



Propuesta metodológica para la detección de Fusarium raza 1, Sigatoka Negra y Picudo Negro en banano mediante el análisis de imágenes multiespectrales de muy alta resolución (UAV).

Albarracín Yugcha, Karen Michelle

Departamento de Ciencias de la Tierra y de la Construcción

Carrera de Ingeniería Geográfica y del Medio Ambiente

Trabajo de titulación, previo a la obtención del título de Ingeniera Geógrafa y del Medio Ambiente

MSC. Leiva González, César Alberto

22 de agosto del 2022



TESIS_ALBARRACIN.pdf

Scanned on: 20:14 August 23, 2022 UTC



Overall Similarity Score



Results Found



Total Words in Text

Identical Words	151
Words with Minor Changes	0
Paraphrased Words	10
Omitted Words	0



Departamento de Ciencias de la Tierra y de la Construcción
Carrera de Ingeniería Geográfica y del Medio Ambiente
Certificación

Certifico que el trabajo de titulación: **“Propuesta metodológica para la detección de Fusarium raza 1, Sigatoka Negra y Picudo Negro en banano mediante el análisis de imágenes multiespectrales de muy alta resolución (UAV)”** fue realizado por la señorita **Albarracín Yugcha, Karen Michelle**; el mismo que cumple con los requisitos legales, teóricos, científicos, técnicos y metodológicos establecidos por la Universidad de las Fuerzas Armadas ESPE, además fue revisado y analizado en su totalidad por la herramienta de prevención y/o verificación de similitud de contenidos; razón por la cual me permito acreditar y autorizar para que se lo sustente públicamente.

Sangolquí, 22 de agosto del 2022

Firma:

.....

Ing. Leiva González, Cesar Alberto, MSc.

C. C.1715416689



Departamento de Ciencias de la Tierra y de la Construcción

Carrera de Ingeniería Geográfica y del Medio Ambiente

Responsabilidad de Autoría

Yo **Albarracín Yugcha, Karen Michelle**, con cédula de ciudadanía n°1726740960, declaro que el contenido, ideas y criterios del trabajo de titulación: **“Propuesta metodológica para la detección de Fusarium raza 1, Sigatoka Negra y Picudo Negro en banano mediante el análisis de imágenes multiespectrales de muy alta resolución (UAV)”** es de mi autoría y responsabilidad, cumpliendo con los requisitos legales, teóricos, científicos, técnicos, y metodológicos establecidos por la Universidad de las Fuerzas Armadas ESPE, respetando los derechos intelectuales de terceros y referenciando las citas bibliográficas.

Sangolquí, 22 de agosto del 2022

Firma:

Albarracín Yugcha, Karen Michelle

C. C.1726740960



Departamento de Ciencias de la Tierra y de la Construcción

Carrera de Ingeniería Geográfica y del Medio Ambiente

Autorización de Publicación

Yo **Albarracín Yugcha, Karen Michelle**, con cédula de ciudadanía n°1726740960, autorizo a la Universidad de las Fuerzas Armadas ESPE publicar el trabajo de titulación: **“Propuesta metodológica para la detección de Fusarium raza 1, Sigatoka Negra y Picudo Negro en banano mediante el análisis de imágenes multiespectrales de muy alta resolución (UAV)”** en el Repositorio Institucional, cuyo contenido, ideas y criterios son de mi responsabilidad.

Sangolquí 22 de agosto del 2022

Firma:

Albarracín Yugcha, Karen Michelle

C. C.1726740960

Dedicatoria

A Dios, por la fortaleza y sabiduría para culminar con éxito esta etapa.

A mis padres Fabián y Mariela, que son el pilar fundamental en mi vida,
que con su amor y confianza me han guiado y son quienes me motivan
a seguir adelante y cumplir todas mis metas.

A mis abuelitas por brindarme un cariño tan sincero, por sus oraciones y
consejos que han sido de gran ayuda para mi vida.

Agradecimiento

A Dios, y a mis padres que nunca dejaron de creer en mí, y siempre están dispuestos a ayudarme a cumplir todas mis metas. Gracias por la paciencia que me han tenido y por apoyarme en todas las decisiones que he tomado a lo largo de mi vida.

A mi abuelita, tíos y toda la familia por su apoyo condicional y por la confianza depositada en mí para cumplir esta meta.

De manera especial quiero agradecer al Ing.Izar Sinde PhD., por su confianza, amistad, paciencia y apoyo. Gracias a sus valiosos conocimientos hicieron posible elaboración y culminación de este proyecto de investigación.

A mis amig@s les agradezco infinitamente por el apoyo que me brindaron y por hacer de la vida universitaria una época inolvidable.

Tabla de contenidos

Dedicatoria	6
Agradecimiento	7
Índice de Tablas.....	11
Índice de figuras.....	12
Resumen.....	14
Abstract.....	15
Capítulo I.....	16
Problema.....	16
Antecedentes	16
Planteamiento del problema.....	18
Justificación e Importancia	19
Descripción del área de estudio	19
Objetivos	21
Objetivo General.....	21
Objetivos Específicos	21
Metas.....	22
Capítulo II.....	23
Marco teórico	23
Teledetección.....	23
Espectro electromagnético	24
Vehículos aéreos no tripulados (VANT – UAV)	25
Agricultura de Precisión.....	26
Musáceas.....	26
Fusarium	26
Picudo Negro	27
Sigatoka Negra	28
Índices de vegetación.....	29
Índice Normalizado Diferencial de Vegetación (NDVI)	29
Índice De Vegetación De Diferencia Normalizada Verde (GNDVI).....	30
Índice De Diferencia Normalizada De Borde Rojo (NDRE).....	30

Índice de clorofila verde (CI_{green}).....	31
Índice de clorofila del borde rojo (CI_{re})	31
Aprendizaje Automático.....	31
Random Forest (RF).....	32
Support Vector Machine (SVM).....	32
Estadística.....	32
Análisis de la Varianza (ANOVA).....	33
Prueba de Shapiro Wilks	33
Prueba de Levene.....	33
Matriz de confusión.....	34
Índice kappa.....	36
Base Legal	37
Capítulo III.....	39
Metodología	39
Generalidades.....	39
Índices de vegetación	40
Métodos de clasificación de imágenes.....	40
Fase I	41
Materiales y Equipos	41
Delimitación del área de estudio	42
Toma de muestras con GPS	44
Planificación de los vuelos.....	48
Ejecución de los vuelos	48
Proceso Fotogramétrico para obtención de imágenes multiespectrales.....	49
Fase II	50
Obtención de los índices de vegetación.....	50
Métodos de clasificación supervisada.....	51
Support Vector Machine (SVM).....	51
Random Forest (RF).....	51
Fase III	53
Análisis estadístico	53

Análisis de la varianza Anova.....	54
Matriz de confusión.....	55
Índice kappa.....	56
Capitulo IV.....	57
Análisis de los resultados	57
Generación de ortomosaicos multiespectrales	57
Generación de índices de vegetación	59
Análisis estadístico de los índices de vegetación calculados.....	61
Análisis exploratorio de datos.....	61
Generación de Clasificación supervisada	65
Support Vector Machine	65
Random Forest	66
Análisis estadístico de la matriz de confusión.....	67
Análisis estadístico del índice kappa.....	69
Plano de afectación de Fusarium raza 1	69
Capítulo V	71
Conclusiones y recomendaciones.....	71
Conclusiones.....	71
Recomendaciones	73
Bibliografía	74

Índice de Tablas

Tabla 1 Valoración del Índice Kappa	37
Tabla 2 Equipos.....	42
Tabla 3 Tipos de Software	42
Tabla 4 Parámetros de los planes de vuelo	48
Tabla 5 Análisis exploratorio parcela 1	61
Tabla 6 Análisis exploratorio parcela 2.....	61
Tabla 7 Prueba de Shapiro Wilks de la parcela 1	62
Tabla 8 Prueba de Shapiro Wilks de la parcela 2	62
Tabla 9 Prueba de Levene de la parcela 1	63
Tabla 10 Prueba de Levene de la parcela 2.....	63
Tabla 11 Prueba de Kruskal-Wallis de la parcela 2.....	64
Tabla 12 Prueba de ANOVA de la parcela 1	65
Tabla 13 Métricas de la matriz de confusión	68
Tabla 14 Índice kappa	69

Índice de figuras

Figura 1 Catastro banano	16
Figura 2 Ubicación Parcela 1.....	20
Figura 3 Ubicación Parcela 2.....	21
Figura 4 Componentes de un sistema de teledetección.....	24
Figura 5 Regiones del espectro electromagnético.....	25
Figura 6 Síntomas del Fusarium.....	27
Figura 7 Síntomas del Picudo Negro.....	28
Figura 8 Síntomas de la Sigatoka Negra.....	29
Figura 9 Matriz de confusión	34
Figura 10 Metodología de desarrollo del proyecto.....	39
Figura 11 Camara Parrot Sequioa montada en Dron Phantom 4.....	41
Figura 12 Delimitación de la parcela 1 ubicada en Caluma.....	43
Figura 13 Delimitación de la parcela 2 ubicada en Patricia Pilar.....	44
Figura 14 Toma de muestras con GPS navegador.....	45
Figura 15 Toma de muestras con GPS navegador en la parcela 1.....	46
Figura 16 Toma de muestras con GPS navegador en la parcela 2.....	47
Figura 17 Diseño del vuelo en la parcela 1	49
Figura 18 Diseño del vuelo en la parcela 2	49
Figura 19 Herramienta “Raster Calculator”	50
Figura 20 Máscara de suelo en la parcela 1	52
Figura 21 Máscara de suelo en la parcela 2	52
Figura 22 Árbol de decisión aplicado a la parcela 1	53
Figura 23 Ortomosaico multiespectral de la parcela 1	57
Figura 24 Ortomosaico multiespectral de la parcela 2.....	58

Figura 25 Índices de vegetación calculados de la parcela 1	59
Figura 26 Índices de vegetación calculados de la parcela 2	60
Figura 27 Resultado de la clasificación de Support Vector Machine parcela 2	66
Figura 28 Resultado de la clasificación de Random Forest en la parcela 1	67
Figura 29 Resultado de la clasificación de Support Vector Machine en la parcela 1	67
Figura 30 Resultado de la clasificación de Random Forest en la parcela 2	68
Figura 31 Plano de afectación de FOC R1 en la parcela 1 ubicada en Caluma.....	70

Resumen

El banano es una planta que se ha visto afectada por varias enfermedades entre las cuales se tiene el Fusarium, Sigatoka Negra y Picudo negro, que provocan pérdidas en la productividad. El proyecto consiste en el desarrollo de una metodología que permita identificar la marchitez por FOC R1, Sigatoka Negra y Picudo Negro en banano, utilizando imágenes multiespectrales basadas en vehículos aéreos no tripulados (UAV). Para ello se trabajó durante los primeros meses del año 2022, en el cantón Caluma para la identificación de Fusarium y, en la parroquia rural Patricia Pilar para la identificación de Sigatoka Negra y Picudo Negro. En la primera fase se realizó la toma de imágenes multiespectrales mediante un vehículo aéreo no tripulado (UAV) a 80 m de altura con una cámara Parrot Sequoia. En la segunda fase se procesaron las imágenes multiespectrales para obtener orto mosaicos que sirvieron como insumos para calcular los índices de vegetación: NDVI, GNDVI, NDRE, CI_{Green} , CI_{RE} ; y aplicación de métodos de clasificación: Random Forest (RF) y Support Vector Machine (SVM). Finalmente, en la fase 3 a los índices de vegetación, se aplicó un análisis de varianza (ANOVA), y se obtuvo como resultado que los índices de vegetación no presentan diferencias significativas en la identificación de FOC R1, Sigatoka Negra y Picudo Negro en las plantas de banano. Se aplicó la matriz de confusión y el índice kappa a los algoritmos de clasificación, dando como resultado que el método más adecuado para identificación de bananos con FOC R1 es el algoritmo de Random Forest con una exactitud del 68%, precisión del 50%, sensibilidad del 75%, e índice kappa de 0.35.

Palabras clave: Fusarium en banano, Sigatoka Negra en banano, Picudo Negro, Índices de Vegetación, métodos de clasificación.

Abstract

Banana is a plant that has been affected by several diseases, including Fusarium, Black Sigatoka and Black Borer Weevil, which cause losses in productivity. Project consists of the development of a methodology to identify FOC R1, Black Sigatoka and Black Borer Weevil wilt in banana using multispectral images based on unmanned aerial vehicles (UAV). For this purpose, work was carried out during the first months of the year 2022, in the canton of Caluma for the identification of Fusarium and in the rural parish of Patricia Pilar for the identification of Black Sigatoka and Black Borer Weevil. In the first phase, multispectral images were taken using an unmanned aerial vehicle (UAV) at an altitude of 80 m with a Parrot Sequoia camera. In the second phase, multispectral images were processed to obtain ortho mosaics that served as inputs to calculate vegetation indices: NDVI, GNDVI, NDRE, CIGreen, CIRE; and application of classification methods: Random Forest (RF) and Support Vector Machine (SVM). The phase 3, an analysis of variance (ANOVA) was applied to the vegetation indices, and the result was that the vegetation indices didn't show significant differences in the identification of FOC R1, Black Sigatoka and Black Borer Weevil in banana plants. Confusion matrix and the kappa index were applied to the classification algorithms, resulting in the Random Forest algorithm being the most suitable method for identification of banana plants with FOC R1 with an accuracy of 68%, precision of 50% and sensitivity of 75%.

Key words: Fusarium in banana, Black Sigatoka in banana, Black Borer Weevil, Vegetation Indices, classification methods.

Capítulo I

Problema

Antecedentes

En Ecuador la producción de banano y plátano es el primer rubro no petrolero en importancia económica. Existen alrededor de 268.410 hectáreas de plantaciones de las cuales el 95% se exporta y abastece a 43 mercados a nivel mundial, lo que representa aproximadamente el 30% del consumo mundial es de banano ecuatoriano (AGROCALIDAD, 2020). En el año 2019 se certificaron 51.114 envíos, siendo los principales destinos Rusia, Turquía e Irak; lo que representa 6'200.971,4 toneladas del producto (AGROCALIDAD, 2016).

En Ecuador el plátano es de alta importancia económica, debido a que la exportación bananera representa el 2% del PIB general y aproximadamente el 35% del PIB agrícola; según el registro del Ministerio de Agricultura y Ganadería (MAG), el Ecuador tiene alrededor de 162.236 hectáreas sembradas de banano, y cuenta con 4.473 productores de esta fruta (Ministerio de Comercio Exterior, 2017) distribuidos de la siguiente manera:

Figura 1

Catastro banano

DISTRIBUCIÓN POR TAMAÑO DE HECTÁREAS	HECTÁREAS SEMBRADAS	NÚMERO DE PRODUCTORES
0-30 (PEQUEÑOS)	35.685	3.480
>30 ≤ 100 (MEDIANOS)	57.486	800
100 O MÁS (GRANDES)	69.063	193
TOTAL	162.236	4.473

Nota. El gráfico representa la distribución por hectáreas sembradas de banano y el número de productores (Ministerio de Comercio Exterior, 2017).

Por la década de los años 60 las plagas que pusieron en alerta a los productores bananeros y que mayor importancia tenían, fueron los nematodos, el Picudo Negro, y la Sigatoka amarilla (INIAP, s.f.). La incidencia de una plaga destructiva para las musáceas como

la de *Fusarium oxysporum f. sp. Cubense* raza 1 (FOC R1), provocó la desaparición de miles de hectáreas de la variedad de banano Gros Michel (AGROCALIDAD, 2020); lo cual fue motivo para que Ecuador procediera a la siembra de la nueva variedad resistente del grupo Cavendish, en reemplazo de la variedad Gros Michel que había demostrado una alta susceptibilidad a la raza 1 (INIAP, s.f.).

El Picudo Negro (*Cosmopolites sordidus*) es el causante de la pérdida desde el 10% hasta 70% de la producción de banano, provocando pérdidas económicas en los pequeños agricultores, quienes dependen de este cultivo para su subsistencia (Espinosa Velepucha et al., 2019). Por otra parte, en Ecuador la enfermedad que más afecta a los cultivos de banano es la Sigatoka negra (*Mycosphaerella fijiensis* Morelet), debido a que su control es difícil y costoso (INIAP, s.f.).

La Agricultura de precisión busca mejorar la producción y optimizar los recursos, una de las principales ventajas que ofrece la agricultura de precisión es maximizar rendimientos en los sitios productivos generando sustentabilidad del medio ambiente, mediante la incorporación de distintas tecnologías en los procesos (Vite Cevallos et al., 2018). Los sensores remotos es una tecnología que permite la detección y evaluación de enfermedades de los cultivos, incluso para detección de *Fusarium*. La tecnología UAV se ha utilizado cada vez más para adquirir imágenes que permitan extraer información fenotípica de cultivos (Ye et al., 2020).

En los últimos años se han desarrollado estudios que involucran la agricultura de precisión como es el caso de (Ye et al., 2020), que indica el uso de imágenes multiespectrales basadas en vehículos aéreos no tripulado (UAV), para identificar zonas bananeras infestadas o no infestadas con la enfermedad del marchitamiento por *Fusarium*. Para ello se utilizó un UAV con una cámara multiespectral de cinco bandas; y se eligieron ocho índices de vegetación (IV) relacionados con la absorción de pigmentos, y los cambios en el crecimiento de las plantas, para determinar las características biofísicas y bioquímicas de las plantas; dando como

resultado que la enfermedad del marchitamiento de Fusarium en banano se puede identificar fácilmente utilizando índices de vegetación.

