



UNIVERSIDAD  
**NACIONAL**  
DE COLOMBIA

# **Variabilidad en la respuesta espectral de especies forestales en un contexto urbano**

**Melissa Maya Mesa**

Universidad Nacional de Colombia  
Facultad de Minas, Departamento Geociencias y Medio Ambiente  
Medellín, Colombia  
2019



# **Variabilidad en la respuesta espectral de especies forestales en un contexto urbano**

**Melissa Maya Mesa**

Tesis o trabajo de investigación presentada(o) como requisito parcial para optar al título  
de:

**Magister en Medio Ambiente y Desarrollo**

Director (a):

Ph.D. Verónica Botero Fernández

Línea de Investigación:

Teledetección y Sistemas de Información Geográfica

Universidad Nacional de Colombia  
Facultad de Minas, Departamento Geociencias y Medio Ambiente  
Medellín, Colombia  
2019



## **Agradecimientos**

A mis padres y Lukas por su apoyo incondicional.

A mi tutora, Verónica Botero Fernández por su acompañamiento y orientación, especialmente por estar siempre presente en el proceso elaboración de este ejercicio.

A Juan Carlos Marín por su acompañamiento y asesoría en los momentos más críticos de este proceso.

A Motos porque sin ti y tu estatura no hubiera sido posible la recolección de datos para llevar a cabo la investigación.

A Ericka Guana por los aportes valiosos en la fase de campo.

A todos los que de alguna manera permitieron materializar el que he denominado “mi sueño”.

Gracias.



## Resumen

La discriminación correcta de especies es fundamental para la gestión del arbolado urbano y su cuantificación mediante inventarios forestales. En el presente estudio se evaluó el potencial de clasificación de 3 especies forestales: *Erythrina fusca*, *Ficus benjamina* y *Terminalia catappa*, aplicando un Análisis Discriminante Lineal a los datos obtenidos a partir de una imagen de alta resolución espacial con 4 bandas espectrales (R, G, B y NIR) y la adquisición de espectros foliares de reflectancia tomados en 3 alturas de la copa. Los individuos muestreados estaban localizados en 3 Campus universitarios de la ciudad de Medellín: Universidad Nacional, Eafit y Universidad de Medellín.

La clasificación a partir de los datos multiespectrales obtuvo mejores resultados con la información espectral sin transformar con una precisión general de 67,25%. Sin embargo, las precisiones individuales para las especies *E. fusca* y *F. benjamina* fueron 87,41% y 83,33%, respectivamente; *T. catappa* no fue discriminada con este tipo de datos. A partir de los datos hiperespectrales, se observó una mejor clasificación con los espectros transformados con Varianza Normal Estándar (SNV), obteniendo una precisión general de 78,69%, y específica de 82,99%, 51,68% y 94,89% para *E. fusca*, *F. benjamina* y *T. catappa*, respectivamente. Las bandas seleccionadas con el algoritmo *Relief* que influyeron en la discriminación espectral de especies correspondieron a una región del espectro visible (400 – 420 nm) y del borde rojo (716 – 745 nm), además de un intervalo en el infrarrojo cercano (841 – 907 nm). Se evidenciaron diferencias estadísticamente significativas ( $P < 0,05$ ) en la reflectancia foliar dentro del perfil vertical del dosel, mostrando un patrón descendente desde la altura superior a la inferior. No se encontraron diferencias estadísticamente significativas en los espectros foliares de acuerdo con la clasificación de los árboles según la distancia y tipo a las fuentes contaminantes (vías). Estos resultados sugieren que los datos hiperespectrales son una fuente potencial de información para la identificación de especies forestales y pueden proporcionar un conocimiento *a priori* de la composición florística en la ejecución de inventarios forestales urbanos.

**Palabras clave:** Discriminación de especies, Datos multiespectrales, Datos hiperespectrales, Análisis Discriminante Lineal, Posición del dosel, Árboles urbanos.

# Contenido

	Pág.
<b>1. Introducción .....</b>	<b>19</b>
1.1 Planteamiento del problema y justificación .....	19
1.2 Antecedentes .....	20
1.3 Pregunta de investigación .....	23
1.4 Hipótesis .....	23
1.5 Objetivo general .....	23
1.6 Objetivos específicos .....	23
<b>2. Marco conceptual y revisión de literatura .....</b>	<b>25</b>
2.1 Conceptos y términos.....	25
2.1.1 Teledetección .....	25
2.1.2 Sensores remotos.....	25
2.1.3 Espectroscopía .....	26
2.1.4 Respuesta espectral típica de la vegetación sana .....	26
2.1.5 Métodos estadísticos para el procesamiento de datos espectrales.....	28
2.2 Antecedentes de la aplicación de la teledetección en el estudio de la vegetación	31
2.2.1 Aplicación de la espectroscopia en la discriminación de especies vegetales .	34
2.2.2 Aplicación de la espectroscopía en especies vegetales caso de estudio: Colombia.....	37
<b>3. Metodología .....</b>	<b>39</b>
3.1 Área de estudio .....	39
3.1.1 Sitios de muestreo .....	39
3.2 Selección de especies forestales .....	40
3.3 Verificación de los inventarios forestales de los sitios de muestreo .....	41
3.4 Toma y procesamiento de datos .....	45
3.4.1 Datos multiespectrales.....	47
3.4.2 Datos hiperespectrales .....	50
<b>4. Resultados y discusión .....</b>	<b>63</b>
4.1 Datos multiespectrales .....	63
4.1.1 Caracterización espectral y discriminación de las especies forestales seleccionadas en el estudio .....	64
4.1.2 Caracterización espectral y discriminación de las especies: <i>E. fusca</i> , <i>F. benjamina</i> y <i>T. catappa</i> .....	73
4.2 Datos hiperespectrales.....	81
4.2.1 Caracterización espectral .....	82
4.2.2 Discriminación espectral de las especies forestales .....	85



---

4.2.3 Variabilidad en la respuesta espectral considerando la posición del dosel y la distancia a una fuente contaminante .....	88
<b>5. Conclusiones y recomendaciones.....</b>	<b>93</b>
5.1 Conclusiones.....	93
5.2 Recomendaciones.....	94
<b>6. Nota aclaratoria .....</b>	<b>96</b>
<b>7. Anexos .....</b>	<b>97</b>



## Lista de figuras

	Pág.
<b>Figura 1–1:</b> Firma espectral característica de la vegetación sana.....	21
<b>Figura 2–1:</b> Firma espectral característica de la vegetación sana.....	28
<b>Figura 2–2:</b> Mapa mundial de la localización de 101 estudios realizados sobre identificación de especies (Puntos rojos).....	34
<b>Figura 3–1:</b> Localización general del área de estudio. ....	40
<b>Figura 3–2:</b> Localización de los árboles seleccionados en el Campus de la Universidad Nacional de Colombia.....	43
<b>Figura 3–3:</b> Localización de los árboles seleccionados en el Campus de la Universidad de Antioquia.....	43
<b>Figura 3–4:</b> Localización de los árboles seleccionados en el Campus de la Universidad de Medellín. ....	44
<b>Figura 3–5:</b> Localización de los árboles seleccionados en el Campus de la Universidad Eafit. ....	44
<b>Figura 3–6:</b> Actividades realizadas para la toma y procesamiento de datos espectrales. ....	46
<b>Figura 3–7:</b> Copas delimitadas de los árboles objeto de estudio. ....	47
<b>Figura 3–8:</b> Resumen del análisis estadístico realizado a partir de los datos multiespectrales. ....	49
<b>Figura 3–9:</b> Equipos de medición empleados. ....	51
<b>Figura 3–10:</b> Localización de los árboles seleccionados para el muestreo hiperespectral en la Universidad Nacional de Colombia de las especies <i>E. fusca</i> , <i>F.</i> <i>benjamina</i> y <i>T. catappa</i> . ....	52
<b>Figura 3–11:</b> Niveles altitudinales de medición. ....	53
<b>Figura 3–12:</b> Resumen del análisis estadístico realizado a partir de los datos hiperespectrales. ....	57
<b>Figura 3–13:</b> Clasificación de los árboles por tipo y distancia a las vías en la UNAL. .	59
<b>Figura 3–14:</b> Clasificación de los árboles por tipo y distancia a las vías en Eafit. ....	60
<b>Figura 3–15:</b> Clasificación de los árboles por tipo y distancia a las vías en la Universidad de Medellín. ....	60
<b>Figura 4–1:</b> Respuesta espectral media de las 10 especies forestales seleccionadas.. .....	66

<b>Figura 4–2:</b>	Respuesta espectral media de las 10 especies forestales seleccionadas por sitios de muestreo. ....	69
<b>Figura 4–3:</b>	Análisis de Componentes Principales con las 10 especies considerando todos los sitios de muestreo y datos transformados con SNV. ....	70
<b>Figura 4–4:</b>	Proporción de la varianza explicada en el PCA con las especies <i>E. fusca</i> , <i>F. benjamina</i> y <i>T. catappa</i> considerando todos los sitios de muestreo y datos transformados con SNV. ....	71
<b>Figura 4–5:</b>	Análisis Discriminante Lineal con las 10 especies considerando todos los sitios de muestreo y datos transformados con SNV. ....	72
<b>Figura 4–6:</b>	Respuesta espectral media de las especies <i>E. fusca</i> , <i>F. benjamina</i> y <i>T. catappa</i> en los 4 sitios de muestreo. ....	74
<b>Figura 4–7:</b>	Variación de la respuesta espectral en función del diámetro de copa (DC) de los árboles seleccionados por especie. ....	76
<b>Figura 4–8:</b>	Análisis de Componentes Principales con las especies <i>E. fusca</i> , <i>F. benjamina</i> y <i>T. catappa</i> considerando todos los sitios de muestreo y datos originales... ..	77
<b>Figura 4–9:</b>	Proporción de la varianza explicada en el PCA con las especies <i>E. fusca</i> , <i>F. benjamina</i> y <i>T. catappa</i> considerando todos los sitios de muestreo y datos originales... ..	78
<b>Figura 4–10:</b>	Análisis Discriminante Lineal con las especies <i>E. fusca</i> , <i>F. benjamina</i> y <i>T. catappa</i> considerando todos los sitios de muestreo y datos sin transformar.....	79
<b>Figura 4–11:</b>	Forma de la copa de las especies especies <i>E. fusca</i> , <i>F. benjamina</i> y <i>T. catappa</i> . ....	80
<b>Figura 4–12:</b>	Respuesta espectral media en cada posición del dosel para las especies (A) <i>E. fusca</i> , (B) <i>F. benjamina</i> y (C) <i>T. catappa</i> . ....	84
<b>Figura 4–13:</b>	Variables o bandas elegidas mediante el algoritmo Relief para realizar el LDA a partir de los datos transformados con SNV de las 3 especies seleccionadas. ....	85
<b>Figura 4–14:</b>	Análisis Discriminante Lineal con los espectros transformados SNV de las 3 especies seleccionadas considerando todos los sitios de muestreo. ....	87
<b>Figura 4–15:</b>	Respuesta espectral media en cada posición del dosel para la especie <i>E. fusca</i> en el sitio de muestreo UDEM.....	89
<b>Figura 7–1:</b>	Análisis Discriminante Lineal con las 10 especies considerando todos los sitios de muestreo y datos transformados con MSC.....	98
<b>Figura 7–2:</b>	Análisis Discriminante Lineal con las especies <i>E. fusca</i> , <i>F. benjamina</i> y <i>T. catappa</i> considerando todos los sitios de muestreo y datos transformados con SNV.....	99
<b>Figura 7–3:</b>	Análisis Discriminante Lineal con las especies <i>E. fusca</i> , <i>F. benjamina</i> y <i>T. catappa</i> considerando todos los sitios de muestreo y datos transformados con MSC. ...	100
<b>Figura 7–4:</b>	Análisis Discriminante Lineal con los espectros transformados MSC de las 3 especies seleccionadas considerando todos los sitios de muestreo. ....	101
<b>Figura 7–5:</b>	Análisis Discriminante Lineal con los espectros sin transformar de las 3 especies seleccionadas considerando todos los sitios de muestreo.....	102
<b>Figura 7–6:</b>	Respuesta espectral media en cada posición del dosel para la especie <i>E. fusca</i> en el sitio de muestreo UNAL.....	103
<b>Figura 7–7:</b>	Respuesta espectral media en cada posición del dosel para la especie <i>E. fusca</i> en el sitio de muestreo EAFIT.....	104

---

<b>Figura 7–8:</b> Respuesta espectral media en cada posición del dosel para la especie <i>F. benjamina</i> en el sitio de muestreo UNAL.....	104
<b>Figura 7–9:</b> Respuesta espectral media en cada posición del dosel para la especie <i>F. benjamina</i> en el sitio de muestreo EAFIT. ....	105
<b>Figura 7–10:</b> Respuesta espectral media en cada posición del dosel para la especie <i>T. catappa</i> en el sitio de muestreo UNAL. ....	105
<b>Figura 7–11:</b> Respuesta espectral media en cada posición del dosel para la especie <i>T. catappa</i> en el sitio de muestreo EAFIT.....	106
<b>Figura 7–12:</b> Respuesta espectral media en cada posición del dosel para la especie <i>T. catappa</i> en el sitio de muestreo UDEM. ....	106



# Lista de tablas

	Pág.
<b>Tabla 2–1:</b> Pigmentos foliares y sus máximos de absorción. ....	27
<b>Tabla 3–1:</b> Listado de especies seleccionadas. ....	41
<b>Tabla 3–2:</b> Número de árboles inventariados en los sitios de muestreo correspondientes a las especies seleccionadas. ....	42
<b>Tabla 3–3:</b> Clasificación de las vías de tránsito adyacentes a los sitios de muestreo..	61
<b>Tabla 3–4:</b> Clasificación de los árboles muestreados con respecto a su localización dentro de los Campus universitarios.....	61
<b>Tabla 4–1:</b> Número de árboles y pixeles considerados en el estudio después de eliminar valores atípicos. ....	63
<b>Tabla 4–2:</b> Matriz de confusión obtenida a partir del LDA con las 10 especies considerando todos los sitios de muestreo y datos transformados con SNV. ....	72
<b>Tabla 4–3:</b> Resumen de la variabilidad en la respuesta espectral de las bandas (B1, B2, B3 y B4) en los diferentes sitios de muestreo.....	75
<b>Tabla 4–4:</b> Matriz de confusión obtenida a partir del LDA con las especies <i>E. fusca</i> , <i>F. benjamina</i> y <i>T. catappa</i> considerando todos los sitios de muestreo y datos originales. ...	78
<b>Tabla 4–5:</b> Número de árboles seleccionados en los sitios de muestreo para la medición de espectros.....	82
<b>Tabla 4–6:</b> Matriz de confusión obtenida a partir del LDA con los espectros transformados SNV de las 3 especies seleccionadas considerando todos los sitios de muestreo. ....	88
<b>Tabla 4–7:</b> Resumen de la variabilidad espectral media de las 3 especies seleccionadas en cada posición del dosel considerando los diferentes sitios de muestreo. ....	90
<b>Tabla 4–8:</b> Resumen de la variabilidad espectral media de las 3 especies seleccionadas según la clasificación de los árboles en los diferentes sitios de muestreo... ..	92
<b>Tabla 7–1:</b> Matriz de confusión obtenida a partir del LDA con las 10 especies considerando todos los sitios de muestreo y datos transformados con MSC.....	97
<b>Tabla 7–2:</b> Matriz de confusión obtenida a partir del LDA con las especies <i>E. fusca</i> , <i>F. benjamina</i> y <i>T. catappa</i> considerando todos los sitios de muestreo y datos transformados con SNV. ....	98

**Tabla 7–3:** Matriz de confusión obtenida a partir del LDA con las especies *E. fusca*, *F. benjamina* y *T. catappa* considerando todos los sitios de muestreo y datos transformados con MSC. .... 99

**Tabla 7–4:** Matriz de confusión obtenida a partir del LDA con los espectros transformados MSC de las 3 especies seleccionadas considerando todos los sitios de muestreo. ....101

**Tabla 7–5:** Matriz de confusión obtenida a partir del LDA con los espectros sin transformar de las 3 especies seleccionadas considerando todos los sitios de muestreo. . ....102



## Lista de fotos

Pág.

<b>Foto 3-1:</b> Medición de espectros foliares. ....	54
<b>Foto 3-2:</b> Hoja empleada para la adquisición de espectros.....	55



# 1. Introducción

## 1.1 Planteamiento del problema y justificación

Los árboles proporcionan diversos beneficios a las ciudades y sus habitantes, mejoran la calidad del aire local y regional, secuestran carbono, son el hábitat de la fauna silvestre, y crean espacios estéticamente más agradables y saludables. Asimismo, reducen el ruido, el efecto de la isla de calor, la escorrentía de las aguas lluvias y el uso de energía para calefacción y enfriamiento (Galenieks, 2017).

Para garantizar la prestación de dichos servicios ecosistémicos por parte de los árboles en entornos urbanos, las administraciones gubernamentales deben realizar una caracterización de las especies existentes que incluya una identificación taxonómica georreferenciada, estado fenológico, fitopatología, edad, entre otros. Implica también, establecer un sistema de monitoreo en términos del estado fitosanitario y la interacción con las diferentes obras de infraestructura, así como, el desarrollo de un mecanismo que permita registrar las intervenciones que se lleven a cabo sobre los individuos, como tala y poda.

Ese nivel de detalle en la información, se constituye en un insumo para que las autoridades ambientales tengan criterios de decisión en los programas de control y manejo del arbolado urbano con respecto a la evaluación de la diversidad de especies y la cuantificación de los valores y servicios ecosistémicos prestados por los individuos arbóreos (Östberg *et al.*, 2013). Permite también, plantear estrategias de gestión a corto, mediano y largo plazo, teniendo el control sobre el mantenimiento, seguimiento y renovación de los árboles (Baró *et al.*, 2014).

En consecuencia, los organismos gubernamentales han empezado a implementar la ejecución de inventarios forestales urbanos, sin embargo, estos son costosos y topan con la dificultad al momento de la recolección, procesamiento y análisis de datos, asociado al conjunto de capacidades técnicas y profesionales especializadas del personal que participa en los inventarios, que normalmente no están disponibles dentro de las instituciones públicas.

Para subsanarlo, las administraciones de las ciudades han optado por contratar personas externas o hacer llamados a voluntarios para realizar el trabajo, esto puede generar problemas en la forma de tomar los datos y en la continuidad de la información, ya que el intervalo de tiempo en el cual se realiza es irregular y los protocolos de medición varían dependiendo de la institución y los criterios de las personas que lo ejecuten. Esta falta de calidad en la toma de datos como su periodicidad, incide en que la información levantada no sea comparable entre periodos de muestreo y pierda pertinencia para la toma de decisiones (Strunk *et al.*, 2016)

## 1.2 Antecedentes

En Colombia el árbol urbano hace poco tiempo consistía en un elemento secundario y no vinculante a los procesos de urbanismo. Hoy en día se cuenta con estudios detallados como los que se han ejecutado en el Valle de Aburrá, los cuales proporcionan instrumentos que integran los componentes para la intervención de los espacios verdes, guían el manejo del arbolado urbano y dan criterios para la planificación. Entre estos se destacan: Árboles ornamentales en el Valle de Aburrá. Elementos de Manejo (Morales & Varón, 2006), Manual de silvicultura urbana para Medellín (Secretaría de Medio Ambiente de Medellín, 2015), Árboles Nativos y Ciudad. Aportes para la silvicultura urbana de Medellín (Alcaldía de Medellín, 2011) y Guía para el manejo del arbolado urbano en el Valle de Aburrá (Área Metropolitana del Valle de Aburrá & Universidad Nacional de Colombia, 2015) (Tovar-Corzo, 2013).

Incluso en los últimos años diferentes ciudades en el país han ejecutado inventarios forestales, los cuales son gestionados a través de geovisores que permiten la administración y visualización de la información del arbolado urbano, como es el caso de Pereira con el portal geográfico SIGPER (disponible en <https://www.sigper.com/arboloadop/>) y SIGAU o Sistema de Información para la Gestión del Arbolado Urbano de la ciudad de Bogotá, administrado por el Jardín Botánico José Celestino Mutis (disponible en <http://www.jbb.gov.co/index.php/sigau>). En ambos portales la consulta de los árboles arroja su respectiva ficha la cual contiene información relacionada con la localización, nombre científico, revisión fitosanitaria, intervenciones, altura, entre otras.

Medellín no es la excepción, la ciudad cuenta con un aplicativo digital desde el año 2011 denominado Sistema de Árbol Urbano (SAU), el cual permite el ingreso, consulta y visualización de todas las intervenciones como podas y trasplantes, realizadas sobre los árboles ubicados en zonas urbanas. Sin embargo, las dificultades de la actualización de la información de esta herramienta y el

monitoreo de los individuos arbóreos plantean la necesidad de encontrar alternativas que permitan agilizar y optimizar el proceso que se viene realizando mediante la ejecución de inventarios forestales tradicionales, debido a que se evidencian errores en la georreferenciación e identificación de los árboles, no se cuenta con la información levantada o no es posible asociar el punto muestreado con el individuo arbóreo, como se muestra en la Figura 1–1, donde los puntos de la localización de algunos árboles en la Universidad Eafit no es posible asociarlos a su respectiva copa, debido a la agrupación de los individuos en este lugar.



**Figura 1–1:** Firma espectral característica de la vegetación sana.  
**Fuente:** Adaptado de Chuvieco (2008).

Por otra parte, los avances en la tecnología han facilitado el desarrollo de diferentes herramientas que permiten la obtención de información espacial de los ecosistemas forestales a partir del uso de la teledetección en aspectos relacionados con la composición, estructura, dinámica, perturbaciones y productividad del bosque; alcanzando a capturar información útil y de calidad a distancia, con la posibilidad de realizar monitoreos y seguimientos periódicos de

manera ágil en una escala adecuada, que posibilita la comprensión de los procesos evaluados. Durante las últimas décadas, los estudios centrados en la identificación de especies arbóreas han aumentado constantemente, sin embargo la brecha entre los resultados obtenidos y los inventarios forestales ejecutados siguen siendo uno de los mayores desafíos en este campo de investigación (Masek *et al.*, 2015).

La caracterización espectral de la vegetación presenta una amplia variedad de aplicaciones en la identificación de especies que responden a las necesidades identificadas desde el sector de la conservación y la administración de los bosques. Según Fassnacht *et al.*, (2016) las aplicaciones en las cuales se ha implementado la herramienta están relacionadas con la evaluación y monitoreo de la biodiversidad, identificación y mapeo del hábitat de la fauna silvestre, y aspectos contemplados en la medición de la productividad expresada en términos de biomasa y volumen en plantaciones y bosques naturales. Dichas aplicaciones mencionadas anteriormente se han centrado en bosques templados y ecosistemas boreales evidenciando la falta de estudios de este tipo en los bosques tropicales, siendo en la ejecución de inventarios de los recursos donde se ha implementado en mayor medida el uso de la respuesta espectral de las especies, inclusive en áreas urbanas en donde se ha presentado como una eficiente alternativa para el análisis rápido y efectivo de la composición e identificación de especies, proporcionando información fundamental para la gestión del arbolado urbano (Jensen *et al.*, 2012).

Numerosos estudios anteriores han permitido la creación de librerías o bibliotecas espectrales (Zomer *et al.*, 2009; Dudley *et al.*, 2015; Hovi *et al.*, 2017), las cuales son una herramienta potencial donde se puede acceder de forma gratuita a la recopilación de la caracterización espectral o firmas espectrales de diferentes coberturas y especies vegetales que facilitan el proceso de identificación automática. Aunque se tiene información para muchas especies forestales, todas estas corresponden a especies de ecosistemas de otras latitudes, para el trópico no se reporta la generación de firmas espectrales de las especies nativas. La construcción de la respuesta espectral típica de determinada especie tropical conformaría un primer acercamiento a la cuantificación de los recursos forestales, mediante un método no convencional.

Por esta razón, en esta investigación se plantea como reto la discriminación de especies forestales en entornos urbanos a partir de información espectral, en el caso específico de este proyecto las que se encuentran localizadas en diferentes Campus universitarios de la ciudad de Medellín.

### **1.3 Pregunta de investigación**

¿Es posible discriminar especies forestales en un contexto urbano utilizando datos espectrales obtenidos a partir de ortofotos disponibles para la zona de estudio y espectroscopía de reflectancia en el visible e infrarrojo cercano?

### **1.4 Hipótesis**

La identificación de especies forestales en contextos urbanos se puede llevar a cabo utilizando datos espectrales obtenidos a partir de ortofotos disponibles para la zona de estudio y espectroscopía de reflectancia en el rango de medición de 400 a 1000 nm, encontrando diferencias espectrales en las especies, generadas por su respuesta fisiológica en el visible (400 – 700 nm) y la estructura biofísica y contenido de agua en el infrarrojo cercano (700 - 1000 nm), además de las condiciones ambientales como la influencia a fuentes contaminantes (vías) que pueden estar afectando a las especies en dicho contexto.

### **1.5 Objetivo general**

Evaluar la variabilidad en la respuesta espectral para cada una de las especies forestales seleccionadas dentro del estudio, considerando sus diferencias intrínsecas y las variables ambientales que puedan afectarlas en un contexto urbano.

### **1.6 Objetivos específicos**

- Construir la curva espectral para las especies forestales seleccionadas a partir de las bandas rojo, verde, azul e infrarrojo del espectro electromagnético teniendo en cuenta la variabilidad de la fisiología y morfología de cada especie.
- Evaluar la capacidad de discriminación de las especies forestales seleccionadas a partir de las curvas de la respuesta espectral de cada una de estas.
- Desarrollar un análisis exploratorio del efecto de la contaminación del aire sobre la variabilidad de la respuesta espectral y su relación con la localización de las especies forestales seleccionadas.





## **2.Marco conceptual y revisión de literatura**

### **2.1 Conceptos y términos**

#### **2.1.1 Teledetección**

Chuvieco (2000) define la percepción remota o teledetección como la técnica que permite la adquisición de información de un objeto o un fenómeno mediante el análisis de datos registrados por un instrumento o sensor que no está en contacto físico con estos. En un contexto ambiental, la teledetección tiene como finalidad identificar y caracterizar los diferentes materiales de la superficie terrestre y los procesos que ocurren en ésta, a partir de la interacción energética que existe entre la tierra y el sensor, que constituye un flujo de radiación electromagnética (Chuvieco, 2008).

#### **2.1.2 Sensores remotos**

Los sensores remotos se clasifican en pasivos y activos. Los sensores pasivos se limitan a captar la energía de la luz solar que es reflejada y emitida por las coberturas de la tierra como consecuencia de sus características. Estos sensores pueden ser agrupados por la cantidad de bandas o intervalos de longitudes de onda del espectro electromagnético en los cuales es posible adquirir información, por tanto se denominan multiespectrales o hiperespectrales. Comúnmente la mayoría de sensores multiespectrales generan imágenes que presentan entre 10 y 12 bandas, caso contrario a las imágenes hiperespectrales que poseen bandas estrechas y continuas que pueden cubrir varias regiones del espectro (mayor a 12 bandas). En cuanto a los sensores activos, estos captan la energía emitida por ellos mismos, después que es reflejada por la superficie terrestre, como es el caso del radar (Khorram *et al.*, 2012).

Existen dos propiedades inherentes a los sensores remotos que definen la precisión en su aplicación como lo son la resolución espacial y espectral. La resolución espacial se refiere al tamaño de objeto más pequeño que puede ser identificado en una imagen o tamaño de pixel, mientras que la resolución

espectral indica el número y ancho de las bandas que puede discriminar el sensor (Ghosh *et al.*, 2014).

