

**MODELO DE PREDICCIÓN DEL RENDIMIENTO DE UN CULTIVO DE CAFÉ
MEDIANTE ANÁLISIS DE IMÁGENES AÉREAS MULTIESPECTRALES**



Universidad
del Cauca

Julián Andrés Bolaños Ortega

Anexos

Director:

PhD(C). Liseth Viviana Campo

Co Director:

PhD. Juan Carlos Corrales

Asesor:

PhD. Cristhian Figueroa

Universidad del Cauca

Facultad de Ingeniería Electrónica y Telecomunicaciones

Departamento de Telemática

Línea de Investigación e-@mbiente

Popayán, Enero de 2023

JULIAN ANDRES BOLAÑOS ORTEGA

**MODELO DE PREDICCIÓN DEL RENDIMIENTO DE UN CULTIVO DE CAFÉ
MEDIANTE ANÁLISIS DE IMÁGENES AÉREAS MULTIESPECTRALES**

**Tesis presentada en a la facultad de Ingeniería Electrónica y
Telecomunicaciones de la Universidad del Cauca para la obtención del
título de**

Magister En:
Ingeniería Telemática

Director:
PhD(C). Liseth Viviana Campo

Co Director:
PhD. Juan Carlos Corrales

Asesor:
PhD. Cristhian Figueroa

Popayán
2023

TABLA DE CONTENIDO

- 1. Correlaciones** **7**
- 2. Gráficas** **11**
- 3. Aplicativo Software** **17**
 - 3.1. Arquitectura del flujo de datos 17
 - 3.1.1. Tabla Planta 18
 - 3.1.2. Tabla Lote 21
- 4. Retos Recolección de muestras** **23**

Lista de Tablas

1.1. Correlaciones del Rendimiento con Volumen, Max Red, Max Blue, variance BlueCriMax y MaxGreen	8
1.2. Correlaciones del Rendimiento con MtviOne, ndviMean, latLaiArea, saviMax, tviMax y ndviMax	8

Anexo 1

Correlaciones

El presente anexo presenta en forma de tabla las correlaciones más importantes comparadas con el rendimiento del cultivo y una gráfica ampliada de correlaciones que incluye algunas variables adicionales que las mostradas en el documento principal. Las tablas 1.1 y 1.2 corresponden a la información calculada mediante el uso de las 373 muestras de rendimiento obtenidas de los cultivos de prueba. Existió un proceso inicial de análisis realizado con 103 muestras que ayudó a explorar el comportamiento de las variables. Las correlaciones obtenidas se muestran en la figura 1.1. En este proceso inicial se evidenció que habían variables ya correlacionadas como el volumen y el ndviMax, pero el resto de variables importantes no lograban valores importantes comparados con los mostrados en las tablas 1.1 y 1.2.

	yields	volumenImagen	volumenCalculado	criMax	areaLateral	alturalmagen	alturaMedida	tviMin	ndviMax	cgreenviMax
yields	1.000000	0.760897	0.754905	0.505919	0.767902	0.719736	0.730056	-0.428355	0.598940	0.456593
volumenImagen	0.760897	1.000000	0.948400	0.427770	0.993002	0.832837	0.709445	-0.583990	0.452548	0.347101
volumenCalculado	0.754905	0.948400	1.000000	0.402826	0.932618	0.715321	0.810952	-0.530451	0.469610	0.341151
criMax	0.505919	0.427770	0.402826	1.000000	0.460574	0.515144	0.466527	-0.771753	0.333457	0.905484
areaLateral	0.767902	0.993002	0.932618	0.460574	1.000000	0.882646	0.739974	-0.609942	0.438408	0.386764
alturalmagen	0.719736	0.832837	0.715321	0.515144	0.882646	1.000000	0.754090	-0.619832	0.352235	0.465195
alturaMedida	0.730056	0.709445	0.810952	0.466527	0.739974	0.754090	1.000000	-0.494318	0.392954	0.464862
tviMin	-0.428355	-0.583990	-0.530451	-0.771753	-0.609942	-0.619832	-0.494318	1.000000	-0.178391	-0.657251
ndviMax	0.598940	0.452548	0.469610	0.333457	0.438408	0.352235	0.392954	-0.178391	1.000000	0.271050
cgreenviMax	0.456593	0.347101	0.341151	0.905484	0.386764	0.465195	0.464862	-0.657251	0.271050	1.000000

Figura 1.1: Matriz de Correlación Ampliada: Contiene la mayoría de las variables involucradas en el proceso

Cuadro 1.1: Correlaciones del Rendimiento con Volumen, Max Red, Max Blue, variance BlueCriMax y MaxGreen

	yields	volImage	maxRed	maxBlue	varianceBlue	criMax	maxGreen
yields	1.00	0.71	0.65	0.64	0.60	0.63	0.61
volImage	0.71	1.00	0.62	0.61	0.62	0.62	0.59
maxRed	0.65	0.62	1.00	0.92	0.80	0.95	0.93
maxBlue	0.64	0.61	0.92	1.00	0.77	0.81	0.82
varianceBlue	0.60	0.62	0.80	0.77	1.00	0.77	0.85
criMax	0.63	0.62	0.95	0.81	0.77	1.00	0.91
maxGreen	0.61	0.59	0.93	0.82	0.85	0.91	1.00