Planteamiento del problema

El banano se ve amenazado constantemente por diversos problemas fitosanitarios que traen como consecuencia pérdidas económicas y productivas. La marchitez por fusarium es una enfermedad del banano causada por especies patógenas de *Fusarium oxysporum*; esta enfermedad es la mayor amenaza para las áreas productoras de banano en todo el mundo. Si no se gestiona adecuadamente puede perjudicar seriamente a la industria bananera de todo un país (Shen et al., 2019).

El Picudo Negro del banano causa hasta un 42% de pérdida de la cosecha, además su presencia es un síntoma de alerta ya que estos insectos pueden trasladar el hongo Fusarium R4T de una planta a otra y de una finca a otra (AGROCALIDAD, 2021; Gold, Kagezi, Night, & Ragama, 2004). La Sigatoka Negra es considerada la plaga más destructiva y con mayor afectación económica en cultivos de banano y plátano; debido a que puede causar pérdidas de hasta 50% en el rendimiento y causar pérdidas del 100% por el deterioro de la calidad del fruto (Guzmán Quesada y Paladines García, s.f.).

Tradicionalmente la detección y discriminación de plagas y enfermedades en cultivos ha sido de forma manual, lo cual requiere mucho tiempo y dinero (Shi et al., 2018), el método tradicional de detección de fusarium y otras enfermedades, consiste en utilizar la inspección visual lo cual consume mucho tiempo y puede llegar a ser ineficiente cuando se trata de monitorear extensiones grandes de cultivo (Liu et al., 2020). El uso de este método no proporciona datos precisos sobre la distribución de las enfermedades lo cual conlleva al uso excesivo de pesticidas, generando pérdidas económicas para los agricultores. Por consiguiente el presente proyecto de investigación tiene como objetivo solucionar dicha problemática

mediante el desarrollo de una metodología que permita la pronta detección de Fusarium Raza 1, Sigatoka Negra y Picudo Negro, mediante el uso de tecnologías geoespaciales.

Justificación e Importancia

El banano es un cultivo comercial muy importante para la exportación y alimento básico en muchos países en desarrollo (FAO, 2016). Desde el punto de vista socioeconómico genera fuentes de trabajo y provee alimentos ricos en energía a la mayoría de la población campesina (INIAP, s.f.) .

Los mayores productores de banano del país son de pequeñas empresas; por tanto, la producción del banano en el país gira principalmente en el ámbito de la economía familiar y la Economía Popular y Solidaria (EPS), lo que le convierte en un sector que coadyuva a la generación de empleo y la reducción de pobreza (Ministerio de Comercio Exterior, 2017). La baja productividad de plátano registrada en el país es consecuencia de problemas bióticos (plagas y enfermedades), abióticos (sequía) y tecnológicos, pues de la superficie total sembrada, solo el 34%, reciben control de plagas, lo que indica que más del 60% no tiene acceso a la tecnología (INEC, 2011; INIAP, s.f.).

El presente proyecto de investigación estaría ligado con Plan Nacional de Desarrollo 2017-2021, en donde el objetivo 6 del segundo eje del plan menciona que se debe: “Desarrollar las capacidades productivas y del entorno, para lograr la soberanía alimentaria y el Buen Vivir Rural”, por lo cual es importante el desarrollo de la investigación y el uso de tecnologías que permitan garantizar la soberanía alimentaria. A partir de esto se desarrollara una metodología que permita la identificación de la pronta detección de Fusarium Raza 1, Sigatoka negra y Picudo negro en banano mediante el uso de tecnologías geoespaciales.

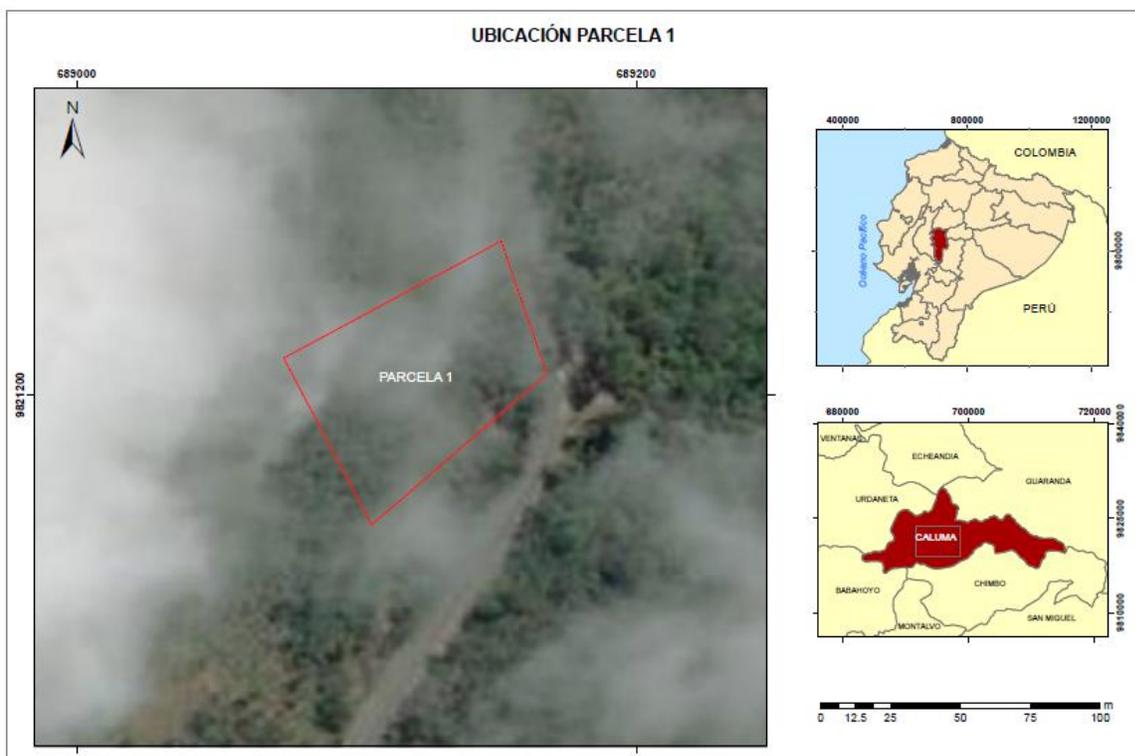
Descripción del área de estudio

El presente proyecto se realizó en 2 zonas de estudio: Parcela 1 con plantación de banano correspondiente a especie Gros Michel, ubicada en el cantón Caluma de la Provincia

de Bolívar (ver Figura 2); y parcela 2 con plantación de plátano correspondiente a la especie Hartón (*Musa paradisiaca*), ubicada en la parroquia rural Patricia Pilar del Cantón Buena Fe de la Provincia de los Ríos (ver Figura 3).

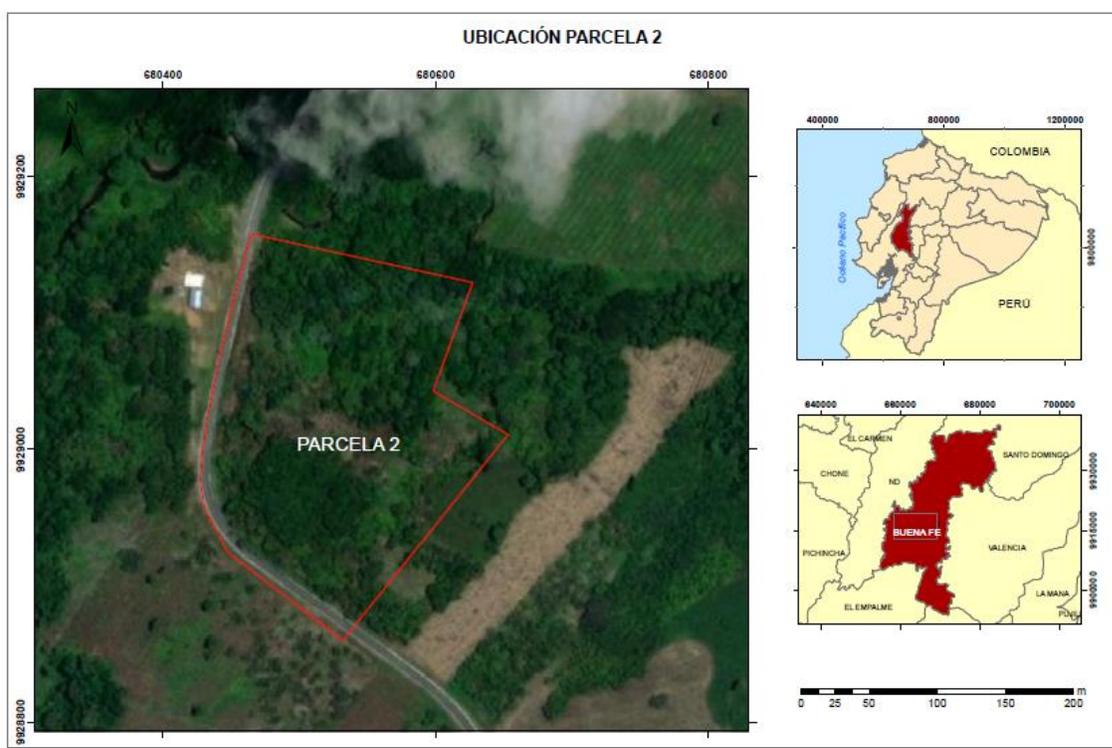
Figura 2

Ubicación Parcela 1



Nota. El gráfico representa la ubicación de la parcela 1

Figura 3

Ubicación Parcela 2

Nota. El gráfico representa la ubicación de la parcela 2.

Objetivos***Objetivo General***

Plantear una metodología para la detección de patologías propias del cultivo de banano mediante índices de vegetación y métodos de clasificación supervisada de imágenes multiespectrales de muy alta resolución tomadas con UAV

Objetivos Específicos

- Capturar imágenes multiespectrales mediante un UAV para obtener información detallada y generar ortomosaicos de las zonas de estudio.

- Obtener los valores de índices de vegetación a partir de los ortomosaicos multispectrales para analizar la afectación de Fusarium Raza 1, Sigatoka Negra y Picudo Negro en los bananos.
- Aplicar métodos de clasificación supervisada (machine learning) para analizar la afectación de Raza 1, Sigatoka negra y picudo negro en banano.
- Realizar análisis estadísticos mediante análisis de varianza ANOVA, matrices de confusión e índice kappa para determinar que índices de vegetación son los más adecuados para detección de Fusarium Raza 1, Sigatoka Negra y Picudo Negro

Metas

- 2 Ortomosaicos multiespectrales de la zona de estudio
- 1 Reporte de los índices de vegetación
- 2 Raster con la clasificación de plantas de banano enfermo y no enfermo.
- 1 Reporte estadístico
- 2 Planos de afectación de Fusarium Raza 1, Sigatoka Negra y Picudo Negro en banano

Capítulo II

Marco teórico

Teledetección

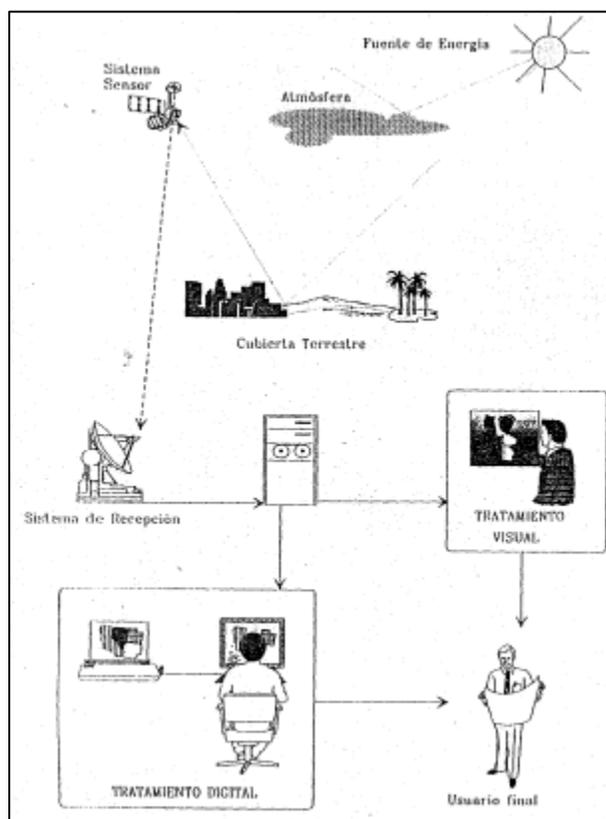
La teledetección espacial es una técnica que permite adquirir datos de la superficie terrestre en forma de imágenes, desde sensores montados sobre plataformas espaciales, suponiendo que entre la tierra y el sensor existe una interacción energética, ya sea por reflexión de energía o de un haz energético artificial (Chuvienco, 1990). El objetivo esencial de la teledetección se basa en la identificación de los materiales de la superficie terrestre y los fenómenos que en ella se operan a través de su signatura espectral (Sacristán Romero, 2006).

Un sistema de teledetección espacial incluye los siguientes elementos:

- Fuente de energía: Puede tratarse de un foco externo al sensor (teledetección pasiva), o un haz energético emitido por este (teledetección activa). Siendo la fuente más importante la energía solar.
- Cubierta terrestre: recibe la señal energética procedente de la fuente de energía, y es reflejada o emitida de acuerdo a sus características físicas.
- Sistema sensor: corresponde al sensor y la plataforma que lo alberga
- Sistema de recepción- comercialización: recibe la información transmitida por la plataforma, dicha información se graba en un formato apropiado.
- Intérprete: encargado de convertir los datos en información temática de interés, esta puede ser visual o digital.
- Usuario final: analiza el documento y es el encargado de dictaminar sobre las consecuencias que de él se deriven (Chuvienco, 1990).

Figura 4

Componentes de un sistema de teledetección



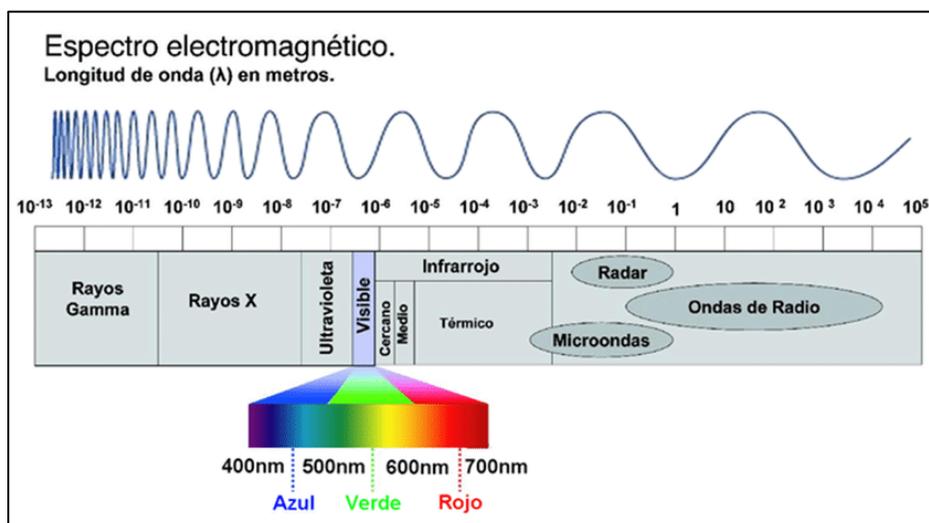
Nota. El gráfico representa los elementos que componen un sistema de teledetección, tomado de: (Chuvieco, 1990).

Espectro electromagnético

Constituido por un conjunto de ondas con diferentes características debido a su distinto comportamiento en el medio de propagación, la misma precedencia o la misma forma de interacción con la materia (Luque Ordóñez, 2012). El espectro electromagnético se divide en una serie de regiones, las más utilizadas en teledetección son la luz visible, el infrarrojo reflejado, el infrarrojo térmico y las microondas (Sarria, 2006).

Figura 5

Regiones del espectro electromagnético



Región	Rango de longitudes
Rayos γ	$<0,03nm$
Rayos X	$0,03nm - 3nm$
Ultravioleta	$3nm - 0,3\mu$
Visible	$0,3\mu - 0,7\mu$
Infrarrojo	$0,7\mu - 300\mu$
Microondas	$300\mu - 20cm$
Radio-TV	$> 200cm$

Nota. El gráfico representa las regiones que el espectro electromagnético abarca, cada una con sus longitudes de onda, tomado de: (Sarria, 2006).

Vehículos aéreos no tripulados (VANT – UAV)

Un vehículo aéreo no tripulado (UAV) del inglés *unmanned aerial vehicle*, también conocido como dron, hace referencia a un avión sin piloto humano a bordo. Su vuelo es controlado de forma autónoma por computadoras o bajo un control remoto de un piloto en tierra o en otro vehículo (Praveen, 2017).

En la actualidad los UAV se han presentado como una herramienta beneficiosa en la agricultura de precisión; las principales ventajas incluyen la capacidad de generar imágenes de muy alta resolución, la capacidad de incorporar una cámara multiespectral que permita la

detección de enfermedades en los cultivos (Gómez Candón et al., 2014). Los UAV se pueden clasificar en dos tipos que son ala fija y multirroto.