Los sensores multiespectrales como Landsat<sup>1</sup>, presentan bandas anchas poco continuas, generando para un amplio rango un valor promedio de la energía reflejada en esa región. Por lo tanto, las variaciones sutiles en la respuesta espectral que diferencian unas especies vegetales de otras se pierden o enmascaran, ya que estas se producen en regiones de longitud onda muy estrecha y específica. Como consecuencia, los sensores hiperespectrales son una buena alternativa para obtener información más detallada en el estudio de la vegetación, como la identificación de especies, a diferencia de los sensores multiespectrales que por sus características solo pueden permitir la discriminación de la vegetación de otras coberturas (Kumar *et al.*, 2002).

### **2.1.3 Espectroscopía**

La espectroscopía implica el estudio de las interrelaciones entre las características espectrales de los objetos y sus atributos biofísicos en el medio ambiente (Milton, 1987). Una vez que la radiación incide sobre la superficie pueden ocurrir diferentes procesos: se puede transmitir, reflejar o absorber.

Con el objetivo de poder comparar las medidas espectrales adquiridas de objetos en diferentes días y condiciones de iluminación, se utiliza la reflectancia, la cual es una medida independiente de las variaciones en la iluminación o esta puede ser calibrada para su modificación. La forma en que un objeto refleja o absorbe la energía conforma un comportamiento único que se denomina firma espectral, en donde la relación entre la reflectancia en función de la longitud de onda permite caracterizar e identificar superficies o materiales, como se busca en el presente trabajo de investigación en el que se emplea este concepto para discriminar especies forestales.

### **2.1.4 Respuesta espectral típica de la vegetación sana**

El espectro de reflectancia de las plantas adultas y sanas se caracteriza por presentar baja reflectancia en el azul (400 – 500 nm) y rojo (600 – 700 nm) y un incremento en el verde (500 – 600 nm), debido a la fuerte absorción de los

---

<sup>1</sup> Serie de satélites diseñados para la observación de la tierra a nivel mundial, puesto en órbita en el año 1972 y renombrado como Landsat en 1975. El proyecto fue generado y es administrado conjuntamente por la NASA y el Servicio Geológico de EE.UU. (USGS), adquiere imágenes multiespectrales de resolución moderada.

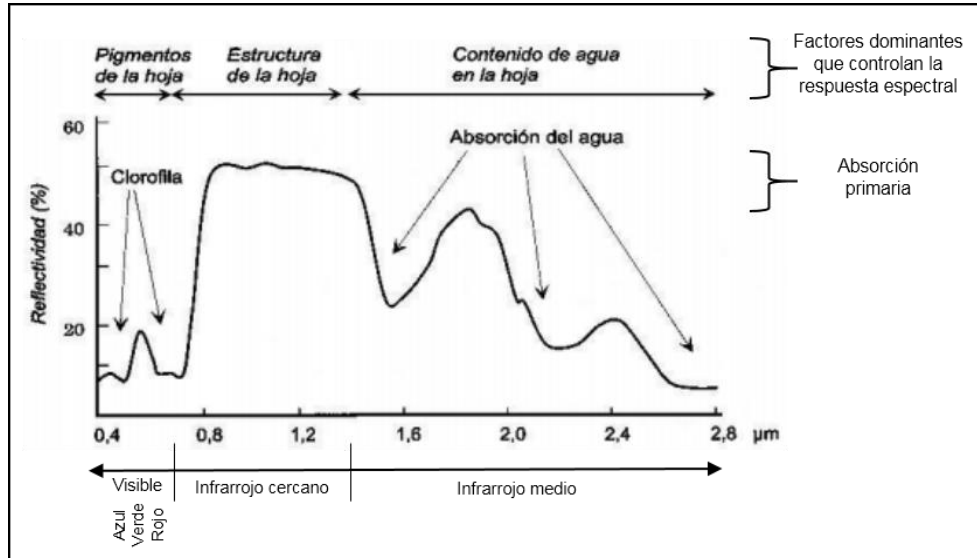
pigmentos foliares o fotosintéticos, los cuales tienen sus picos máximos de absorción en la luz azul y roja, como se evidencia en la Tabla 2–1, de ahí, que comúnmente las plantas sean de color verde. Cabe resaltar que, la clorofila domina la respuesta espectral ya que es de 5 a 10 veces más abundante este pigmento que los demás carotenoides. Como consecuencia, las plantas utilizan la luz visible para realizar el proceso de la fotosíntesis en donde convierten el dióxido de carbono y el agua en compuestos orgánicos para su mantenimiento y crecimiento (Suarez *et al.*, 2015)

**Tabla 2–1:** Pigmentos foliares y sus máximos de absorción.

<b>Tipo de pigmento</b>	<b>Longitud de onda máxima absorción (nm)</b>
Clorofila a	420, 490, 660
Clorofila b	435, 643
$\beta$ caroteno	425, 455, 480
$\alpha$ caroteno	420, 440, 470
Luteína	425, 450, 475

**Fuente:** Adaptado de Kumar *et al.*, (2002).

En comparación con la región de la luz visible, la vegetación presenta alta reflectancia en el infrarrojo cercano (700 – 1300 nm) determinada por el volumen de espacios aéreos intercelulares en la estructura interna de la hoja (mesófilo) y las discontinuidades entre las paredes celulares. Por otra parte, la reflectancia en el infrarrojo medio (1300 – 2500 nm) es menor que en el cercano, en esta región se dan fuertes absorciones de energía debido al agua contenida en la planta y en menor medida por otros contenidos bioquímicos foliares como lignina, celulosa, almidón y diferentes tipo de proteínas (Peñuelas & Filella, 1998). En la Figura 2–1, se presenta el comportamiento espectral típico de la vegetación sana y fotosintéticamente activa descrito anteriormente.



**Figura 2-1:** Firma espectral característica de la vegetación sana.  
**Fuente:** Adaptado de Chuvieco (2008).

### 2.1.5 Métodos estadísticos para el procesamiento de datos espectrales

El uso de técnicas espectrales para la discriminación de especies forestales han incrementado su potencial con el desarrollo de diferentes métodos estadísticos, que han permitido el procesamiento de información espectral, bien sea a partir de datos multispectrales o hiperspectrales. Esta evolución se ha visto facilitada por los avances en la capacidad computacional y la creación de algoritmos disponibles para su utilización, como las librerías de R (Lu & Weng, 2007).

El procesamiento y la manipulación inicial de los datos son fundamentales para llevar a cabo una clasificación correcta de las especies, usualmente este proceso contempla: detección y eliminación de valores atípicos, pre – tratamiento de los datos o aplicación de diferentes transformaciones, reducción de la dimensionalidad de las variables, y finalmente la aplicación de los métodos de clasificación. A continuación, se describe los métodos que fueron empleados en este estudio, con el fin de establecer en qué consisten y por qué fueron utilizados.

#### ▪ Detección y eliminación de valores atípicos

Hawkins (1980) define al valor atípico como “una observación que se desvía considerablemente de otras observaciones despertando sospechas de que fue generada por un mecanismo diferente”.

Los valores atípicos en datos espectrales se consideran multivariados, debido a que la medición se realiza en diferentes longitudes de onda y debe llevarse a cabo considerando toda la observación o espectro. La presencia de valores atípicos es el producto de una falla del equipo de medición, alteración en el proceso de adquisición de la respuesta espectral, o una combinación de ambas. Los llamados “outliers” o valores atípicos generan resultados y análisis equivocados, como consecuencia de un sesgo en la media de los datos, este fenómeno es denominado “enmascaramiento” y hace que las observaciones correctas se consideren erróneamente como atípicas, o viceversa, una observación atípica se tome como normal (Thennadil *et al.*, 2018).

En la actualidad existen numerosos métodos para la detección de valores atípicos, una cantidad considerable de dichos métodos se basan en la distancia y técnicas de visualización (LI *et al.*, 2016). En esta investigación se empleó para la identificación de valores atípicos en datos multiespectrales la distancia de Mahalanobis y para los datos hiperespectrales un Análisis de Componentes Principales (PCA).

La distancia de Mahalanobis es un concepto introducido en el año 1936 y ampliamente conocido, esta técnica clasifica un punto como un valor atípico basándose en la distribución de chi-cuadrado, a través de la medición de la distancia entre un punto y la distribución a la que pertenece ese punto (Filzmoser & Gschwandtner, 2017). Por otra parte, el Análisis de Componentes Principales (PCA) representa una de las técnicas más empleadas y útiles dentro de los métodos multivariados para la detección de valores atípicos, donde se utiliza la agrupación de los datos en su respectiva gráfica para determinar *a priori* la presencia de “outliers”, los cuales son seleccionados posteriormente con otros criterios estadísticos basados en la dispersión (Liu *et al.*, 2018).

#### ▪ **Pre – tratamiento de los datos espectrales**

El pre-tratamiento de los datos espectrales es la actividad más importante antes de llevar a cabo la aplicación del método de clasificación. Este proceso se realiza principalmente con el objetivo de filtrar el ruido ambiental presente en los espectros, y en ciertos casos eliminar el error sistemático humano y del equipo de medición. Las técnicas de pre-tratamiento más utilizadas en la espectroscopía se encuentran agrupadas en 2 categorías: métodos de corrección de dispersión y derivados espectrales (Rinnan *et al.*, 2009). En este estudio se aplicaron transformaciones que corresponden a los métodos de corrección.

La Corrección de Dispersión Multiplicativa (MSC) es la transformación más utilizada para el procesamiento de datos espectrales de la región del infrarrojo cercano, este concepto fue introducido por primera vez por Martens *et al.*, en el año 1983 y consiste en el cálculo de un espectro medio o de referencia para todo el conjunto de datos, y posteriormente el ajuste de cada espectro individual mediante la creación de una línea recta cuyos parámetros son coeficientes estimados para la corrección de las contribuciones aditivas y multiplicativas de la dispersión (Encina-Zelada *et al.*, 2017).

De otro lado, la Varianza normal estándar (SNV) es la segunda transformación más empleada para corregir los efectos de la dispersión de la luz en los espectros del infrarrojo cercano (Barnes *et al.*, 1989).. Este método simple se aplica para cada observación o espectro individual, al cual se le resta su promedio y se divide por su respectiva desviación estándar (Leandro *et al.*, 2019). Si se compara con MSC, esta transformación presenta una ventaja en su aplicación, dado que no requiere un espectro de referencia para realizar la corrección.

#### ▪ **Reducción de la dimensionalidad de las variables**

Una de las mayores limitaciones del procesamiento y manipulación de los datos espectrales está relacionada con la gran cantidad variables, además de la alta correlación espectral indicativa de la existencia de un problema de multicolinealidad entre dichas variables (Zhang *et al.*, 2017).

Por lo tanto, antes de realizar una clasificación de especies a partir de datos espectrales se requiere reducir la dimensionalidad de los datos. Para esto en esta investigación se utilizó el Análisis de Componentes Principales (PCA) debido a que es un método basado en la transformación ortogonal ampliamente empleado para identificar un número reducido de variables no correlacionadas que se denominan “componentes principales”, del conjunto de variables originales. En el PCA las primeras componentes principales abarcan la mayor varianza, además tienen una mayor señal/ruido, lo que indica que una buena elección de éstas elimina el problema de la dimensionalidad y puede contribuir a la corrección del ruido de los datos (Xu *et al.*, 2018).

#### ▪ **Método de clasificación**

Para la discriminación de especies existen una gran cantidad de métodos que se han revisado en detalle con el enfoque de clasificación a partir de datos espectrales (Lu & Weng, 2007). En este caso se utilizó el Análisis Discriminante Lineal (LDA) el cual es un método de clasificación supervisada aplicado con mayor frecuencia, cuyo objetivo se centra en encontrar un modelo que permita

maximizar la separación entre las diferentes clases. El LDA presenta algunas ventajas en comparación con otros métodos de clasificación, están relacionadas con la interpretación sencilla de las diferencias encontradas entre las clases y la aceptación de un número considerable de variables de entrada; asimismo se reporta en la literatura una desventaja asociada con la capacidad limitada cuando se tienen problemas de multicolinealidad entre las variables (Cortés *et al.*, 2019).

## **2.2 Antecedentes de la aplicación de la teledetección en el estudio de la vegetación**

La teledetección y los sensores remotos han demostrado ser una potencial herramienta para apoyar la gestión sostenible de los recursos forestales, aunque tradicionalmente los estudios realizados de los procesos ecológicos que se desarrollan en los ecosistemas boscosos han sido basados en trabajo de campo y manipulaciones experimentales, la teledetección se ha convertido en una valiosa fuente de información para el monitoreo y levantamiento de cartografía de los bosques en escalas locales, regionales y nacionales (Avtar *et al.*, 2017).

En el manejo forestal la aplicación de la percepción remota inició con la interpretación manual de fotografías aéreas. En la actualidad, en general son cuatro los atributos ecológicos claves del bosque en los que la teledetección ha desarrollado aplicaciones importantes, estos corresponden a: estructura, dinámica de las perturbaciones, productividad y composición. A continuación, se mencionan brevemente cada uno de estos aspectos, haciendo énfasis en el siguiente título en la composición e identificación de especies mediante la percepción remota, por ser el tema central de este estudio.

La estructura está relacionada con la organización de los árboles individuales en el paisaje, además, de la forma en que el dosel llena el espacio. Algunas de las variables estructurales que se pueden inferir a partir de la teledetección incluyen la altura, densidad del fuste, cubierta, distribución del dosel y biomasa aérea (Masek *et al.*, 2015), donde la demanda por la estimación de esta última se ha incrementado notablemente, ya que no requiere métodos destructivos ni invasivos, como la tala (Dittmann *et al.*, 2017). A partir del procesamiento de imágenes de muy alta resolución disponibles en Google Earth y del satélite IKONOS<sup>2</sup>, ha sido posible estimar con una precisión relativamente alta la biomasa

---

<sup>2</sup> Es un satélite comercial de imágenes de alta resolución para la observación de la tierra de la compañía DigitalGlobe. Fue lanzado en el año 1999 y tuvo operación hasta el 2015, capturando imágenes multiespectrales (rojo, verde, azul e infrarrojo cercano) con 4m de resolución espacial y

de bosques tropicales en comparación con los datos de referencia de biomasa aérea calculada a partir de un modelo alométrico regional (Ploton *et al.*, 2012).

Actualmente, la mayor exactitud en la estimación de las variables de la estructura de los bosques se presenta cuando se usa de manera sinérgica la tecnología LIDAR (Light Detection and Ranging) con sensores pasivos, tanto multiespectrales como hiperespectrales (Zolkos *et al.*, 2013). Sin embargo, la cobertura espacial y temporal es restringida, para esto, Pflugmacher *et al.*, (2012) lograron medir con precisión las perturbaciones en el tiempo en un bosque de coníferas mixto en el este de Oregón (Estados Unidos), a partir de los perfiles espectrales de las series temporales anuales de Landsat, con el objetivo de caracterizar los cambios en algunas variables estructurales de la vegetación a corto y largo plazo.

Los procesos que han generado disminución en el crecimiento o mortalidad en los árboles también han sido objeto de monitoreo por parte de la teledetección, ya sean por causas naturales como los incendios forestales, ataque de insectos, enfermedades, eventos de tormentas y/o procesos asociados a la acción antrópica como la deforestación y la introducción de especies invasoras. Lin *et al.*, (2016) han reportado exitosamente un modelo espacio-temporal que utiliza imágenes del sensor de cámara infrarroja del satélite HJ-1B<sup>3</sup> para detectar incendios forestales activos, donde el algoritmo propuesto se aplicó en bosques de China. Incluso, otros estudios han incorporado un mayor número de variables relacionadas con las perturbaciones del bosque, por ejemplo Guindon *et al.*, (2014) mediante imágenes de MODIS<sup>4</sup> (Moderate-Resolution Imaging Spectroradiometer) del año 2000 al 2011, desarrollaron un modelo para detectar cambios anuales producto de incendios, cosechas o inundaciones de los bosques boreales en Canadá.

---

pancromáticas con 1m de resolución (Recuperado de la base de datos de las misiones satelitales disponible en <https://directory.eoportal.org/web/eoportal/satellite-missions>).

<sup>3</sup> Corresponde al segundo satélite de la misión China “Hai Yang” para la observación del océano. Actualmente está activo y fue lanzado en el año 2007 con el objetivo de examinar el color del océano y temperatura de la superficie del mar, así como el monitoreo del medio ambiente (Recuperado de la base de datos de las misiones satelitales disponible en <https://directory.eoportal.org/web/eoportal/satellite-missions>).

<sup>4</sup> Es un sensor de la NASA que cuenta con 36 bandas espectrales y se encuentra a bordo de los satélites TERRA lanzado en 1999 y AQUA en 2002 para la observación de la tierra, con 20 años de adquisición de datos ha sido ampliamente utilizado para modelar el comportamiento de los incendios forestales a escala regional y global, considerando tanto el tamaño y la duración como la dirección de estos (Recuperado de la base de datos de las misiones satelitales disponible en <https://directory.eoportal.org/web/eoportal/satellite-missions>).



La productividad del bosque es medida en términos de Productividad Primaria Bruta, la cual se relaciona con la fijación de carbono llevada a cabo por la vegetación, en consecuencia, constituye un servicio ecosistémico al proporcionar biomasa para la generación de energía y capturar y almacenar carbono. Con el propósito de evaluar y monitorear el estado de los sumideros de carbono frente al aumento de las emisiones y el cambio climático, ha surgido una alternativa desde la teledetección en la cual satélites como GOSAT<sup>5</sup> y OCO-2<sup>6</sup> pueden medir la tasa fotosintética global y localizar las reacciones de estrés de las plantas en una resolución espacial hasta de 10,5 km, a partir de la medición de la fluorescencia de la clorofila, es decir, de la luz reemitida de la clorofila de la hoja durante el proceso de la fotosíntesis (Parazoo *et al.*, 2014).

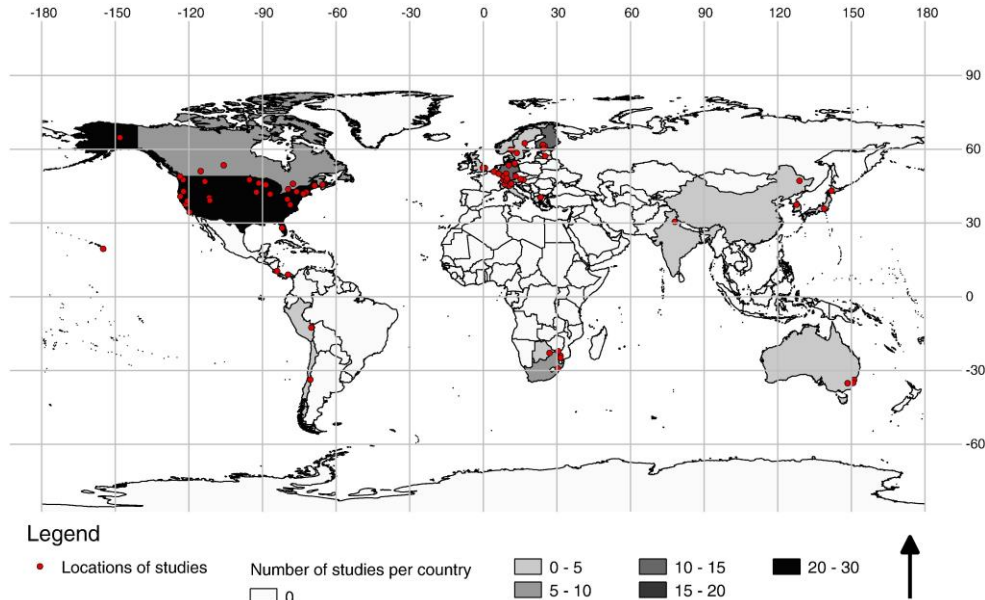
Por otra parte, el conocimiento sobre la composición de especies arbóreas de los bosques es información valiosa para todas aquellas personas que toman decisiones sobre estos, la correcta determinación taxonómica de las especies es una actividad clave para mantener alta calidad en la conservación y manejo forestal. En los últimos 35 años, se ha incrementado constantemente el número de estudios enfocados en la identificación de especies, la mayoría se han realizado en Norteamérica, en Estados Unidos y Canadá, y Europa especialmente en Alemania y los países Escandinavos (Figura 2–2).

Los ecosistemas objeto de análisis están representados principalmente por bosques templados, ecosistemas boreales y finalmente por bosques tropicales donde se presenta el mayor vacío de la información. En la literatura, se observa una tendencia reportada de utilizar en las investigaciones más de un tipo de sensor, en tal caso la combinación se lleva a cabo con un sistema activo y otro pasivo, predominan los ensayos realizados con datos hiperespectrales o espectroscopia de imágenes, en un segundo lugar se encuentran los sensores multiespectrales, que van desde moderada resolución como Landsat, hasta muy alta como IKONOS. ( Fassnacht *et al.*, 2016).

---

<sup>5</sup> Misión de la Agencia Japonesa de Exploración Aeroespacial (JAXA) lanzada en el año 2009 con el objetivo de estudiar el transporte de los gases de efecto invernadero y su aporte al cambio climático (Recuperado de la base de datos de las misiones satelitales disponible en <https://directory.eoportal.org/web/eoportal/satellite-missions>).

<sup>6</sup> Satélite de la NASA puesto en funcionamiento en el año 2014 con el propósito de mapear las fuentes y sumideros de dióxido de carbono en la tierra (Recuperado de la base de datos de las misiones satelitales disponible en <https://directory.eoportal.org/web/eoportal/satellite-missions>).



**Figura 2–2:** Mapa mundial de la localización de 101 estudios realizados sobre identificación de especies (Puntos rojos).  
**Fuente:** Tomado de Fassnacht *et al.*, (2016).

### 2.2.1 Aplicación de la espectroscopia en la discriminación de especies vegetales

En los bosques tropicales, las investigaciones que utilizan datos hiperespectrales se encuentran en los inicios de su desarrollo, los adelantos se han restringido debido a la falta de disponibilidad a este tipo de información, la complejidad de los bosques tropicales, además, del acceso limitado a los datos de campo para la validación de los resultados obtenidos (Sánchez-Azofeifa *et al.*, 2009).

A principios de la década de los 80's, con el desarrollo de la espectrometría de imágenes (Goetz, *et al.*, 1985), la percepción remota hiperespectral ha logrado posicionarse como una herramienta potencial que puede mejorar la eficiencia en la identificación de especies de plantas en comparación con la identificación convencional en el trabajo de campo (Adam *et al.*, 2010), mediante el uso combinado del análisis multivariado que incluye pre-procesamientos, los cuales disminuyen las perturbaciones sutiles y ocultas, al igual que suavizan los espectros para una mejor discriminación e identificación de estos. Además, de la aplicación de una o más técnicas multivariadas como Mínimos Cuadrados Parciales (PCS), Análisis de Componentes Principales (PCA) y Análisis Discriminante Lineal (LDA) (Wenz, 2018).

La clasificación de especies a partir de la espectroscopia, se ha enfocado en el uso de las características de la absorción de la energía realizada por las plantas

para discriminar las diferentes especies, identificando los cambios en la respuesta espectral en la región del visible (400 – 700 nm) y el infrarrojo cercano (700 – 1300 nm).

El estudio realizado por Clark *et al.*, (2005) en la Estación Biológica La Selva (Costa Rica), representa el primer trabajo publicado en el ecosistema de bosque tropical sobre la clasificación automática de 7 especies de árboles emergentes, considerando escalas de pixel, hoja y copa; mediante imágenes hiperespectrales de alta resolución espacial. En esta investigación se tomaron tanto espectros de reflectancia en laboratorio como con sensores aerotransportados, encontrando que la precisión y las regiones óptimas del espectro para la clasificación de especies varía con la escala. En cuanto a la precisión, ésta disminuye desde las escalas de hojas medidas en laboratorio hasta el nivel de pixel y copa medidas desde el sensor aerotransportado, y en general las bandas del infrarrojo cercano (700 – 1327 nm) fueron consistentes para la discriminación en todas las escalas consideradas. Concluyendo que este tipo de imágenes de alta resolución espectral y espacial tienen un potencial importante para la discriminación automatizada de las especies.

Otro estudio ejecutado en la Estación Biológica La Selva (Costa Rica), realizó la discriminación de 5 especies de árboles a escala de dosel, en el cual reportaron la dificultad en la identificación de especies debido a la variabilidad intra e interespecies de la reflectancia espectral para las especies de árboles seleccionadas (Zhang *et al.*, 2006). Por otra parte, Castro-Esau *et al.*, (2006) analizaron los factores específicos que controlan la variabilidad espectral, teniendo en cuenta no solo la variedad intra e interespecies de la reflectancia, sino también la variabilidad entre sitios y temporadas de muestreo, ya que el experimentó consideró 7 lugares en Costa Rica, Panamá y México dominados por bosque seco tropical, obteniendo resultados sin éxito en la clasificación de especies entre sitios y temporadas diferentes.

Recientemente Clark & Roberts (2012), exploraron la clasificación de 7 especies de árboles en bosque húmedo tropical con datos hiperespectrales adquiridos a escalas de tejidos (hojas y corteza) y dosel. Esta investigación se consideró fuera de lo convencional, primero por analizar los espectros de corteza, además de un análisis tradicional del espectro de las hojas. En segundo lugar, por utilizar un método clasificador diferente a los comúnmente empleados para la discriminación de especies en bosque tropical denominado “*Random Forests*”, el cual tiene la ventaja de ser un método no paramétrico, por tanto, obvia el principio de normalidad. Finalmente, en lugar de recurrir a las bandas espectrales como variables predictoras hicieron uso de algunas métricas hiperespectrales que

correspondían a las propiedades de la vegetación, ya que a partir de la región espectral del infrarrojo es posible detectar rasgos estructurales y compuestos foliares (Harrison *et al.*, 2018). Por ejemplo, el contenido de agua de la hoja y dosel (Ullah *et al.*, 2014), nitrógeno (Serbin *et al.*, 2012), lignina y celulosa (Meerdink *et al.*, 2016).

De esta manera y siguiendo la hipótesis que afirma que las diferencias en las propiedades químicas y estructurales crean firmas espectrales únicas y detectables para la clasificación de especies, Asner *et al.*, (2014) analizaron la reflectancia de hojas para 17 bosques a lo largo de un gradiente altitudinal y de tipo de suelo en el norte, centro y sur de Perú (Amazonas), vinculando los rasgos químicos al comportamiento espectral de las hojas (400 – 2500 nm). Esto se logró utilizando el método estadístico de Regresión parcial de mínimos cuadrados con el Análisis de la suma de cuadrados residual predicha (PLSR – PRESS). Como resultado, se obtuvo un alto porcentaje de precisión en la discriminación de especies, sin embargo, reportaron una gran variabilidad entre los rasgos químicos y espectrales de las especies. Donde los rasgos químicos que intervienen en la producción primaria se relacionaron con el cambio altitudinal, mientras que los nutrientes determinaron la variación espectral con respecto al cambio de fertilidad en el suelo.

La espectroscopia en la región visible e infrarrojo cercano también se ha venido implementado en hojas secas y madera, a diferencia de los estudios hasta ahora presentados, los cuales han sido ejecutados sobre hojas frescas. Actualmente, el Amazonas cuenta con un gran número de investigaciones en las que se han discriminado con éxito especies filogenéticamente muy relacionadas. Por ejemplo, Durgante *et al.*, (2013) identificaron 10 especies de los géneros *Eschweilera* y *Corythophora* de la familia *Lecythidaceae*, a partir de especímenes disponibles en herbarios de Brasil, mediante lecturas del infrarrojo cercano por transformación de Fourier (FT-NIR). Utilizando estos mismos tipos de lecturas, Lang *et al.*, (2015) identificaron 13 especies amazónicas de los géneros *Protium* y *Crepidosperrum* de la familia *Burseraceae*, a partir de muestras de hojas secas en diferentes etapas de desarrollo, juveniles y adultos. Incluso, Botelho (2017) aplicó la herramienta de espectroscopía para material seco de especies con muestras provenientes de varios sitios del Amazonas identificando especies con alta precisión.

