Aporte Santiaguino*, 14(1), 78–91.
<https://doi.org/10.32911/as.2021.v14.n1.761>
- Smida, Z., Cucala, L., Gannoun, A., & Durif, G. (2022). A Wilcoxon-Mann-Whitney spatial scan statistic for functional data. *Computational Statistics & Data Analysis*, 167(107378), 107378.
<https://doi.org/10.1016/j.csda.2021.107378>
- Song, X.-P., Huang, W., Hansen, M. C., & Potapov, P. (2021). An evaluation of Landsat, Sentinel-2, Sentinel-1 and MODIS data for crop type mapping. *Science of Remote Sensing*, 3.
<https://doi.org/10.1016/j.srs.2021.100018>
- Sporchia, F., Taherzadeh, O., & Caro, D. (2021). Stimulating environmental degradation: A global study of resource use in cocoa, coffee, tea and tobacco supply chains. *Current Research in Environmental Sustainability*, 3(100029). <https://doi.org/10.1016/j.crsust.2021.100029>
- Titti, G., Sarretta, A., Lombardo, L., Crema, S., Pasuto, A., & Borgatti, L. (2022). Mapping susceptibility with open-source tools: A new plugin for

- QGIS. *Frontiers in earth science*, 10.
<https://doi.org/10.3389/feart.2022.842425>
- USGS-U.S. Geological Survey. (s/f). *EarthExplorer*. USGS - U.S. Geological Survey. <https://earthexplorer.usgs.gov/>
- Valero Medina, J. A., & Alzate Atehortúa, B. E. (2019). Comparison of maximum likelihood, support vector machines, and random forest techniques in satellite images classification. *Tecnura*, 23(59), 13–26.
<https://doi.org/10.14483/22487638.14826>
- Vega, A., Trujillo, R., Cortéz, L., Paruelo, J., Volante, J., Mosciaro, J., Morales Poclava, M., Vale, L., Castrillo, S., Sawchik, J., Tiscornia, G., Fuente, M., & Maldonado, I. (2015). Expansión agrícola en Argentina, Bolivia, Paraguay, Uruguay y Chile entre 2000-2010. Caracterización espacial mediante series temporales de índices de vegetación. *RIA. Revista de investigaciones agropecuarias*, 41(2), 179–191.
<https://www.redalyc.org/articulo.oa?id=86441580011>
- Vera Rojas, N. J. (2020). *Análisis de uso de suelo a partir de imágenes satelitales sentinel 2 en el cantón Buena Fe, provincia de Los Ríos* [UTEQ].
<http://repositorio.uteq.edu.ec/handle/43000/6186>
- Viteri-Salazar, O., & Toledo, L. (2020). The expansion of the agricultural frontier in the northern Amazon region of Ecuador, 2000–2011: Process, causes, and impact. *Land Use Policy*, 99(104986), 104986.
<https://doi.org/10.1016/j.landusepol.2020.104986>
- Vura, S., Patil, P., & Patil, S. B. (2023). A study of different compression algorithms for multispectral images. *Materials Today: Proceedings*, 80, 2193–2197. <https://doi.org/10.1016/j.matpr.2021.06.175>
- Waśniewski, A., Hościło, A., Zagajewski, B., & Moukétou-Tarazewicz, D. (2020). Assessment of Sentinel-2 satellite images and Random Forest classifier for Rainforest mapping in Gabon. *Forests*, 11(9), 941.
<https://doi.org/10.3390/f11090941>

- Weiss, M., Jacob, F., & Duveiller, G. (2020). Remote sensing for agricultural applications: A meta-review. *Remote Sensing of Environment*, 236, 111402. <https://doi.org/10.1016/j.rse.2019.111402>
- Xiao, C., Li, P., Feng, Z., Liu, Y., & Zhang, X. (2020). Sentinel-2 red-edge spectral indices (RESI) suitability for mapping rubber boom in Luang Namtha Province, northern Lao PDR. *International Journal of Applied Earth Observation and Geoinformation: ITC Journal*, 93. <https://doi.org/10.1016/j.jag.2020.102176>
- Yang, Y., Chen, Y., Yang, K., Cermak, J., & Chen, Y. (2021). High-resolution aerosol retrieval over urban areas using sentinel-2 data. *Atmospheric Research*, 264, 105829. <https://doi.org/10.1016/j.atmosres.2021.105829>
- Zepeda-Jazo, I. (2018). Manejo sustentable de plagas agrícolas en México. *AGRICULTURA, SOCIEDAD Y DESARROLLO*, 15(1), 99–108. https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1870-54722018000100099