Cuadro 1.2: Correlaciones del Rendimiento con MtviOne, ndviMean, laiLatArea, saviMax, tviMax y ndviMax

	yields	mtviOne Mean	ndviMean	laiLatArea	saviMax	tviMax	ndviMax
yields	1.00	0.53	0.49	0.69	0.49	0.49	0.55
mtviOne Mean	0.53	1.00	0.76	0.45	0.65	0.65	0.68
ndviMean	0.49	0.76	1.00	0.39	0.46	0.46	0.61
laiLatArea	0.69	0.45	0.39	1.00	0.55	0.55	0.58
saviMax	0.49	0.65	0.46	0.55	1.00	1.00	0.95
tviMax	0.49	0.65	0.46	0.55	1.00	1.00	0.95
ndviMax	0.55	0.68	0.61	0.58	0.95	0.95	1.00

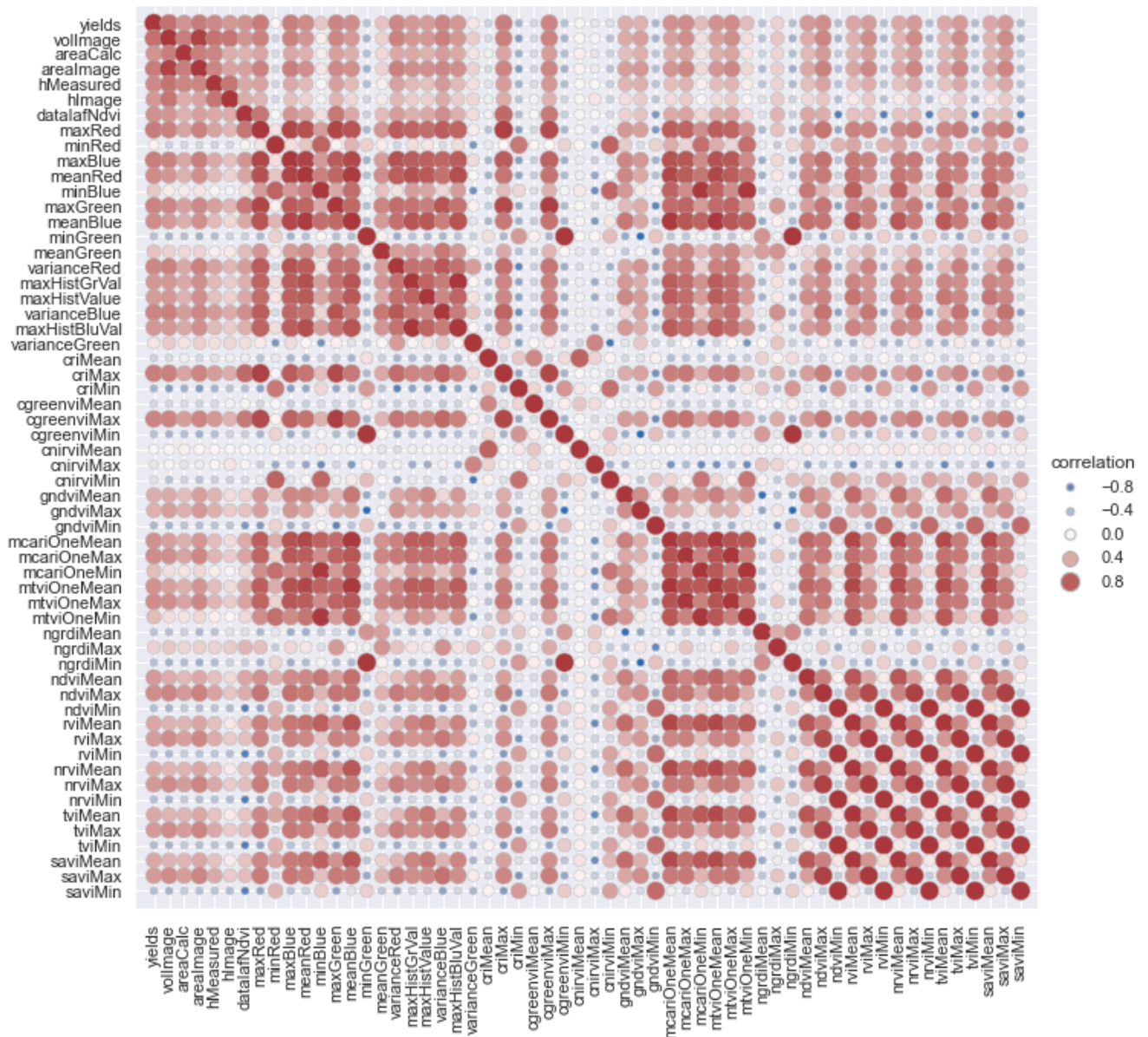


Figura 1.2: Matriz de Correlación Ampliada: Contiene la mayoría de las variables involucradas en el proceso

Anexo 2

Gráficas

El presente anexo presenta gráficas adicionales que involucran comportamientos de variables con el rendimiento.

La gráfica 2.1 presenta la distribución que tiene el el rendimiento respecto y sus valores de referencia.