Aunque la mayoría de estudios para la discriminación de plantas se basan en hojas frescas y secas, otras investigaciones han argumentado las ventajas de usar muestras de ramas o corteza, debido a que las hojas presentan mayor variabilidad en la respuesta espectral de acuerdo a la etapa de desarrollo en que

se encuentren, ya que influye en el grosor de la pared celular y la composición química. Por consiguiente, Lang *et al.*, (2017) realizaron una clasificación de especies amazónicas, mediante un Análisis Discriminante Lineal (LDA), con el fin de analizar la eficiencia y comparar los resultados obtenidos con los dos tipos de muestras, obteniendo buena precisión con ambos niveles de medición.

Por otra parte, Hadlich *et al.*, (2018) según los resultados obtenidos en su estudio proponen que la toma de datos hiperespectrales de la corteza en campo beneficiarían los inventarios forestales, gracias a la reducción de costos y mejoramiento en la precisión de la asignación de nombres botánicos a las especies. Al mismo tiempo que el trabajo de campo se llevaría a cabo de una manera más eficiente por el fácil acceso al tronco en comparación con las hojas, y la posibilidad de disminuir el número de muestras que deben trasladarse al herbario para su identificación.

Igualmente, la reflectancia del infrarrojo cercano ha permitido identificar especies maderables en peligro de extinción y contribuir al monitoreo de tala ilegal, como es el caso de los múltiples estudios reportados para caoba (*Swietenia macrophylla*), en los que se han unido esfuerzos para rastrear el origen de la madera en países como Bolivia, Brasil, Guatemala, México y Perú (Bergo *et al.*, 2016; Braga *et al.*, 2011; Silva *et al.*, 2018).

### **2.2.2 Aplicación de la espectroscopía en especies vegetales caso de estudio: Colombia**

En Colombia son pocos los estudios de espectroscopia realizados, sin embargo, los mayores esfuerzos se concentran en el sector agrícola del país, con énfasis en la nutrición de grandes extensiones cultivos. Un caso de estudio corresponde a CENICAÑA, donde el equipo de investigación desarrolló una metodología para la determinación de nitrógeno en tejido foliar de la caña mediante el uso de la reflectancia en la región del infrarrojo cercano (1100 – 2500 nm), obteniendo resultados exitosos dado su similitud con los reportes de los análisis de laboratorio (Larrahondo *et al.*, 2001).

Botero (2009), presentó en su trabajo de grado para optar al título de magíster un experimento en donde utilizó espectroscopía de reflectancia (400 – 1050 nm) con el fin de determinar las diferentes concentraciones nutricionales del tejido foliar en plantas de rábano cultivadas y muestras de hojas de banano provenientes de la zona bananera de Urabá, mediante la correlación de las variables medidas tanto fisiológicas como morfológicas con las características de la respuesta espectral. Con esta técnica fue posible estimar todas las variables consideradas (concentración de fósforo, potasio y calcio, porcentaje de humedad, etc.) en el estudio con alta precisión en la determinación, en la mayoría de los casos los resultados fueron iguales a los obtenidos con los métodos

tradicionalmente realizados en laboratorio. Por tanto, el autor concluye que esta alternativa es una herramienta potencial en explotaciones agropecuarias comerciales, ya que los resultados se obtienen en corto tiempo, con alta precisión, sin requerir análisis adicionales y a bajos costos.

Por otra parte, CENIPALMA con el objetivo de desarrollar modelos predictivos del comportamiento temporal de los fenómenos asociados a la sanidad vegetal, mediante el procesamiento de imágenes adquiridas por sensores remotos, desarrolló un algoritmo el cual relaciona la información adquirida por el sensor MODIS con la información de series históricas del ataque por insectos defoliadores en el año 2007. Para esto se caracterizaron los patrones de las curvas espectrales, obteniendo series temporales de índices de vegetación, que permiten relacionar eventos de defoliación en términos de niveles de defoliación. De esta forma, la respuesta espectral del infrarrojo cercano permitió monitorear la sanidad de cultivos de palma de aceite a través del tiempo y en escala regional (Araque & Jiménez, 2009)

Finalmente, en el campo de la identificación de especies vegetales Camacho (2014) en su tesis de maestría desarrolló una metodología en escala 1:10.000 que facilita y apoya el mapeo de comunidades de frailejón, mediante la generación de curvas espectrales y la clasificación automática de imágenes SPOT 5 utilizando redes neuronales. Cabe resaltar que, esta investigación llevada a cabo en el país es la más relacionada con el presente estudio, sin embargo, la discriminación de especie (frailejón) se realizó a nivel de comunidad agrupando áreas extensas de cubrimiento, de ahí el interés de ejecutar experimentos a nivel de especie.

## **3. Metodología**

### **3.1 Área de estudio**

El proyecto se llevó a cabo en el municipio de Medellín (Antioquia), en la subregión del Valle de Aburrá. Las autoridades ambientales urbanas de Medellín, tanto la Alcaldía como el Área Metropolitana del Valle de Aburrá, en los últimos años han venido implementando diferentes políticas orientadas a la siembra de árboles y conservación del arbolado urbano.

Como consecuencia, la ciudad cuenta con un inventario forestal detallado gestionado a través de un aplicativo de Sistemas de Información Geográfica denominado Sistema de Árbol Urbano (SAU) el cual permite la consulta y registro de los árboles urbanos, así que cada uno de ellos cuenta con información relacionada con las características, especie, registro fotográfico, intervenciones, medidas dasométricas, entre otras. Por lo tanto, el presente estudio contó con una base de datos inicial que permitió conocer con anticipación la diversidad de los árboles de la ciudad en términos de riqueza de especies y su abundancia relativa con el fin de realizar una pre – selección.

#### **3.1.1 Sitios de muestreo**

Se seleccionaron 4 sitios de muestreo porque la mayoría de éstos contaban con inventarios forestales detallados y actualizados al año 2016, en donde los árboles estaban mejor georreferenciados con un margen de error aproximadamente de 2 m, en comparación con el resto del arbolado de Medellín disponible en la plataforma del SAU, los cuales presentaban errores que superaban dicho valor para la mayoría de los individuos consultados. Por otra parte, fueron sitios seguros para llevar a cabo trabajo de campo y muestreos que incluyeron toma de datos mediante equipos que requieren algunas condiciones para su uso como el suministro constante de energía eléctrica. Los sitios correspondieron a los Campus de la Universidad Nacional de Colombia (Núcleo Río, Volador y

Robledo), Universidad de Antioquia, Universidad de Medellín y Universidad de Eafit (Figura 3–1).

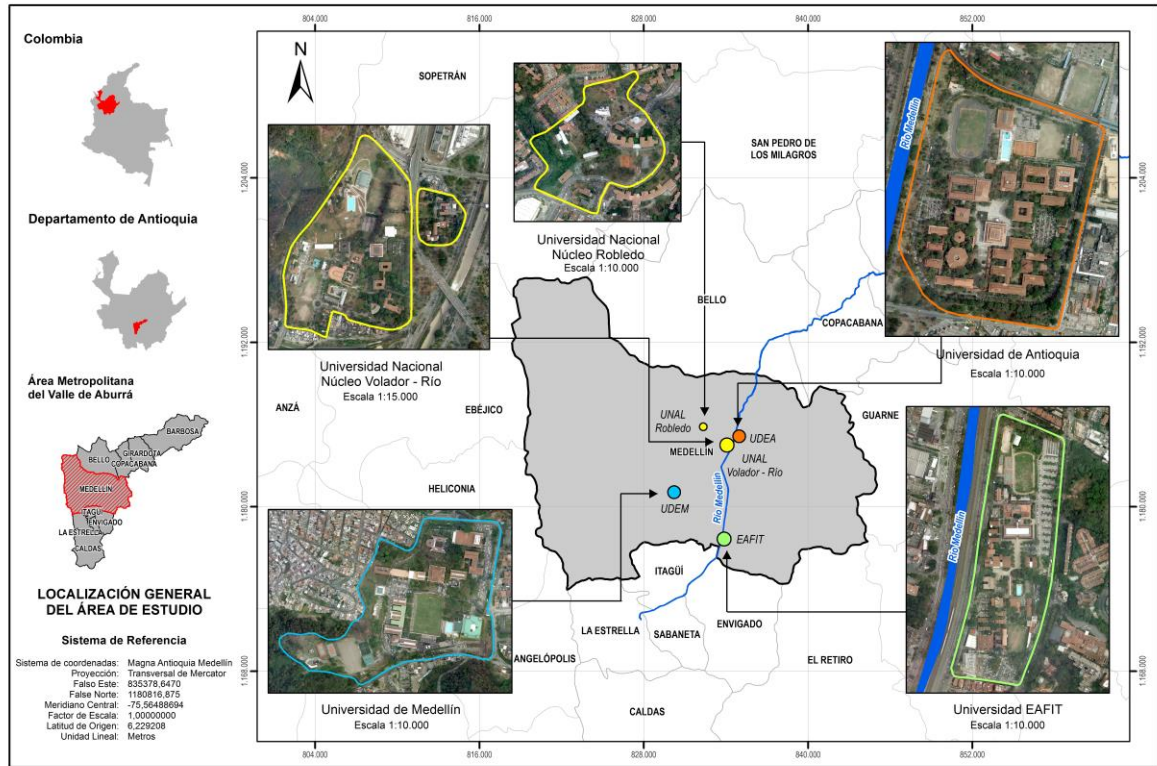


Figura 3–1: Localización general del área de estudio.

### 3.2 Selección de especies forestales

Se revisó conjuntamente la base de datos disponible en el geoportal SAU y el Catálogo Virtual de Flora del Valle de Aburrá de la Universidad EIA (2019), para identificar la diversidad de las especies que conforman el arbolado urbano de Medellín, teniendo en cuenta los inventarios forestales de los sitios de muestreo previamente seleccionados, con el fin de garantizar la presencia de las especies y una abundancia representativa de éstas en los Campus universitarios.

Además, se consideró la importancia ecológica y ambiental al momento de seleccionar las especies, asimismo características como la altura, diámetro, coloración de las hojas, permanencia del follaje durante el año, períodos de floración y fructificación. De esta manera, se aseguró la accesibilidad a los árboles determinada por sus dimensiones y la posibilidad de realizar diferentes muestreos en los tiempos establecidos en el estudio al encontrar material vegetal disponible para la toma de datos. En la Tabla 3–1, se presentan las 10 especies forestales seleccionadas para el desarrollo de la investigación, el nombre



científico y común, la abreviatura que se utilizó para su identificación, al igual que la familia botánica a la cual pertenecen.

**Tabla 3–1:** Listado de especies seleccionadas.

<b>Familia botánica</b>	<b>Nombre científico</b>	<b>Nombre común</b>	<b>Abreviatura</b>
<i>Boraginaceae</i>	<i>Cordia alliodora</i>	Nogal cafetero	CA
<i>Fabaceae</i>	<i>Erythrina fusca</i>	Búcaro	EF
<i>Moraceae</i>	<i>Ficus benjamina</i>	Laurel, Falso Laurel, Caucho	FB
<i>Fabaceae</i>	<i>Pithecellobium dulce</i>	Chiminango	PD
<i>Bignoniaceae</i>	<i>Spathodea campanulata</i>	Tulipán africano, Miona	SC
<i>Meliaceae</i>	<i>Swietenia macrophylla</i>	Caoba, Palo santo	SM
<i>Myrtaceae</i>	<i>Syzygium malaccense</i>	Pero de agua, Pomarrosa	SyM
<i>Polygonaceae</i>	<i>Triplaris americana</i>	Varasanta	TA
<i>Combretaceae</i>	<i>Terminalia catappa</i>	Almendro	TC
<i>Bignoniaceae</i>	<i>Tabebuia rosea</i>	Guayacán rosado	TR

### 3.3 Verificación de los inventarios forestales de los sitios de muestreo

Para este estudio se utilizó una imagen de alta resolución que cubre el Área Metropolitana del Valle de Aburrá, generada por el Instituto Geográfico Agustín Codazzi –IGAC- en el año 2010 para el Proyecto Cartografía de Antioquia “CARTOANTIOQUIA”<sup>7</sup>, la cual tiene una resolución espacial de 0,15 m y una resolución espectral de 4 bandas, representadas por las bandas del visible (rojo - B1, verde -B2 y azul – B3) y una en el infrarrojo cercano – B4, indispensable para realizar análisis con la vegetación.

Teniendo como base la imagen de alta resolución espacial se generaron mapas con la localización de los árboles establecidas en los inventarios forestales suministrados en los sitios de muestreo, con los cuales se hicieron visitas a los Campus y se verificaron todos los individuos pertenecientes a las especies seleccionadas. En caso tal que la universidad no contará con un inventario, a cada árbol se le asignó un código de identificación correspondiente a las iniciales del nombre de la universidad y la especie seguido por un número consecutivo.

<sup>7</sup> El Proyecto “CartoAntioquia” se llevó a cabo en el marco del Convenio interadministrativo 4085 de 2009 para la generación de la cartografía del departamento de Antioquia, firmado por Gobernación de Antioquia, IGAC, Alcaldía de Medellín, Área Metropolitana del Valle de Aburrá, Empresas Públicas de Medellín – EPM-, Instituto para el Desarrollo de Antioquia – IDEA – e ISAGEN.

Además, se realizaron metadatos con las observaciones generales pertinentes como por ejemplo el estado fitosanitario. Cuando el árbol no fue posible identificarlo en la imagen debido a su reciente fecha de siembra o presentó alguna condición en particular como enfermedad o muerte se excluyó del estudio. Finalmente, se obtuvo una base de datos consolidada de los árboles existentes pertenecientes a las especies forestales seleccionadas en los sitios de muestreo. En la Tabla 3–2, se presenta el número total de árboles inventariados (665) en cada uno de los sitios de muestreo, correspondientes a las 10 especies forestales seleccionadas para este estudio y en la Figura 3–2, Figura 3–3, Figura 3–4 y Figura 3–5 se muestra su respectiva localización en los Campus universitarios, producto del trabajo de campo.

**Tabla 3–2:** Número de árboles inventariados en los sitios de muestreo correspondientes a las especies seleccionadas.

Especie	SITIO				
	UNAL	UDEA	EAFIT	UDEM	TOTAL
<i>Cordia alliodora</i>	11	11	-	-	22
<i>Erythrina fusca</i>	86	16	17	24	143
<i>Ficus benjamina</i>	19	53	54	-	126
<i>Pithecellobium dulce</i>	37	19	1	9	66
<i>Spathodea campanulata</i>	34	31	8	18	91
<i>Swietenia macrophylla</i>	6	19	-	-	25
<i>Syzygium malaccense</i>	6	3	-	4	13
<i>Triplaris americana</i>	17	12	2	1	32
<i>Terminalia catappa</i>	40	13	4	17	74
<i>Tabebuia rosea</i>	18	14	40	1	73
<b>TOTAL</b>	274	191	126	74	665

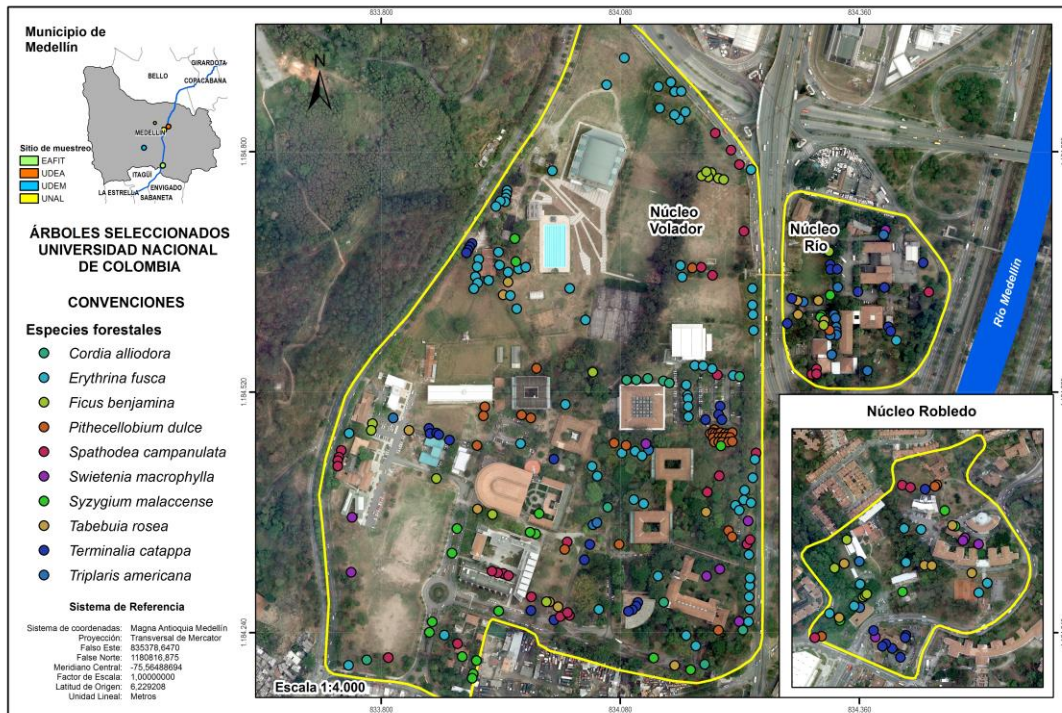


Figura 3–2: Localización de los árboles seleccionados en el Campus de la Universidad Nacional de Colombia.

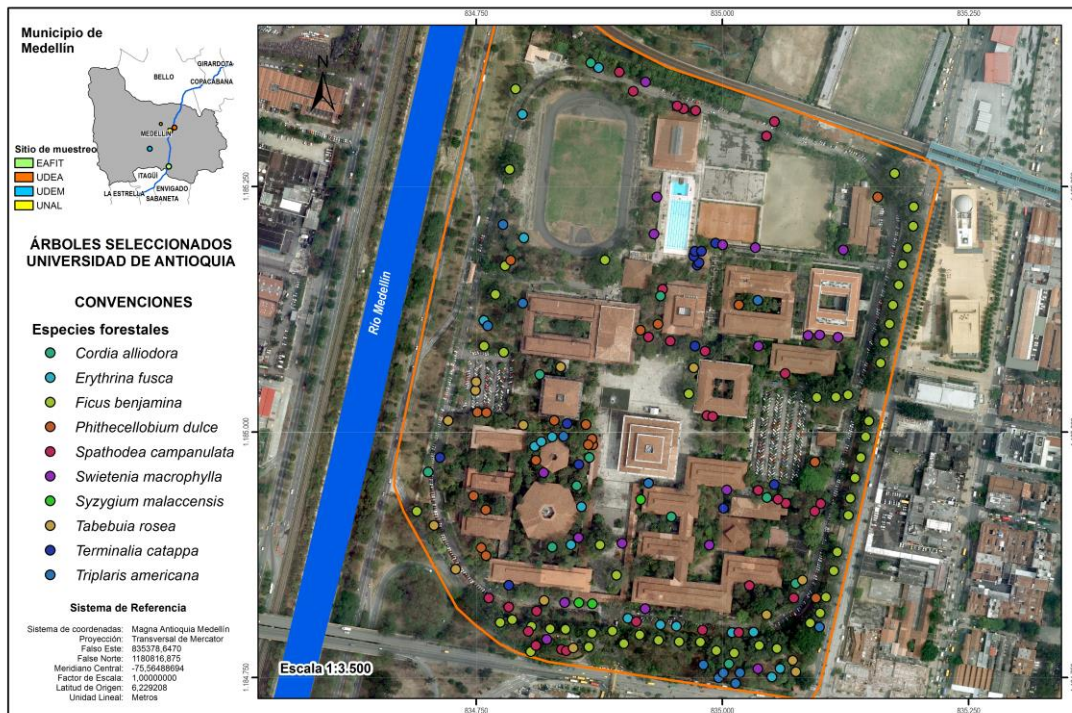
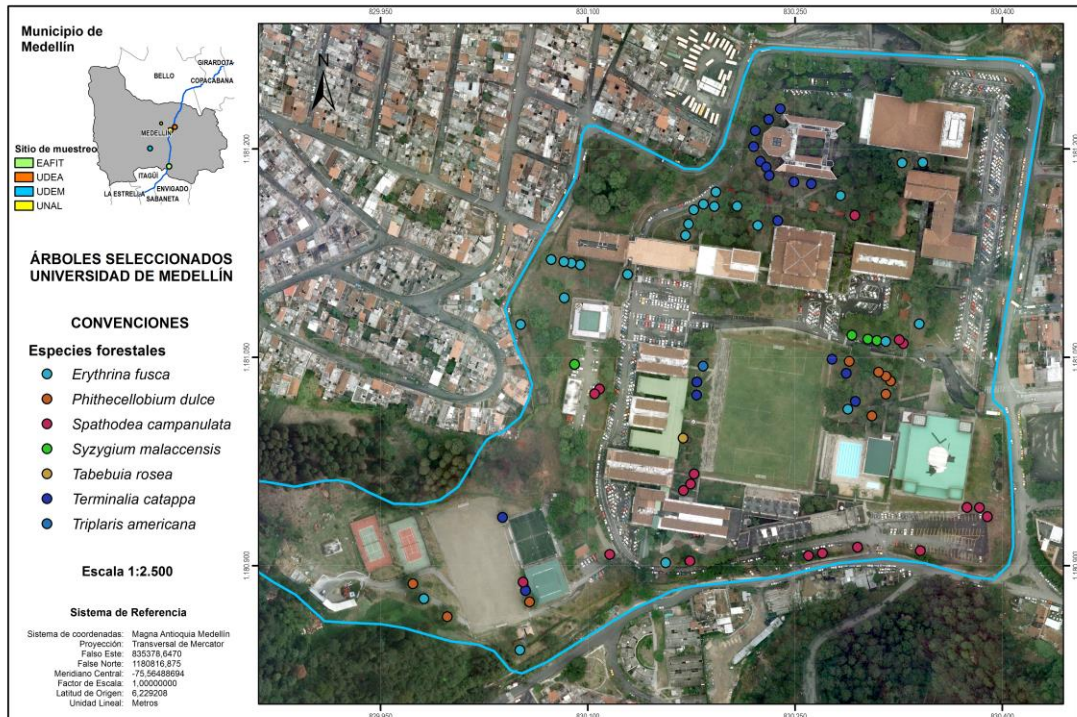
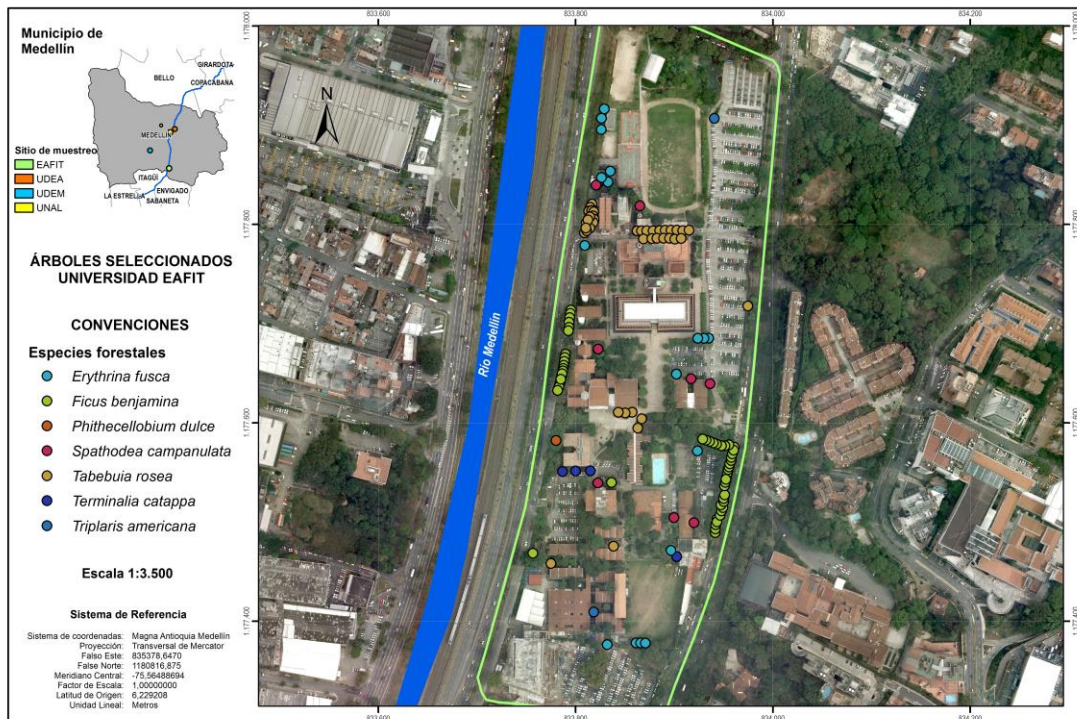


Figura 3–3: Localización de los árboles seleccionados en el Campus de la Universidad de Antioquia.





