

**ESTIMACIÓN DE LA SALUD DE CULTIVOS DE CAFÉ
MEDIANTE ANÁLISIS DE IMÁGENES
MULTIESPECTRALES CAPTURADAS EN VEHÍCULOS
AÉREOS NO TRIPULADOS**



Trabajo de Grado en
Modalidad de Trabajo de investigación

**NATALIA ARTEAGA LÓPEZ
CARLOS ANDRÉS DELGADO CALVACHE**

Director: PhD. Cristhian Nicolás Figueroa Martínez
Codirector: PhD. Juan Carlos Corrales
Asesor: Mg. Juan Fernando Casanova

Universidad del Cauca
**Facultad de Ingeniería Electrónica y Telecomunicaciones
Departamento de Telemática
Popayán, octubre de 2021**

ESTIMACIÓN DE LA SALUD DE CULTIVOS DE CAFÉ MEDIANTE ANÁLISIS DE IMÁGENES MULTIESPECTRALES CAPTURADAS EN VEHÍCULOS AÉREOS NO TRIPULADOS

Natalia Arteaga López
Carlos Andrés Delgado Calvache

Trabajo de grado presentado a la Facultad de Ingeniería Electrónica y
Telecomunicaciones de la
Universidad del Cauca para obtener el título de: Ingeniero en Electrónica y
Telecomunicaciones

Director: PhD. Cristhian Nicolás Figueroa Martínez
Codirector: PhD. Juan Carlos Corrales
Asesor: Mg. Juan Fernando Casanova

Departamento de Telemática
Facultad de Ingeniería Electrónica y Telecomunicaciones
Universidad del Cauca
Popayán, Cauca, 2021

Agradecimientos

Queremos agradecer a Dios por nuestras vidas, a nuestras familias y amigos, especialmente a nuestros padres y hermanos por su incondicional apoyo y acompañamiento en esta etapa. Agradecemos a la familia Cerón López por su cordialidad y acompañamiento. Queremos agradecer a nuestro director el ingeniero Cristhian Figueroa, nuestro codirector Juan Carlos Corrales y nuestro asesor Juan Fernando Casanova por su ayuda y guía constante en nuestra investigación.

Resumen

El uso de vehículos aéreos no tripulados (VANT) equipados con cámaras espectrales se ha incrementado en los últimos años, especialmente en el sector agrícola, debido a que permite a los agricultores e investigadores analizar el estado de un cultivo en cuanto a su salud, nutrientes, crecimiento, epidemias, entre otros parámetros. En Colombia el sector cafetero afronta varios retos, como la necesidad de incrementar la productividad, el rendimiento y la calidad del café. En este sentido el presente trabajo estimó el estado de la salud de un cultivo de café de variedad Castilla ubicado en San Joaquín, Tambo, Cauca con el fin de apoyar la toma de decisiones de los caficultores. Para ello se midieron datos de clorofila en campo con el dispositivo CCM-200 plus, se capturaron imágenes multiespectrales con la cámara MAPIR SURVEY 3 aerotransportada en un VANT SOLO 3DR y se generaron datos sintéticos con el fin de aumentar el conjunto de datos. Se obtuvieron 6 índices de vegetación, los cuales, en conjunto con los valores de clorofila se modelaron a través de la implementación de regresiones lineales simples y múltiples, árboles de decisión, máquina de vectores, bosque aleatorios y k vecinos más cercanos. El modelo que presentó el mejor desempeño con el error cuadrático medio más bajo fue el entrenado con máquina de vectores de soporte. También, se encontró que los índices que demostraron mejor desempeño en los modelos fueron el CVI, GNDVI y GCI, los cuales son ampliamente utilizados en la agricultura para estimar la clorofila de las plantas.

Keywords— Agricultura inteligente, aprendizaje automático, café, imágenes multiespectrales, índices de vegetación, VANT.

Abstract

The use of unmanned aerial vehicles (UAVs) equipped with spectral cameras has increased in recent years, especially in the agricultural sector, because it allows farmers and researchers to analyze the state of a crop in terms of its health, nutrients, growth, epidemics, among other parameters. In Colombia, the coffee sector faces several challenges, such as the need to increase coffee productivity, yield, and quality. In this sense, the present work estimated the health status of a Castilla variety coffee crop located in San Joaquín, Tambo, Cauca to support coffee growers' decision-making. For this, chlorophyll data were measured in the field with the CCM-200 plus device, multispectral images were captured with the airborne MAPIR SURVEY 3 camera in a VANT SOLO 3DR and synthetic data were generated to increase the data set. 6 vegetation indices were obtained, which, together with the chlorophyll values, were modeled through the implementation of simple and multiple linear regressions, decision trees, vector machine, random forest, and k closest neighbors. The model that presented the best performance and gave the lowest error is the one trained with a vector machine, an error mae = 7.85, in addition it was found that the indices that demonstrated better performance in the different models were the CVI, GNDVI and GCI, the which are widely used to estimate chlorophyll.

Keywords—coffee, machine learning, multispectral images, smart agriculture, vegetation indices, UAV.

Contenido

1. Introducción.....	16
1.1. Planteamiento del Problema.....	16
1.2. Antecedentes	17
1.3. Motivación.....	17
1.4. Objetivos.....	18
1.4.1. Objetivo general	18
1.4.2. Objetivos específicos	18
2. Estado del arte.....	20
2.1. Definiciones y conceptos fundamentales	20
2.1.1. Propiedades de la hoja.....	20
2.1.2. Métodos para obtener datos de la salud de la planta medidos en campo....	20
2.1.3. Agricultura de precisión	21
2.1.4. Teledetección	21
2.1.5. Vehículos aéreos no tripulados.....	21
2.1.6. Espectro electromagnético.....	22
2.1.7. Tipos de Imágenes espectrales	22
2.1.8. Índices de vegetación	22
2.2. Análisis bibliométrico	23
2.2.1. Fases metodológicas y herramienta de software.....	23
2.2.2. Búsqueda de conjuntos de datos	26
2.2.3. Análisis de resultados.....	27
2.3. Trabajos relacionados	34
2.3.1. Estudios realizados en Colombia.....	35
2.3.2. Estudios realizados en otros países.....	37
2.4. Conclusiones del estado del arte	38
3. Materiales y métodos.....	40
3.1. Materiales	40
3.2. Diseño experimental.....	42
3.2.1. Captura de datos.....	43
3.2.2. Preprocesamiento de datos	55
3.2.3. Modelado de datos.....	59
4. Conclusiones	74
5. Trabajos de futuros	77

LISTA DE FIGURAS

ILUSTRACIÓN 1. FLUJO DE TRABAJO DEL ANÁLISIS BIBLIOMÉTRICO. IMAGEN ADAPTADA DE [31].....	24
ILUSTRACIÓN 2. COMPRESIÓN Y DISTRIBUCIÓN DEL MAPA ESTRATÉGICO. FUENTE [31].....	25
ILUSTRACIÓN 3. FILTRO DE ETAPAS. FUENTE PROPIA	27
ILUSTRACIÓN 4. PRIMER PERIODO. FUENTE PROPIA	28
ILUSTRACIÓN 5. SEGUNDO PERIODO. FUENTE PROPIA.....	31
ILUSTRACIÓN 6. DOCUMENTOS DE WOS Y SCOPUS SOBRE IMÁGENES MULTIESPECTRALES DE CULTIVOS DE CAFÉ. FUENTE PROPIA	35
ILUSTRACIÓN 7. ETAPAS MODELO CRIPS-DM. FUENTE [250].	42
ILUSTRACIÓN 8. ETAPAS DISEÑO EXPERIMENTAL. FUENTE ADAPTADA DE [250].	43
ILUSTRACIÓN 9. METODOLOGÍA PARA LA CAPTURA DE DATOS. FUENTE PROPIA.....	44
ILUSTRACIÓN 10. MONTAJE DEL VANT. FUENTE PROPIA	45
ILUSTRACIÓN 11. PUNTOS DE POLÍGONO (ROJOS) Y PUNTOS DE CONTROL (VERDES). FUENTE PROPIA.	47
ILUSTRACIÓN 12. NIVEL DE SALUD POR NDVI. FUENTE [262].	48
ILUSTRACIÓN 13. MAPEO DEL CULTIVO DEL TERRENO. FUENTE PROPIA.	48
ILUSTRACIÓN 14. RASTER CALCULATOR DE QGIS. FUENTE PROPIA.	52
ILUSTRACIÓN 15. DISTRIBUCIÓN DE LAS PLANTAS EN EL TERRENO. FUENTE [266]	53
ILUSTRACIÓN 16. MUESTREO DE PUNTOS POR HOJA. FUENTE ADAPTADA DE [269].	54
ILUSTRACIÓN 17. REVISIÓN DE DATOS. FUENTE PROPIA.....	55
ILUSTRACIÓN 18. ZONA DE MEDIDA DIBUJO. FUENTE ADAPTADA DE [272].	57
ILUSTRACIÓN 19. DIFERENCIA COLOR ENTRE ZONA ALTA (A) Y ZONA MEDIA (B). FUENTE ADAPTADA DE [273].	57
ILUSTRACIÓN 20. REPORTE DE RESULTADOS DE DATOS SINTÉTICOS OBTENIDOS EN GRETAL. FUENTE CAPTURA DE LA HERRAMIENTA GRETAL [281].	58
ILUSTRACIÓN 21. REPORTE DE CORRELACIÓN ENTRE DATOS DE ENTRENAMIENTO Y SINTÉTICOS. FUENTE PROPIA.....	58
ILUSTRACIÓN 22. MODELADO DE DATOS. FUENTE PROPIA.....	59
ILUSTRACIÓN 23. RANGOS DE ESTADOS DE NUTRIENTES ADECUADOS PARA LAS HOJAS DE CAFÉ. FUENTE [252].....	60
ILUSTRACIÓN 24. MAPA DE CALOR DE CORRELACIÓN DE PEARSON. FUENTE PROPIA.	62
ILUSTRACIÓN 25. DISTRIBUCIÓN ERROR SIN DATOS SINTÉTICOS (A) Y CON DATOS SINTÉTICOS (B). FUENTE PROPIA.	63
ILUSTRACIÓN 26. DATOS REALES (AZUL) Y ESTIMADOS (ROJO) Y CORRELACIÓN DE PEARSON. FUENTE PROPIA.....	63
ILUSTRACIÓN 27. DISTRIBUCIÓN ERROR ÁRBOL DE DECISIÓN CON DATOS INICIALES (A) E INTEGRANDO DATOS SINTÉTICOS (B). FUENTE PROPIA	65
ILUSTRACIÓN 28. DATOS REALES (AZUL) Y ESTIMADOS (ROJO) Y CORRELACIÓN DE PEARSON. FUENTE PROPIA.	65
ILUSTRACIÓN 29. DISTRIBUCIÓN DEL ERROR DE DATOS DE ENTRENAMIENTO INICIALES (A) Y DATOS SINTÉTICOS (B) CON MÁQUINAS DE VECTORES. FUENTE PROPIA.....	67
ILUSTRACIÓN 30. DATOS REALES (AZUL) Y ESTIMADOS (ROJO) Y CORRELACIÓN DE PEARSON. FUENTE PROPIA.	67
ILUSTRACIÓN 31. DISTRIBUCIÓN DEL ERROR DE DATOS INICIALES (A) Y DATOS SINTÉTICOS (B) CON BOSQUE ALEATORIO. FUENTE PROPIA.	69
ILUSTRACIÓN 32. DATOS REALES (AZUL) Y ESTIMADOS (ROJO) Y CORRELACIÓN DE PEARSON. FUENTE PROPIA.	69
ILUSTRACIÓN 33. DISTRIBUCIÓN DEL ERROR DE DATOS INICIALES (A) Y DATOS SINTÉTICOS (B) CON K VECINOS MÁS CERCANOS. FUENTE PROPIA.	71
ILUSTRACIÓN 34. DATOS REALES (AZUL) Y ESTIMADOS (ROJO) Y CORRELACIÓN DE PEARSON. FUENTE PROPIA.	71
ILUSTRACIÓN 35. VALOR REAL Y VALOR DE PREDICCIÓN SEGÚN CADA MODELO. FUENTE PROPIA.	73

LISTA DE TABLAS

TABLA 1. CAMBIO DE CUADRANTE POR TEMAS. FUENTE PROPIA.....	33
TABLA 2. RESUMEN COMPARATIVO DE TRABAJOS RELACIONADOS. FUENTE PROPIA.	38
TABLA 3. ESPECIFICACIONES CÁMARA MAPIR SURVEY 3. FUENTE [238]	40
TABLA 5. ESPECIFICACIONES DEL CLOROFILÓMETRO CCM 200-PLUS [249].	41
TABLA 6. VALORES ÍNDICES DE VEGETACIÓN. FUENTE PROPIA.....	50
TABLA 7. VARIACIÓN ESTÁNDAR POR NÚMERO DE MUESTRAS. FUENTE PROPIA.....	53
TABLA 8. DESVIACIÓN ESTÁNDAR PROMEDIO DE PUNTOS DENTRO DE CADA HOJA. FUENTE PROPIA.	54
TABLA 9. CORRELACIÓN DE PEARSON DE ÍNDICES DE VEGETACIÓN. FUENTE PROPIA	57
TABLA 10. VALORES SPAD Y CCI POR NIVEL DE SALUD. FUENTE ADAPTADA DE DONG T ET AL., [249].	60
TABLA 11. CORRELACIÓN ENTRE ÍNDICES DE VEGETACIÓN Y CLOROFILA. FUENTE PROPIA.....	61
TABLA 12. ERROR DE DATOS DE ENTRENAMIENTO Y DATOS SINTÉTICOS CON REGRESIÓN LINEAL MÚLTIPLE. FUENTE PROPIA.....	62
TABLA 13. ERROR DE DATOS DE ENTRENAMIENTO Y DATOS SINTÉTICOS POR CADA ÍNDICE DE VEGETACIÓN CON REGRESIÓN LINEAL SIMPLE. FUENTE PROPIA.	64
TABLA 14. ERROR ÁRBOL DE DECISIÓN CON DATOS INICIALES Y CON DATOS SINTÉTICOS. FUENTE PROPIA.....	65
TABLA 15. ERROR DE DATOS DE ENTRENAMIENTO Y DATOS SINTÉTICOS POR CADA ÍNDICE DE VEGETACIÓN CON ÁRBOL DE DECISIÓN. FUENTE PROPIA.....	66
TABLA 16. ERROR DE DATOS DE ENTRENAMIENTO Y DATOS SINTÉTICOS CON MÁQUINAS DE VECTORES. FUENTE PROPIA.	66
TABLA 17. ERROR DE DATOS DE ENTRENAMIENTO Y DATOS SINTÉTICOS POR CADA ÍNDICE DE VEGETACIÓN CON MÁQUINA DE VECTORES. FUENTE PROPIA.....	68
TABLA 18. ERROR DE DATOS DE ENTRENAMIENTO Y DATOS SINTÉTICOS CON BOSQUE ALEATORIO. FUENTE PROPIA.	68
TABLA 19. ERROR DE DATOS DE ENTRENAMIENTO Y DATOS SINTÉTICOS POR CADA ÍNDICE DE VEGETACIÓN CON BOSQUE ALEATORIO. FUENTE PROPIA	70
TABLA 20. ERROR DE DATOS DE ENTRENAMIENTO Y DATOS SINTÉTICOS CON K VECINOS MÁS CERCANOS. FUENTE PROPIA	70
TABLA 21. ERROR DE DATOS DE ENTRENAMIENTO Y DATOS SINTÉTICOS POR CADA ÍNDICE DE VEGETACIÓN CON K VECINOS MÁS CERCANOS. FUENTE PROPIA.....	72
TABLA 22. COMPARACIÓN ERRORES Y CORRELACIÓN ENTRE LOS MODELOS. FUENTE PROPIA. 72	
TABLA 23. ERRORES POR MODELOS E ÍNDICES DE VEGETACIÓN. FUENTE PROPIA.	73

LISTA DE ACRONIMOS

CCI	<i>Chlorophyll Content Index</i>
CVI	<i>Chlorophyll Vegetation Index</i>
GCI	<i>Green Chlorophyll Index</i>
GNDVI	<i>Green Normalized Difference Vegetation Index</i>
NDVI	<i>Normalized Difference Vegetation Index</i>
NRVI	<i>Normalized Ratio Vegetation Index</i>
NIR	<i>Near Infrared</i>
PA	<i>Precision Agriculture</i>
PCA	<i>Principal Component Analysis</i>
RVI	<i>Ratio Vegetation Index</i>
SAVI	<i>Soil Adjusted Vegetation Index</i>
SLR	<i>Systematic Literature Reviews in Software Engineering</i>
SPAD	<i>Soil Plant Analysis Development</i>
UAV	<i>Unmanned Aircraft Vehicle</i>
VANT	<i>Vehículo Aéreo No Tripulado</i>
VI	<i>Vegetation Indexes</i>

ANEXOS

ANEXO A: Configuración del vuelo del VANT.

ANEXO B: Creación de ortomosaicos.

ANEXO C: Obtención de los índices de vegetación con QGIS.

ANEXO D: Archivos con los índices de vegetación obtenidos.

ANEXO E: Desviación estándar de los datos de clorofila por planta.

ANEXO F: Desviación estándar de los datos de clorofila por hoja.

ANEXO G: Datos de entrada para los modelos.

ANEXO H: Reportes de generación de datos sintéticos con Gretel.

ANEXO I: Ejemplo de procedimiento realizado con redes neuronales.

ANEXO J: Artículo y Poster.

ANEXO K: Artículo resultados de la investigación.

ANEXO L: Imágenes multiespectrales obtenidas

CAPÍTULO 1

1. Introducción

1.1. Planteamiento del Problema

El consumo de café a nivel mundial ha incrementado en los últimos años, obligando a países líderes en la producción de café a mejorar su rendimiento. Colombia es el único de los cinco primeros países productores de café que aún no ha podido superar el nivel de producción del año 1991, siendo este, el máximo valor alcanzado en los últimos 25 años [1]. Esta baja productividad del café en Colombia es resultado de diferentes problemáticas, tales como: cambio climático, epidemias, falta de apoyo económico por parte de entidades gubernamentales, falta de innovación y tecnología, mal manejo de recursos, entre otros [2]–[5]. La calidad y cantidad de la producción de café dependen del estado de salud de las plantas, el cual, generalmente es monitoreado de forma visual por los agricultores; esta observación directa de las plantas es propensa a errores y no permite conocer de forma global y rápida el estado de salud del cultivo [6].

Surge entonces, la necesidad de emplear herramientas tecnológicas económicas que sean capaces de reconocer la salud en los cultivos con mayor rapidez y precisión. En respuesta a esa necesidad tecnológica, se encuentra la agricultura de precisión (*PA*, por sus siglas en inglés). La *PA* es una estrategia de gestión agrícola, que apoya al agricultor en sus actividades y en la toma de decisiones asociadas a la producción de cultivos al utilizar diferentes técnicas de monitoreo [7]. Esas técnicas permiten observar grandes y pequeñas zonas de cultivo ya que utilizan tecnologías como los sistemas de información geográfica, satélites, visión por computador, Vehículos Aéreos No Tripulados (*VANT*), entre otros.

En cuanto a la visión por computador, la fuente más usada de imágenes terrestres para la *PA* es el satélite, pero este presenta algunas limitaciones tales como su alta sensibilidad a la variabilidad de las condiciones climáticas, el bajo monitoreo en tiempo real y la dificultad de obtener resoluciones espaciotemporales adecuadas. En consecuencia, ha incrementado el uso de *VANT* [8]–[10], ya que representan un enfoque prometedor para superar las limitaciones de la teledetección satelital [11]. El uso de *VANT* equipados con cámaras espectrales permite ver parámetros de la vegetación como el nivel del verdor, cantidad de agua y la diferencia espectral entre la vegetación sana y la enferma [12], entre otros; para ello se realizan estudios cuantitativos mediante diversas técnicas como el cálculo de índices de vegetación (*VI*, por sus siglas en inglés) [13], [14], que sirven como indicadores del estado de salud del cultivo [15], [16].

Existen dos tipos principales de imágenes útiles para el análisis de la salud en las plantas, las hiperespectrales y las multiespectrales. Las primeras, poseen varias bandas espectrales de información a través de todo el espectro electromagnético y se utilizan frecuentemente en aplicaciones de PA y teledetección [17]–[20], sin embargo son muy costosas para agricultores pequeños o pequeñas empresas de teledetección [21]. En contraste, las imágenes multiespectrales son más económicas porque utilizan menos bandas espectrales, pero también permiten analizar la información del espectro electromagnético con un nivel aceptable de precisión [22].

¿Es posible estimar el estado de salud de pequeños cultivos de café en el departamento del Cauca utilizando imágenes capturadas por cámaras multiespectrales de bajo costo transportadas en VANT?

1.2 Antecedentes

Gracias al proyecto IoT-Agro¹ de la universidad del Cauca tenemos la oportunidad de trabajar con las cámaras multiespectrales Agrocám y Survey 3. Estas cámaras poseen las bandas rojo, verde e infrarrojo cercano (NIR, por sus siglas en inglés), con las cuales podemos hallar índices de vegetación como el Índice de vegetación de diferencia normalizada (NDVI, por sus siglas en inglés), el Índice de vegetación de diferencia normalizada verde (GNDVI, por sus siglas en inglés) y el Índice de vegetación ajustado al suelo (SAVI, por sus siglas en inglés), los cuales ayudan a analizar el estado de salud de un cultivo.

1.3 Motivación

En Colombia, los caficultores se enfrentan a varios desafíos, entre ellos la falta de apoyo económico por parte de entidades gubernamentales, recursos limitados y bajo acceso a tecnologías que optimicen sus actividades [23]. Estas limitaciones se pueden abordar desde diferentes puntos, sin embargo, desde el ámbito tecnológico es necesaria una solución eficiente, económica, accesible y que sea fácil de implementar.

Una de las tecnologías más usadas y menos costosas para optimizar la gestión agrícola, son los VANT, ya que estos equipados con cámaras multiespectrales permiten ver la diferencia espectral entre la vegetación sana y la enferma [13], [24]. Al hacer uso de VANT equipados con cámaras multiespectrales para analizar cultivos de café se debe tener en cuenta la variedad de suelos, entornos y el tratamiento que se le da. Por consiguiente, pequeñas variaciones en dichas condiciones pueden representar un impacto en las diferentes características presentes en los cultivos y el análisis de su reflectancia [15], [25], [26]. Por

¹ IoT-Agro: Alternativas Innovadoras de Agricultura Inteligente para sistemas productivos agrícolas del departamento del Cauca soportado en entornos de IoT [312].

ejemplo, es diferente analizar cultivos de café bajo sistemas agroforestales a cultivos de libre exposición ya que crecen en diferentes microclimas.

Por estas razones, la presente investigación se enfocó en el estudio del estado de salud de cultivos de café a través del análisis de imágenes obtenidas de las cámaras multiespectrales Agrocám y Survey 3 transportadas en VANT, con el fin de crear herramientas que a futuro permitan apoyar al caficultor en el monitoreo de la salud general de sus cultivos de una forma global y rápida, ahorrando recursos, tiempo y dinero.

1.4 Objetivos

1.4.1 Objetivo general

Estimar el estado de salud de cultivos de café mediante el análisis de imágenes multiespectrales capturadas por vehículos aéreos no tripulados.

1.4.2 Objetivos específicos

1. Calcular índices de vegetación de un cultivo de café a partir de imágenes capturadas por cámaras multiespectrales transportadas en VANT.
2. Crear un modelo basado en aprendizaje automático que permita relacionar los índices de vegetación con datos de la salud de la planta medidos en campo.
3. Evaluar la precisión del modelo creado para conocer la salud de un cultivo.

1.5 Estructura del documento

Este documento ha sido dividido en los capítulos descritos a continuación.

- Capítulo 1: expone las bases del presente trabajo de grado, presenta el planteamiento detallado del problema, la motivación del tema del trabajo y define los objetivos y aportes investigativos del trabajo.
- Capítulo 2: describe una base conceptual para entrar en contexto en torno a los conceptos básicos y artículos relacionados que tienen que ver con la agricultura, las cámaras multiespectrales y las aeronaves no tripuladas.
- Capítulo 3: presenta la recolección bibliométrica con la herramienta Scimart y la revisión sistemática de literatura (SLR, por sus siglas en inglés) [27] en ingeniería de software a través de la metodología de Kitchenham [28], [29].
- Capítulo 4: presenta el desarrollo del modelo de estimación de la salud de cultivos de café.

- Capítulo 5: presenta la experimentación y resultados, donde se hacen pruebas del funcionamiento del algoritmo.
- Capítulo 6: se presentan las conclusiones y trabajos futuros de la investigación.

En este capítulo se explicaron las bases y fundamentos del presente trabajo de grado, se realizó el planteamiento del problema, la motivación, justificación, además de los objetivos y aportes investigativos del trabajo.

CAPÍTULO 2

2. Estado del arte

En este capítulo se expone el análisis bibliométrico y sistemático realizado. El cual se focalizó en las investigaciones relacionadas con el empleo del espectro electromagnético en el sector agrícola. Para el análisis bibliométrico se utilizó la herramienta SciMAT [27], [28], [30] y para la revisión sistemática de literatura se siguieron las guías propuestas por Kitchenham [28], [30].

2.1. Definiciones y conceptos fundamentales

2.1.1 Propiedades de la hoja

La detección remota de la cobertura de cultivos se fundamenta en algunas propiedades que tienen las hojas al interactuar con la radiación solar. Se presentan a continuación estas propiedades de acuerdo con el trabajo de Gómez, K [15].

- Estructura de la hoja: en el campo de la teledetección, muchas de las aplicaciones en la agricultura dependen de la comprensión de las propiedades espectrales de las plantas y cada una de sus partes. Aunque la estructura de la hoja no es igual en todas las plantas, existen ciertos elementos comunes a la mayoría, como la epidermis superior, la epidermis inferior, el tejido empalizado y el tejido esponjoso [15].
- Comportamiento espectral de la hoja viva: las características espectrales de la hoja cambian con la madurez o estrés en las plantas, estos cambios aparentemente ocurren simultáneamente en las regiones visibles e infrarrojo, lo cual permite realizar estudios de vegetación como por ejemplo: conocer el estado de madurez y salud de la planta, saber que zonas poseen superficie vegetativa viva e identificar si hay escasez de humedad [15].

2.1.2 Métodos para obtener datos de la salud de la planta medidos en campo

Para la obtención de datos de la salud de la planta existen diversos métodos, uno de ellos es el cálculo del índice del área foliar [31]. Otros métodos requieren de pruebas en laboratorio, por ejemplo, calcular el contenido hídrico relativo [32] y el nivel de clorofila presente en una planta utilizando dimetilsulfóxido [33]. También existe la posibilidad de determinar los niveles

de clorofila a través de una medición portátil con dispositivos electrónicos como lo es el CCM-200 plus [34].

2.1.3 Agricultura de precisión

La Agricultura de Precisión (*PA*, por sus siglas en inglés) es el conjunto de técnicas y tecnologías que se aplican al trabajo realizado entorno a la agricultura, con el fin de optimizar el rendimiento, precisión y eficacia de producción de los cultivos. Para ello utiliza la teledetección usando dispositivos remotos como satélites, vehículos aéreos no tripulados (VANT), sensores, imágenes y datos geográficos; que reúnen información de utilidad para el agricultor, además de ser amigables con el medio ambiente [35]–[37].

2.1.4 Teledetección

Permite adquirir y recolectar información global a distancia [38]. Al no tener que entrar en contacto físico con el entorno, se observan ventajas de optimización de tiempo y cuidado del medio ambiente, ya que no requiere de pruebas invasivas al cultivo. Algunos de los métodos más utilizados para la teledetección son los satelitales, cámaras espectrales, VANT e imágenes laser (*LIDAR, Light Detection and Ranging o Laser Imaging Detection and Ranging*), entre otros, en donde se pueden evaluar valores de temperatura, humedad y luminosidad [39]–[41].

2.1.5 Vehículos aéreos no tripulados

Los VANT, pueden ser controlados de forma remota y ofrecen una gran cantidad de aplicaciones en diferentes sectores como agricultura, seguridad, monitoreo ambiental, cobertura y desarrollo de vegetación, entre otros [42]. También, al combinarse con sensores de imágenes, rango y posicionamiento forman un sistema que brinda grandes ventajas y beneficios en las actividades agrícolas; al adquirir de forma rápida y fácil datos de campo para el monitoreo y manejo de recursos, nutrientes, herbicidas, entre otros [9], [11], [12].

En este trabajo de grado se hizo uso del VANT Solo de 3DR, el cual funciona con cuatro motores y cuatro hélices. A través del uso de un computador se controla la navegación, la altitud y las comunicaciones del vuelo; las entradas de control se reciben mediante la red de Wi-Fi segura de 3DR Link. Solo cuenta con un procesador de 1GHz, un motor de 880 KV, batería de 14.7 V 5200-mAh L polímero de Litio, dimensiones de 25 cm x 46 cm y un peso de 1,5 kg [43].

2.1.6 Espectro electromagnético

Es un conjunto de radiaciones electromagnéticas que difieren entre sí por su frecuencia. El espectro de un objeto se refleja y absorbe en la distribución característica de su radiación electromagnética [44].

2.1.7 Tipos de Imágenes espectrales

- Imagen multiespectral: es un tipo de imagen que captura datos dentro de rangos de longitud de onda específicos a través del espectro electromagnético. En la agricultura de precisión se utiliza la banda del infrarrojo cercano y el rojo visible [45].
- Imágenes hiperespectrales: es una imagen que posee varias bandas espectrales de información, contiguas entre sí, a través de todo el espectro electromagnético [46]. Estas imágenes tienen diferentes usos, tales como: identificación y vigilancia de cultivos, identificación de minerales, calidad y seguridad de alimentos, entre otros [19], [20].

2.1.8 Índices de vegetación

Los índices de vegetación (*VI, Vegetation Indexes*) son producto de varios valores espectrales que son operados para producir un simple valor que indique la cantidad o vigor de vegetación dentro de un píxel [15]. Se presentan a continuación algunos de los índices de vegetación más relevantes de acuerdo con el reporte de ADSU [47].

- Índice de vegetación de diferencia normalizada (*NDVI, Normalized Difference Vegetation Index*): el NDVI es útil para detectar la capacidad viva, verde o fotosintética de las plantas. Para esto mide la diferencia entre el infrarrojo cercano (que refleja fuertemente la vegetación) y la luz roja (que la vegetación absorbe). Es un índice útil para identificar rápidamente las áreas de vegetación y su condición. Este índice puede utilizarse para analizar los cultivos en todas las etapas del crecimiento [48]–[50].
- Índice de vegetación de diferencia normalizada verde (*GNDVI, Green Normalized Difference Vegetation Index*): el GNDVI es sensible a las variaciones de clorofila en el cultivo. Es útil para evaluar biomasa, nitrógeno, actividad fotosintética, madurez o estrés en las plantas. Para ello opera la banda visible verde y la banda infrarrojo cercano [51], [52].

- Índice de vegetación de cociente normalizado (*NRVI, Normalized Ratio Vegetation Index*): el NRVI reduce los efectos del suelo y la topografía, factores de luminosidad y los efectos atmosféricos [12].
- Índice de vegetación simple (*RVI, Ratio Vegetation Index*): el RVI es sensible a la topografía del terreno y poco sensible a la luminosidad [12].
- Índice de clorofila verde (*GCI, Green Chlorophyll Index*): el GCI se utiliza para estimar el contenido de clorofila de las hojas teniendo en cuenta las bandas verdes e infrarrojas cercanas [10].
- Índice de vegetación de clorofila (*CVI, Chlorophyll Vegetation Index*): el CVI presenta sensibilidad al contenido de clorofila. Se utiliza del principio a mitad del ciclo de crecimiento de un cultivo. La mayor sensibilidad del índice a la concentración de clorofila en la hoja se debe al trabajar con los colores rojo y verde [53].

2.2. Análisis bibliométrico

2.2.1. Fases metodológicas y herramienta de software

Para realizar el análisis bibliométrico se utilizó la herramienta de mapeo científico de código abierto SciMAT [30]. El mapeo científico permite a los científicos realizar análisis de la evolución del entorno científico [54] y también muestra cómo los campos, documentos y autores individuales se relacionan entre sí [55], [56]. El mapeo científico realizado se centró en estudiar los avances en el análisis de imágenes multiespectrales obtenidas por VANT de cultivos. La Ilustración 1 muestra los pasos seguidos, los cuales son descritos a continuación.

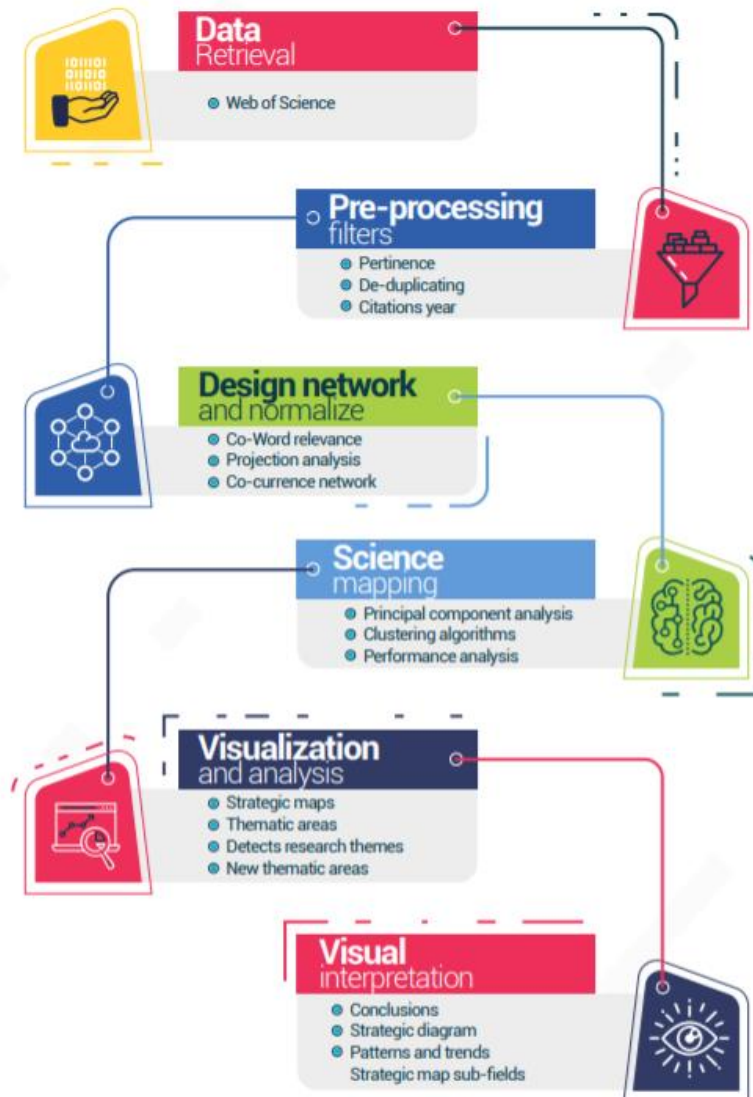


Ilustración 1. Flujo de trabajo del análisis bibliométrico. Imagen adaptada [30].

Recuperación de datos: esta investigación eligió la base de datos bibliográfica Web of Science (WoS) para construir un conjunto de datos, dado que es una de las bases de datos científicas más importantes. En este punto, la revisión entregó 10.999 documentos relacionados con la cadena de búsqueda "AGRICULTURA Y (ESPECTRAL O MUSLTISPECTRAL O HIPERSPECTRAL)", entre los años 2008 y 2019. Cabe resaltar que este revisión se realizó al inicio del trabajo de grado, en 2019, por eso los estudios se analizaron hasta ese año.

- Preprocesamiento: para realizar el preprocesamiento es necesaria la reducción de datos, para ello se establece un umbral de frecuencia mínima, dicho umbral es de tres citas, es decir, para que los documentos sean tenidos en cuenta, estos deben haber sido citados 3 veces.

- Diseño de red de unidades de análisis y normalización: se pueden establecer varias relaciones entre las unidades de análisis, estas se pueden representar como una red; donde autores, revistas, referencias, términos o documentos, se representan como nodos. Se estableció la red, en base a la concurrencia, con el fin de obtener aquellas que manejan relaciones por documentos, autores, términos y referencias.
- Mapeo científico: en la construcción del mapa, el algoritmo de agrupación permite identificar clústeres de temas y sus agrupaciones asociadas, incluidos los requisitos analizados y la información bibliográfica relacionada.
- Esquemas de visualización y análisis: los investigadores pueden observar el comportamiento de publicaciones, autores y referencias en diferentes áreas relacionadas con el uso de VANT y el análisis de imágenes multiespectrales, aplicadas en la agricultura. El análisis se realiza detectando patrones y tendencias en temas de investigación relacionados, para ello, los autores agrupan palabras y términos similares o estrechamente relacionados. Se elaboró un diagrama mostrando los grupos organizados por circunstancias, centralidad y densidad de las esferas en un plano cartesiano como se observa en la Ilustración 2. La centralidad mide el grado de conexión de una red con otra y el nivel de participación.
- Interpretación visual: los resultados fueron identificados de acuerdo con los temas, áreas, campos, disciplinas y esquemas teóricos sobre el conocimiento y la evolución científica del uso de VANT y análisis de imágenes multiespectrales aplicadas en agricultura.

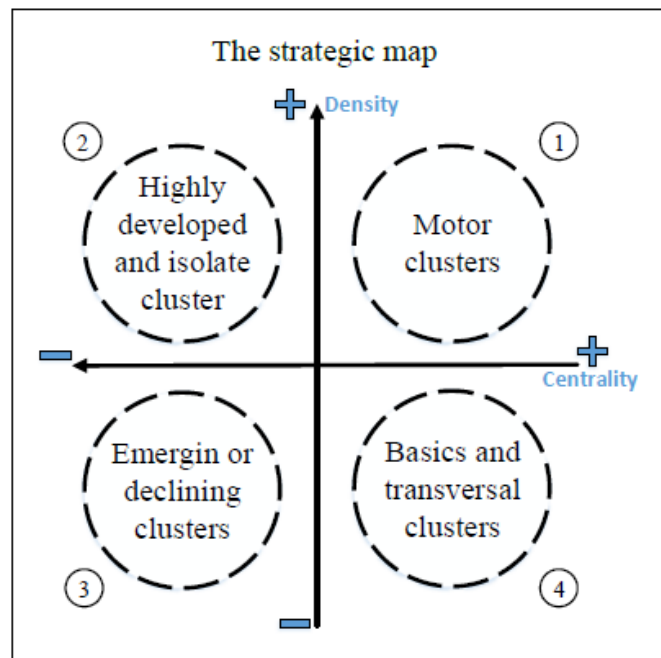


Ilustración 2. Comprensión y distribución del mapa estratégico. Fuente [30]

2.2.2. Búsqueda de conjuntos de datos

La cadena de búsqueda “AGRICULTURA Y (ESPECTRAL O MULTISPECTRAL O HIPERSPECTRAL)” ejecutada en el motor de búsqueda avanzada de la base de datos WoS entregó 10,999 documentos. Este corpus fue cargado en la herramienta SciMAT para desarrollar la metodología propuesta en el apartado anterior. La Ilustración 3, muestra el número de documentos resultantes de los filtros.