ANEXOS

Anexos

Anexo 1

Metadatos completos de la imagen 1 de Sentinel-2

Summary	
Date:	2019-08-15T15:36:21.024Z
Filename:	S2A_MSIL1C_20190815T153621_N0208_R068_T17MPV_2 0190815T222516.SAFE
Identifier:	S2A_MSIL1C_20190815T153621_N0208_R068_T17MPV_2 0190815T222516
Instrument:	MSI
Satellite:	Sentinel-2
Size:	523.86 MB
Product	
Cloud cover percentage:	13.742
Datastrip identifier:	S2A_OPER_MSI_L1C_DS_MTI__20190815T222516_S2019 0815T153622_N02.08
Datatake sensing start:	2019-08-15T15:36:21.024Z
Degraded ancillary data percentage:	0.0
Degraded MSI data percentage:	0
Footprint:	<pre><gml:PolygonsrsName="http://www.opengis.net/gml/srs/epsg.xml#4326" xmlns:gml="http://www.opengis.net/gml"> <gml:outerBoundaryIs> <gml:LinearRing> <gml:coordinates>-0.993166101164354,-79.85373 0.850069697749979,-79.82184 -0.701618985745763,-79.78897 -0.553110753375327,-79.75607 0.404495984471584,-79.723236 -0.25575616884828,-79.69064 -0.106912060607654,-79.65814 0,-79.634766 0,-79.11493 -0.992852760559312,-79.114655 0.993166101164354,-79.85373</gml:coordinates> </gml:LinearRing> </gml:outerBoundaryIs> </gml:Polygon></pre>
Format:	SAFE
Format correctness:	PASSED
General quality:	PASSED
Generation time:	2019-08-15T22:25:16.000000Z
Geometric quality:	PASSED

Granule identifier:	S2A_OPER_MSI_L1C_TL_MTI__20190815T222516_A021658_T17MPV_N02.08
Ingestion Date:	2019-08-16T00:25:18.584Z
JTS footprint:	MULTIPOLYGON (((-79.85373 -0.993166101164354, -79.114655 -0.992852760559312, -79.11493 0, -79.634766 0, -79.65814 -0.106912060607654, -79.69064 -0.25575616884828, -79.723236 -0.404495984471584, -79.75607 -0.553110753375327, -79.78897 -0.701618985745763, -79.82184 -0.850069697749979, -79.85373 -0.993166101164354)))
Level-1C PDI Identifier:	S2A_OPER_MSI_L1C_TL_MTI__20190815T222516_A021658_T17MPV_N02.08
Mission datatake id:	GS2A_20190815T153621_021658_N02.08
Orbit number (start):	21658
Pass direction:	DESCENDING
Processing baseline:	02.08
Processing level:	Level-1C
Product type:	S2MSI1C
Radiometric quality:	PASSED
Relative orbit (start):	68
Sensing start:	2019-08-15T15:36:21.024Z
Sensing stop:	2019-08-15T15:36:21.024Z
Sensor quality:	PASSED
Tile Identifier:	17MPV
Tile Identifier horizontal order:	MV17P

Instrument

Instrument abbreviation:	MSI
Instrument mode:	INS-NOBS
Instrument name:	Multi-Spectral Instrument

Platform

NSSDC identifier:	2015-028A
Satellite name:	Sentinel-2
Satellite number:	A

Nota: Copernicus by the European Union's Earth Observation Programme (Copernicus Open Access Hub, s.f.)

Anexo 2

Metadatos completos de la imagen 2 de Sentinel-2

Summary	
Date:	2019-08-20T15:36:29.024Z
Filename:	S2B_MSIL1C_20190820T153629_N0208_R068_T17MPU_2 0190820T202350.SAFE
Identifier:	S2B_MSIL1C_20190820T153629_N0208_R068_T17MPU_2 0190820T202350
Instrument:	MSI
Satellite:	Sentinel-2
Size:	667.35 MB
Product	
Cloud cover percentage:	29.477.399.999.999.900
Datastrip identifier:	S2B_OPER_MSI_L1C_DS_MTI__20190820T202350_S2019 0820T154239_N02.08
Datatake sensing start:	2019-08-20T15:36:29.024Z
Degraded ancillary data percentage:	0.0
Degraded MSI data percentage:	0
Footprint:	<pre> <gml:PolygonsrsName="http://www.opengis.net/gml/srs/epsg .xml#4326"xmlns:gml="http://www.opengis.net/gml"><gml:outerBoundaryls> <gml:LinearRing> <gml:coordinates>1.897482275386293,-80.05121 1.834367448292868,-80.03732 -1.685812558585486,- 80.004456 -1.537242122446741,-79.97159 1.388624105601028,-79.93878 -1.239978256968598,- 79.906006 -1.091349700449564,-79.8732 0.942781736155273,-79.8403 -0.90415271374235,- 79.83176 -0.903875955648899,-79.114685 1.896723233843827,-79.11389 -1.897482275386293,- 80.05121</gml:coordinates></gml:LinearRing> </gml:outerBoundaryls> </gml:Polygon> </pre>
Format:	SAFE
Format correctness:	PASSED
General quality:	PASSED
Generation time:	2019-08-20T20:23:50.000000Z
Geometric quality:	PASSED