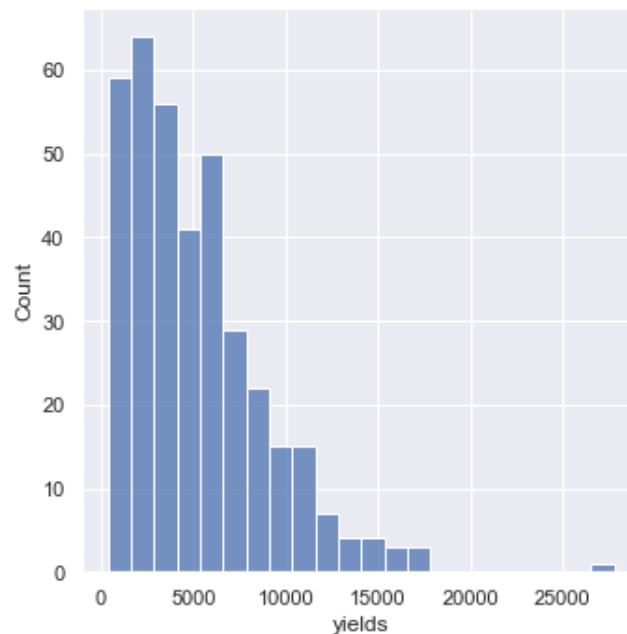


Figura 2.1: Gráfico de Dispersión del rendimiento

La figura 2.2 compara los valores de la banda azul en sus máximos, medios y mínimos con el rendimiento.

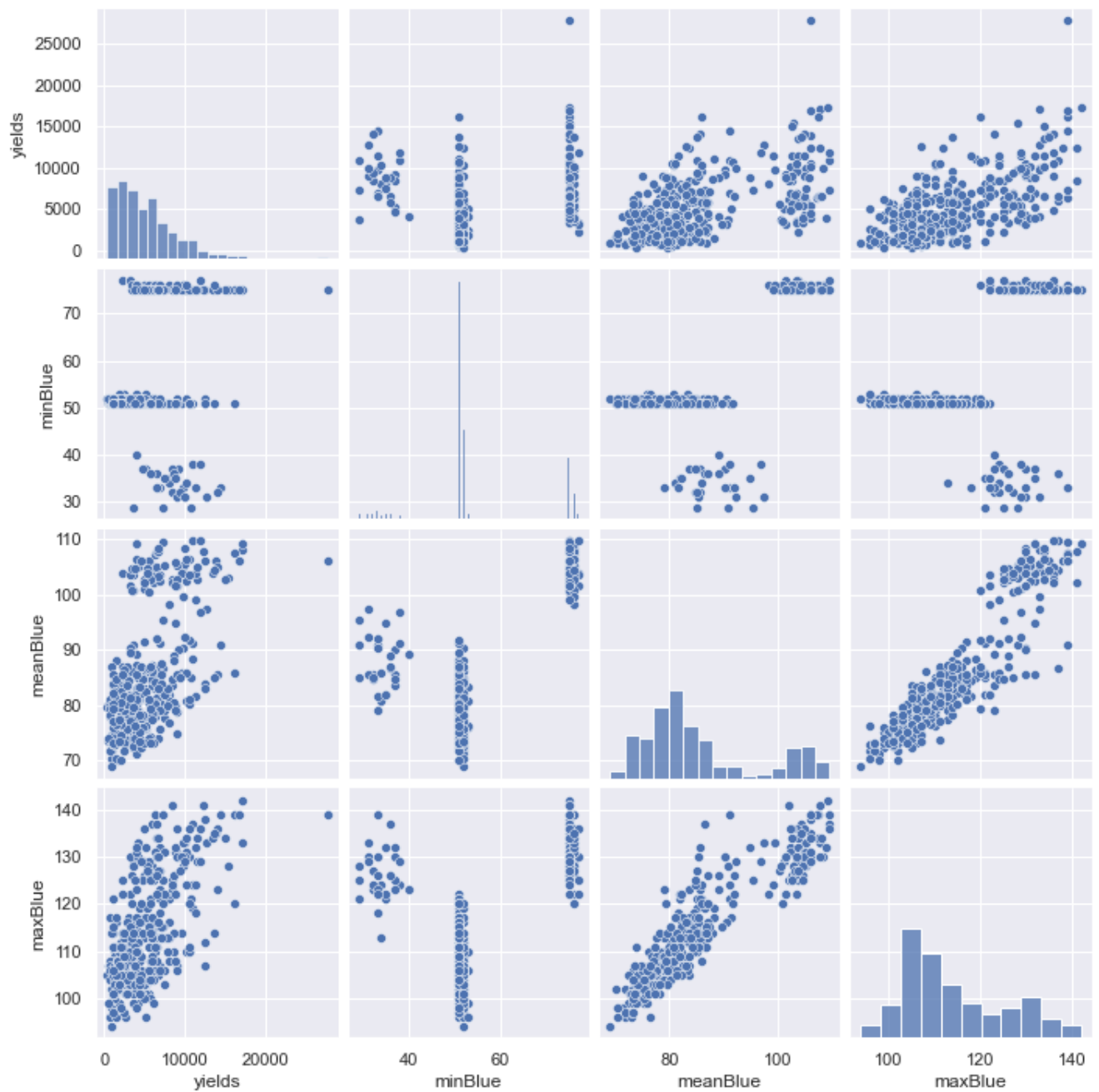


Figura 2.2: Bandas Azules comparadas con el rendimiento

La figura 2.3 compara los valores de la banda verde en sus máximos, medios y mínimos con el rendimiento.