Agricultura de Precisión

El concepto de agricultura de precisión (AP) se estableció a fines de la década de 1970, cuando el sistema de posicionamiento global (GPS) fue desarrollado. La agricultura de precisión permite el manejo específico de un área de cultivo, utiliza herramientas tecnológicas como el posicionamiento global, sistema de información (SIG), sensores remotos, dispositivos de distribución de riego, sensores climatológicos de cultivo, procesamiento avanzado de información, y telecomunicaciones (Naiqian et al., 2002).

El uso de AP permite tener en cuenta la variabilidad espacial y temporal dentro del terreno de cultivo para tomar decisiones sobre el momento, la calidad y cantidad de insumos que se debe administrar, que permite mejorar la gestión de la empresa agropecuaria en aspectos económicos y ambientales (Gómez et al., 2016; Leiva, 2003).

Musáceas

Las musáceas son plantas monocotiledóneas de gran tamaño, dentro de las familias de las musáceas encontramos los bananos y plátanos. Los bananos y plátanos son cultivos perennes debido a que los vástagos año tras año siguen brotando de una única mata. Los brotes tienen un crecimiento enérgico y pueden producir un racimo maduro en menos de un año (FAO, 2016; Arias et al., 2004).

Fusarium

Fusarium oxysporum es un hongo que se presenta principalmente como saprófito en el suelo, o también puede presentarse como patógeno especializado, denominado de forma especial (f.sp.) (De Granda et al., 2001). Los aislamientos patógenos causan el marchitamiento por fusarium de varios cultivos agrícolas, según la planta hospedante se subdividen en formas especiales (Fourie et al., 2009).

La marchitez por *Fusarium* en banano es conocida como mal de Panamá, es considerada una de las enfermedades más destructivas y graves del banano en todo el mundo (Getha y Vikineswary, 2002) . Debido a la translocación de agua en los tejidos vasculares los primeros síntomas del marchitamiento incluyen la clorosis, necrosis y la muerte y caída de las hojas (Lin et al., 2009).

Figura 6

Síntomas del *Fusarium*



Nota. El gráfico representa los principales síntomas de *Fusarium* en banano, tomado de: (AGROCALIDAD, 2021)

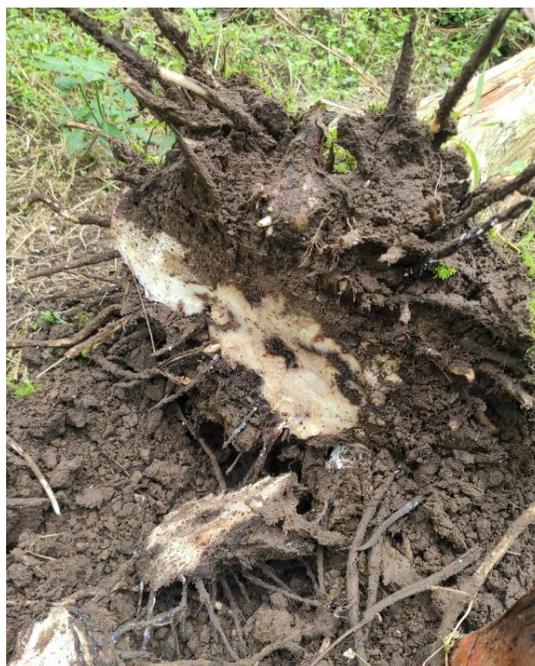
Picudo Negro

El Picudo Negro *Cosmopolites sordidus* en Ecuador es considerada una de las plagas principales en los cultivos de plátano y banano (Gold et al., 2005). Este insecto plaga destruye

el rizoma causando el volcamiento de plantas, en más de un 10%, y provocando debilitamiento fuerte en el resto de plantas (Espinosa et al., 2004). Las larvas se desarrollan dentro del corno formando túneles que perforan el corno interrumpiendo el transporte de nutrientes y agua a la planta. (Miranda et al., 2019)

Figura 7

Síntomas del Picudo Negro



Nota. El gráfico representa la perforación del corno, producido por el desarrollo de las larvas.

Sigatoka Negra

La Sigatoka Negra, es causada por el hongo Ascomycete *Mycosphaerella fijiensis*, es una enfermedad foliar que afecta la producción de musáceas (plátano y banano) a nivel mundial (Alvarez et al., 2013). La Sigatoka Negra como consecuencia de la afectación del tejido foliar, reduce la fotosíntesis y afecta el crecimiento de la planta (Guzmán Quesada y Paladines García, s.f.).

Figura 8**Síntomas de la Sigatoka Negra**

Nota. El gráfico representa la afectación de la Sigatoka Negra en banano

Índices de vegetación

Los índices de vegetación (IV) son combinaciones de bandas espectrales que realzan el contraste de la vegetación (que tiene una alta reflectancia) y atenúan detalles del suelo desnudo, estructuras construidas, etc. (Earth Observing System, 2022). Los IV permiten determinar el efecto sobre la biomasa de sequías, plagas, desmontes, incendios, granizo o déficit de nutrientes; los valores bajos de los índices de vegetación, usualmente indican vegetación poco vigorosa, mientras que los valores altos indican vegetación muy vigorosa (Piedad Rubio et al., 2020).

Índice Normalizado Diferencial de Vegetación (NDVI)

EL NDVI mide la biomasa fotosintéticamente activa de las plantas, es considerado uno de los más adecuados para el seguimiento del desarrollo de la vegetación. Su principal ventaja

es que puede ser utilizado durante toda la temporada de producción de cultivos a excepción de cuando la cubierta vegetal sea demasiado escasa (Earth Observing System, 2022). Se obtiene a partir de la siguiente formula:

$$NDVI = \frac{NIR - R}{NIR + R} \quad (1)$$

En donde:

R = Banda roja

NIR = Banda del infrarrojo cercano

Índice De Vegetación De Diferencia Normalizada Verde (GNDVI)

El GNDVI mide el contenido de clorofila con mayor precisión, es una modificación del NDVI debido a que sustituye el verde por el rojo visible (540 a 570 nm) (Earth Observing System, 2022). Se obtiene a partir de la siguiente formula:

$$GNDVI = \frac{NIR - GREEN}{NIR + GREEN} \quad (2)$$

En donde:

NIR = Banda del infrarrojo cercano

GREEN = Banda verde

Índice De Diferencia Normalizada De Borde Rojo (NDRE)

De acuerdo con (Earth Observing System, 2022) “El NDRE combina las bandas espectrales del infrarrojo cercano (NIR) y una banda específica para el estrecho rango entre el rojo visible y la zona de transición (la llamada región del borde rojo). Este índice de vegetación se aplica a la cubierta vegetal de alta densidad”. Se obtiene a partir de la siguiente formula:

$$NDRE = \frac{NIR - RED\ EDGE}{NIR + RED\ EDGE} \quad (3)$$

En donde:

NIR = Banda del infrarrojo cercano

RED EDGE = Borde rojo

Índice de clorofila verde (CI_{green}).

Es usado para calcular el contenido total de clorofila en las hojas, se fundamenta en un modelo de regresión lineal, se obtiene a partir de la siguiente formula: (Gitelson et al., 2005).

$$CI_{green} = \left(\frac{NIR}{G} \right) - 1 \quad (4)$$

En donde:

NIR = Banda del infrarrojo cercano

G = Banda del verde

Índice de clorofila del borde rojo (CI_{re})

Permite obtener una estimación del contenido de clorofila de las hojas a partir de la tasa de reflectividad en las bandas NIR y de límite rojo. Se obtiene a partir de la siguiente formula (Esri, s.f.).

$$CI_{RE} = \left(\frac{NIR}{RED\ EDGE} \right) - 1 \quad (5)$$

En donde:

NIR = Banda del infrarrojo cercano

RED EDGE = Banda del borde rojo

Aprendizaje Automático

Es parte de la inteligencia artificial, una forma de aprendizaje automático que involucra la identificación de patrones, mediante un análisis de datos que permiten la creación de algoritmos o modelos para generar predicciones de datos (Calderón Romero y Hurtado Cortes,

2019). Existe una variedad de algoritmos que pueden mejorar el aprendizaje utilizando nuevos datos de entrada. Estos se clasifican en 2 grupos: aprendizaje supervisado y aprendizaje no supervisado (Shalev-Shwartz y Ben-David, 2014)

Random Forest (RF)

De acuerdo con (Breiman, 2001) “El modelo bosques aleatorios son una combinación de predictores de árboles de modo que cada árbol depende de los valores de un vector aleatorio muestreado de forma independiente y con la misma distribución para todos los árboles del bosque”; se define con la ecuación:

$$\{h(x, \theta_k), k = 1, \dots\}$$

(6)

En donde:

x = Variable de entrada

h = Clasificador

$\{\theta_k\}$ = Variables predictoras aleatorias distribuidas (Breiman, 2001)

Support Vector Machine (SVM)

La máquina de vectores de soporte (SVM) es un método de clasificación supervisada. SVM pertenece a la familia de los clasificadores lineales puesto que inducen separadores lineales o hiperplanos en espacios de características de muy alta dimensionalidad, a pesar de que se pueden adaptar de forma sencilla para actuar como clasificadores no lineales mediante la aplicación de una función o kernel no lineal sobre los datos de entrada (Rojas, 2009,p. 51).

Estadística

Es una ciencia que adquiere,organiza, y analiza datos o informacion numerica ,facilitando la toma de decisiones (Lind et al., 2012). la necesidad de analizar grandes conjuntos de datos permitieron el desarrollo de programas estandarizados que contiene una

gama amplia de herramientas estadísticas, que en la actualidad son aplicados en diferentes áreas de estudio. (Berenson et al., 2006).

Análisis de la Varianza (ANOVA)

Se denomina análisis de la varianza ANOVA a la comparación de tres o más medias poblacionales, con el fin de determinar si estas pueden ser iguales. (Lind et al., 2012).

Actualmente es el método estadístico más utilizado para las pruebas de hipótesis

Si se desea emplear ANOVA, se deben cumplir las siguientes hipótesis:

- Las poblaciones deben ser independientes
- Las poblaciones siguen una distribución normal
- Las poblaciones tienen desviaciones estándares iguales (homocedasticidad) (Lind et al., 2012).

Prueba de Shapiro Wilks

Los procedimientos estadísticos dependen de la normalidad de la población, por lo cual es importante recurrir a una prueba de normalidad. El test de Shapiro Wilks es utilizado en muestras menores o iguales a 50 observaciones, una vez calculada la media y la varianza y se rechaza la hipótesis nula de normalidad si el estadístico es menor al p-valor crítico. (Flores Tapia y Flores Cevallos, 2021)

Prueba de Levene

La condición de homogeneidad de varianzas, indica que las varianzas de la variable dependiente que se comparen deben ser iguales. Para comprobar la homogeneidad de varianzas se utiliza la prueba de Levene en la cual el p-valor debe arrojar una significancia mayor de 0,05 para que se cumpla la hipótesis nula. (Rubio Hurtado y Berlanga Silvente, 2012)

Matriz de confusión

La matriz de confusión es un indicador que permite visualizar mediante una tabla de doble entrada la calidad de los datos de un clasificador, La diagonal de la matriz nos indica la coincidencia por categoría entre la cantidad de píxeles reales y los de la clasificación, mientras que los datos restantes a la diagonal muestran aquellos que se confunden con otras categorías (Borrás et al., 2017).

Figura 9

Matriz de confusión

		Predicción	
		Positivos	Negativos
Observación	Positivos	Verdaderos Positivos (VP)	Falsos Negativos (FN)
	Negativos	Falsos Positivos (FP)	Verdaderos Negativos (VN)

Nota. El gráfico representa la matriz de confusión y los valores a considerar en esta herramienta, tomado de: (Porta, 2006)

En la matriz se tiene los siguientes valores a considerar:

- **VP** (Verdaderos positivos): instancias correctamente reconocidas por el sistema.
- **VN** (Verdaderos negativos): instancias que son negativas y correctamente reconocidas por el sistema.
- **FN** (Falsos negativos): instancias que son positivas y que el sistema los clasifica como negativos.
- **FP** (Falsos positivos): instancias que son negativas pero el sistema los clasifica como positivos (Porta, 2006).

Las ecuaciones que se pueden utilizar para validar la clasificación son las que se muestran a continuación. (Porta, 2006)

Exactitud.

$$\text{Exactitud} = \frac{VP + VN}{Total} \quad (7)$$

En donde:

VP = Verdaderos positivos

VN= Verdaderos negativos

Tasa de error.

$$\text{Tasa de error} = \frac{FP + FN}{Total} \quad (8)$$

En donde:

FP = Falsos positivos

FN= falsos negativos

Sensibilidad.

$$\text{Sensibilidad} = \frac{VP}{Total\ positivos(VP + FN)} \quad (9)$$

En donde:

VP = Verdad positivos

FN= Falsos negativos

Especificidad.

$$\text{Especificidad} = \frac{VN}{Total\ negativos(VN + FP)} \quad (10)$$

En donde:

FP = Falsos positivos

VN= Verdad negativos

Precisión.

$$Precisión = \frac{VP}{Total\ clasificados\ positivos(VP + FP)}$$

(11)

En donde:

VP = Verdad positivos

FP= Falsos positivos

Índice kappa

El índice kappa es utilizado para evaluar la coincidencia de métodos categóricos que tienen dos o más clases, este estadístico muestra la proporción de acuerdos observados con respecto del máximo de acuerdos posibles más allá del azar (Castañeda et al., 2018).

El índice Kappa se calcula a partir de la siguiente ecuación adaptada de (Abraira, 2001):

$$K = \frac{\left(\frac{VP + VN}{N}\right) - \left(\frac{TP \times TCP + TN \times TCN}{N^2}\right)}{1 - \left(\frac{TP \times TCP + TN \times TCN}{N^2}\right)}$$

(12)

En donde:

VP = Falsos positivos

VN= Falsos negativos

TN=Total negativos

TP= Total positivos

TCP= Total de clasificados positivos

TCN= Total de clasificados negativos

N= Total de valores.

La escala de valoración del índice Kappa puede ser obtenida a partir de la Tabla 1

Tabla 1

Valoración del Índice Kappa

Valor Kappa	Grado de acuerdo
<0.00	Sin acuerdo
0.00-0.20	Insignificante
0.21-0.40	Mediano
0.41-0.60	Moderado
0.61-0.80	Sustancial
0.81-1.00	Casi perfecto

Nota: Esta tabla representa la escala de valoración del índice kappa, adaptada de: (Landis y Koch, 1977).

Base Legal

La base legal del presente proyecto se encuentra respaldada por la Constitución de la República del Ecuador, en el capítulo tercero de la Soberanía Alimentaria, donde el Art.281 menciona que la que “La soberanía alimentaria constituye un objetivo estratégico y una obligación del Estado para garantizar que las personas, comunidades, pueblos y nacionalidades alcancen la autosuficiencia de alimentos sanos y culturalmente apropiado de forma permanente”. Por otra parte menciona que es responsabilidad del estado “Asegurar el desarrollo de la investigación científica y de la innovación tecnológica apropiadas para garantizar la soberanía alimentaria” (CONSTITUCIÓN DE LA REPUBLICA DEL ECUADOR [Const.], 2008).

De la misma manera la Ley Orgánica del Régimen de Soberanía Alimentaria, en el capítulo III, llamado Investigación, Asistencia Técnica y Diálogo de Saberes, donde el Art. 9, menciona que el “Estado asegurará y desarrollará la investigación científica y tecnológica en materia agroalimentaria, con el principal objetivo de mejorar la calidad nutricional de los

alimentos, así como proporcionar una asistencia técnica, sustentada en un diálogo e intercambio de saberes con los pequeños y medianos productores” (LORSA, 2009).