- Primer filtro: se seleccionaron los clústeres generados por la herramienta SciMat relacionados con el tema de investigación. Se eliminaron las referencias con menos de cinco citas, excepto las publicadas a partir del 2018. Dando como resultado un total de 856 documentos.
- Segundo filtro: se seleccionaron los títulos que estaban relacionados con el tema de investigación, obteniendo 372 documentos.
- Tercer filtro: el estudio fue enfocado en el uso de tecnologías como los VANT y las imágenes multiespectrales por lo cual se optó por eliminar aquellos que hicieran uso de imágenes satelitales o imágenes hiperespectrales.
- Cuarto filtro: se seleccionaron los documentos con mayor relación con el tema de investigación. De los 125 documentos obtenidos sólo uno fue realizado en Colombia, más no hablaba de café, por esa razón se procedió a realizar la búsqueda mencionada en la sección 2.3 Trabajos relacionados.

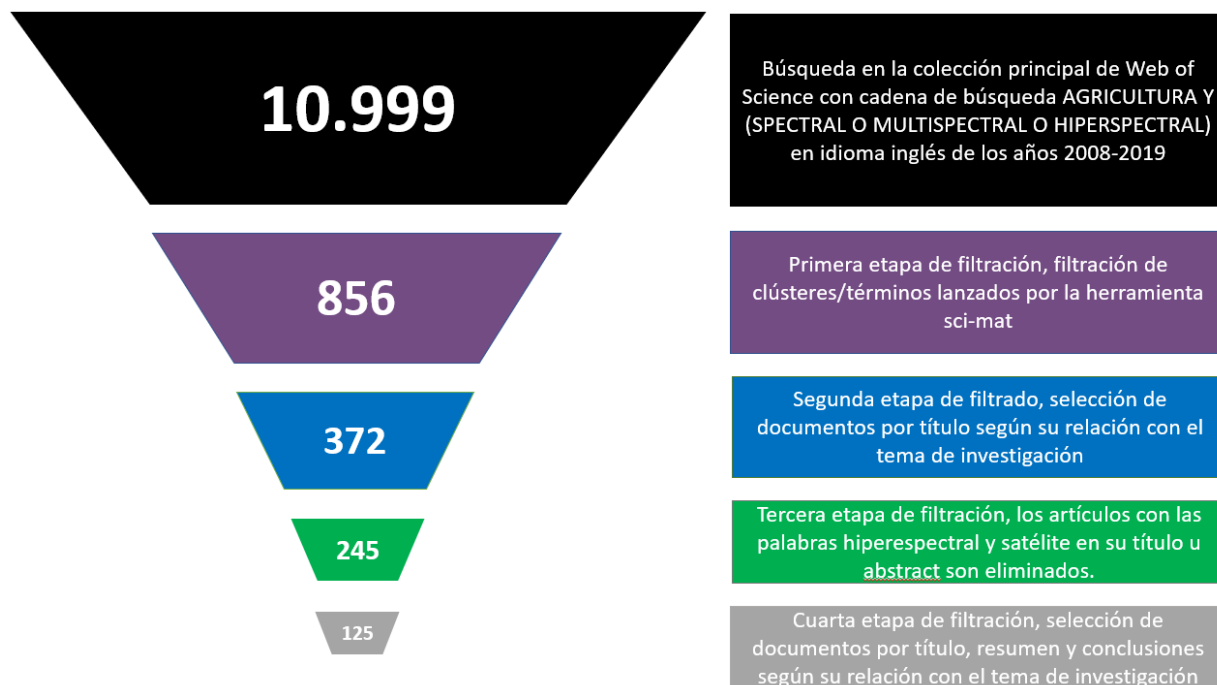


Ilustración 3. Filtro de etapas. Fuente propia

2.2.3. Análisis de resultados

En diciembre de 2015, la Convención Marco de las Naciones Unidas sobre el Cambio Climático (CMNUCC) firmó el Acuerdo de París [57], cuyo objetivo es mantener el aumento de la temperatura mundial por debajo de 2°C en los niveles preindustriales y limitar el aumento de temperatura a 1,5 °C. Por ello, el Panel Intergubernamental sobre Cambio Climático (IPCC) publicó el informe especial sobre calentamiento global de 1,5 °C, donde proponen la introducción de energías renovables, una mejor gestión de residuos y la reducción de la demanda energética, entre otras cosas, con el fin de fortalecer la respuesta global a la amenaza del cambio climático [58].

El cambio climático y las medidas que toman los gobiernos para enfrentarlo son fundamentales para la agricultura y su gestión. Por lo anterior, el acuerdo de París fue la referencia para dividir el análisis en dos periodos de tiempo, 2008-2015 y 2016-2019.

Según el contexto desarrollado se explican a continuación los mapas estratégicos presentes según los periodos de tiempo en la Ilustración 4 y la Ilustración 5.

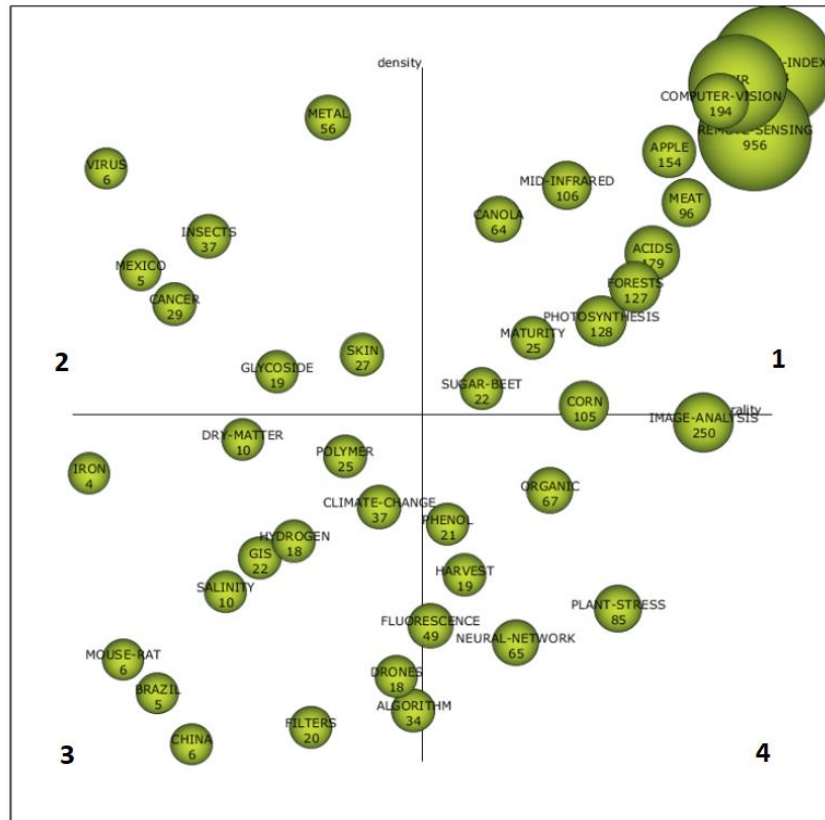


Ilustración 4. Primer periodo. Fuente propia

En el primer cuadrante se encuentran los temas motores, aparecen los términos “índices de vegetación” y “teledetección” los cuales fueron descritos en la sección 2.1 Definiciones y conceptos fundamentales. Términos como la “espectroscopía”, “infrarrojo medio” y “visión por computadora” representan tecnologías no invasivas utilizadas para el análisis de cultivos o terrenos en la agricultura, las cuales no contaminan ni destruyen el medio ambiente y son consideradas como rápidas, confiables y precisas [10], [59], [60]. La revisión bibliográfica realizada con los resultados arrojados por SciMat destaca el uso de estas tecnologías, por ejemplo, en el análisis de la calidad y madurez de la vegetación, además permiten monitorear componentes como el carbono, calcio y nitrógeno en diferentes terrenos, cultivos y alimentos [61]–[70]. También, aparecen palabras representativas de los cultivos de canola [71]–[75], carne [76]–[78], manzana [79]–[81], bosques [82]–[84], remolacha [85]–[87] o maíz [88], [89], los cuales son tipos de vegetación o alimentos altamente estudiados.

Los temas de alto desarrollo encontrados en el segundo cuadrante son más especializados. Por ejemplo se agrupan investigaciones en torno al metal, en donde se analizan el comportamiento de ligandos y complejos metálicos mediante la síntesis y caracterización de metales como el zinc, el cobre y el cobalto [90]–[92]. En concordancia con lo visto en el primer cuadrante, la

espectroscopía es ampliamente utilizada en diferentes campos. Por ejemplo, en la salud, es aplicable para detectar e identificar alteraciones y cambios en los tejidos biológicos relacionados con el cáncer y fármacos para combatirlo [93], [94].

Por otro lado, diversas investigaciones se centran en conocer la calidad y estado de salud en productos alimenticios, analizando si han sido afectadas por virus [95] o por insectos [96]–[99]. Partes importantes de análisis de las plantas para detectar dichas problemáticas, son la piel de los alimentos [100]–[102], las hojas y el suelo del cultivo; como se menciona en los términos encontrados en el primer cuadrante del segundo periodo.

La separación de periodos de tiempo están estrechamente relacionados con el cambio climático, los periodos de sequías pueden llevar a la muerte a muchos árboles, por ende, las imágenes multiespectrales pueden ayudar a encontrar tendencias para identificar las medidas de mitigación que se pueden realizar [103]. Además, aparecen componentes como el hierro [104], hidrogeno [105], [106] y el nitrógeno nuevamente, por ejemplo, se puede conocer o estimar la cantidad de nitrógeno presente en una planta al establecer una relación entre la materia seca y el contenido de nitrógeno en las hojas [107].

El tercer cuadrante recalca la importancia de la teledetección y los países que sobresalen en este campo investigativo, como lo son Brasil [108]–[111] y China [112]–[114]. Se destaca en la teledetección el uso de tecnologías como los sistemas de información geográfica [115]–[117], VANT [118]–[121] y las imágenes espectrales ya que permiten una fácil manipulación de una gran cantidad de datos y superar las limitaciones de las medidas en tierra. Las imágenes hiperespectrales en conjunto con el uso de filtros permiten crear un sistema para la captura de imágenes de alta calidad con parámetros óptimos de trabajo [122].

En el ámbito de la teledetección y la agricultura de precisión, la espectroscopía es bastante utilizada. Por ejemplo, se destaca la identificación de polímeros en objetos que precisan de conservación, es difícil por los métodos convencionales ya que son destructivos, por esta razón el uso de espectroscopia de infrarrojo cercano puede ser una solución viable para este proceso. Este mecanismo requiere una base de datos amplia para tener la capacidad de distinguir una mayor cantidad de polímeros, y tiene limitaciones para objetos con una reflexión limitada en la región NIR [123], [124].

Un término relevante en el tercer cuadrante es “algoritmo”. Los algoritmos pueden ser utilizados para un sinfín de propósitos. Algunos ejemplos de uso en la agricultura son: detectar la contaminación por Frass (excremento de insectos) [125]; estimar el volumen forestal, utilizando una combinación de datos ópticos remotos y de escáner laser [126]; estudiar las variaciones en la

fotosíntesis, el rendimiento y el crecimiento de la planta con diferente concentración de nitrógeno [127] e identificar unidades litológicas [128], [129].

También, el tercer cuadrante muestra el término “salinidad”, ya que en algunas ocasiones es necesario el uso de agua salada para hidratar los cultivos, esto puede generar reacciones negativas en las plantas dependiendo del nivel de tolerancia, por esta razón es importante conocer en qué momento el nivel de sal no es adecuado [130]–[132].

Como último término a analizar en el tercer cuadrante, se encuentra la palabra “rata” (mouse-rat). Al usar una técnica de imagen de reflectancia difusa multiespectral para calcular los coeficientes óptico corticales de los cerebros de rata expuestos, se logra conocer los cambios espaciotemporales de la hemodinámica cerebral y la morfología del tejido cortical al mismo tiempo, esto se podría aplicar en cirugía cerebral, para monitorear las funciones cerebrales [133].

Finalmente, en el cuarto cuadrante, el análisis de imagen es uno de los términos más abarcados. Este hace parte de todo sistema de teledetección, pero puede hacerse de muchas formas y para diferentes propósitos, como el cálculo de clorofila en las hojas de un cultivo a partir de imágenes RGB [134], o para mapear y caracterizar vegetación [135], [136]. También es posible identificar y caracterizar alimentos como la carne con la ayuda de imágenes multiespectrales [137]. Las imágenes y datos espectrales facilitan el trabajo de monitorear el estado de los cultivos en sus diferentes etapas, por ejemplo, al analizar lotes tanto en cosecha como en pos-cosecha se puede controlar el estado de envejecimiento de los lotes usando espectroscopia NIR [138], [139]. Por otro lado, las imágenes espectrales en conjunto con algoritmos de aprendizaje supervisado y no supervisado permiten identificar el estrés, maduración y calidad de un cultivo, entre otros parámetros [140]–[142].

En los algoritmos, las redes neuronales se pueden emplear para separar alimentos según su tiempo de maduración [143], identificar el estado de salud [144] o predecir el rendimiento y crecimiento de un cultivo; con datos de las malezas, elevación e imágenes multiespectrales [145]. Continuando con el análisis de alimentos, los métodos de fluorescencia tienen potencial para evaluar y cuantificar de forma rápida y no destructiva el estado de salud, maduración, calidad y parámetros microbianos de varios productos crudos y procesados. Todo esto con el fin de ayudar a la industria alimentaria a reducir costos y mejorar los controles y seguimiento de alimentos [146]–[150].

El suelo es un factor clave en la agricultura, por ejemplo, el carbono orgánico del suelo (COS) tiene una relación directa con la sustentabilidad de cultivos presentes en ese espacio, por esta razón es importante estimar los niveles de COS en el suelo con tecnologías rápidas y eficientes [64], [151]–[153].

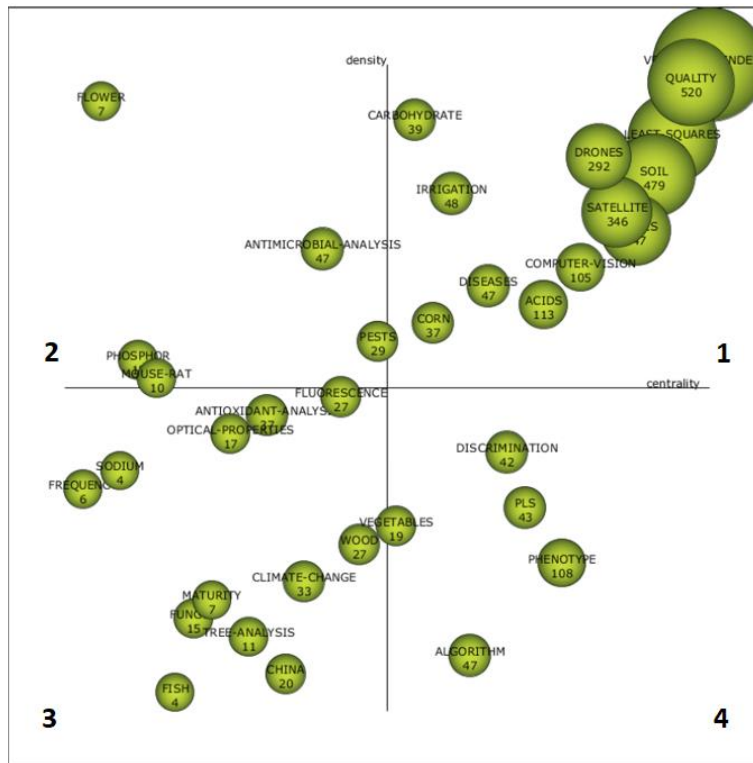


Ilustración 5. Segundo periodo. Fuente propia

En el primer cuadrante del segundo periodo, como método de análisis, se destaca la regresión de mínimos cuadrados parciales (PLSR). Este tiene diferentes aplicaciones en el contexto de la agricultura, principalmente para analizar componentes específicos de productos alimentarios [154]–[156]. Otra tecnología de alto uso en la teledetección aplicada a la agricultura son los satélites. Las imágenes satelitales tienen un gran potencial en el seguimiento y caracterización de enfermedades de cultivos e insectos que afectan grandes áreas [157], [158]. Al procesar dichas imágenes se puede obtener la caracterización y cuantificación de atributos de un terreno, optimizando así el trabajo y toma de decisiones de los agricultores [159]–[161].

La calidad, salud y seguridad de los alimentos son de gran importancia para las personas. Llevar un estilo de vida saludable con una nutrición adecuada requiere que los productos consumidos estén en excelentes condiciones y procurar la protección del medio ambiente, por ejemplo, al reducir el desperdicio de agua en el riego [60], [162], [163] o al detectar y predecir enfermedades de forma no invasiva [164]–[166]. La industria alimentaria busca crear productos de alta calidad utilizando tecnologías sostenibles [167]–[170] para evaluar el suelo del terreno o las hojas de los cultivos para estimar salud o componentes esenciales como carbono, nitrógeno y humedad [171]–[175].

En el segundo cuadrante, se encuentran temas tratados previamente, las plagas, en relación con las enfermedades e insectos, afectan la calidad y el estado sanitario de los cultivos, generando gran preocupación para la humanidad. Por lo tanto, varias investigaciones se han centrado en la creación y desarrollo de técnicas de detección más allá de los métodos tradicionales para proporcionar una evaluación rápida y no destructiva de la vegetación [99], [66], [100] - [102].

En el grupo “flower”, el girasol es uno de los términos más comunes. Un ejemplo es el uso de espectroscopía e imágenes VIS o NIR para analizar la salud de las hojas de girasoles de una manera no destructiva, rápida y consistente [176], [177].

Como se ha mencionado, el análisis de los componentes y elementos de un cultivo es un motor principal en la agricultura; el fósforo, el nitrógeno y el carbono, son elementos comunes para monitorear y analizar los cultivos [178], [179]. Temáticas aisladas, en donde también se analizan elementos particulares, como el estudio de la actividad antimicrobiana para aplicaciones en medicina [180],[181] y las propiedades antioxidantes de hierbas, alimentos y cultivos [182], [183]. Al aislar los componentes de los hongos, como las depsidonas y los metabolitos, se investigan compuestos obtenidos con propiedades antibacterianas, antioxidantes y citotóxicas, que ayudan a crear un arsenal contra bacterias resistentes a múltiples fármacos [180], [184], [185].

En el tercer cuadrante, de temáticas emergentes o en declive, se encuentran varios métodos y tecnologías aplicadas para el estudio de materia biológica. Investigadores analizan las propiedades ópticas dependientes del tiempo y del espectro del material biológico para evaluar su firmeza, salud o calidad [186]–[189]. El uso de la frecuencia con fines de investigación es extenso ya que permite evaluar e identificar el comportamiento de diferentes fenómenos y seres. Por ejemplo, la frecuencia puede utilizarse para rastrear cómo cambia el estado de la planta según las condiciones ambientales [190] para caracterizar y diferenciar cultivos [191].

Se pueden observar otros términos como la madera, los árboles y el pescado. A continuación, se expone lo principal de cada uno. Para la madera se suelen usar imágenes espectrales para analizar las modificaciones químicas y cambios en la composición de la madera cuando se expone a diferentes ambientes [192]–[195]. Por otro lado, diferentes investigaciones analizan los componentes y aspectos del árbol para caracterizar y diferenciar unas especies de otras. Además, pueden monitorear y estimar el estrés hídrico, índice de área foliar, clorofila y biomasa, entre otros [196], [197], [206], [198]–[205].

El pescado es un alimento muy demandado a nivel mundial, de ahí la importancia de analizar su calidad e inocuidad. La espectroscopia, las técnicas de aprendizaje automático, el análisis de imágenes multispectrales e hiperespectrales representan alternativas rápidas, precisas y no invasivas para monitorear, estimar o predecir factores relacionados con la calidad y salud de los peces [155], [207]–[210].

Por último, en el cuarto cuadrante, en el área de temas transversales, se pueden observar términos como hortalizas; tratando de analizar la calidad, inocuidad y desempeño de alimentos, como frutas y verduras, aplicando tecnologías como espectroscopia infrarroja media FT (FT-MIR) [211], imágenes espectrales [167] y luces LED [212]. Por otro lado, están los temas generales como la discriminación, el fenotipo y los algoritmos. Los algoritmos de aprendizaje automático como se ha mencionado anteriormente se han aplicado para analizar, discriminar y estimar diferentes factores y atributos de un terreno, vegetación o alimento. Para discriminar plantas enfermas de alimentos sanos, de baja calidad o alterados, diferentes investigaciones han analizado el estrés de la planta, su salud, madurez y calidad a través de imágenes hiperespectrales [166], [213]–[215]. En la aplicación de fenotipado de plantas de alto rendimiento [216], el uso de reflectancia espectral [217], [218], VANT [219]–[221], entre otras tecnologías, se desarrollan soluciones no intrusivas y con seguimiento de alto rendimiento de las características fisiológicas de las plantas.

Tabla 1. Cambio de cuadrante por temas. Fuente propia

Tema	Cuadrante 1° periodo	Cuadrante 2° periodo
VEGETATION-INDEXES	1	1
COMPUTER-VISION	1	1
MATURITY	1	3
CLIMATE-CHANGE	3	3
ALGORITHM	3	4
DRONES	3	1

En la Tabla 1 se observa el cambio de cuadrante por tema principal relacionado con el tema de investigación propuesto en este trabajo. El tema, índices de vegetación continúa siendo un tema motor al igual que la visión por computadora, lo que indica que son temas con relevancia y de interés investigativo.

La madurez, al pasar del 1° al 3° cuadrante demuestra ser un tema en declive por lo cual está dejando de ser un foco importante de investigación. Por el contrario, drones pasa del 3° al 1°, lo cual denota que es un tema en crecimiento que terminó siendo un tema motor, esto implica que es una temática en auge, con potencial crecimiento y que es un buen tema para seguir investigando. Cabe destacar que el factor cambio climático, fundamento de la división temporal, sigue en el mismo cuadrante (crecimiento o declive), al ser algo importante y de gran impacto al planeta se podría concluir que es un tema en crecimiento.

Por último, como se mencionó en conclusión a la Ilustración 3, ninguno de los 125 documentos resultantes estudiaba el café, y fueron realizados en Colombia, por esa razón se procedió a realizar la búsqueda mencionada en la sección 2.3. Trabajos relacionados.

2.3. Trabajos relacionados

Adicionalmente al análisis bibliométrico realizado se ejecutó la búsqueda bibliográfica con el filtro de búsqueda “*Multispectral and Coffee*” sin restricción de años, en las bases de datos WOS y Scopus, se encontraron 21 y 30 documentos respectivamente, de los cuales se descartaron aquellos documentos que no se relacionaban con el tema de investigación o estaban duplicados, obteniendo así, solamente 22 estudios que analizaban cultivos de café a través de imágenes multiespectrales. De estos estudios ninguno fue realizado en Colombia, como puede observarse en la Ilustración 6.

Por esta razón, se expande la investigación al revisar los 536 resultados de la cadena de búsqueda “Multispectral and Coffee and Colombia” en Google Scholar, sin embargo, solamente se encontraron dos estudios enfocados en cultivos de café en Colombia [6], [222], en los cuales, ninguno utilizó una cámara multiespectral comercial, probada a nivel industrial, sino una construcción científica en laboratorio basada en Raspberry Pi² [6] e imágenes satelitales [222]. Cabe destacar que en dicha búsqueda se encontraron investigaciones de imágenes multiespectrales en Colombia de otros cultivos que contribuyeron a la construcción de esta sección. A continuación, se presenta la descripción de los trabajos más destacados en el estudio de imágenes multiespectrales de cultivos agrícolas, clasificados en: realizados en Colombia y realizados en otros países.

² Raspberry Pi es un minicomputador de placa simple, creado inicialmente para incentivar la enseñanza de la informática.

Filtro "Multispectral and coffee" Wos y SCOPUS

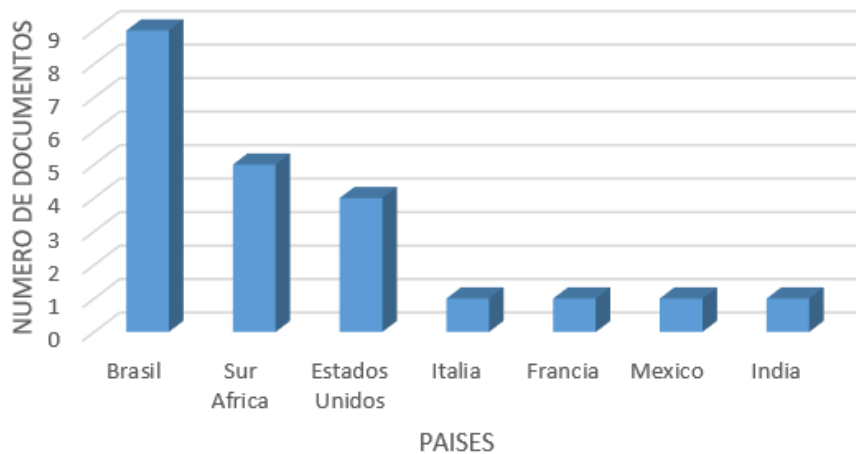


Ilustración 6. Documentos de WOS y Scopus sobre imágenes multiespectrales de cultivos de café. Fuente propia

2.3.1 Estudios realizados en Colombia

En esta sección se presentan estudios de cultivos agrícolas que hacen uso de VANT e imágenes multiespectrales realizados en Colombia.

2.3.1.1 Estudios de café

- En la investigación de J. M. Castrillón Cuervo *et al.*, [223] desarrollaron un algoritmo de segmentación (lineal, bayesiano y vecinos más cercanos) de granos de café a través de la adquisición de imágenes de granos sanos y defectuosos. Determinaron las mayores diferencias entre tres clases de granos, sanos, fermento e inmaduro. A diferencia de esta investigación, nosotros nos enfocamos en las hojas de los cultivos de café, no en el grano ni su estado de maduración.
- En la investigación de C. Castrillón *et al.*, [224] desarrollaron un algoritmo que permite segmentar los granos de café lavado e identificar si han sido afectados por la broca del café al analizar la representación del color de crominancia azul y roja. En comparación nuestro trabajo no evalúa la salud del cultivo de café a través del análisis del grano, sino de la hoja.
- En el trabajo de J.A Bolaños *et al.*, [6] se desarrolló un sistema con una Raspberry pi para la toma de imágenes de cultivos del municipio de Timbío. Los investigadores calcularon el índice de vegetación NDVI, con el fin de estimar la salud de dicho cultivo; para ello crearon un sistema desarrollado en laboratorio con una Raspberry pi cámara NOIR (RGB+NIR) y una Raspberry PI Cámara RGB. Cabe resaltar las siguientes diferencias: se calculó más de un índice de vegetación, no sólo el NDVI y se analizó la correlación de dichos índices con niveles de

clorofila, se usó una cámara multispectral comercial para corroborar la estimación de salud realizada con modelos de aprendizaje automático, entre otros.

- En la investigación de S. Bolaños et al., [222] se mapeó y clasificó el área de diferentes sistemas de café (monocultivo, policultivo y sistemas agroforestales) en el departamento de Huila. Realizaron una clasificación supervisada utilizando el algoritmo de vecino más cercano en un enfoque basado en objetos; para ello utilizaron la información espacial y espectral obtenida al integrar imágenes de satélite y mapas topográficos a través de sistemas de información geográfica (SIG).

2.3.1.2 Estudios en otros cultivos diferentes al café:

- En el artículo de J.P Rojas et al., [225] desarrollaron un sistema (Hardware y Software) para la toma y el procesamiento de imágenes multispectrales de cultivos de arroz en sus diferentes etapas y para dos tipos diferentes de siembra: Santa Rosa para tierras bajas y Palmira para tierras altas. Para esto construyeron un mosaico georreferenciado con las imágenes capturadas y calcularon los índices de vegetación (RVI, NDVI, GNDVI, DVI, CTVI, TVI y MSAVI).
- En el trabajo de C.A. Devia et al., [226], los autores calcularon 7 índices de vegetación (RVI, NDVI, GNDVI, DVI, TVI, CTVI y MSAVI) para estimar la biomasa aérea (materia orgánica viva por encima del suelo [227]) del arroz en diferentes etapas de cultivo. También compararon los resultados de los métodos K-Means y redes neuronales artificiales (*ANN, Artificial Neural Network*) al momento de realizar las estimaciones dentro del algoritmo general de búsqueda estocástica. Nuestro estudio estima la salud, no la biomasa de cultivos de café a través del uso de regresiones lineales y redes neuronales artificiales con datos físicos de clorofila y los índices de vegetación.
- El trabajo de J. Navia et al., [8] presentó un enfoque de la teledetección y la agricultura de precisión para monitorear campos de cultivo de arroz en Bogotá. Gracias a este enfoque los autores proporcionaron a los agricultores una herramienta integrada para medir y evaluar la vegetación verde viva mediante el ensamblaje de un mosaico de imágenes multispectrales del terreno.
- En la investigación de A.R Johnsen et al., [228] fueron evaluados los problemas más comunes que suceden durante el desarrollo del cultivo de papa en el departamento de Cundinamarca tales como, estrés por exceso de sales minerales y exceso de humedad en el suelo. Para esto, utilizaron la respuesta espectral de los cultivos obtenida a partir de imágenes de infrarrojo cercano (NIR) de alta resolución. Los resultados obtenidos demostraron que la respuesta espectral permite identificar características

de la vegetación y problemas en el cultivo de papa.

- A.K Torres Galindo et al., [229] desarrollaron un sistema con una Raspberry Pi B capaz de capturar imágenes multiespectrales en las bandas RGBN (rojo, verde, azul e infrarrojo cercano) y una aplicación que permite calcular el índice de vegetación NDVI para estimar la salud de la vegetación de la palma de aceite de la altillanura de la Orinoquía.

2.3.2 Estudios realizados en otros países

- P. Pereira Coltri et al., [230] utilizaron imágenes multiespectrales del satélite GeoEye-1 con bandas RGB para obtener índices de vegetación como GNDVI y NDVI. Entre sus resultados, se destaca que encontraron una buena correlación entre las bandas y la biomasa del café, y que esta relación es mucho más fuerte al trabajar con los índices de vegetación. De manera similar concluyeron que se puede tener una correcta estimación del stock de carbono en los cultivos de café de Minas Gerais, Brasil.
- En el estudio de A. Chemura et al., [231] realizado en “Jersey Tea and Coffee Estates” en el Distrito de Chipinge, Zimbabwe demostró el valor de los datos multiespectrales de Sentinel-2 MSI y su relación con los índices de vegetación en el modelado de nitrógeno foliar de café a escala de paisaje. Identificaron las variables que mejor se desempeñaron en la predicción de N para mapear y medir los niveles que pueden usarse para la caracterización y administración de cultivos de café con el fin de lograr una excelente productividad.
- La investigación de D.B Marin et al., [232] evaluó el potencial del sensor multiespectral OLS / TIRS Landsat 8 para la identificación y monitorización espacial y temporal del café (*Coffea arabica*) afectado por el tizón bacteriano en un campo de café comercial en el estado de Minas Gerais, Brasil.
- L.F. Johnson et al., [233] recolectaron imágenes multiespectrales con un VANT Pathfinder-Plus sobre un cafetal comercial de Hawai durante la temporada de cosecha 2002. Estas imágenes fueron utilizadas para monitorear el estado de madurez del cafetal. El índice de madurez medio por campo fue correlacionado significativamente con los recuentos en tierra registrados por el productor y con la fecha de cosecha final.
- El objetivo del estudio de C. Gómez et al., [234] fue identificar áreas adecuadas para la hibridación interespecífica del café en Nueva Caledonia. Para esto, los investigadores generaron un mapa predictivo al entrenar un sistema de red neuronal, esta red utilizó parámetros ambientales basados en el campo, predictores de detección remota mediante imágenes pancromáticas (blanco y negro) de Quickbird e

imágenes multiespectrales.

- El estudio realizado por F. S. Kawakubo et al., [235] en Minas Gerais, Brasil tuvo como objetivo contribuir al mapeo y clasificación de los cultivos de café (café de producción, café mixto y café viejo) utilizando imágenes multiespectrales con una resolución espacial de 23.5 m. Las imágenes multiespectrales fueron tomadas por el instrumento Linear Imaging Self Scanner a bordo del sistema satelital Indian Remote Sensing.
- El estudio realizado por A. Chemura et al., [236] en Chipinge, Zimbabwe evaluaron el potencial de Landsat 8 OLI para desarrollar mapas temáticos específicos por edad del café. Concluyeron que es útil para el desarrollo de dichos mapas y que la banda NIR de OLI fue la banda más importante en la discriminación intraclase del café.
- El estudio realizado en Brasil de G. D. Martins et al., [237] identificó la presencia de nematodos a través de mediciones biofísicas y espectrales de las hojas de café; ninguno de los parámetros biofísicos empleados como: Índice Foliar, biomasa y clorofila discriminaron eficientemente el estado de salud de las hojas. Por el contrario, la clasificación multiespectral (rojo, borde rojo e infrarrojo cercano) definió la distribución espacial de plantas de café sanas, moderadamente infectadas y severamente infectadas a través de con una precisión general del 78% y un coeficiente Kappa de 0,71.

2.4 Conclusiones del estado del arte

Tabla 2. Resumen comparativo de trabajos relacionados. Fuente propia.

Artículo	Ubicación		Cultivo		Fuente de dato		
	Colombia	Exterior	Café	Otro	Cámara multiespectral	Satelital	Otro
[222]	X		x			x	
[223]	X		x				x
[224]	X		x				x
[225]	X			x	x		
[8]	X			x	x		
[226]	X			x	x		
[229]	X			x			x
[6]	X		x				x
[228]	X			x	x		
[230]		x	x			x	
[231]		x	x			x	
[232]		x	x			x	
[233]		x	x		x		
[234]		x	x			x	

[235]		x	x			x	
[236]		x	x			x	
[237]		x	x			x	
Este estudio	X		X		X		

Deducciones obtenidas a partir de la Tabla 2, la cual presenta de forma sintetizada lo encontrado en los trabajos relacionados:

En Colombia se encontraron pocos estudios realizados sobre cultivos de café; pero, a diferencia de nuestro estudio, estos no realizan un análisis de datos de los índices de vegetación y clorofila con técnicas de aprendizaje automático (regresiones líneas y redes neuronales artificiales) para estimar la salud.

A nivel internacional, se encuentran varios trabajos científicos sobre el estudio de cultivos de café. No obstante, la mayoría de estos estudios hacen uso de imágenes satelitales [230]–[232],[234]–[237]. En contraste, el presente estudio pretende usar imágenes capturadas por cámaras multiespectrales transportadas en VANT. Por último, este trabajo presenta un estudio del estado de salud de cultivos de café en el departamento del Cauca, Colombia a través de imágenes capturadas por la cámara multiespectral Survey 3 transportada en VANT.

En este capítulo se explicaron los pasos realizados para elaborar el estado del arte, el análisis bibliométrico realizado en la herramienta SciMat y la descripción de la base conceptual para entrar en contexto en torno a los conceptos básicos y artículos relacionados con la agricultura, las cámaras multiespectrales y los VANT.

CAPÍTULO 3

3. Materiales y métodos

A continuación, se describen los materiales utilizados para el desarrollo de la presente investigación.

3.1. Materiales

Los materiales y equipos utilizados en el desarrollo del trabajo de investigación serán mencionados y descritos a continuación:

- **Computador:** Un computador portátil HP con 4GB de memoria RAM y procesador AMD A6-5200 APU Radeon (TM) HD Graphics 2.00 GHz, en el cual se realizaron todos los procesamientos correspondientes en Python.
- **Cámara MAPIR Survey 3:** La cámara incluye un receptor GPS USB externo para geo-etiquetar automáticamente cada imagen capturada. Su sensor de 12MP, como se ve en la Tabla 3, y su lente nítido sin ojo de pez, le permite capturar fácilmente medios aéreos. Captura fotos en el disparador predeterminado del temporizador o se puede enviar una señal PWM a través del puerto HDMI. Este modelo cuenta con bandas Rojo Verde + NIR (RGN, NDVI), por lo cual ve luz infrarroja cercana de 850 nm, roja de 660 nm y verde de 550 nm [238].

- Tabla 3. Especificaciones cámara MAPIR Survey 3. Fuente [238]

Resolución	12 MP (4,000 x 3,000 px), 8MP
Óptica	87° HFOV (19mm) f/2.8
Velocidad de captura	RAW+JPG:2.75 segundos/foto. JPG: 1.5 segundos/foto.
GSD	5.5 cm/px (2.17in/px) a 120 m (~400 ft) AGL
Peso	76g con batería
Dimensiones	59 x 41.5 x 36mm
Batería	Removable Li-ion (1200mAh)
Almacenamiento	Micro SD (128GB) (64GB Card ≈ 15,000 JPG, 2,200 RAW+JPG)
Conectividad	PWM vía HDMI
GPS	ublox UBX-G7020-KT

- **Vehículo aéreo no tripulado:** SOLO 3DR es un VANT pequeño que funciona con cuatro motores y cuatro hélices. Este contiene módulos que controlan la navegación, la altitud y las comunicaciones del vuelo; además,

posee unas entradas de control mediante la red de Wi-Fi segura de 3DR Link. Cuenta con un procesador de 1GHz, un motor de 880 KV, batería de 14.7 V 5200-mAh L polímero de Litio, dimensiones de 25 cm x 46 cm y un peso de 1,5 kg [43].

- **QGIS:** es un sistema de información geográfica con licencia GNU. Es un proyecto oficial de Open Source Geospatial Foundation (OSGeo) [239].
- **Agisoft Metashape:** es un software fotogramétrico que implementa algoritmos de procesamiento de imágenes, con técnicas de fotogrametría digital para tareas de reconstrucción, visualización, topografía y mapeo 3D con el fin de desarrollar nuevas herramientas [240].
- **Colaboratory:** es una herramienta que permite ejecutar código en la nube con lenguaje Python, dando acceso a instalaciones fáciles y acceso gratuito a GPU [241].
- **Programa Mission planner:** es una aplicación gratuita de código abierto que se puede utilizar como una utilidad de configuración o como un complemento de control dinámico para vehículos autónomos [242].
- **Clorofilómetro CCM 200-PLUS:** es un aparato compacto, manual, abastecido por una pila alcalina de 9V. El CCM 200-PLUS cuenta con una sección para la evaluación de muestras localizada en la parte superior del instrumento y una pantalla de cristal líquida (LCD) en donde se ve el valor del índice de contenido de clorofila (*CCI, por sus siglas en inglés*). Este dispositivo representa ser una herramienta útil para estimar de manera no destructiva las concentraciones de clorofila [243]–[246]. Las especificaciones se encuentran en la Tabla 4.

Tabla 4. Especificaciones del clorofilómetro CCM 200-plus [249].

Parámetros de medida	653 nm (Clorofila) & 931 nm (NIR).
Area de medida	9.52 mm (3/8") diámetro (71 mm ²)
Resolución	1 unidad CCI
Detector	Fotodiodo de silicio con amplificador integral para medición de absorbancia, monitoreo de potencia y compensación de temperatura.
Almacenamiento	8 Mbytes de memoria para hasta 160.000 mediciones de datos, o 94.000 con entradas de datos GPS adicionales.
Modos	Un solo punto, seleccionable por el usuario con promedios de 1 a 30 puntos y un protocolo estadístico de 10 a 30 puntos que solicita reemplazar los puntos de datos más allá de una desviación estándar de 2 sigma.
Interfaz de usuario	Pantalla gráfica de 128 x 32 píxeles, señales de pitido, teclas para: configuración, protocolos de

	medición, diagnóstico, control de datos y entrada de líneas de comentarios alfanuméricos por parte del usuario.
I/O:	USB 1.1 y RS-232; salida de archivos multipunto, salida única bajo demanda, compatible con NMEA 0183 para la entrada de datos GPS.
Rango de temperatura	0-50 °C.
Fuente de poder	9V batería alcalina
Intervalo de auto apagado	4 minutos
Tamaño	152(L) x 82(W) x 25(D) mm

3.2. Diseño experimental

Para el diseño experimental se adaptó el modelo CRISP-DM [247]. Este modelo consta de 5 fases, las cuales se pueden observar en la Ilustración 7. La primera fase, involucra la comprensión del negocio y entender los objetivos del proyecto; la segunda fase, plantea la recolección de los datos; la tercera fase, la preparación de los datos, lo cual implica un primer acercamiento en el estudio y entendimiento de los datos obtenidos y la cuarta fase, corresponde a las técnicas y el procesamiento aplicado a los datos, los cuales se evalúan en la quinta fase (Evaluación).