Anexos:

Granule identifier:	S2B_OPER_MSI_L1C_TL_MTI__20190820T202350_A0128 21_T17MPU_N02.08
Ingestion Date:	2019-08-20T21:44:30.140Z
JTS footprint:	MULTIPOLYGON (((-80.05121 -1.897482275386293, - 79.11389 -1.896723233843827, -79.114685 - 0.903875955648899, -79.83176 -0.90415271374235, - 79.8403 -0.942781736155273, -79.8732 - 1.091349700449564, -79.906006 -1.239978256968598, - 79.93878 -1.388624105601028, -79.97159 - 1.537242122446741, -80.004456 -1.685812558585486, - 80.03732 -1.834367448292868, -80.05121 - 1.897482275386293)))
Level-1C PDI Identifier:	S2B_OPER_MSI_L1C_TL_MTI__20190820T202350_A0128 21_T17MPU_N02.08
Mission datatake id:	GS2B_20190820T153629_012821_N02.08
Orbit number (start):	12821
Pass direction:	DESCENDING
Processing baseline:	02.08
Processing level:	Level-1C
Product type:	S2MSI1C
Radiometric quality:	PASSED
Relative orbit (start):	68
Sensing start:	2019-08-20T15:36:29.024Z
Sensing stop:	2019-08-20T15:36:29.024Z
Sensor quality:	PASSED
Tile Identifier:	17MPU
Tile Identifier horizontal order:	MU17P

Instrument

Instrument abbreviation:	MSI
Instrument mode:	INS-NOBS
Instrument name:	Multi-Spectral Instrument

Platform

NSSDC identifier:	2017-013A
Satellite name:	Sentinel-2
Satellite number:	B

Nota: Copernicus by the European Union's Earth Observation Programme (Copernicus Open Access Hub, s.f.)

Anexo 3

Informe completo de la separabilidad espectral de ROIs (Superior e Inferior) de la imagen Sentinel-2