**Figura 3–4:** Localización de los árboles seleccionados en el Campus de la Universidad de Medellín.

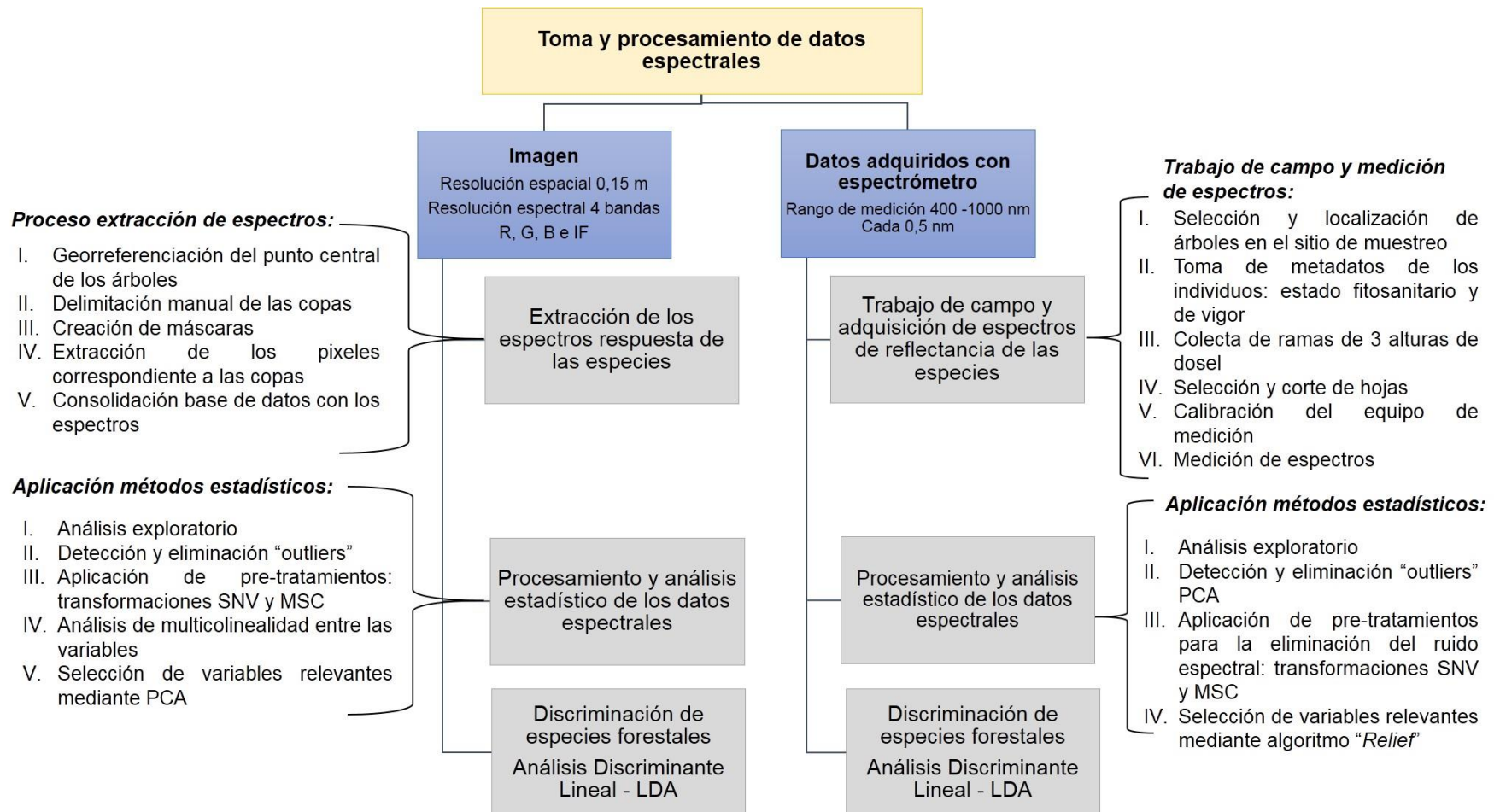


**Figura 3–5:** Localización de los árboles seleccionados en el Campus de la Universidad Eafit.

### **3.4 Toma y procesamiento de datos**

El objetivo principal de este trabajo de investigación fue evaluar la variabilidad de la respuesta espectral para la discriminación de especies, abordando el problema desde dos escalas diferentes, relacionadas con el origen y las características de los datos espectrales. La primera usando datos multiespectrales extraídos a nivel de copa de los árboles a partir de la imagen de alta resolución espacial pero de baja resolución espectral, y la segunda empleando los datos hiperespectrales tomados a nivel foliar con un espectrómetro portátil de cada una de las especies, considerando 3 alturas del dosel de los árboles. Este espectrómetro tiene un intervalo de medición de 200 a 1100 nm, el cual toma datos cada 0,5 nm.

En las secciones siguientes se detalla el proceso de adquisición, procesamiento y análisis estadístico de la información espectral utilizada y en la Figura 3–6, se presenta una síntesis de las actividades realizadas en este estudio, abordadas desde las dos escalas mencionadas anteriormente, es decir, tanto para los datos multiespectrales como hiperespectrales.



**Figura 3–6:** Actividades realizadas para la toma y procesamiento de datos espectrales.



### 3.4.1 Datos multiespectrales

- **Extracción y procesamiento de los datos**

Con base en el inventario forestal realizado para las 10 especies forestales seleccionadas en los sitios de muestreo, se generó un archivo (shapefile) con el punto georreferenciado central de cada uno de los 665 árboles identificados con su código ID correspondiente, y a partir del cual mediante el software ArcGIS 10.6 se delimitó manualmente su copa dejando por fuera 2 o más píxeles del límite para evitar el efecto borde que se puede generar, y de esta manera asegurar en cierta medida que los píxeles que fueron delimitados pertenecen en su gran mayoría a la copa del árbol y no al suelo u otras superficies (Figura 3–7).

Se hizo uso del módulo *Spatial Analyst Tools* de ArcGIS para extraer las áreas de la imagen pertenecientes a los polígonos de las copas de los árboles inventariados, mediante la herramienta *Extract By Mask*. Posteriormente, los archivos resultantes se convirtieron a formato texto a través de las herramientas de conversión del mismo software. Finalmente, por medio de un código elaborado en lenguaje *Python* se consolidó una base de datos con todos los árboles del estudio, que tuvo en cuenta: sitio de muestreo, ID del árbol, valor del píxel en las 4 bandas (B1, B2, B3 y B4), posición X y Y del píxel y la especie correspondiente. Resaltando que la copa de un determinado árbol está representada por una cantidad considerable de píxeles.



**Figura 3–7:** Copas delimitadas de los árboles objeto de estudio.

### ▪ **Análisis estadístico**

En primer lugar, se llevó a cabo un análisis exploratorio de los datos mediante la construcción de las gráficas de los espectros de las 10 especies forestales seleccionadas, empleando el promedio por árbol de los valores de los píxeles de cada una de las 4 bandas (rojo, verde, azul e infrarrojo). Este análisis previo permitió examinar visualmente la distribución de los datos y detectar de manera preliminar algunos valores extremos, los cuales fueron eliminados.

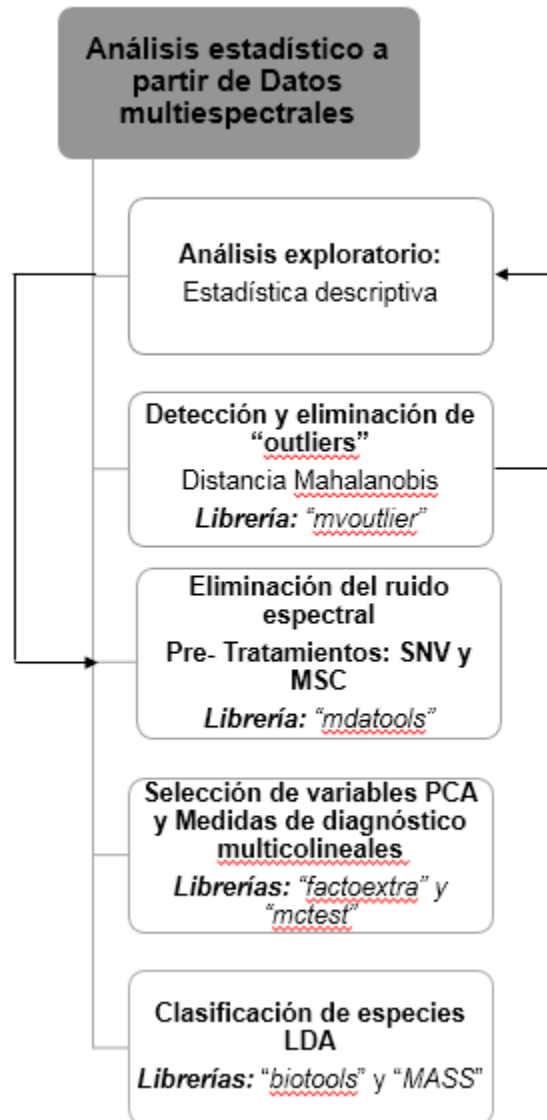
Posteriormente, se realizó la búsqueda de “outliers” por medio del Gráfico del cuartil ajustado disponible en la librería de R denominada “*mvoutlier*”, la cual contiene varios métodos para la detección de valores atípicos multivariados. En particular el método seleccionado se basa en las distancias de Mahalanobis y se representa gráficamente mediante la distribución de Chi – Cuadrado, mostrando por fuera del umbral los valores atípicos encontrados.

Una vez que se eliminaron los datos atípicos se aplicaron diferentes pretratamientos a los datos que consistieron en realizar las transformaciones: Variación normal estándar (SNV) y Corrección de dispersión multiplicativa (MSC) para minimizar el ruido y suavizar la respuesta espectral. Las transformaciones se hicieron sobre la base de datos completa, sin realizar agrupamientos por características como la especie o sitio de muestreo, mediante la librería de R llamada “*mdatools*”, ampliamente utilizada para el preprocesamiento, exploración y análisis de datos multivariados.

Luego, del pretratamiento de los datos se realizó el Análisis Discriminante Lineal (LDA) para evaluar la posible separación entre las 10 especies forestales consideradas, teniendo en cuenta que previamente se seleccionaron las bandas o variables claves que presentaron mayor influencia en la discriminación espectral de las especies, recogieron la mayor parte de la información o variabilidad de los datos y no fueron combinación lineal una de las otras. Este proceso se realizó para los datos multiespectrales agrupándolos conjuntamente por tipo de datos (datos crudos, transformados con SNV y MSC) y sitios de muestreo teniendo en cuenta las 10 especies seleccionadas y las especies con mejores resultados obtenidos a nivel grupal, utilizando de manera simultánea el Análisis de Componentes Principales (PCA) y su respectivo gráfico biplot, y la aplicación de medidas de diagnóstico multicolineales individuales que incluyeron: Varianza Factor de Inflación (VIF) y Método de Leamer disponibles en la librería “*mctest*” de R.

Cabe resaltar que, el LDA se ejecutó mediante la librería “*biotools*” y “*MASS*” en el software libre R versión 3.5.1.





**Figura 3–8:** Resumen del análisis estadístico realizado a partir de los datos multiespectrales.

- **Variabilidad de la respuesta espectral en las bandas (B1, B2, B3 y B4) en los diferentes sitios de muestreo**

Adicionalmente, se realizó un análisis de varianza de una vía (ANOVA) para evaluar la existencia de diferencias estadísticas de las medias entre las bandas, con el fin de evidenciar la variabilidad de la respuesta espectral media de las especies forestales en los sitios de muestreo. Se probó la hipótesis nula ( $H_0$ ) de no haber diferencias significativas entre las bandas espectrales en los sitios de muestreo, es decir, medias iguales, y como hipótesis alternativa ( $H_a$ ) la existencia

de diferencias significativas o medias diferentes ( $H_0: B_1=B_2=B_3=B_4$  y  $H_a: B_1 \neq B_2 \neq B_3 \neq B_4$ ).

Una vez que se determinaron la existencia de diferencias en las medias mediante una ANOVA estadísticamente significativa ( $p < 0,05$ ), se realizaron pruebas post hoc de Tukey (HSD) para establecer qué pares de grupos fueron diferentes entre sí, teniendo en cuenta las posibles combinaciones entre bandas (B1, B2, B3 y B4) por sitio de muestreo (UNAL, UDEA, EAFIT y UDEM) de cada una de las especies seleccionadas. Por ejemplo: B1 UNAL vs. B1 UDEA, B1 UNAL vs. B1 EAFIT, B1 UNAL vs. B1 UDEM, B1 UDEA vs. B1 EAFIT, B1 UDEA vs. B1 UDEM y B1 EAFIT vs. B1 UDEM.

Las ANOVAS se llevaron a cabo en el software libre para el análisis de datos científicos PAST (PAleontological STatistics) Versión 3.25.

### 3.4.2 Datos hiperespectrales

- **Equipos de medición y software de adquisición de datos**

Se midió la respuesta espectral de las especies vegetales seleccionadas en términos de reflectancia (VIS/NIR) mediante un espectrómetro portátil Ocean Optics USB2000 + que cuenta con intervalo de medición de 200 a 1100 nm en el cual captura datos cada 0,5 nm, es alimentado por una fuente de luz halógena de tungsteno (HL-2000-HP) que proporciona la capacidad de tomar datos en el visible e infrarrojo cercano en un rango de longitud de onda de 360 a 2400 nm, utilizando una sonda de reflexión de grado premium con una fibra de 600  $\mu\text{m}$  (QR600-7-VIS-125 °F) y un rango de medición de 400 a 2100 nm, al igual que un estándar de reflectancia difusa (WS-1) que presenta una reflectividad mayor al 98% en el rango de 250 a 1500 nm. Los espectros fueron adquiridos con el software *OceanView* el cual permite la configuración manual de los parámetros de adquisición y representa gráficamente los resultados obtenidos, esta aplicación fue desarrollada por la empresa *Ocean Optics*, así como los equipos mencionados anteriormente (Figura 3–9).



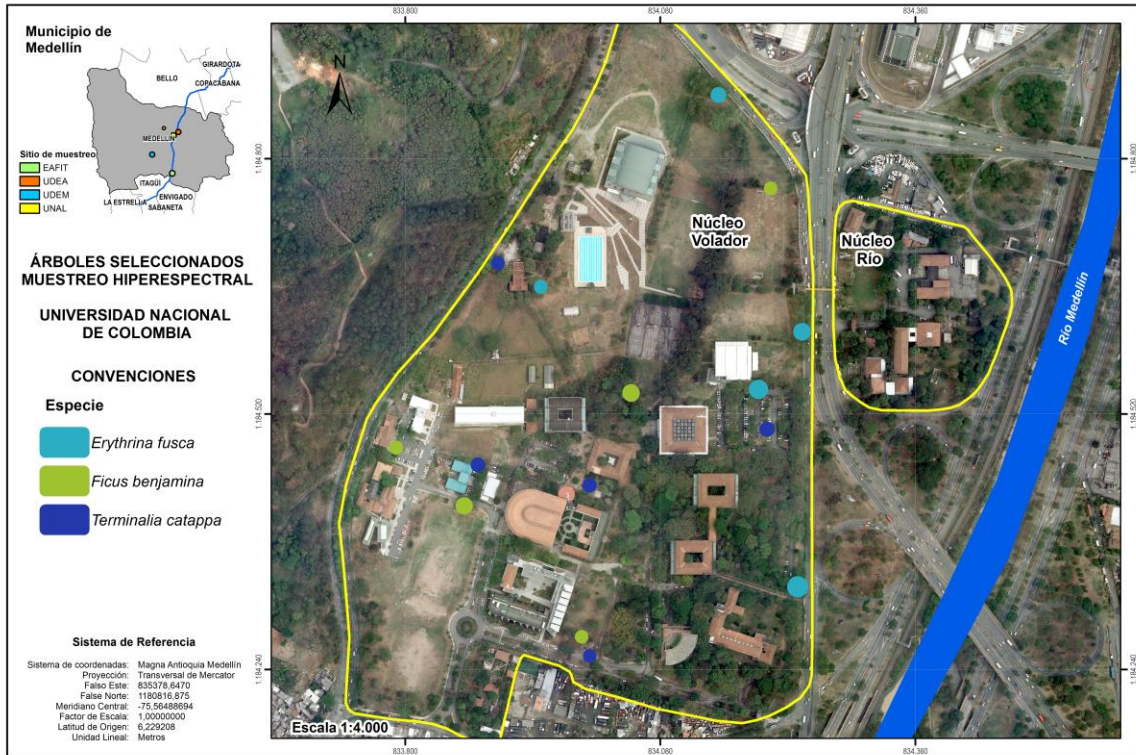
**Figura 3–9:** Equipos de medición empleados.

**Fuente:** Página oficial Ocean Optics disponible en <https://oceanoptics.com/>.

#### ▪ Trabajo de campo y toma de muestras

Se seleccionaron en promedio 5 árboles adultos al azar de cada especie en los sitios de muestreo con base en los 665 árboles identificados en el inventario forestal realizado previamente, y se visualizaron en la imagen de alta resolución espacial con el fin de hacer un primer reconocimiento. Cabe resaltar que, durante este proceso de selección se tuvo en cuenta la posición de los árboles dentro del Campus con respecto a las vías de tránsito aledañas, con el objetivo de abarcar la variación que se podría presentar en la respuesta espectral producto de una mayor o menor exposición a fuentes contaminantes, como lo son en este caso los vehículos y motocicletas. Por tanto, los individuos objeto de estudio se distribuyeron en el exterior e interior del Campus, como se muestra a manera de ejemplo en la Figura 3–10, en donde se presentan la localización de los árboles seleccionados para el muestreo en la Universidad Nacional de Colombia correspondientes a las especies *E. fusca*, *F. benjamina* y *T. catappa*.

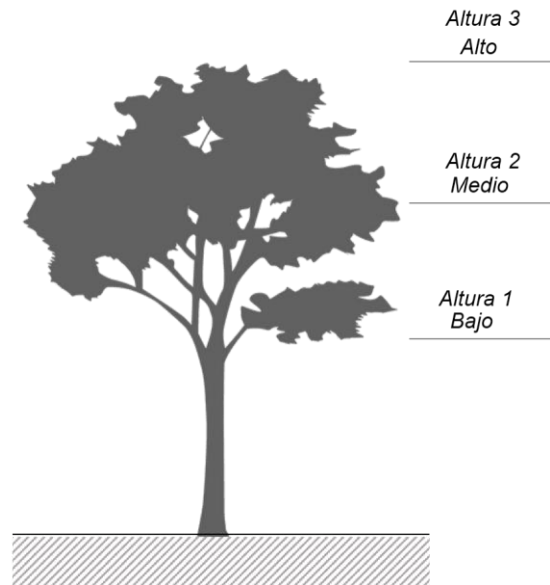
Para la adquisición de espectros se instalaron los equipos de medición en cada sitio de muestreo, en donde se estabilizó la fuente de luz durante 15 minutos dejándola encendida. Posteriormente, se calibró el espectrómetro con ayuda del estándar de reflectancia difusa (WS-1).



**Figura 3–10:** Localización de los árboles seleccionados para el muestreo hiperespectral en la Universidad Nacional de Colombia de las especies *E. fusca*, *F. benjamina* y *T. catappa*.

Luego, mediante el código de identificación asignado se localizaron cada uno de los árboles en campo y se tomaron datos como fecha, hora y condiciones climáticas del muestreo, al igual que observaciones que incluyeron las notas o metadatos referentes a cada individuo muestreado con respecto a su entorno, estado fitosanitario y de vigorosidad. Resaltando que los muestreos y trabajo de campo se realizaron del 6 al 11 agosto del año 2018, tiempo en el cual se presentaron días muy soleados y sin eventos de precipitación.

Las mediciones se llevaron a cabo en 3 niveles altitudinales de la copa: alto – Altura 3, medio – Altura 2 y bajo – Altura 1; como se muestra en la Figura 3–11. El número de niveles altitudinales se determinaron en función de los estándares forestales para la caracterización de doseles reportados en investigaciones anteriores, cuyos informes coinciden en la importancia de considerar la variación de las características en la estratificación vertical (Wilkes *et al.*, 2016).



**Figura 3–11:** Niveles altitudinales de medición.

Por tanto, para estos 3 niveles se colectó una rama de la parte más exterior de la copa con ayuda de una tijera corta ramas de varias extensiones, las cuales a su vez sirvieron para estimar las alturas de colecta. Cabe resaltar que, en la mayoría de los casos debido a la altura total del árbol la medición no se logró realizar en el nivel “Alto” sino a una altura menor.

El proceso de la colecta se hizo desde el nivel “Bajo – Altura 1” hasta el “Alto– Altura 3” marcando la rama con una cinta de papel identificada con el código del árbol correspondiente, luego se llevó a cabo su transporte hasta el sitio de localización del espectrómetro, teniendo en cuenta que no transcurrieran más de 30 minutos entre la recolección de la hoja y la medición. Autores como Khavaninzadeh *et al.*, (2015), incluso han reportado que la reflectancia de la hoja permanece estable hasta 4 horas después de la cosecha del árbol.

En cada uno de los niveles altitudinales se cortaron 5 hojas evidentemente sanas con unas tijeras podadoras y se realizaron las mediciones mediante la fibra óptica en el lado derecho del haz de la hoja, obteniendo 3 espectros por cada una de estas, separados por una distancia aproximada de 0,5 a 1 cm. Se tomó el registro en el centro de la hoja evitando las nervaduras y pigmentaciones anómalas en las estructuras vegetales. De esta manera a cada árbol se le midieron 45 espectros en total, por tanto, cada especie contó con 225 espectros en cada sitio de muestreo (Foto 3-1).

Por otra parte, debido a la baja relación señal/ruido para las longitudes de onda después de 1000 nm y antes de 400 nm, los espectros se recortaron a un rango



de 400 – 1000 nm para su posterior análisis, en este rango del espectro el equipo midió la reflectancia en las hojas cada 0,5 nm aproximadamente. El tiempo de integración utilizado en este estudio fue de 3 ms, un promedio de 10 lecturas por medición y un valor de 5 para el “*Boxcar width*”, el cual promedia para cada medición 5 datos a la derecha e izquierda y de esta manera suaviza la curva espectral final.



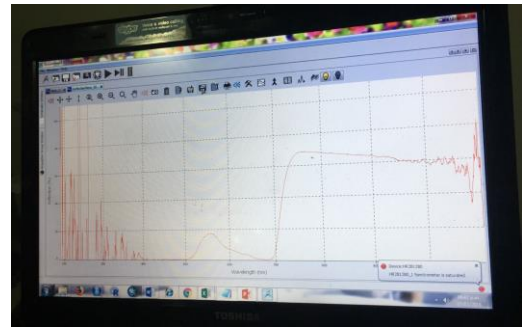
A. Toma de muestras



B. Ramas colectadas para medición



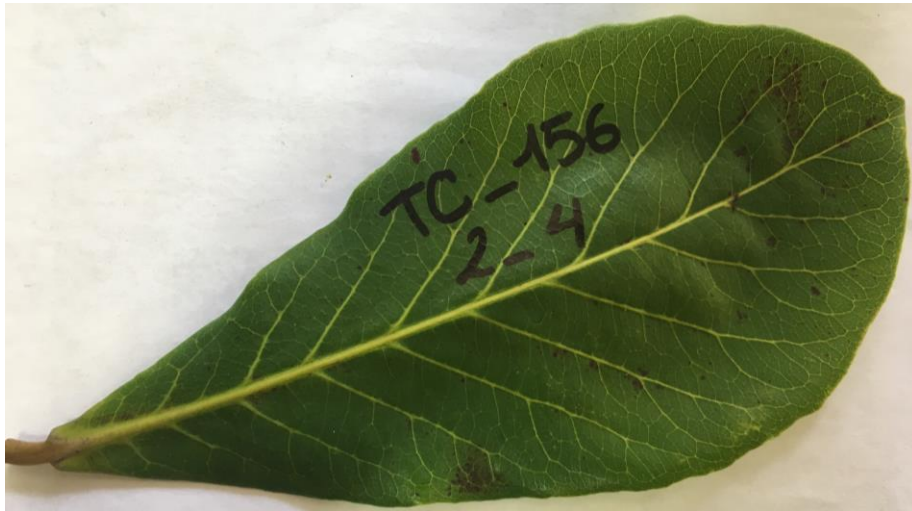
C. Hojas empleadas en la medición



D. Respuesta espectral de la vegetación

**Foto 3-1:** Medición de espectros foliares.

Finalmente, cada hoja fue marcada en el lado izquierdo del haz con las iniciales de la especie, código de identificación asignado en el inventario, nivel altitudinal (1- Bajo, 2 – Medio y 3 – Alto) y número de hoja (1 al 5) separados por un guion bajo como se muestra en el ejemplo: TC\_156\_2\_4, el cual hace referencia a la especie *T. catappa*, árbol 156, Altura 2 o Bajo de la hoja número 4. Lo anterior se empleó para la identificación de la muestra tanto de su fotografía como el archivo resultante de la adquisición del espectro (Foto 3-2).



**Foto 3-2:** Hoja empleada para la adquisición de espectros.

#### ▪ **Análisis estadístico**

Inicialmente se identificaron los espectros con comportamientos anómalos por medio de la visualización de su respectiva gráfica para cada una de las especies forestales consideradas en el estudio, en este caso los espectros se construyeron agrupándolos por sitio, con el fin de abarcar en el análisis exploratorio la variabilidad que los datos pudieran presentar y llevar a cabo un reconocimiento previo de la distribución de estos.

Después, se realizó la detección y eliminación de valores atípicos con la técnica de Análisis de Componentes Principales (PCA), los cuales se hicieron por sitios (UNAL, EAFIT y UDEM), especie y altura (Altura 1, Altura 2 y Altura 3), considerando todas las posibles combinaciones que surgieron de dichas características de muestreo. Posteriormente, con la base de datos consolidada se aplicaron 2 métodos para la corrección de la línea base de datos espectrales que incluyeron las transformaciones Variación normal estándar (SNV) y Corrección de dispersión multiplicativa (MSC), con la finalidad de eliminar los efectos de dispersión aditiva y multiplicativa de los espectros. Este proceso se desarrolló de igual manera en los datos multiespectrales, por tanto, la librería y software empleados fueron los mismos (*'mdatools'* disponible en R).

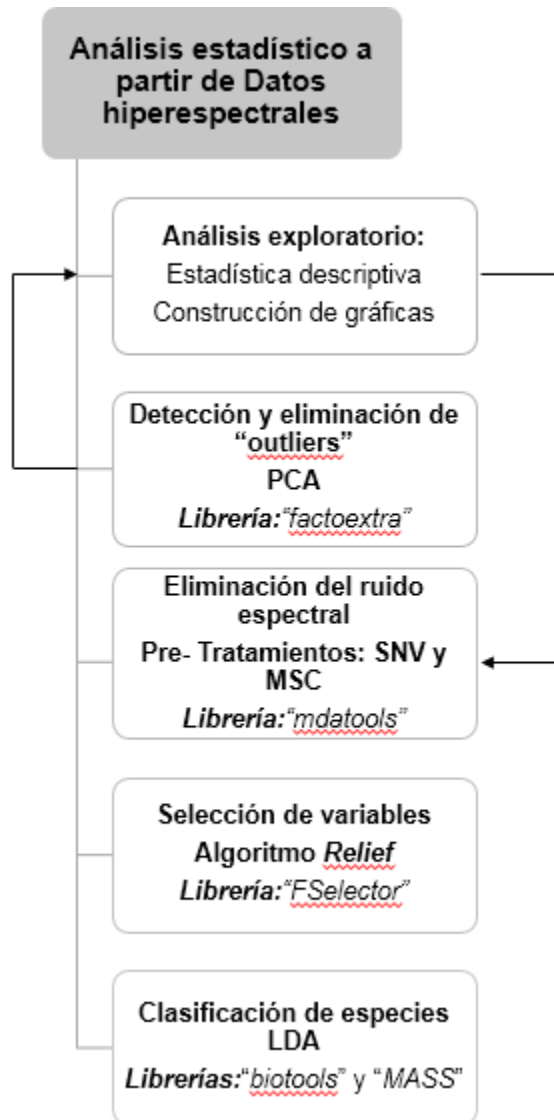
Por otra parte, los datos hiperespectrales tenían 1363 longitudes de onda (desde 400 nm hasta 1000 nm, tomados cada 0,5 nm), por tanto, se utilizó un método para identificar y eliminar la mayor cantidad de información irrelevante y redundante que fuera posible, calificando la importancia de las variables en la clasificación del resultado, en este caso de la discriminación de las especies

forestales. El método de filtro utilizado fue el algoritmo Relief disponible en la librería “*FSelector*” de R, el cual se aplicó para todas las longitudes de onda en intervalos de 10 nm para los datos crudos y transformados con SNV y MSC, considerando las 3 especies simultáneamente, y de a 2 especies en los múltiples arreglos que se generaron. Las longitudes de onda seleccionadas correspondieron a todas aquellas variables que obtuvieron un valor de peso de importancia  $> 0,05$  (Kira & Aai, 1992).

Finalmente, se llevó a cabo el Análisis Discriminante Lineal (LDA) para la posible identificación de las especies forestales, teniendo en cuenta las variables o longitudes de onda seleccionadas previamente, y las bases de datos establecidas para la aplicación del algoritmo Relief, es decir, los diferentes tipos de datos considerando todas las especies de manera conjunta y las combinaciones de a 2 especies que pudieron resultar. Adicionalmente, los LDA se realizaron por sitios de muestreo, con el objetivo de evaluar el comportamiento del método de clasificación frente a la variabilidad determinada por diferentes condiciones ambientales e intrínsecas de las hojas de los árboles muestreados. Esto se implementó en R usando los paquetes “*biotools*” y “*MASS*”.