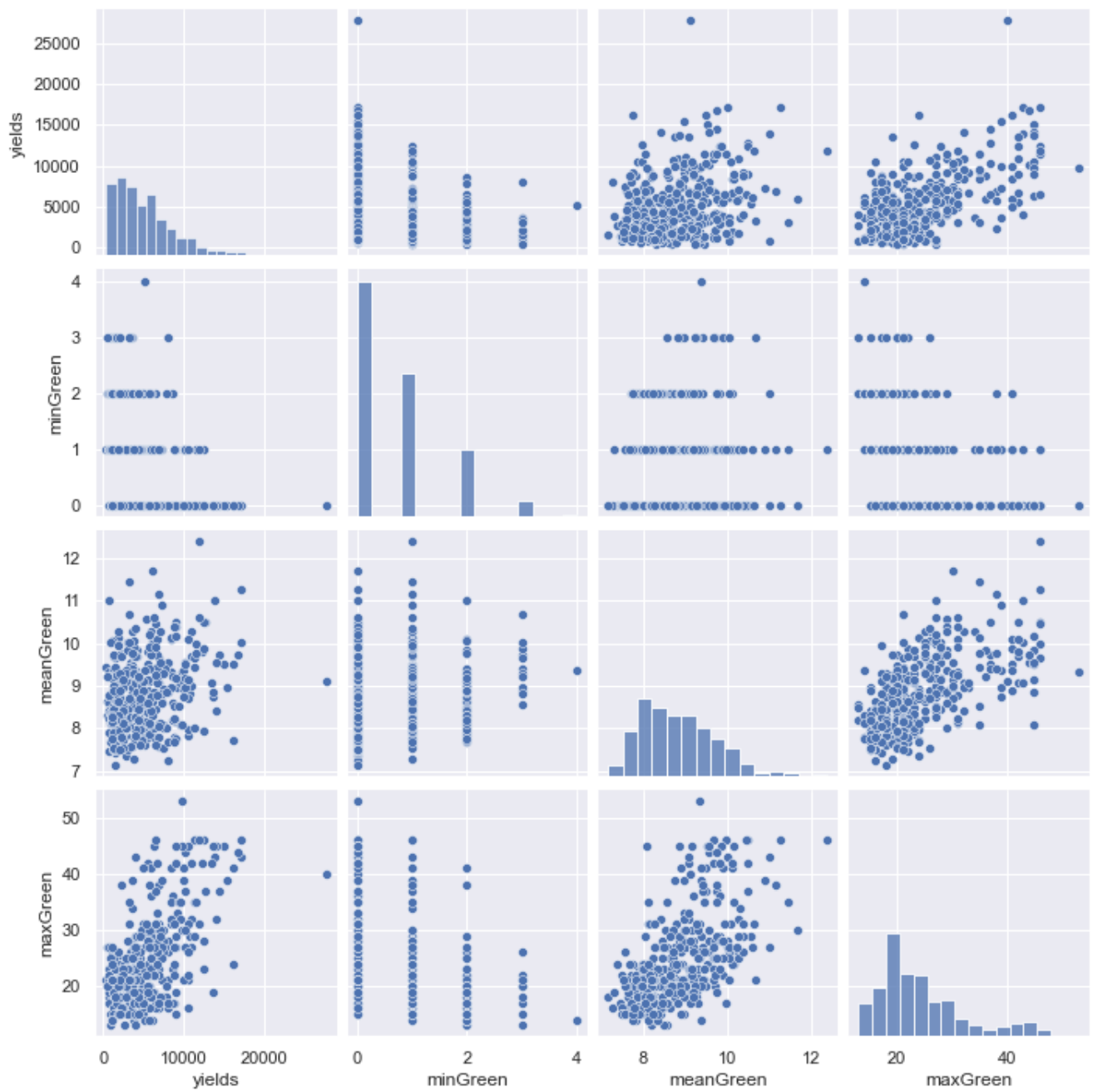


Figura 2.3: Bandas Verdes comparadas con el rendimiento

La figura 2.4 compara los valores de la banda roja en sus máximos, medios y mínimos con el rendimiento.