El color azul hace referencia al estado del arte, el color verde agrupa las fases de análisis y preparación de los datos y el color gris, las etapas de modelación y evaluación.

1. Comprensión del negocio	Esta primera fase se enfoca en la comprensión de los objetivos de proyecto.
2. Estudio y comprensión de los datos	Esta fase inicia con la recolección de datos inicial.
3. Análisis de los datos y selección de características	Esta fase incluye las actividades necesarias para la preparación de los datos, es decir, la transformación de los datos que se usarán.
4. Modelado	En esta fase, se seleccionan y aplican las técnicas de modelado.
5. Evaluación	Evaluación de las técnicas de modelado

Ilustración 7. Etapas modelo CRIPS-DM. Fuente [247].

En la Ilustración 8, se observan las etapas del diseño experimental adaptado uniendo algunas fases del modelo CRIPS-DM. En la primera etapa “Captura de datos”, se encuentra el entendimiento de la problemática que se ve reflejada en

el estado del arte; la segunda etapa “Preprocesamiento de datos”, corresponde a la unión de la segunda y tercera fase del modelo CRISP-DM, en la cual se capturan y recolectan los datos, se filtran y se generan datos sintéticos, con el fin de obtener el conjunto de datos final con los que serán entrenados los modelos y la tercera etapa “Modelado de datos”, representa la implementación de modelos de aprendizaje automático y la evaluación del error de los mismos.

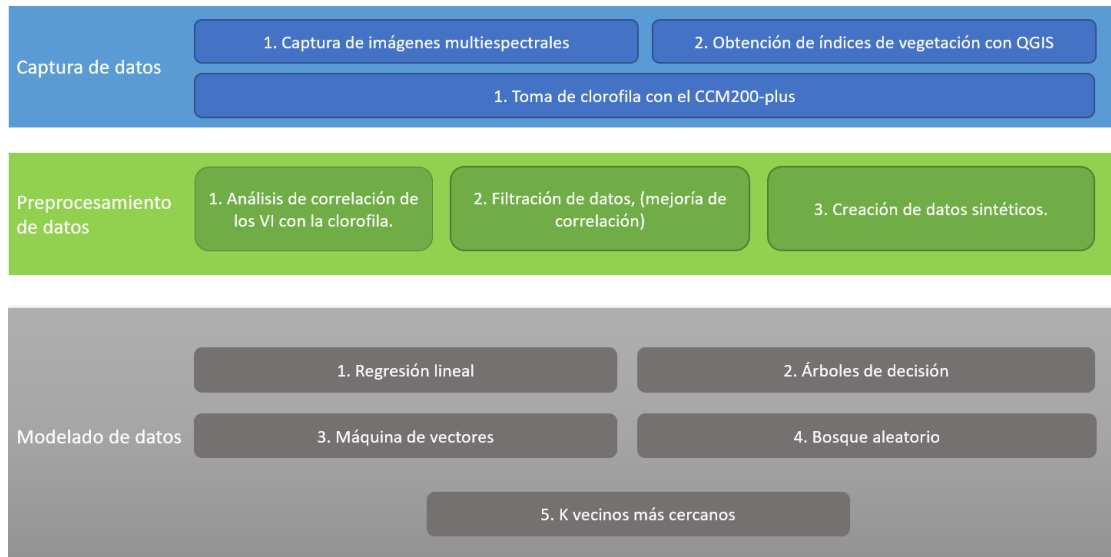


Ilustración 8. Etapas diseño experimental. Fuente adaptada de [247].

3.2.1. Captura de datos

Para la captura de los datos, se realizaron tres procesos como se observa en la Ilustración 9. El primer paso fue la programación del vuelo del VANT en el software Mission Planner y la captura de las imágenes multiespectrales por medio de la cámara mapir survey3 aerotransportada en el VANT Solo 3DR (Este VANT tiene compatibilidad con el programa *Mission Planner* y permite configurar la captura automática de fotos); el segundo paso fue el procesamiento de las imágenes en los programas *agisoft metashape* y QGIS para la obtención de los índices de vegetación y en el tercer paso se tomaron muestras de clorofila en las plantas de café con el clorofilómetro CCM-200 PLUS, el cual fue usado como método no destructivo para determinar el contenido de clorofila en las plantas [248], [249].

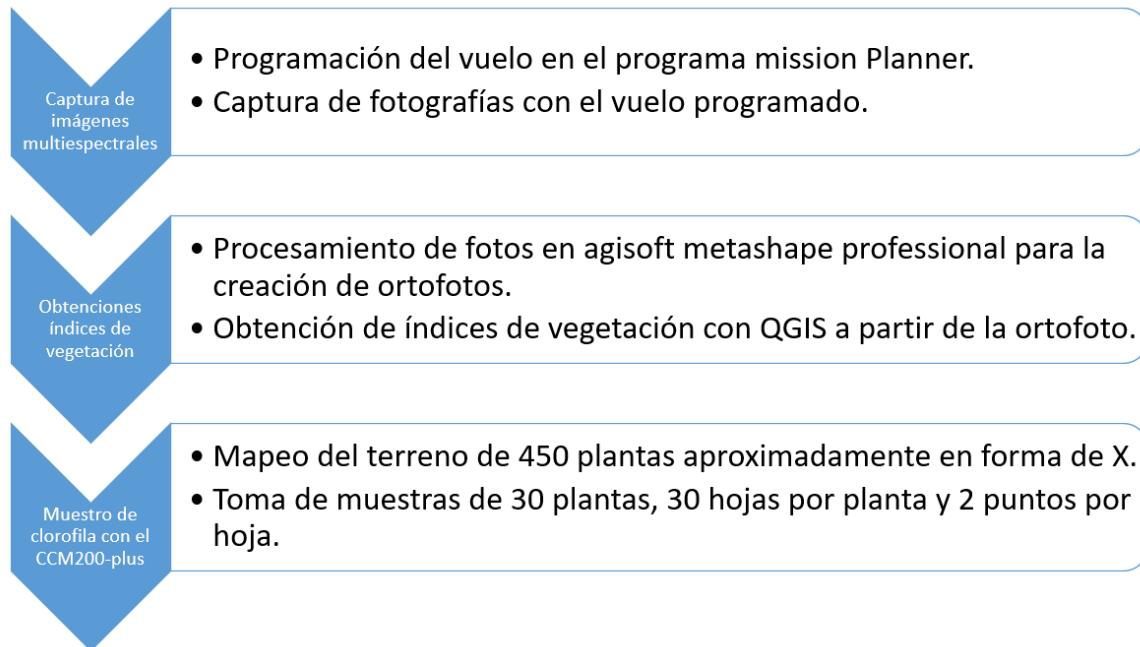


Ilustración 9. Metodología para la captura de datos. Fuente propia.

3.2.1.1. Metodología para la captura de imágenes multiespectrales

Existen diferentes parámetros que intervienen en el proceso de captura de las imágenes usando VANT, los cuales están definidos por el alcance de las herramientas utilizadas y su acoplamiento. Entre los más importantes se destacan la altura, autonomía y velocidad del vuelo del VANT, el intervalo de captura de la cámara, el área a cubrir, y la resolución de la cámara [250].

Para acoplar la cámara Mapir Survey con el VANT y generar la captura programada en el vuelo, se siguieron los pasos de la guía “Activar la cámara Survey2 usando la señal PWM de 3DR SOLO” que se encuentra disponible en línea [251]. Para el montaje físico mostrado en la Ilustración 10, se utilizaron amarras para evitar movimientos de la cámara que produzcan distorsiones en las imágenes, además se acopló un módulo GPS para habilitar la realización de vuelos autónomos. De acuerdo con la guía mencionada, se utilizó la herramienta Mission Planner para la configuración del vuelo del VANT, en la cual se generó una misión que luego fue cargada al VANT 3DRSolo.



Ilustración 10. Montaje del VANT. Fuente propia

Para realizar una correcta recolección de información a partir de imágenes es importante tener en cuenta las limitaciones del sistema, como lo es: la resolución, distancia focal, velocidad de captura, entre otros [252]. En el presente estudio para establecer el plan de vuelo utilizaron las siguientes características físicas de la cámara Survey3W.

- Ancho del sensor: 6.2496mm
- Alto del sensor: 4.6872mm
- Distancia focal: 3.37mm
- Resolución: 12MP (4032*3024 pixeles)
- Velocidad de captura: 0.00575 segundos

Con los anteriores parámetros fue calculada la distancia de muestreo del suelo (GSD por sus siglas en inglés), la cual es la distancia entre el centro de dos pixeles medidos en el suelo. Entre mayor sea la distancia GSD habrá mayor cobertura en la foto; pero, con menor detalle; y entre menor sea la distancia GSD, la cobertura será menor; pero, con mayor detalle en cada fotografía. La Ecuación 1 representa dicha relación

$$H_m = \frac{ImW * GSD * F}{SW * 100}$$

Ecuación 1

Donde:

H: altura del vuelo en metros

ImW: ancho de la foto en pixeles

GSD: distancia de muestreo del suelo en centímetros/pixel

SW: ancho del sensor en milímetros

F: distancia focal en milímetros

A partir de la Ecuación 1 se obtiene el valor de GSD como se puede observar en la Ecuación 2.

$$GSD = \frac{SW * 100 * H_m}{ImW * F}$$

Ecuación 2

Los valores más comunes de GSD utilizados en la agricultura de precisión varían de 2cm a 15cm [253]. Tomando estos valores de referencia, la altura mínima sería de 44m y la altura máxima sería de 327m. Sin embargo, a pesar de tener valores de referencia, estos no fueron aplicados. En primera instancia, porque la altura máxima de 327m excede la altura legal permitida en Colombia regulada por la Unidad Administrativa Especial de Aeronáutica Civil de 123m [254]. Por otro lado, la altura mínima utilizada fue de 25m (GSD =1.2cm), esto con el fin de buscar la mejor resolución en las fotos teniendo en cuenta los obstáculos del terreno (otro tipo de árboles). Cabe resaltar que el terreno de análisis es un cultivo de café de variedad Castillo ubicado en una finca en San Joaquín, Tambo, Cauca.

Al delimitar el área del cultivo de café, sobre la cual se efectúa el plan de vuelo automático en Mission planner (área roja de la Ilustración 11), éste no permitió una altura superior a 80m. Por dicha razón, y para no estar sobre el límite establecido por el programa, la altura máxima usada fue de 75m.

Después de determinar las alturas del vuelo (25m y 75m), Mission Planner genera la mejor misión de vuelo teniendo en cuenta las especificaciones de la cámara, como por ejemplo, la altura, un traslape frontal de 80% y un traslape lateral de 75% [255], [256]. Las misiones de vuelo generan puntos de espera (puntos verdes) con coordenadas en los cuales el VANT enviará una señal PWM a la cámara, para que esta tome una fotografía. Este proceso de configuración de vuelo se describe detenidamente en el ANEXO A, además las imágenes multiespectrales capturadas en los vuelos se encuentran disponibles en un repositorio de GitHub³.

³ Repositorio de Github disponible en [313].



Ilustración 11. Puntos de polígono (rojos) y puntos de control (verdes).
Fuente propia.

Las imágenes multispectrales obtenidas fueron procesadas en la herramienta software *agisoft metashape* para la creación del ortomosaico (representación proyectada ortogonal del terreno en forma de foto [257]). El proceso necesario para generar el ortomosaico, a partir de las imágenes multispectrales individuales, se encuentra descrito detalladamente en el ANEXO B.

3.2.1.2. Obtención de los índices de vegetación

Los índices de vegetación están relacionados con las condiciones de un cultivo, al analizar los valores espectrales de las diferentes bandas (rojo, verde, infrarrojo cercano) se puede determinar el vigor, nivel de clorofila y/o salud de la vegetación en determinado píxel de la imagen. Por ejemplo, se destaca el NDVI, en donde los valores fluctúan entre -1 y +1; a mayor valor numérico, más vigor tendrá la cobertura, como se encuentra representado Ilustración 12 [258]. Cabe destacar que los índices de vegetación obtenidos son: NDVI, GNDVI, RVI, GCI, NRVI y CVI.

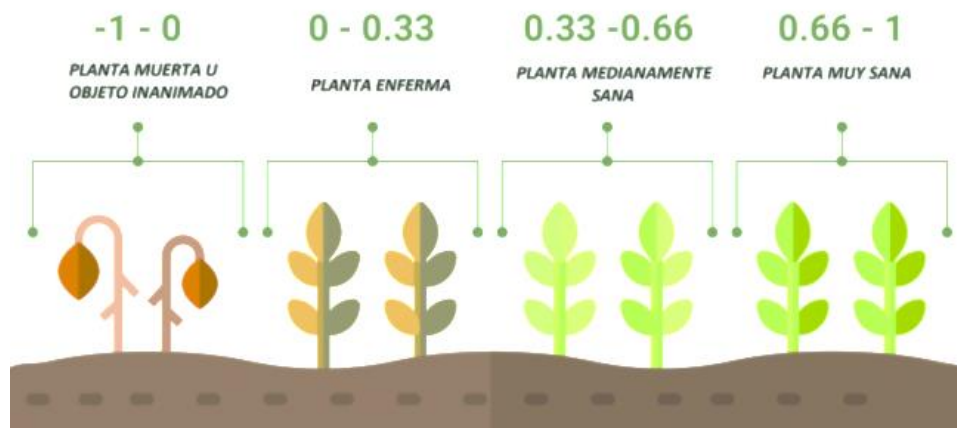


Ilustración 12. Nivel de salud por NDVI. Fuente [258].

Para obtener los índices de vegetación NDVI, GNDVI, RVI, GCI, NRVI y CVI se cargaron las ortofotos en el programa QGIS, se dibuja un recuadro de cada árbol muestreado (figura roja) del terreno, como se observa en la Ilustración 13. Además, se ve el cultivo de café segmentado en 3 secciones, 3 hileras con 5 filas y cada fila con 30 cafetos, dando un total de 450 cafetos. El mapeo del cultivo del terreno se describe detalladamente en la sección 3.2.1.3 Muestreo de clorofila, en donde se explica la selección del mapeo en forma de X y el número de muestras a tomar.

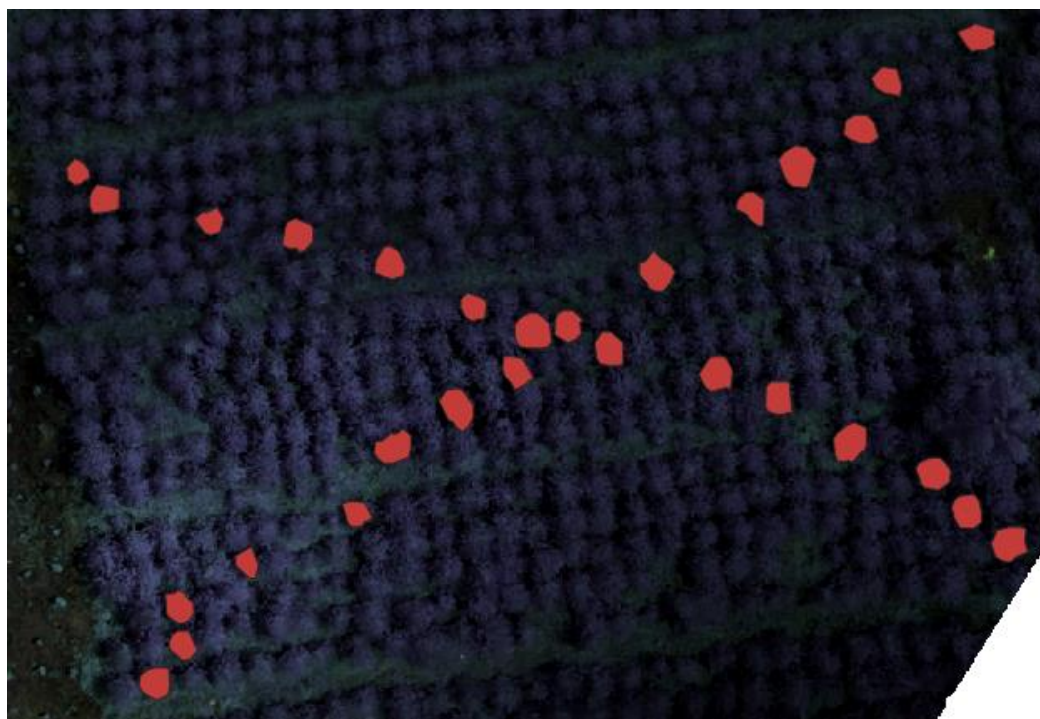


Ilustración 13. Mapeo del cultivo del terreno. Fuente propia.

Posteriormente, se agregaron las fórmulas correspondientes a los índices de vegetación, en el apartado de '*raster calculator*' del software QGIS (La configuración del software QGIS está descrita en el ANEXO C). Como se observa en la Ilustración 14, en el panel de operaciones se colocan las ecuaciones relacionadas al cálculo de los índices de vegetación. Los valores calculados para los índices de vegetación seleccionados (NDVI, CVI, GNDVI, NRVI3, NRVI, RVI y GCI), se anexan al valor de clorofila promedio por planta, lo cual se ve representado por la columna MEDIA de valor 1. Este valor es resultado de promediar los valores de clorofila por hoja de la planta, que se encuentran en la columna MEDIA con valor igual a 0. Además, están los datos de fecha y hora de la muestra tomada y del vuelo del VANT. Igualmente se registra la zona donde fue tomada la muestra, la cual indica la sección en donde se tomó dicha muestra, si es de la parte alta o media como se representa en la Ilustración 18 y se explica detalladamente en la sección de filtrado de datos. Cada muestra posee identificadores únicos que se sacan según el vuelo, fila y columna de la planta. Los resultados obtenidos y las columnas descritas se observan en la Tabla 5, y los datos completos de los vuelos se encuentran en el ANEXO D.

Tabla 5. Valores índices de vegetación. Fuente propia

ID	VUELO	FILA	PLANTA	MEDIA	HORA CLL	FECHA	HORA VUELO	ZONA	CLL (CCI)	NDVI	CVI	GNDVI	NRVI	RVI	GCI	NRVI3
112	1	1	2	1	7:00	02-06-21	11:30	MEDIA	114.51	0.227	1.458	0.519	0.227	1.589	1.519	0.011
112	1	1	2	0	7:00	02-06-21	11:30	MEDIA	114.9							
112	1	1	2	0	7:01	02-06-21	11:30	ALTA	82.7							
112	1	1	2	0	7:01	02-06-21	11:30	ALTA	74.9							
112	1	1	2	0	7:02	02-06-21	11:30	MEDIA	124.9							
112	1	1	2	0	7:03	02-06-21	11:30	MEDIA	123.1							
112	1	1	2	0	7:03	02-06-21	11:30	MEDIA	115.8							
112	1	1	2	0	7:04	02-06-21	11:30	MEDIA	152.8							
112	1	1	2	0	7:05	02-06-21	11:30	MEDIA	135.3							
112	1	1	2	0	7:05	02-06-21	11:30	ALTA	92.9							
112	1	1	2	0	7:06	02-06-21	11:30	MEDIA	116.8							
112	1	1	2	0	7:07	02-06-21	11:30	MEDIA	148.6							
112	1	1	2	0	7:07	02-06-21	11:30	MEDIA	147.8							
112	1	1	2	0	7:07	02-06-21	11:30	ALTA	82							

112	1	1	2	0	7:08	02-06-21	11:30	MEDIA	140.3							
112	1	1	2	0	7:09	02-06-21	11:30	MEDIA	146.4							
112	1	1	2	0	7:10	02-06-21	11:30	MEDIA	108.6							
112	1	1	2	0	7:10	02-06-21	11:30	MEDIA	110							
112	1	1	2	0	7:11	02-06-21	11:30	MEDIA	155.6							
112	1	1	2	0	7:12	02-06-21	11:30	MEDIA	118.6							
112	1	1	2	0	7:13	02-06-21	11:30	MEDIA	121.8							
112	1	1	2	0	7:13	02-06-21	11:30	MEDIA	123.6							
112	1	1	2	0	7:14	02-06-21	11:30	ALTA	41.2							
112	1	1	2	0	7:14	02-06-21	11:30	MEDIA	118.1							
112	1	1	2	0	7:15	02-06-21	11:30	MEDIA	125.9							
112	1	1	2	0	7:16	02-06-21	11:30	ALTA	50.2							
112	1	1	2	0	7:16	02-06-21	11:30	ALTA	79.5							

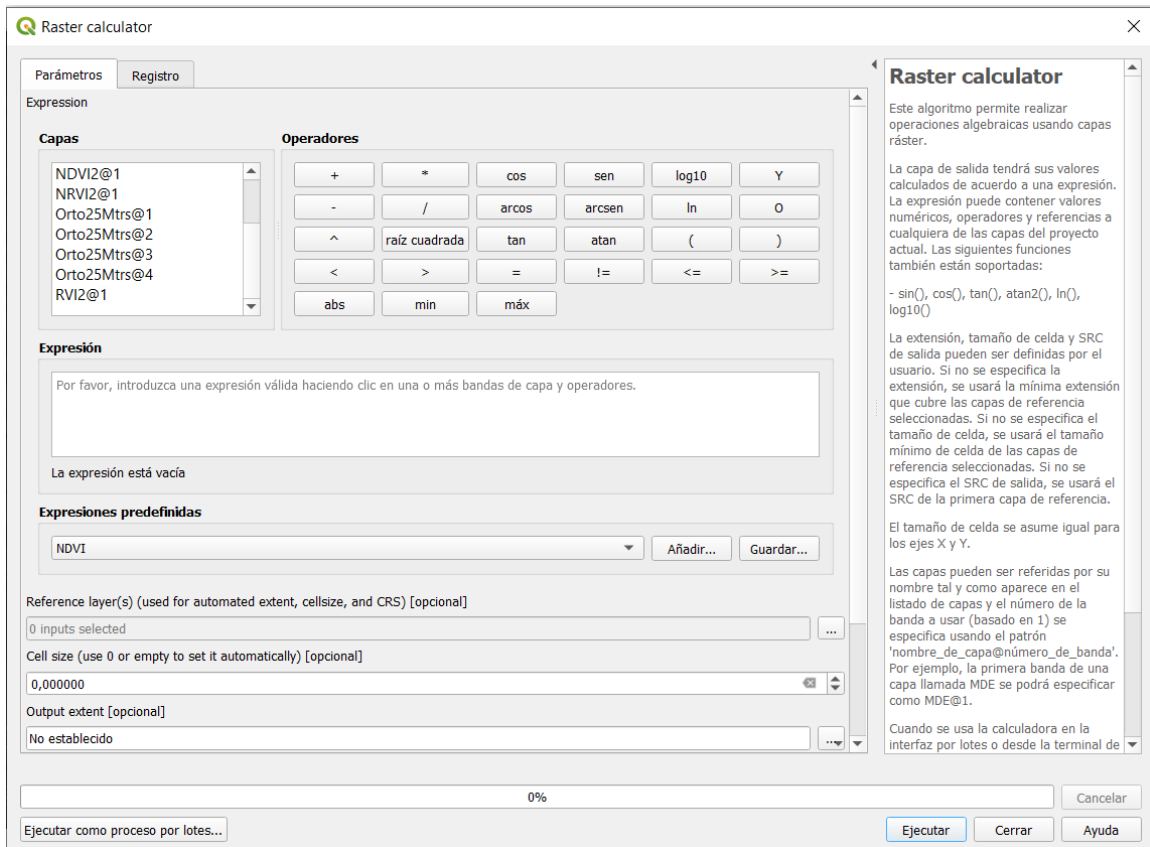


Ilustración 14. Raster calculator de QGIS. Fuente propia.

3.2.1.3. Muestreo de clorofila

Como se ha mencionado anteriormente, los índices de vegetación están relacionados con diversos componentes de las plantas, la clorofila es uno de ellos. Por ejemplo, el índice CVI analiza las bandas NIR, verde y roja para estimar el contenido de clorofila [259], en donde reluce la banda roja (550/850 nm) al presentar una alta correlación con la concentración de clorofila [260]. Por esta razón, se realizó el muestreo de clorofila en las plantas de café usando el clorofilómetro CCM-200 PLUS (tercer paso de la Ilustración 9). Además, en paralelo con el muestreo de clorofila, se capturaron imágenes multiespectrales para encontrar la correlación entre dichos datos. Las tomas de clorofila del terreno se hicieron en un periodo de 8 horas (entre las 7 am y las 3 pm) y el vuelo fue realizado a las 12 del mediodía.

Para realizar el muestreo físico de clorofila en el cultivo, se seleccionó un área con aproximadamente 450 plantas de café de variedad castilla de la misma edad (2 años) organizadas en 15 hileras, cada hilera con 30 plantas. Al analizar la documentación encontrada para muestreo foliar de cultivos de café [261]–[264], se establece un rango de muestreo de 10-30 plantas, tomando por cada planta 15-30 hojas.

Con base en lo anterior, se tomó como referencia el límite superior, es decir, se hizo un muestreo de 30 cafetos y de cada uno de los cafetos se tomaron muestras de 30 hojas, las cuales estaba ubicadas en zona media y alta del cafeto. Por cada una de las 30 hojas se tomaron medidas en dos puntos (centro y exterior) de la hoja. El muestreo se realizó de manera semi-aleatoria en forma de X, como se puede observar en el cuarto elemento de la Ilustración 15 y en la Ilustración 13, con base en lo expuesto por J. Lozano [262].

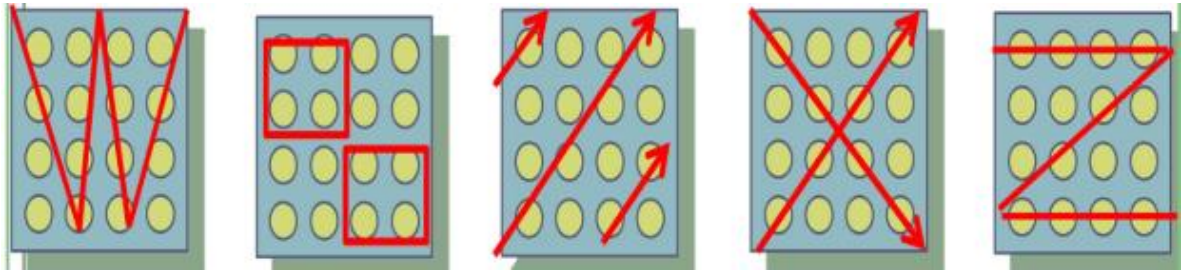


Ilustración 15. Distribución de las plantas en el terreno. Fuente [262]

Arriba mencionado, en la literatura encontrada se determinó que el número de muestras por cafeto debía ser alrededor de 30 hojas [253]–[256]. En consecuencia, el presente estudio, para corroborar dicho valor (30), varió el número de muestras de hojas desde 10 hasta 70, encontrando también, que el número adecuado de muestras a tomar es 30 hojas, ya que este es el punto en donde la desviación estándar no se reduce significativamente como se observa en la Tabla 6 (Los datos completos que se usaron para calcular la desviación estándar se encuentran en el ANEXO E). Además, se observó que la desviación estándar es alta entre las hojas de una misma planta, es decir, hay gran dispersión de los datos. Lo anterior se debe a varios factores, por ejemplo, la diferencia de color entre las hojas de las ramas altas y medias, la posición de las hojas en la rama, los puntos de la hoja en donde se toman las muestras de clorofila, el clima, el error humano, el ángulo en el cual está dando el sol al momento de la captura de la imagen multiespectral, entre otros.

Tabla 6. Variación estándar por número de muestras. Fuente propia

N° hojas	σ CCI
10	58.22
20	55.70

30	55.04
40	55.56
50	54.89
60	54.45
70	54.43

Otro factor que contribuye al incremento de la desviación estándar de la clorofila entre las hojas es que incluso existen variaciones entre los valores de clorofila de puntos cercanos en la misma hoja; como se puede observar en la Tabla 7, existe una gran desviación estándar entre puntos diferentes de una misma hoja, como se observa en la Ilustración 16, se tomaron seis puntos repartidos entre externos y medios de la hoja. Los datos que se usaron para calcular la desviación estándar se encuentran en el ANEXO F. Esta alta desviación estándar presente en los valores de clorofila en los diferentes puntos de la hoja y entre las hojas en sí, además de los mencionados anteriormente, se debe a que el cloroplasto no es uniforme ⁴; en otras palabras, la pigmentación verde en los cloroplastos varía.



Ilustración 16. Muestreo de puntos por hoja. Fuente adaptada de [265].

Tabla 7. Desviación estándar promedio de puntos dentro de cada hoja.
Fuente propia.

⁴ Orgánulo presente en las plantas que contiene la clorofila y realiza la fotosíntesis [314]

σ CCI por hoja
47.9
15.73
13.85
6.7
5.8
10.05
4.94
4.09
22.11
19.62
19.89
26.54
27.34
15.73

3.2.1.4 Preprocesamiento de datos

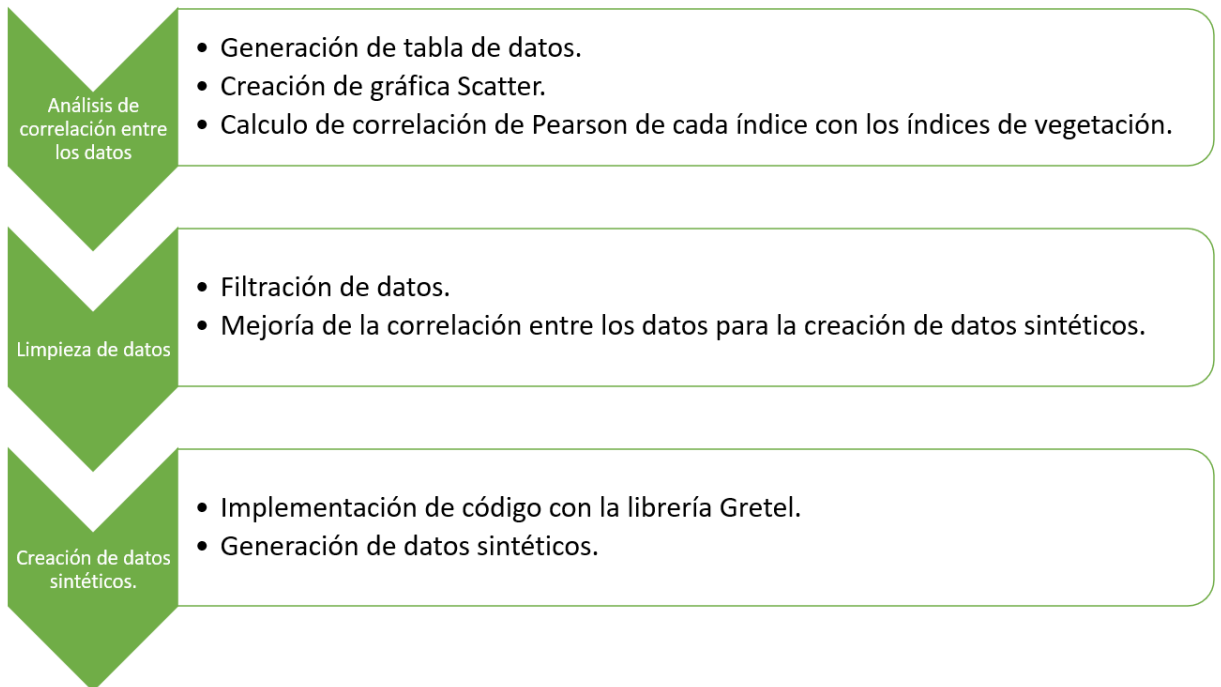


Ilustración 17. Revisión de datos. Fuente propia.

En la etapa de preprocesamiento, se tienen 3 subprocesos como se observa en la Ilustración 17, el análisis de correlación, el filtrado de los datos y la creación de los datos sintéticos. Los datos obtenidos (índices de vegetación y concentración de clorofila) se encuentran en el ANEXO G.

- **Filtrado de datos**

En primera instancia, se analiza la correlación de Pearson, la cual mide la relación entre los datos, entre más se acerquen a 1, mejor es la relación de las variables [266], [267]. En la Tabla 8 se puede observar que la correlación inicial con todos los datos es muy baja, por lo cual es importante analizar los datos y filtrar la información según los atributos de categorización expuestos en la Tabla 5. Por esta razón, se definieron 3 filtros para la información obtenida. El primero es la zona de la planta la cual indica la sección en donde se tomó la muestra, zona alta o zona media como se representa en la Ilustración 18, ya que representa una diferencia en ubicación y color; se observa en la Ilustración 19, una diferencia de verdor entre la zona alta (color verde claro) y la zona media (color verde oscuro), entre más claro sea el verde de la hoja medida, menor es el valor de clorofila dado; el segundo filtro es la hora y el tercero la fecha, los cuales indican la información temporal de la toma de cada muestra en las plantas.

El filtro compuesto con el cual la correlación de Pearson mejoró, fue aquel en donde se filtró por zona media de la planta, ya que es la zona más amplia, con mayor cantidad de hojas y más visible en la foto. La fecha con mejor comportamiento fue el 06/02/2021, esto se debe a que el 28/03/2021 se nubló por secciones al momento de tomar las fotografías, además se filtró en un intervalo de hora de 08:00 am a 04:00 pm para estar en una franja horaria más cerca del vuelo del VANT, el cual fue realizado a las 11:30 am.

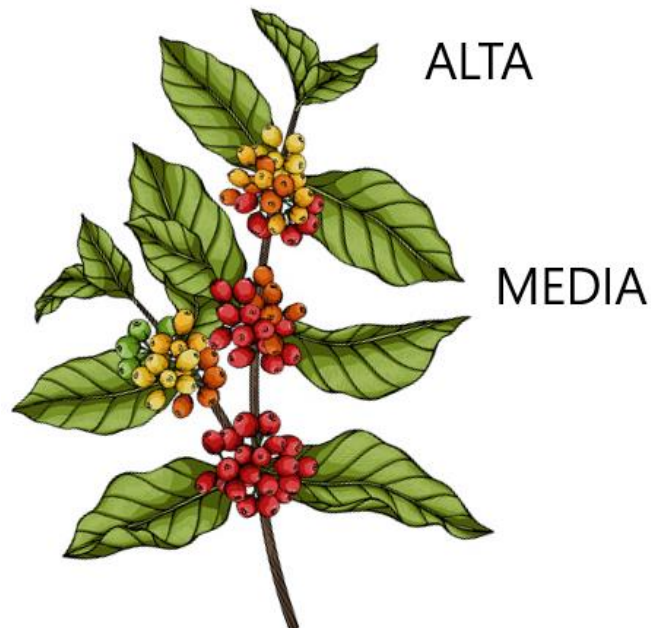


Ilustración 18. Zona de medida dibujo. Fuente adaptada de [268].



Ilustración 19. Diferencia color entre zona alta (a) y zona media (b). Fuente adaptada de [269].

Tabla 8. Correlación de Pearson de índices de vegetación. Fuente propia

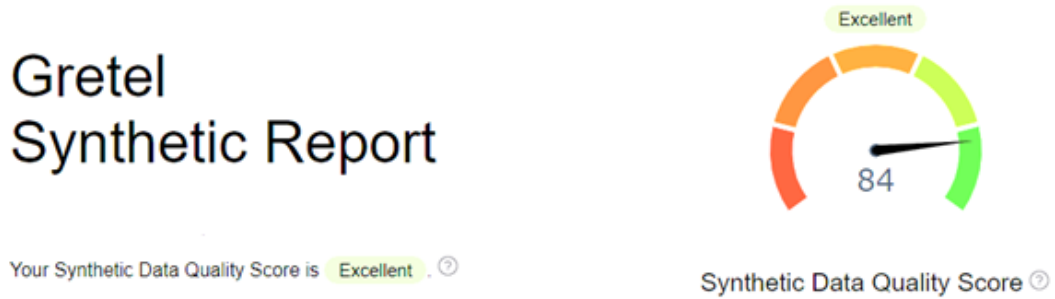
	ANTES DEL FILTRO	DESPUÉS DE FILTRO
NDVI	0,08671	0.596
CVI	-0,01813	0.222
GNDVI	-0,00187	0.272
NRVI	0,08671	0.596
RVI	0,091016	0.602
GCI	-0,00172	0.272

En la Tabla 8 se observa que después del proceso de filtrado, la correlación de Pearson mejoró, ya que los valores obtenidos se acercan un poco más a 1, en comparación a los datos sin manejo de filtros. No obstante, la información se redujo, por lo cual, para aumentar el tamaño del conjunto de datos y superar las limitantes de tiempo para tomar más muestras en campo, se optó por una alternativa usada en el ámbito del aprendizaje automático, la generación de datos sintéticos [270]–[274].

- **Generación de datos sintéticos**

Para generar los datos sintéticos se usó la librería de *Gretel Synthetics*, la cual permite generar datos sintéticos a través del uso de enlaces a la API REST de Gretel [275]. Gretel tiene diferentes configuraciones, donde se probaron las principales, las cuales fueron: la predeterminada, recuento de registros bajo, datos numéricos y recuento de campos alto [276]. Cada una de estas configuraciones fue probada, con el fin de buscar aquella que diera el mejor resultado, es decir, que el reporte generado por la API de Gretel diera un puntaje en la calidad de los datos “excelente”, como el que se

observa en la Ilustración 20 [277]. En el ANEXO H se encuentran los resultados obtenidos con sus respectivos datos.



Data Summary Statistics



Ilustración 20. Reporte de resultados de datos sintéticos obtenidos en Gretel. Fuente captura de la herramienta Gretel [277].

En las estadísticas de los datos en la Ilustración 20, se pueden observar tres identificadores: estabilidad de correlación de campo (95), estabilidad de estructura profunda (75) y estabilidad de distribución de campo (83). Los valores observados indican el puntaje de los datos en ese identificador, entre mayor sea, indica una mejor calidad de datos, se consideran buenos valores 70-100, siendo 100 el mejor valor.

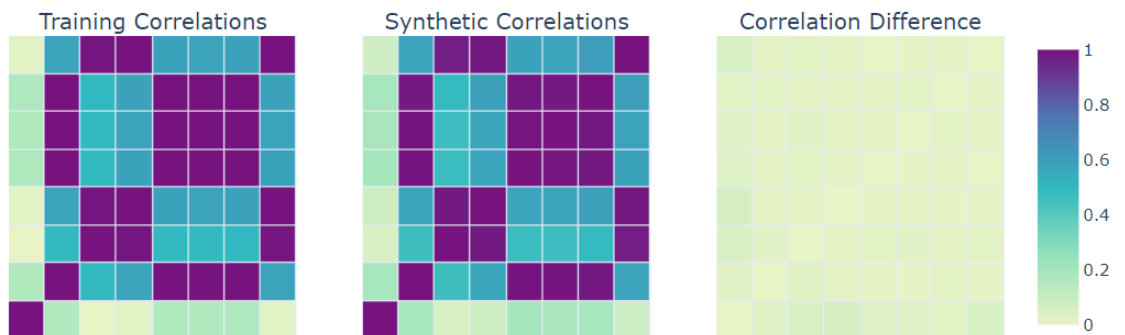


Ilustración 21. Reporte de correlación entre datos de entrenamiento y sintéticos. Fuente propia

Para hallar la estabilidad de correlación de campo, se calcula la correlación en los datos de entrenamiento (Training Correlations de la Ilustración 21) y

en los datos sintéticos (Synthetic Correlations Correlations de la Ilustración 21). Posteriormente, se calcula la diferencia entre estos valores, dando como resultado la diferencia de la correlación (Correlation Difference), la cual se promedia en todos los campos, entre menor sea la diferencia de correlación, es decir, un valor más cercano a 0 (color amarillo) como se observa en la escala de colores de la Ilustración 21, mejor será la calidad de los datos obtenidos, lo cual es importante para el entrenamiento de técnicas de aprendizaje automático.