En la Figura 3–12, se presenta un resumen del análisis estadístico que se llevó a cabo a partir de los datos hiperespectrales, en el cual se observa tanto el procedimiento realizado como la librería del software R empleada entre comillas y letra cursiva.





**Figura 3–12:** Resumen del análisis estadístico realizado a partir de los datos hiperespectrales.

#### ▪ Impacto de la estratificación del dosel en la respuesta espectral

Se llevó a cabo el test de Kruskal - Wallis como alternativa no paramétrica al análisis de varianza de una vía (ANOVA) para verificar la existencia de diferencias estadísticas entre la reflectancia espectral media de los 3 niveles altitudinales, para los datos crudos agrupados por sitios de muestreo y especie. Se tomó el nivel altitudinal de la copa como un factor fijo y se probó la hipótesis nula ( $H_0$ ) que no había diferencias significativas en la respuesta espectral de las hojas en los estratos muestreados, por tanto, las medias serían iguales, y como

hipótesis alternativa ( $H_a$ ) estas serían diferentes ( $H_o: A1 = A2 = A3$  y  $H_a: A1 \neq A2 \neq A3$ ).

Cuando el resultado del test fue estadísticamente significativo ( $p < 0,05$ ) se realizaron pruebas post hoc de Mann - Whitney por el método Holm por ser el más recomendado, con el fin de resaltar qué grupos específicamente fueron diferentes, es decir, encontrar las diferencias por pares en la reflectancia espectral para todas las combinaciones de pares posibles de los 3 niveles altitudinales de la copa, como por ejemplo: Altura 1 vs. Altura 3, Altura 1 vs. Altura 2, Altura 2 vs. Altura 3 (Carvalho *et al.*, 2013).

Todos los análisis estadísticos descritos anteriormente se realizaron utilizando R 3.5.1 para Windows.

- **Análisis de la relación entre la variabilidad espectral y la localización de los árboles**

Para analizar el efecto de la localización de los árboles y la posible influencia de la contaminación del aire sobre la reflectancia espectral, se realizó una clasificación de los individuos muestreados de acuerdo con la proximidad a las vías de tránsito más cercanas y las características de dichas vías relacionadas con el flujo vehicular.

La calificación de las distancias de las vías se llevó a cabo midiendo cada uno de los árboles considerados en el muestreo con respecto a las vías de tránsito circundantes, dejando como medida final la menor distancia que se midió, que en este caso corresponde al escenario más crítico. Una vez tomadas todas las distancias se hizo una clasificación por medio de una distribución de frecuencias en 3 clases: A, B Y C; estandarizando dichos rangos para todos los sitios de muestreo. En donde la letra A corresponde a los árboles más cercanos a la vía (0 a 30 m), la letra B a los intermedios (30 – 60 m) y la letra C para los más alejados (> 60 m).

Por otra parte, se identificaron las vías adyacentes a cada uno de los Campus universitarios y se clasificaron de acuerdo con el volumen y tipo de vehículos que usualmente transitan por estas vías. Adicionalmente, se revisó la normatividad establecida por la Secretaría de Movilidad de la Alcaldía de Medellín mediante la Resolución 2585 de 2016 “Por medio de la cual se actualiza la clasificación y prelación de las vías urbanas de la ciudad de Medellín” con el objetivo de validar la clasificación de las vías realizadas previamente. Las vías se calificaron con un número del 1 al 4, en donde el 1 correspondió a las vías con mayor volumen de vehículos que incluyeron carros particulares, motos, camiones y volquetas; y 4

aquellas vías con poco flujo vehicular representado por carros y motos (Tabla 3–3). En la Figura 3–13, Figura 3–14 y Figura 3–15 se presenta la clasificación de los árboles en la Universidad Nacional de Colombia, Universidad Eafit y Universidad de Medellín, respectivamente, de acuerdo a los criterios descritos anteriormente.

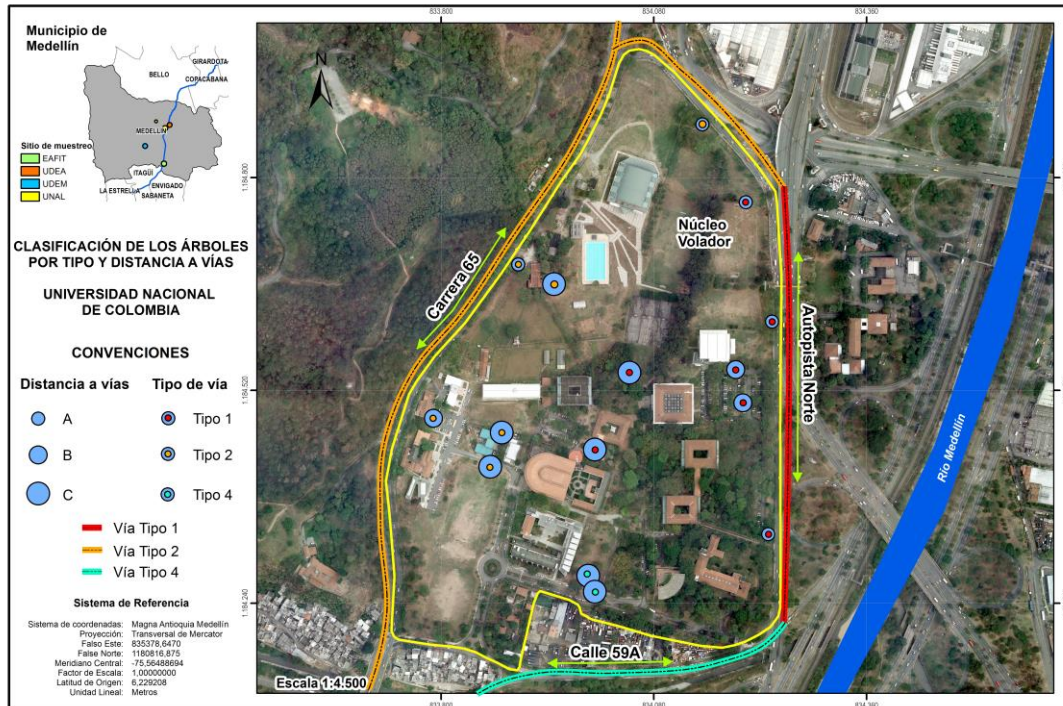


Figura 3–13: Clasificación de los árboles por tipo y distancia a las vías en la UNAL.



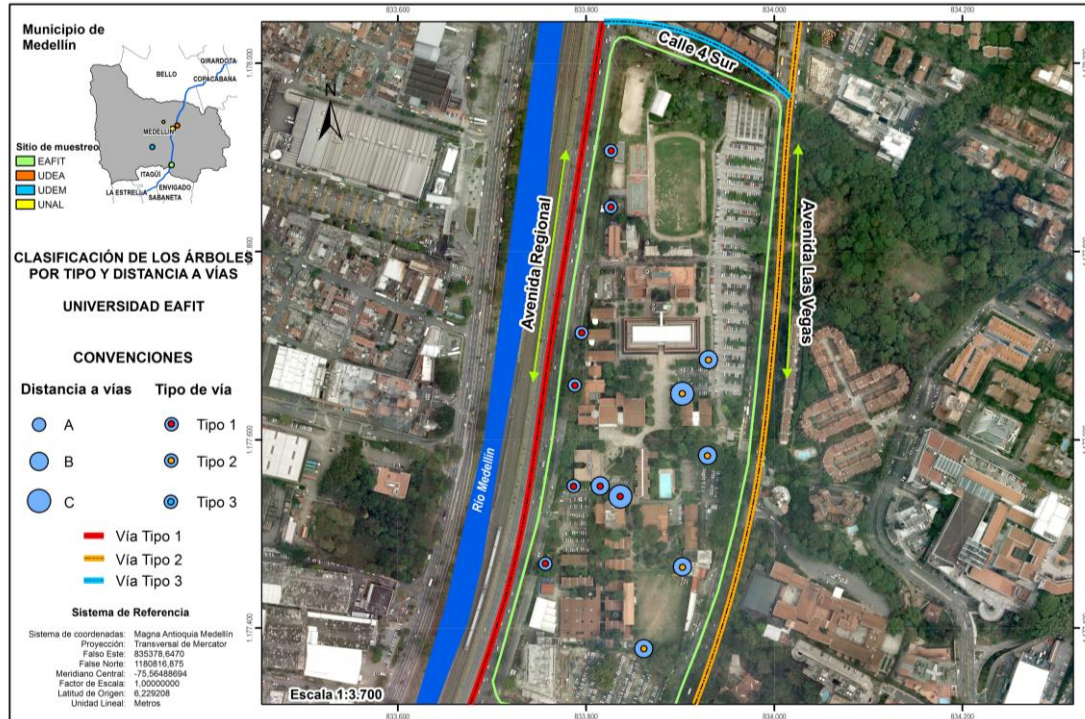


Figura 3–14: Clasificación de los árboles por tipo y distancia a las vías en Eafit.

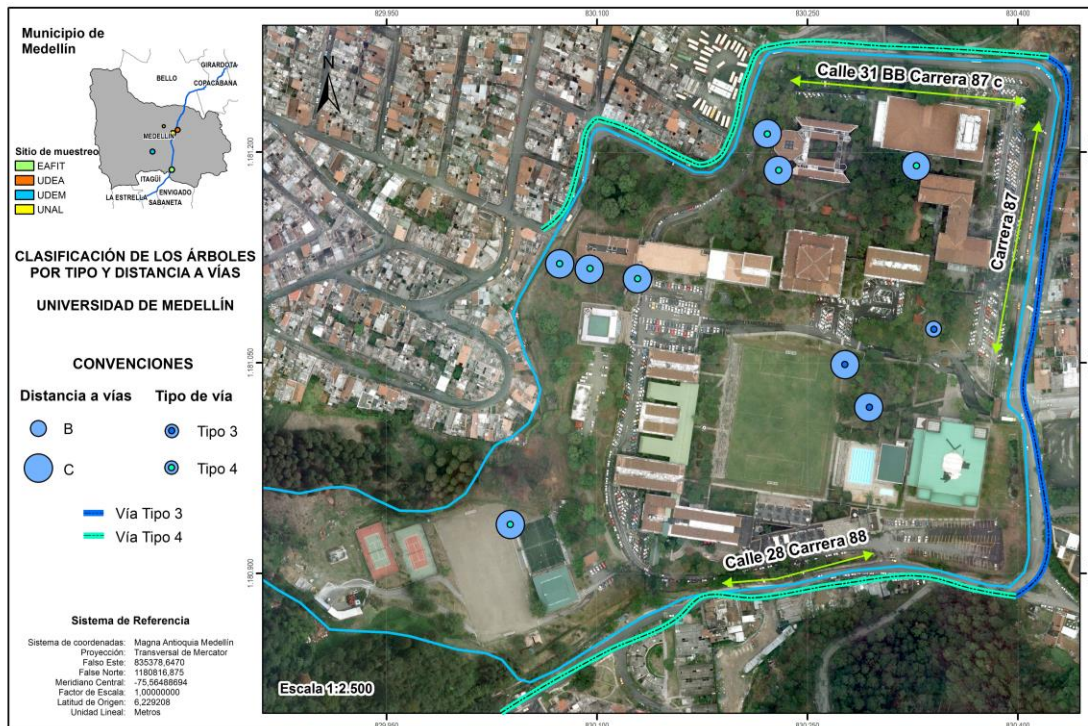


Figura 3–15: Clasificación de los árboles por tipo y distancia a las vías en la Universidad de Medellín.

A partir de los criterios descritos anteriormente, se obtuvo como resultado la siguiente clasificación de los árboles considerados en el muestreo, bajo las categorías de Crítica, Intermedia y Baja. Como se muestra en la Tabla 3–4.

**Tabla 3–3:** Clasificación de las vías de tránsito adyacentes a los sitios de muestreo.

Sitio de muestreo	Vías	Clasificación de vías según Secretaría de Movilidad (Resolución 2585 de 2016)	Calificación de vías
Universidad Nacional de Colombia	Autopista Norte	Autopista	1
	Carrera 65	Arteria	2
	Calle 59A	Vía local y ordinaria	4
Universidad EAFIT	Avenida Regional	Autopista	1
	Avenida Las Vegas	Arteria	2
	Calle 4 Sur	Arteria - Corredor	3
Universidad de Medellín	Carrera 87	Arteria	3
	Calle 31 BB Carrera 87 c	Vía local y ordinaria	4
	Calle 28 Carrera 88	Vía local y ordinaria	4

**Tabla 3–4:** Clasificación de los árboles muestreados con respecto a su localización dentro de los Campus universitarios.

Calificación de vías	Calificación de la distancia a las vías	Calificación de los árboles	Categoría de clasificación de los árboles
1	A	1A	Crítica
	B	1B	Intermedia
	C	1C	Intermedia
2	A	2A	Crítica
	B	2B	Intermedia
	C	2C	Baja
3	A	3A	Intermedia
	B	3B	Intermedia
	C	3C	Baja
4	A	4A	Baja
	B	4B	Baja
	C	4C	Baja

Una vez clasificados los árboles se llevó a cabo el test de Kruskal – Wallis como alternativa no paramétrica al análisis de varianza de una vía (ANOVA), para

verificar la existencia de diferencias estadísticas entre la reflectancia espectral media con respecto a localización de los árboles.

Este método se aplicó para los datos crudos agrupados por sitios de muestreo y especie, donde se probó la hipótesis nula ( $H_0$ ) de no haber diferencias significativas en la reflectancia espectral de las hojas de árboles en diferentes localizaciones, es decir, medias iguales, y como hipótesis alternativa ( $H_a$ ) la existencia de diferencias significativas o medias diferentes ( $H_0: C=I=B$  y  $H_a: C \neq I \neq B$ ). Donde la C denota la categoría de clasificación Crítica que corresponde a los árboles más cercanos a vías con alto flujo vehicular, la letra I corresponde a la categoría Intermedia la cual hace referencia a los árboles localizados bajo la influencia de vías con un nivel medio de flujo vehicular y la B corresponde a la categoría Baja a la cual pertenecen generalmente árboles influenciados por vías denominadas locales y ordinarias que pueden encontrarse a distancias entre 0 y > 120 m.

Después de un test de Kruskal – Wallis significativo ( $p < 0,05$ ), se realizaron pruebas post hoc de Mann - Whitney por el método Holm, las cuales se calcularon con el fin de encontrar las diferencias en la reflectancia espectral media de las hojas para todos los pares posibles (Crítica vs. Intermedia, Crítica vs. Baja, Intermedia vs. Baja) de las 3 localizaciones en las que se clasificaron los árboles.

## 4. Resultados y discusión

En este capítulo se presentan los resultados obtenidos a partir del procesamiento y análisis estadístico de los datos multiespectrales e hiperespectrales.

Como se mencionó en la metodología esta investigación se abordó desde dos escalas diferentes, relacionadas con el origen y las características de los datos espectrales. Por tanto, en esta sección, inicialmente se presenta el tratamiento y procesamiento de datos multiespectrales extraídos de una imagen de alta resolución espacial pero baja resolución espectral (4 bandas – R, G, B e IF). A partir de estos resultados y su respectivo análisis estadístico se seleccionaron un grupo menor de especies que se estudiaron empleando datos hiperespectrales tomados con un espectrómetro portátil, los resultados obtenidos con este proceso conforman la segunda parte de la presente sección.

Finalmente, se desarrolla un análisis estadístico con los datos hiperespectrales en el que se evaluó la variabilidad en la respuesta espectral de las especies forestales seleccionadas considerando la posición vertical del dosel y la distancia a una fuente contaminante (vías).

### 4.1 Datos multiespectrales

Los análisis realizados a partir de los datos multiespectrales en esta investigación contemplaron 10 especies forestales localizadas en 4 Campus universitarios del municipio de Medellín, representadas por 665 individuos (356.391 observaciones o píxeles). Después de llevar a cabo la detección y eliminación de valores atípicos multivariados fueron considerados 652 árboles (339.012 observaciones o píxeles) (Tabla 4–1).

**Tabla 4–1:** Número de árboles y píxeles considerados en el estudio después de eliminar valores atípicos.

Especie	Sitio					
	UNAL	UDEA	EAFIT	UDEM	Total Individuos	Total Píxeles
<i>Cordia alliodora</i>	11	10	-	-	21	9.513
<i>Erythrina fusca</i>	86	16	17	24	143	120.513

Especie	Sitio					
	UNAL	UDEA	EAFIT	UDEM	Total Individuos	Total Pixeles
<i>Ficus benjamina</i>	19	53	54	-	126	94.289
<i>Pithecellobium dulce</i>	37	19	1	9	66	34.910
<i>Spathodea campanulata</i>	32	31	6	18	87	27.148
<i>Swietenia macrophylla</i>	6	18	-	-	24	6.106
<i>Syzygium malaccense</i>	6	2	-	3	11	3.490
<i>Triplaris americana</i>	17	12	1	1	31	6.209
<i>Terminalia catappa</i>	40	13	3	17	73	22.652
<i>Tabebuia rosea</i>	18	13	38	1	70	14.182
<b>Total</b>	272	187	120	73	652	339.012

A partir de esta información se llevó a cabo la caracterización espectral y la discriminación de las especies, temas que se abordaron considerando las 10 especies mencionadas anteriormente y una selección de un menor de éstas con el fin de evaluar el desempeño del método de clasificación. El proceso de selección de un menor de especies se describe con detalle en la sección 4.1.2.

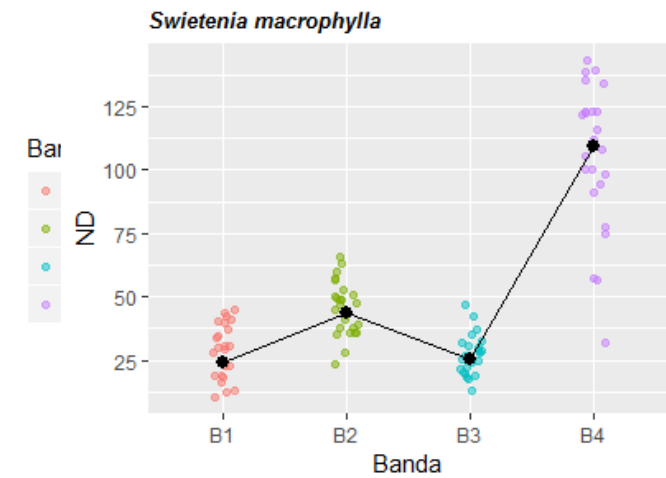
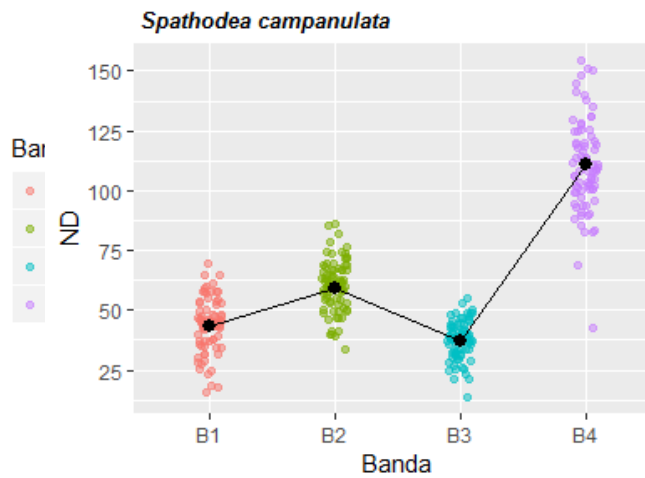
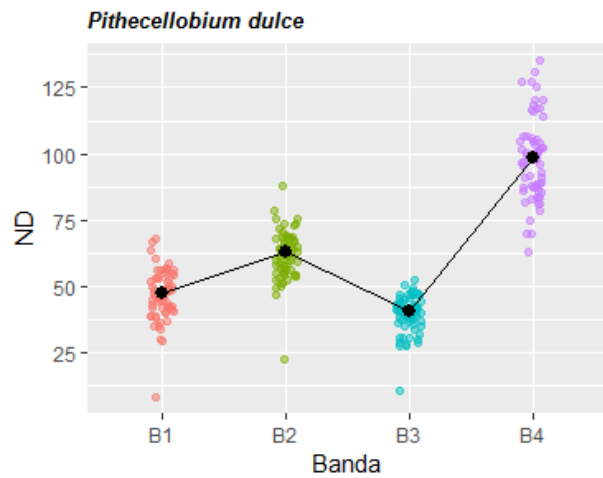
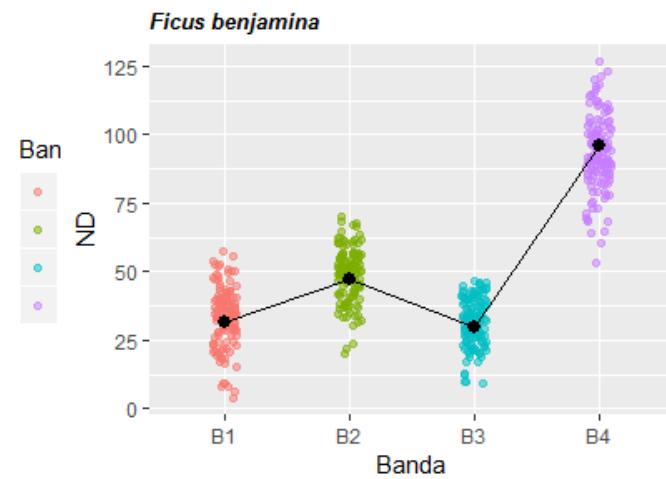
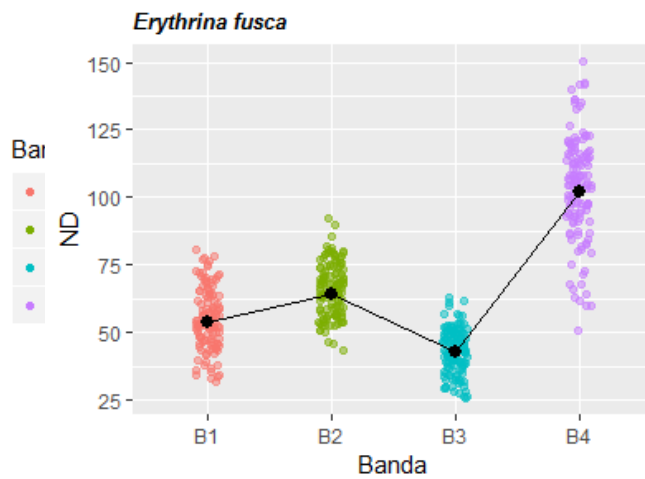
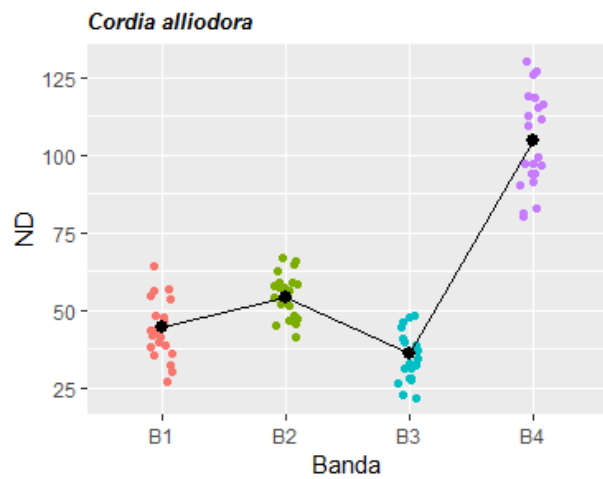
#### 4.1.1 Caracterización espectral y discriminación de las especies forestales seleccionadas en el estudio

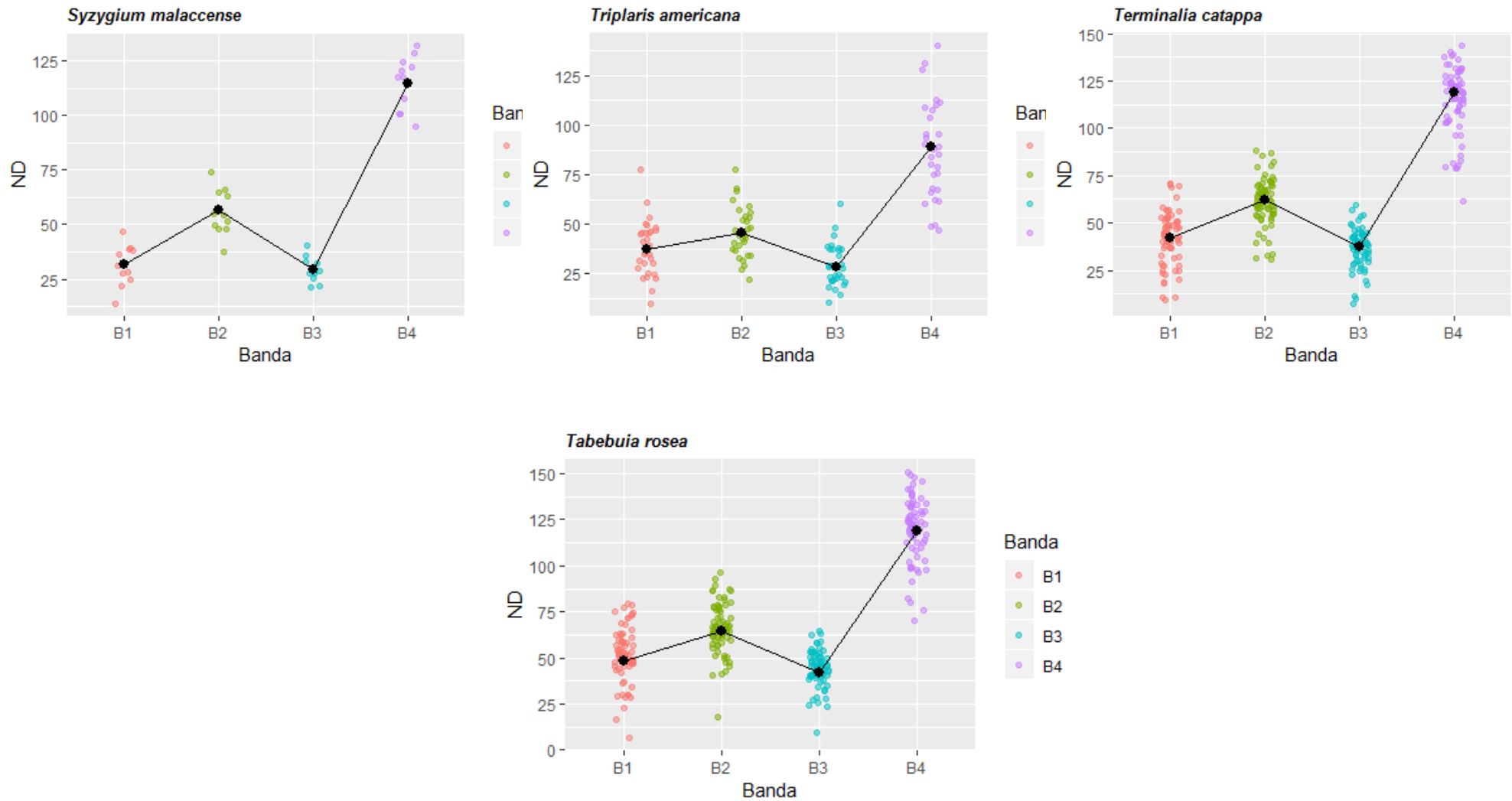
##### ▪ Caracterización espectral

Se construyeron las gráficas de la respuesta espectral media para las 10 especies seleccionadas en el estudio, teniendo en cuenta todos los sitios de muestreo conjuntamente. Como se muestra en la Figura 4–1, en donde todas las especies presentaron alta variabilidad en la respuesta espectral de cada banda, las cuales representan el promedio de un intervalo amplio del espectro.