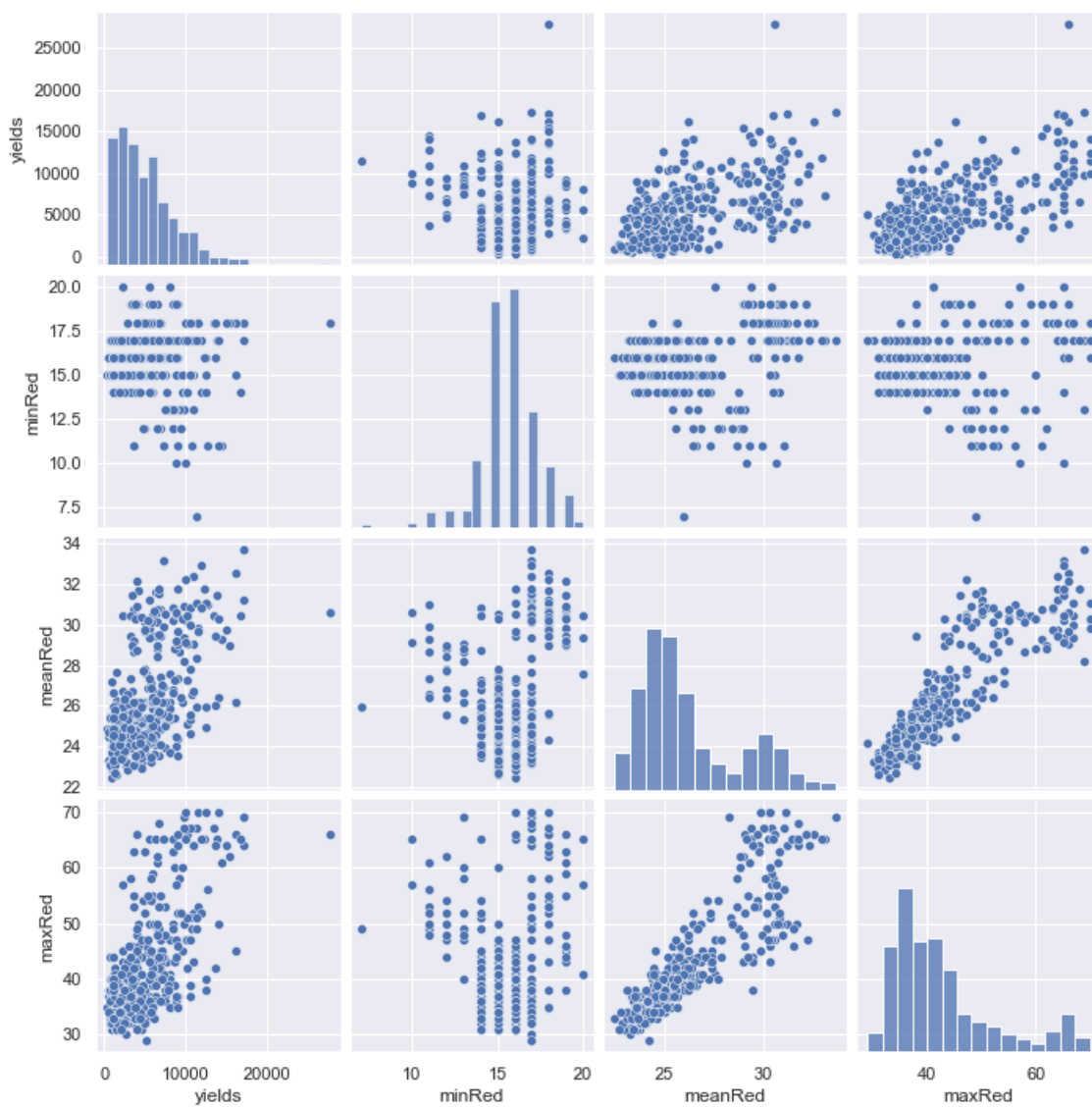


Figura 2.4: Bandas Rojas comparadas con el rendimiento

La figura 2.6 compara los valores de NDVI en sus máximos, medios y mínimos con el rendimiento.