<p>Input File: 2019_08_15_Stack_plus_Reort.emp.tif ROI Name: (Jeffries-Matusita, Transformed Divergence)</p> <p>ZA_Humano: ZC_Banano: (1.97312941 2.00000000) ZC_Cacao: (1.99981198 2.00000000) ZC_palma: (1.93815874 2.00000000) ZC_otros: (1.97577875 2.00000000) Suelo_desnudo: (1.82167252 2.00000000) Cuerpos_agua: (1.99711016 2.00000000) Nubes: (1.94203656 2.00000000)</p> <p>ZC_Banano: ZA_Humano: (1.97312941 2.00000000) ZC_Cacao: (1.98352175 2.00000000) ZC_palma: (1.32903022 1.48471259) ZC_otros: (0.68354598 0.77360640) Suelo_desnudo: (1.61129775 1.94410199) Cuerpos_agua: (1.99882840 2.00000000) Nubes: (1.98892649 2.00000000)</p> <p>ZC_Cacao: ZA_Humano: (1.99981198 2.00000000) ZC_Banano: (1.98352175 2.00000000) ZC_palma: (1.98657363 2.00000000) ZC_otros: (1.95764629 2.00000000) Suelo_desnudo: (1.99330397 2.00000000) Cuerpos_agua: (1.9998261 2.00000000) Nubes: (1.99993988 2.00000000)</p> <p>ZC_palma: ZA_Humano: (1.93815874 2.00000000) ZC_Banano: (1.32903022 1.48471259) ZC_Cacao: (1.98657363 2.00000000) ZC_otros: (1.48053228 1.76381436) Suelo_desnudo: (1.56560161 1.95974339) Cuerpos_agua: (1.99832728 2.00000000) Nubes: (1.98814341 2.00000000)</p> <p>ZC_otros: ZA_Humano: (1.97577875 2.00000000) ZC_Banano: (0.68354598 0.77360640) ZC_Cacao: (1.95764629 2.00000000) ZC_palma: (1.48053228 1.76381436) Suelo_desnudo: (1.58244681 1.91326267) Cuerpos_agua: (1.99796131 2.00000000) Nubes: (1.99178047 2.00000000)</p> <p>Suelo_desnudo: ZA_Humano: (1.82167252 2.00000000) ZC_Banano: (1.61129775 1.94410199) ZC_Cacao: (1.99330397 2.00000000) ZC_palma: (1.56560161 1.95974339) ZC_otros: (1.58244681 1.91326267) Cuerpos_agua: (1.99821666 2.00000000) Nubes: (1.94179839 2.00000000)</p> <p>Cuerpos_agua: ZA_Humano: (1.99711016 2.00000000) ZC_Banano: (1.99882840 2.00000000) ZC_Cacao: (1.9998261 2.00000000) ZC_palma: (1.99832728 2.00000000) ZC_otros: (1.99796131 2.00000000) Suelo_desnudo: (1.99821666 2.00000000) Nubes: (1.99958429 2.00000000)</p> <p>Nubes: ZA_Humano: (1.94203656 2.00000000) ZC_Banano: (1.98892649 2.00000000) ZC_Cacao: (1.99993988 2.00000000) ZC_palma: (1.98814341 2.00000000) ZC_otros: (1.99178047 2.00000000) Suelo_desnudo: (1.94179839 2.00000000) Cuerpos_agua: (1.99958429 2.00000000)</p> <p>Pair Separation (least to most);</p> <p>ZC_Banano and ZC_otros - 0.68354598 ZC_Banano and ZC_palma - 1.32903022 ZC_palma and ZC_otros - 1.48053228 ZC_palma and Suelo_desnudo - 1.56560161 ZC_otros and Suelo_desnudo - 1.58244681 ZC_Banano and Suelo_desnudo - 1.61129775 ZA_Humano and Suelo_desnudo - 1.82167252 ZA_Humano and ZC_palma - 1.93815874 Suelo_desnudo and Nubes - 1.94179839 ZA_Humano and Nubes - 1.94203656 ZC_Cacao and ZC_otros - 1.95764629 ZA_Humano and ZC_Banano - 1.97312941 ZA_Humano and ZC_otros - 1.97577875 ZC_Banano and ZC_Cacao - 1.98352175 ZC_Cacao and ZC_palma - 1.98657363 ZC_palma and Nubes - 1.98814341 ZC_Banano and Nubes - 1.98892649 ZC_otros and Nubes - 1.99178047 ZC_Cacao and Suelo_desnudo - 1.99330397 ZA_Humano and Cuerpos_agua - 1.99711016 ZC_otros and Cuerpos_agua - 1.99796131 Suelo_desnudo and Cuerpos_agua - 1.99821666 ZC_palma and Cuerpos_agua - 1.99832728 ZC_Banano and Cuerpos_agua - 1.99882840 Cuerpos_agua and Nubes - 1.99958429 ZA_Humano and ZC_Cacao - 1.99981198 ZC_Cacao and Cuerpos_agua - 1.9998261 ZC_Cacao and Nubes - 1.99993988</p>	<p>Input File: 2019_08_20_Stack_corte.tif ROI Name: (Jeffries-Matusita, Transformed Divergence)</p> <p>ZA_Humano: ZC_Banano: (1.98990007 1.99999633) ZC_Cacao: (1.99998576 2.00000000) ZC_palma: (1.88765645 1.99983155) ZC_otros: (1.72064429 1.97005931) Suelo_desnudo: (1.32544954 1.56881498) Cuerpos_agua: (1.74671900 1.97612922) Nubes: (1.68916961 1.94688048)</p> <p>ZC_Banano: ZA_Humano: (1.98990007 1.99999633) ZC_Cacao: (1.99998576 2.00000000) ZC_palma: (1.88771679 1.98423020) ZC_otros: (1.75190896 1.96146753) Suelo_desnudo: (1.98157168 1.99997850) Cuerpos_agua: (1.99426826 1.99999863) Nubes: (1.92386870 1.99951784)</p> <p>ZC_Cacao: ZA_Humano: (1.99998576 2.00000000) ZC_Banano: (1.99998327 2.00000000) ZC_palma: (1.99989153 2.00000000) ZC_otros: (1.99447147 2.00000000) Suelo_desnudo: (1.99990557 2.00000000) Cuerpos_agua: (1.99999988 2.00000000) Nubes: (1.99950201 2.00000000)</p> <p>ZC_palma: ZA_Humano: (1.88765645 1.99983155) ZC_Banano: (1.88771679 1.98423020) ZC_Cacao: (1.99989153 2.00000000) ZC_otros: (1.77333284 1.99825902) Suelo_desnudo: (1.93199177 1.99999387) Cuerpos_agua: (1.97573437 1.99996463) Nubes: (1.91602770 1.99994028)</p> <p>ZC_otros: ZA_Humano: (1.72064429 1.97005931) ZC_Banano: (1.75190896 1.96146753) ZC_Cacao: (1.99447147 2.00000000) ZC_palma: (1.77333284 1.99825902) Suelo_desnudo: (1.72381395 1.94780308) Cuerpos_agua: (1.96362478 1.99973670) Nubes: (1.37015568 1.55952499)</p> <p>Suelo_desnudo: ZA_Humano: (1.32544954 1.56881498) ZC_Banano: (1.98157168 1.99997850) ZC_Cacao: (1.99990557 2.00000000) ZC_palma: (1.93199177 1.99999387) ZC_otros: (1.72381395 1.94780308) Cuerpos_agua: (1.55087096 1.88507452) Nubes: (1.66751891 1.91982728)</p> <p>Cuerpos_agua: ZA_Humano: (1.74671900 1.97612922) ZC_Banano: (1.99426826 1.99999863) ZC_Cacao: (1.99999988 2.00000000) ZC_palma: (1.97573437 1.99996463) ZC_otros: (1.96362478 1.99973670) Suelo_desnudo: (1.55087096 1.88507452) Nubes: (1.83188186 1.98345381)</p> <p>Nubes: ZA_Humano: (1.68916961 1.94688048) ZC_Banano: (1.92386870 1.99951784) ZC_Cacao: (1.99950201 2.00000000) ZC_palma: (1.91602770 1.99994028) ZC_otros: (1.37015568 1.55952499) Suelo_desnudo: (1.66751891 1.91982728) Cuerpos_agua: (1.83188186 1.98345381)</p> <p>Pair Separation (least to most);</p> <p>ZA_Humano and Suelo_desnudo - 1.32544954 ZC_otros and Nubes - 1.37015568 Suelo_desnudo and Cuerpos_agua - 1.55087096 Suelo_desnudo and Nubes - 1.66751891 ZA_Humano and Nubes - 1.68916961 ZA_Humano and ZC_otros - 1.72064429 ZC_otros and Suelo_desnudo - 1.72381395 ZA_Humano and Cuerpos_agua - 1.74671900 ZC_Banano and ZC_otros - 1.75190896 ZC_palma and ZC_otros - 1.77333284 Cuerpos_agua and Nubes - 1.83188186 ZA_Humano and ZC_palma - 1.88765645 ZC_Banano and ZC_palma - 1.88771679 ZC_palma and Nubes - 1.91602770 ZC_Banano and Nubes - 1.92386870 ZC_palma and Suelo_desnudo - 1.93199177 ZC_otros and Cuerpos_agua - 1.96362478 ZC_palma and Cuerpos_agua - 1.97573437 ZC_Banano and Suelo_desnudo - 1.98157168 ZA_Humano and ZC_Banano - 1.98990007 ZC_Banano and Cuerpos_agua - 1.99426826 ZC_Cacao and ZC_otros - 1.99447147 ZC_Cacao and Nubes - 1.99950201 ZC_Banano and ZC_Cacao - 1.99998327 ZC_Cacao and ZC_palma - 1.99989153 ZC_Cacao and Suelo_desnudo - 1.99990557 ZA_Humano and ZC_Cacao - 1.99998576 ZC_Cacao and Cuerpos_agua - 1.99999988</p>
--	--