Esto evidencia la dificultad de los datos multiespectrales para capturar información precisa que puede estar ocurriendo en puntos específicos, dado que este tipo de datos se caracteriza por generar un valor promedio de la respuesta espectral a partir de la medición de un fragmento grande del espectro, perdiendo de esta manera la particularidad y generando un aumento considerable en la variabilidad.



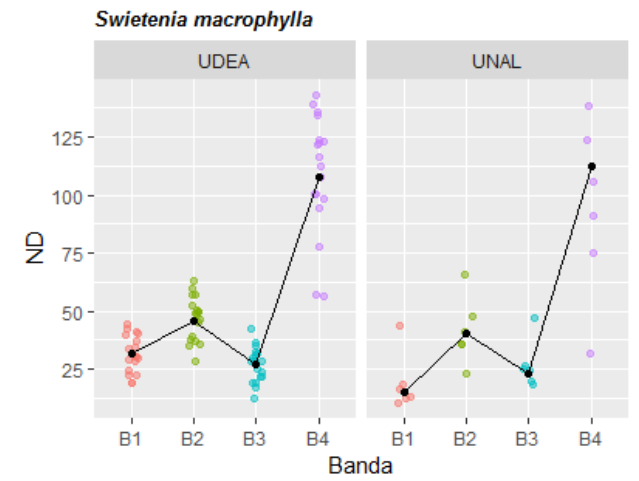
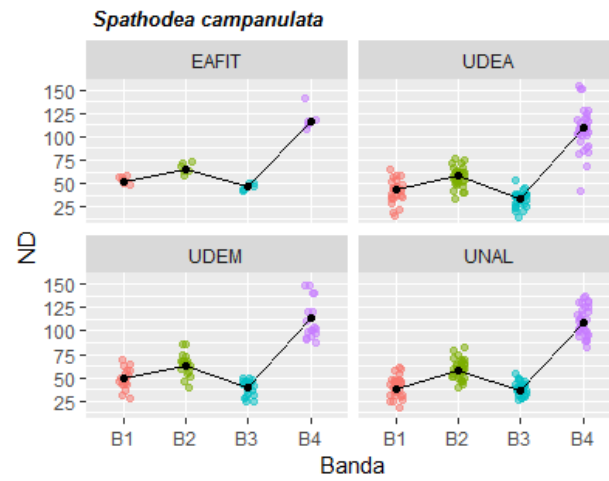
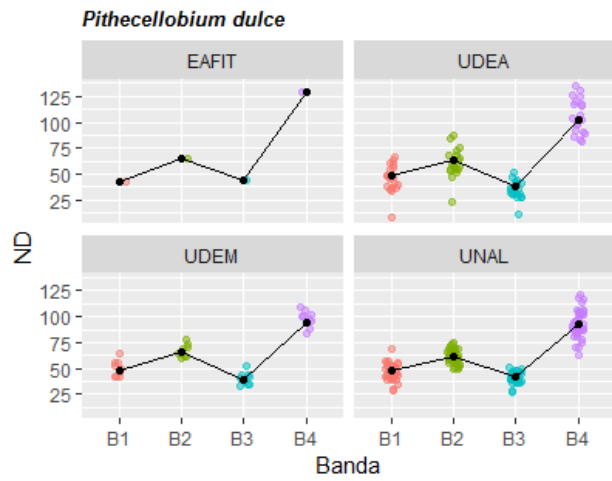
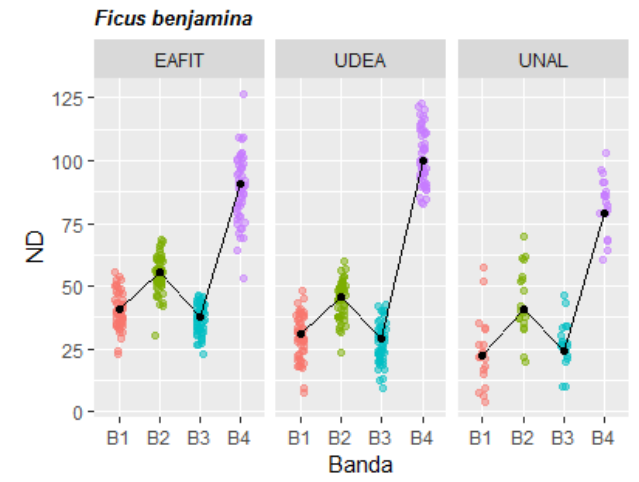
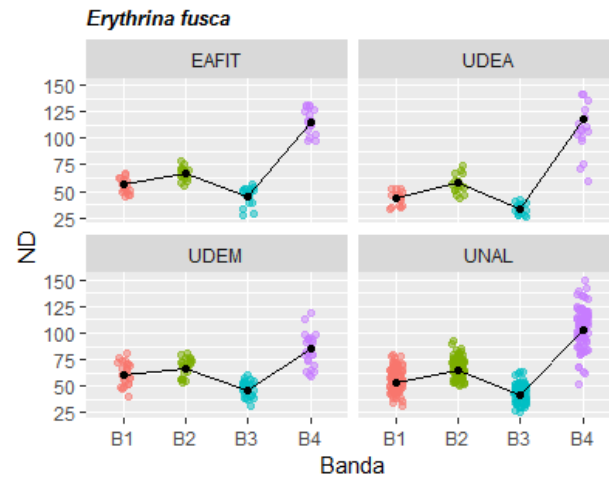
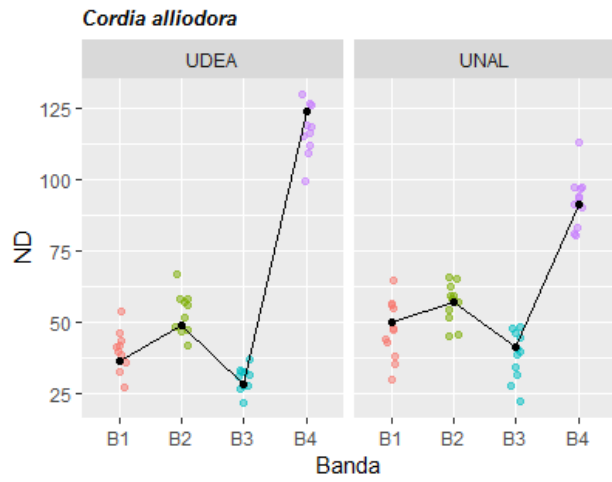


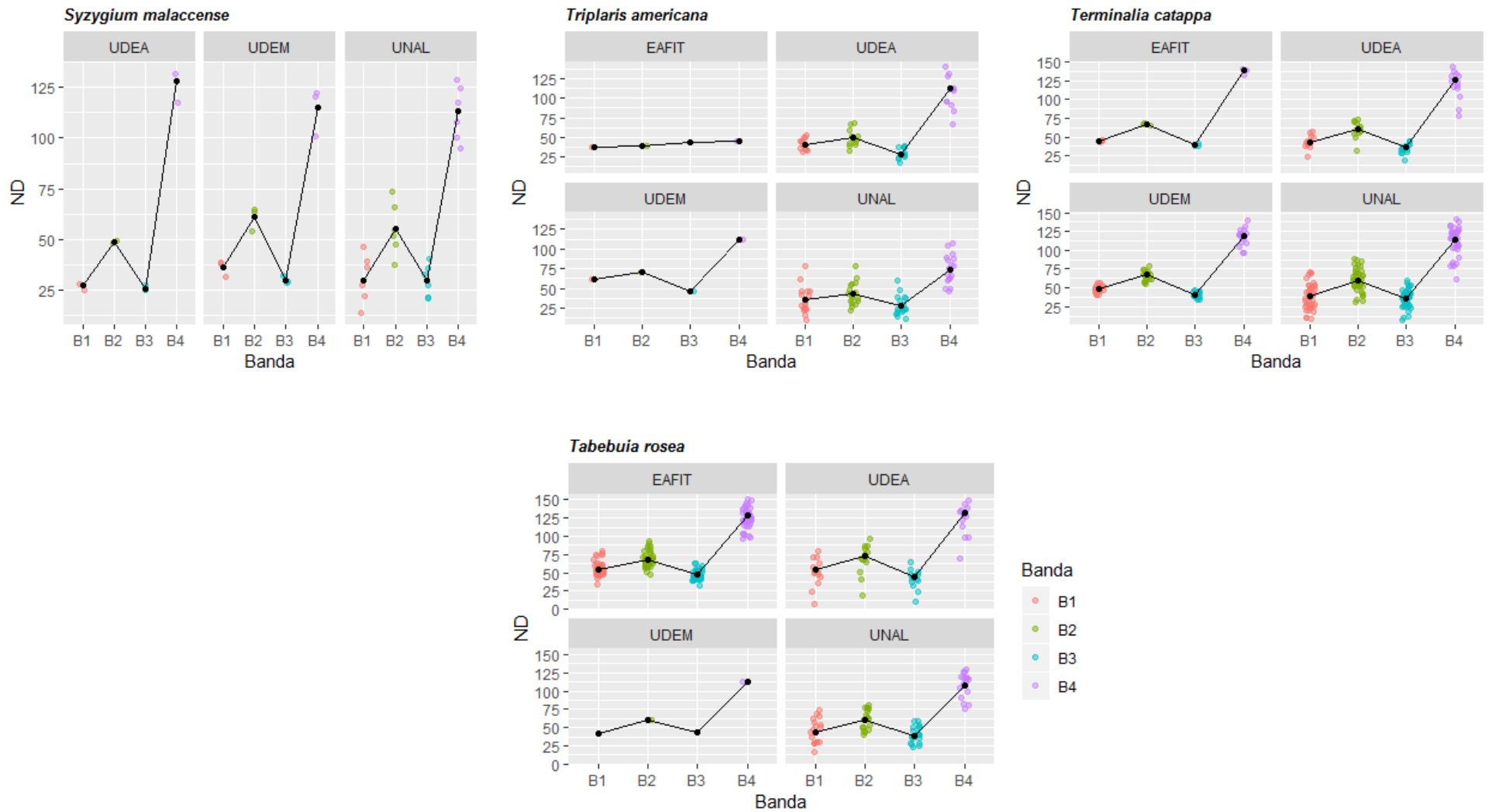


**Figura 4-1:** Respuesta espectral media de las 10 especies forestales seleccionadas.

La alta variabilidad encontrada en la respuesta espectral promedio de las especies y reflejada en la dispersión de su respectiva gráfica, se relacionó en primera instancia con las condiciones ambientales en las que se encontraban los individuos correspondientes a las especies seleccionadas en el estudio, de acuerdo a la influencia de la localización geográfica de los Campus universitarios y al manejo que le dan al arbolado en cada uno de éstos. Por tanto, la respuesta espectral media de cada especie fue discriminada para cada Campus, como se muestra en la Figura 4–2.

Cabe aclarar que, en este estudio los datos multiespectrales se presentaron en números digitales (ND), ya que no se contó con los metadatos de la imagen para realizar la transformación a reflectancia.



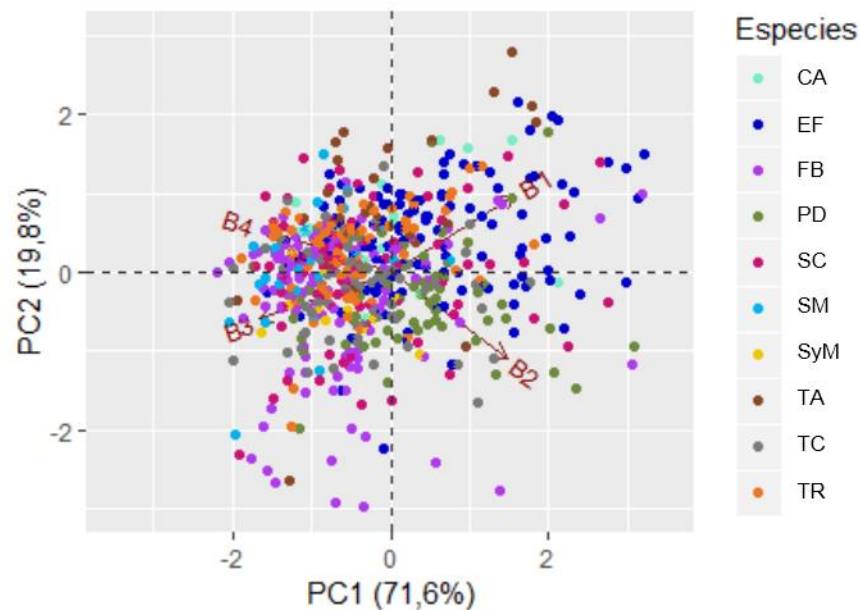


**Figura 4–2:** Respuesta espectral media de las 10 especies forestales seleccionadas por sitios de muestreo.

### ▪ Discriminación espectral de las especies forestales

Se realizaron diferentes pre – tratamientos que consistieron en la aplicación de las transformaciones SNV y MSC, con el objetivo de minimizar el ruido y suavizar la respuesta espectral de este tipo de datos. De igual manera, el análisis exploratorio inicial con Análisis de Componentes Principales (PCA) y posterior clasificación de las 10 especies seleccionadas se llevaron a cabo con los datos originales y transformados.

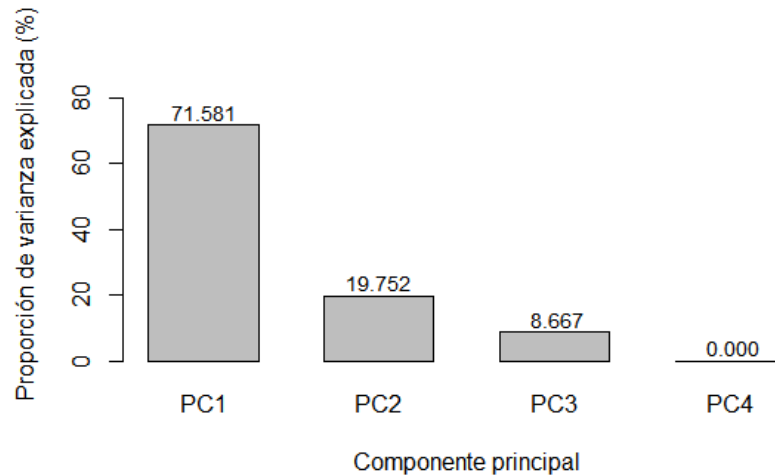
Los datos multiespectrales presentaron alta colinealidad entre sus variables lo que se evidenció adicionalmente en los PCA (Figura 4–3), por tanto, se les aplicó un test diagnóstico de multicolinealidad y posteriormente se eliminaron las bandas que aportaron información redundante. Es preciso señalar que cuando se analizaron los datos sin transformar y transformados en los diferentes sitios de muestreo, se eliminaron las bandas B3 y B4 para los datos transformados con MSC, y sólo B3 para los datos transformados con SNV, en todos los casos. Este patrón se presentó debido a que estas bandas están altamente correlacionadas con otras que proporcionan mayor varianza en el PCA.



**Figura 4–3:** Análisis de Componentes Principales con las 10 especies considerando todos los sitios de muestreo y datos transformados con SNV.

La proporción de varianza explicada por los componentes principales PC1 (71,6%) y PC2 (19,8%) abarcó más del 90% de la variabilidad total de los datos (Figura 4–4), lo que indicó la robustez del modelo. Aunque es un análisis

exploratorio inicial, en la gráfica del PCA (Figura 4–3), no se observaron agrupaciones definidas para las especies, excepto para *E. fusca* y *F. benjamina*, las cuales presentaron una separabilidad difusa, comportamiento coherente con los resultados mostrados más adelante por el Análisis Discriminante Lineal (LDA).



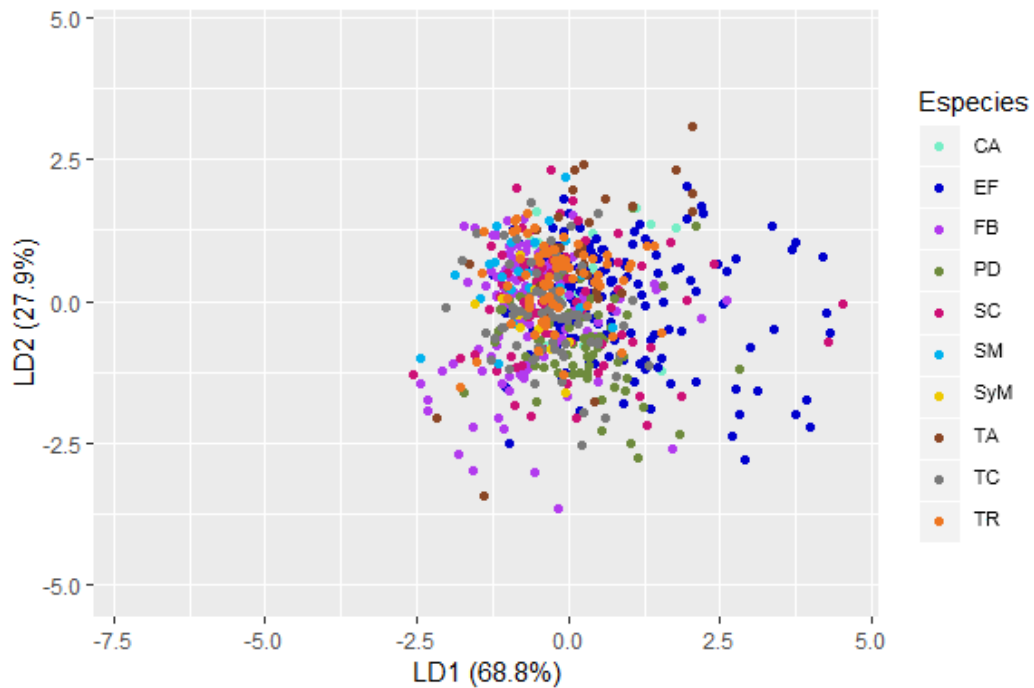
**Figura 4–4:** Proporción de la varianza explicada en el PCA con las especies *E. fusca*, *F. benjamina* y *T. catappa* considerando todos los sitios de muestreo y datos transformados con SNV.

El Análisis Discriminante Lineal (LDA) arrojó mejores resultados cuando se realizó con los datos transformados con SNV (LDA con MSC ver Anexo A), sin embargo, la precisión general del método de clasificación fue baja con 32,67%, logrando discriminar con éxito 2 especies: *E. fusca* y *F. benjamina* con una precisión de 72,73% y 86,51% respectivamente, cuando se tuvieron en cuenta todos los sitios de muestreo (Tabla 4–2). Este comportamiento se evidencia en la gráfica resultante del LDA que se presenta en la Figura 4–5, donde no es posible diferenciar o separar las clases conformadas por la información espectral de las especies restantes.

El LDA también se aplicó teniendo en cuenta la información espectral de cada sitio de muestreo específico, en los cuales discriminó a lo sumo 4 especies. Los valores más altos de precisión se obtuvieron para las especies *E. fusca* y *F. benjamina*, coincidiendo con las discriminadas cuando se consideraron todos los sitios de muestreo.

**Tabla 4–2:** Matriz de confusión obtenida a partir del LDA con las 10 especies considerando todos los sitios de muestreo y datos transformados con SNV.

Especie	CA	EF	FB	PD	SC	SM	SY	TA	TC	TR	Total	Precisión (%)
CA	0	9	12	0	0	0	0	0	0	0	21	0,00
EF	0	104	39	0	0	0	0	0	0	0	143	72,73
FB	0	17	109	0	0	0	0	0	0	0	126	86,51
PD	0	36	30	0	0	0	0	0	0	0	66	0,00
SC	0	27	60	0	0	0	0	0	0	0	87	0,00
SM	0	3	21	0	0	0	0	0	0	0	24	0,00
SY	0	0	11	0	0	0	0	0	0	0	11	0,00
TA	0	18	13	0	0	0	0	0	0	0	31	0,00
TC	0	16	57	0	0	0	0	0	0	0	73	0,00
TR	0	20	50	0	0	0	0	0	0	0	70	0,00
<b>Total</b>	0	250	402	0	0	0	0	0	0	0	652	32,67



**Figura 4–5:** Análisis Discriminante Lineal con las 10 especies considerando todos los sitios de muestreo y datos transformados con SNV.

A las especies *E. fusca* y *F. benjamina* les correspondieron el mayor número de individuos, las demás especies no fueron representativas de la población total, situación que se vio reflejada en la discriminación de estas especies y las



precisiones reportadas para éstas, sugiriendo dificultades estadísticas relacionadas con el tamaño de la muestra.

El tamaño reducido de las muestras de cada especie es uno de los criterios que dificulta en mayor medida la comparación con los resultados obtenidos por otros autores; sin embargo, en casos donde la cantidad de datos fue muy baja, resultó desafiante para la clasificación e incluso imposible cuando se contaba con un gran número de variables predictoras (Fassnacht *et al.*, 2016).

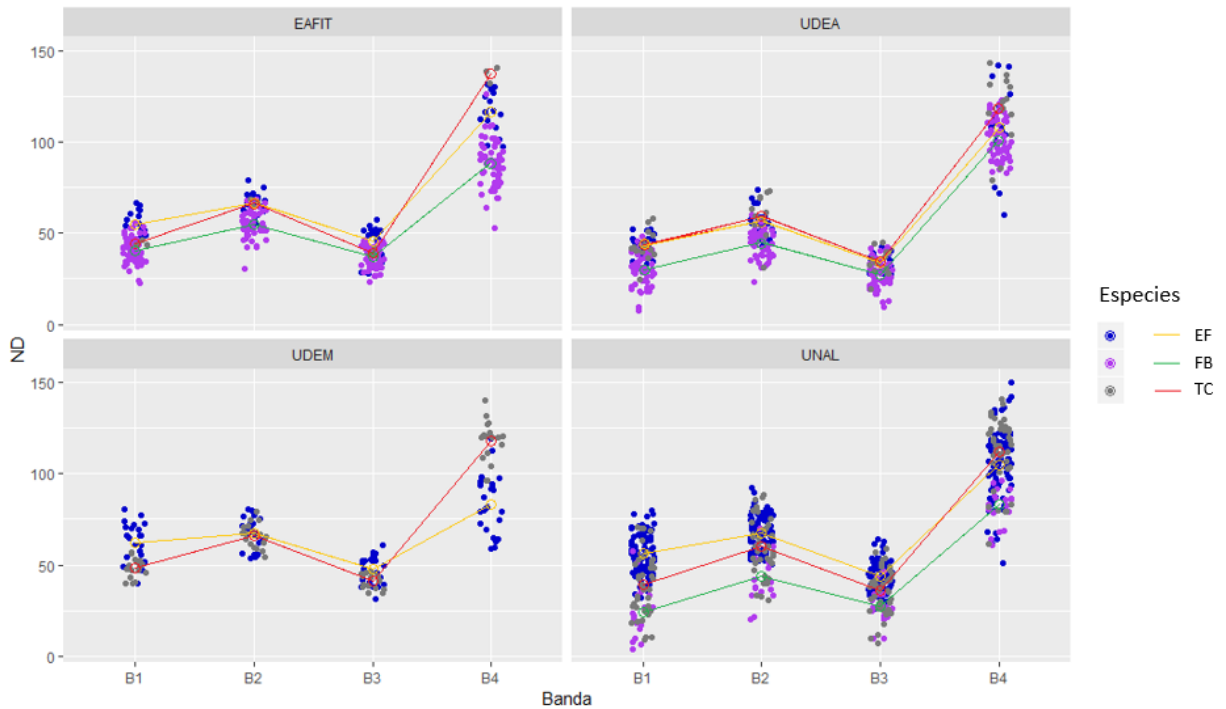
#### **4.1.2 Caracterización espectral y discriminación de las especies: *E. fusca*, *F. benjamina* y *T. catappa***

El mayor número de individuos y los porcentajes de precisión más elevados, se constituyeron en los criterios de selección de las especies -*E. fusca* y *F. benjamina*. Se eligió una tercera especie con un comportamiento espectral intermedio en comparación con las demás especies, como se evidenció en la Figura 4–5, correspondió a *T. catappa*; esto se realizó con el objetivo de verificar la potencia del método de clasificación cuando se tiene un número más reducido de especies. Por tanto, a continuación, se presenta la caracterización espectral y la clasificación de éstas especies a partir de datos multiespectrales.

##### **▪ Caracterización espectral**

La media de la banda B1 fue de 54,15; 43,82 y 31,41 para la especie *E. fusca*, *T. catappa* y *F. benjamina*, respectivamente. Para la banda B2 se observan valores medios similares para las especies *E. fusca* y *T. catappa* con 64,62 y 63,10, excepto en la UNAL en donde se presentan pequeñas diferencias para estas especies, con 67,62 y 60,04. En el caso de la banda B3 se evidencia un comportamiento similar que el de la banda B2, en donde *E. fusca* y *T. catappa* tienen medias similares (42,83 y 37,92), aunque con mayor dispersión en los sitios de muestreo correspondientes a UNAL, EAFIT y UDEM en los cuales se presentan diferencias en promedio de 7,15. La banda B4 exhibe valores mayores para la especie *T. catappa* con una media de 121,47, seguida por *E. fusca* con 103,26 y *F. benjamina* con 90,84, registrando este patrón para todos los sitios de muestreo.

Como se evidencia en la Figura 4–6, la mayor dispersión se observa en los sitios de muestreo correspondientes a UNAL y UDEA con desviaciones estándar de 16,16 y 10,45, respectivamente. Particularmente, la especie *T. catappa* obtuvo, una desviación estándar de 15,61 (UNAL) y 11,58 (UDEA).



**Figura 4-6:** Respuesta espectral media de las especies *E. fusca*, *F. benjamina* y *T. catappa* en los 4 sitios de muestreo.

Cabe resaltar que, la banda B4 presenta mayor dispersión en la respuesta espectral para las 3 especies seleccionadas en todos los sitios de muestreo, en comparación con las demás bandas, donde la desviación estándar en promedio fue de 11,72, 13,39 y 17,64 para las especies *F. benjamina*, *T. catappa* y *E. fusca* respectivamente.

La variabilidad encontrada se corroboró mediante un análisis de medias de la respuesta espectral de cada especie, teniendo en cuenta las posibles combinaciones de los sitios de muestreo. Para *E. fusca* se encontraron diferencias estadísticamente significativas ( $P < 0,05$ ) en la UDEA en las bandas B1 y B2 con respecto a los demás sitios de muestreo, en la banda B3 con respecto a EAFIT y UDEM, de igual forma, en la banda B4 de la UDEM con respecto a los demás sitios (Tabla 4-3).