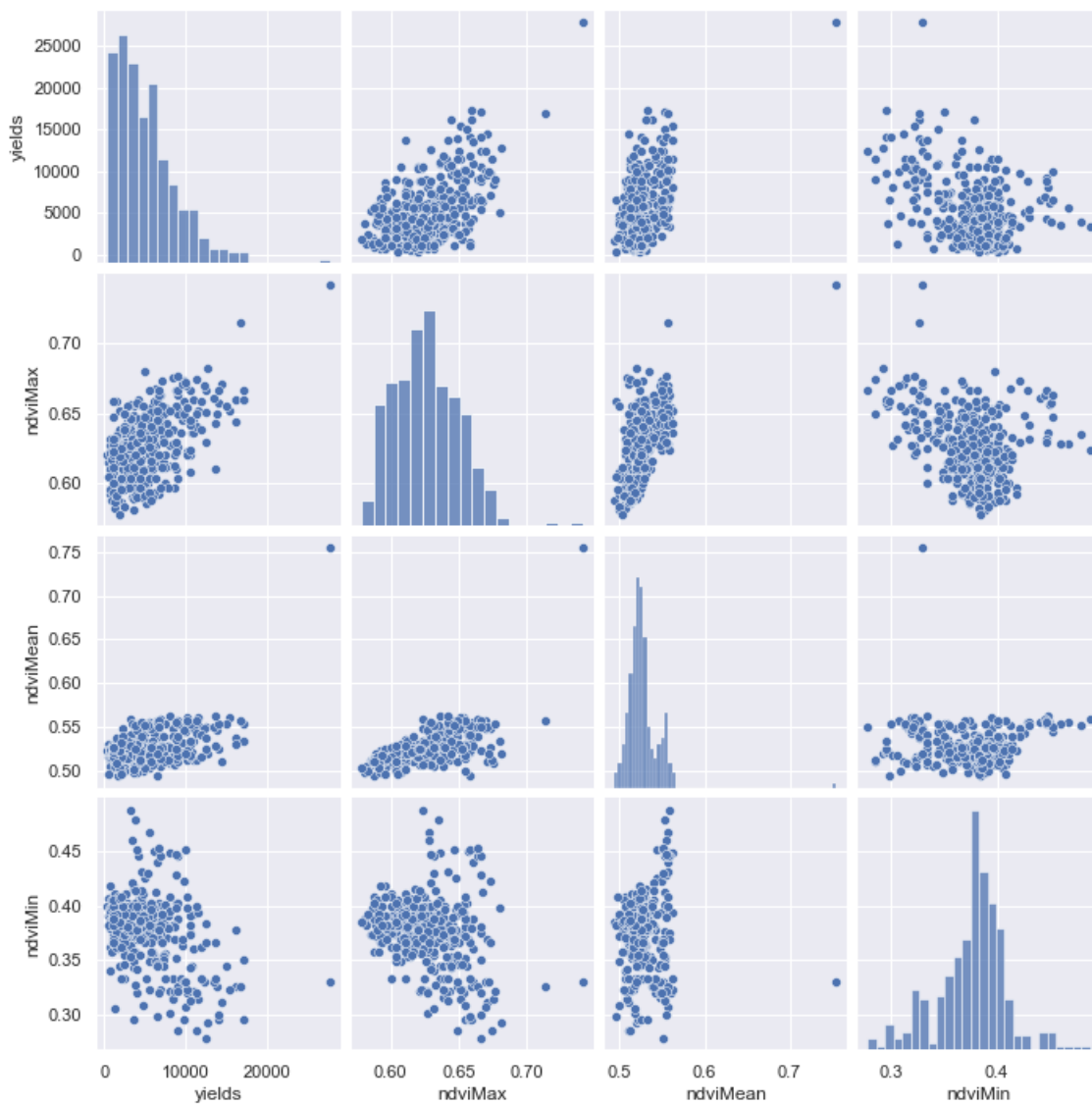


Figura 2.5: NDVI comparado con el rendimiento

La figura 2.6 analiza la dispersión de NDVI gráficos de BoxPlot.

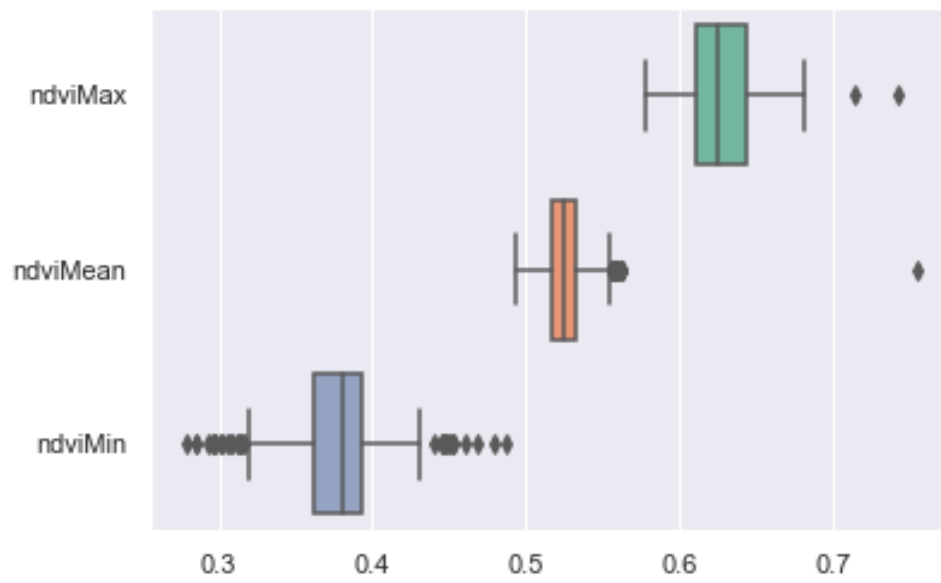


Figura 2.6: NDVI boxplot

Anexo 3

Aplicativo Software

El presente anexo presenta una explicación extensa sobre las tecnologías y las herramientas involucradas en la solución software para la recolección manual de muestras.

3.1. Arquitectura del flujo de datos

El proceso general del flujo de datos es mostrado en la figura 3.1.

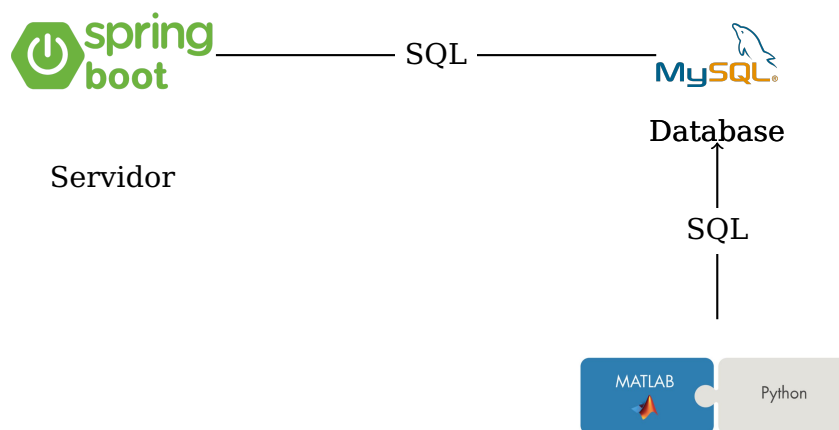


Figura 3.1: Arquitectura de base de datos. Fuente: propia

El servidor para la creación de la interfaz y el servicio REST que permite el intercambio de la información con los usuarios de la aplicación móvil se basa en el motor de Spring Boot en lenguaje Java. Este motor contiene las herramientas necesarias para poder generar un API de forma segura y eficiente. De igual forma, el proceso de analítica de datos está construido en lenguaje Python, permitiendo generar e insertar las plantas individuales y facilitando herramientas de post procesamiento

para el análisis de datos. La base de datos involucrada corresponde a un motor de base de datos relacional MySQL versión 5.7.