Nota: Autor (2023)

Anexo 4

Toma de coordenadas para áreas de entrenamiento y generación de ROIs para zonas de cultivo de interés y otros



Nota: Autor (2023)

Anexo 5

Toma de coordenadas para áreas de entrenamiento al azar y generación de ROIs para proceso de validación con zonas de cultivo de cacao y banano



Nota: Autor (2023)

Anexo 6

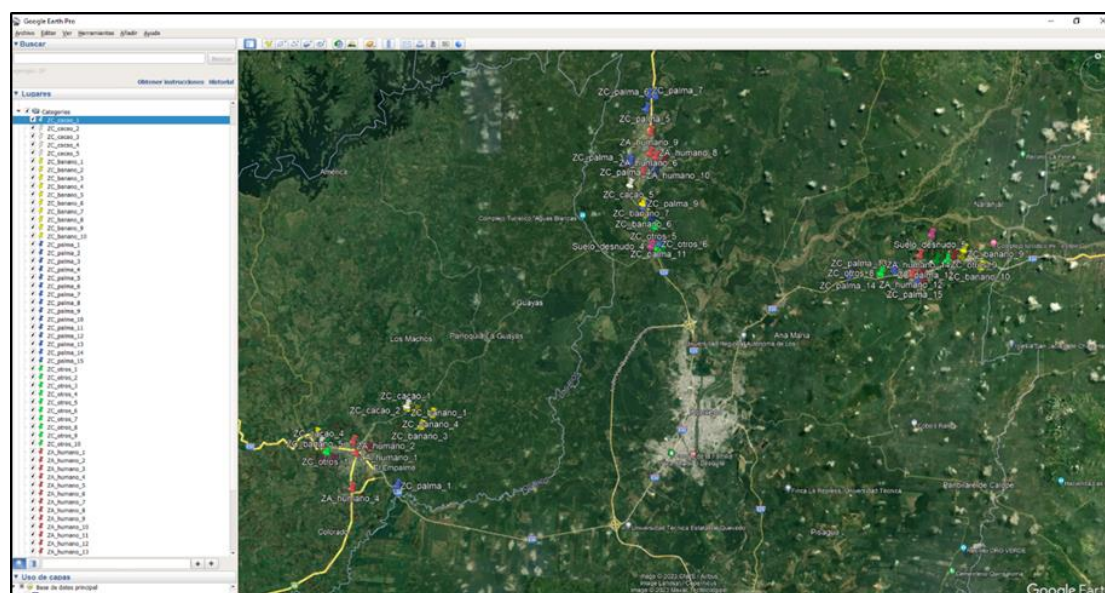
Toma de coordenadas para áreas de entrenamiento al azar y generación de ROIs para proceso de validación con zonas de otras coberturas y asentamientos humanos