Capítulo III

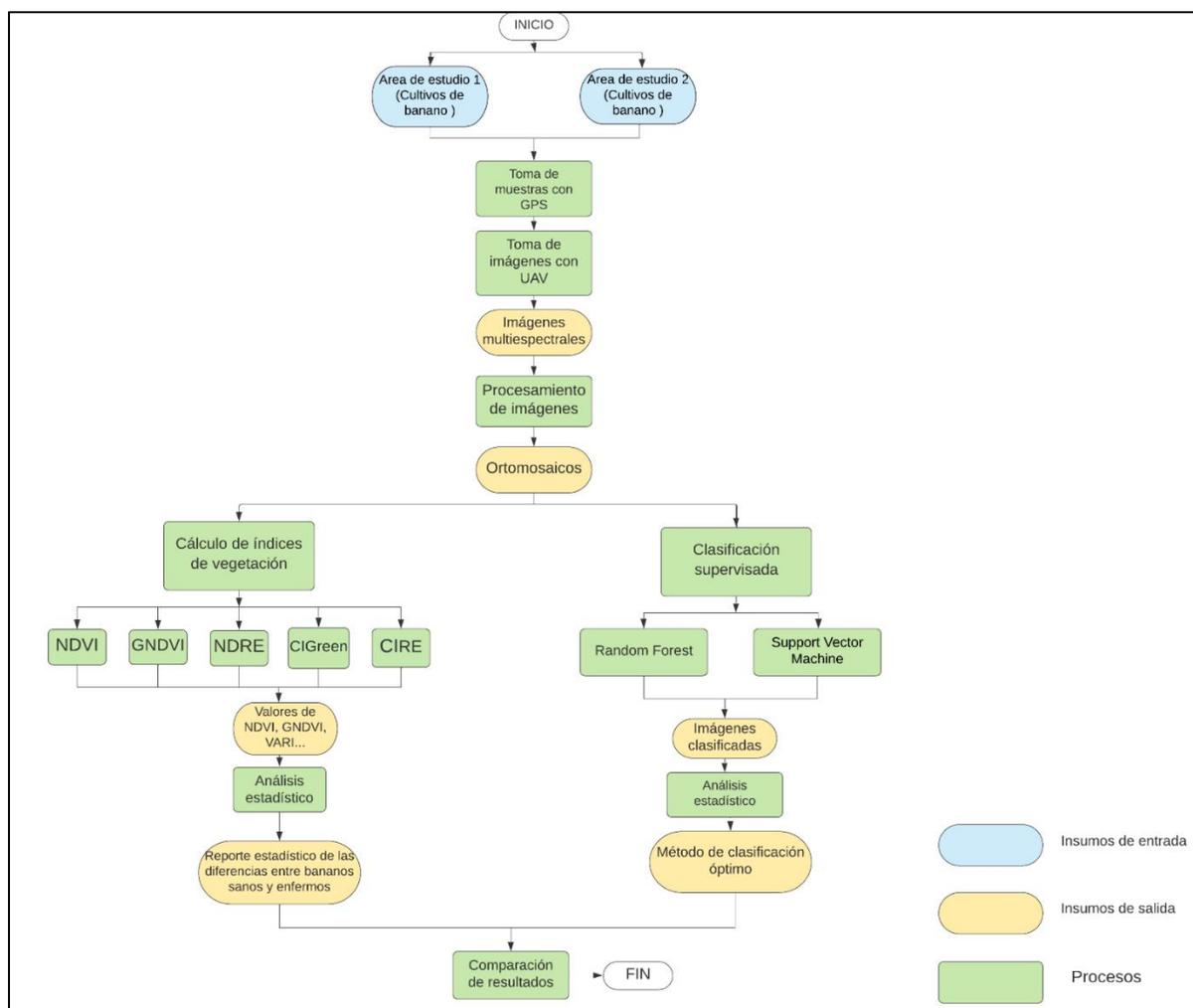
Metodología

Generalidades

El proyecto se desarrolló en 2 áreas de estudio: Parcela 1 ubicada en el cantón Caluma, y la finca 2 ubicada la parroquia rural del Cantón Buena Fe de la provincia de los Ríos, la toma de datos se realizó en los primeros meses del 2022, aplicando la misma metodología en las 2 zonas de estudio, la cual tendrá 3 fases (ver Figura 10)

Figura 10

Metodología de desarrollo del proyecto



Nota. El gráfico representa la metodología aplica en el proyecto de investigación.

En la fase uno se realizó el reconocimiento de las zonas de estudio para la toma de muestras con GPS de los bananos con presencia o ausencia de enfermedad (FOC R1, Sigatoka Negra o Picudo Negro), estos datos se depuraron y se generó una base de datos de cada zona de estudio. Conjuntamente, con el muestreo en campo se realizaron planes de vuelo para obtener imágenes multiespectrales, mediante dron Phantom IV y cámara multiespectral Parrot Sequoia.

En la fase dos las imágenes multiespectrales obtenidas fueron procesadas y corregidas con un software fotogramétrico para obtener ortomosaicos, con estos insumos se procedió a utilizar dos tipos de metodología, la primera fue el cálculo de índices de vegetación y la segunda fue la aplicación de algoritmos de machine learning para realizar clasificación supervisada.

Los índices de vegetación y los métodos de clasificación supervisada utilizados en este estudio fueron los siguientes:

Índices de vegetación

Se aplicaron 5 índices de vegetación a los ortomosaicos obtenidos de las 2 zonas de estudio, los cuales fueron:

- Índice Normalizado Diferencial de Vegetación (NDVI)
- Índice De Vegetación De Diferencia Normalizada Verde (GNDVI)
- Índice del borde rojo de diferencia normalizada (NDRE)
- Índice de clorofila verde (CI_{green})
- Índice de clorofila del borde rojo (CI_{RE})

Métodos de clasificación de imágenes

Se aplicaron 2 métodos de clasificación de imágenes para conocer que método genera mejor resultado en cuanto a la identificación de FOC R1, Sigatoka Negra, y Picudo Negro, los cuales fueron:

- Algoritmo Random Forest (RF)
- Support Vector Machine (SVM)

En la fase 3, con los valores de los índices de vegetación obtenidos, se realizó un análisis de varianza ANOVA, mientras que a los métodos de clasificación se les aplicó la matriz de confusión e índice kappa. Finalmente, se determinó que metodología permite alcanzar mejores resultados respecto a la identificación de FOC R1, Sigatoka Negra, y Picudo Negro en banano.

Fase I

Materiales y Equipos

Para el desarrollo del proyecto de investigación el principal equipo utilizado para la obtención de información fue una cámara multispectral Parrot Sequoia montada en el dron Phantom IV (ver Figura 11), también se utilizaron equipos y distintos tipos de software que facilitaron el procesamiento y obtención de resultados (ver Tabla 2 y 3).

Figura 11

Camara Parrot Sequioa montada en Dron Phantom 4



Nota. El gráfico representa la cámara y el dron utilizados en el presente proyecto de investigación.

Tabla 2*Equipos*

Nº	Equipos	Descripción
1	UAV	1 DJI Phantom 4
2	Camara multiespectral	1 camara modelo Parrot Sequoia
3	Tarjeta de calibración	1 Mica Sense Calibrates Reflectance Panel
4	GPS navegador	Garmin Etrex
5	Computador	Laptop Toshiba
6	Celulares	Iphone 7- Xiaomi Mi 9 Se

Nota. Esta tabla muestra los equipos utilizados en el presente proyecto de investigación.

Tabla 3*Tipos de Software*

Nº	Software	Descripción
1	Pix4D	Software para la planificación de vuelo
2	DJI Ground Station Pro	Software para la planificación de vuelo
3	Agisoft Metashape	Software para el procesamiento fotogramétrico
4	ArcGIS - ArcMap	Software comercial SIG
5	ENVI	Software para el análisis geoespacial
6	InfoStat	Software para el análisis estadístico

Nota. Esta tabla muestra los tipos de software utilizados en el presente proyecto de investigación.

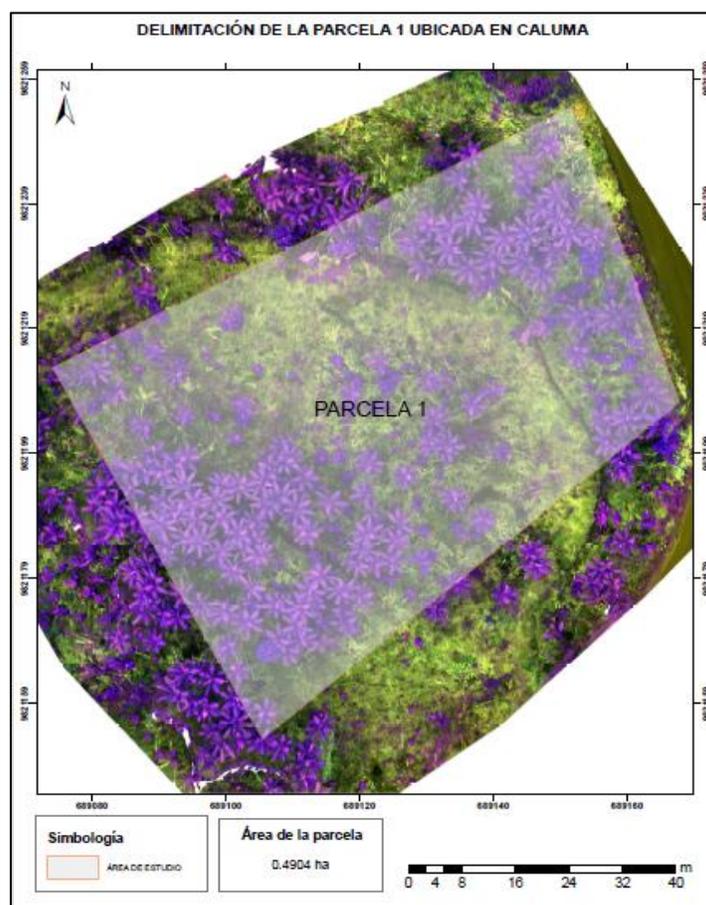
Delimitación del área de estudio

Las zonas de estudios involucradas en el proyecto fueron 2: parcela 1 con plantación de banano la cual se encuentra infectada con Fusarium Raza 1, ubicada en el cantón Caluma de

la Provincia de Bolívar, con una superficie de 0.4904 hectáreas; y parcela 2 con plantación de plátano correspondiente a la especie Hartón (*Musa paradisiaca*), infectada con Sigatoka Negra y Picudo Negro, ubicada en la parroquia rural del Cantón Buena Fe de la provincia de los Ríos, con una superficie de 4.2918 hectáreas (Ver Figura 12 y 13).

Figura 12

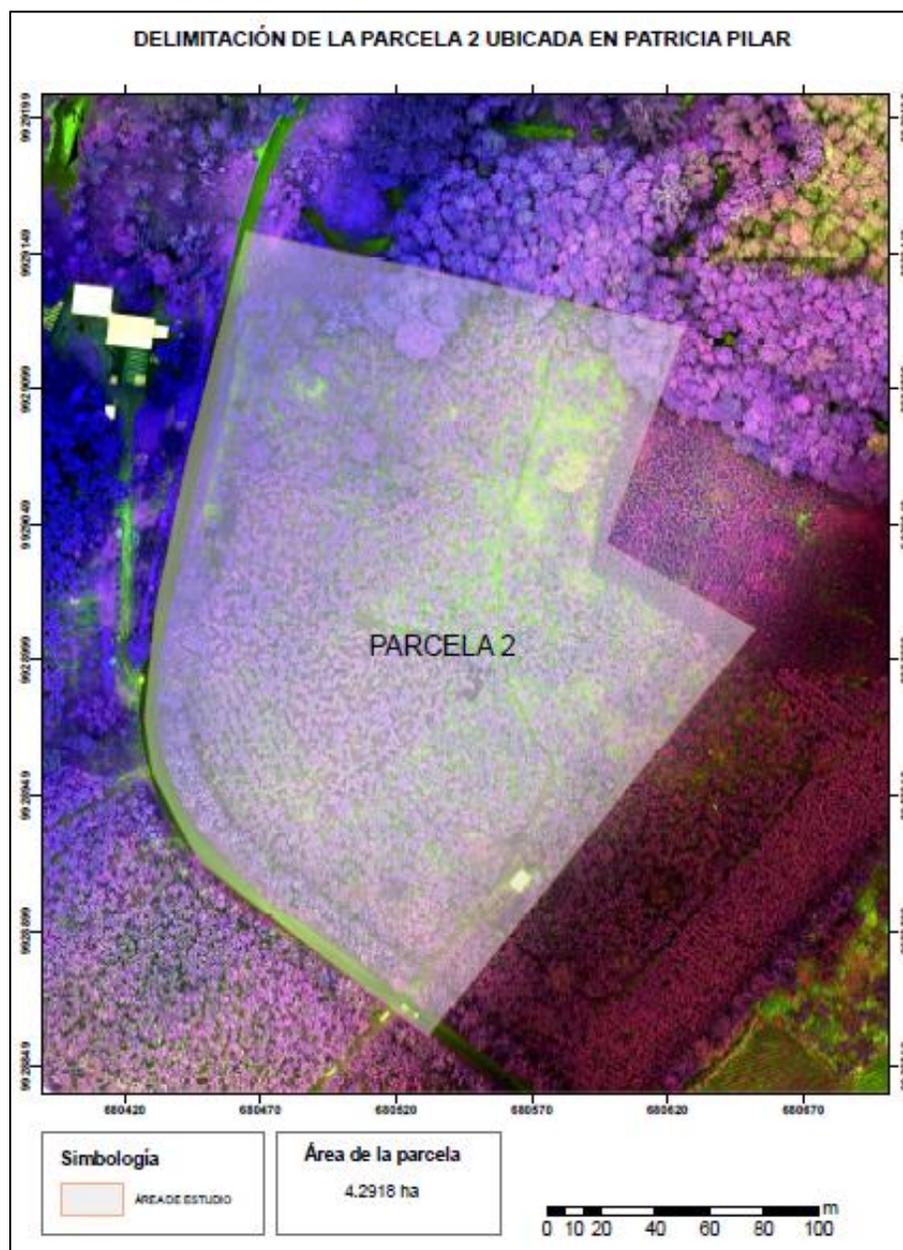
Delimitación de la parcela 1 ubicada en Caluma



Nota: El gráfico representa el área de la parcela 1 involucradas en el proyecto de investigación.

Figura 13

Delimitación de la parcela 2 ubicada en Patricia Pilar



Nota: El gráfico representa las parcelas involucradas en el proyecto de investigación.

Toma de muestras con GPS

Se identificaron y georreferenciaron con un GPS navegador las plantas que presentan sintomatología de Fusarium Raza 1 en la parcela ubicada en Caluma, y plantas con

sintomatología de Sigatoka Negra y Picudo Negro en la parcela ubicada en Patricia Pilar, con el fin de obtener un shp de puntos (ver Figura 14,15 y 16).

Figura 14

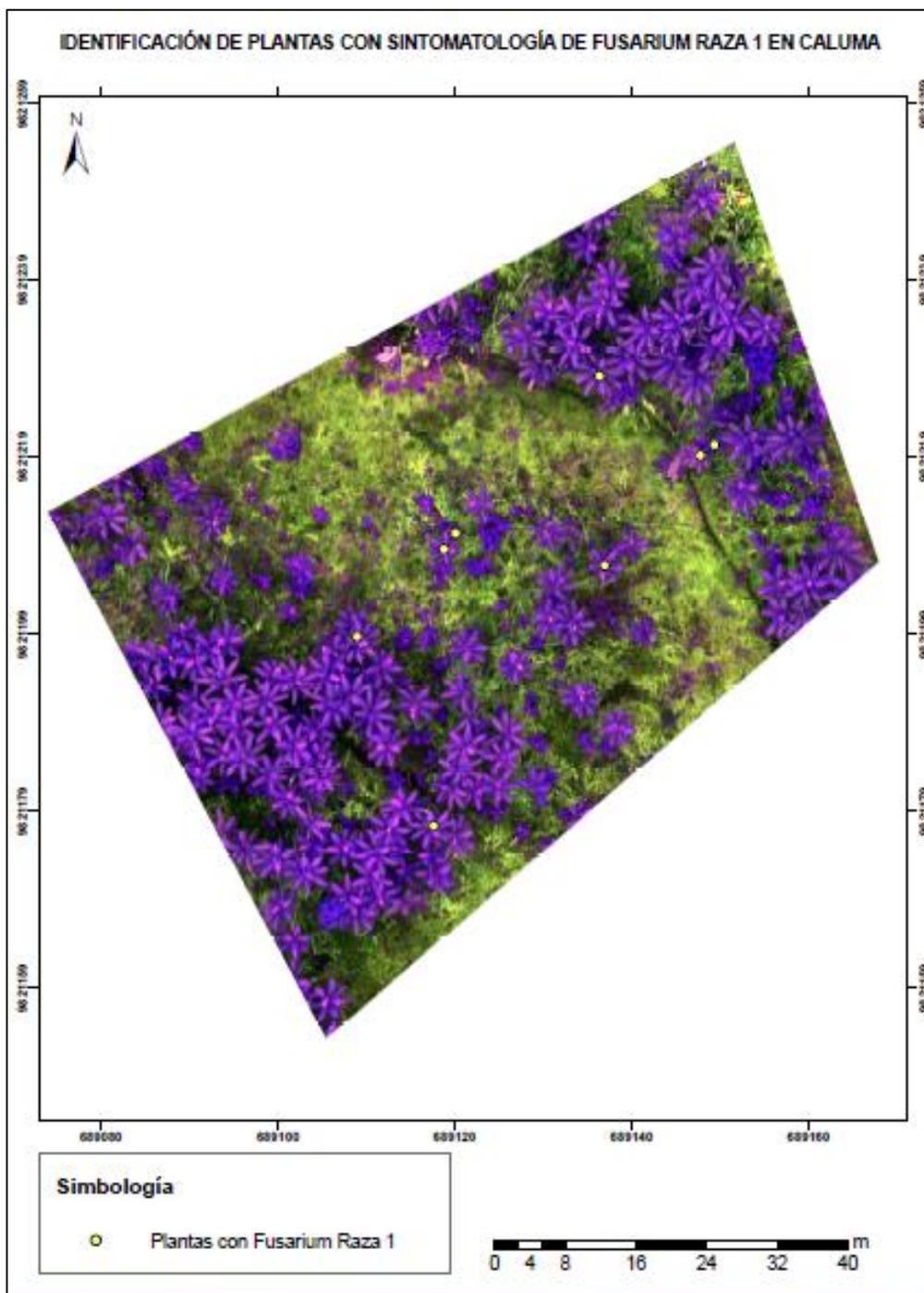
Toma de muestras con GPS navegador



Nota: El gráfico representa la toma de muestras de plantas con sintomatología.