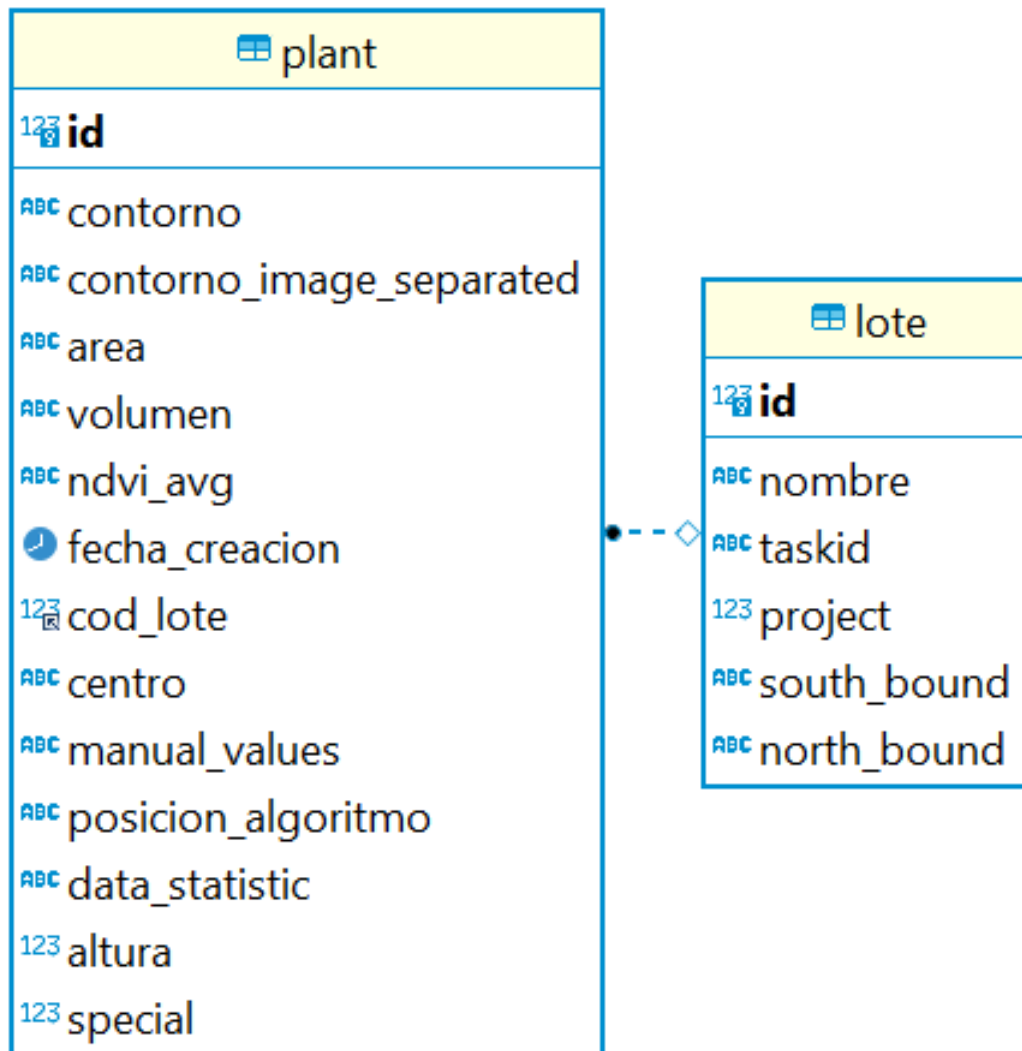


Figura 3.2: Diagrama ER de base de datos

La figura 3.2 expone las tablas que estuvieron involucradas en el proceso de obtención de información.

3.1.1. Tabla Planta

Esta tabla está destinada a almacenar las plantas ya individualizadas y detectadas por el algoritmo de segmentación. Presenta los siguientes campos:

- **contorno:** El contorno corresponde a un arreglo de coordenadas geográficas que establecen la posición de la planta detectada en un mapa en formato

WGS84.