Para calcular la estabilidad de estructura profunda, Gretel compara un análisis de componentes principales (*PCA, por sus siglas en inglés*) con los datos originales y luego con los datos sintéticos, una explicación más detallada sobre PCA se encuentra en la documentación de Gretel [277].

Por último, la estabilidad de distribución de campo denota que tan parecida es la distribución de campo en los datos sintéticos con respecto a los originales, para conocer la calidad de los datos se utiliza la distancia de Jensen-Shannon. Cuanto menor sea este valor, mejor serán los datos. Para mayor información se puede consultar la explicación de la documentación de Gretel [277].

3.2.1.5 Modelado de datos

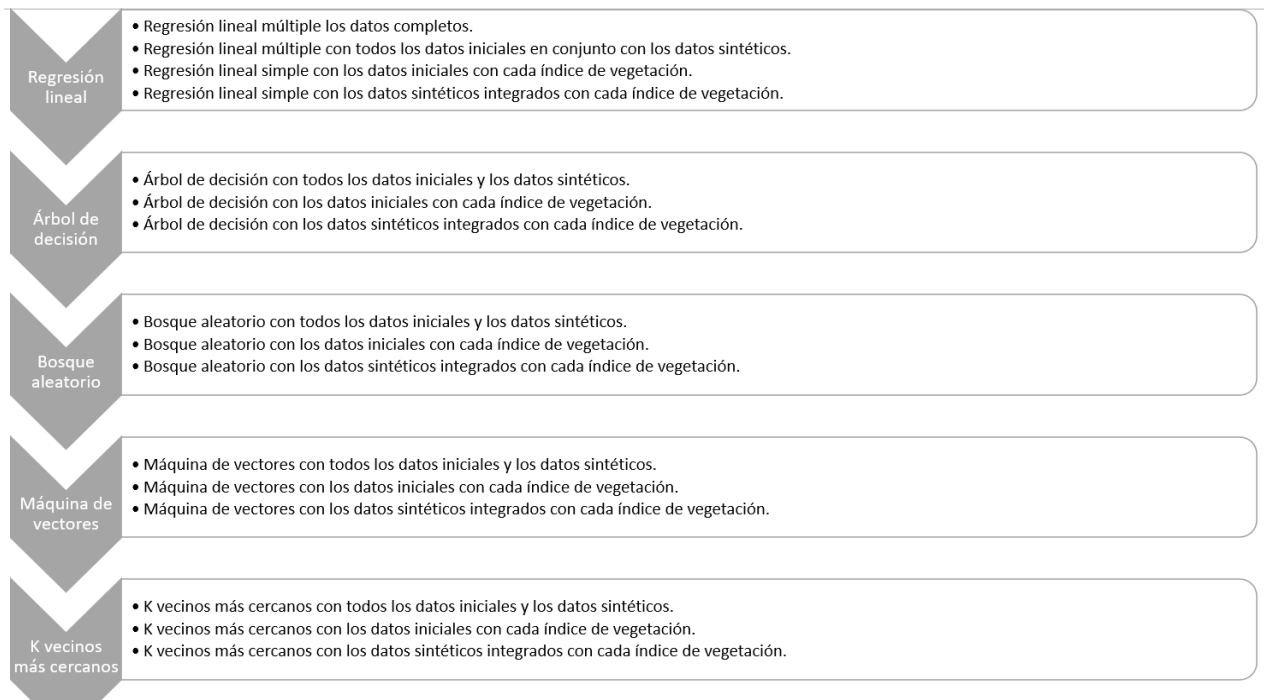


Ilustración 22. Modelado de datos. Fuente propia.

La clorofila sirve como indicador del estado de salud, nutrientes y productividad de las plantas [246], [278]–[280]. Establecer un esquema de clasificación del estado de salud y nutrición de la planta con base en los

valores de clorofila medidos in situ es posible [249], [281]–[284], un ejemplo de ello se observa en la Ilustración 23 en donde, entre mayor es el valor dado por el clorofilómetro SPAD, mejor será el estado de salud de la planta [285]. Para expresar la relación del estado de salud de la planta con los valores CCI, se utilizó la relación entre SPAD y CCI expresado en el estudio de Dong T. et al., [249], el cual se ve expuesto en la Tabla 9, ya que tiene en cuenta varios tipos de cultivos.

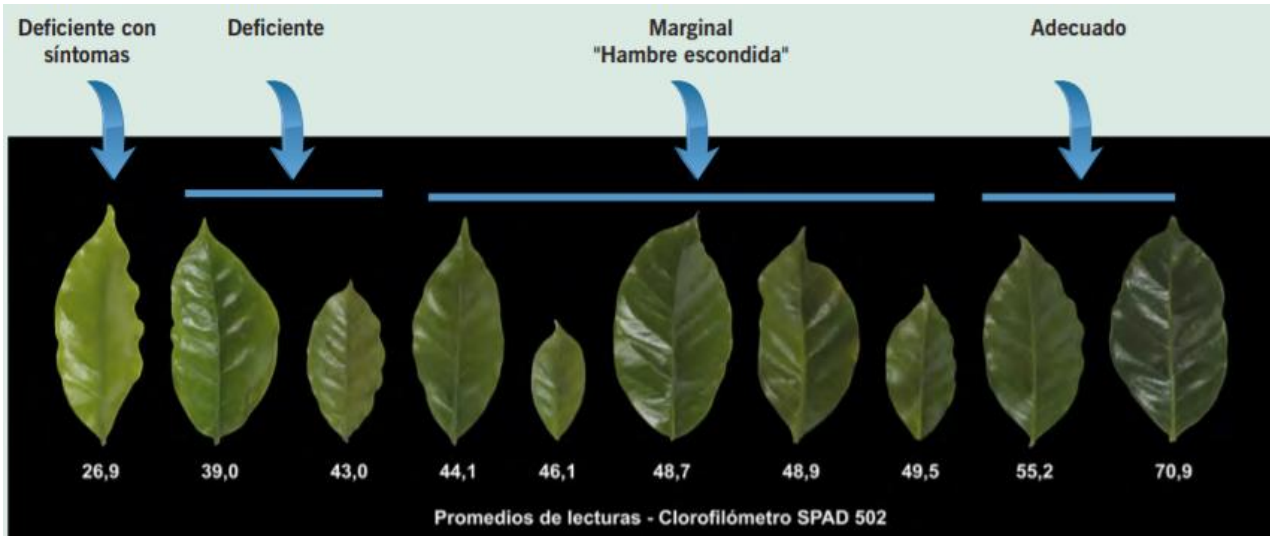


Ilustración 23. Rangos de estados de nutrientes adecuados para las hojas de café. Fuente [249]

Tabla 9. Valores SPAD y CCI por nivel de salud. Fuente adaptada de Dong T et al., [249].

CATEGORÍA	SPAD 502 (SPAD)	CCM-200 (CCI)
Deficiente	$X < 43$	$X < 25$
Marginal/Normal	$44 < X > 55$	$26 < X > 54$
Adecuado	$X > 56$	$X > 55$

Teniendo claro y definiendo los valores de SPAD y CCI para el estudio, se necesita una herramienta adecuada para realizar la estimación de variables y componentes vegetativos, estudios previos han demostrado la efectividad y popularidad del uso de las regresiones lineales, árboles de decisión, bosques aleatorios, máquinas de vectores y k vecinos más cercanos para la estimación de variables y componentes vegetativos [48], [60], [286]–[291], además de ser útiles para conjuntos de datos pequeños [292], [293]. Por estas razones, en el presente estudio fueron implementados dichos modelos tanto para los datos iniciales sin datos sintéticos como para aquellos con los datos sintéticos. Los códigos implementados se encuentran en nuestro

repositorio de github ⁵.

Como se observa en la Ilustración 22, en el proceso de modelado de datos, se aplicaron cinco modelos; regresión lineal (simple y múltiple), árbol de decisión, bosque aleatorio, máquina de vectores y k vecinos más cercanos. Para cada uno de ellos se utilizó validación cruzada, ya que permite evaluar los modelos mediante su entrenamiento en subconjuntos de datos. Estos se dividen en datos de entrada disponibles (80%) y datos complementarios (20%), los datos utilizados se encuentran en el ANEXO G [294]. Para el entrenamiento de dichos modelos y la segmentación en los subconjuntos de datos, se utilizó la librería *train_test_split()* de *scikit-learn*, ya que minimiza el sesgo en los procesos de evaluación y validación [295].

- **Modelado con regresión lineal**

La regresión lineal simple o múltiple es comúnmente aplicada para estimar el valor de una variable según el valor de otra a través de una recta que refleja el comportamiento de las variables [296]. En este caso la variable dependiente, es decir, la que se quiere predecir es el nivel de clorofila CCI, y las variables o la variable independiente serán dadas por los índices de vegetación hallados.

Para evitar la multicolinealidad se analizó la correlación entre las variables, esto se puede observar en la Tabla 10 y en la Ilustración 24 (Entre más claro sea el color referente a la línea de color ubicada en el extremo derecho, mayor es la correlación). Posteriormente, se eliminaron aquellas que se veían representadas en otra, es decir, que presentarán una correlación mayor a 0.9, por esta razón se realizó la regresión lineal múltiple para el RVI y GCI.

Para entrenar los modelos de regresión lineal simple y regresión lineal múltiple se utilizó de forma aleatoria el 80% de los datos y para los datos de pruebas el 20% restante. La implementación de la regresión lineal utilizada corresponde a la librería *sklearn* de python.

Tabla 10. Correlación entre índices de vegetación y clorofila. Fuente propia

	CLL	NDVI	CVI	GNDVI	NRVI	RVI	GCI	NRVI3
CLL	1.0000	0.5966	0.4298	0.4674	0.5966	0.6001	0.4674	0.5980
NDVI	0.5966	1.0000	0.6100	0.6754	1.0000	0.9995	0.6754	0.9999
CVI	0.4298	0.6100	1.0000	0.9962	0.6100	0.6247	0.9962	0.6145
GNDVI	0.4674	0.6754	0.9962	1.0000	0.6754	0.6890	1.0000	0.6796
NRVI	0.5966	1.0000	0.6100	0.6754	1.0000	0.9995	0.6754	0.9999
RVI	0.6001	0.9995	0.6247	0.6889	0.9995	1.0000	0.6890	0.9995
GCI	0.4674	0.6754	0.9962	1.0000	0.6754	0.6890	1.0000	0.6796
NRVI3	0.5980	0.9999	0.6145	0.6796	0.9999	0.9995	0.6796	1.0000

⁵ Repositorio de github <https://github.com/NataliaArteaga/MultispectralImagesCoffee>.

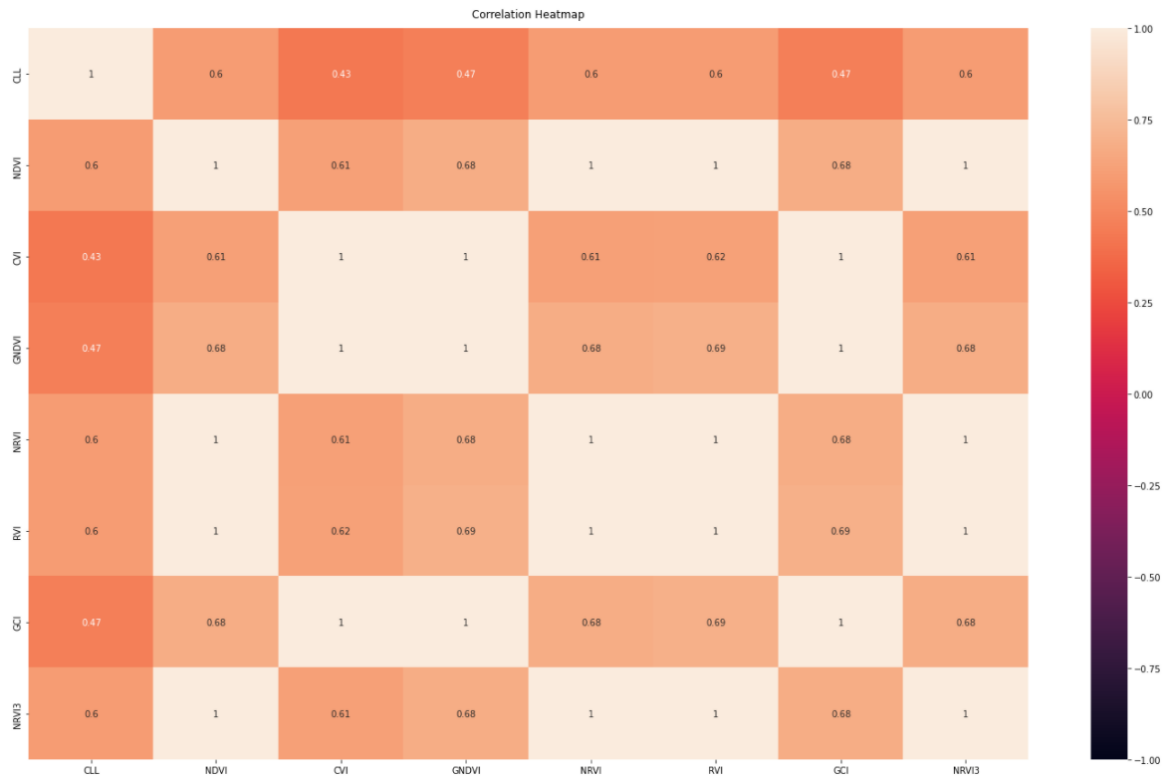


Ilustración 24. Mapa de calor de correlación de Pearson. Fuente propia.

En la Tabla 11 se registran el MSE y el MAE de la regresión lineal múltiple. Para la regresión lineal múltiple (y para los siguientes modelos) el valor a predecir fue el índice de contenido de clorofila (CCI) y las variables independientes los índices de vegetación (RVI y GCI). Además, se implementó una regresión lineal simple para cada uno de los índices de vegetación para analizar su comportamiento individual. Como se observa en la Tabla 12, el mejor comportamiento de la regresión lineal se observó en el NRVI al tener el error más bajo, esto puede deberse a que este índice reduce los efectos topográficos, de iluminación e implicaciones climáticas [15].

Tabla 11. Error de datos de entrenamiento y datos sintéticos con regresión lineal múltiple. Fuente propia

	INICIAL	CON DATOS SINTETICOS
MSE	163.998	111.954
MAE	10.614	8.575

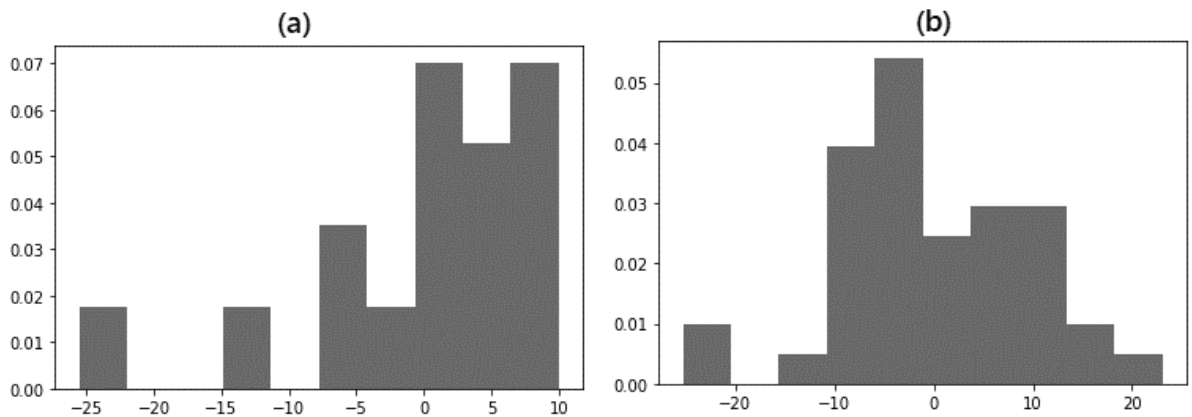
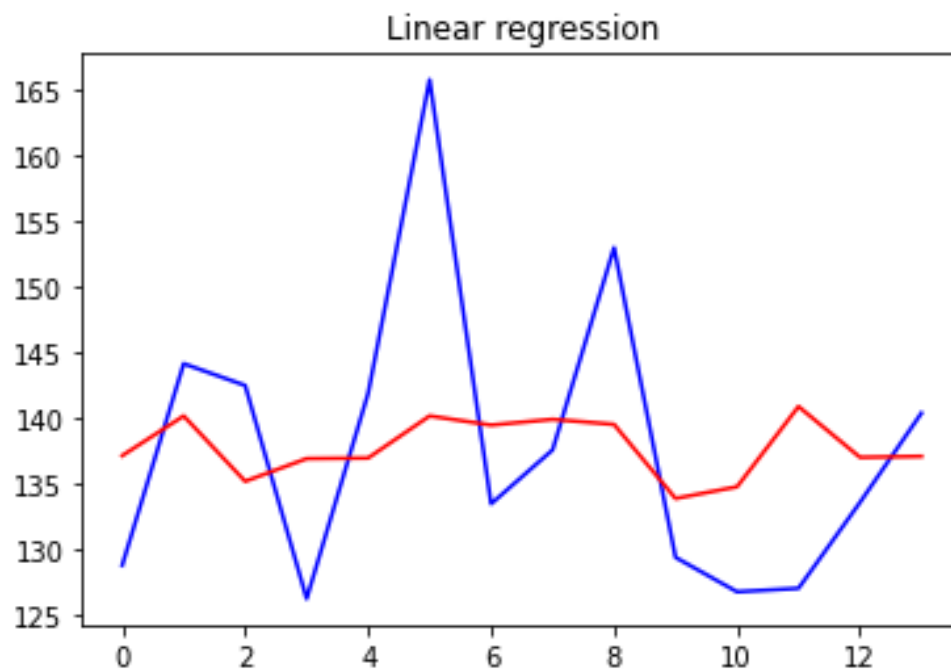


Ilustración 25. Distribución error sin datos sintéticos (a) y con datos sintéticos (b). Fuente propia.



Pearsons correlation: 0.410

Ilustración 26. Datos reales (azul) y estimados (rojo) y correlación de Pearson. Fuente propia

Al aumentar el número de muestras, como se observa en la Ilustración 25, la distribución del error se acerca más a la distribución gaussiana, la cual es la distribución normal y esperada al analizar el error de un modelo [297], [298]. En la Ilustración 26, se visualiza la comparación de los valores reales y los estimados, en el eje X se encuentra el identificador y en el eje Y el valor de clorofila, además se obtiene una correlación de Pearson de 0.41, por lo cual hay una correlación positiva moderada entre los valores reales y estimados del modelo, entre mayor sea la correlación de los datos y menor el error

presentado, mejor es el funcionamiento del modelo.

Tabla 12. Error de datos de entrenamiento y datos sintéticos por cada índice de vegetación con regresión lineal simple. Fuente propia.

		INICIAL	DATOS SINTÉTICOS
NDVI	MAE	10.6102	8.4129
	MSE	163.0969	108.682
CVI	MAE	10.5688	8.4148
	MSE	162.321	116.220
GNDVI	MAE	10.5640	8.44
	MSE	162.421	115.56
NRVI	MAE	10.5404	8.4031
	MSE	160.45113	108.084
RVI	MAE	10.5916	8.41612
	MSE	162.022	108.308
GCI	MAE	10.5668	8.4506
	MSE	161.6797	115.56

Cabe resaltar que inicialmente se hizo el mismo procedimiento y aplicación de la regresión lineal simple y múltiple para analizar el comportamiento sin el filtro realizado por zona y fecha mencionado, la diferencia se encuentra en que se realizó una limpieza de datos externos manualmente y además se modeló el conjunto de datos con redes neuronales artificiales; los resultados obtenidos no fueron los mejores en ese momento, por lo cual se tomó la decisión de realizar el proceso de filtración y explorar nuevos algoritmos, ya que las redes neuronales no eran las adecuadas. Lo mencionado anteriormente se encuentra descrito en el ANEXO I.

- **Modelado con árbol de decisión**

Los árboles de decisión son algoritmos que al igual que las regresiones relacionan distintas variables con el fin de predecir otra, utilizados principalmente cuando la variable de respuesta es numérica. Los árboles de decisión poseen una estructura que está formada por ramas y nodos. Los nodos internos representan las características que se tienen en cuenta para tomar una decisión, las ramas representan la decisión en función de una condición y los nodos finales el resultado de la decisión, es decir, representa la estimación esperada [299], [300]. Para la implementación se utilizó la librería sklearn con la configuración por defecto [301]. El algoritmo se entrenó con los datos iniciales y el conjunto de datos integrando los datos sintéticos, se observa en la Tabla 13 que nuevamente el error disminuyó al incluirlos y en la Ilustración 27 mejoró la distribución gaussiana (que es lo esperado) al incorporar los datos sintéticos.

Tabla 13. Error árbol de decisión con datos iniciales y con datos sintéticos.
Fuente propia

	INICIAL	CON DATOS SINTETICOS
MAE	11.71	10.87
MSE	262.73	208.05

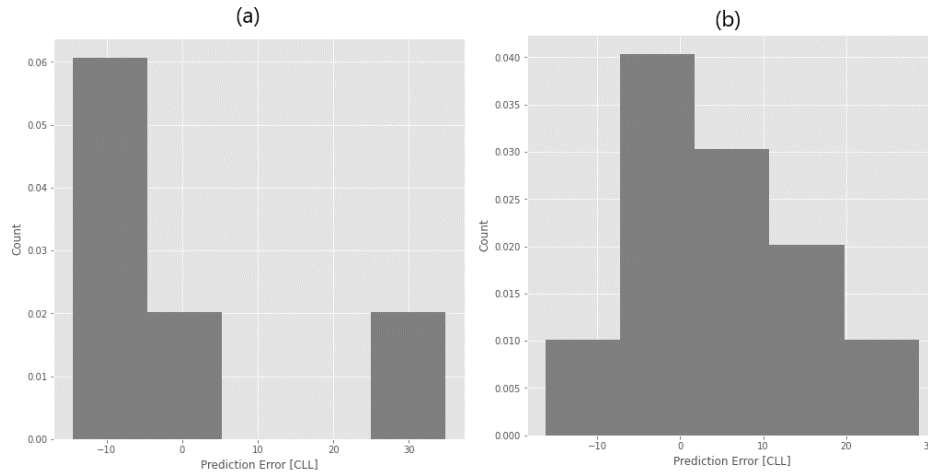


Ilustración 27. Distribución error árbol de decisión con datos iniciales (a) e integrando datos sintéticos (b). Fuente propia

En la gráfica de la Ilustración 28 se observa la correlación de Pearson obtenida entre los valores reales y estimados, la cual es sustancial o fuerte con un valor de 0.502, lo cual indica una fuerte conexión en la estimación del valor.

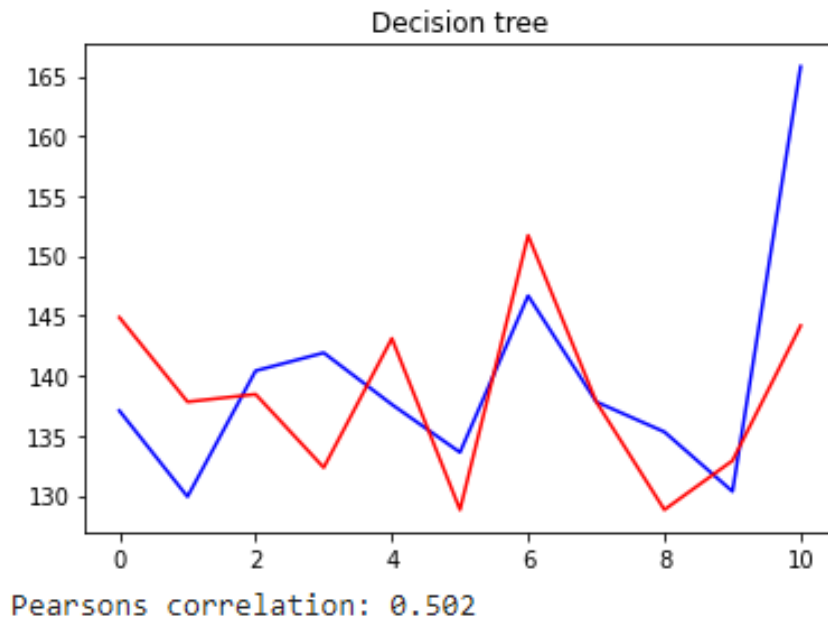


Ilustración 28. Datos reales (azul) y estimados (rojo) y correlación de Pearson. Fuente propia.

La Tabla 14 enseña el error obtenido con cada índice de vegetación con los datos iniciales y con los datos sintéticos, en este caso, quien demostró el mejor desempeño es el CVI, el cual es un índice adecuado al ser usado para estimar el contenido de clorofila de las hojas [302], [303].

Tabla 14. Error de datos de entrenamiento y datos sintéticos por cada índice de vegetación con árbol de decisión. Fuente propia.

		INICIAL	DATOS SINTÉTICOS
NDVI	MAE	13.9162	11.69
	MSE	290.13	242.01
CVI	MAE	11.18	9.84
	MSE	264.73	175.79
GNDVI	MAE	11.13	11.04
	MSE	221.12	209.56
NRVI	MAE	13.92	11.70
	MSE	289.87	241.79
RVI	MAE	12.42	11.11
	MSE	243.36	213.79
GCI	MAE	10.48	11.16
	MSE	225.81	219.582

- **Modelado con máquina de vectores**

La máquina de vectores (SVM, *por sus siglas en inglés*) es un algoritmo de aprendizaje supervisado, en el cual, a partir de un conjunto de datos de entrenamiento, se configura un modelo que prediga la clase de un nuevo dato. Para ello, el algoritmo obtiene un hiperplano que representa de la mejor manera posible el comportamiento de los datos, separándolos entre si al categorizar y segmentar los datos según su ubicación en el hiperplano [304], [305]. El rendimiento depende de la configuración de los parámetros del Kernel, es por ello que se debe usar librerías de programación que tengan bien desarrolladas la implementación de este algoritmo, como Scikit Learn [306]. Se observa en la Tabla 15 y en la Ilustración 29 que el error mejoró tanto en su valor como en la forma de la distribución al incluir los datos sintéticos, al igual que en los modelos anteriores.

Tabla 15. Error de datos de entrenamiento y datos sintéticos con máquinas de vectores. Fuente propia.

	INICIAL	CON DATOS SINTETICOS
MAE	10.129	7.849
MSE	188.77	106.491

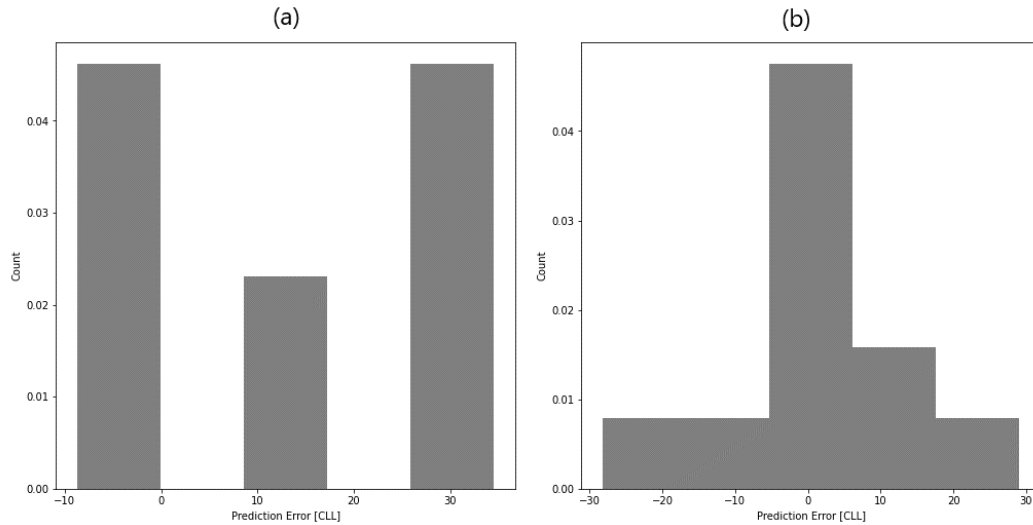


Ilustración 29. Distribución del error de datos de entrenamiento iniciales (a) y datos sintéticos (b) con máquinas de vectores. Fuente propia.

En la Ilustración 30 se observa una alta correlación entre los valores predichos y reales con un valor de 0.764, lo cual refleja ser un modelo adecuado para el fin de esta investigación debido a su alta correlación presente.

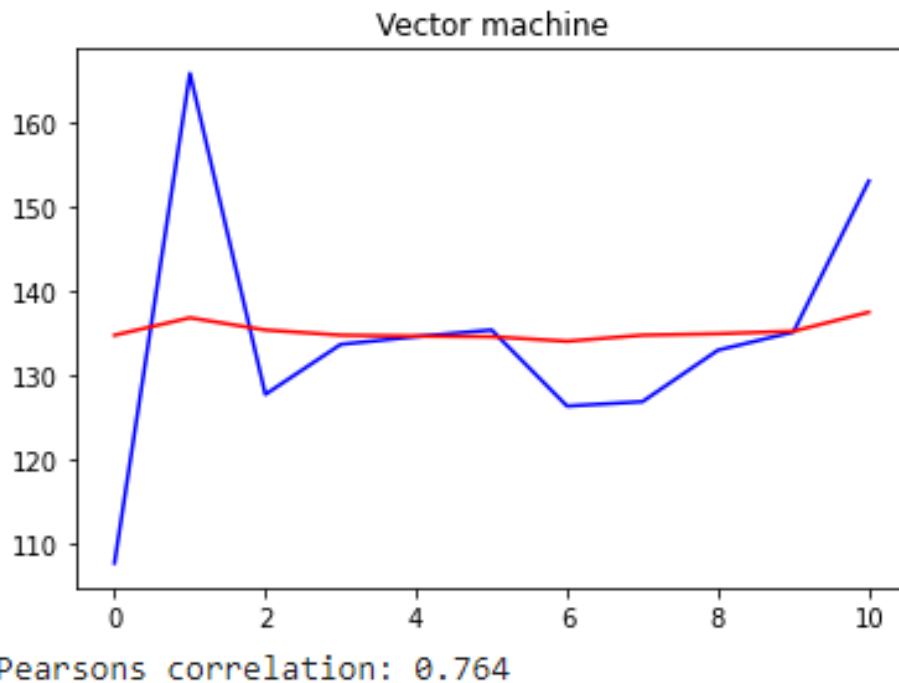


Ilustración 30. Datos reales (azul) y estimados (rojo) y correlación de Pearson. Fuente propia.

La Tabla 16 enseña los errores obtenidos por el modelo en cada índice de vegetación con los datos iniciales y con los datos sintéticos, en este caso el

índice con mejor comportamiento fue el GNDVI, el cual es más sensible a la variación de la clorofila en el cultivo que el índice NDVI, este índice es adecuado para estimar el vigor del cultivo [307], cabe resaltar que el CVI y GCI, índices enfocados en la estimación de clorofila tuvieron un error bastante próximo al GNDVI, lo que concuerda con lo encontrado en el análisis de la correlación de Pearson.

Tabla 16. Error de datos de entrenamiento y datos sintéticos por cada índice de vegetación con máquina de vectores. Fuente propia

		INICIAL	DATOS SINTÉTICOS
NDVI	MAE	10.26	8.44
	MSE	192.69	117.18
CVI	MAE	10.03	7.96
	MSE	188.83	108.31
GNDVI	MAE	10.014	7.95
	MSE	187.178	108.81
NRVI	MAE	10.32	8.50
	MSE	195.18	118.93
RVI	MAE	10.17	8.45
	MSE	190.50	117.26
GCI	MAE	10.07	7.96
	MSE	190.23	108.61

- **Modelado con bosque aleatorio**

Los bosques aleatorios seleccionan al azar características para construir varios árboles de decisión y luego promediar los resultados. Los bosques aleatorios generalmente evitan el exceso de adaptación, ningún árbol ve todos los datos de entrenamiento, esto hace que cada árbol se entrene con distintas muestras de datos para un mismo problema; al promediar y combinar dichos resultados, unos errores se compensan con otros y se tiene una predicción que generaliza mejor [308]. Nuevamente para la implementación del modelo se utilizó la librería sklearn.

Al integrar los datos sintéticos en el conjunto de datos que alimentan el modelo, se ve una mejoría del error en la Tabla 17 y en la distribución observada en la Ilustración 31.

Tabla 17. Error de datos de entrenamiento y datos sintéticos con bosque aleatorio. Fuente propia.

	INICIAL	CON DATOS SINTÉTICOS
MAE	9.05	8.79
MSE	154.70	130.98

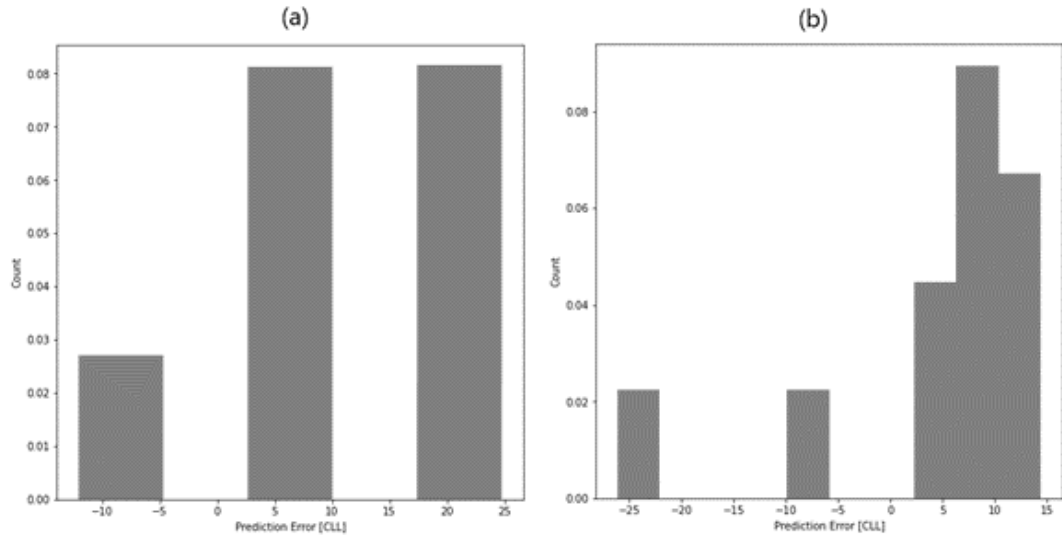


Ilustración 31. Distribución del error de datos iniciales (a) y datos sintéticos (b) con bosque aleatorio. Fuente propia.

Para el modelo de bosque aleatorio el conjunto de datos analizados, es decir, los datos reales y estimados se asemejan entre sí, como se observa en la Ilustración 32 arrojó una correlación de 0.693, la cual es una correlación muy alta y representa ser un buen modelo de predicción al igual que la máquina de vectores.

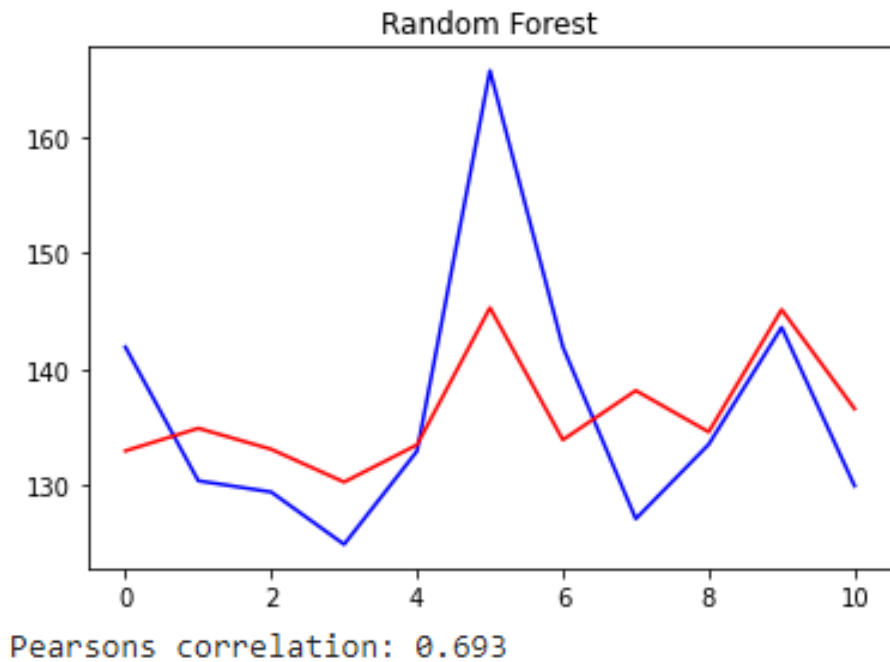


Ilustración 32. Datos reales (azul) y estimados (rojo) y correlación de Pearson. Fuente propia.

El índice que mostró el error más bajo en el modelo bosque aleatorio es el CVI como se observa en la Tabla 18, un índice que es utilizado regularmente

para estimar la clorofila.

Tabla 18. Error de datos de entrenamiento y datos sintéticos por cada índice de vegetación con bosque aleatorio. Fuente propia

		INICIAL	DATOS SINTÉTICOS
NDVI	MAE	11.41	10.25
	MSE	207.19	177.50
CVI	MAE	10.50	8.848
	MSE	192.59	137.78
GNDVI	MAE	9.52	9.49
	MSE	163.13	155.409
NRVI	MAE	11.18	10.27
	MSE	200.74	177.30
RVI	MAE	10.68	9.64
	MSE	185.16	162.17
GCI	MAE	9.39	9.38
	MSE	163.22	157.02

- **Modelado con K vecinos más cercanos**

K vecinos más cercanos (KNN, por sus siglas en inglés) es un algoritmo de aprendizaje supervisado. Este algoritmo no hace suposiciones sobre la funcionalidad y busca memorizar las instancias de aprendizaje que se utilizan para usarlo en la fase predictiva. Una pequeña cantidad de vecinos suele demostrar mayor flexibilidad, bajo sesgo, pero alta varianza. Por el contrario, un gran número de vecinos una varianza más baja pero un sesgo más alto.

El algoritmo clasifica cada nuevo dato que ingresa en el grupo que corresponda, dependiendo del número de vecinos más cercanos de un grupo o de otro. Los datos parecidos están próximos y los que no lo son están alejados entre sí. Por lo tanto, la distancia entre dos casos es una medida de disimilaridad. Para usos de predicción de valores continuos se utiliza la media o el valor objetivo medio de los vecinos más cercanos [309]–[311].

En la Tabla 19 se registran los errores obtenidos con los datos iniciales y con los datos sintéticos del modelo de la librería sklearn. Nuevamente se demuestra una mejoría en el error al modelar el algoritmo con los datos sintéticos. Esto también se observa al mejorar la distribución gaussiana obtenida en la Ilustración 33.