Figura 15

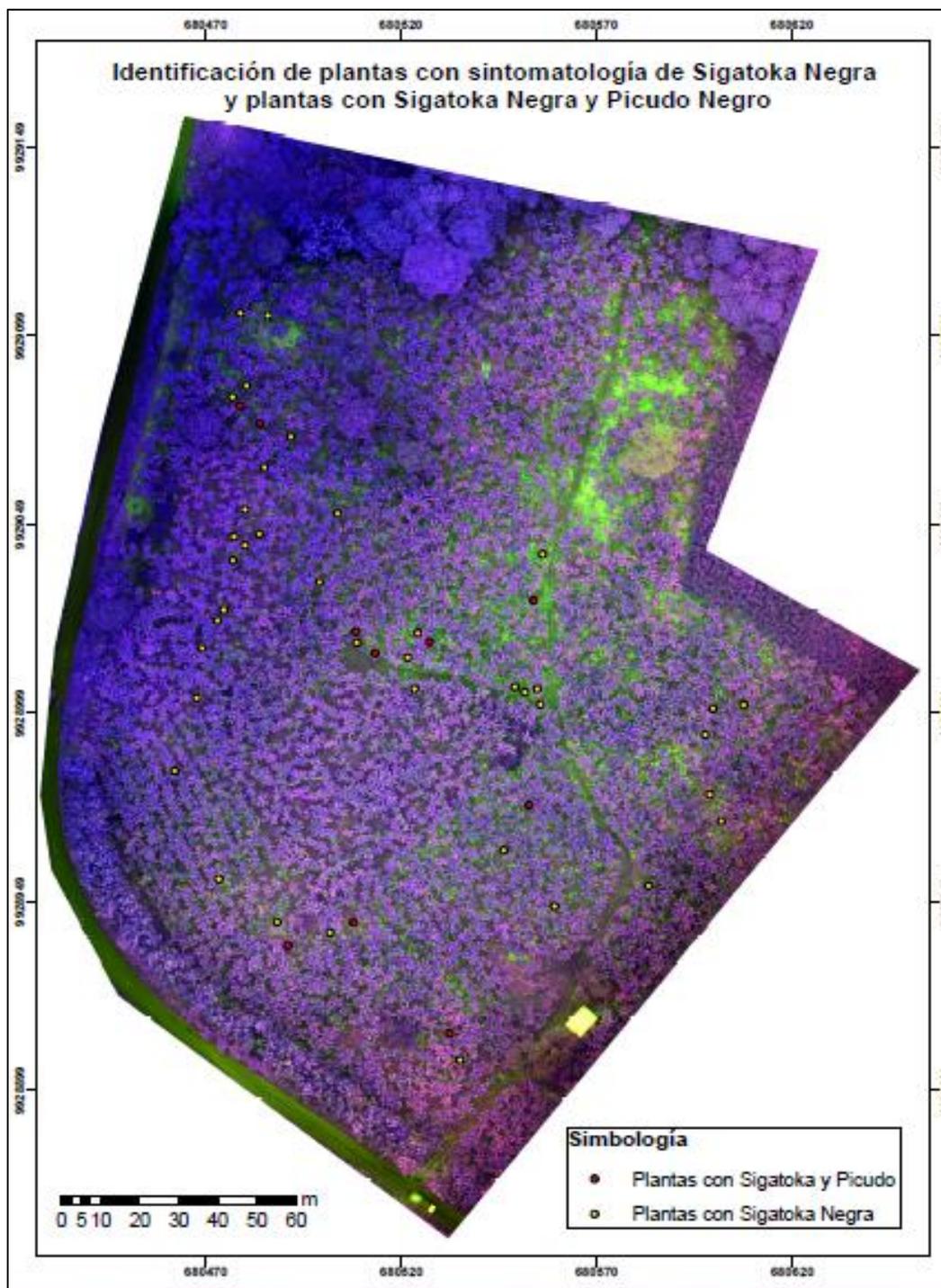
Toma de muestras con GPS navegador en la parcela 1



Nota: El gráfico representa la toma de muestras de plantas con sintomatología de FOC R1 en la parcela 1.

Figura 16

Toma de muestras con GPS navegador en la parcela 2



Nota: El gráfico representa la toma de muestras de plantas con sintomatología en la parcela 2.

Planificación de los vuelos

Para obtener las imágenes multiespectrales, mediante el uso de la cámara Parrot Sequoia montada en el dron Phantom IV, se realizaron diferentes planes de vuelo de acuerdo a cada área de estudio.

Los planes de vuelo se ejecutaron en la aplicación Pix4D para la parcela 1 y la aplicación DJI Ground Station Pro para la parcela 2. En ambas aplicaciones se ingresaron parámetros como: el traslape longitudinal, el traslape transversal, altura de vuelo y área del terreno, Adicionalmente, las aplicaciones calculan automáticamente parámetros como: tiempo de vuelo cantidad de líneas, GSD y velocidad (ver Tabla 4).

Tabla 4

Parámetros de los planes de vuelo

Parámetros	Parcela 1	Parcela 2
Área del terreno	0.4904	4.2918
Altura de vuelos	80 m	80 m
Traslape Longitudinal	90%	80%
Traslape Transversal	90%	80%
Tiempo de Vuelo	3 m 35 s	15 m 52 s
Líneas de vuelo	5	13
GSD	3.50 cm/ pixel	3.44cm/pixel

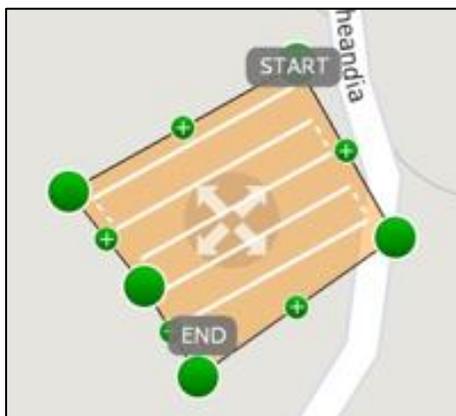
Nota: La tabla indica los parámetros ingresados y generados en las aplicaciones para planificación de los vuelos.

Ejecución de los vuelos

Una vez configurados los vuelos en cada aplicación con los parámetros mencionados en la tabla anterior, se realizaron los dos vuelos como se indican en la Figura 17 y 18.

Figura 17

Diseño del vuelo en la parcela 1



Nota: Esta figura representa los vuelos realizados en la parcela ubicada en Caluma.

Figura 18

Diseño del vuelo en la parcela 2



Nota: Esta figura representa los vuelos realizados en la parcela ubicada en Patricia Pilar

Proceso Fotogramétrico para obtención de imágenes multiespectrales

En el proceso fotogramétrico de las imágenes multiespectrales se utilizó el software Agisoft Methasape con la plantilla sistema multicamara para cada parcela.

Primero se ingresaron las imágenes multiespectrales y las imágenes de la calibración de reflectancia; se realizó la calibración radiométrica ingresando los valores de reflectancia correspondientes a cada banda que se encuentran en el certificado del panel. Se ejecutó la

orientación interior y exterior de las imágenes, teniendo como resultado una nube de puntos. Posteriormente, se creó una nube de puntos densa que permite obtener un ortomosaico de mejor calidad. Finalmente, se creó un modelo digital de elevaciones (DEM) que fue el insumo para la generación de los ortomosaicos multispectrales.

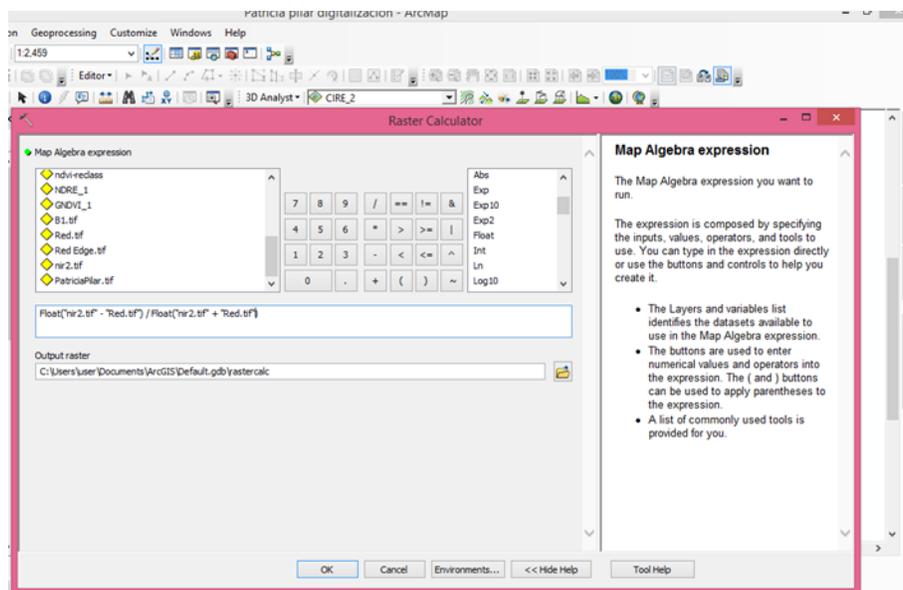
Fase II

Obtención de los índices de vegetación

En el software ArcMap se cargaron los ortomosaicos multispectrales, y mediante la herramienta “Raster Calculator” se calcularon los índices de vegetación (NDVI, GNDVI, NDRE, CI_{Green} , y CIRE) en función de las ecuaciones 1, 2, 3, 4 y 5 (ver Figura 19). Para extraer los valores de los píxeles de cada planta con sintomatología, se utilizó la herramienta “Zonal Statistics as table”.

Figura 19

Herramienta “Raster Calculator”



Nota: Esta figura representa la herramienta utilizada para el cálculo de los 5 índices de vegetación.

Métodos de clasificación supervisada

En el software ENVI 5.3 se utilizó la herramienta “Layer Stack” para compilar las 4 bandas multiespectrales (Green, Red, Red Edge y Nir) dando como resultado una sola imagen de cada zona de estudio, que fue utilizada para la clasificación supervisada.

Support Vector Machine (SVM)

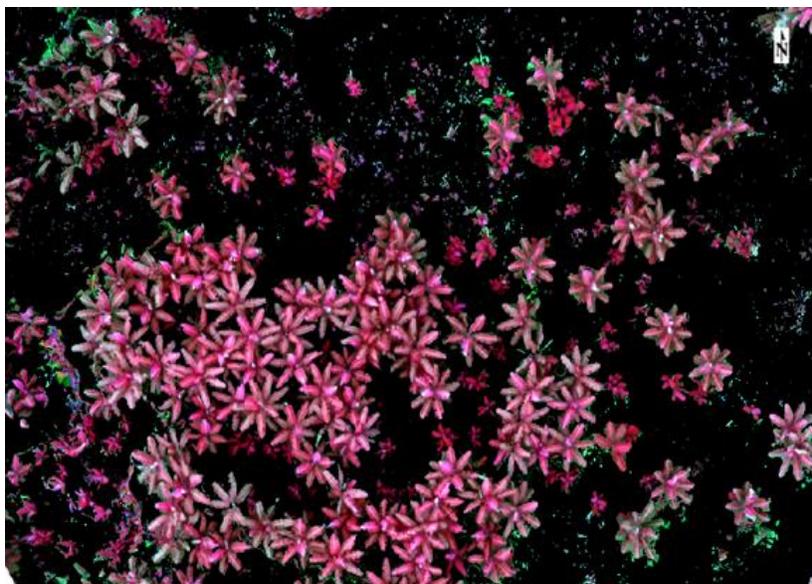
Para la clasificación de máquina de vectores de soporte (SVM), mediante la herramienta ROI se generaron las muestras de plantas con y sin sintomatología que fueron identificadas en campo, una vez obtenidas las muestras, con la herramienta “Support Vector Machine Classification” se realizó la clasificación supervisada en la cual se ingresaron la imagen, las muestras y los parámetros de clasificación.

Random Forest (RF)

Para la clasificación de Random Forest se utilizó la imagen compilada por la herramienta Layer Stack, y se aplicó una máscara de suelo, para ayudar a que la clasificación sea más óptima, debido a que el objetivo del estudio está enfocado en clasificar solo las plantas, el mismo proceso se realizó en las dos parcelas (ver Figura 20 y 21).

Figura 20

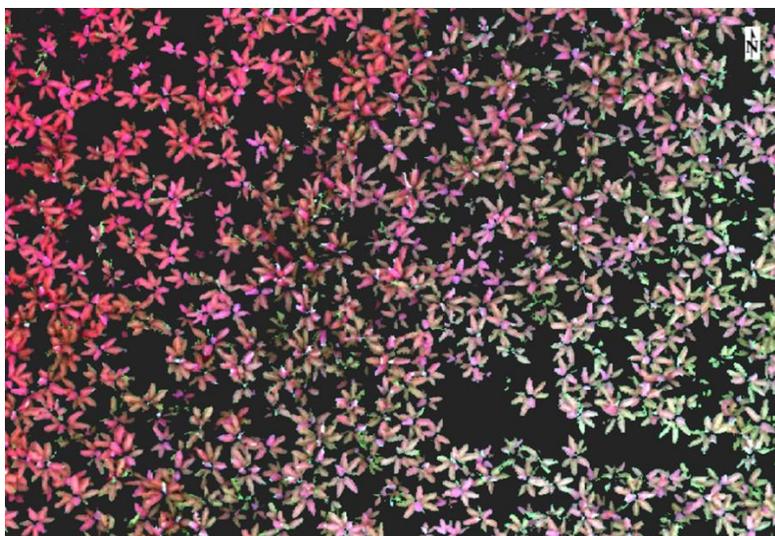
Máscara de suelo en la parcela 1



Nota: Esta figura representa la máscara de suelo que fue aplicada a la parcela ubicada en Caluma.

Figura 21

Máscara de suelo en la parcela 2



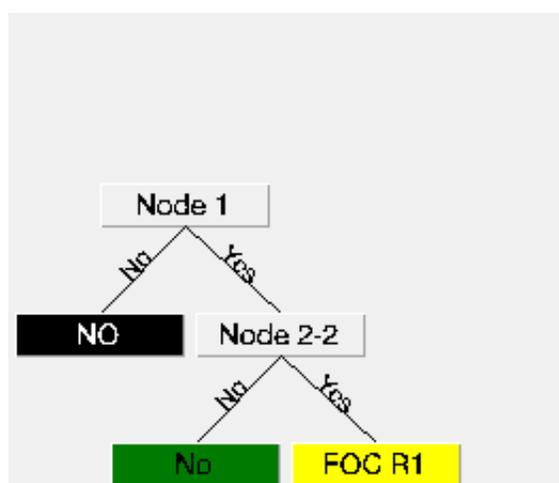
Nota: Esta figura representa la máscara de suelo que fue aplicada a la parcela ubicada en Patricia Pilar.

Debido a que en el suelo se encontraban hojas de banano y vegetación, al realizar la máscara de suelo esta no se eliminó por completo, por lo cual se generó un árbol de decisión de dos nodos partiendo como nodo principal la clasificación de planta de banano y con única respuesta de si o no, una vez clasificadas las plantas, el segundo nodo se basa en determinar si las plantas presentan la sintomatología de FOC R1 (ver Figura 22).

Con la herramienta Crosshairs se visualizó el valor de los píxeles de las manchas que presentan las hojas con sintomatología, con estos valores se estableció un rango que fue utilizado en el nodo.

Figura 22

Árbol de decisión aplicado a la parcela 1



Nota: Esta figura representa los nodos aplicados en el árbol de decisión.

Fase III

Análisis estadístico

El análisis de varianza (ANOVA) tiene como requisito cumplir tres supuestos, que permiten validar los resultados obtenidos en el análisis estadístico. Las pruebas que se mencionan a continuación se encuentran con un nivel de significancia de 0.05 y un nivel de confianza del 95%, lo cual establece que se rechaza la hipótesis nula si el p-valor es menor o igual al nivel de significancia, caso contrario no se rechaza la hipótesis nula.

Prueba de Normalidad. *Se aplicó la prueba de Shapiro Wilks para poblaciones menores o iguales a 50, con el fin de conocer si los datos tienen una distribución normal; esta prueba se basa en las siguientes hipótesis:*

H₀: Los datos se ajustan a una distribución normal

H₁: Los datos no se ajustan a una distribución normal.

Prueba de Homogeneidad. *Se aplicó la prueba de Levene para conocer si existe homogeneidad entre las varianzas, esta prueba se basa en las siguientes hipótesis:*

H₀: Las varianzas presentan homogeneidad

H₁: Las varianzas no presentan homogeneidad.

Análisis de la varianza Anova.

Se comprobó que los datos obtenidos de la parcela 1, ubicada en Caluma, siguen una distribución normal y sus varianzas son homogéneas, por lo cual se aplicó la prueba ANOVA, teniendo en cuenta las siguientes hipótesis:

H₀: Los índices de vegetación no permiten diferenciar plantas de banano con FOC R1 y plantas de banano sin FOC R1

H₁: Los índices de vegetación permiten diferenciar plantas de banano con FOC R1 y plantas de banano sin FOC R1

Prueba de Kruskal-Wallis. Se comprobó que los datos obtenidos de la parcela 2, ubicada en Patricia Pilar, no cumplen con la prueba de normalidad por lo cual no se pudo aplicar la prueba ANOVA y se utilizó la prueba no paramétrica de Kruskal Wallis, teniendo en cuenta las siguientes hipótesis:

H₀: No existe diferencia significativa entre plantas de bananos con Sigatoka Negra y bananos con Picudo Negro y Sigatoka Negra.

H₁: Existe diferencia significativa entre plantas de bananos con Sigatoka Negra y bananos con Picudo Negro y Sigatoka Negra.