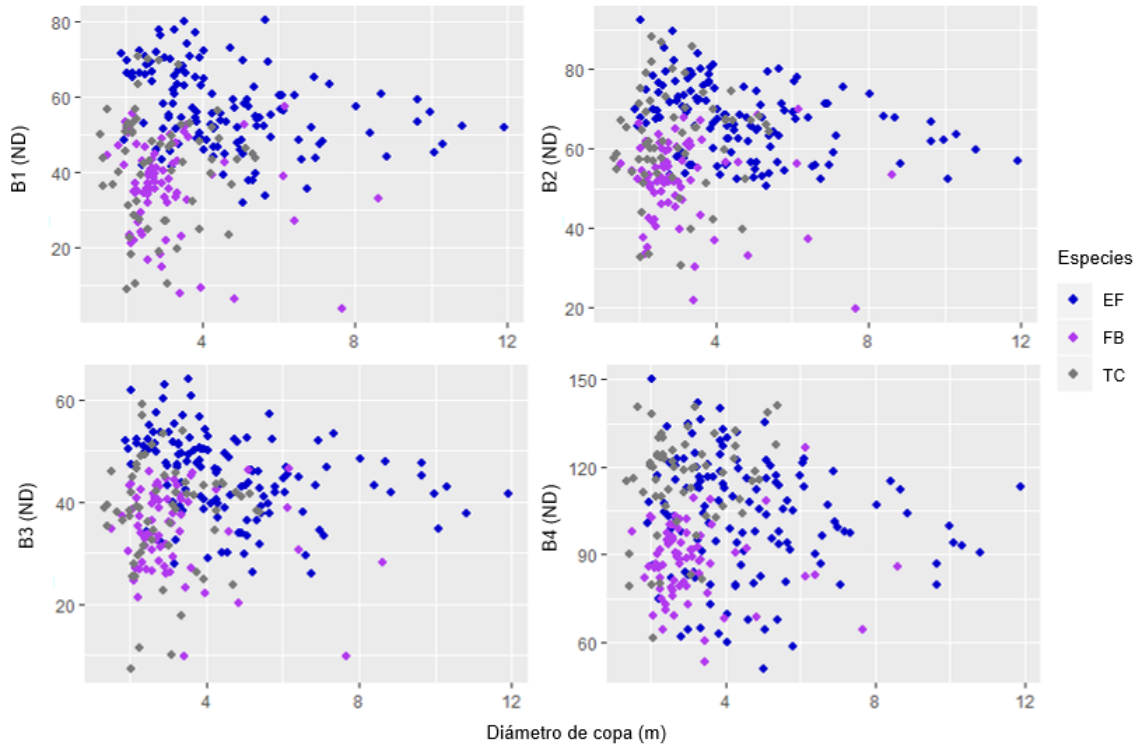
**Tabla 4–3:** Resumen de la variabilidad en la respuesta espectral de las bandas (B1, B2, B3 y B4) en los diferentes sitios de muestreo.

Especies	Banda	Valor P de la combinación de los sitios de muestreo*					
		UNAL vs. UDEA	UNAL vs. EAFIT	UNAL vs. UDEM	UDEA vs. EAFIT	UDEA vs. UDEM	EAFIT vs. UDEM
<i>E. fusca</i>	B1	5,735E-05	0,975	0,073	0,007	5,934E-07	0,148
	B2	9,243E-05	0,983	1,000	0,009	0,002	0,994
	B3	1,384E-05	0,898	0,144	1,337E-04	4,288E-07	0,762
	B4	0,977	0,123	3,988E-06	0,520	4,480E-04	5,158E-07
<i>F. benjamina</i>	B1	0,047	1,080E-09	-	3,849E-08	-	-
	B2	0,931	2,114E-05	-	7,039E-08	-	-
	B3	0,999	2,786E-06	-	1,812E-10	-	-
	B4	2,229E-07	0,129	-	2,120E-06	-	-
<i>T. catappa</i>	B1	0,777	0,907	0,097	0,999	0,739	0,971
	B2	1,000	0,811	0,364	0,814	0,518	1,000
	B3	0,978	0,940	0,270	0,883	0,291	0,991
	B4	0,205	0,005	0,426	0,048	0,636	0,021

\*En color azul se muestran los valores  $P < 0,05$ , los cuales presentan diferencias estadísticamente significativas.

Para la especie *F. benjamina* se presentaron diferencias estadísticamente significativas ( $P < 0,05$ ) para las bandas B1, B2 y B3 de EAFIT con relación a los sitios de muestreo UNAL y UDEA. Además, las bandas B1 y B4 mostraron diferencias estadísticas significativas ( $P < 0,05$ ) entre los sitios UNAL y UDEA. También, se encontraron diferencias estadísticamente significativas ( $P < 0,05$ ) en *T. catappa* en la banda B4 de EAFIT con respecto a los demás sitios de muestreo.

La variabilidad visualizada en la respuesta espectral media a partir de datos multiespectrales, se relacionó con el diámetro de copa (DC) de los árboles seleccionados, como una posible estimación de la edad y tamaño de estos. Sin embargo, en la Figura 4–7 evidenció que el DC no está influenciando la variabilidad espectral encontrada, dado que en la distribución de los datos no se observó un patrón representativo, sugiriendo que la variabilidad puede asociarse a otras variables que no fueron evaluadas en este caso.



**Figura 4-7:** Variación de la respuesta espectral en función del diámetro de copa (DC) de los árboles seleccionados por especie.

La alta variabilidad encontrada en la respuesta espectral en estudios similares, la han vinculado con el estado fitosanitario, la variabilidad natural de cada especie (Waser *et al.*, 2014) y la influencia en la cantidad de radiación reflejada debido a la interacción con los árboles vecinos (Korpela *et al.*, 2011).

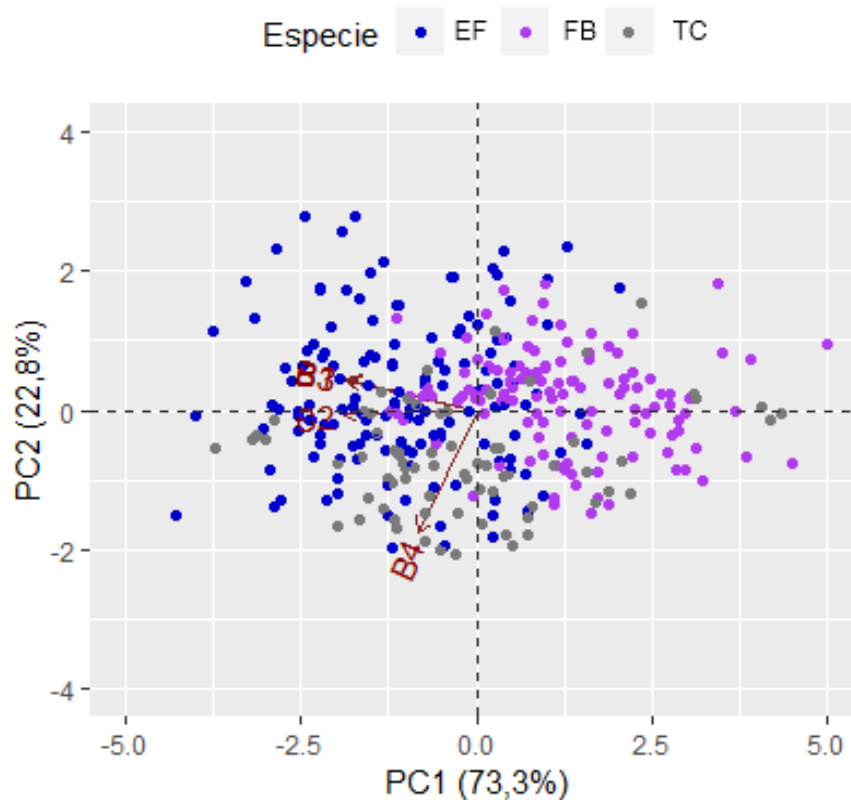
#### ▪ **Discriminación de las especies *E. fusca*, *F. benjamina* y *T. catappa***

Se realizó la discriminación de especies considerando los datos transformados (SNV y MSC) y sin transformar, como se planteó en el capítulo metodológico. Se eligió el mejor tratamiento que correspondió al escenario que tuvo en cuenta los datos sin transformar de las especies seleccionadas, ya que se observó una mayor capacidad de discriminación del método y mayor facilidad para realizar la interpretación de los resultados al considerar los datos originales (LDA con SNV y MSC ver Anexo A).

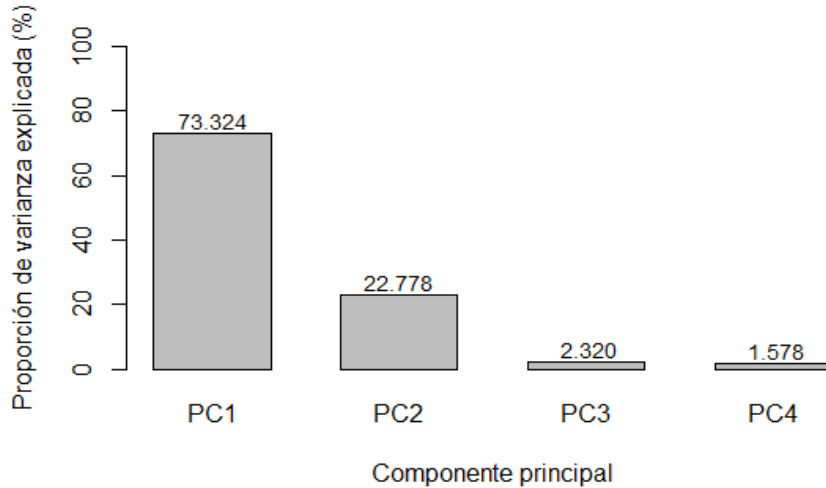
Como se muestra en la Figura 4-8, el PCA evidenció la colinealidad existente entre las bandas B1 y B3. Dicho comportamiento se corroboró mediante la aplicación de un test diagnóstico de multicolinealidad y posteriormente se eliminó la banda B3. Es preciso señalar que cuando se analizaron los datos sin transformar y transformados mediante SNV y MSC en los diferentes sitios de

muestreo, se eliminaron las bandas B3 y/o B4 en todos los casos. Este patrón se presentó debido a que estas bandas están altamente correlacionadas con otras que proporcionan mayor varianza en el PCA.

La proporción de varianza explicada por los componentes principales PC1 (73,3%) y PC2 (22,8%) abarcó casi el 100% de la variabilidad total de los datos (Figura 4–9), lo que indicó la robustez del modelo. Aunque es un análisis exploratorio inicial, en la gráfica del PCA se observaron agrupaciones definidas para las especies *E. fusca* y *F. benjamina*, y difusa para *T. catappa*, comportamiento coherente con los resultados arrojados más adelante por el Análisis Discriminante Lineal (LDA).



**Figura 4–8:** Análisis de Componentes Principales con las especies *E. fusca*, *F. benjamina* y *T. catappa* considerando todos los sitios de muestreo y datos originales.

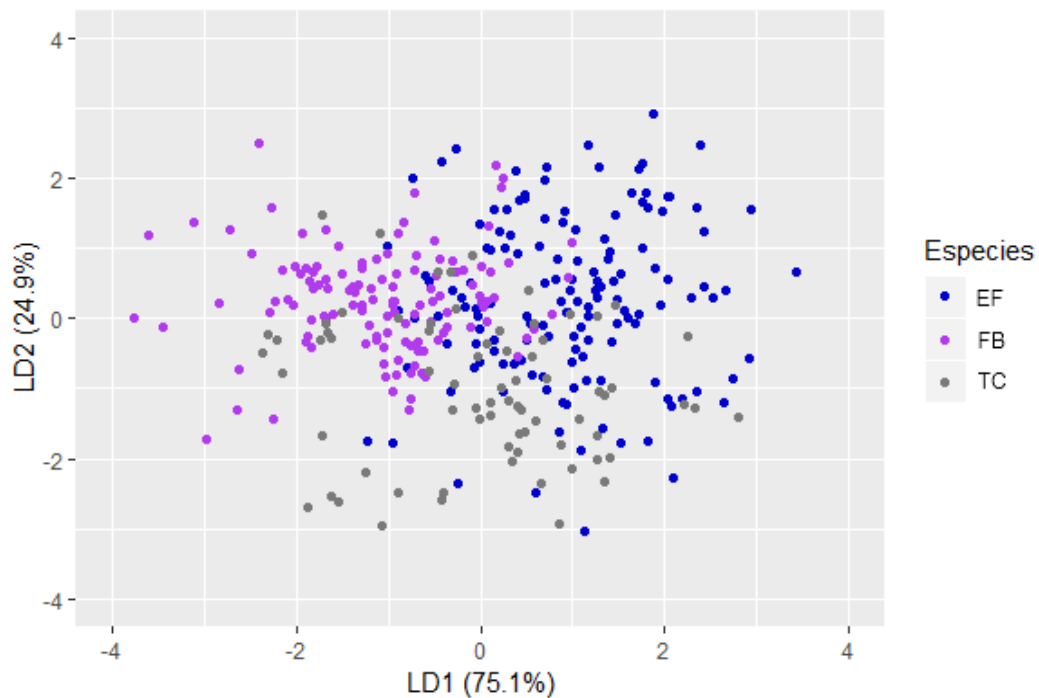


**Figura 4–9:** Proporción de la varianza explicada en el PCA con las especies *E. fusca*, *F. benjamina* y *T. catappa* considerando todos los sitios de muestreo y datos originales.

Los resultados obtenidos con el LDA a partir de los datos originales alcanzaron una precisión general de 67,25%. Sin embargo, las precisiones individuales aumentaron notablemente para las especies *E. fusca* y *F. benjamina* con 87,41% y 83,33% respectivamente. Ninguno de los 73 árboles de la especie *T. catappa* fue clasificado correctamente, como se muestra en la matriz de confusión (Tabla 4–4 y Figura 4–10).

**Tabla 4–4:** Matriz de confusión obtenida a partir del LDA con las especies *E. fusca*, *F. benjamina* y *T. catappa* considerando todos los sitios de muestreo y datos originales.

Pronóstico	EF	FB	TC	Total	Precisión (%)
EF	125	18	0	143	87,41
FB	21	105	0	126	83,33
TC	44	29	0	73	0,00
<b>Total</b>	190	152	0	342	67,25

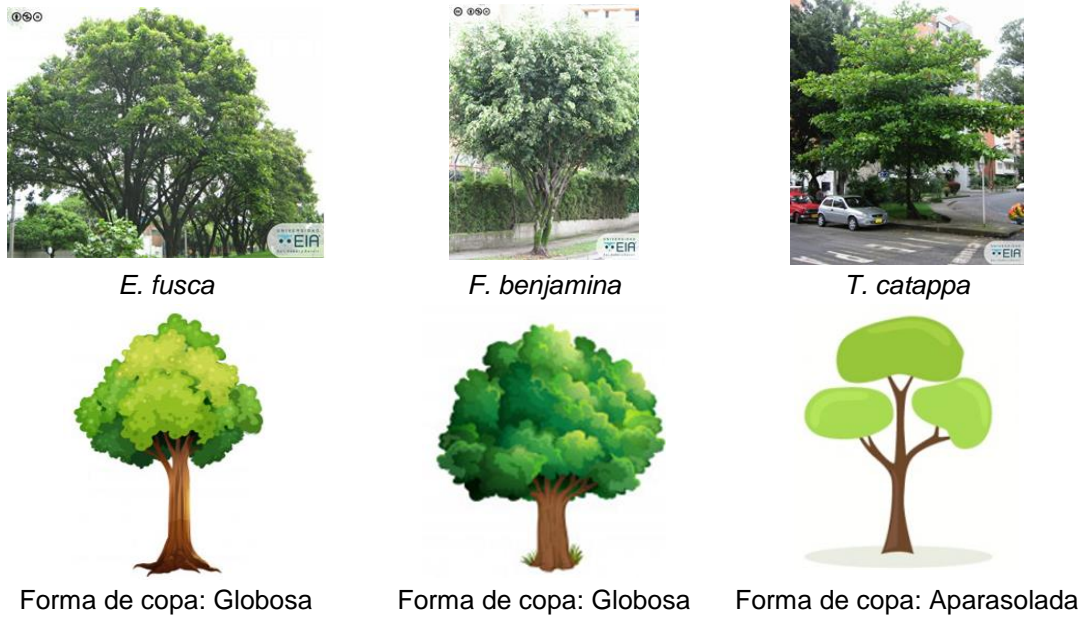


**Figura 4–10:** Análisis Discriminante Lineal con las especies *E. fusca*, *F. benjamina* y *T. catappa* considerando todos los sitios de muestreo y datos sin transformar.

Los bajos valores de precisión en la clasificación de *T. catappa* se conservaron en los escenarios donde se tuvo en cuenta las 3 especies en todos los sitios de muestreo, esto puede estar relacionado con las longitudes de onda con las que se contaron para la discriminación, ya que Heikkinen *et al.*, (2010) han confirmado un incremento en la precisión del 5 a 13% en la clasificación de algunas especies cuando consideraron una quinta banda cercana al espectro de borde rojo.

De igual manera, se han encontrado limitaciones en la clasificación de algunas especies, las cuales se asociaron a una estructura forestal compleja, con la existencia de varios estratos y copas pequeñas que en muchos de los casos no correspondieron a los árboles emergentes (Immitzer *et al.*, 2012). Estas características mencionadas anteriormente conducen a la presencia de píxeles mixtos, los cuales causan clasificaciones erróneas.

Esto es evidente en *T. catappa*, donde los resultados obtenidos en la discriminación posiblemente están relacionados con la arquitectura de la copa y su forma aparasolada. La estratificación de *T. catappa* puede involucrar píxeles con mayor variabilidad espectral o píxeles filtrados que corresponden a materiales artificiales del entorno urbano. (Pu & Landry, 2012). (Figura 4–11).



**Figura 4–11:** Forma de la copa de las especies *E. fusca*, *F. benjamina* y *T. catappa*.

**Fuente:** Fotografías tomadas del Catálogo de flora del Valle de Aburrá. Universidad EIA.

De otro lado, asociado al entorno urbano y sus características, Zhang & Hu (2012) observaron mejores precisiones en la clasificación de especies, debido a la baja competencia entre los individuos que se da en estos entornos, dado que los árboles tienen más probabilidades de desarrollar su tamaño y forma determinada por la genética, reduciendo la variabilidad espectral de los píxeles.

Otro aspecto que dificultó la clasificación de la especie *T. catappa* se denomina “señal de fondo”, consiste en que los espectros de las copas se mezclaron con diversos espectros de la superficie. Esto se hizo evidente en el presente estudio, por el tamaño reducido de la copa de algunos árboles de esta especie, a pesar de la resolución espacial alta de la imagen empleada. Vale aclarar que los principios aceptados para la resolución espacial mínima de la imagen, sugieren que una copa debe estar cubierta por al menos 4 píxeles para poder identificar la especie con precisión. Este principio no se cumplió, ya que las copas estaban representadas por un número mayor de píxeles (Jensen, 2007).

La clasificación de especies a partir de datos multiespectrales, se ha beneficiado a partir del 2009 por la utilización de las imágenes del satélite WorldView-2, las cuales presentan mayor resolución espectral (4 bandas adicionales de las estándar y resolución espacial de 2m). En contraste, (Chen, 2011) y Ke *et al.*, (2010), emplearon la información satelital de IKONOS -4 m- y QuickBird -2,5 m- respectivamente, con una alta resolución espacial pero menor resolución espectral, dado que solo cuentan con 4 bandas (rojo, verde, azul e infrarrojo).



cercano); como fueron las características espectrales de la imagen empleada en esta investigación.

Otras condiciones que dificultan y/o impiden la discriminación de especies, están relacionadas con: el estado de salud, diferencias en la iluminación y apertura de los árboles, estado fenológico, condiciones propias del sitio, edad del árbol y la influencia de la radiación reflejada de otras superficies, además de la variabilidad natural de cada especie (Paz-Kagan *et al.*, 2019).

Para ilustrar en este sentido el estudio de Waser *et al.*, (2014), se enfocó en la variabilidad espectral aportada por el estado fitosanitario de los árboles, clasificando 7 especies de árboles diferentes y 4 niveles de infestación en la especie *Fraxinus excelsior*, obteniendo precisiones altas en la clasificación de especies de 83% y en los niveles de infestación del 73%. Además, sugirieron el uso de espectroscopía con el fin de adquirir más información espectral y evaluar otros parámetros biofísicos que posiblemente permitan realizar la clasificación con mayor precisión.

Otros autores consideraron las diferencias en la estructura física de los árboles, influenciada por las condiciones del sitio y sus vecinos, reportaron un aumento significativo en la precisión general, de 58% a 78%, e incluso llegando a un 87% de precisión, cuando la clasificación la realizaron a partir de imágenes multiespectrales y medidas de textura, relacionadas principalmente con las sombras internas de la copa, ramificación y propiedades del follaje como tamaño, densidad y reflectividad (Gini *et al.*, 2018).

En síntesis, la mayoría de las investigaciones revisadas incluyendo la presente, se han centrado en evaluar el potencial de un determinado sensor y/o método de clasificación de especies de árboles, sin valorar la multiplicidad de factores que dificultan y en ocasiones impiden la discriminación exitosa de especies (Fassnacht *et al.*, 2016).

## 4.2 Datos hiperespectrales

La aplicación de técnicas de espectroscopía se llevó a cabo dado que a pesar de la alta resolución espacial de la imagen empleada en el análisis a partir de los datos multiespectrales, la resolución espectral limitó la separabilidad de las especies. Por tanto, se evaluó el potencial de dichas técnicas para la discriminación de especies.

Para esto se seleccionaron 3 de las 10 especies consideradas en el inventario forestal ejecutado para este estudio. *Erythrina fusca*, *Ficus benjamina* y

*Terminalia catappa* fueron elegidas debido a que estaban presentes en la mayoría de los Campus universitarios, contaban con una abundancia homogénea en estos sitios (disponibilidad de individuos para el muestreo) y representan especies importantes para la prestación de servicios ecosistémicos en entornos urbanos.

Los resultados encontrados en la discriminación de estas 3 especies a partir de los datos multiespectrales, también fue considerado un criterio de decisión. No obstante, es necesario resaltar que por el hecho que hayan sido separables o confundibles con este tipo de datos, no significa que se comporten de igual manera en la clasificación a partir de datos hiperespectrales.

Una vez que *E. fusca*, *F. benjamina* y *T. catappa* se seleccionaron podría haberse estado descartando otras especies separables con espectroscopía de detalle, sin embargo, el levantamiento de los espectros foliares de reflectancia a través de el espectrómetro es una actividad dispendiosa, por tanto, con el fin de reducir el volumen de trabajo y acotar el alcance del estudio se redujo de 10 a 3 especies.

#### 4.2.1 Caracterización espectral

En total se midieron 1.710 espectros, de las especies *E. fusca*, *F. benjamina* y *T. catappa*, los cuales se redujeron a 1.567 después de llevar a cabo la detección y eliminación de “valores atípicos” (Tabla 4–5). La Figura 4–12 presenta la reflectancia media de cada una de las 3 especies seleccionadas considerando todos los sitios de muestreo, agrupada por la altura del dosel en la que se colectaron las muestras.

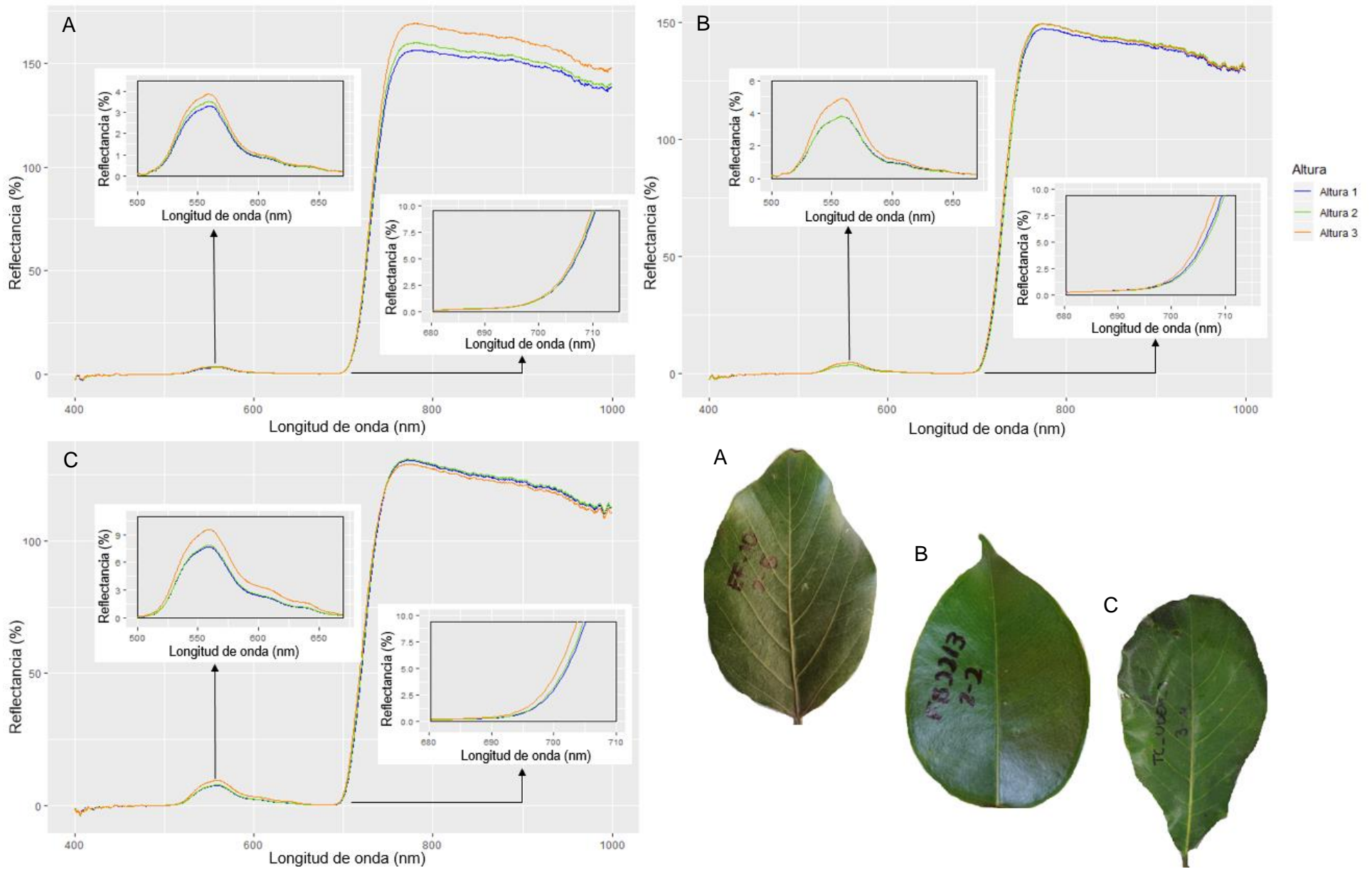
**Tabla 4–5:** Número de árboles seleccionados en los sitios de muestreo para la medición de espectros.

Especie	SITIO					
	UNAL		EAFIT		UDEM	
	Árboles	Espectros	Árboles	Espectros	Árboles	Espectros
<i>Erythrina fusca</i>	5	225	5	225	5	225
<i>Ficus benjamina</i>	5	225	5	225	-	-
<i>Terminalia catappa</i>	5	225	3	135	5	225
<b>TOTAL</b>	15	675	13	585	10	450
<b>TOTAL ÁRBOLES</b>	38					
<b>TOTAL ESPECTROS</b>	1710					

Las mayores diferencias se encontraron en las regiones del espectro correspondientes al visible (500 – 600 nm), “borde rojo” (680 – 710 nm) y el

infrarrojo cercano (750 – 1000 nm), confirmando los resultados de Li *et al.*, (2013), que relacionan estas diferencias con las propiedades bioquímicas y biofísicas del tejido de las plantas - contenido de agua, pigmentos fotosintéticos, carbohidratos estructurales y la morfología de la hoja. Estos rasgos foliares usualmente varían a lo largo del perfil vertical del dosel, de ahí que sea posible observar dichas variaciones en las alturas consideradas (Liu *et al.*, 2015).

En general, la reflectancia espectral media de las hojas presentó un patrón descendente desde la altura 3 hasta la altura 1 (nivel alto > nivel medio > nivel bajo) en el espectro visible. En el espectro del “borde rojo” las diferencias en la reflectancia media en los diferentes niveles altitudinales no fueron tan marcadas como en el espectro visible, el patrón descendente se conservó de la misma manera desde la altura 3 hasta la altura 1, para las 3 especies seleccionadas. En la región del espectro correspondiente al infrarrojo cercano, no se encontró un patrón representativo de la reflectancia media en las alturas del dosel, en ninguna de las especies estudiadas.