Tabla 19. Error de datos de entrenamiento y datos sintéticos con k vecinos más cercanos. Fuente propia

	INICIAL	CON DATOS SINTETICOS
MAE	9.71	8.247

MSE	152.67	120.74
-----	--------	--------

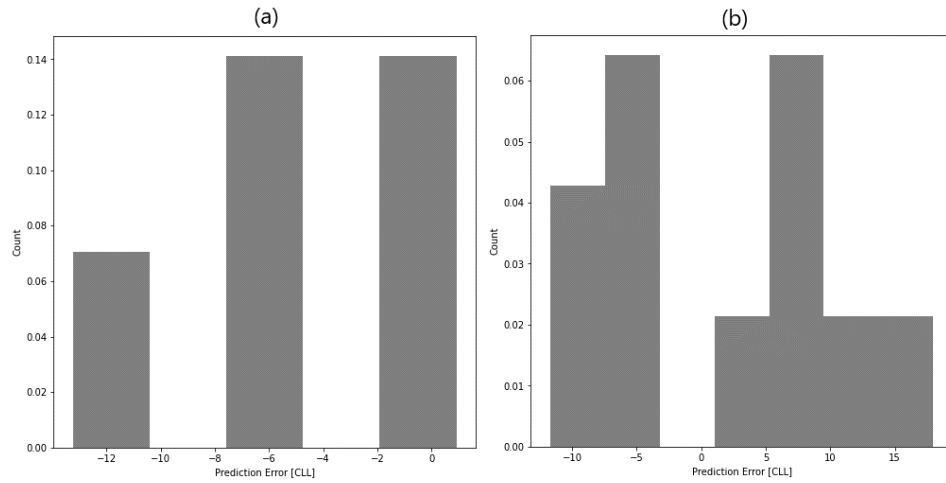


Ilustración 33. Distribución del error de datos iniciales (a) y datos sintéticos (b) con k vecinos más cercanos. Fuente propia.

La Ilustración 34 muestra en una gráfica una línea azul que vislumbra los valores reales versus la línea roja que son los valores estimados, además se observa la correlación dada, la cual es de 0.446, la cual es moderada y no tan eficiente en comparación con los modelos anteriores. Entre mayor sea la cercanía entre los puntos al valor real (azul), mejor es el comportamiento del modelo.

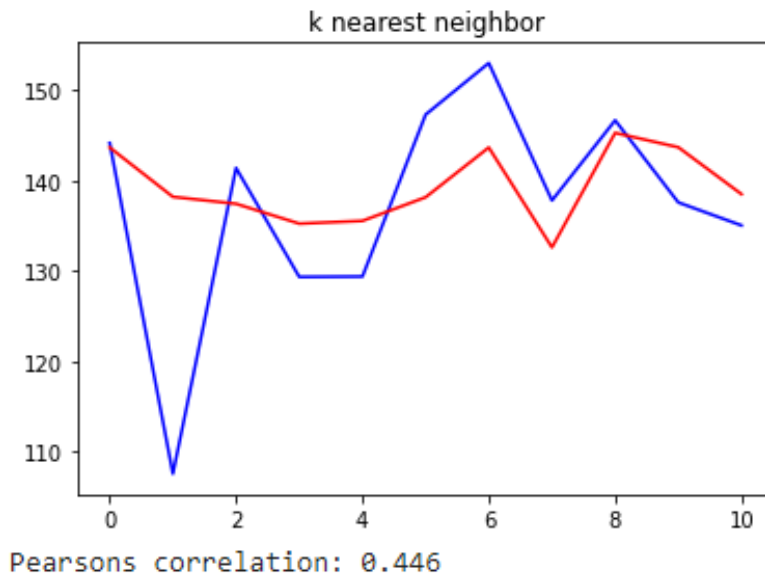


Ilustración 34. Datos reales (azul) y estimados (rojo) y correlación de Pearson. Fuente propia.

En la Tabla 20 se observa la mejoría al integrar los datos sintéticos para cada índice de vegetación, destaca en este caso el GNDVI, GCI y CVI

nuevamente, como en el algoritmo de máquina de vectores.

Tabla 20. Error de datos de entrenamiento y datos sintéticos por cada índice de vegetación con k vecinos más cercanos. Fuente propia

		INICIAL	DATOS SINTÉTICOS
NDVI	MAE	10.43	9.43
	MSE	180.25	140.70
CVI	MAE	10.88	8.30
	MSE	170.32	111.51
GNDVI	MAE	9.96	8.25
	MSE	154.14	115.68
NRVI	MAE	10.44	9.42
	MSE	180.22	140.97
RVI	MAE	10.35	9.27
	MSE	175.38	135.91
GCI	MAE	10.02	8.26
	MSE	156.55	17.15

- **Comparación entre modelos**

En la Tabla 21 se observa que el modelo que obtuvo el error más bajo al tener todos los datos de entrenamiento es el registrado por máquina de vectores de soporte, con un error de 7.85 y una correlación de 0.58. En la Ilustración 35 se observa el comportamiento de los datos estimados de cada modelo en comparación con los datos reales (línea azul), el eje X representa el identificador del dato y el eje Y el valor de clorofila.

Tabla 21. Comparación errores y correlación entre los modelos. Fuente propia.

MODELO	MAE	MSE	PEARSON CORR
Regresión lineal	8.58	111.95	0.48
Árbol de decisión	10.87	208.05	0.42
Máquina de vectores	7.85	106.49	0.58
Bosque aleatorio	8.79	130.98	0.49
K vecinos más cercanos	8.26	120.74	0.51

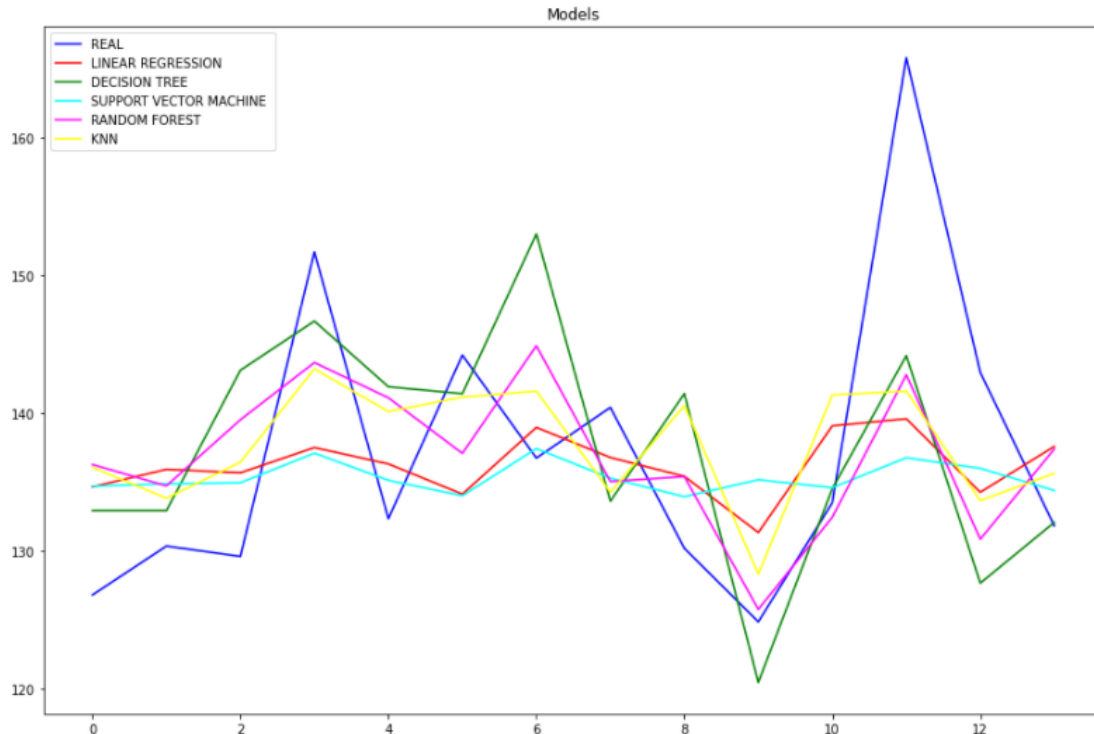


Ilustración 35. Valor real y valor de predicción según cada modelo. Fuente propia.

El índice con mejor desempeño es el CVI como se observa en la columna Media de la Tabla 22, el cual fue uno de los más destacados en cada modelo. El CVI como se ha mencionado en secciones anteriores es uno de los índices más usados para estimar clorofila, al igual que el GCI y GNDVI, los cuales también resaltan por desempeñarse mejor que los demás. El modelo que presentó el error más bajo es máquina de vectores, debido a la estandarización que este genera en la creación del hiperplano y la segmentación de datos en ella, mitigando estimaciones con valores extremos, otro modelo con buen desempeño es el bosque aleatorio, el cual al ser un conjunto de árboles de decisión que no se ven entre sí, mitiga el error y crea una mejor generalización en la estimación de los datos.

Tabla 22. Errores por modelos e índices de vegetación. Fuente propia.

		Regresión lineal	Árbol de decisión	Máquina de vectores	Bosque aleatorio	K vecinos más cercanos	Media
NDVI	MAE	8.412	11.69	8.44	10.25	9.43	9.64
	MSE	108.68	242.01	117.18	177.50	140.70	157.21
CVI	MAE	8.414	9.84	7.96	8.848	8.30	8.67
	MSE	116.22	175.79	108.31	137.78	111.51	129.92
GNDVI	MAE	8.44	11.04	7.95	9.49	8.23	9.03

	MSE	115.56	209.56	108.81	155.41	115.68	141
NRVI	MAE	8.403	11.70	8.50	10.27	9.42	9.66
	MSE	108.08	241.79	118.93	177.30	140.97	157.41
RVI	MAE	8.416	11.11	8.45	9.64	9.27	9.38
	MSE	108.31	213.79	117.26	162.17	135.91	147.49
GCI	MAE	8.45	11.16	7.96	9.38	8.26	9.04
	MSE	115.56	219.582	8.44	157.02	17.15	103.55

En este capítulo se describió la metodología aplicada y sus fundamentos para la captura de los datos (imágenes multiespectrales y la clorofila en planta con el CCM-200 plus), además de su procesamiento. Después del procesamiento, los datos finales alimentaron los modelos de regresión lineal, árboles de decisión, bosque aleatorio, máquina de vectores de soporte y k vecinos más cercanos. Cada uno de estos modelos fue evaluado con los datos sin y con los datos sintéticos con los errores MAE y MSE y el análisis de la correlación de Pearson.

4. Conclusiones

Al analizar los errores en los diferentes modelos, sus valores se pueden justificar por varios factores. En primera instancia, los datos de la clorofila presentan una desviación estándar significativa tanto a nivel de hoja como de planta, lo que aumenta la dispersión entre los datos obtenidos. Además, se deben tener en cuenta los tiempos de muestreo largos de los parámetros físicos respecto al tiempo del vuelo del VANT, ya que tomar las muestras físicas llevaba un tiempo aproximado de 7 - 8 horas y el vuelo del VANT se llevaba a cabo durante 10 minutos en una hora específica del día (generalmente entre las 11:00am y 01:00pm). Otro aspecto importante es que la variación climática afecta la respuesta de clorofila y la reflectancia capturada por la cámara, ya que, en uno de los vuelos, el día estaba despejado y en el otro existió nubosidad intermitente. (Se realizaron más vuelos, pero estos fueron descartados al no poder tomar todos los datos de forma completa ya que las lluvias afectaron la toma de muestras físicas, además el paro nacional del 28 de abril de 2021 y la emergencia sanitaria debida a la pandemia por COVID-19 afectaban la movilización para realizar más vuelos).

Por estas razones fue importante tener categorizada y segmentada la información, para poder realizar el análisis exploratorio de datos pertinente y modelar los datos de una mejor forma. Para ello, es importante tomar los datos en una hora determinada, preferiblemente al medio día en donde el sol se encuentra en el punto más alto y el nivel de reflectancia es el más adecuado. Tener un esquema de muestreo es importante, definir la sección y puntos de medición a nivel de hoja y de planta permite relacionar mejor los datos obtenidos.

Teniendo en cuenta los aspectos que afectaron el experimento, al observar los errores obtenidos, la correlación y el nivel de precisión obtenido, el modelo que más se destaca es la máquina de vectores, en el cual los valores atípicos tienen menos impacto debido a la segmentación de hiperplanos. Otro modelo que se destaca por su buen rendimiento y correlación es el bosque aleatorio, el cual al promediar los resultados de árboles de decisión que no se ven entre sí, disminuye el error que pueden generar datos extremos. Al momento de revisar la multicolinealidad a través del análisis de la correlación entre los índices de vegetación se destaca la división entre dos grupos, el primero relaciona los índices NDVI, RVI y NRVI. El segundo grupo es el que contiene los índices GNDVI, GCI y CVI. Para la mayoría de los modelos, exceptuando la regresión lineal, el segundo grupo destaca por obtener los errores más bajos, y tiene todo el sentido, ya que como se ha expresado en varias ocasiones, según la literatura son índices que se utilizan para estimar el contenido de clorofila. En la regresión lineal el índice que se destacó fue el NRVI (aunque fue cuestión de decimales), el cual es similar al NDVI, pero además reduce los efectos de la topografía, la iluminación y los efectos atmosféricos, lo cual ayudó en momento en que las nubes nublaban la captura de las imágenes multiespectrales.

Para finalizar, la salud del cultivo en general se estima como adecuada en cada planta tomando de referencia el valor de clorofila de referencia y el estimado, ya que

este supera los 54 CCI en cada planta. En términos generales la cámara mapir Survey aerotransportada en VANT fue útil para estimar la salud de cultivos de café, no obstante, se aconseja tener más indicativos (carbono, nitrógeno, abono, entre otros) a parte de la clorofila medida con el CCM 200 PLUS para alimentar los modelos de predicción.

LIMITACIONES

Como se ha mencionado en diferentes secciones, en el desarrollo de esta investigación existieron varias limitaciones, una de las más grandes fue la dificultad de no poder ir frecuentemente a tomar muestras con el clorofilómetro y los vuelos del VANT, esto debido a la coyuntura por la que está pasando el mundo debida a la pandemia por COVID-19, sumado a la problemática socio cultural de Colombia y el Cauca en el marco del paro nacional. Lo anterior en combinación con la variabilidad climática, en donde en varias ocasiones se presentaron lluvias, lo que no permitió continuar con la captura de datos, lo que influyó en que el volumen de datos obtenido no fuera de la cantidad esperada. Esto en conjunto con las variables climáticas tales como: lluvia, nubosidad, viento, entre otras, hicieron que el vuelo del VANT y la concentración de clorofila se vieran afectadas, lo cual altera la relación que se quiere establecer entre los índices de vegetación (calculados a partir de las imágenes multiespectrales) y el valor CCI, que se toma manualmente, proceso que tiene una duración aproximada de 8 horas, espaciado aproximadamente por 4 horas del vuelo del VANT.

Otro inconveniente que se tuvo en el desarrollo del proyecto fue la falta de acceso a un cultivo tecnificado que permitiera tener en cuenta otras variables que se relacionan con la salud del cultivo como por ejemplo el nitrógeno, abono utilizado y carbono, lo cual permitiría tener más variables y datos a relacionar para entrenar los modelos, no sólo la clorofila capturada por el CCM200. Las restricciones de acceso fueron atribuidas a la pandemia y al paro nacional.

APORTES

En el presente trabajo de investigación se destacan los siguientes aportes:

- Reporte de la revisión bibliométrica realizada (CAPÍTULO 2).
- Poster y artículo titulado 'Uso de UAV equipados con cámaras multiespectrales para el análisis de cultivos de café' IEEE COLCOM (*Colombian Conference on Communications and Computing*) del 7-8 de agosto 2020 (ANEXO J).
- Tutoriales para el manejo de las herramientas QGIS (ANEXO A) y *mission planner* para el procesamiento de imágenes y de sistemas aéreos no tripulados (ANEXO B).
- Repositorio GIT con los códigos implementados para los modelos de aprendizaje automático.
- Artículo que contiene los resultados de la investigación (ANEXO K).
- Banco de imágenes multiespectrales recolectadas (ANEXO L).

5. Trabajos de futuros

Teniendo en cuenta las oportunidades de investigación que se abren con el desarrollo de este proyecto de investigación, se proponen los siguientes trabajos futuros.

1. Con el fin de aumentar el tamaño de datos, es importante tomar un número mayor de muestras, de ser posible con varios equipos (más dispositivos para medir variables físicas en campo). Es importante, tener en cuenta otras variables físicas como el nitrógeno, cantidad y tipo de abono utilizado, carbono, entre otras, de un lote tecnificado con el objetivo de mejorar la estimación y tener en cuenta más factores que influyen en ella. Esto acompañado de otros modelos de entrenamiento como vecinos cercanos, árbol de decisiones, bosque aleatorio, entre otros.
2. Otra posibilidad, es, si se tiene el acceso, trabajar con cultivos enfermos segmentados por tipo de enfermedad, con el fin de analizar dicho cultivo y comportamiento a través de las imágenes multiespectrales y diferentes parámetros medidos en tierra, esto permite caracterizar el cultivo y realizar segmentación automática.
3. Una variante que se puede agregar a este estudio y a lo mencionado en los puntos anteriores de esta sección, es realizar los análisis con cámaras de mayor alcance y más bandas para tener en cuenta más índices de vegetación y una mejor resolución. Además, el tener acceso a un geolocalizador de precisión que acompañe las medidas en tierra, ayudaría a mejorar la calidad y rectificación del ortomosaico.
4. Las etapas del café son prolongadas, por lo cual se propone realizar un análisis de los diferentes ciclos del café, para validar la relación entre los valores de clorofila en cada etapa del crecimiento y el estado de salud.
5. Realizar el estudio para diferentes cultivos agrícolas representativos de Colombia como el plátano, el banano, aguacate, caña, entre otros.
6. Un aspecto importante para validar y realizar pruebas es utilizar nuevos datos de entrada para analizar el comportamiento de los modelos, es decir, utilizar generalización, ya que tomar nuevos datos según el panorama nacional y el tiempo no eran viables en nuestro caso.
7. Crear un prototipo en el cual se suban las fotografías multiespectrales capturadas por VANT y arroje directamente el estado de salud de la planta y/o cultivo de café.

REFERENCIAS

- [1] International Coffee Organization, "International Coffee Organization - What's New." [Online]. Available: <http://www.ico.org/>. [Accessed: 31-Oct-2019].
- [2] J. Avelino *et al.*, "The coffee rust crises in Colombia and Central America (2008–2013): impacts, plausible causes and proposed solutions," *Food Security*, vol. 7, no. 2, pp. 303–321, 2015.
- [3] M. M. Machado-Vargas, C. I. Nicholls-Estrada, and L. A. Ríos-Osorio, "Social-ecological resilience of small-scale coffee production in the porce river basin, antioquia (Colombia)," *Idesia*, vol. 36, no. 3, pp. 141–151, 2018.
- [4] C. J. Arias Barrera and J. F. Prada Hernández, "La producción cafetera y su impacto en el crecimiento económico del departamento del Huila, Colombia," *Ánfora*, vol. 24, no. 42, pp. 45–66, 2017.
- [5] K. D. Rodríguez Vargas, "Evolución de la producción cafetera y su contribución a la competitividad en el periodo 2010-2017," Universidad católica de Colombia, 2018.
- [6] J. A. Bolaños, L. Campo, and J. C. Corrales, "Characterization in the Visible and Infrared Spectrum of Agricultural Crops from a Multirotor Air Vehicle," vol. 687, P. Angelov, J. A. Iglesias, and J. C. Corrales, Eds. Cham: Springer International Publishing, 2018, pp. 29–43.
- [7] National Research Council, *Precision Agriculture in the 21st Century: Geospatial and Information Technologies in Crop Management*. Washington, DC: The National Academies Press, 1997.
- [8] J. Navia, I. Mondragon, D. Patino, and J. Colorado, "Multispectral mapping in agriculture: Terrain mosaic using an autonomous quadcopter UAV," *2016 International Conference on Unmanned Aircraft Systems (ICUAS)*, IEEE, pp. 1351–1358, Jun-2016.
- [9] T. Duan, S. C. Chapman, Y. Guo, and B. Zheng, "Dynamic monitoring of NDVI in wheat agronomy and breeding trials using an unmanned aerial vehicle," *Field Crops Research*, vol. 210, no. March, Elsevier, pp. 71–80, Aug-2017.
- [10] S. Candiago, F. Remondino, M. De Giglio, M. Dubbini, and M. Gattelli, "Evaluating Multispectral Images and Vegetation Indices for Precision Farming Applications from UAV Images," *Remote Sensing*, vol. 7, no. 4, pp. 4026–4047, 02-Apr-2015.
- [11] E. Pino V., "Los drones una herramienta para una agricultura eficiente: un futuro de alta tecnología," *Idesia (Arica)*, vol. 37, pp. 75–84, 2019.
- [12] P. M. Aguayo, "Apuntes de Teledetección: Índices de vegetación," Chile, 2013.

- [13] S. Kharuf-Gutierrez, L. Hernandez-Santana, R. Orozco-Morales, O. de la C. Aday Diaz, and I. Delgado Mora, "Análisis de imágenes multiespectrales adquiridas con vehículos aéreos no tripulados," *Ingeniería Electrónica, Automática y Comunicaciones*, vol. 39, scielocu, pp. 79–91, 2018.
- [14] Z. Liu, J. Shi, L. Zhang, and J. Huang, "Discrimination of rice panicles by hyperspectral reflectance data based on principal component analysis and support vector classification," *Journal of Zhejiang University SCIENCE B*, vol. 11, no. 1, pp. 71–78, 2010.
- [15] M. Katty Gómez Almonte, "Índice de vegetación en áreas del bosque seco del noroeste del Perú a partir de imágenes satelitales," Universidad de Piura, 2005.
- [16] J. Wijidechakul, S. Sasaki, Y. Kiyoki, and C. Koopipat, "UAV-based multispectral image analysis system with semantic computing for agricultural health conditions monitoring and real-time management," *Proceedings - 2016 International Electronics Symposium, IES 2016*, IEEE, pp. 459–464, 2017.
- [17] J. C. Rodríguez Villamizar, "Sistemas multiespectrales e hiperespectrales para la observación del territorio : análisis y aplicación a la prospección de hidrocarburos.," 2015.
- [18] D. Wu and D.-W. Sun, "Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment: A review — Part I: Fundamentals," *Innovative Food Science & Emerging Technologies*, vol. 19, pp. 1–14, Jul-2013.
- [19] Y.-Y. Pu, Y.-Z. Feng, and D.-W. Sun, "Recent Progress of Hyperspectral Imaging on Quality and Safety Inspection of Fruits and Vegetables: A Review," *Comprehensive Reviews in Food Science and Food Safety*, vol. 14, no. 2, pp. 176–188, Mar-2015.
- [20] M. J. Khan, H. S. Khan, A. Yousaf, K. Khurshid, and A. Abbas, "Modern Trends in Hyperspectral Image Analysis: A Review," *IEEE Access*, vol. 6, no. c, pp. 14118–14129, 2018.
- [21] F. López-Granados, "Weed detection for site-specific weed management: mapping and real-time approaches," *Weed Research*, vol. 51, no. 1, pp. 1–11, Feb-2011.
- [22] C. Zeng, D. J. King, M. Richardson, and B. Shan, "Fusion of Multispectral Imagery and Spectrometer Data in UAV Remote Sensing," *Remote Sensing*, vol. 9, no. 7, p. 696, 06-Jul-2017.
- [23] Ministerio de Hacienda y Crédito Público, "Presupuesto Ciudadano 2019 Ministerio de Hacienda y Crédito Público," Bogotá, 2018.
- [24] H. C. and M. O. S. Park, A. Nolan, D. Ryu, S. Fuentes, E. Hernandez, "Estimation of crop water stress in a nectarine orchard using high-resolution

imagery from unmanned aerial vehicle (UAV) S.," *Remote Sens.*, vol. 9, no. 8, p. 828, Aug. 2015.

- [25] D. M. Gates, H. J. Keegan, J. C. Schleiter, and V. R. Weidner, "Spectral Properties of Plants," *Appl. Opt.*, vol. 4, no. 1, p. 11, 1965.
- [26] A. A. Gitelson, Y. Gritz, and M. N. Merzlyak, "Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non destructive chlorophyll assessment in higher plant leaves," *J. Plant Physiol.*, vol. 160, pp. 271–282, 2003.
- [27] J. Pérez Rave, *Revisión sistemática de literatura en ingeniería*. IDINNOV, 2019.
- [28] D. Carrizo and C. Moller, "Methodological structures of systematic literature review in software engineering: A systematic mapping study," *Ingeniare*, vol. 26. Universidad de Tarapaca, pp. 45–54, 2018.
- [29] B. Kitchenham, "Procedures for Performing Systematic Reviews," 2004.
- [30] M. J. Cobo, A. G. López-Herrera, E. Herrera-Viedma, and F. Herrera, "SciMAT: A new science mapping analysis software tool," *J. Am. Soc. Inf. Sci. Technol.*, vol. 63, no. 8, pp. 1609–1630, Aug. 2012.
- [31] G. Valencia A, "Relación entre el índice de área foliar y la productividad del cafeto," *Cenicafé (Colombia)*v. 24(4) p. 79-89, 1973.
- [32] E. A. Bacelar, D. L. Santos, J. M. Moutinho-Pereira, B. C. Gonçalves, H. F. Ferreira, and C. M. Correia, "Immediate responses and adaptative strategies of three olive cultivars under contrasting water availability regimes: Changes on structure and chemical composition of foliage and oxidative damage," *Plant Science*, vol. 170, no. 3, pp. 596–605, Mar-2006.
- [33] J. D. Hiscox and G. F. Israelstam, "A method for the extraction of chlorophyll from leaf tissue without maceration," *Canadian Journal of Botany*, vol. 57, no. 12, pp. 1332–1334, 15-Jun-1979.
- [34] optisciences, "CCM-200." [Online]. Available: <https://www.optisci.com/ccm-200.html>. [Accessed: 02-Jun-2020].
- [35] A. González, G. Amarillo, M. Amarillo, and F. Sarmiento, "Drones Aplicados a la Agricultura de Precisión," *Publicaciones e Investig.*, vol. 10, pp. 23–37, Mar. 2016.
- [36] A. Ahmad and A. M. Samad, "Aerial mapping using high resolution digital camera and unmanned aerial vehicle for Geographical Information System," in *Proceedings - CSPA 2010: 2010 6th International Colloquium on Signal Processing and Its Applications*, 2010.
- [37] B. Brisco, R. J. Brown, T. Hirose, H. Mc Naim, and K. Staenz, "Precision agriculture and the role of remote sensing: A review," *Can. J. Remote Sens.*, vol. 24, no. 3, pp. 315–327, 1998.

- [38] J. F. F. Sabins, "Remote sensing: Principles and interpretation." W.H. Freeman and Company, New York, NY, 01-Jan-1986.
- [39] S. K. Balasundram, "A Review: The Role of Remote Sensing in Precision Agriculture," 2010.
- [40] V. Andreo and D. V. Andreo, "Remote Sensing and Geographic Information Systems in Precision Farming LST-Land Surface Temperature time series View project MEW 9 Seminario Remote Sensing and Geographic Information Systems in Precision Farming," 2013.
- [41] A. Vibhute and S. K. Bodhe, "Application of Image Processing in Agriculture: A Survey," *Int. J. Comput. Appl.*, vol. 52, pp. 0975 – 8887, 2012.
- [42] G. Pajares, "Overview and Current Status of Remote Sensing Applications Based on Unmanned Aerial Vehicles (UAVs)," *Photogrammetric Engineering & Remote Sensing*, vol. 81, no. 4, pp. 281–330, 01-Apr-2015.
- [43] PlanetaDrones, "Drone 3DR Solo - Todo lo que debes saber - Planeta drones." [Online]. Available: <https://planetadrones.es/drone-3dr-solo/>. [Accessed: 07-Nov-2019].
- [44] J. Luque Ordóñez, "Espectro electromagnético y espectro radioeléctrico," *Man. Form. ACTA, ISSN 1888-6051, N^o. 62, 2012, págs. 17-31*, no. 62, pp. 17–31, 2012.
- [45] S. Sistemas & Telemática ; Universidad Icesi Kharuf-Gutierrez, R. Orozco-Morale, O. Aday, E. Pineda, and S. Kharuf-Gutierrez, "Multispectral aerial image processing system for precision agriculture," *Sist. Telemática*, vol. 16, no. 47, pp. 45–58, 2018.
- [46] A. Roman-Gonzalez and N. I. Vargas-Cuentas, "Análisis de imágenes hiperespectrales," *Revista Ingeniería & Desarrollo*, no. 9, pp. 14–17, 2013.
- [47] ADSU Filmaciones, "Herramientas de Análisis." [Online]. Available: <http://sitioftp.com/ADSU/herramientas-analisis.pdf>. [Accessed: 02-Nov-2019].
- [48] F. A. Vega, F. C. Ramírez, M. P. Saiz, and F. O. Rosúa, "Multi-temporal imaging using an unmanned aerial vehicle for monitoring a sunflower crop," *Biosystems Engineering*, vol. 132, pp. 19–27, Apr-2015.
- [49] J. Geipel, J. Link, J. Wirwahn, and W. Claupein, "A Programmable Aerial Multispectral Camera System for In-Season Crop Biomass and Nitrogen Content Estimation," *Agriculture*, vol. 6, no. 1, p. 4, Jan. 2016.
- [50] A. B. Potgieter *et al.*, "Multi-Spectral Imaging from an Unmanned Aerial Vehicle Enables the Assessment of Seasonal Leaf Area Dynamics of Sorghum Breeding Lines," *Front. Plant Sci.*, vol. 8, no. September, pp. 1–11, Sep. 2017.
- [51] C. Lelong, P. Burger, G. Jubelin, B. Roux, S. Labbé, and F. Baret, "Assessment of Unmanned Aerial Vehicles Imagery for Quantitative

Monitoring of Wheat Crop in Small Plots,” *Sensors*, vol. 8, no. 5, pp. 3557–3585, May 2008.

- [52] H. Zheng *et al.*, “Evaluation of RGB, Color-Infrared and Multispectral Images Acquired from Unmanned Aerial Systems for the Estimation of Nitrogen Accumulation in Rice,” *Remote Sensing*, vol. 10, no. 6, p. 824, 25-May-2018.
- [53] L. Peroni Venancio *et al.*, “Potential of using spectral vegetation indices for corn green biomass estimation based on their relationship with the photosynthetic vegetation sub-pixel fraction,” *Agric. Water Manag.*, vol. 236, p. 106155, Jun. 2020.
- [54] E. C. M. Noyons, H. F. Moed, and A. F. J. Van Raan, “Integrating research performance analysis and science mapping,” *Scientometrics*, vol. 46, no. 3, pp. 591–604, 1999.
- [55] H. Small, “Visualizing science by citation mapping,” *J. Am. Soc. Inf. Sci.*, vol. 50, no. 9, pp. 799–813, Jan. 1999.
- [56] M. A. Maleka, P. Budhwarb, and B. S. Reiche, “Sources of support and expatriation: A multiple stakeholder perspective of expatriate adjustment and performance in Malaysia,” *Int. J. Hum. Resour. Manag.*, vol. 26, no. 2, pp. 258–276, Jan. 2015.
- [57] Naciones Unidas, “Convención Marco sobre el Cambio Climático,” 2015.
- [58] Intergovernmental Panel on Climate Change, “Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty,” 2018.
- [59] X. Zhou *et al.*, “Predicting grain yield in rice using multi-temporal vegetation indices from UAV-based multispectral and digital imagery,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 130, pp. 246–255, Aug-2017.
- [60] M. Romero, Y. Luo, B. Su, and S. Fuentes, “Vineyard water status estimation using multispectral imagery from an UAV platform and machine learning algorithms for irrigation scheduling management,” *Comput. Electron. Agric.*, vol. 147, no. January, pp. 109–117, Apr. 2018.
- [61] M. Nagle, B. Mahayothee, P. Rungpichayapichet, S. Janjai, and J. Müller, “Effect of irrigation on near-infrared (NIR) based prediction of mango maturity,” *Scientia Horticulturae*, vol. 125, no. 4, pp. 771–774, 2010.
- [62] P. Williams, P. Geladi, G. Fox, and M. Manley, “Maize kernel hardness classification by near infrared (NIR) hyperspectral imaging and multivariate data analysis,” *Analytica Chimica Acta*, vol. 653, no. 2, pp. 121–130, 2009.
- [63] A. Xie, D. W. Sun, Z. Xu, and Z. Zhu, “Rapid detection of frozen pork quality without thawing by Vis-NIR hyperspectral imaging technique,” *Talanta*, vol. 139, Elsevier, pp. 208–215, 2015.

- [64] C. Gomez, R. A. Viscarra Rossel, and A. B. McBratney, "Soil organic carbon prediction by hyperspectral remote sensing and field vis-NIR spectroscopy: An Australian case study," *Geoderma*, vol. 146, no. 3–4, pp. 403–411, 2008.
- [65] A. Tellaeché, G. Pajares, X. P. Burgos-Artizzu, and A. Ribeiro, "A computer vision approach for weeds identification through Support Vector Machines," *Applied Soft Computing Journal*, vol. 11, no. 1, pp. 908–915, 2011.
- [66] F. M. Satoru Koda, Keiichi Mochida, Komaki Inoue, Takashi Hirayama, Shojiro Tanaka, Ryuei Nishii, "Computer vision-based phenotyping for improvement of plant productivity: a machine learning perspective," *The Japanese journal of gastroenterological surgery*, vol. 8, no. 5, pp. 529–530, 1975.
- [67] B. Zhang *et al.*, "Principles, developments and applications of computer vision for external quality inspection of fruits and vegetables: A review," *Food Research International*, vol. 62, Elsevier Ltd, pp. 326–343, 2014.
- [68] P. Jackman, D. W. Sun, and P. Allen, "Recent advances in the use of computer vision technology in the quality assessment of fresh meats," *Trends in Food Science and Technology*, vol. 22, no. 4, Elsevier Ltd, pp. 185–197, 2011.
- [69] J. M. Soriano-Disla, L. J. Janik, R. A. Viscarra Rossel, L. M. MacDonald, and M. J. McLaughlin, "The performance of visible, near-, and mid-infrared reflectance spectroscopy for prediction of soil physical, chemical, and biological properties," *Applied Spectroscopy Reviews*, vol. 49, no. 2, pp. 139–186, 2014.
- [70] E. Anastasaki *et al.*, "Differentiation of saffron from four countries by mid-infrared spectroscopy and multivariate analysis," *European Food Research and Technology*, vol. 230, no. 4, pp. 571–577, 2010.
- [71] K. Theodoridou and P. Yu, "Application Potential of ATR-FT/IR Molecular Spectroscopy in Animal Nutrition: Revelation of Protein Molecular Structures of Canola Meal and Presscake, As Affected by Heat-Processing Methods, in Relationship with Their Protein Digestive Behavior and Utili," *Journal of Agricultural and Food Chemistry*, vol. 61, no. 23, pp. 5449–5458, 2013.
- [72] H. Xin and P. Yu, "Chemical profile, energy values, and protein molecular structure characteristics of biofuel/bio-oil co-products (Carinata Meal) in comparison with canola meal," *Journal of Agricultural and Food Chemistry*, vol. 61, no. 16, pp. 3926–3933, 2013.
- [73] Samadi, K. Theodoridou, and P. Yu, "Detect the sensitivity and response of protein molecular structure of whole canola seed (yellow and brown) to different heat processing methods and relation to protein utilization and availability using ATR-FT/IR molecular spectroscopy with chemometrics," *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, vol. 105, pp. 304–313, 2013.

- [74] S. Abeysekera, Samadi, and P. Yu, "Response and sensitivity of lipid related molecular structure to wet and dry heating in Canola tissue," *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, vol. 90, Elsevier B.V., pp. 63–71, 2012.
- [75] H. Xin and P. Yu, "Using ATR-FT/iR to detect carbohydrate-related molecular structure features of carinata meal and their in situ residues of ruminal fermentation in comparison with canola meal," *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, vol. 114, Elsevier B.V., pp. 599–606, 2013.
- [76] F. Tao and Y. Peng, "A method for nondestructive prediction of pork meat quality and safety attributes by hyperspectral imaging technique," *Journal of Food Engineering*, vol. 126, Elsevier Ltd, pp. 98–106, 2014.
- [77] Z. Xiong, A. Xie, D. W. Sun, X. A. Zeng, and D. Liu, "Applications of Hyperspectral Imaging in Chicken Meat Safety and Quality Detection and Evaluation: A Review," *Critical Reviews in Food Science and Nutrition*, vol. 55, no. 9, pp. 1287–1301, 2015.
- [78] H. Pu, D. W. Sun, J. Ma, and J. H. Cheng, "Classification of fresh and frozen-thawed pork muscles using visible and near infrared hyperspectral imaging and textural analysis," *Meat Science*, vol. 99, pp. 81–88, 2014.
- [79] G. ElMasry, N. Wang, C. Vigneault, J. Qiao, and A. ElSayed, "Early detection of apple bruises on different background colors using hyperspectral imaging," *LWT - Food Science and Technology*, vol. 41, no. 2, pp. 337–345, 2008.
- [80] Y. Peng and R. Lu, "Analysis of spatially resolved hyperspectral scattering images for assessing apple fruit firmness and soluble solids content," *Postharvest Biol. Technol.*, vol. 48, no. 1, pp. 52–62, 2008.
- [81] W. Huang, J. Li, Q. Wang, and L. Chen, "Development of a multispectral imaging system for online detection of bruises on apples," *J. Food Eng.*, vol. 146, pp. 62–71, 2015.
- [82] G. Alberti *et al.*, "A LiDAR-based approach for a multi-purpose characterization of alpine forests: An italian case study," *IForest*, vol. 6, no. 1, pp. 156–168, 2013.
- [83] Y. Hirata, R. Tabuchi, P. Patanaponpaiboon, S. Pongparn, R. Yoneda, and Y. Fujioka, "Estimation of aboveground biomass in mangrove forests using high-resolution satellite data," *Journal of Forest Research*, vol. 19, no. 1, pp. 34–41, 2014.
- [84] J. Mohammadi and S. Shataee, "Possibility investigation of tree diversity mapping using Landsat ETM+ data in the Hyrcanian forests of Iran," *Remote Sensing of Environment*, vol. 114, no. 7, Elsevier B.V., pp. 1504–1512, 2010.
- [85] L. Pan, Q. Zhu, R. Lu, and J. M. McGrath, "Determination of sucrose content in sugar beet by portable visible and near-infrared spectroscopy," *Food*

Chemistry, vol. 167, Elsevier Ltd, pp. 264–271, 2015.