Matriz de confusión

Para analizar los resultados obtenidos en la clasificación supervisada, y determinar que algoritmo de clasificación es el más adecuado, se realizó la matriz de confusión para cada método utilizado y para cada parcela.

Se realizó un conteo para relacionar las predicciones realizadas por el algoritmo de clasificación y las instancias en cada clase real, las instancias utilizadas en la parcela 1 y parcela 2 se muestran a continuación:

- VP (parcela 1): cantidad de plantas con FOC R1 que fueron identificadas correctamente.
- VN (parcela 1): cantidad de plantas sanas que fueron clasificadas correctamente
- FP (parcela 1): cantidad de plantas sanas que fueron clasificadas como plantas con FOC R1.
- FN (parcela 1): cantidad de plantas con FOC R1 que fueron clasificadas como plantas sanas.
- VP (parcela 2): cantidad de plantas con Sigatoka Negra que fueron identificadas correctamente.
- VN (parcela 2): cantidad de plantas con Sigatoka Negra y Picudo Negro que fueron clasificadas correctamente
- FP (parcela 2): cantidad de plantas con Sigatoka Negra y Picudo Negro que fueron clasificadas como plantas solo con Sigatoka Negra.
- FN (parcela 2): cantidad de plantas con Sigatoka Negra que fueron clasificadas como Sigatoka Negra y Picudo Negro.

A partir de los resultados obtenidos en cada matriz se procedió a calcular las métricas establecidas en las ecuaciones 7, 8, 9,10 y 11.

Índice kappa

Utilizando la ecuación 12 a cada algoritmo de clasificación se le aplico el índice kappa, para determinar que algoritmo es el más adecuado para la identificación de plantas que presentan la sintomatología.

Capítulo IV

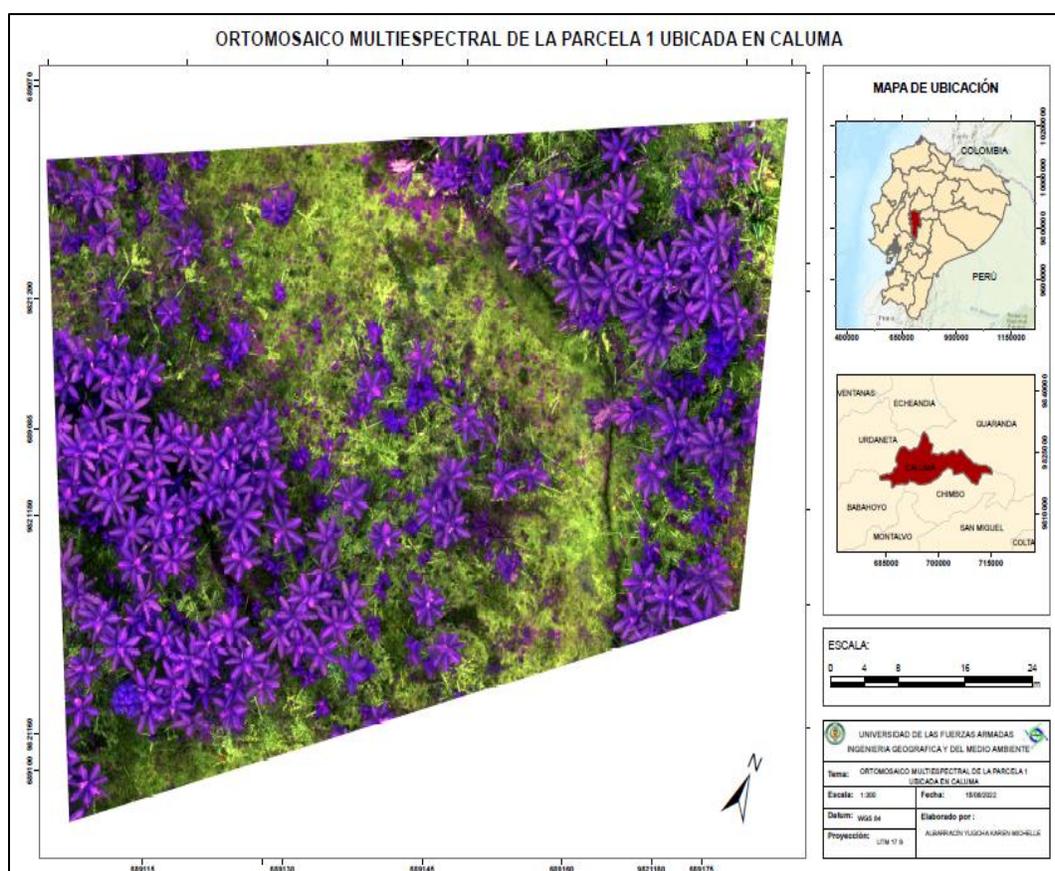
Análisis de los resultados

Generación de ortomosaicos multiespectrales

Se generó 1 ortomosaicos multiespectral de cada zona de estudio, que fueron insumos para realizar la clasificación supervisada (ver figura 23 y 24).

Figura 23

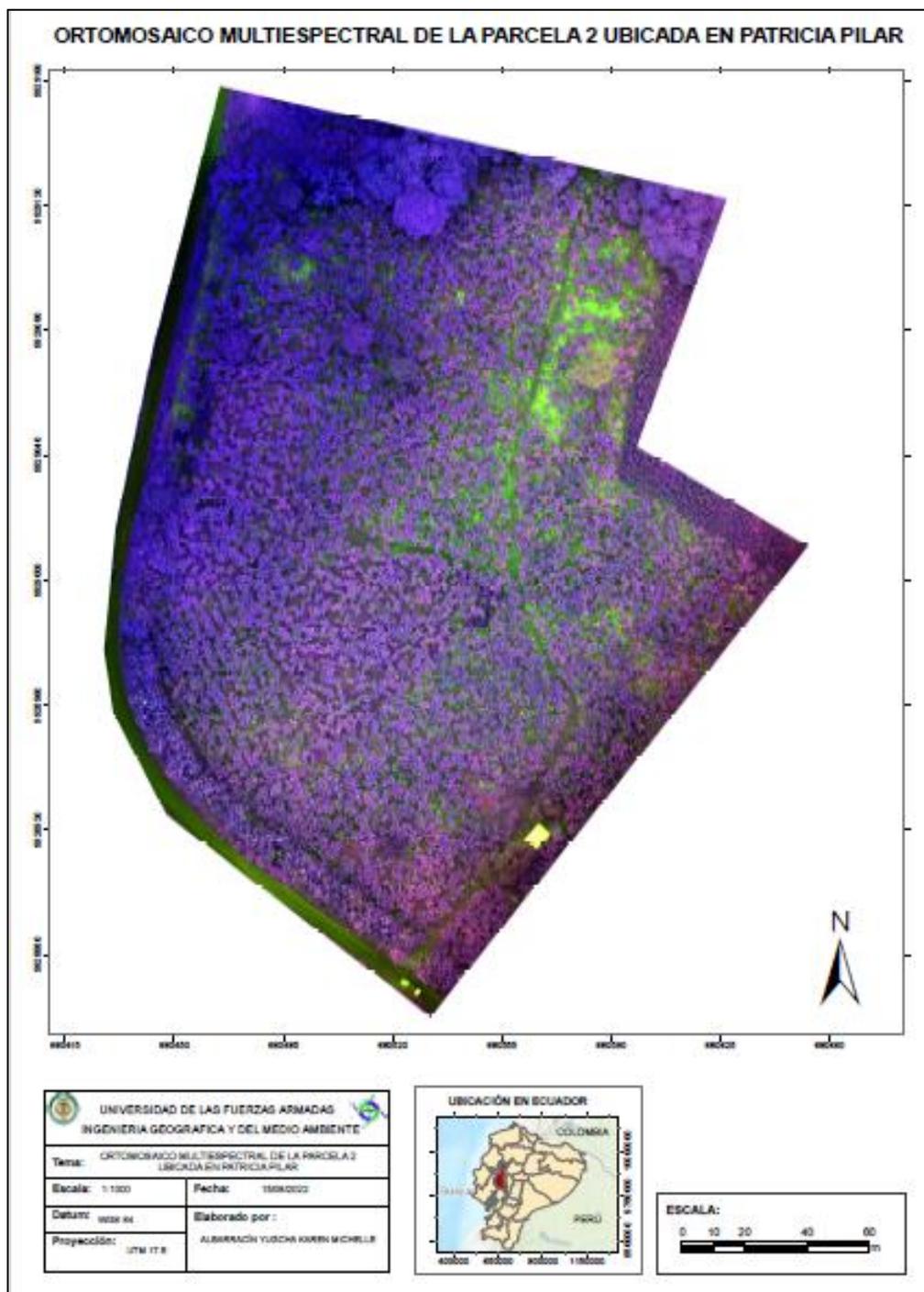
Ortomosaico multiespectral de la parcela 1



Nota: El gráfico representa el ortomosaico multiespectral de la parcela 1, ubicada en Caluma.

Figura 24

Ortomosaico multiespectral de la parcela 2



Nota: El gráfico representa el ortomosaico multiespectral de la parcela ubicada en Patricia Pilar.

Generación de índices de vegetación

A partir de las bandas multiespectrales, se calcularon los cinco índices de vegetación (NDVI, GNDVI, NDRE, Cl_{green}, CIRE) para cada zona de estudio, se generó una base de datos con los índices de vegetación calculados de acuerdo a las plantas muestreadas (ver Figura 25 y 26).

Figura 25

Índices de vegetación calculados de la parcela 1

ÍNDICES DE VEGETACIÓN						
FID	Estado	NDVI	GNDVI	NDRE	Cl _{green}	CIRE
0	Con FOC R1	0.8711	0.7125	0.1978	5.1486	0.4972
1	Con FOC R1	0.8504	0.6866	0.2017	4.5201	0.5156
2	Con FOC R1	0.8570	0.6887	0.1952	4.5195	0.4895
3	Con FOC R1	0.8541	0.6826	0.1895	4.4099	0.4751
4	Con FOC R1	0.8239	0.6376	0.1579	3.6111	0.3783
5	Con FOC R1	0.8627	0.7073	0.2353	4.9892	0.6225
6	Con FOC R1	0.8812	0.6936	0.1686	4.6546	0.4142
7	Con FOC R1	0.8393	0.6813	0.1982	4.3684	0.4975
8	Sin FOC R1	0.8676	0.7069	0.1836	4.8852	0.4519
9	Sin FOC R1	0.8357	0.6743	0.1907	4.1886	0.4732
10	Sin FOC R1	0.8581	0.6871	0.2017	4.4933	0.5079
11	Sin FOC R1	0.8805	0.7014	0.1820	4.8303	0.4493
12	Sin FOC R1	0.8580	0.6927	0.1915	4.6436	0.4793
13	Sin FOC R1	0.8649	0.6926	0.1864	4.6232	0.4603
14	Sin FOC R1	0.8602	0.6939	0.1914	4.6049	0.4745
15	Sin FOC R1	0.8843	0.6972	0.1754	4.7467	0.4282
16	Sin FOC R1	0.8512	0.6611	0.1599	3.9661	0.3857
17	Sin FOC R1	0.8866	0.7063	0.1720	4.9899	0.4177
18	Sin FOC R1	0.8638	0.6907	0.1768	4.5778	0.4317
19	Sin FOC R1	0.8491	0.6881	0.2024	4.5085	0.5122
20	Sin FOC R1	0.8472	0.6827	0.1998	4.4100	0.5035
21	Sin FOC R1	0.8992	0.7044	0.1658	4.9916	0.4039
22	Sin FOC R1	0.8626	0.6777	0.1777	4.2872	0.4338
23	Sin FOC R1	0.8728	0.7081	0.2023	4.9597	0.5095
24	Sin FOC R1	0.8611	0.6891	0.1769	4.5270	0.4318

Nota: El gráfico representa los índices de vegetación calculados de la parcela ubicada en Caluma.

Figura 26

Índices de vegetación calculados de la parcela 2

ÍNDICES DE VEGETACIÓN						
FID	Enfermedad	NDVI	GNDVI	NDRE	CIRE	Cigreen
0	sigatoka	0.3660	0.0657	-0.0037	-0.0059	0.1453
1	Ambas	0.2039	-0.0390	-0.0144	-0.0269	-0.0692
2	sigatoka	0.4812	0.2293	0.0441	0.0969	0.6216
3	sigatoka	0.4218	0.0633	0.0029	0.0112	0.1416
4	sigatoka	0.4199	0.0588	-0.0172	-0.0317	0.1369
5	sigatoka	0.2176	-0.0043	0.0429	0.0914	0.0041
6	Ambas	0.2131	-0.0415	0.0247	0.0524	-0.0739
7	sigatoka	0.2505	-0.0234	0.0466	0.0992	-0.0387
8	sigatoka	0.2622	-0.0147	0.0701	0.1528	-0.0154
9	sigatoka	0.2728	0.0571	0.0713	0.1556	0.1352
10	sigatoka	0.4781	0.2169	0.0322	0.0684	0.5777
11	sigatoka	0.5410	0.2482	0.0287	0.0615	0.6824
12	Ambas	0.4699	0.1083	0.0025	0.0078	0.2487
13	sigatoka	0.4054	0.1038	-0.0032	-0.0036	0.2393
14	sigatoka	0.4521	0.1396	0.0029	0.0099	0.3375
15	sigatoka	0.4499	0.1367	-0.0056	-0.0092	0.3319
16	sigatoka	0.4192	0.0852	-0.0069	-0.0123	0.1915
17	sigatoka	0.4144	0.0987	0.0009	0.0032	0.2352
18	sigatoka	0.4431	0.1454	0.0098	0.0212	0.3545
19	sigatoka	0.3721	0.1390	0.0269	0.0578	0.3337
20	sigatoka	0.3427	0.0498	0.0099	0.0209	0.1122
21	Ambas	0.2409	-0.0138	0.0033	0.0081	-0.0219
22	sigatoka	0.2669	-0.0115	-0.0069	-0.0123	-0.0161
23	sigatoka	0.3309	0.0675	-0.0049	-0.0060	0.1531
24	sigatoka	0.4227	0.1477	0.0115	0.0257	0.3596
25	sigatoka	0.4120	0.1239	0.0101	0.0230	0.2981
26	sigatoka	0.2657	0.0084	0.0414	0.0873	0.0282
27	Ambas	0.1186	-0.0561	0.0159	0.0349	-0.0973
28	sigatoka	0.2094	-0.0173	0.0244	0.0540	-0.0245
29	sigatoka	0.3057	-0.0246	0.1112	0.2528	-0.0403
30	Ambas	0.2531	0.0448	0.0558	0.1194	0.1040
31	sigatoka	0.2315	-0.0328	0.0053	0.0117	-0.0597
32	sigatoka	0.2231	-0.0087	0.0686	0.1491	-0.0087
33	sigatoka	0.1521	-0.1121	-0.0124	-0.0233	-0.1985
34	sigatoka	0.1635	-0.0582	0.0102	0.0222	-0.1032
35	sigatoka	0.1454	-0.0218	0.0342	0.0728	-0.0302
36	sigatoka	0.1747	0.0016	0.0316	0.0674	0.0160
37	sigatoka	0.1065	-0.1156	-0.0119	-0.0222	-0.2025
38	Ambas	0.1611	-0.0202	0.0257	0.0550	-0.0262
39	sigatoka	0.2209	-0.0111	0.0020	0.0050	-0.0102
40	Ambas	0.4595	0.1431	0.0093	0.0206	0.3437
41	sigatoka	0.2181	-0.0173	-0.0043	-0.0076	-0.0273
42	sigatoka	0.2157	-0.0340	-0.0143	-0.0271	-0.0623
43	Ambas	0.2093	0.0069	0.0030	0.0079	0.0248
44	sigatoka	0.3853	0.0704	0.1308	0.3028	0.1635
45	Ambas	0.2932	0.0046	0.0838	0.1847	0.0185
46	sigatoka	0.3202	0.1068	0.1479	0.3513	0.2593
47	sigatoka	0.2428	-0.0198	0.0570	0.1222	-0.0355
48	sigatoka	0.3379	0.0546	0.0156	0.0336	0.1281

Nota: El gráfico representa los índices de vegetación calculados de la parcela 1, ubicada en Patricia Pilar.

Análisis estadístico de los índices de vegetación calculados

Análisis exploratorio de datos

En la Tablas 5 y 6 se observa el análisis exploratorio aplicado a los 5 índices de vegetación calculados en cada zona de estudio.

Tabla 5

Análisis exploratorio parcela 1

Índice	Media	Mínimo	Máximo	Desviación Estándar	Varianza
NDVI	0.8617	0.8239	0.8991	0.0169	0.000
GNDVI	0.6897	0.6376	0.7125	0.0162	0.000
NDRE	0.1872	0.1579	0.2353	0.0169	0.000
CIRE	4.5782	3.6111	5.1486	0.3445	0.119
Cl _{green}	0.4658	0.3783	0.6225	0.0522	0.003

Nota: La tabla indica el análisis exploratorio al cual fueron sometidos los índices de vegetación de la parcela ubicada en Caluma.

Tabla 6

Análisis exploratorio parcela 2

Índice	Media	Mínimo	Máximo	Desviación Estándar	Varianza
NDVI	0.3058	0.1065	0.5410	0.1131	0.013
GNDVI	0.0414	-0.1156	0.2482	0.0840	0.007
NDRE	0.0247	-0.0172	0.1479	0.0372	0.001
CIRE	0.0558	-0.0317	0.3513	0.0836	0.007
Cl _{green}	0.1136	-0.2025	0.6824	0.2003	0.040

Nota: La tabla indica el análisis exploratorio al cual fueron sometidos los índices de vegetación de la parcela 2.