Nota: Autor (2023)

Anexo 7

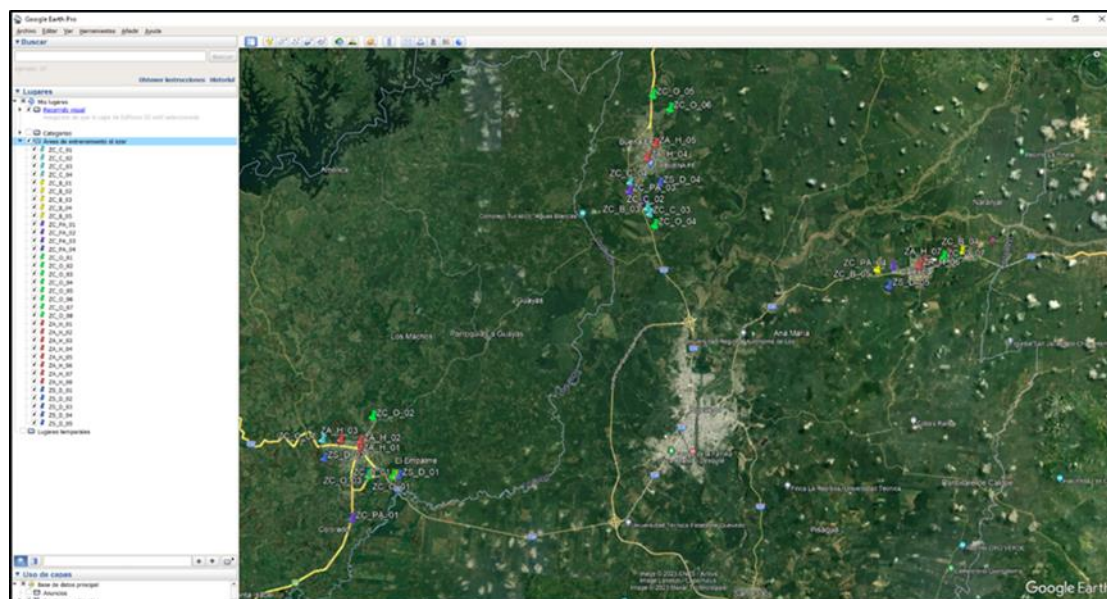
Determinación de áreas de entrenamiento para generación de ROIs en la zona de estudio



Nota: Autor (2023)

Anexo 8

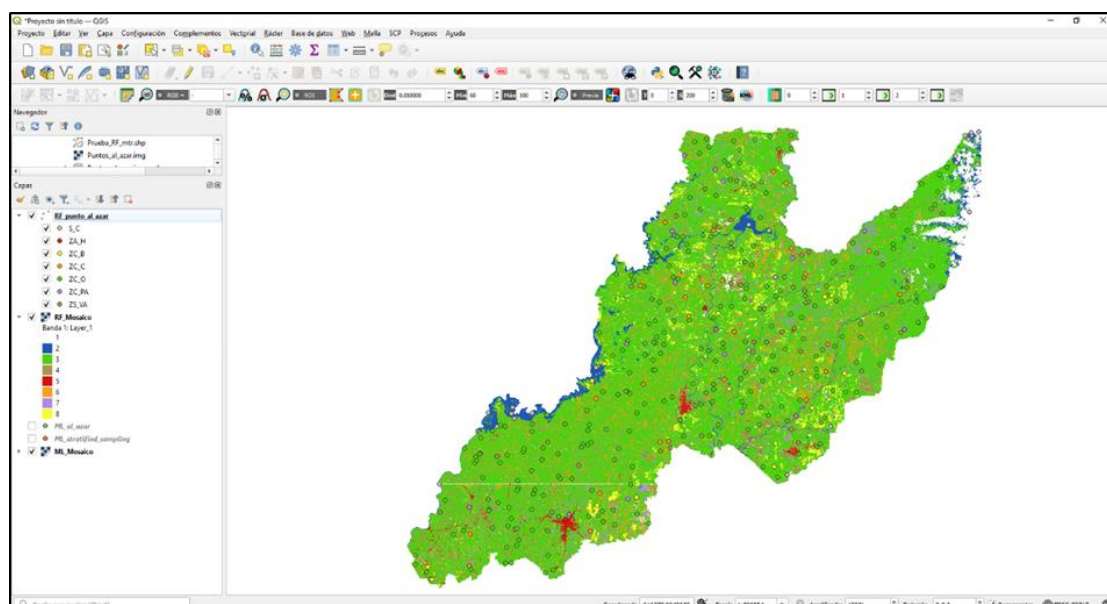
Determinación de áreas de entrenamiento al azar para generación de ROIs y posterior análisis estadístico