- [86] L. Pan, R. Lu, Q. Zhu, J. M. McGrath, and K. Tu, "Measurement of moisture, soluble solids, sucrose content and mechanical properties in sugar beet using portable visible and near-infrared spectroscopy," *Postharvest Biology and Technology*, vol. 102, Elsevier India Pvt Ltd, pp. 42–50, 2015.
- [87] A. K. Mahlein, U. Steiner, H. W. Dehne, and E. C. Oerke, "Spectral signatures of sugar beet leaves for the detection and differentiation of diseases," *Precision Agriculture*, vol. 11, no. 4, pp. 413–431, 2010.
- [88] A. Kross, H. McNairn, D. Lapen, M. Sunohara, and C. Champagne, "Assessment of RapidEye vegetation indices for estimation of leaf area index and biomass in corn and soybean crops," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 34, no. 1, pp. 235–248, 2015.
- [89] J. Geipel, J. Link, and W. Claupein, "Combined spectral and spatial modeling of corn yield based on aerial images and crop surface models acquired with an unmanned aircraft system," *Remote Sensing*, vol. 6, no. 11, pp. 10335–10355, 2014.
- [90] X. Ran, L. Wang, Y. Lin, J. Hao, and D. Cao, "Syntheses, characterization and biological studies of zinc(II), copper(II) and cobalt(II) complexes with Schiff base ligand derived from 2-hydroxy-1-naphthaldehyde and selenomethionine," *Appl. Organomet. Chem.*, vol. 24, no. 10, pp. 741–747, 2010.
- [91] O. B. Chanu, A. Kumar, A. Ahmed, and R. A. Lal, "Synthesis and characterisation of heterometallic trinuclear copper(II) and zinc(II) complexes derived from bis(2-hydroxy-1-naphthaldehyde) oxaloyldihydrazone," *J. Mol. Struct.*, vol. 1007, pp. 257–274, 2012.
- [92] X. Ran, L. Wang, D. Cao, Y. Lin, and J. Hao, "Synthesis, characterization and in vitro biological activity of cobalt(II), copper(II) and zinc(II) Schiff base complexes derived from salicylaldehyde and D,L-selenomethionine," *Applied Organometallic Chemistry*, vol. 25, no. 1, pp. 9–15, 2011.
- [93] E. I. Galanzha, J. W. Kim, and V. P. Zharov, "Nanotechnology-based molecular photoacoustic and photothermal flow cytometry platform for in-vivo detection and killing of circulating cancer stem cells," *Journal of Biophotonics*, vol. 2, no. 12, pp. 725–735, 2009.
- [94] R. Gasper, T. Mijatovic, A. Bénard, A. Derenne, R. Kiss, and E. Goormaghtigh, "FTIR spectral signature of the effect of cardiostimulatory steroids with antitumoral properties on a prostate cancer cell line," *Biochim. Biophys. Acta - Mol. Basis Dis.*, vol. 1802, no. 11, pp. 1087–1094, 2010.
- [95] R. A. Naidu, E. M. Perry, F. J. Pierce, and T. Mekuria, "The potential of spectral reflectance technique for the detection of Grapevine leafroll-associated virus-3 in two red-berried wine grape cultivars," *Comput. Electron. Agric.*, vol. 66, no. 1, pp. 38–45, 2009.

- [96] C. B. Singh, D. S. Jayas, J. Paliwal, and N. D. G. White, "Detection of insect-damaged wheat kernels using near-infrared hyperspectral imaging," *J. Stored Prod. Res.*, vol. 45, no. 3, pp. 151–158, 2009.
- [97] S. Adelabu, O. Mutanga, and E. Adam, "Evaluating the impact of red-edge band from Rapideye image for classifying insect defoliation levels," *ISPRS J. Photogramm. Remote Sens.*, vol. 95, pp. 34–41, 2014.
- [98] C. B. Singh, D. S. Jayas, J. Paliwal, and N. D. G. White, "Identification of insect-damaged wheat kernels using short-wave near-infrared hyperspectral and digital colour imaging," *Comput. Electron. Agric.*, vol. 73, no. 2, pp. 118–125, 2010.
- [99] R. Moschetti, R. P. Haff, S. Saranwong, D. Monarca, M. Cecchini, and R. Massantini, "Nondestructive detection of insect infested chestnuts based on NIR spectroscopy," *Postharvest Biol. Technol.*, vol. 87, pp. 88–94, 2014.
- [100] E. Zamora-Rojas, B. Aernouts, A. Garrido-Varo, W. Saeys, D. Pérez-Marín, and J. E. Guerrero-Ginel, "Optical properties of pig skin epidermis and dermis estimated with double integrating spheres measurements," *Innov. Food Sci. Emerg. Technol.*, vol. 20, pp. 343–349, 2013.
- [101] E. Zamora-Rojas *et al.*, "Understanding near infrared radiation propagation in pig skin reflectance measurements," *Innov. Food Sci. Emerg. Technol.*, vol. 22, pp. 137–146, 2014.
- [102] F. Vaudelle and J. P. L'Huillier, "Influence of the size and skin thickness of apple varieties on the retrieval of internal optical properties using Vis/NIR spectroscopy: A Monte Carlo-based study," *Comput. Electron. Agric.*, vol. 116, pp. 137–149, 2015.
- [103] C. Stone, T. Penman, and R. Turner, "Managing drought-induced mortality in *Pinus radiata* plantations under climate change conditions: A local approach using digital camera data," *For. Ecol. Manage.*, vol. 265, pp. 94–101, 2012.
- [104] V. M. Sellitto *et al.*, "Comparing two different spectroscopic techniques for the characterization of soil iron oxides: Diffuse versus bi-directional reflectance," *Geoderma*, vol. 251–252, no. 1–2, pp. 2–9, 2015.
- [105] N. Prieto *et al.*, "At line prediction of PUFA and biohydrogenation intermediates in perirenal and subcutaneous fat from cattle fed sunflower or flaxseed by near infrared spectroscopy," *Meat Science*, vol. 90, no. 1, Elsevier Ltd, pp. 43–51, 2013.
- [106] N. Prieto, M. E. R. Dugan, O. López-Campos, T. A. McAllister, J. L. Aalhus, and B. Uttaro, "Near infrared reflectance spectroscopy predicts the content of polyunsaturated fatty acids and biohydrogenation products in the subcutaneous fat of beef cows fed flaxseed," *Meat Science*, vol. 90, no. 1, Elsevier Ltd., pp. 43–51, 2012.
- [107] X. Yao *et al.*, "Using leaf dry matter to quantify the critical nitrogen dilution

- curve for winter wheat cultivated in eastern China," *F. Crop. Res.*, vol. 159, pp. 33–42, 2014.
- [108] C. M. Souza *et al.*, "Ten-year landsat classification of deforestation and forest degradation in the brazilian amazon," *Remote Sens.*, vol. 5, no. 11, pp. 5493–5513, 2013.
- [109] D. Lu, M. Batistela, and E. Moran, "Change Detection in the Brazilian Amazon," *Photogramm Eng Remote Sens.*, vol. 74, no. 4, pp. 421–430, 2008.
- [110] D. Lu, M. Batistella, E. Moran, S. Hetrick, D. Alves, and E. Brondizio, "Fractional forest cover mapping in the brazilian amazon with a combination of modis and tm images," *Int. J. Remote Sens.*, vol. 32, no. 22, pp. 7131–7149, 2011.
- [111] C. Brannstrom, W. Jepson, A. M. Filippi, D. Redo, Z. Xu, and S. Ganesh, "Land change in the Brazilian Savanna (Cerrado), 1986-2002: Comparative analysis and implications for land-use policy," *Land use policy*, vol. 25, no. 4, pp. 579–595, 2008.
- [112] F. Li *et al.*, "Estimating N status of winter wheat using a handheld spectrometer in the North China Plain," *F. Crop. Res.*, vol. 106, no. 1, pp. 77–85, 2008.
- [113] K. Yu, F. Li, M. L. Gnyp, Y. Miao, G. Bareth, and X. Chen, "Remotely detecting canopy nitrogen concentration and uptake of paddy rice in the Northeast China Plain," *ISPRS J. Photogramm. Remote Sens.*, vol. 78, pp. 102–115, 2013.
- [114] W. Koppe *et al.*, "Evaluating multispectral and hyperspectral satellite remote sensing data for estimating winter wheat growth parameters at regional scale in the North China plain," *Photogramm. Fernerkundung, Geoinf.*, vol. 2010, no. 3, pp. 167–178, 2010.
- [115] C. T. Morehart, "Mapping ancient chinampa landscapes in the Basin of Mexico: A remote sensing and GIS approach," *J. Archaeol. Sci.*, vol. 39, no. 7, pp. 2541–2551, 2012.
- [116] S. Deng, M. Kato, Q. Guan, N. Yin, and M. Li, "Interpretation of forest resources at the individual tree level at Purple Mountain, Nanjing City, China, using WorldView-2 imagery by combining GPS, RS and GIS technologies," *Remote Sens.*, vol. 6, no. 1, pp. 87–110, 2013.
- [117] M. Jayanthi, "Monitoring brackishwater aquaculture development using multi-spectral satellite data and GIS - a case study near Pichavaram mangroves south-east coast of India," *Indian J. Fish.*, vol. 58, no. 1, pp. 85–90, 2011.
- [118] P. J. Zarco-Tejada, V. González-Dugo, and J. A. J. Berni, "Fluorescence, temperature and narrow-band indices acquired from a UAV platform for water stress detection using a micro-hyperspectral imager and a thermal

- camera," *Remote Sens. Environ.*, vol. 117, pp. 322–337, Feb. 2012.
- [119] J. Bendig *et al.*, "Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 39, no. 1, pp. 79–87, Jul. 2015.
- [120] P. J. Zarco-Tejada, M. L. Guillén-Climent, R. Hernández-Clemente, A. Catalina, M. R. González, and P. Martín, "Estimating leaf carotenoid content in vineyards using high resolution hyperspectral imagery acquired from an unmanned aerial vehicle (UAV)," *Agric. For. Meteorol.*, vol. 171–172, pp. 281–294, Apr. 2013.
- [121] K. Uto, H. Seki, G. Saito, and Y. Kosugi, "Characterization of rice paddies by a UAV-mounted miniature hyperspectral sensor system," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 6, no. 2, pp. 851–860, 2013.
- [122] J. Gómez-Sanchis, D. Lorente, E. Soria-Olivas, N. Aleixos, S. Cubero, and J. Blasco, "Development of a Hyperspectral Computer Vision System Based on Two Liquid Crystal Tuneable Filters for Fruit Inspection. Application to Detect Citrus Fruits Decay," *Food Bioprocess Technol.*, vol. 7, no. 4, pp. 1047–1056, 2014.
- [123] V. Šuštar *et al.*, "Identification of historical polymers using Near-Infrared Spectroscopy," *Polym. Degrad. Stab.*, vol. 47, no. 5, pp. 341–347, 2014.
- [124] A. Bilici, I. Kaya, and M. Yildirim, "A comparative study of 9,9-bis(4-aminophenyl)fluorene polymers prepared by catalytic and non-catalytic oxidative polymerisation methods," *Eur. Polym. J.*, vol. 47, no. 5, pp. 1005–1017, 2011.
- [125] C.-C. Yang *et al.*, "Development of multispectral imaging algorithm for detection of frass on mature red tomatoes," *Postharvest Biol. Technol.*, vol. 93, pp. 1–8, Jul. 2014.
- [126] S. Shataee, H. Weinaker, and M. Babanejad, "Plot-level forest volume estimation using airborne laser scanner and TM data, comparison of boosting and random forest tree regression algorithms," in *Procedia Environmental Sciences*, 2011, vol. 7, pp. 68–73.
- [127] Z. Duli, K. R. Reddy, V. G. Kakani, and J. J. Read, "Remote-sensing algorithms for estimating nitrogen uptake and nitrogen-use efficiency in cotton," *Acta Agric. Scand. Sect. B Soil Plant Sci.*, vol. 60, no. 6, pp. 500–509, 2010.
- [128] B. Waske, J. A. Benediktsson, K. Árnason, and J. R. Sveinsson, "Mapping of hyperspectral AVIRIS data using machine-learning algorithms," *Can. J. Remote Sens.*, vol. 35, no. sup1, pp. S106–S116, Jan. 2009.
- [129] S. Adelabu, O. Mutanga, E. Adam, and M. A. Cho, "Exploiting machine learning algorithms for tree species classification in a semiarid woodland

using RapidEye image,” *J. Appl. Remote Sens.*, vol. 7, no. 1, p. 073480, Nov. 2013.

- [130] R. Rud, M. Shoshany, and V. Alchanatis, “Spectral indicators for salinity effects in crops: A comparison of a new green indigo ratio with existing indices,” *Remote Sens. Lett.*, vol. 2, no. 4, pp. 289–298, 2011.
- [131] H. R. Matinfar, S. K. Alavi Panah, F. Zand, and K. Khodaei, “Detection of soil salinity changes and mapping land cover types based upon remotely sensed data,” *Arab. J. Geosci.*, vol. 6, no. 3, pp. 913–919, 2013.
- [132] S. Hamzeh *et al.*, “Estimating salinity stress in sugarcane fields with spaceborne hyperspectral: Vegetation indices,” *Int. J. Appl. Earth Obs. Geoinf.*, vol. 21, no. 1, pp. 282–290, 2012.
- [133] I. Nishidate *et al.*, “Evaluation of Cerebral Hemodynamics and Tissue Morphology of in Vivo Rat Brain Using Spectral Diffuse Reflectance Imaging,” *Appl. Spectrosc.*, vol. 71, no. 5, pp. 866–878, 2017.
- [134] S. P. Yadav, Y. Ibaraki, and S. Dutta Gupta, “Estimation of the chlorophyll content of micropropagated potato plants using RGB based image analysis,” *Plant Cell, Tissue Organ Cult.*, vol. 100, no. 2, pp. 183–188, Feb. 2010.
- [135] M. Karlson *et al.*, “Tree crown mapping in managed woodlands (Parklands) of semi-arid West Africa using WorldView-2 imagery and geographic object based image analysis,” *Plant Cell, Tissue Organ Cult.*, vol. 67, no. 1, pp. 22643–22669, Feb. 2014.
- [136] S. Dupuya, G. Lainé, J. Tassin, and J. M. Sarrailh, “Characterization of the horizontal structure of the tropical forest canopy using object-based LiDAR and multispectral image analysis,” *Int. J. Appl. Earth Obs. Geoinf.*, vol. 25, no. 1, pp. 76–86, 2013.
- [137] A. I. Ropodi, D. E. Pavlidis, F. Mohareb, E. Z. Panagou, and G. J. E. Nychas, “Multispectral image analysis approach to detect adulteration of beef and pork in raw meats,” *Food Res. Int.*, vol. 67, pp. 12–18, 2015.
- [138] G. Giovanelli, N. Sinelli, R. Beghi, R. Guidetti, and E. Casiraghi, “NIR spectroscopy for the optimization of postharvest apple management,” *Postharvest Biol. Technol.*, vol. 87, pp. 13–20, 2014.
- [139] K. Koenig *et al.*, “Comparative classification analysis of post-harvest growth detection from terrestrial LiDAR point clouds in precision agriculture,” *ISPRS J. Photogramm. Remote Sens.*, vol. 104, pp. 112–125, 2014.
- [140] A. R. Johnsen *et al.*, “Detection of early plant stress responses in hyperspectral images,” *ISPRS J. Photogramm. Remote Sens.*, vol. 49, no. 6, pp. 2261–2274, 2014.
- [141] A. R. Johnsen, B. P. Horgan, B. S. Hulke, and V. Cline, “Evaluation of remote sensing to measure plant stress in creeping bentgrass (*Agrostis stolonifera* L.) fairways,” *Crop Sci.*, vol. 49, no. 6, pp. 2261–2274, 2009.

- [142] Y. C. Yang, D. W. Sun, N. N. Wang, and A. Xie, "Real-time evaluation of polyphenol oxidase (PPO) activity in lychee pericarp based on weighted combination of spectral data and image features as determined by fuzzy neural network," *Talanta*, vol. 139, pp. 198–207, 2015.
- [143] M. Prevolnik *et al.*, "Classification of dry-cured hams according to the maturation time using near infrared spectra and artificial neural networks," *Meat Sci.*, vol. 96, no. 1, pp. 14–20, 2014.
- [144] Z. Y. Liu, H. F. Wu, and J. F. Huang, "Application of neural networks to discriminate fungal infection levels in rice panicles using hyperspectral reflectance and principal components analysis," *Comput. Electron. Agric.*, vol. 72, no. 2, pp. 99–106, 2010.
- [145] P. A. Gutiérrez, F. López-Granados, J. M. Peña-Barragán, M. Jurado-Expósito, M. T. Gómez-Casero, and C. Hervás-Martínez, "Mapping sunflower yield as affected by *Ridolfia segetum* patches and elevation by applying evolutionary product unit neural networks to remote sensed data," *Comput. Electron. Agric.*, vol. 60, no. 2, pp. 122–132, 2008.
- [146] H. J. He and D. W. Sun, "Microbial evaluation of raw and processed food products by Visible/Infrared, Raman and Fluorescence spectroscopy," *Trends Food Sci. Technol.*, vol. 46, no. 2, pp. 199–210, 2015.
- [147] D. L. Betemps *et al.*, "Non-destructive evaluation of ripening and quality traits in apples using a multiparametric fluorescence sensor," *J. Sci. Food Agric.*, vol. 92, no. 9, pp. 1855–1864, 2012.
- [148] C. C. Yang *et al.*, "Red to far-red multispectral fluorescence image fusion for detection of fecal contamination on apples," *J. Food Eng.*, vol. 108, no. 2, pp. 312–319, 2012.
- [149] K. Yu, G. Leufen, M. Hunsche, G. Noga, X. Chen, and G. Bareth, "Investigation of leaf diseases and estimation of chlorophyll concentration in seven barley varieties using fluorescence and hyperspectral indices," *Remote Sens.*, vol. 6, no. 1, pp. 64–86, 2013.
- [150] K. Briling *et al.*, "UV-induced fluorescence spectra and lifetime determination for detection of leaf rust (*Puccinia triticina*) in susceptible and resistant wheat (*Triticum aestivum*) cultivars," *J. Food Eng.*, vol. 92, no. 9, pp. 337–345, 2012.
- [151] S. M. O'Rourke *et al.*, "Comparing different multivariate calibration methods for the determination of soil organic carbon pools with visible to near infrared spectroscopy," *Geoderma*, vol. 166, no. 3–4, pp. 198–205, 2011.
- [152] G. N. Nóbrega *et al.*, "Evaluation of methods for quantifying organic carbon in mangrove soils from semi-arid region," *J. Soils Sediments*, vol. 15, no. 2, pp. 282–291, 2015.
- [153] S. M. O'Rourke and N. M. Holden, "Optical sensing and chemometric

- analysis of soil organic carbon - a cost effective alternative to conventional laboratory methods?," *Soil Use Manag.*, vol. 27, no. 2, pp. 143–155, 2011.
- [154] T. Liu *et al.*, "Non-destructive assessment of ascorbic acid in apples using near-infrared (NIR) spectroscopy together with partial least squares (PLS) regression," *Acta Hortic.*, vol. 1208, pp. 447–454, 2018.
- [155] J. H. Cheng and D. W. Sun, "Partial Least Squares Regression (PLSR) Applied to NIR and HSI Spectral Data Modeling to Predict Chemical Properties of Fish Muscle," *Food Eng. Rev.*, vol. 9, no. 1, pp. 36–49, 2017.
- [156] H. Lee, M. S. Kim, H. S. Lim, E. Park, W. H. Lee, and B. K. Cho, "Detection of cucumber green mottle mosaic virus-infected watermelon seeds using a near-infrared (NIR) hyperspectral imaging system: Application to seeds of the 'Sambok Honey' cultivar," *Biosyst. Eng.*, vol. 148, pp. 138–147, 2016.
- [157] A. M. West *et al.*, "Using multi-date satellite imagery to monitor invasive grass species distribution in post-wildfire landscapes: An iterative, adaptable approach that employs open-source data and software," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 59, pp. 135–146, 2017.
- [158] J. Zhang, Y. Huang, L. Yuan, G. Yang, L. Chen, and C. Zhao, "Using satellite multispectral imagery for damage mapping of armyworm (*Spodoptera frugiperda*) in maize at a regional scale," *Pest Manag. Sci.*, vol. 72, no. 2, pp. 335–348, 2016.
- [159] J. A. M. Demattê, C. T. Fongaro, R. Rizzo, and J. L. Safanelli, "Geospatial Soil Sensing System (GEOS3): A powerful data mining procedure to retrieve soil spectral reflectance from satellite images," *Remote Sens. Environ.*, vol. 212, no. August 2017, pp. 161–175, 2018.
- [160] Y. Feng, D. Lu, E. Moran, L. Dutra, M. Calvi, and M. de Oliveira, "Examining Spatial Distribution and Dynamic Change of Urban Land Covers in the Brazilian Amazon Using Multitemporal Multisensor High Spatial Resolution Satellite Imagery," *Remote Sens.*, vol. 9, no. 4, p. 381, 2017.
- [161] C. Wei, J. Huang, L. R. Mansaray, Z. Li, W. Liu, and J. Han, "Estimation and mapping of winter oilseed rape LAI from high spatial resolution satellite data based on a hybrid method," *Remote Sens.*, vol. 9, no. 5, 2017.
- [162] C. Romero-Trigueros *et al.*, "Effects of saline reclaimed waters and deficit irrigation on Citrus physiology assessed by UAV remote sensing," *Agric. Water Manag.*, vol. 183, pp. 60–69, 2017.
- [163] S. Vanino *et al.*, "Capability of Sentinel-2 data for estimating maximum evapotranspiration and irrigation requirements for tomato crop in Central Italy," *Remote Sens. Environ.*, vol. 215, no. October 2017, pp. 452–470, 2018.
- [164] A. Lowe, N. Harrison, and A. P. French, "Hyperspectral image analysis techniques for the detection and classification of the early onset of plant

disease and stress,” *Plant Methods*, vol. 13, no. 1, pp. 1–12, 2017.

- [165] S. Thomas *et al.*, “Benefits of hyperspectral imaging for plant disease detection and plant protection: a technical perspective,” *J. Plant Dis. Prot.*, vol. 125, no. 1, pp. 5–20, 2018.
- [166] Y. Sun, K. Wei, Q. Liu, L. Pan, and K. Tu, “Classification and discrimination of different fungal diseases of three infection levels on peaches using hyperspectral reflectance imaging analysis,” *Sensors (Switzerland)*, vol. 18, no. 4, 2018.
- [167] Y. Lu, Y. Huang, and R. Lu, “Innovative hyperspectral imaging-based techniques for quality evaluation of fruits and vegetables: A review,” *Appl. Sci.*, vol. 7, no. 2, 2017.
- [168] Y. Liu, H. Pu, and D. W. Sun, “Hyperspectral imaging technique for evaluating food quality and safety during various processes: A review of recent applications,” *Trends Food Sci. Technol.*, vol. 69, pp. 25–35, 2017.
- [169] T. Yaseen, D. W. Sun, and J. H. Cheng, “Raman imaging for food quality and safety evaluation: Fundamentals and applications,” *Trends Food Sci. Technol.*, vol. 62, pp. 177–189, 2017.
- [170] R. Siche, R. Vejarano, V. Aredo, L. Velasquez, E. Saldaña, and R. Quevedo, “Evaluation of Food Quality and Safety with Hyperspectral Imaging (HSI),” *Food Eng. Rev.*, vol. 8, no. 3, pp. 306–322, 2016.
- [171] A. Morellos *et al.*, “Machine learning based prediction of soil total nitrogen, organic carbon and moisture content by using VIS-NIR spectroscopy,” *Biosyst. Eng.*, vol. 152, pp. 104–116, 2016.
- [172] A. C. Dotto, R. S. D. Dalmolin, A. ten Caten, and S. Grunwald, “A systematic study on the application of scatter-corrective and spectral-derivative preprocessing for multivariate prediction of soil organic carbon by Vis-NIR spectra,” *Geoderma*, vol. 314, no. May 2017, pp. 262–274, 2018.
- [173] F. Cao, F. Liu, H. Guo, W. Kong, C. Zhang, and Y. He, “Fast detection of sclerotinia sclerotiorum on oilseed rape leaves using low-altitude remote sensing technology,” *Sensors (Switzerland)*, vol. 18, no. 12, 2018.
- [174] L. Huang *et al.*, “Potential of visible and near-infrared hyperspectral imaging for detection of *Diaphania pyloalis* larvae and damage on mulberry leaves,” *Sensors (Switzerland)*, vol. 18, no. 7, pp. 1–16, 2018.
- [175] A. Y. Khaled, S. Abd Aziz, S. K. Bejo, N. M. Nawi, and I. Abu Seman, “Spectral features selection and classification of oil palm leaves infected by Basal stem rot (BSR) disease using dielectric spectroscopy,” *Comput. Electron. Agric.*, vol. 144, no. August 2017, pp. 297–309, 2018.
- [176] F. López-Granados, J. Torres-Sánchez, A. Serrano-Pérez, A. I. de Castro, F.-J. Mesas-Carrascosa, and J.-M. Peña, “Early season weed mapping in sunflower using UAV technology: variability of herbicide treatment maps

against weed thresholds," *Precis. Agric.*, vol. 17, no. 2, pp. 183–199, Apr. 2016.

- [177] A. J. Steidle Neto, D. C. Lopes, F. A. C. Pinto, and S. Zolnier, "Vis/NIR spectroscopy and chemometrics for non-destructive estimation of water and chlorophyll status in sunflower leaves," *Biosyst. Eng.*, vol. 155, pp. 124–133, Mar. 2017.
- [178] G. R. Mahajan, R. N. Pandey, R. N. Sahoo, V. K. Gupta, S. C. Datta, and D. Kumar, "Monitoring nitrogen, phosphorus and sulphur in hybrid rice (*Oryza sativa* L.) using hyperspectral remote sensing," *Precis. Agric.*, vol. 18, no. 5, pp. 736–761, 2017.
- [179] Y. Zhang *et al.*, "Energy transfer and tunable emission of Ca₁₄Al₁₀Zn₆O₃₅:Bi³⁺, Sm³⁺ phosphor," *Mater. Res. Bull.*, vol. 100, no. November 2017, pp. 56–61, 2018.
- [180] S. R. M. Ibrahim, E. S. Elkhayat, G. A. A. Mohamed, S. M. Fat'hi, and S. A. Ross, "Fusarithioamide A, a new antimicrobial and cytotoxic benzamide derivative from the endophytic fungus *Fusarium chlamydosporium*," *Biochem. Biophys. Res. Commun.*, vol. 479, no. 2, pp. 211–216, 2016.
- [181] Y. Z. Zhu, J. W. Liu, X. Wang, I. H. Jeong, Y. J. Ahn, and C. J. Zhang, "Anti-BACE1 and antimicrobial activities of steroidal compounds isolated from marine urechis uncinatus," *Mar. Drugs*, vol. 16, no. 3, 2018.
- [182] S. Felhi, A. Daoud, H. Hajlaoui, K. Mnafigui, N. Gharsallah, and A. Kadri, "Solvent extraction effects on phytochemical constituents profiles, antioxidant and antimicrobial activities and functional group analysis of *Ecballium elaterium* seeds and peels fruits," *Food Sci. Technol.*, vol. 37, no. 3, pp. 483–492, 2017.
- [183] N. Ahmad, N. Ahmad, and A. Rab, "Light-induced biochemical variations in secondary metabolite production and antioxidant activity in callus cultures of *Stevia rebaudiana* (Bert)," *J. Photochem. Photobiol. B Biol.*, vol. 154, pp. 51–56, 2016.
- [184] W. W. May Zin *et al.*, "Antibacterial and antibiofilm activities of the metabolites isolated from the culture of the mangrove-derived endophytic fungus *Eurotium chevalieri* KUFA 0006," *Phytochemistry*, vol. 141, pp. 86–97, 2017.
- [185] J. Ouyang *et al.*, "Mollicellins O–R, four new depsidones isolated from the endophytic fungus *Chaetomium* sp. EEF-10," *Molecules*, vol. 23, no. 12, pp. 1–11, 2018.
- [186] M. L. Askoura, F. Vaudelle, and J. P. L'Huillier, "Use of steady-state imaging setup for assessing the internal optical properties of non-spherical apple samples," *Comput. Electron. Agric.*, vol. 157, no. August 2018, pp. 181–188, 2019.

- [187] A. López-Maestresalas *et al.*, “Bulk Optical Properties of Potato Flesh in the 500–1900 nm Range,” *Food Bioprocess Technol.*, vol. 9, no. 3, pp. 463–470, 2016.
- [188] W. Wang, M. Huang, Q. Zhu, and J. Qin, “Optical property inversion of biological materials using Fourier series expansion and LS-SVM for hyperspectral imaging,” *Inverse Probl. Sci. Eng.*, vol. 26, no. 7, pp. 1019–1036, 2018.
- [189] N. Bluvshstein *et al.*, “Broadband optical properties of biomass-burning aerosol and identification of brown carbon chromophores,” *J. Geophys. Res.*, vol. 122, no. 10, pp. 5441–5456, 2017.
- [190] T. S. Magney, L. A. Vierling, J. U. H. Eitel, D. R. Huggins, and S. R. Garrity, “Response of high frequency Photochemical Reflectance Index (PRI) measurements to environmental conditions in wheat,” *Remote Sens. Environ.*, vol. 173, pp. 84–97, 2016.
- [191] A. W. Njoroge *et al.*, “Frequency and time pattern differences in acoustic signals produced by *Prostephanus truncatus* (Horn) (Coleoptera: Bostrichidae) and *Sitophilus zeamais* (Motschulsky) (Coleoptera: Curculionidae) in stored maize,” *J. Stored Prod. Res.*, vol. 69, pp. 31–40, 2016.
- [192] M. Traoré, J. Kaal, and A. Martínez Cortizas, “Application of FTIR spectroscopy to the characterization of archeological wood,” *Spectrochim. Acta - Part A Mol. Biomol. Spectrosc.*, vol. 153, pp. 63–70, 2016.
- [193] G. Capobianco *et al.*, “Protective behaviour monitoring on wood photo-degradation by spectroscopic techniques coupled with chemometrics,” *Spectrochim. Acta - Part A Mol. Biomol. Spectrosc.*, vol. 172, pp. 34–42, 2017.
- [194] G. Bonifazi *et al.*, “A new approach for the modelling of chestnut wood photo-degradation monitored by different spectroscopic techniques,” *Environ. Sci. Pollut. Res.*, vol. 24, no. 16, pp. 13874–13884, 2017.
- [195] D. Cirule *et al.*, “Spectral Sensitivity of Thermally Modified and Unmodified Wood,” *BioResources*, vol. 11, no. 1, pp. 324–335, 2016.
- [196] D. Kachamba, H. Ørka, T. Gobakken, T. Eid, and W. Mwase, “Biomass Estimation Using 3D Data from Unmanned Aerial Vehicle Imagery in a Tropical Woodland,” *Remote Sens.*, vol. 8, no. 11, p. 968, Nov. 2016.
- [197] C. Camino, P. J. Zarco-Tejada, and V. Gonzalez-Dugo, “Effects of heterogeneity within tree crowns on airborne-quantified SIF and the CWSI as indicators of water stress in the context of precision agriculture,” *Remote Sens.*, vol. 10, no. 4, pp. 1–18, 2018.
- [198] D. Gómez-Candón, N. Virlet, S. Labbé, A. Jolivot, and J.-L. Regnard, “Field phenotyping of water stress at tree scale by UAV-sensed imagery: new

insights for thermal acquisition and calibration,” *Precis. Agric.*, vol. 17, no. 6, pp. 786–800, Dec. 2016.

- [199] L. Liu, N. C. Coops, N. W. Aven, and Y. Pang, “Mapping urban tree species using integrated airborne hyperspectral and LiDAR remote sensing data,” *Remote Sens. Environ.*, vol. 200, pp. 170–182, Oct. 2017.
- [200] N. Verma, D. Lamb, N. Reid, and B. Wilson, “Comparison of Canopy Volume Measurements of Scattered Eucalypt Farm Trees Derived from High Spatial Resolution Imagery and LiDAR,” *Remote Sens.*, vol. 8, no. 5, p. 388, May 2016.
- [201] G. Vaglio Laurin, N. Puletti, Q. Chen, P. Corona, D. Papale, and R. Valentini, “Above ground biomass and tree species richness estimation with airborne lidar in tropical Ghana forests,” *International Journal of Applied Earth Observation and Geoinformation*, vol. 52, Elsevier B.V., pp. 371–379, 2016.
- [202] M. Karlson, M. Ostwald, H. Reese, H. R. Bazié, and B. Tankoano, “Assessing the potential of multi-seasonal WorldView-2 imagery for mapping West African agroforestry tree species,” *Int. J. Appl. Earth Obs. Geoinf.*, vol. 50, pp. 80–88, Aug. 2016.
- [203] S. Madonsela *et al.*, “Multi-phenology WorldView-2 imagery improves remote sensing of savannah tree species,” *Int. J. Appl. Earth Obs. Geoinf.*, vol. 58, pp. 65–73, Jun. 2017.
- [204] G. Omer, O. Mutanga, E. Abdel-Rahman, and E. Adam, “Empirical Prediction of Leaf Area Index (LAI) of Endangered Tree Species in Intact and Fragmented Indigenous Forests Ecosystems Using WorldView-2 Data and Two Robust Machine Learning Algorithms,” *Remote Sens.*, vol. 8, no. 4, p. 324, Apr. 2016.
- [205] A. Axelsson, E. Lindberg, and H. Olsson, “Exploring Multispectral ALS Data for Tree Species Classification,” *Remote Sens.*, vol. 10, no. 3, p. 183, Jan. 2018.
- [206] J. Guerra-Hernández *et al.*, “Comparison of ALS- and UAV(SfM)-derived high-density point clouds for individual tree detection in Eucalyptus plantations,” *Int. J. Remote Sens.*, vol. 39, no. 15–16, pp. 5211–5235, Aug. 2018.
- [207] J. L. Xu, C. Riccioli, and D. W. Sun, “Development of an alternative technique for rapid and accurate determination of fish caloric density based on hyperspectral imaging,” *J. Food Eng.*, vol. 190, pp. 185–194, 2016.
- [208] J. H. Cheng *et al.*, “Developing a multispectral imaging for simultaneous prediction of freshness indicators during chemical spoilage of grass carp fish fillet,” *J. Food Eng.*, vol. 182, pp. 9–17, 2016.
- [209] G. ElMasry, N. Nakazawa, E. Okazaki, and S. Nakauchi, “Non-invasive sensing of freshness indices of frozen fish and fillets using pretreated

excitation-emission matrices," *Sensors Actuators, B Chem.*, vol. 228, pp. 237–250, 2016.

- [210] J. H. Cheng, D. W. Sun, G. Liu, and Y. N. Chen, "Developing a multispectral model for detection of docosahexaenoic acid (DHA) and eicosapentaenoic acid (EPA) changes in fish fillet using physarum network and genetic algorithm (PN-GA) method," *Food Chem.*, vol. 270, pp. 181–188, 2019.
- [211] S. Bureau, D. Cozzolino, and C. J. Clark, "Contributions of Fourier-transform mid infrared (FT-MIR) spectroscopy to the study of fruit and vegetables: A review," *Postharvest Biol. Technol.*, vol. 148, no. October 2018, pp. 1–14, 2019.
- [212] G. Cocetta *et al.*, "Light use efficiency for vegetables production in protected and indoor environments," *Eur. Phys. J. Plus*, vol. 132, no. 1, 2017.
- [213] H. Pu, D. Liu, L. Wang, and D. W. Sun, "Soluble Solids Content and pH Prediction and Maturity Discrimination of Lychee Fruits Using Visible and Near Infrared Hyperspectral Imaging," *Food Analytical Methods*, vol. 9, no. 1, pp. 235–244, 2015.
- [214] X. Feng, Y. Zhao, C. Zhang, P. Cheng, and Y. He, "Discrimination of transgenic maize kernel using NIR hyperspectral imaging and multivariate data analysis," *Sensors (Switzerland)*, vol. 17, no. 8, 2017.
- [215] N. Susič *et al.*, "Discrimination between abiotic and biotic drought stress in tomatoes using hyperspectral imaging," *Sensors Actuators, B Chem.*, vol. 273, pp. 842–852, 2018.
- [216] S. C. Kefauver *et al.*, "Comparative UAV and field phenotyping to assess yield and nitrogen use efficiency in hybrid and conventional barley," *Front. Plant Sci.*, vol. 8, no. October, pp. 1–15, 2017.
- [217] A.-K. Mahlein, "Plant Disease Detection by Imaging Sensors – Parallels and Specific Demands for Precision Agriculture and Plant Phenotyping," *Plant Dis.*, vol. 100, no. 2, pp. 241–251, Feb. 2016.
- [218] C. R. Yendrek *et al.*, "High-throughput phenotyping of maize leaf physiological and biochemical traits using hyperspectral reflectance," *Plant Physiol.*, vol. 173, no. 1, pp. 614–626, 2017.
- [219] Z. Khan, V. Rahimi-Eichi, S. Haefele, T. Garnett, and S. J. Miklavcic, "Estimation of vegetation indices for high-throughput phenotyping of wheat using aerial imaging," *Plant Methods*, vol. 14, no. 1, pp. 1–11, 2018.
- [220] M. Tattaris, M. P. Reynolds, and S. C. Chapman, "A Direct Comparison of Remote Sensing Approaches for High-Throughput Phenotyping in Plant Breeding," *Front. Plant Sci.*, vol. 7, no. AUG2016, pp. 1–9, Aug. 2016.
- [221] A. Haghghattalab *et al.*, "Application of unmanned aerial systems for high throughput phenotyping of large wheat breeding nurseries," *Plant Methods*, vol. 12, no. 1, pp. 1–15, Jun. 2016.

- [222] S. Bolanos, "Integrating GIS and remote sensing for coffee mapping Sandra," in *XII SELPER Symposium*, 2006, pp. 1–11.
- [223] J. M. Castrillón Cuervo, J. R. Sanz Uribe, and P. J. Ramos Giraldo, "Técnicas de Visión Artificial para la Identificación de Dos Defectos del Café Lavado," *Cenicafé*, vol. 69, no. 1, pp. 83–90, 2018.
- [224] C. Castrillón, U. Sanz, G. Ramos, and P. J. Algoritmo, "Algoritmo para la identificación de café lavado afectado por la broca del café," *Cenicafé*, vol. 68, no. 2, pp. 7–19, 2017.
- [225] J. P. Rojas *et al.*, "Aerial mapping of rice crops using mosaicing techniques for vegetative index monitoring," *2018 International Conference on Unmanned Aircraft Systems, ICUAS 2018*, IEEE, pp. 846–855, 2018.
- [226] C. A. Devia *et al.*, "High-Throughput Biomass Estimation in Rice Crops Using UAV Multispectral Imagery," *Journal of Intelligent and Robotic Systems: Theory and Applications*, *Journal of Intelligent & Robotic Systems*, pp. 573–589, 2019.
- [227] Forest Resources Assessment, "Actualización de la evaluación de los recursos forestales mundiales a 2005." [Online]. Available: <http://www.fao.org/3/ae156s/ae156s03.htm#TopOfPage>. [Accessed: 07-Nov-2019].
- [228] V. A. B. Meneses, J. M. Téllez, and D. F. A. Velasquez, "Uso De Drones Para El Analisis De Imágenes Multiespectrales En Agricultura De Precisión," *Ciencia y Tecnología Alimentaria*, vol. 13, no. 1, pp. 28–40, 2015.
- [229] A. K. Torres Galindo, A. F. Gómez Rivera, and A. F. Jiménez López, "Desarrollo de un sistema multiespectral para aplicaciones en agricultura de precisión usando dispositivos embebidos," *Sistemas y Telemática*, vol. 13, no. 33, pp. 27–44, 2015.
- [230] P. Pereira Coltri, J. Zullo, R. Ribeiro Do Valle Goncalves, L. A. S. Romani, and H. S. Pinto, "Coffee crop's biomass and carbon stock estimation with usage of high resolution satellites images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, no. 3, pp. 1786–1795, 2013.
- [231] A. Chemura, O. Mutanga, J. Odindi, and D. Kutuywayo, "Mapping spatial variability of foliar nitrogen in coffee (*Coffea arabica* L.) plantations with multispectral Sentinel-2 MSI data," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 138, International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS), pp. 1–11, 2018.
- [232] D. B. Marin, M. de Carvalho Alves, E. A. Pozza, L. L. Belan, and M. L. de Oliveira Freitas, "Multispectral radiometric monitoring of bacterial blight of coffee," *Precision Agriculture*, no. 0123456789, Springer US, 2018.
- [233] L. F. Johnson, S. R. Herwitz, B. M. Lobitz, and S. E. Dunagan, "Feasibility of

monitoring coffee field ripeness with airborne multispectral imagery,” vol. 20, no. 6, pp. 845–850, 2004.