Al realizar el análisis exploratorio en la parcela 1 el índice que mayor desviación estándar representa es el de CIRE, mientras que en la parcela 2 el que mayor desviación

estándar presenta es el de CI_{green} , es importante tener en cuenta que mientras grande sea el valor de la desviación estándar, existe una mayor dispersión de datos con respecto a la media.

En la tabla 7 y 8 se muestran los resultados obtenidos en la prueba de normalidad de Shapiro Wilks para cada zona de estudio.

Tabla 7

Prueba de Shapiro Wilks de la parcela 1

Índice	Media	D.E.	W*	p-valor
NDVI	0	0.02	0.96	0.7661
GNDVI	0	0.02	0.93	0.2200
NDRE	0	0.02	0.98	0.9532
CIRE	0	0.34	0.96	0.6918
CI_{green}	0	0.05	0.98	0.9718

Nota: La tabla indica los resultados obtenidos en la prueba de Shapiro Wilks de la parcela 1, ubicada en Caluma.

Al aplicar la prueba de Shapiro Wilks a los índices calculados en la plantas de la parcela 1, se evidencia que los índices de vegetación presentan una distribución normal, siendo todos los p-valor mayores a 0.05.

Tabla 8

Prueba de Shapiro Wilks de la parcela 2

Índice	Media	D.E.	W*	p-valor
NDVI	0.31	0.11	0.94	0.110
GNDVI	0.04	0.08	0.95	0.118
NDRE	0.02	0.04	0.85	<0.0001
CIRE	0.06	0.08	0.82	<0.0001
CI_{green}	0.11	0.2	0.92	0.010

Nota: La tabla indica los resultados obtenidos en la prueba de Shapiro Wilks de la parcela 2.

Al aplicar la prueba de Shapiro Wilks a los índices calculados en la plantas de la parcela 2, se evidencia que los índices de vegetación NDVI y GNDVI presentan una distribución normal, debido a que su p-valor es de 0.110 y 0.118 respectivamente; mientras que los índices de vegetación NDRE, CIRE y Clgreen no presentan una distribución normal, siendo su p-valor <0.0001, <0.0001 y 0.010 respectivamente.

Por lo tanto solo los índices de vegetación de NDVI y GNDVI presentan una distribución normal, ya que su p-valor es mayor a 0,05.

Los resultados obtenidos en la prueba de Levene para cada zona de estudio se muestran en la Tabla 9 y 10.

Tabla 9

Prueba de Levene de la parcela 1

Índice	SC	gl	CM	F	p-valor
NDVI	6.10E-06	1	6.10E-06	0.06	0.816
GNDVI	1.30E-04	1	1.30E-04	0.98	0.332
NDRE	1.40E-04	1	1.40E-04	1.31	0.263
Cire	0.04	1	0.04	0.69	0.415
Cl green	1.40E-03	1	1.40E-03	1.26	0.272

Nota: La tabla indica los resultados obtenidos en la prueba de Levene de la parcela 1, ubicada en Caluma.

Al aplicar la prueba de Levene a los cinco índices calculados en la plantas de la parcela 1, se evidencia que todos los índices de vegetación presentan homogeneidad, debido a que todos los p-valor son mayores a 0.05.

Tabla 10

Prueba de Levene de la parcela 2

Índices	SC	gl	CM	F	P-VALOR
NDVI	6.60E-04	1	6.60E-04	0.2	0.656

Índices	SC	gl	CM	F	P-VALOR
GNDVI	3.30E-03	1	3.30E-03	1.47	0.231
NDRE	5.40E-04	1	5.40E-04	0.93	0.341
CIRE	3.00E-03	1	3.00E-03	0.93	0.341
Cl _{green}	0.02	1	0.02	1.49	0.229

Nota: La tabla indica los resultados obtenidos en la prueba de Levene de la parcela 2, ubicada en Patricia Pilar.

Al aplicar la prueba de Levene a los índices calculados en la plantas de la parcela 2, se evidencia que los índices de vegetación presentan homogeneidad, debido a que todos los p-valor son mayores a 0.05.

Debido a que los datos de la parcela 2 ubicada en Patricia Pilar no cumplen con el supuesto de normalidad, fueron sometidos a la prueba de Kruskal-Wallis, para determinar si existe diferencia significativa entre plantas de bananos con Sigatoka Negra y bananos con Picudo Negro y Sigatoka Negra, los resultados obtenidos en esta prueba se muestran en la Tabla 11.

Tabla 11

Prueba de Kruskal-Wallis de la parcela 2

Índice	SC	gl	CM	F	p-valor
NDVI	6.60E-04	1	6.60E-04	0.2	0.1433
GNDVI	3.30E-03	1	3.30E-03	1.47	0.2148
NDRE	5.40E-04	1	5.40E-04	0.93	0.9802
CIRE	3.00E-03	1	3.00E-03	0.93	0.9407
Cl _{green}	0.02	1	0.02	1.49	0.2148

Nota: La tabla indica los resultados obtenidos en la prueba de Kruskal-Wallis de la parcela 2, ubicada en Patricia Pilar.

Al aplicar la prueba de Kruskal-Wallis a los índices calculados en las plantas de la parcela 2, se evidencia que todos los p-valores son mayores a 0.05. Por lo tanto no existe diferencia significativa entre plantas de bananos con Sigatoka Negra y bananos con Picudo y Sigatoka Negra.

Debido a que los datos de la parcela 1 ubicada en Caluma si cumplen con el supuesto de normalidad y homogeneidad, se aplicó la prueba ANOVA, para determinar si los índices de vegetación permiten diferenciar plantas de banano con FOC R1 y plantas de banano sin FOC R1, los resultados obtenidos en esta prueba se muestran en la tabla 12.

Tabla 12

Prueba de ANOVA de la parcela 1

Índice	SC	gl	CM	F	p-valor
NDVI	5.40E-04	1	5.40E-04	1.95	0.1758
NDRE	4.00E-04	1	4.00E-04	1.42	0.2462
Cl _{green}	0.03	1	0.03	0.25	0.6252
CIRE	4.90E-03	1	4.90E-03	1.88	0.1841

Nota: La tabla indica los resultados obtenidos en la prueba ANOVA de la parcela 1, ubicada en Caluma

Al aplicar la prueba ANOVA a los índices calculados en las plantas de la parcela 1, se evidencia que los p- valores son mayores a 0.05. Por lo tanto los índices de vegetación no permiten diferenciar plantas de banano con FOC R1 y plantas de banano sin FOC R1.

Generación de Clasificación supervisada

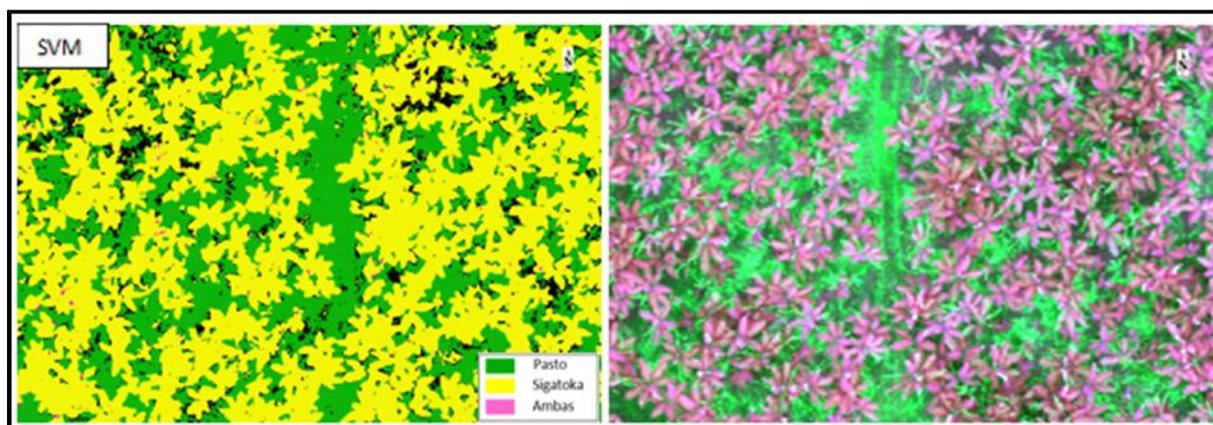
Support Vector Machine

Se aplicó la clasificación supervisada a las dos parcelas, sin embargo el algoritmo de Support Vector Machine solo dio resultados en la parcela 2 (ver Figura 27), debido a que en la

imagen de la parcela 1 clasificó toda la imagen con FOC R1 incluyendo el suelo y plantas sanas.

Figura 27

Resultado de la clasificación de Support Vector Machine parcela 2



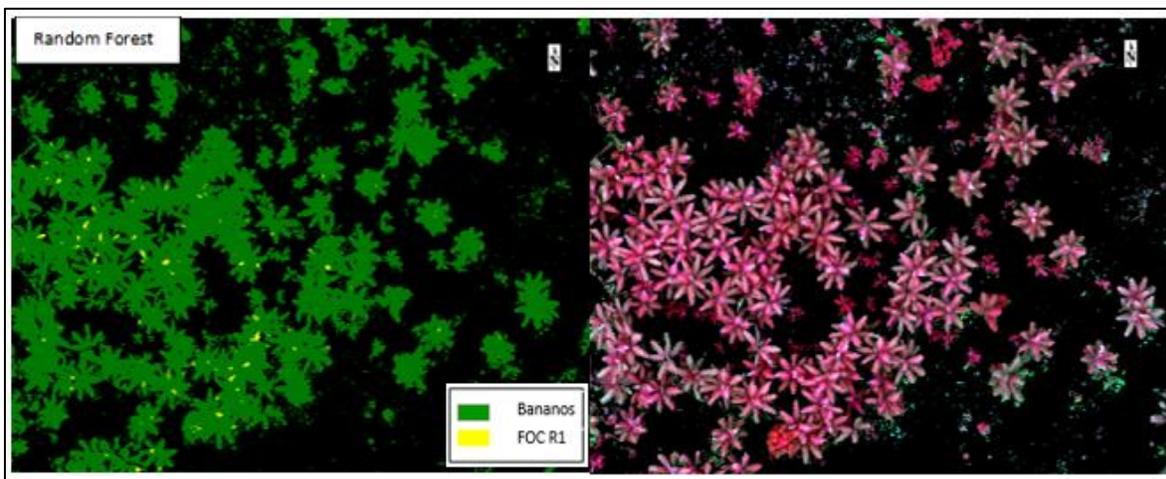
Nota: El gráfico representa el resultado de la clasificación en la parcela 2 ubicada Patricia Pilar.

Random Forest

El uso de este algoritmo de clasificación solo dio resultados en la parcela 1 (ver Figura 28), debido a que en la parcela 2 solo se evidenció la clasificación de la planta sin mostrar la sintomatología que esta presenta. Para realizar esta clasificación se deben tomar los rangos de las hojas con sintomatología, por lo cual es importante que la enfermedad del banano presente su sintomatología en las hojas y de esta manera se pueda aplicar correctamente el algoritmo de clasificación.

Figura 28

Resultado de la clasificación de Random Forest en la parcela 1



Nota: El gráfico representa el resultado de la clasificación en la parcela 2 ubicada en Caluma.

Análisis estadístico de la matriz de confusión

Se generó la matriz de confusión para cada zona de estudio, para obtener las predicciones realizadas por el algoritmo y las instancias en cada clase real (ver Figura 29 y 30).

Figura 29

Resultado de la clasificación de Support Vector Machine en la parcela 1

Support Vector Machine			
n=49		Predicción	
		Sigatoka y Picudo	Sigatoka
Observación	Sigatoka y Picudo	1 (VP)	9 (FN)
	Sigatoka	1 (FP)	38 (VN)

Nota: El gráfico representa los resultados obtenidos en la matriz de confusión.

Figura 30

Resultado de la clasificación de Random Forest en la parcela 2

Random Forest			
n=25		Predicción	
		FOC R1	Sana
Observación	FOC R1	6 (VP)	2 (FN)
	Sana	6 (FP)	11 (VN)

Nota: El gráfico representa los resultados obtenidos en la matriz de confusión.

Los valores obtenidos de las métricas de la matriz de confusión, permitieron visualizar el desempeño de cada algoritmo aplicado en las diferentes zonas de estudio (ver Tabla 13).

Tabla 13

Métricas de la matriz de confusión

Algoritmo de clasificación	Exactitud	Precisión	Sensibilidad	Tasa de error
Random Forest	0.680	0.5	0.75	0.320
Support Vector Machine	0.796	0.5	0.1	0.204

Nota: La tabla indica los resultados obtenidos de la exactitud, precisión sensibilidad y tasa de error derivados de las métricas calculadas.

Para el algoritmo de clasificación de Random Forest aplicado en la parcela 1, se obtuvo una exactitud de 68% y una precisión de 50% lo cual indica que el algoritmo de clasificación es más exacto que preciso, también presenta una sensibilidad del 75% lo cual representan los casos positivos bien clasificados, por lo tanto este algoritmo si es útil para la clasificación.

En el algoritmo de Support Vector Machine aplicado en la parcela 2, se obtuvo una exactitud del 80% y una precisión del 50% lo que indica que los datos presentan una buena

exactitud, sin embargo al analizar la sensibilidad el algoritmo presenta una sensibilidad muy baja (10%), lo que indica que el algoritmo de clasificación no es útil, debido a que el algoritmo clasificó a casi todas las plantas con Sigatoka sin identificar las plantas que presentan Sigatoka y Picudo Negro.

Análisis estadístico del índice kappa

Los resultados del índice kappa aplicado a los dos algoritmos se encuentran en la Tabla 14, de acuerdo a la Tabla 1 en la cual se indica la valoración del índice kappa, el algoritmo de Support Vector Machine presenta una valoración insignificante, mientras que el algoritmo de Random Forest presenta una valoración mediana.

Tabla 14

Índice kappa

Parcela	Algoritmo de clasificación	Índice kappa
1	Random Forest	0.351
2	Support Vector Machine	0.106

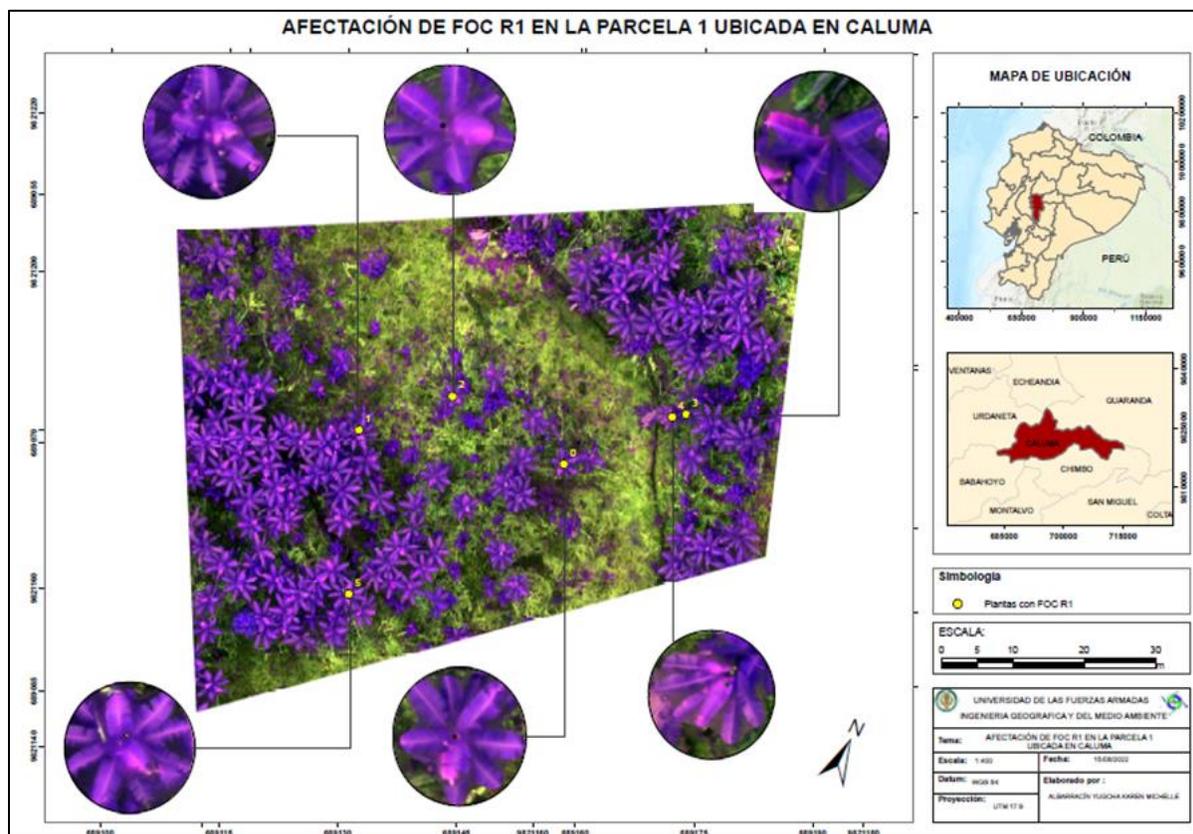
Nota: La tabla indica los resultados obtenidos del índice kappa aplicados a los algoritmos de clasificación.