Nota: Autor (2023)

Anexo 9

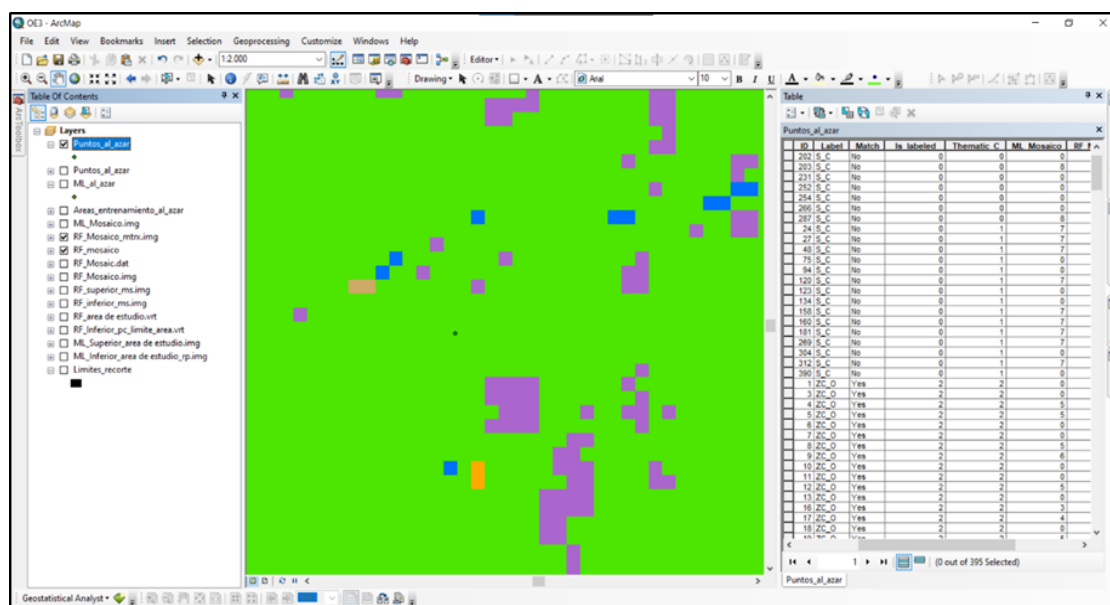
Generación de puntos al azar para método estadístico



Nota: Autor (2023)

Anexo 10

Obtención de datos aleatorios para examinar las diferencias estadísticas mediante la utilización de la función "Extract Multi Values to Point"



Nota: Autor (2023)

Anexo 11

Datos obtenidos a partir de herramienta SAGA de QGIS para el análisis de matriz de confusión y Coeficiente Kappa

	CLASS	2	6	1	5	3	7	4	SumUser	AccUser
1	2	260	6	0	3	2	2	0	273	95.238095
2	6	1	10	0	0	0	0	0	11	90.909091
3	1	2	0	12	0	0	0	0	14	85.714286
4	5	5	0	0	23	0	0	0	28	82.142857
5	3	1	0	0	0	26	0	0	27	96.296296
6	7	1	0	0	0	0	13	0	14	92.857143
7	4	0	0	0	0	0	0	3	3	100
8	0	1	0	0	0	1	0	0	0	0
9	0	270	16	12	26	28	15	3		
10	0	96.296296	62.5	100	88.461538	92.857143	86.666667	100		

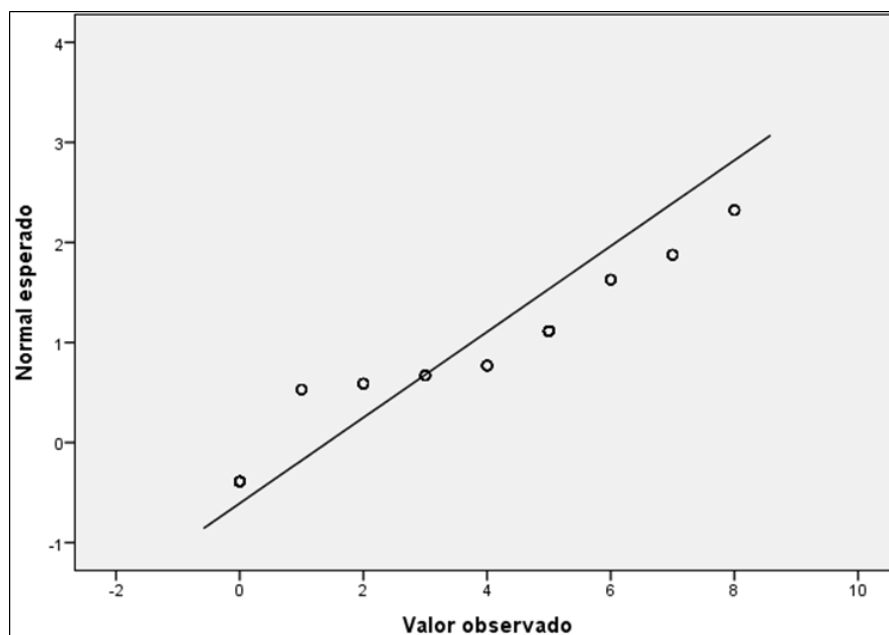
	Class	SumRef	AccProd	SumClassified	AccUser
1	2	270	0.962963	273	0.952381
2	6	16	0.625	11	0.909091
3	1	12	1	14	0.857143
4	5	26	0.884615	28	0.821429
5	3	28	0.928571	27	0.962963
6	7	15	0.866667	14	0.928571
7	4	3	1	3	1