- **contorno image separated:** Esta variable establece los cuatro puntos límite donde se encuentra la imagen de la planta individual, esto en formato WGS84.
- **area,volumen:** Estas variables corresponden a las medidas físicas obtenidas usando las librerías de Python y el API de WebODM respectivamente. El volumen finalmente fue desestimado por su inexactitud, pero hubo un proceso realizado para la obtención de dicha información basado en el cálculo de volumen de WebODM.
- **ndvi avg:** Esta variable inicialmente estaba contemplada para insertar los valores promedio del NDVI, sin embargo, se convirtió en la columna donde se insertan todas las variables obtenidas de los índices de vegetación en formato JSON de la siguiente forma:

```
1 {
2   "criMean": 0.9997435364500818,
3   "criMax": 1.625517788981259,
4   "criMin": 0.5541537916981565,
5   "cgreenviMean": 0.9991199623142342,
6   "cgreenviMax": 2.825730145713745,
7   "cgreenviMin": 0,
8   "cnirviMean": 0.999826208281424,
9   "cnirviMax": 1.314637324967752,
10  "cnirviMin": 0.5830130745509161,
11  "gndviMean": 0.7926492702643498,
12  "gndviMax": 1,
13  "gndviMin": 0.5652173913043478,
14  "mcariOneMean": 60.756272235112384,
15  "mcariOneMax": 89.75999999999999,
16  "mcariOneMin": 18.119999999999994,
17  "mtviOneMean": 60.756272235112384,
18  "mtviOneMax": 89.75999999999999,
19  "mtviOneMin": 18.119999999999994,
20  "ngrdiMean": -0.46398226970315204,
21  "ngrdiMax": -0.15384615384615385,
22  "ngrdiMin": -1,
23  "ndviMean": 0.5262694216897984,
24  "ndviMax": 0.673469387755102,
25  "ndviMin": 0.36585365853658536,
26  "rviMean": 3.24056307830495,
27  "rviMax": 5.125,
```

```
28 "rviMin": 2.1538461538461537,  
29 "nrviMean": 0.5262694216897984,  
30 "nrviMax": 0.673469387755102,  
31 "nrviMin": 0.36585365853658536,  
32 "tviMean": 1.0129235048183085,  
33 "tviMax": 1.083267920578793,  
34 "tviMin": 0.9305125783870873,  
35 "saviMean": 0.7858924212778536,  
36 "saviMax": 1.0050761421319796,  
37 "saviMin": 0.5454545454545454  
38 }
```

- **cod lote:** Este elemento permite identificar los lotes establecidos por id, esta relación tiene una llave foránea con la tabla lote.
- **centro:** Corresponde al coordenada geográfica del centro de la planta definido en WGS 84.
- **manual values:** Pertenece a la información manual recolectada por la aplicación. También es definida en formato JSON de la siguiente forma:

```
1 {  
2 "altura": "192",  
3 "diametro": "172",  
4 "status": "3",  
5 "numRamasHeader": "34",  
6 "numNodesHeader": "398",  
7 "numBeansHeader": "400",  
8 "numRamasFooter": "16",  
9 "numNodesFooter": "647",  
10 "numBeansFooter": "435"  
11 }
```

- **posición algoritmo:** Es una variable de control para establecer en parte del algoritmo de detección iba un proceso para luego poder re iniciarlo si algo fallaba. También para identificar la planta insertada dentro del flujo general.
- **data statistic:** Corresponde a la data estadística de las bandas calculada mediante el proceso de recorrer y medir cada una de las bandas mientras se calculaban los índices de vegetación almacenada en formato JSON como se muestra:

```
1 {
```

```
2  "maxRed": 52,  
3  "minRed": 13,  
4  "maxBlue": 124,  
5  "meanRed": 29.01014297088656,  
6  "minBlue": 38,  
7  "maxGreen": 32,  
8  "meanBlue": 91.13365873922524,  
9  "minGreen": 0,  
10 "meanGreen": 10.268979125558207,  
11 "varianceRed": 29.895590532066183,  
12 "maxHistGrVal": 32,  
13 "maxHistValue": 106,  
14 "varianceBlue": 12.045213328061612,  
15 "maxHistBluVal": 32,  
16 "varianceGreen": 380.93910349678094,  
17 "maxHistRedIndexValue": 2  
18 }
```

- **altura:** Corresponde a los valores de altura que fueron calculados por el programa de Matlab basado en el área y el DSM.
- **special:** Algunas plantas fueron marcadas como especiales solo para fines del programa sin efectos en los resultados finales.

3.1.2. Tabla Lote

Como el proceso de análisis puede estar involucrado más de un lote dentro del proceso, se define la tabla lote para lograr identificar las características que corresponden a cada uno. Los campos de la tabla lote se definen a continuación

- **Nombre:** El nombre del lote permite identificar más fácilmente al programador al lograr relacionar el id con la planta.
- **taskid:** Identificador único en el software WebODM de orthophoto de donde se extrajo el lote.
- **project:** Proyecto exacto en WebODM donde se extrajo el lote.
- **south y north bounds:** El software de la librería LeafletJS requiere unos límites en coordenadas geográficas para permitir mostrar los orthophotos mediante sub imágenes y facilitar la carga en los dispositivos móviles. Estos límites se deben guardar para tenerlos como referencia y hacer un llamado correcto al API de tiles de WebODM.

Todo el código del servidor web y los scripts de la generación de base de datos están disponibles en el enlace: <https://github.com/julianandres/coffeunicauca>

Anexo 4

Retos Recolección de muestras

Durante el tiempo de recolección de datos existieron retos los cuales deben ser tenidos en cuenta al realizar este tipo de actividades, estos son descritos a continuación:

- Problemas de conexión Wifi dada la extensión de terreno, fue necesario conectar un cable UTP de 50 metros aproximadamente para lograr cubrir la extensión del terreno.
- Falta de disponibilidad de conexión a la energía. Se necesita una conexión a la red eléctrica lo más cerca posible al sitio de la prueba.
- Problemas de batería en los móviles al realizar la prueba.
- Picaduras por parte de hormigas a las personas que recogieron las muestras.