Plano de afectación de Fusarium raza 1

Como resultado se obtuvo que el único algoritmo de clasificación útil fue el de Random Forest, a partir de esta clasificación se presenta el plano de afectación de FOC R1 de la parcela 1 Ubicada en Caluma (ver Figura 31). En el plano de afectación de FOC R1 en bananos se puede evidenciar las 6 muestras que fueron clasificadas correctamente con la sintomatología.

Figura 31

Plano de afectación de FOC R1 en la parcela 1 ubicada en Caluma.



Nota: El gráfico representa la afectación de FOC R1 en la parcela 1 ubicada en Caluma.

Capítulo V

Conclusiones y recomendaciones

Conclusiones

Se obtuvo un ortomosaico de cada zona de estudio a partir de la cámara multiespectral Parrot Sequoia montada en el UAV DJI Phantom IV, con el fin de elegir los índices de vegetación y el algoritmo de clasificación más adecuado, que permitan identificar Fusarium raza 1 en la parcela 1, la cual está ubicada en Caluma, y la identificación de Sigatoka Negra y Picudo Negro en la parcela 2 ubicada en Patricia Pilar.

Utilizando las 4 bandas (Green, Red, Red Edge y Nir) que proporciona la cámara multiespectral se calcularon 5 índices de vegetación: NDVI, GNDVI, NDRE, CIRE y CI_{green} . En la parcela 1 los índices de NDVI, GNDVI, y CI_{green} al ser analizados por separado brindaron valores cercanos a la realidad, debido a que las plantas con sintomatología presentan valores más bajos, mientras que en las plantas sin sintomatología estos valores aumentan; sin embargo estos valores al ser analizados estadísticamente no establecen una diferencia significativa entre las plantas con FOC R1 y plantas sin FOC R1.

En la parcela 2 los 5 índices (NDVI, GNDVI, NDRE, CIRE Y CI_{green}) brindaron valores cercanos a la realidad, ya que las plantas con afectación por las dos enfermedades (Sigatoka Negra y Picudo Negro) presentan valores más bajos que las plantas que presentan solo Sigatoka Negra; sin embargo estos valores al ser analizados estadísticamente no presentan diferencia significativa entre plantas con Sigatoka Negra y Plantas con Sigatoka Negra y Picudo Negro.

En cada zona de estudio se aplicaron 2 algoritmos de clasificación: Support Vector Machine y Random Forest. En la parcela 1 el algoritmo de clasificación supervisada más adecuado fue el de Random Forest, el cual obtuvo una exactitud del 68%, una precisión del 50%, una sensibilidad del 75% y un índice kappa de 0.35.

En la parcela 2, el algoritmo de Support Vector Machine fue el que identificó las 3 clases, sin embargo al ser analizado estadísticamente se concluye que no es útil. Aunque la exactitud (80%) y la precisión (50%) son aceptables, la sensibilidad (10%) no permite aceptar el algoritmo, debido a que este método clasificó a las plantas con Sigatoka Negra y Picudo Negro como plantas solo con Sigatoka Negra.

Recomendaciones

Para la identificación de FOC R1 y Sigatoka Negra se recomienda utilizar un espectroradiómetro directamente en las hojas que presenta la sintomatología, teniendo en cuenta que los datos espectrales sean tomados en varias hojas con sintomatología, con el fin de obtener un conjunto de datos más robustos.

Se recomienda replicar la metodología en una plantación de banano que presente Sigatoka Negra y otra enfermedad que su afectación se evidencie directamente a las hojas de banano.

Se recomienda replicar la metodología en una parcela que no sea muy densa, y que los bananos se encuentren distribuidos correctamente, para evitar el cruce de hojas entre los bananos. También es importante que la parcela se encuentre limpia y sin hojas de bananos en el suelo, ya que esto no permite realizar una clasificación correcta.

Se recomienda que para la toma de muestras con GPS, esta sea reemplazada por otro método, que presenta una mejor precisión y de esta manera se puede obtener coordenadas más precisas de los bananos que presentan la sintomatología.

Bibliografía

- Abraira, V. (2001). El Índice kappa. *SEMERGEN-Medicina de Familia*, 27(5), 247-249.
[https://doi.org/https://doi.org/10.1016/S1138-3593\(01\)73955-X](https://doi.org/https://doi.org/10.1016/S1138-3593(01)73955-X)
- AGROCALIDAD. (2016). *Banano*. Oferta Exportable Ecuatoriana - Banano:
https://www.agrocalidad.gob.ec/BPA/?dt_portfolios=banano
- AGROCALIDAD. (2020). *RECOMENDACIONES PARA PREVENIR EL INGRESO DE Fusarium oxysporum f.sp. cubense RAZA 4 Tropical*. Quito. <https://www.agrocalidad.gob.ec/wp-content/uploads/2020/05/foc11.pdf>
- AGROCALIDAD. (2021). *Ecuador sin Fusarium R4T*. https://www.agrocalidad.gob.ec/wp-content/uploads/2021/05/ROTAFOLIO-A1_compressed.pdf
- Alvarez, E., Pantoja, A., Gañan, L., & Ceballo, G. (2013). *La Sigatoka negra en plátano y banano: Guía para el reconocimiento y manejo de la enfermedad, aplicado a la agricultura familiar*. FAO.
- Arias, P., Dankers, C., Liu, P., & Pilkauskas, P. (2004). *LA ECONOMÍA MUNDIAL DEL BANANO 1985-2002*. ROMA: FAO.
<https://www.fao.org/3/y5102s/y5102s00.htm#Contents>
- Berenson, M., Levine, D. M., & Krehbiel, T. C. (2006). *Estadística para administración* (4 ed.). México: Prentice Hall.
- Borrás, 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). <https://doi.org/10.4995/raet.2017.7133>
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45, 5-32.
<https://doi.org/https://doi.org/10.1023/A:1010933404324>

- Calderón Romero, A., & Hurtado Cortes, H. (2019). Machine learning en la detección de enfermedades en plantas. *Tecnología Investigación y Academia*, 7(2), 55-61.
<https://revistas.udistrital.edu.co/index.php/tia/article/view/15685>
- Castañeda, M., Raul, & Ortiz Davila, Á. (2018). Máquinas de soporte vectorial para clasificación supervisada de imágenes en bases de datos espaciales. *Revista Ibérica de Sistemas e Tecnologías de Información*, E17, 538-551.
https://www.researchgate.net/publication/341106050_Maquinas_de_soporte_vectorial_para_clasificacion_supervisada_de_imagenes_en_bases_de_datos_espaciales
- Chuvieco, E. (1990). *Fundamentos de teledetección espacial*. Madrid: Rialp, S.A.
- CONSTITUCIÓN DE LA REPUBLICA DEL ECUADOR [Const.]. (2008). *Artículo 281*. Quito: Quito: Ediciones Legales .
- De Granda, E., Orozco De Amezquita, M. C., Bautista Mendoza, G. R., & Valencia Zapata, H. A. (2001). Fusarium Oxysporum el hongo que nos falta conocer. *Acta Biológica Colombiana*, 6(1), 7-25.
<https://revistas.unal.edu.co/index.php/actabiol/article/view/63462>
- Earth Observing System. (18 de 02 de 2022). *Earth Observing System*. Índices De Vegetación En La Agricultura Digital: <https://eos.com/es/blog/indices-de-vegetacion/>
- Espinosa Velepucha, Y., Quevedo Guerrero, J. N., & García Batista, R. M. (2019). Determinación de la eficiencia de diferentes trampas para el control de picudo negro (Cosmopolites Sordidus G.) en banano orgánico. *Revista Científica Agroecosistemas*, 7(1), 171-180. <https://aes.ucf.edu.cu/index.php/aes>
- Espinosa, A., Vivas, L., Lara, E., & Pico, J. (2004). *Manejo del picudo negro (Cosmopolites sordidus Germar) con el hongo Entomopatogeno Beauveria bassiana(Bals.) Vull.*
<https://repositorio.iniap.gob.ec/bitstream/41000/2016/1/iniaplsbd2004m.pdf>

- Esri. (s.f.). *ArcGIS Pro*. Clre: <https://pro.arcgis.com/es/pro-app/latest/arcpy/image-analyst/cire.htm>
- FAO. (2016). *Todo sobre los bananos: lo que debería saber acerca de esta fruta tropical*. Retrieved 08 de 01 de 2022, from <https://www.fao.org/zhc/detail-events/es/c/447827/>
- Flores Tapia, C. E., & Flores Cevallos, K. L. (2021). Pruebas para comprobar la normalidad de datos en procesos productivos: andersondarling, ryan-joiner, shapiro-wilk y kolmogórov-smirnov. *Societas. Revista de Ciencias Sociales y Humanísticas*, 23(2). <http://portal.amelica.org/ameli/jatsRepo/341/3412237018/index.html>
- Fourie, G., Steenkamp, E. T., Gordon, T. R., & Viljoen, A. (2009). Evolutionary Relationships among the *Fusarium oxysporum* f. sp. *cubense* Vegetative Compatibility Groups. *Applied and Environmental Microbiology*, 75(14), 4770-4781. <https://doi.org/https://doi.org/10.1128/AEM.00370-09>
- Getha, K., & Vikineswary, S. (2002). Antagonistic effects of *Streptomyces violaceusniger* strain G10 on *Fusarium oxysporum* f.sp. *cubense* race 4: Indirect evidence for the role of antibiosis in the antagonistic process. *Journal of Industrial Microbiology and Biotechnology*, 28(6), 303-310. <https://doi.org/https://doi.org/10.1038/sj/jim/7000247>
- Gitelson, A., Viña, A., Ciganda, V., Donald C, R., & Timothy, A. (2005). Remote estimation of canopy chlorophyll content in crops. *GEOPHYSICAL RESEARCH LETTERS*, 2-3.
- Gold, C., Kagezi, G., Night, G., & Ragama, P. (2004). The effects of banana weevil, *Cosmopolites sordidus*, damage on highland banana growth, yield and stand duration in Uganda. *Annals of Applied Biology*, 145(3), 263-269. <https://doi.org/https://doi.org/10.1111/j.1744-7348.2004.tb00382.x>
- Gold, C., Ragama, P., Coe, R., & Rukazambuga, N. (2005). Selection of assessment methods for evaluating banana weevil *Cosmopolites sordidus* (Coleoptera: Curculionidae)

- damage on highland cooking banana (*Musa* spp., genome group AAA-EA). *Bulletin of Entomological Research*, 115-123. <https://doi.org/doi:10.1079/BER2004341>
- Gómez Candón, D., De Castro, A., & López Granados, F. (2014). Assessing the accuracy of mosaics from unmanned aerial vehicle (UAV) imagery for precision agriculture purposes in wheat. *Precision Agric*, 44-56. <https://doi.org/https://doi.org/10.1007/s11119-013-9335-4>
- Gómez, A., Clavijo, F., & Jimenez, A. (2016). *Agricultura de Precisión y Sensores Multiespectrales Aerotransportados*.
- Guzmán Quesada, M., & Paladines García, R. (s.f.). *CropLife Latín America*. Retrieved 20 de 02 de 2022, from Sigatoka Negra: <https://www.croplifela.org/es/plagas/listado-de-plagas/sigatoka-negra>
- INEC. (2011). *Reporte estadístico del Sector Agropecuario*. https://www.ecuadorencifras.gob.ec/wp-content/descargas/Presentaciones/espac_2010.pdf
- INIAP. (s.f.). *Banano, plátano y otras musáceas*. <http://www.iniap.gob.ec/pruebav3/banano-platano-y-otras-musaceas/>
- Landis, J., & Koch, G. (1977). La medición de la concordancia del observador para datos categóricos. *Biometría*, 33(1), 159-174. <https://doi.org/https://doi.org/10.2307/2529310>
- Leiva, F. (2003). La agricultura de precisión: una producción más sostenible y competitiva con visión futurista. https://www.researchgate.net/publication/228425520_La_agricultura_de_precision_una_produccion_mas_sostenible_y_competitiva_con_vision_futurista
- Lin, Y. H., Chang, J. Y., Liu, E. T., Chao, C. P., Huang, J. W., & Linda Chang, P. F. (2009). Development of a molecular marker for specific detection of *Fusarium oxysporum* f. sp.

- cubense race 4. *European Journal of Plant Pathology*(123), 353-365.
<https://doi.org/https://doi.org/10.1007/s10658-008-9372-4>
- Lind, D. A., Marchal, W. G., & Wathen, S. A. (2012). *Estadística aplicada a los negocios y la economía* (15 ed.). McGRAW-HILL.
- Liu, L., Dong, Y., Huang, W., Du, X., & Ma, H. (2020). Monitoring Wheat Fusarium Head Blight Using Unmanned Aerial Vehicle Hyperspectral Imagery. *Remote Sensing*, 12(22).
<https://doi.org/https://doi.org/10.3390/rs12223811>
- Luque Ordóñez, J. (2012). Espectro Electromagnético y espectro radioeléctrico. *ACTA*(62), 17-31. <https://www.acta.es/recursos/revista-digital-manuales-formativos/13-062>
- Ministerio de Comercio Exterior. (2017). *Informe sector Bananero Ecuatoriano*.
<https://www.produccion.gob.ec/wp-content/uploads/2019/06/Informe-sector-bananero-espa%C3%B1ol-04dic17.pdf>
- Miranda, I., Gracia Perera, D., & Rodríguez, M. (2019). Meta-análisis de las estrategias para el manejo de *Cosmopolitis sordidus* Guermer en *Musa* spp. *Revista de Protección Vegetal*, 34(2). <http://scielo.sld.cu/pdf/rpv/v34n2/2224-4697-rpv-34-02-e08.pdf>
- Naiqian, Z., Wang, M., & Wang, N. (2002). Precision agriculture- A worldwide overview. *Computers and Electronics in Agriculture* , 36(2-3), 113-132.
[https://doi.org/http://dx.doi.org/10.1016/S0168-1699\(02\)00096-0](https://doi.org/http://dx.doi.org/10.1016/S0168-1699(02)00096-0)
- Piedad Rubio, A., Hernández López, D., Lárraga Altamirano, H., & Zacarías González, E. (2020). TELEDETECCIÓN EN LA AGRICULTURA DE PRECISIÓN:. *Revista de divulgación científica y tecnológica*, 6(2), 46-58.
<https://www.eumed.net/uploads/revistas/4f75a0794aa27a6c69b687652d5a122b.pdf#page=47>
- Porta, J. (2006). Técnicas cuantitativas para la extracción de términos en un corpus. En *Terminología y Sociedad del conocimiento*. Peter Lang.

- Praveen, E. (2017). A Review Paper on Unmanned Aerial Vehicle (U.A.V.). *INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT)*, 5(23).
<https://www.ijert.org/a-review-paper-on-unmanned-aerial-vehicle-u-a-v>
- Rubio Hurtado, M. J., & Berlanga Silvente, V. (2012). Cómo aplicar las pruebas paramétricas bivariadas t de Student y ANOVA en SPSS. Caso práctico. *REIRE*, 5(2).
<https://doi.org/10.1344/reire2012.5.2527>
- Sacristán Romero, F. (2006). La Teledetección satelital y los sistemas de protección ambiental. *AquaTIC*(24), 13-41. http://www.revistaaquatic.com/aquatic/pdf/24_02.pdf
- Sarria, F. (2006). *Sistemas de Información Geográfica*. Retrieved 14 de 1 de 2022, from Teledetección: <https://www.um.es/geograf/sigmur/temariohtml/>
- Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding Machine Learning: From Theory to Algorithms*, (1 ed.). Cambridge University Press.
- Shen, Z., Xue, C., Penton, C. R., Thomashow, L. S., Zhang, N., Wang, B., . . . Shen, Q. (2019). Suppression of banana Panama disease induced by soil microbiome reconstruction through an integrated agricultural strategy. *Soil Biology and Biochemistry*, 128, 164-174.
<https://doi.org/https://doi.org/10.1016/j.soilbio.2018.10.016>
- Shi, Y., Huang, W., Ye, H. R., Xing, N., Geng, Y., Dong, Y., & Peng, D. P. (2018). Partial Least Square Discriminant Analysis Based on Normalized Two-Stage Vegetation Indices for Mapping Damage from Rice Diseases Using PlanetScope Datasets. *Sensors (Basel)*, 18(6). <https://doi.org/https://dx.doi.org/10.3390/s18061901>
- Vite Cevallos, H. A., Vargas Collaguazo, L. A., & Vargas Collaguazo, J. D. (2018). Uso de índices espectrales en la agricultura de precisión caso de estudio campus de la Facultad Técnica de Machala. *Alternativas*, 19(1), 71-79.
<https://doi.org/http://dx.doi.org/10.23878/alternativas.v19i1.195>

Ye, H., Huang, W., Huang, S., Cui, B., Dong, Y., Guo, A., . . . Jin, Y. (2020). Recognition of Banana Fusarium Wilt Based on UAV Remote Sensing. *Remote Sens*, 12(6), 938.

<https://doi.org/https://doi.org/10.3390/rs12060938>

Ye, H., Huang, W., Shanyu, H., Cui, B., Dong, Y., Guo, A., . . . Jin, Y. (2020). Identification of banana fusarium wilt using supervised classification algorithms with UAV-based multi-spectral imagery. *Int J Agric & Biol Eng*, 13(3), 136-142.

<https://doi.org/10.25165/j.ijabe.20201303.5524>