	NAME	VALUE
1	Kappa	0.860818
2	Overall Accura	0.937838

Nota: Autor (2023)

Anexo 12

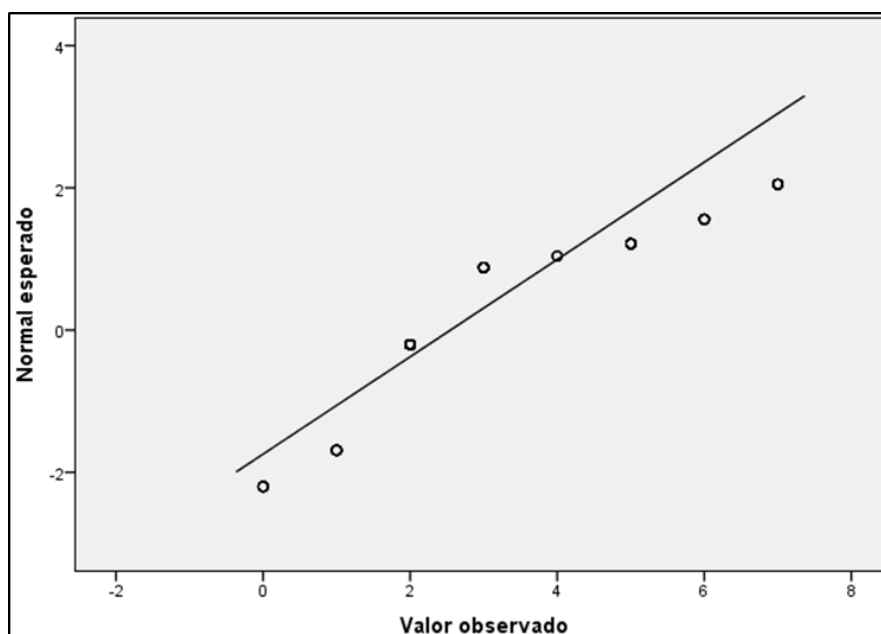
Gráfico Q-Q para normalidad de los datos de Maximum Likelihood



Nota: Autor (2023)

Anexo 13

Gráfico Q-Q para normalidad de los datos de Random Forest



Nota: Autor (2023)

RESUMEN

Se analiza la expansión e intensificación de la agricultura no regulada e insostenible, principalmente en América Latina, enfocándose en Ecuador, donde la degradación del suelo está estrechamente relacionada con prácticas agrícolas insostenibles. Destaca la importancia de una planificación territorial exhaustiva y el uso de teledetección para evaluar cambios en el uso del suelo y la cobertura vegetal. La investigación se centra en la zona norte de las provincias del Guayas y de Los Ríos en Ecuador, enfocándose en el procesamiento de imágenes satelitales para detectar cultivos de banano, cacao y palma. Utiliza una metodología experimental y cuantitativa para comparar la efectividad de distintas metodologías de clasificación, evaluando la precisión del método de inteligencia artificial Bosques Aleatorios usando imágenes del satélite Sentinel-2. Los resultados del estudio incluyen el preprocesamiento de imágenes satelitales, el cálculo de índices espectrales NDVI y RESI, y la clasificación supervisada de las imágenes de Sentinel-2 utilizando los métodos de Máxima Verosimilitud y Bosques Aleatorios. Se realizaron análisis estadísticos detallados, incluyendo la validación de resultados mediante matrices de confusión y el cálculo del coeficiente Kappa. En conclusión, el manuscrito ofrece una comprensión profunda de cómo la teledetección y la inteligencia artificial pueden ayudar en la detección y clasificación de cultivos en áreas específicas, contribuyendo a la gestión agrícola y la planificación territorial sostenible.

Palabras clave: Maximum Likelihood, Random Forest, Cultivos, NDVI, RESI.



<http://www.editorialgrupo-aea.com>



[Editorial Grupo AeA](#)



[editorialgrupoaea](#)



[Editorial Grupo AEA](#)

ISBN: 978-9942-651-17-4



9 789942 651174