- [234] C. Gomez *et al.*, “Use of high-resolution satellite imagery in an integrated model to predict the distribution of shade coffee tree hybrid zones,” *Remote Sensing of Environment*, vol. 114, no. 11, Elsevier Inc., pp. 2731–2744, 2010.
- [235] F. S. Kawakubo and R. P. Pérez Machado, “Mapping coffee crops in southeastern Brazil using spectral mixture analysis and data mining classification,” *International Journal of Remote Sensing*, vol. 37, no. 14, pp. 3414–3436, 2016.
- [236] A. Chemura and O. Mutanga, “Developing detailed age-specific thematic maps for coffee (*Coffea arabica* L.) in heterogeneous agricultural landscapes using random forests applied on Landsat 8 multispectral sensor,” *Geocarto International*, vol. 32, no. 7, pp. 759–776, 2017.
- [237] G. D. Martins, M. D. L. B. T. Galo, and B. S. Vieira, “Detecting and Mapping Root-Knot Nematode Infection in Coffee Crop Using Remote Sensing Measurements,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 12, pp. 5395–5403, 2017.
- [238] MAPIR, “Cámara Survey3W - Rojo + Verde + NIR (RGN, NDVI) - CÁMARA MAPIR.” [Online]. Available: <https://www.mapir.camera/products/survey3w-camera-red-green-nir-rgn-ndvi>. [Accessed: 02-Nov-2019].
- [239] QGIS, “QGIS.” [Online]. Available: <https://www.qgis.org/es/site/>. [Accessed: 14-Jun-2021].
- [240] Agisoft, “Agisoft Metashape.” [Online]. Available: <https://www.agisoft.com/>. [Accessed: 14-Jun-2021].
- [241] Colaboratory, “Colaboratory.” [Online]. Available: <https://colab.research.google.com/notebooks/welcome.ipynb?hl=es-AR>. [Accessed: 14-Jun-2021].
- [242] ARDUPILOT, “Mission Planner.” [Online]. Available: <https://ardupilot.org/planner/>. [Accessed: 14-Jun-2021].
- [243] R. Novoa S.-A. and N. Villagrán A., “Evaluación de un instrumento medidor de clorofila en la determinación de niveles de nitrógeno foliar en maíz,” *Agríc. Técnica*, vol. 62, no. 1, Jan. 2002.
- [244] N. Knighton and B. Bugbee, “A Comparison of Opti-Sciences CCM-200 Chlorophyll Meter and the Minolta SPAD 502 Chlorophyll Meter.”
- [245] C. Santiago Castañeda, P. José Almanza-merchán, E. Hernando Pinzón, G. Eduardo Cely-reyes, and P. Antonio Serrano-cely, “Estimación de la concentración de clorofila mediante métodos no destructivos en vid (*Vitis vinifera* L.) cv. Riesling Becker,” *Ciencias Hortícolas*, vol. 12, no. 2, pp. 329–337.

- [246] A. R. Reis, J. L. Favarin, E. Malavolta, J. L. Júnior, and M. F. Moraes, "Photosynthesis, chlorophylls, and SPAD readings in coffee leaves in relation to nitrogen supply," *Commun. Soil Sci. Plant Anal.*, vol. 40, no. 9–10, pp. 1512–1528, May 2009.
- [247] A. Azevedo and M. F. Santos, "KDD, SEMMA and CRISP-DM: a parallel overview," *IADS - DM*, 2008.
- [248] R. Callejas, E. Kania, A. Contreras, C. Peppi, and L. Morales, "Evaluación de un método no destructivo para estimar las concentraciones de clorofila en hojas de variedades de uva de mesa," *Idesia*, vol. 31, no. 4, pp. 19–26, 2013.
- [249] C. S. Castañeda, P. J. Almanza-Merchán, E. H. Pinzón, G. E. Cely, and P. A. Serrano, "Estimación de la concentración de clorofila mediante métodos no destructivos en vid (*Vitis vinifera* L.) cv. Riesling Becker," *Rev. Colomb. Ciencias Hortícolas*, vol. 12, no. 2, pp. 329–337, 2018.
- [250] J. Ortega, "Caracterización en el espectro visible e infrarrojo de cultivos agrícolas desde un vehículo aéreo multirrotor," 2017.
- [251] MAPIR, "Trigger Survey2 Camera Using PWM Signal From 3DR SOLO - MAPIR CAMERA." [Online]. Available: <https://www.mapir.camera/pages/trigger-survey2-cameras-from-3dr-solo>. [Accessed: 14-Feb-2021].
- [252] B. Ruzgiene, "Requirements for aerial photography," *L. Use Law Zo. Dig.*, vol. 30, no. 3, pp. 75–79, 2004.
- [253] M. Elarab, A. M. Ticlavilca, A. F. Torres-Rua, I. Maslova, and M. McKee, "Estimating chlorophyll with thermal and broadband multispectral high resolution imagery from an unmanned aerial system using relevance vector machines for precision agriculture," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 43, pp. 32–42, Dec. 2015.
- [254] Unidad Administrativa Especial de Aeronáutica Civil, "Reglamentos Aeronáuticos de Colombia," 2020. [Online]. Available: https://www.aerocivil.gov.co/normatividad/RAC/RAC_91_-_Reglas_Generales_de_Vuelo_y_Operacion.pdf. [Accessed: 14-Jun-2021].
- [255] S. Heriberto and V. Bilbao, "FOTOGRAFIA AEREA," 2005.
- [256] B. T. Fraser and R. G. Congalton, "Issues in Unmanned Aerial Systems (UAS) Data Collection of Complex Forest Environments," *Remote Sens. 2018, Vol. 10, Page 908*, vol. 10, no. 6, p. 908, Jun. 2018.
- [257] P. Nuñez Calleja, "Comparativa de software para la realización de ortofotos a partir de imágenes obtenidas por drones," Universidad de Oviedo, 2016.
- [258] E. G. Manrique, "Índice de vegetación, aplicación del NDVI," *Av. y aplicaciones*, pp. 217–219, 1999.

- [259] M. Vincini, E. Frazzi, and P. D'Alessio, "A broad-band leaf chlorophyll vegetation index at the canopy scale," in *Precision Agriculture*, 2008, vol. 9, no. 5, pp. 303–319.
- [260] J. S. Schepers, T. M. Blackmer, W. W. Wilhelm, and M. Resende, "Transmittance and reflectance measurements of corn leaves from plants with different nitrogen and water supply," *J. Plant Physiol.*, vol. 148, no. 5, pp. 523–529, Jan. 1996.
- [261] J. Avelino, E. Treminio, F. Casanoves, S. Vilchez, J. Cárdenas, and C. Lizardo, "Guía para la vigilancia de la roya del café (*Hemileia vastatrix*)," 2019.
- [262] J. Lozano, "Guía para la toma de muestra foliar," pp. 8–9, 2006.
- [263] E. Héctor Iván Trujillo, A. Luis Fernando Aristizábal, P. Alex Enrique Bustillo, and Q. Mauricio Jiménez, "Evaluación de métodos para cuantificar poblaciones de broca del café, *Hypothenemus hampei* (Ferrari) (Coleoptera: Curculionidae: Scolytinae), en fincas de caficultores experimentadores," *Rev. Colomb. Entomol.*, vol. 32, no. 1, pp. 39–44, 2006.
- [264] L. R. M. Díaz, "Actividades de monitoreo y manejo de la roya del café," 2014.
- [265] "Hoja de café." [Online]. Available: https://www.freepik.es/fotos-premium/hoja-cafe-sobre-fondo-blanco_2630216.htm. [Accessed: 21-May-2021].
- [266] H. Lalinde *et al.*, "Sobre el uso adecuado del coeficiente de correlación de Pearson: definición, propiedades y suposiciones," 2018.
- [267] A. Carrasquilla-Batista *et al.*, "Regresión lineal simple y múltiple: aplicación en la predicción de variables naturales relacionadas con el crecimiento microalgal," *Rev. Tecnol. en Marcha*, vol. 29, no. 8, pp. 33–45, Dec. 2016.
- [268] IStock, "Ilustración De Dibujo De Flores Y Hojas De Café Con Arte Lineal Sobre Fondos Blancos y más Vectores Libres de Derechos de Cafeto." [Online]. Available: <https://www.istockphoto.com/es/vector/ilustración-de-dibujo-de-flores-y-hojas-de-café-con-arte-lineal-sobre-fondos-blancos-gm1203209083-345730167>. [Accessed: 09-Aug-2021].
- [269] 123RF, "Hoja de café." [Online]. Available: https://es.123rf.com/imagenes-de-archivo/hoja_de_cafe.html?sti=mqeu16b51k8q1eb0jm%7C. [Accessed: 29-Aug-2021].
- [270] J. Tremblay *et al.*, "Training Deep Networks with Synthetic Data: Bridging the Reality Gap by Domain Randomization," 2018.
- [271] A. Kortylewski, B. Egger, A. Schneider, T. Gerig, A. Morel-Forster, and T. Vetter, "Analyzing and Reducing the Damage of Dataset Bias to Face Recognition with Synthetic Data," 2019.
- [272] Y. Toda *et al.*, "Training instance segmentation neural network with synthetic

datasets for crop seed phenotyping,” *Commun. Biol.*, vol. 3, no. 1, pp. 1–12, Dec. 2020.

- [273] T. A. Le, A. G. Baydin, R. Zinkov, and F. Wood, “Using synthetic data to train neural networks is model-based reasoning,” in *Proceedings of the International Joint Conference on Neural Networks*, 2017, vol. 2017-May, pp. 3514–3521.
- [274] F. S. Saleh, “Effective Use of Synthetic Data for Urban Scene Semantic Segmentation ★,” 2018.
- [275] “Gretel.” [Online]. Available: <https://gretel.ai/>. [Accessed: 19-May-2021].
- [276] “gretel-blueprints/config_templates/gretel/synthetics at main · gretelai/gretel-blueprints.” [Online]. Available: https://github.com/gretelai/gretel-blueprints/tree/main/config_templates/gretel/synthetics. [Accessed: 19-May-2021].
- [277] Gretel, “Interpreting the Gretel Synthetics Report - Gretel.ai.” [Online]. Available: <https://docs.gretel.ai/deep-dives/gretel-synthetics/interpreting-the-gretel-synthetics-report#field-correlation-stability>. [Accessed: 19-May-2021].
- [278] K. TAKAYAMA and H. NISHINA, “Chlorophyll Fluorescence Imaging of the Chlorophyll Fluorescence Induction Phenomenon for Plant Health Monitoring,” *Environ. Control Biol.*, vol. 47, no. 2, pp. 101–109, 2009.
- [279] A. K. Dey, M. Sharma, and M. R. Meshram, “An Analysis of Leaf Chlorophyll Measurement Method Using Chlorophyll Meter and Image Processing Technique,” in *Procedia Computer Science*, 2016, vol. 85, pp. 286–292.
- [280] D. Pavlovic, B. Nikolic, S. Djurovic, H. Waisi, A. Andjelkovic, and D. Marisavljevic, “Chlorophyll as a measure of plant health: Agroecological aspects,” *Pestic. i fitomedicina*, vol. 29, no. 1, pp. 21–34, 2014.
- [281] Cenicafe, “Evaluación temprana de la deficiencia del nitrógeno en café y aplicaciones.”
- [282] T. Dong *et al.*, “Assessment of Portable Chlorophyll Meters for Measuring Crop Leaf Chlorophyll Concentration,” *Remote Sens.*, vol. 11, p. 2706, 2019.
- [283] R. Callejas, E. Kania, A. Contreras, C. Peppi, and L. Morales, “Evaluación de un método no destructivo para estimar las concentraciones de clorofila en hojas de variedades de uva de mesa,” *Idesia*, vol. 31, no. 4, pp. 19–26, 2013.
- [284] A. T. Netto, E. Campostrini, J. G. De Oliveira, and R. E. Bressan-Smith, “Photosynthetic pigments, nitrogen, chlorophyll a fluorescence and SPAD-502 readings in coffee leaves,” *Sci. Hortic. (Amsterdam)*, vol. 104, no. 2, pp. 199–209, Mar. 2005.
- [285] K. Minolta, “A lightweight handheld meter for measuring the chlorophyll content of leaves without causing damage to plants. SPAD-502Plus.”

- [286] T. Poblete, S. Ortega-Farías, M. Moreno, and M. Bardeen, "Artificial Neural Network to Predict Vine Water Status Spatial Variability Using Multispectral Information Obtained from an Unmanned Aerial Vehicle (UAV)," *Sensors*, vol. 17, no. 11, p. 2488, Oct. 2017.
- [287] X. E. Pantazi, A. A. Tamouridou, T. K. Alexandridis, A. L. Lagopodi, J. Kashefi, and D. Moshou, "Evaluation of hierarchical self-organising maps for weed mapping using UAS multispectral imagery," *Comput. Electron. Agric.*, vol. 139, pp. 224–230, Jun. 2017.
- [288] J. Behmann, A.-K. Mahlein, T. Rumpf, C. Römer, and L. Plümer, "A review of advanced machine learning methods for the detection of biotic stress in precision crop protection," *Precis. Agric.*, vol. 16, no. 3, pp. 239–260, Jun. 2015.
- [289] G. A. Lobos, I. Matus, A. Rodriguez, S. Romero-Bravo, J. L. Araus, and A. del Pozo, "Wheat genotypic variability in grain yield and carbon isotope discrimination under Mediterranean conditions assessed by spectral reflectance," *J. Integr. Plant Biol.*, vol. 56, no. 5, pp. 470–479, May 2014.
- [290] A. J. Steidle Neto, D. C. Lopes, F. A. C. Pinto, and S. Zolnier, "Vis/NIR spectroscopy and chemometrics for non-destructive estimation of water and chlorophyll status in sunflower leaves," *Biosyst. Eng.*, vol. 155, pp. 124–133, 2017.
- [291] P. Rajkumar, N. Wang, G. Elmasry, G. S. V. Raghavan, and Y. Garipey, "Studies on banana fruit quality and maturity stages using hyperspectral imaging," *J. Food Eng.*, vol. 108, no. 1, pp. 194–200, Jan. 2012.
- [292] S. Sharma and V. Sharma, "Performance of Various Machine Learning Classifiers on Small Datasets with Varying Dimensionalities," *Circ. Comput. Sci.*, vol. 1, pp. 30–35, 2016.
- [293] J. P. Maheswari, "Breaking the curse of small datasets in Machine Learning," 2018. [Online]. Available: <https://towardsdatascience.com/breaking-the-curse-of-small-datasets-in-machine-learning-part-1-36f28b0c044d>. [Accessed: 08-Sep-2021].
- [294] C. Schaffer and S. A. H. C. Edu, "Selecting a classification method by cross-validation," *Mach. Learn. 1993 131*, vol. 13, no. 1, pp. 135–143, Oct. 1993.
- [295] R. Python, "Divida su conjunto de datos con train_test_split () de scikit-learn." [Online]. Available: <https://realpython.com/train-test-split-python-data/>. [Accessed: 20-Sep-2021].
- [296] M. H. Diskin, "Definition and Uses of the Linear Regression Model," *Water Resour. Res.*, vol. 6, no. 6, pp. 1668–1673, Dec. 1970.
- [297] L. Tian, T. Cai, E. Goetghebeur, and L. J. Wei, "Model Evaluation Based on the Distribution of Estimated Absolute Prediction Error Model Evaluation Based on the Distribution of Estimated Absolute Prediction Error."

- [298] D. Chicco, M. J. Warrens, and G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation," *PeerJ Comput. Sci.*, vol. 7, pp. 1–24, Jul. 2021.
- [299] C. Kingsford and S. L. Salzberg, "What are decision trees?," *Nat. Biotechnol.* 2008 269, vol. 26, no. 9, pp. 1011–1013, Sep. 2008.
- [300] L. Rokach and O. Maimon, "Decision Trees," *Data Min. Knowl. Discov. Handb.*, pp. 165–192, May 2005.
- [301] Sklearn, "Decision Tree Regressor." [Online]. Available: <https://scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeRegressor.html>. [Accessed: 09-Sep-2021].
- [302] J. J. Negron-Granados, R. Legarda-Sáenz, and V. Uc-Cetina, "Approach for the Classification of Polinífera Vegetation Using Multispectral Images and Neural Networks," *Res. Comput. Sci.*, vol. 147, no. 5, 2018.
- [303] M. E. Arboit and D. S. Maglione, "Análisis multitemporal y multiespacial del índice de vegetación de diferencia normalizada (NDVI) y del índice de vegetación ajustado al suelo (SAVI) en centros urbanos forestados y oasis irrigados, con climas secos," vol. 109, pp. 13–60, 2019.
- [304] M. Awad and R. Khanna, "Support Vector Regression," *Effic. Learn. Mach.*, pp. 67–80, 2015.
- [305] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Stat. Comput.* 2004 143, vol. 14, no. 3, pp. 199–222, Aug. 2004.
- [306] Aprende IA, "Vectores de Soporte Regresión." [Online]. Available: <https://aprendeia.com/maquina-de-vectores-de-soporte-regresion-teoria/>. [Accessed: 09-Sep-2021].
- [307] M. A. Gilabert, J. Gonzalez-Piqueras, and J. García-Haro, "Acerca de los índices de vegetación," *Teledeteccion*, 1997.
- [308] H. Wang, Q. Yilihamu, M. Yuan, H. Bai, H. Xu, and J. Wu, "Prediction models of soil heavy metal(loid)s concentration for agricultural land in Dongli: A comparison of regression and random forest," *Ecol. Indic.*, vol. 119, p. 106801, Dec. 2020.
- [309] S. Götz, N. Zimmermann, D. Engelhardt, and H. Bernhardt, "Simulation of Agricultural Logistic Processes with k-Nearest Neighbors Algorithm," *Agric. Eng. Int. CIGR J.*, Apr. 2015.
- [310] A. Mucherino, P. Papajorgji, and P. M. Pardalos, "A survey of data mining techniques applied to agriculture," *Oper. Res.* 2009 92, vol. 9, no. 2, pp. 121–140, Jun. 2009.
- [311] S. Bafandeh, I. And, and M. Bolandraftar, "Application of K-Nearest Neighbor (KNN) Approach for Predicting Economic Events: Theoretical Background,"

J. Eng. Res. Appl. www.ijera.com, vol. 3, pp. 605–610, 2013.

- [312] “Proyecto: IoT-Agro.” [Online]. Available: <https://www.iot-agro.com/proyecto/>. [Accessed: 25-May-2020].
- [313] N. Arteaga and C. Delgado, “Multispectral Images Coffee.” [Online]. Available: <https://github.com/NataliaArteaga/MultispectralImagesCoffee>. [Accessed: 22-May-2021].
- [314] H. C. S. Nascimento and R. A. Marengo, “SPAD-502 readings in response to photon fluence in leaves with different chlorophyll content,” *Rev. Ceres*, vol. 57, no. 5, pp. 614–620, Oct. 2010.

ANEXO A: Configuración del vuelo del VANT

Este anexo corresponde a un video tutorial de configuración del vuelo del VANT SOLO 3D en el programa Mission Planner, donde se muestra cómo crear el polígono correspondiente al área de estudio, la configuración de las especificaciones requeridas para el vuelo y toda la información necesaria para realizar las tomas. El video se encuentra en la siguiente URL: https://drive.google.com/drive/folders/1c7zjNkOnTEH_Cz8xJQlpF2HJPYZGeT85?usp=sharing

ANEXO B: Creación de ortomosaicos

Este anexo corresponde a un video tutorial en donde se encuentra el proceso requerido para una correcta creación de ortomosaico a partir de imágenes individuales georreferenciadas que son primero procesadas y calibradas para posteriormente ser corregirlas espacialmente con el fin de unificarlas en una sola imagen. Este anexo se encuentra en

<https://drive.google.com/drive/folders/1onwRHlccvdUuUz-4o6R3ibIY2rI2Ncno?usp=sharing>

ANEXO C: Obtención de los índices de vegetación con QGIS

Este anexo es un video tutorial que contiene la guía para realizar el cálculo de funciones matemáticas utilizando las bandas disponibles en las imágenes de entradas, en este caso las funciones tienen como resultado índices de vegetación por zonas, estas zonas son delimitadas por polígonos y los promedios son calculados gracias a las librerías propias de QGIS. La dirección es [https://drive.google.com/drive/folders/1xCxZECImYjdGLVh0kdhqFCshbUB2d7K ?usp=sharing](https://drive.google.com/drive/folders/1xCxZECImYjdGLVh0kdhqFCshbUB2d7K?usp=sharing)

ANEXO D: Archivos con los índices de vegetación obtenidos

Después de la generación de las ortofotos y su procesamiento en el programa QGIS, se obtienen como resultado los archivos de configuración para cargar en el programa y un archivo (por cada cálculo de índice de vegetación) con los identificadores del valor que se dan según la ubicación de la planta, la media, el valor mínimo, el valor máximo, el promedio y todos los cálculos que se seleccionaron en la herramienta QGIS. URL de acceso <https://drive.google.com/drive/folders/1ffrJkDtRux6bC4BLDsk7cx28i7P7RsFo?usp=sharing>

ANEXO E: Desviación estándar de los datos de clorofila por planta

El documento desviaciónEstandar.xlsx posee los datos de clorofila tomados de una sola planta en diferentes zonas (baja, media y alta), para validar la desviación estándar según el número de datos a tomar en cuenta por planta. Estos datos fueron utilizados para generar la tabla 7 de la monografía. URL https://drive.google.com/drive/folders/1_7YzklmSLKO-6wm-PFR-k4A_93XyNunD?usp=sharing

ANEXO F: Desviación estándar de los datos de clorofila por hoja

El documento desviaciónEstandarPorHoja.xlsx posee los datos de clorofila tomados de seis puntos de una misma hoja. Se tomaron en total catorce hojas, dando por ende un total de 84 muestras para analizar la desviación estándar por hoja. Estos datos fueron utilizados para generar la tabla 8 de la monografía. Este documento se encuentra en el siguiente enlace:

https://drive.google.com/drive/folders/1KKSFoJi3lOdYt3xV0w2dPn0rSv-7lO_X?usp=sharing

ANEXO G: Datos de entrada para los modelos

Se realizaron dos procedimientos, el primero con el proceso de filtrado que se menciona en el capítulo 3, el cual tiene un archivo con todos los índices de vegetación y valor de clorofila (CLL con unidades CCI, NDVI, CVI, NRVI, RVI, GCI, GNDVI adimensionales) con los datos iniciales y otro integrando los datos sintéticos. Además, si se requiere, en una carpeta hay archivos por cada índice de vegetación en contraste con el CLL con sólo los datos iniciales y otro integrando los datos sintéticos.

Por otro lado, el segundo proceso, el cual implica la limpieza de datos extremos, en el archivo datosCorrelacionIndices.xlsx (ubicado en) contiene los datos mencionados en el primer procedimiento, con el análisis de correlación inicial y la correlación al remover datos extremos. URL de acceso <https://drive.google.com/drive/folders/1FiTTiplmHeZIsziOmy6LsDYJX9uOnP?usp=sharing>

ANEXO H: Reportes de generación de datos sintéticos con Gretel

Contiene los reportes de las pruebas realizadas en la API de Gretel con diferentes configuraciones soportadas en la librería (default, low and field count), donde se visualiza la correlación y calidad de los datos generados en comparación con los datos originales (Ilustración 20 e Ilustración 21), también se agrega un archivo con extensión .csv con los datos generados para alimentar los diferentes modelos expuestos en el capítulo 3. Las columnas que contiene son: CLL, NDVI, GCI, GNDVI, RVI, NRVI, CVI y NRVI3. Estos reportes se encuentran en <https://drive.google.com/drive/folders/16vwfn2tc2P8yVhN1Wr4FXVi6x45tLD2d?usp=sharing>

ANEXO I: Ejemplo de procedimiento realizado con redes neuronales

Se describe el segundo procedimiento realizado con los datos obtenidos del anexo H, datos con los cuales se modeló una red neuronal artificial y regresión lineal simple y múltiple tanto con el conjunto de datos iniciales como con aquellos integrando los datos sintéticos generados por la API de Gretel. En este anexo se encuentran los modelos entrenados con los datos sin ningún tipo de filtro inicial y sus resultados. También se puede encontrar en la siguiente URL de acceso <https://drive.google.com/drive/folders/1isZzZlBh4x9bsaF1aUpljYmPN3iiLS7m?usp=sharing>

Limpieza de datos

A continuación, se condensan dichos resultados a través de una gráfica de dispersión con la ecuación de la recta dada por 'y'. Al observar datos extremos (en amarillo). Como se representa en la Ilustración 1 y la Ilustración 1, se procedió a limpiar dichos valores, esto se realizó de forma manual ya que eran pocos valores. La limpieza se realizó con el fin de mejorar la correlación de Pearson y el R^2 los cuales miden la relación entre los datos, entre más se acerquen a 1, mejor es la relación de las variables.

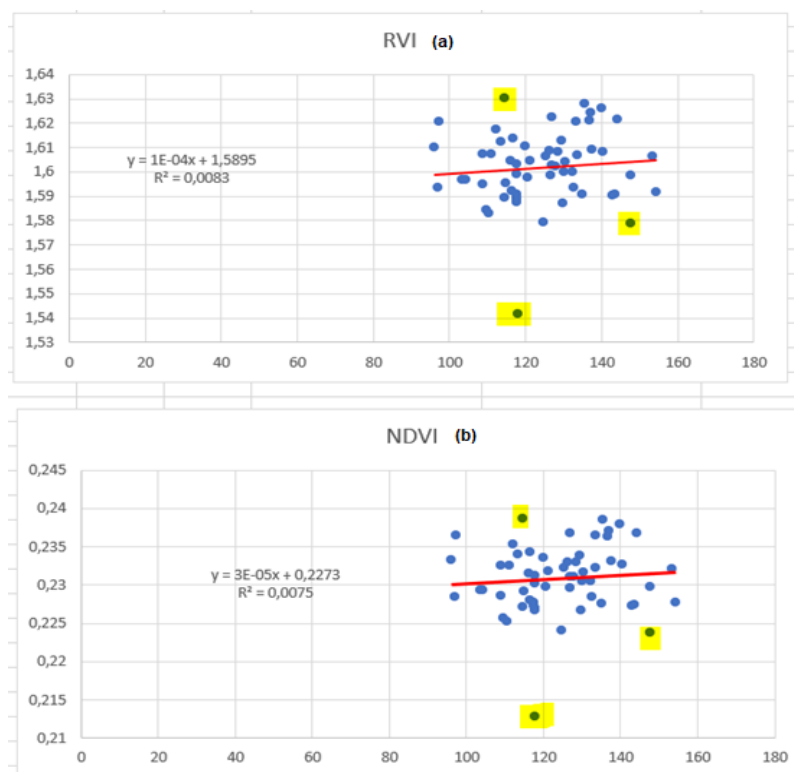


Ilustración 1. Gráfica de dispersión con valores extremos para RVI(a) y NDVI(b).
Fuente propia

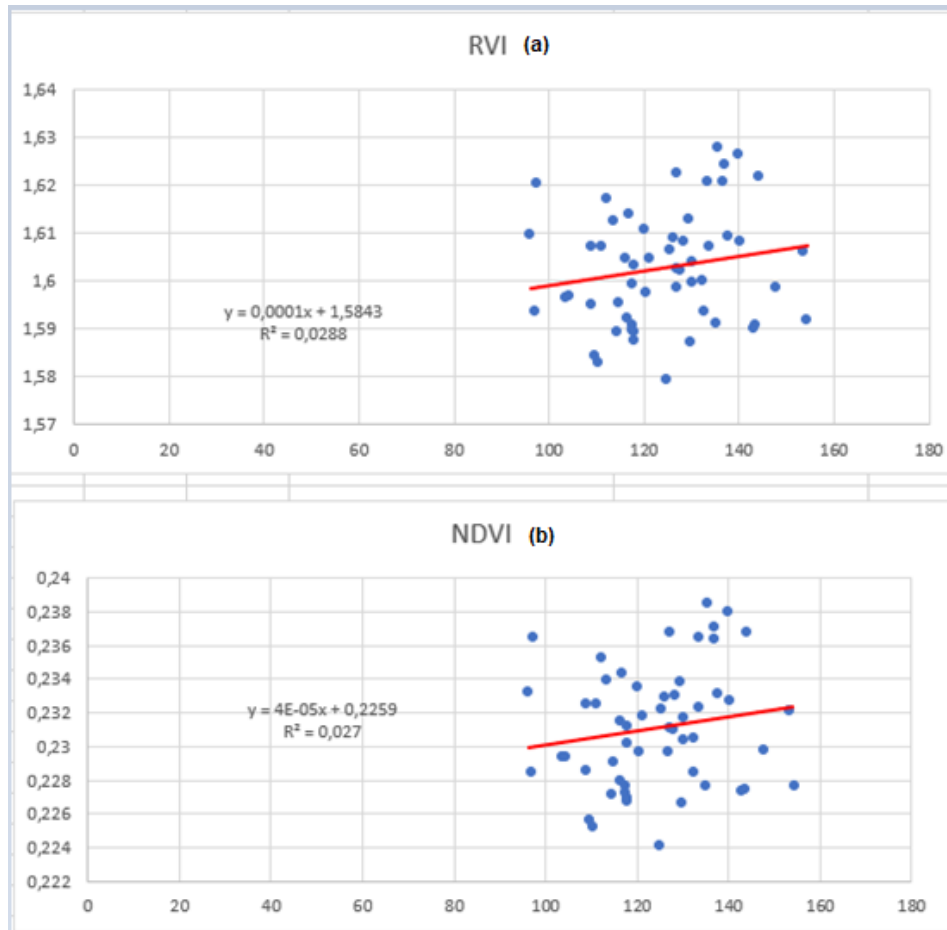


Ilustración 2. Gráfica de dispersión después de eliminar valores extremos para RVI(a) y NDVI(b). Fuente propia.

Tabla 1. Correlación de Pearson de índices de vegetación. Fuente propia

	ANTES DE LIMPIEZA	DESPUÉS DE LIMPIEZA
NDVI	0,08671	0,164307
CVI	-0,01813	-0,00014
GNDVI	-0,00187	0,021876
NRVI	0,08671	0,164307
RVI	0,091016	0,16979
GCI	-0,00172	0,022046

En la Tabla 1 se observa que después de la limpieza la correlación de Pearson mejoró ya que los valores obtenidos se acercan un poco más al 1 en comparación a los datos sin limpieza de extremos.

Para aumentar el tamaño del conjunto de datos obtenidos inicialmente y superar las limitantes de tiempo y espacio para tomar más muestras en campo, se optó por una alternativa usada en el ámbito del aprendizaje automático, la generación de datos sintéticos [1]–[5].

Regresión lineal

Tabla 2. Error regresión lineal con y sin datos sintéticos. Fuente propia

	INICIAL	CON DATOS SINTETICOS
MSE	225.41	199.66
MAE	12.01	11.43

En la Tabla 2 se registran el MSE y el MAE de la regresión lineal múltiple tanto con los datos iniciales en contraste con los datos integrando los datos sintéticos. Para la regresión lineal múltiple el valor a predecir fue el índice de contenido de clorofila (CCI) y las variables independientes los índices de vegetación (NDVI, CVI, GNDVI, NRVI3, NRVI, RVI y GCI) por planta. Además, se implementó una regresión lineal simple para cada uno de los índices de vegetación. Como se observa en la Tabla 3, el mejor comportamiento de la regresión lineal se observó en el NRVI al tener los errores de error más bajo.

Tabla 3. Error de datos de entrenamiento y datos sintéticos por cada índice de vegetación con regresión lineal simple. Fuente propia.

		INICIAL	CON DATOS SINTETICOS
NDVI	MAE	9.35	10.77
	MSE	140.90	177.97
CVI	MAE	9.24	10.97
	MSE	133.63	182.26
GNDVI	MAE	9.26	10.8
	MSE	134.74	178.37
NRVI	MAE	8.97	10.74
	MSE	135.67	177.03
RVI	MAE	9.04	10.83
	MSE	133.46	179.74
GCI	MAE	9.16	10.91
	MSE	132.62	179.82

Se observa que la regresión lineal múltiple mostrada en la Tabla 2 mejora al evaluarla con los datos sintéticos, no obstante, al evaluar cada índice por separado condensada en la Tabla 3, el error aumenta un poco en comparación con los datos iniciales.

Red neuronal artificial

Las redes neuronales son sistemas computacionales, inspirados en las neuronas presentes en el cerebro, las cuales se conectan entre sí buscando establecer inteligencia artificial para entender datos del mundo real [6]. Para la configuración de la red neuronal se configuraron 3 métodos de activación y de corrección en base a la librería Keras [7].

- Funciones de activación

Sigmoid: La función de activación (FA) Sigmoid no lineal es una función diferenciable acotada y definida para valores de entrada reales, con derivadas positivas y con cierto grado de suavidad [8].

Tanh: La función tangente hiperbólica conocida es una función centrada en cero, con rango entre -1 y 1 [9]. La función tanh es más usada en comparación con la función sigmoidea, ya que proporciona un mejor rendimiento de entrenamiento para redes neuronales multicapa [8], [10].

Relu: La FA Relu es la más utilizada [11], [12] ya que permite el aprendizaje rápido en las redes neuronales y es fácil de optimizar con métodos de descenso de gradiente por su comportamiento casi lineal [13].

- Métodos de corrección

Una devolución de llamada (callback) se puede realizar acciones en varias etapas del entrenamiento con el fin de mejorar el entrenamiento. A continuación se explican brevemente las utilizadas [14].

EarlyStopping: Esta función lo que realiza es detener el entrenamiento cuando la métrica monitoreada ha dejado de mejorar.

ModelCheckpoint: Permite ejecutar un modelo en un intervalo en donde se va guardando el mejor resultado obtenido.

ReduceLRonPlateau: Ha diferencia del EarlyStopping, este método no detiene el entrenamiento, sino que reduce la tasa de aprendizaje cuando una métrica ha dejado de mejorar, el límite se establece en base a las épocas de 'paciencia'.

Se realizó una prueba inicial con los datos para observar cual era la mejor configuración de activación y de corrección. Como se observa en la Tabla 4, la mejor configuración es en la que se usa el método de activación Relu, lo cual valida lo expuesto en la literatura [15], [16] y la función de corrección EarlyStopping.

Tabla 4. Error con diferentes funciones de activación y de corrección. Fuente propia.

CONFIGURACIÓN		MAE	MSE
Activación	Callback		
Relu	ReduceLRonPlateau	10.01	155.28
Relu	ModelCheckpoint	10.47	164.48
Relu	EarlyStopping	9.92	148.67
Tanh	ReduceLRonPlateau	11.05	175.39
Tanh	ModelCheckpoint	11.06	177.98
Tanh	EarlyStopping	11.05	175.09
Sigmoid	ReduceLRonPlateau	11.05	175.45

Sigmoid	ModelCheckpoint	11.05	177.46
Sigmoid	EarlyStopping	11.05	175.09

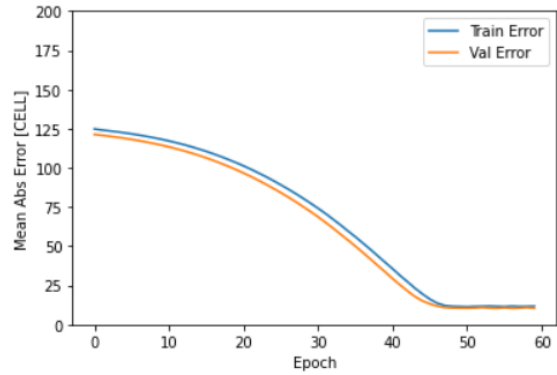
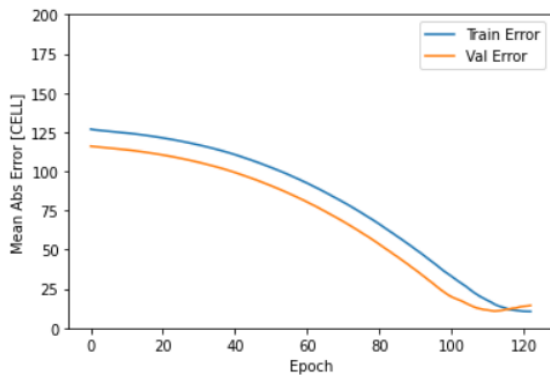


Ilustración 3. Pérdida de entrenamiento y de validación. Fuente propia

La pérdida de validación fue monitoreada y evaluada contra la pérdida de entrenamiento para asegurar que no exista sobreajuste, lo cual sucede si la pérdida de validación es mayor que la de entrenamiento [17]. Se puede observar que los valores de pérdida de validación no fueron significativamente mayores que las pérdidas de entrenamiento como se denota en la Ilustración 3. **Error! No se encuentra el origen de la referencia..** Se encuentra una mayor relación cuando se entrena el modelo con los datos sintéticos.

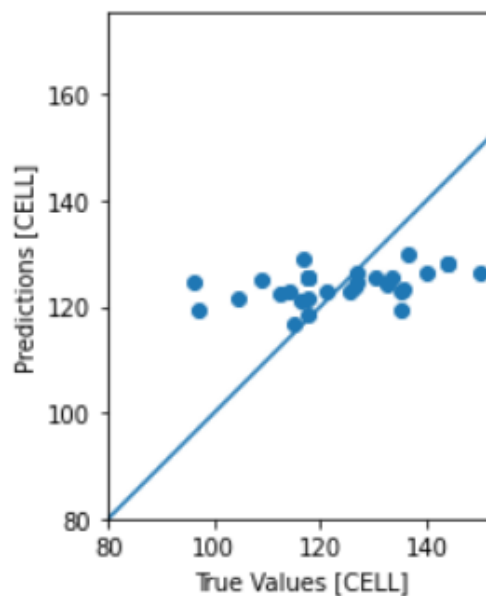
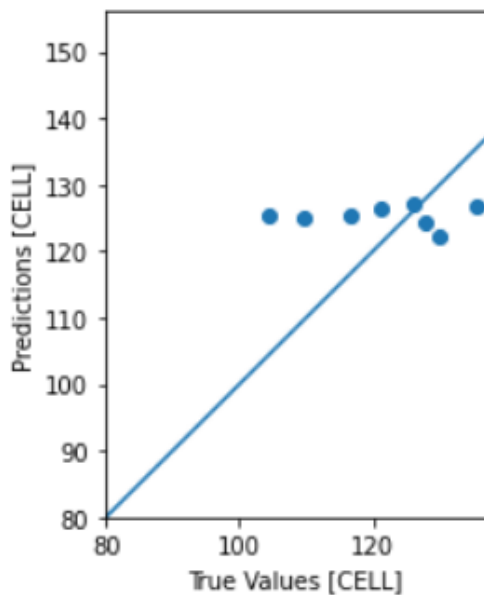


Ilustración 4. Distribución de valores respecto a la recta de predicción. Fuente propia

En la Ilustración 4 se observa la distribución de los valores a validar en relación con la recta de predicción. Entre más cerca estén los valores estimados a la recta de predicción, mejor habrá sido el modelo.

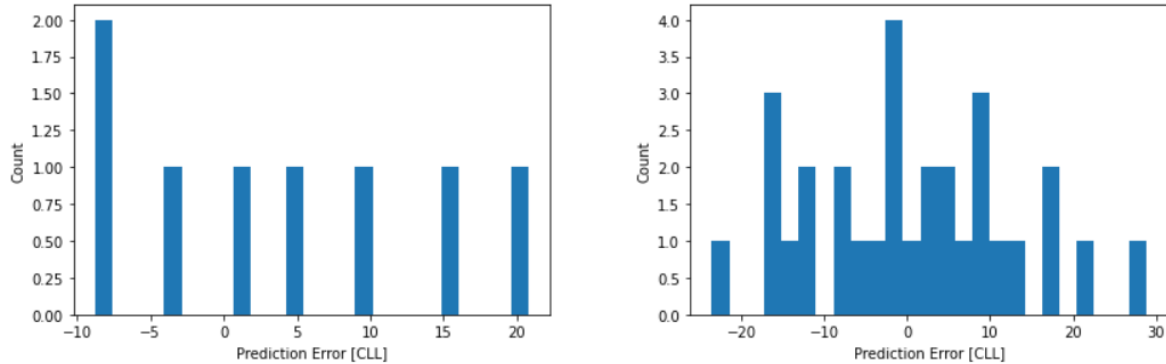


Ilustración 5. Distribución del error. Fuente propia.

Al aumentar el número de muestras, como es esperado y se observa en la Ilustración 5 la distribución del error debería acercarse a la distribución gaussiana, la cual es la distribución normal y esperada al analizar el error de un modelo [18], [19].

Se realizó el procedimiento realizado previamente para analizar el comportamiento de cada uno de los índices de vegetación con la configuración que mostró el mejor desempeño (activación es relu y su callback de corrección es EarlyStopping).

Cuando se entrenan con todos los datos (con todos los índices de vegetación), se observa que la red neuronal artificial ofrece una mejor estimación al tener un error mae y mse más pequeño en comparación con la regresión lineal múltiple. Además, se analizó el comportamiento de la red neuronal artificial con el conjunto de datos inicial Ilustración 6(a) en contraste con el conjunto de datos integrando los datos sintéticos generados Ilustración 6(b).

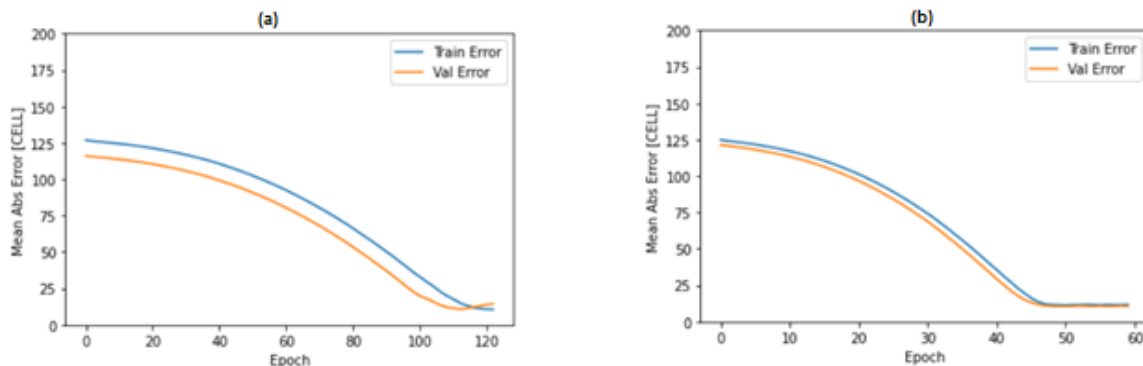


Ilustración 6. Pérdida de entrenamiento y de validación sin datos sintéticos (a) y con datos sintéticos (b). Fuente propia.

La pérdida de validación fue monitoreada y evaluada contra la pérdida de entrenamiento para asegurar que no exista sobreajuste, lo cual sucede si la pérdida de validación es mayor que la de entrenamiento [17]. Se puede observar que los valores de pérdida de validación no fueron significativamente mayores que las pérdidas de entrenamiento como se denota en la Ilustración 6.

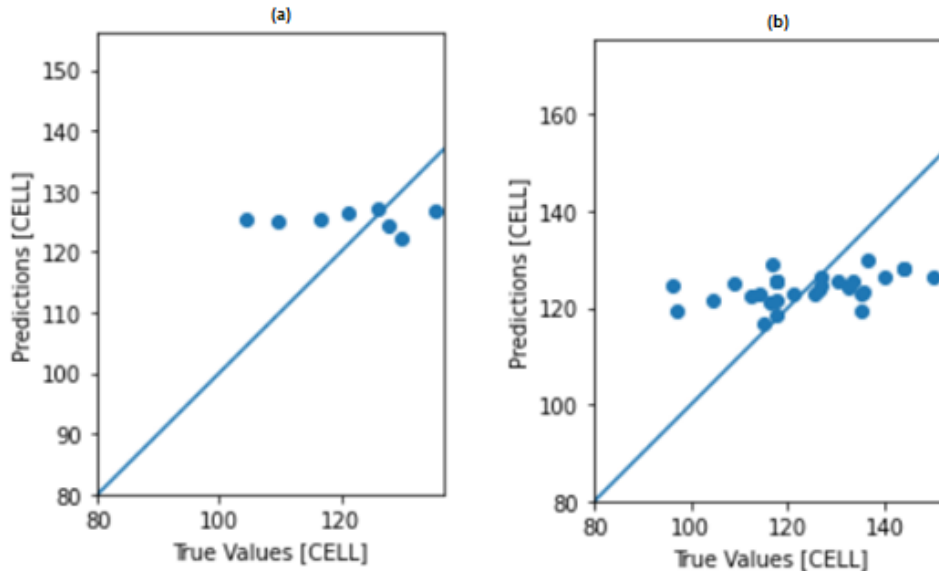


Ilustración 7. Distribución de valores respecto a la recta de predicción. Fuente propia

En la Ilustración 7, se observa la distribución de los valores a validar en relación con la recta de predicción. Entre más cerca estén los valores estimados a la recta de predicción, mejor habrá sido el modelo. Se observa una mejoría en la cantidad de datos cercanos a la recta cuando se entrenan con los datos sintéticos (b) en contraste con los datos iniciales (a).

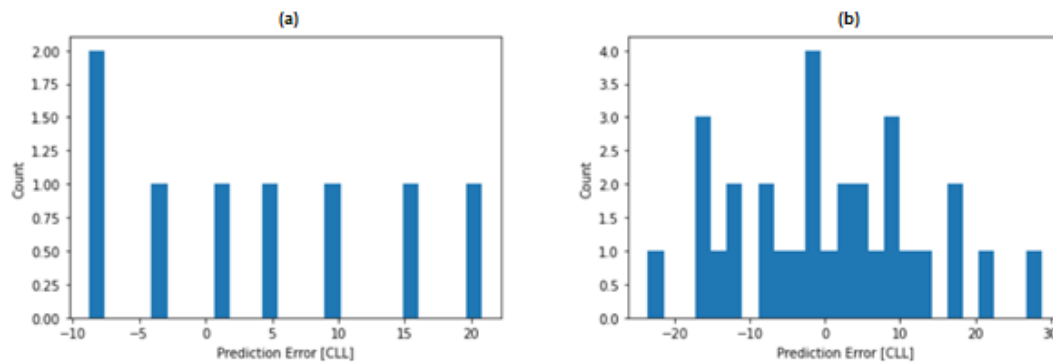


Ilustración 8. Distribución del error. Fuente propia.

Al aumentar el número de muestras, como es esperado y se observa en la Ilustración 8, la distribución del error se acerca a la distribución gaussiana cuando se agregan los datos sintéticos Ilustración 8 (b), mejorando respecto a la Ilustración 8 (a), la cual es la distribución normal y esperada al analizar el error de un modo.

Se ejecutó el procedimiento realizado previamente para analizar el comportamiento de cada uno de los índices de vegetación con la configuración que mostró el mejor desempeño (activación es relu y su callback de corrección es EarlyStopping).

De manera similar, se aplica el análisis efectuado anteriormente para la Ilustración 9, contrastando el comportamiento de los datos iniciales en comparación con los datos integrando los datos sintéticos. El mejor comportamiento fue dado por el GNDVI como se observa en la comparación de errores de la Tabla 5. **Error! No se encuentra el origen de la referencia.** El comportamiento en la Ilustración 9 Ilustración 9 del GNDVI representa la mejoría del modelo al comparar los datos iniciales Ilustración 9(a) con los datos integrando los datos sintéticos generados Ilustración 9(b).

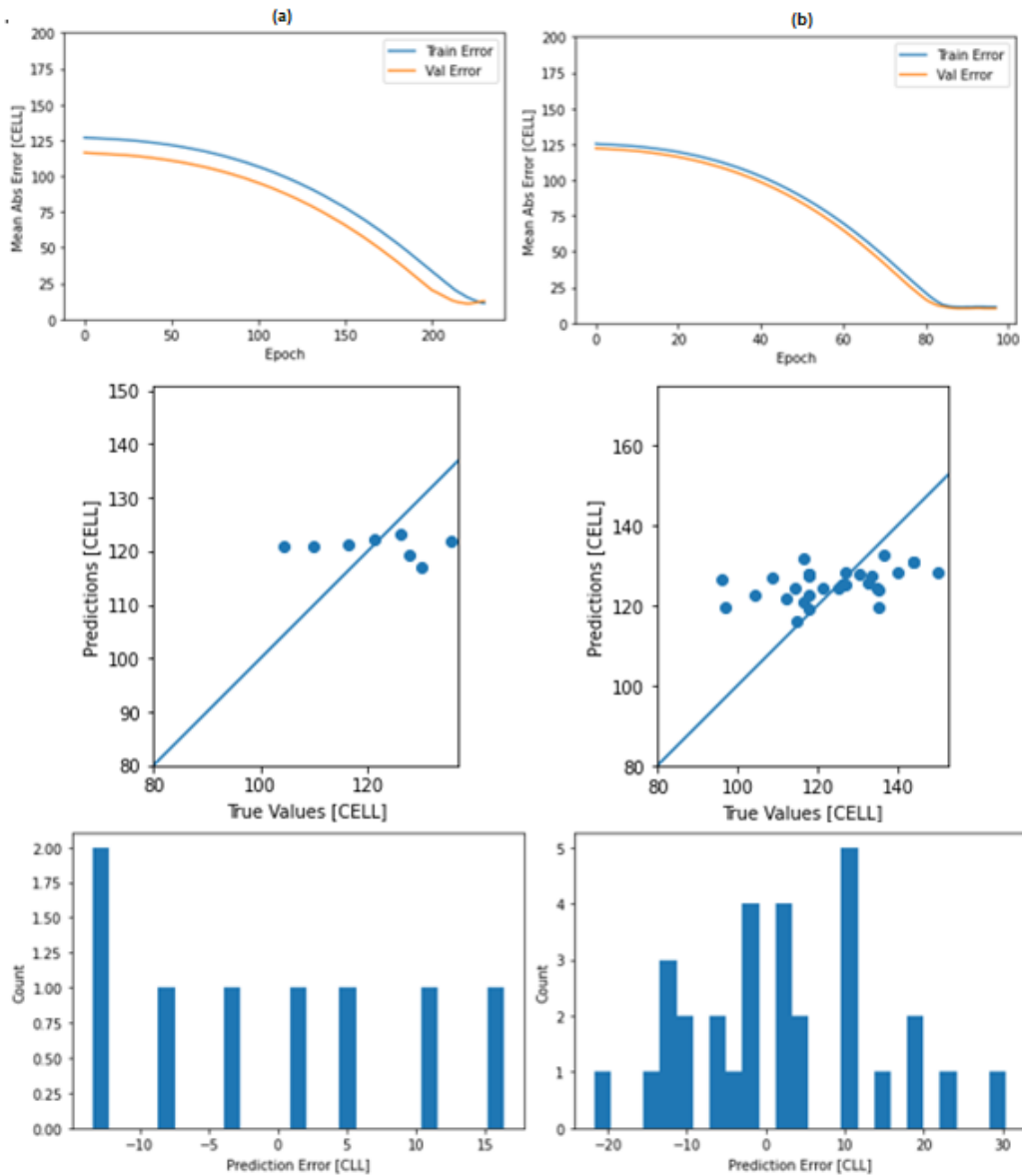


Ilustración 9. Resultados de predicción para GNDVI. Fuente propia.

Tabla 5. Error por cada índice de vegetación en cada modelo. Fuente propia

	Regresión lineal		Red neuronal artificial	
	mae	mse	mae	mse
NDVI	10.77	177.97	11.01	178.67
GNDVI	10.80	178.37	9.69	147.15
NRVI	10.74	177.03	10.87	171.65
CVI	10.97	182.26	9.99	157.70
GCI	10.91	179.82	10.46	164.50
RVI	10.83	179.74	10.33	159.57

El procedimiento descrito anteriormente se realizó para cada índice de vegetación como se muestran de la Ilustración 10 a la Ilustración 15.

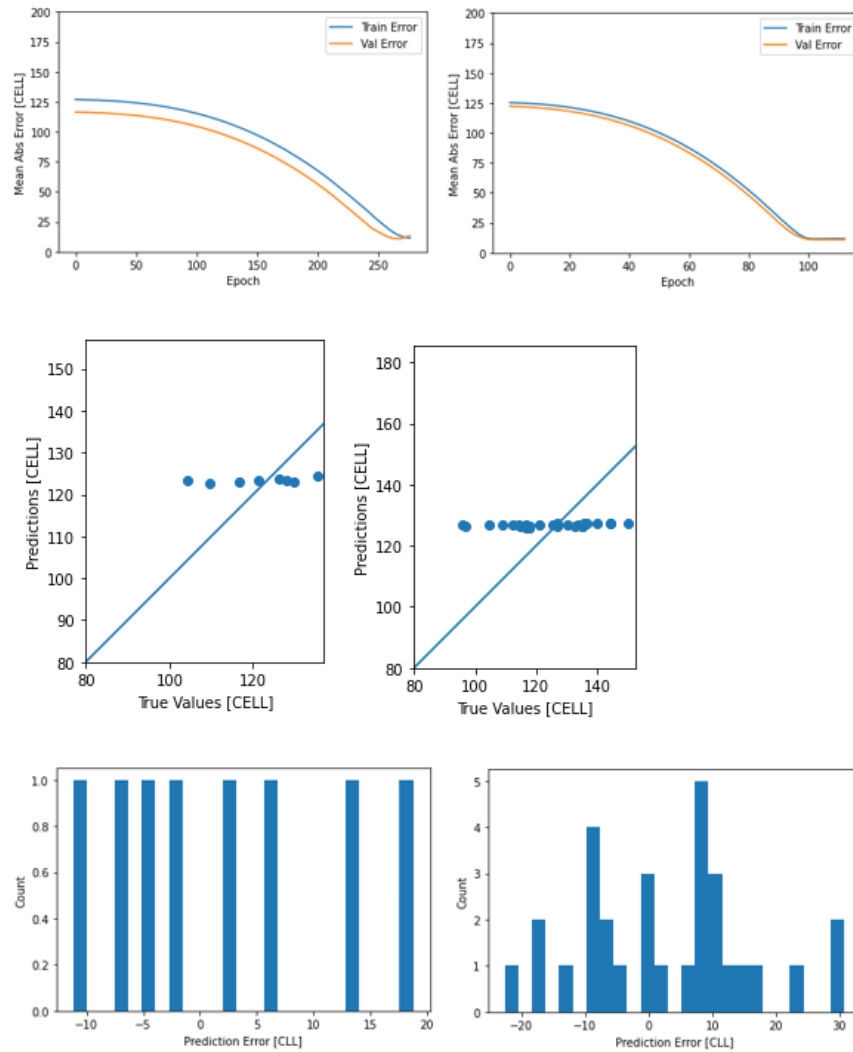


Ilustración 10. Resultados de predicción para NDVI. Fuente propia.

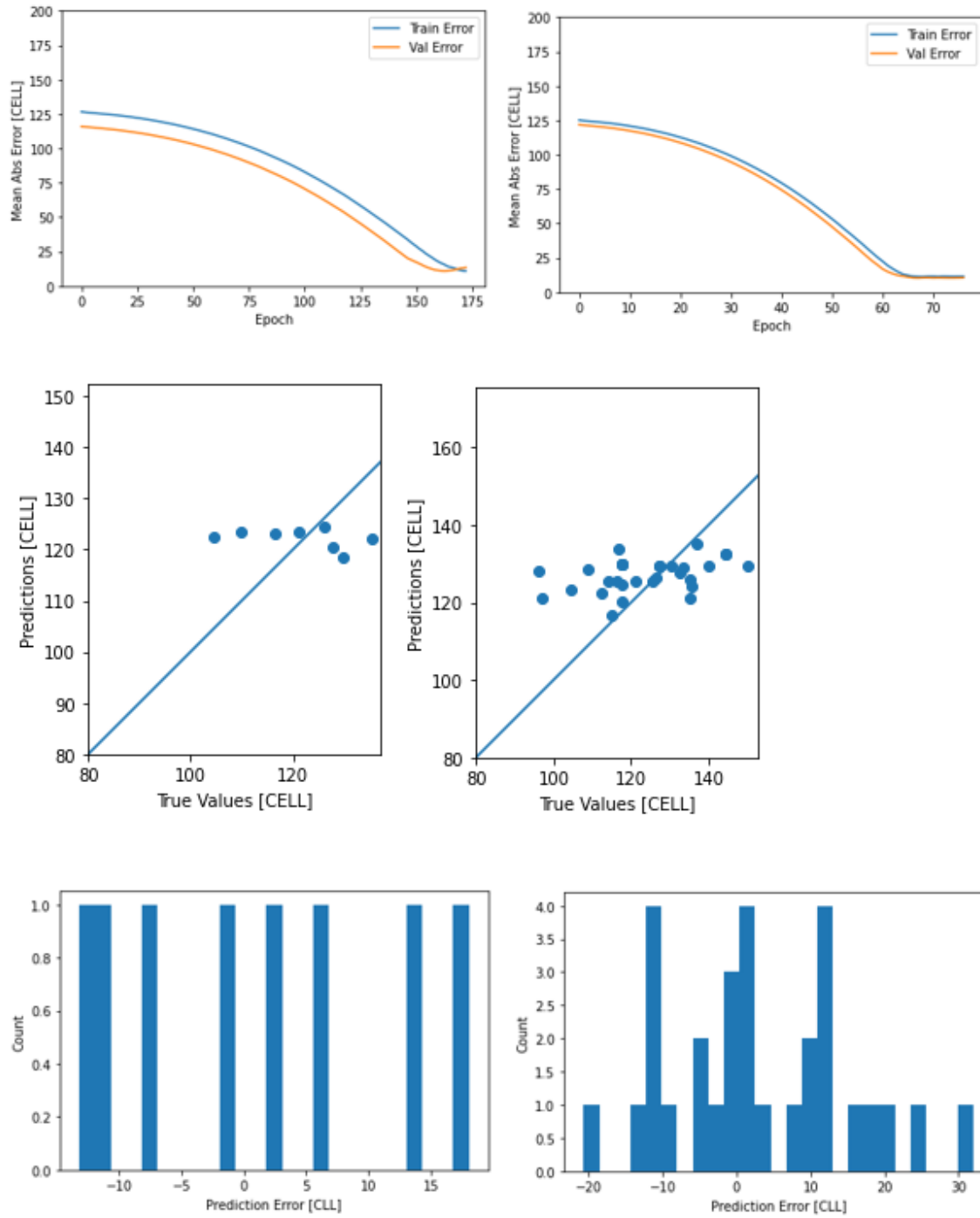


Ilustración 11. Resultados de predicción para CVI. Fuente propia.

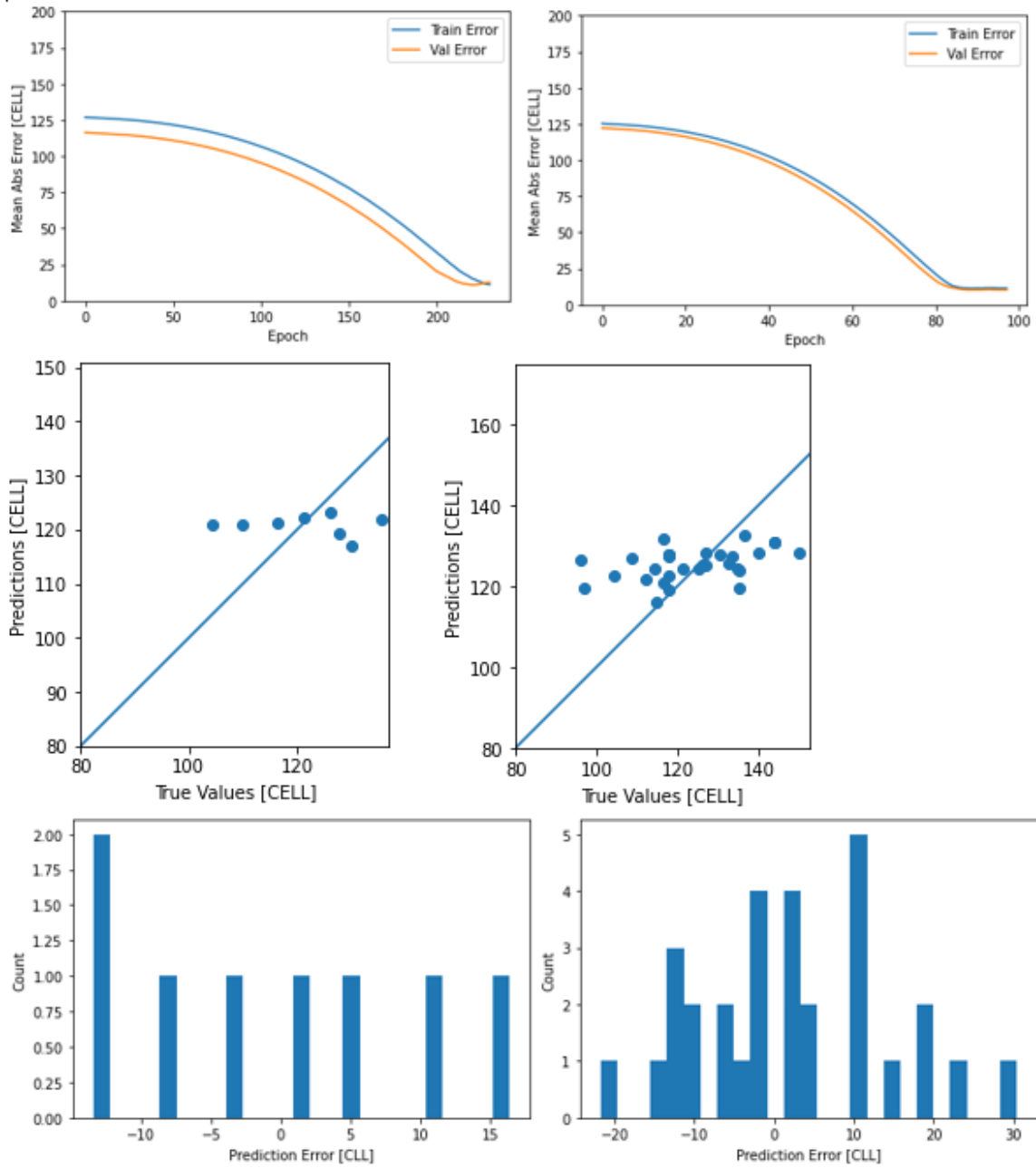


Ilustración 12. Resultados de predicción para GNDVI. Fuente propia.

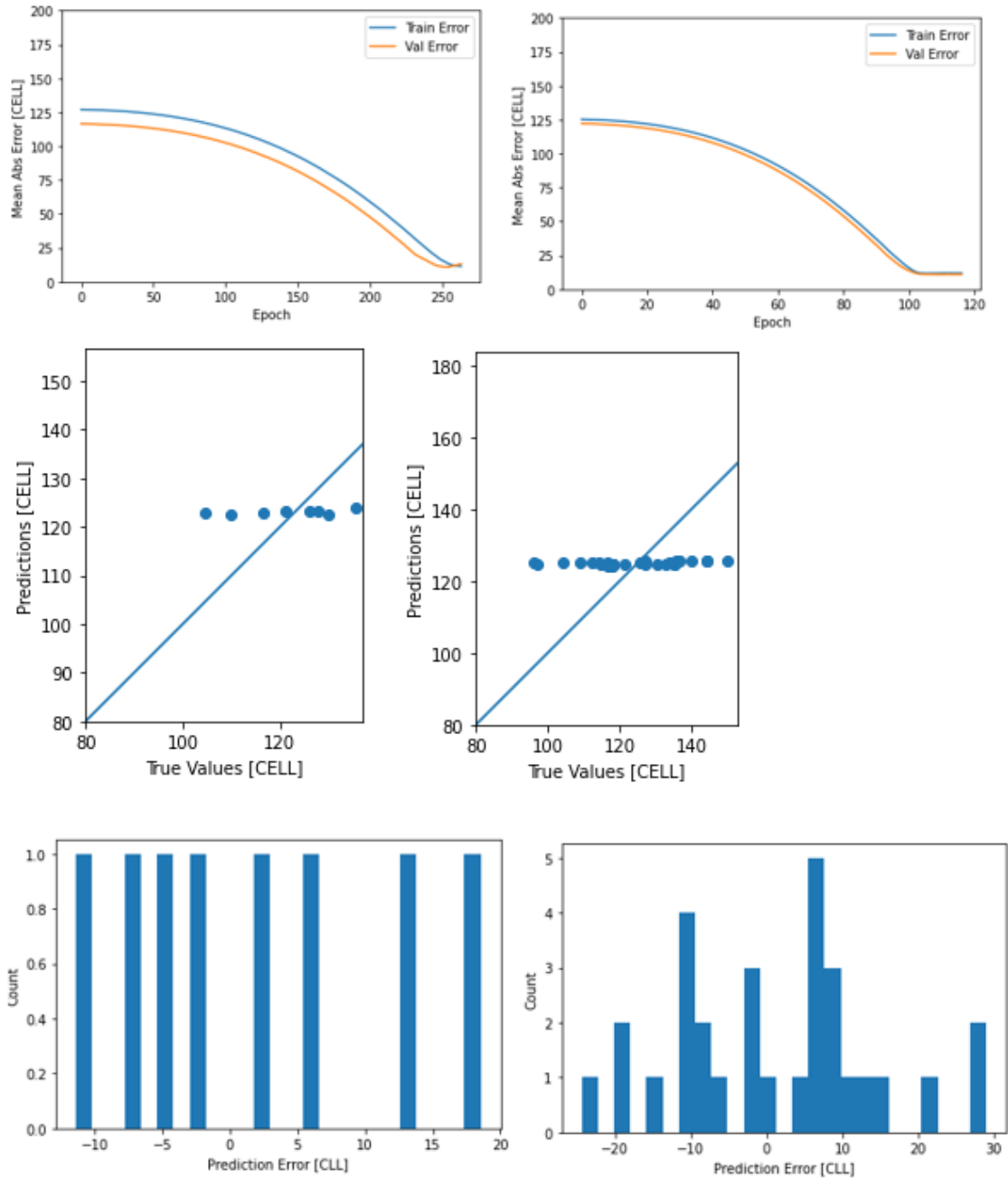


Ilustración 13. Resultados de predicción para NRVI. Fuente propia.

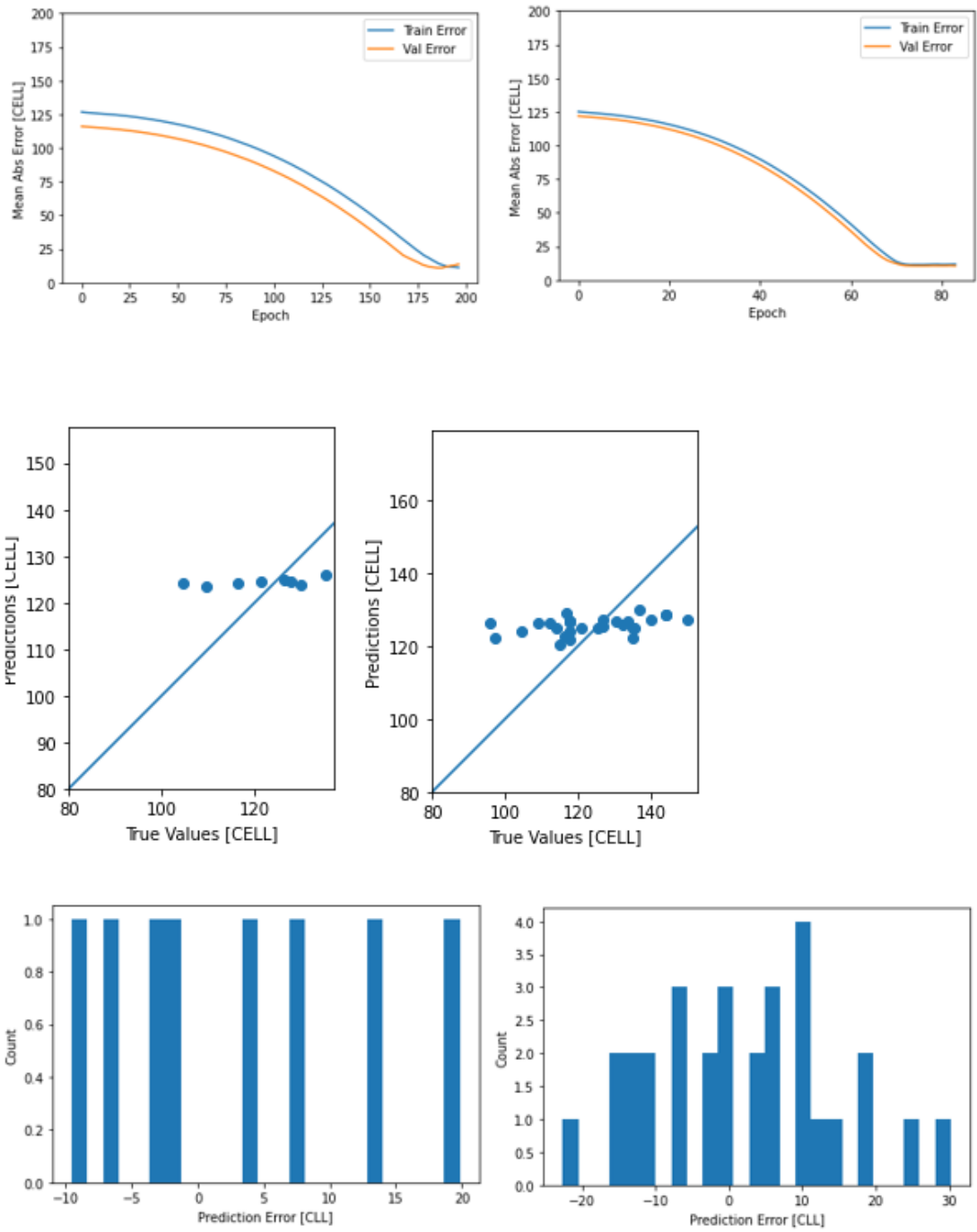


Ilustración 14. Resultados de predicción para RVI. Fuente propia.

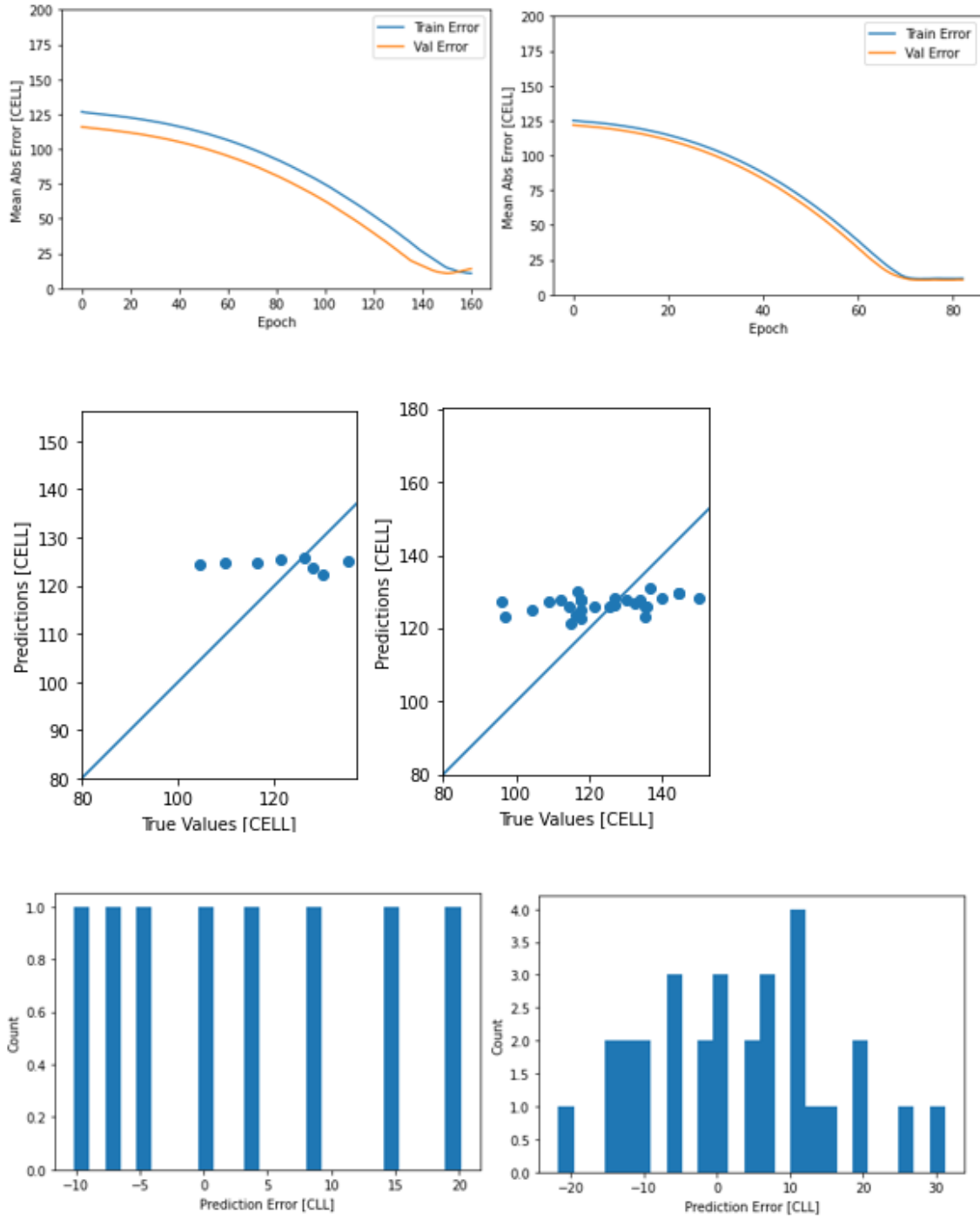


Ilustración 15. Resultados de predicción para GCI. Fuente propia.

ANEXO J: Artículo y Poster

Poster y artículo titulado 'Uso de UAV equipados con cámaras multiespectrales para el análisis de cultivos de café' IEEE COLCOM (*Colombian Conference on Communications and Computing*) del 7-8 de agosto 2020. También se pueden encontrar en

<https://drive.google.com/drive/folders/1oZ3G2l82romPiLMzaLfokMbKNqRyhL1A?usp=sharing>

ANEXO K: Artículo resultados de la investigación

Artículo que contiene los resultados de la investigación, consta de introducción, estado del arte, materiales y métodos y conclusiones. Este artículo también se encuentra en <https://drive.google.com/drive/folders/11A-FEbQjFIG4n5br0PtZA3OLmKWzErW?usp=sharing>

ANEXO L: Imágenes Multiespectrales Obtenidas

Banco de imágenes multiespectrales de un cultivo de café de variedad Castillo ubicado en una finca en San Joaquín, Tambo, Cauca. Las imágenes se obtuvieron de la cámara MAPIR SURVEY 3 aerotransportada en un VANT SOLO 3DR entre los meses de febrero y marzo de 2021, las cuales se utilizaron y se procesaron el programa agisoft metashape y posteriormente QGIS. El banco de imágenes se encuentra en

https://drive.google.com/drive/folders/1GEHuU3yGkkcitbcP_HZV1Sd9vx2RldQ1?usp=sharing

REFERENCIAS

- [1] J. Tremblay *et al.*, “Training Deep Networks with Synthetic Data: Bridging the Reality Gap by Domain Randomization,” 2018.
- [2] A. Kortylewski, B. Egger, A. Schneider, T. Gerig, A. Morel-Forster, and T. Vetter, “Analyzing and Reducing the Damage of Dataset Bias to Face Recognition with Synthetic Data,” 2019.
- [3] Y. Toda *et al.*, “Training instance segmentation neural network with synthetic datasets for crop seed phenotyping,” *Commun. Biol.*, vol. 3, no. 1, pp. 1–12, Dec. 2020.
- [4] T. A. Le, A. G. Baydin, R. Zinkov, and F. Wood, “Using synthetic data to train neural networks is model-based reasoning,” in *Proceedings of the International Joint Conference on Neural Networks, 2017*, vol. 2017-May, pp. 3514–3521.
- [5] F. S. Saleh, “Effective Use of Synthetic Data for Urban Scene Semantic Segmentation *,” 2018.
- [6] Enzyme, “¿Qué son las redes neuronales y cómo se aplican?” [Online]. Available: <https://blog.enzymeadvisinggroup.com/redes-neuronales-que-son-y-aplicaciones>. [Accessed: 21-May-2021].
- [7] Keras, “Keras: the Python deep learning API.” [Online]. Available: <https://keras.io/>. [Accessed: 21-May-2021].
- [8] J. Han and C. Moraga, “The influence of the sigmoid function parameters on the speed of backpropagation learning,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 1995, vol. 930, pp. 195–201.
- [9] K. G. Kim, “Deep learning,” *Nature*, vol. 29, no. 7553, pp. 1–73, 2019.
- [10] R. M. Neal, “Connectionist learning of belief networks,” *Artif. Intell.*, vol. 56, no. 1, pp. 71–113, Jul. 1992.
- [11] V. Nair and G. E. Hinton, “Rectified Linear Units Improve Restricted Boltzmann Machines.”
- [12] P. Ramachandran, B. Zoph, and Q. V. Le, “Searching for Activation Functions,” *arXiv*, Oct. 2017.
- [13] E. Thoret, T. Andriillon, D. Léger, and D. Pressnitzer, “Probing machine-learning classifiers using noise, bubbles, and reverse correlation,” *bioRxiv*. bioRxiv, p. 2020.06.22.165688, 23-Jun-2020.
- [14] Keras, “Callbacks.” [Online]. Available: <https://keras.io/api/callbacks/>. [Accessed: 20-May-2021].
- [15] M. D. Zeiler *et al.*, “On rectified linear units for speech processing,” in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, 2013*, pp. 3517–3521.
- [16] G. E. Dahl, T. N. Sainath, and G. E. Hinton, “Improving deep neural networks for LVCSR using rectified linear units and dropout,” in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, 2013*, pp. 8609–8613.
- [17] R. Roberts, L. Inzerillo, and G. Di Mino, “Exploiting data analytics and deep learning systems to support pavement maintenance decisions,” *Appl. Sci.*, vol. 11, no. 6, p. 2458, Mar. 2021.

- [18] L. Tian, T. Cai, E. Goetghebeur, and L. J. Wei, "Model Evaluation Based on the Distribution of Estimated Absolute Prediction Error Model Evaluation Based on the Distribution of Estimated Absolute Prediction Error."
- [19] Physics, "Error Analysis." [Online]. Available: http://felix.physics.sunysb.edu/~allen/252/PHY_error_analysis.html. [Accessed: 20-May-2021].