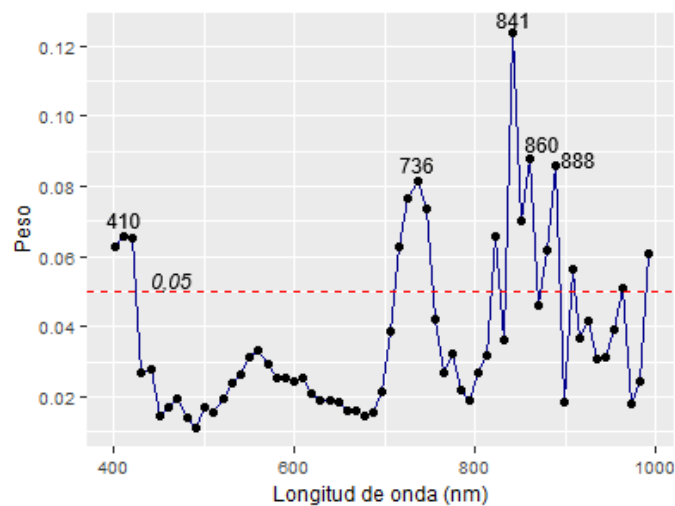
**Figura 4-12:** Respuesta espectral media en cada posición del dosel para las especies (A) *E. fusca*, (B) *F. benjamina* y (C) *T. catappa*.

### 4.2.2 Discriminación espectral de las especies forestales

La clasificación de especies a partir de datos hiperespectrales se llevó a cabo de forma similar al proceso realizado con los datos multispectrales. En este caso se utilizaron los espectros originales y transformados (SNV y MSC) de las especies *E. fusca*, *F. benjamina* y *T. catappa*. Se plantearon diferentes escenarios para la aplicación del método de clasificación, considerando las posibles combinaciones que surgieron de los sitios de muestreo, tipo de pretratamiento y las especies seleccionadas tomadas en conjunto y en pares.

Debido a las precisiones obtenidas y la consistencia de los resultados en los diferentes escenarios establecidos, el mejor pretratamiento fue SNV (LDA con MSC y datos sin transformar ver Anexo B), siendo la transformación que logró eliminar la mayor cantidad de ruido originado por la variabilidad propia de los datos espectrales, el equipo empleado y las condiciones ambientales dadas en la medición. A continuación, se presentan los resultados encontrados teniendo en cuenta los espectros transformados con SNV de las 3 especies seleccionadas en todos los sitios de muestreo.

Se realizó una reducción de variables aplicando el algoritmo *Relief* con el que se seleccionaron 16 longitudes de onda, las cuales obtuvieron un peso  $> 0,05$ , donde las bandas claves que influyeron en la discriminación espectral de especies correspondieron a una región del espectro visible (VIS) alrededor del azul (400 – 420 nm) y la región del borde rojo (716 – 745 nm), además de un intervalo en el infrarrojo cercano (NIR) (841 – 907 nm) (Figura 4–13).



**Figura 4–13:** Variables o bandas elegidas mediante el algoritmo Relief para realizar el LDA a partir de los datos transformados con SNV de las 3 especies seleccionadas.

Estas importantes longitudes de onda coinciden con las regiones del espectro influenciadas por las características de absorción de los pigmentos fotosintéticos de las plantas y el contenido de agua dentro de la estructura celular (Liu *et al.*, 2015). Lo anterior apoya la teoría de que las variaciones en estos factores producen un comportamiento espectral típico de las especies que permiten su separación en la región del VIS y NIR.

La mayoría de los estudios revisados que se enfocaron en la selección de bandas, coinciden en la importancia de la región del visible y el borde rojo para clasificar especies, hacen énfasis alrededor del pico verde entre 450 – 590 nm por presentar una correlación relativamente baja con respecto a las bandas vecinas del espectro visible, que resultó fundamental para la discriminación de especies en contextos urbanos (Alonzo *et al.*, 2016; Fassnacht *et al.*, 2014; Van Deventer *et al.*, 2015).

Estos estudios reportaron la importancia de las bandas del infrarrojo de onda corta (SWIR) en la discriminación de especies. Alonzo *et al.*, (2014) asociaron el poder discriminatorio de dicha región espectral con el contenido de lignina y celulosa que se encuentra en la materia vegetal foliar en cantidades y concentraciones diferentes para cada especie. En esta investigación no se adquirieron espectros en el SWIR, dado las características de los equipos de medición empleados, lo que posiblemente hubiera contribuido a una mejor clasificación de las especies.

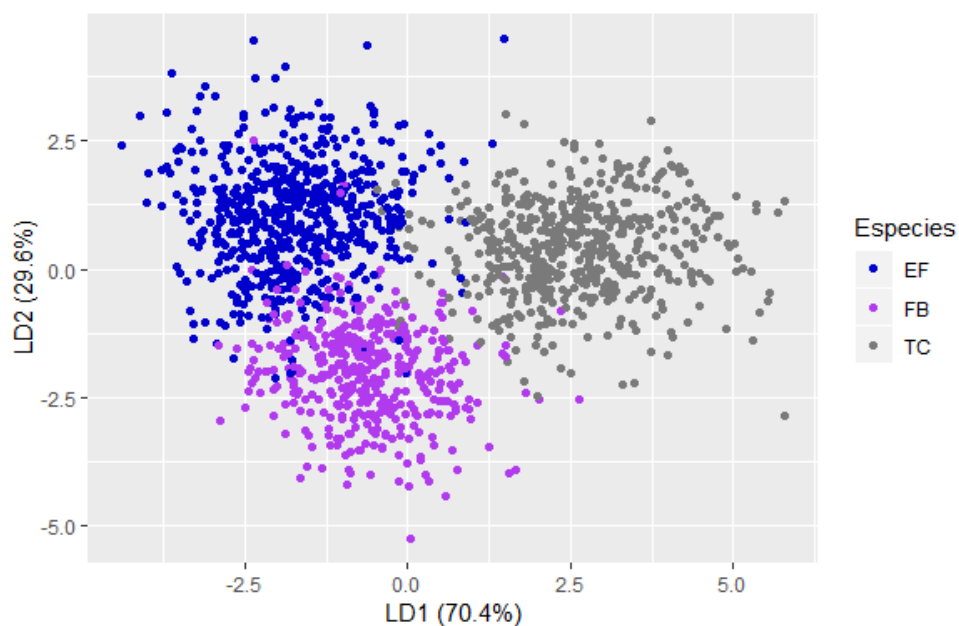
Con respecto al NIR, un estudio realizado en el trópico confirma los resultados obtenidos en esta investigación con relación a la importancia de las bandas de esta región del espectro, argumentando que la capacidad discriminatoria de las especies en el intervalo 744 nm – 1313 nm está asociada a los cambios en la fenología y el contenido del agua (Clark *et al.*, 2005). Caso contrario, Alonzo *et al.*, (2014) afirmaron en su investigación que la región del NIR no fue adecuada para la discriminación, debido a una alta variación en las bandas involucradas en la separabilidad de las especies influenciadas por los cambios en el contenido de los pigmentos fotosintéticos.

El LDA realizado a partir de los datos hiperespectrales utilizando los espectros transformados con SNV mejoró considerablemente la clasificación de las especies, aumentando aproximadamente en 11 puntos porcentuales la precisión general (78,69%). Cabe resaltar que, la especie *T. catappa* pasó de no ser clasificada a partir de los datos multiespectrales a tener una precisión de 94,89%, donde de 528 árboles fueron clasificados correctamente 501 (Tabla 4–6). En la Figura 4–14, se observa como el LDA realizado a partir de los datos hiperespectrales permitió alcanzar una mejor separación de las 3 especies

consideradas. También, se evidenció el bajo desempeño del método de clasificación con *F. benjamina*, al confundir esta especie principalmente con *E. fusca* seguida por *T. catappa*, ya que, de un total de 416 espectros, 183 fueron clasificados como *E. fusca* y 18 como *T. catappa*.

El bajo desempeño en la clasificación de *F. benjamina* que se evidenció con el mejor pretratamiento, conservó el patrón con los espectros originales y transformados con MSC, lo cual puede estar asociado a la falta de variables relevantes para la discriminación que obtuvieron un peso  $> 0,05$ , cuando se aplicó el algoritmo *Relief* con las 3 especies, alcanzando valores significativamente más altos en la precisión una vez que se realizó este proceso para las especies pareadas.

Estos resultados sugieren que la selección de variables predictoras realizada por el algoritmo *Relief* fue más eficiente cuando se realizó el análisis por pares de especies, lo que se ve reflejado en la precisión de la clasificación (Marín-Ortiz J.C., comunicación personal, 10 de mayo de 2019).



**Figura 4–14:** Análisis Discriminante Lineal con los espectros transformados SNV de las 3 especies seleccionadas considerando todos los sitios de muestreo.

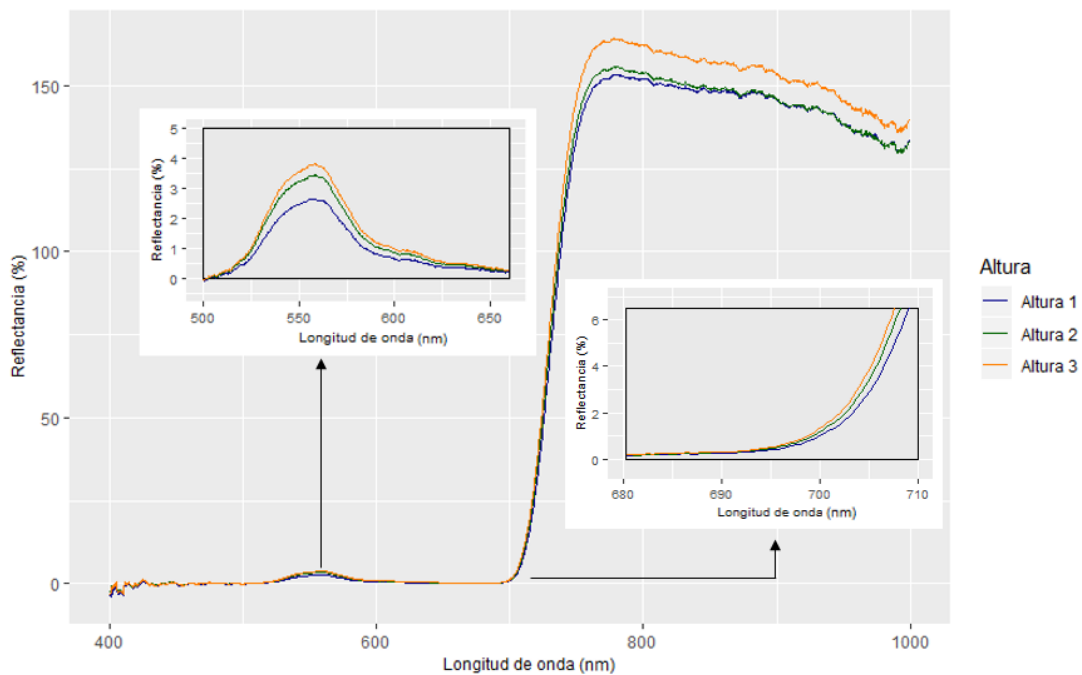
**Tabla 4–6:** Matriz de confusión obtenida a partir del LDA con los espectros transformados SNV de las 3 especies seleccionadas considerando todos los sitios de muestreo.

Pronóstico	EF	FB	TC	Total	Precisión (%)
EF	517	104	2	623	82,99
FB	183	215	18	416	51,68
TC	0	27	501	528	94,89
Total	700	346	521	1567	78,69

#### 4.2.3 Variabilidad en la repuesta espectral considerando la posición del dosel y la distancia a una fuente contaminante

En general, la reflectancia espectral media mostró un patrón descendente desde la altura 3 hasta la altura 1 (nivel alto > nivel medio > nivel bajo) en el espectro visible (500 – 650 nm). Esta tendencia descendente de la reflectancia en el VIS fue más pronunciada para las especies en el siguiente orden: *E. fusca*, *T. catappa* y *F. benjamina*. Aunque en el espectro de “borde rojo” (680 – 710 nm) las diferencias en la reflectancia media en los diferentes niveles altitudinales no fueron tan marcadas como en el espectro visible, el patrón descendente se conservó de la misma manera desde la altura 3 hasta la altura 1 (nivel alto > nivel medio > nivel bajo) cuando se consideraron cada una de las 3 especies seleccionadas en los diferentes sitios de muestreo. Estos patrones descritos se muestran en la Figura 4–12, para cada una de las 3 especies considerando todos los sitios de muestreo, y de manera ilustrativa en la Figura 4–15 para la especie *E. fusca* en la UDEM. En el Anexo C se presenta las demás gráficas de las especies por cada sitio de muestreo.





**Figura 4–15:** Respuesta espectral media en cada posición del dosel para la especie *E. fusca* en el sitio de muestreo UDEM.

Una prueba post hoc de Mann – Whitney evidenció las diferencias estadísticamente significativas ( $P < 0,05$ ) en la mayoría de los espectros de las hojas entre todas las alturas del dosel consideradas para la especie *E. fusca*, excepto en el resultado obtenido en la comparación de los espectros de la Altura 1 con la 2 en la Universidad EAFIT.

La especie *T. catappa* mostró diferencias estadísticas significativas ( $P < 0,05$ ) para todas las posibles combinaciones de alturas del dosel en los sitios de muestreo UNAL y EAFIT. Adicionalmente, no se encontraron variaciones significativas en los espectros de las hojas de la Altura 1 y 2 de la UDEM, además de cuando se consideraron todos los sitios de muestreo conjuntamente y las diferentes capas del dosel.

Finalmente, no se encontraron diferencias estadísticamente significativas ( $P > 0,05$ ) para los espectros de las hojas de *F. benjamina* en las 3 alturas del dosel en todos los sitios de muestreo (Tabla 4–7).

**Tabla 4–7:** Resumen de la variabilidad espectral media de las 3 especies seleccionadas en cada posición del dosel considerando los diferentes sitios de muestreo.

Especies	Sitio	Valor P de la combinación de las posiciones del dosel*		
		Altura 1 vs. Altura 2	Altura 1 vs. Altura 3	Altura 2 vs. Altura 3
<i>E. fusca</i>	UNAL	3,9E-08	2,2E-15	5,1E-13
	EAFIT	0,320	3E-07	2,0E-07
	UDEM	0,028	1,7E-10	2,0E-06
	TODOS LOS SITIOS	0,017	4,0E-11	1,3E-08
<i>F. benjamina</i>	UNAL	0,100	0,550	0,550
	EAFIT	0,986	0,054	0,054
	UDEM	-	-	-
	TODOS LOS SITIOS	0,350	0,280	0,820
<i>T. catappa</i>	UNAL	0,010	0,010	6,3E-06
	EAFIT	1,4E-11	< 2E-16	0,030
	UDEM	0,570	3,5E-06	6,2E-07
	TODOS LOS SITIOS	1,000	1,000	1,000

\*En color azul se muestran los valores  $P < 0,05$ , los cuales presentan diferencias estadísticamente significativas.

Caso contrario al patrón observado en el espectro visible en esta investigación, Gara *et al.*, (2018) obtuvieron que la reflectancia media medida en las hojas de 4 especies disminuyó significativamente desde la altura del dosel inferior a la superior en un estudio llevado a cabo en una sala experimental bajo condiciones ambientales controladas. Estos resultados se asociaron con un aumento en el contenido de nitrógeno y carbono desde el dosel inferior hasta el superior, lo que mejoró la capacidad fotosintética y la eficiencia en el uso de la luz en un 20% y 30% respectivamente en la capa de dosel más expuesta a la radiación.

Las variaciones en el comportamiento espectral evidenciadas en la Tabla 4–7 para cada una de las 3 especies, pueden atribuirse a las diversas estrategias de crecimiento desarrolladas por estas especies que se relacionan con la asignación y distribución de nutrientes a través del perfil vertical del dosel (Li & Wang, 2013). Sin embargo, los contenidos de los nutrientes foliares dependen también del estado fenológico de la planta (Coble *et al.*, 2016) y el microclima que se genera en las capas del dosel (Meerdink *et al.*, 2016).

Adicionalmente, las hojas del dosel superior alcanzaron el punto de saturación lumínica, a pesar de que las hojas estaban expuestas a radiaciones solares altas no hay un incremento notorio de la tasa fotosintética. Los cloroplastos han excedido su capacidad de absorción de luz, por tanto, aumentó significativamente la reflectancia en comparación con las otras alturas del dosel, lo que se evidenció en una menor reflectancia y mayor absorción en la altura 1 y 2 de los pigmentos foliares, producto de un incremento considerable en la tasa de fotosíntesis.

En conclusión, las diferencias observadas en la reflectancia media de las 3 especies en el perfil vertical del dosel, autores como Coble *et al.*, (2016) las han relacionado con las estrategias que emplean las plantas para optimizar los procesos metabólicos, incluyendo la fotosíntesis, la cual involucra principalmente nutrientes como carbono y nitrógeno.

No obstante, en el caso de *F. benjamina* donde no se encontraron diferencias estadísticamente significativas ( $P > 0,05$ ) en los espectros tomados en las diferentes alturas de las copas de los árboles, puede relacionarse con una estructura de dosel más cerrado que permite que la luz se distribuya con mayor uniformidad en el perfil vertical de este. Además, en el presente trabajo, se analizaron los espectros completos entre el rango de 400 a 1000 nm, en comparación con el estudio de Gara *et al.*, (2018), en el cual realizaron los análisis de medias parcialmente, correspondientes a un rango espectral específico, donde las diferencias estadísticamente encontradas fueron reportadas para el VIS y/o NIR en particular.

Además, y similar a los resultados obtenidos en este trabajo, Khavaninzadeh *et al.*, (2015) reportaron que la reflectancia de la hoja medida en 3 bandas espectrales (R, G y B) aumentó significativamente desde la altura de dosel inferior hasta la superior en un estudio realizado en un área industrial con altos niveles de contaminación, equivalente a las condiciones ambientales de Medellín. Los autores plantearon que el patrón descendente de la reflectancia desde la parte superior del dosel hasta la inferior estaba relacionado con la degradación de la clorofila, que ocurre debido al estrés generado por diferentes factores bióticos y abióticos, y varía de acuerdo con la altura de la copa por la contaminación del aire.

Por tanto, se verificaron los cambios en los espectros de las hojas como un potencial indicador de la deformación de la vegetación debido a la contaminación del aire, relacionada con la distancia y tipo de fuentes contaminantes (vías). Sin embargo, no se encontraron diferencias estadísticamente significativas ( $P > 0,05$ ) en las propiedades ópticas en la mayoría de las hojas de las 3 especies (Tabla 4–

8). Esto puede estar asociado a la falta de diferentes tipos de hábitats urbanos que ofrecieran un gradiente de niveles de contaminación del aire contrastantes. Diferentes autores llevaron a cabo la evaluación de la calidad del hábitat urbano utilizando las características espectrales de varias especies localizadas en áreas urbanas y suburbanas con niveles de contaminación del aire diferentes (Khavanin *et al.*, 2013; Brackx *et al.*, 2017).

**Tabla 4–8:** Resumen de la variabilidad espectral media de las 3 especies seleccionadas según la clasificación de los árboles en los diferentes sitios de muestreo.

Especies	Sitio	Valor P de la combinación de las posibles posiciones de los árboles		
		Crítica vs. Intermedia	Crítica vs. Baja	Intermedia vs. Baja
<i>E. fusca</i>	UNAL	0,360	0,170	0,360
	EAFIT	0,081	0,050	0,547
	UDEM	-	-	0,071
	TODOS LOS SITIOS	0,360	0,360	0,520
<i>F. benjamina</i>	UNAL	0,240	0,240	0,960
	EAFIT	0,140	-	-
	UDEM	-	-	-
	TODOS LOS SITIOS	0,310	0,280	0,280
<i>T. catappa</i>	UNAL	0,180	0,420	0,420
	EAFIT	0,520	-	-
	UDEM	-	-	-
	TODOS LOS SITIOS	0,270	0,270	0,700

\*En color azul se muestran los valores  $P < 0,05$ , los cuales presentan diferencias estadísticamente significativas.

## 5. Conclusiones y recomendaciones

### 5.1 Conclusiones

Con base en los resultados obtenidos, se concluye que:

- Los datos multiespectrales empleados en el presente estudio presentaron alta variabilidad y no obtuvieron buenos resultados en la clasificación de las especies seleccionadas cuando se tiene un tamaño de muestra bajo.
- La imagen empleada para la extracción de los datos multiespectrales presentó alta resolución espacial, sin embargo, la baja resolución espectral limitó la discriminación de las especies seleccionadas, debido a que cada banda de este tipo de datos corresponde al valor promedio de la respuesta espectral en un intervalo grande del espectro, por tanto, se suprime variaciones sutiles que pueden estar determinando la separabilidad de las especies.
- Los datos hiperespectrales permitieron realizar una selección de las longitudes de onda específicas con las cuales se obtuvo mayor precisión en la clasificación de las especies seleccionadas, en comparación con los resultados observados a partir de los datos multiespectrales.
- Las diferencias estadísticamente significativas observadas en la reflectancia media en el perfil vertical del dosel de las 3 especies seleccionadas demuestran la existencia de un cambio en las características anatómicas y fisiológicas de las hojas dentro de la copa de los árboles.
- No se evidenciaron diferencias estadísticamente significativas en los espectros de reflectancia medidos en las hojas de las especies seleccionadas de acuerdo con la clasificación de los árboles realizada previamente según la distancia y el tipo de fuentes contaminantes (vías).
- Los datos hiperespectrales permiten adquirir una gran cantidad de datos en un período relativamente corto. Sin embargo, su procesamiento es complejo y requiere de un tiempo considerable para su análisis.

- Los resultados obtenidos con el ejercicio de esta investigación indican que la aplicación de técnicas espectrales para la identificación preliminar de especies en inventarios forestales es un tema con mucho potencial, sin embargo, la adquisición de espectros en campo es dispendiosa y el análisis estadístico de la información requiere gran volumen de trabajo que actualmente no se justifica desde la relación costo – beneficio.
- El espectrómetro empleado para tomar los datos hiperespectrales es adecuado para adquirir espectros a escala de hoja y/o individuo, mientras que las imágenes capturadas mediante un sensor remoto poseen mayor potencial para llevar a cabo análisis en extensiones geográficas más amplias, considerando una escala de bosque o paisaje.
- Los datos hiperespectrales son una fuente potencial de información para la identificación de especies forestales a través de taxonomía integrativa, la cual puede ser utilizada con el fin de mejorar la fiabilidad y proporcionar un conocimiento *a priori* de la composición florística en la ejecución de inventarios forestales urbanos.

## 5.2 Recomendaciones

Dado el potencial observado en la clasificación de especies a partir de datos hiperespectrales en el presente estudio, se recomienda en investigaciones posteriores automatizar el proceso de adquisición de espectros desde un vehículo aéreo no tripulado (UAV) o comúnmente denominado dron, con el fin de optimizar el proceso de obtención de la información en vuelos programados a través de un sensor cuyo rango espectral incluya el infrarrojo de onda corta, además del espectro visible e infrarrojo cercano, debido al aporte de las bandas relevantes de dicha región para la discriminación de especies.

Se recomienda realizar futuras investigaciones en las que se caractericen espectralmente otras especies forestales presentes en el área de estudio, los productos generados por este estudio y todos los que se lleven a cabo en el área de estudio deberán conformar una biblioteca de firmas espectrales disponible para diferentes usuarios, con el fin de consolidar una base de datos para la automatización de la búsqueda de especies en el municipio de Medellín.

En esta investigación se evaluó el potencial de clasificación de especies de árboles a partir de datos hiperespectrales y multispectrales utilizando LDA, por tanto, un trabajo futuro deberá medir variables adicionales de importancia para el fenómeno de estudio, como el estado fenológico o edad del árbol, que permiten

comprender la variabilidad de la respuesta espectral, la cual impide o apoya el proceso de discriminación de especies.

Otro punto que se debe considerar en futuros estudios es examinar el efecto de la posición vertical del dosel en los rasgos funcionales y las propiedades espectrales de las hojas en diferentes especies, con el objetivo de comprender la fisiología, funcionamiento y estructura de las plantas frente a diferentes condiciones ambientales.

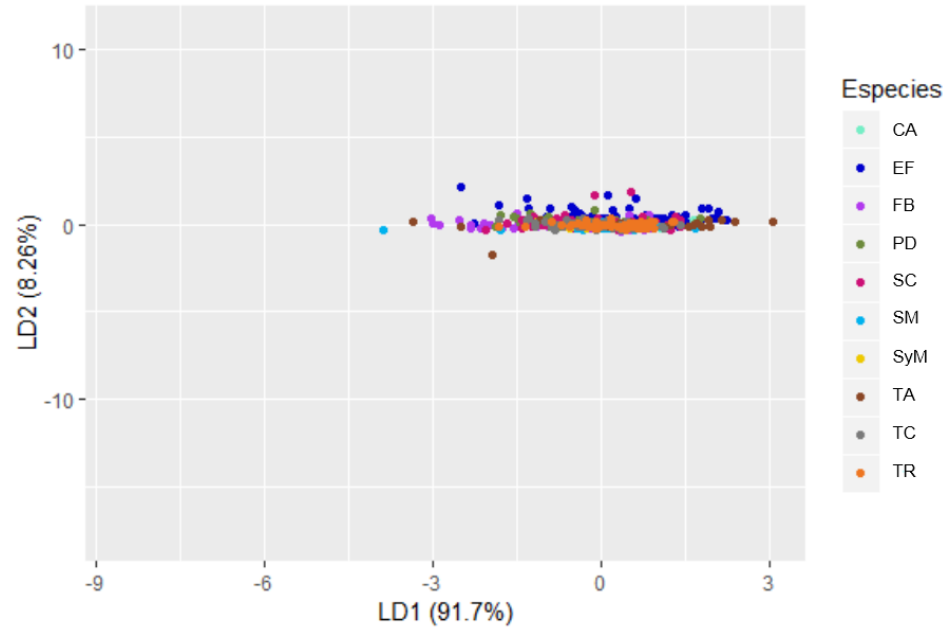
Finalmente, puesto que no se encontraron diferencias estadísticamente significativas en los espectros foliares de acuerdo con la clasificación de los árboles realizada según la distancia y el tipo de fuentes contaminantes (vías), se propone incluir en el área de estudio sitios que evidencien un gradiente de niveles de contaminación del aire contrastantes, por ejemplo, árboles localizados en el perímetro urbano de Medellín y otros en corregimientos como Santa Elena, San Cristóbal o San Sebastián de Palmitas. Esto permitiría valorar la calidad del hábitat a partir de características espectrales de múltiples especies.

## **6. Nota aclaratoria**

Es necesario aclarar que las gráficas elaboradas para este trabajo de investigación a partir de los datos hiperespectrales presentan valores atípicos en la región del infrarrojo cercano (750 – 1.000 nm), los cuales fueron percibidos después de realizar la caracterización y discriminación espectral, debido a que el equipo (espectrómetro portátil) no se encontraba calibrado antes de llevar a cabo la adquisición de los espectros de reflectancia. Sin embargo, la región visible del espectro (400 – 750 nm) evidenció valores de reflectancia dentro de los estándares normales, lo que permitió efectuar correctamente la discriminación espectral de las especies forestales seleccionadas.





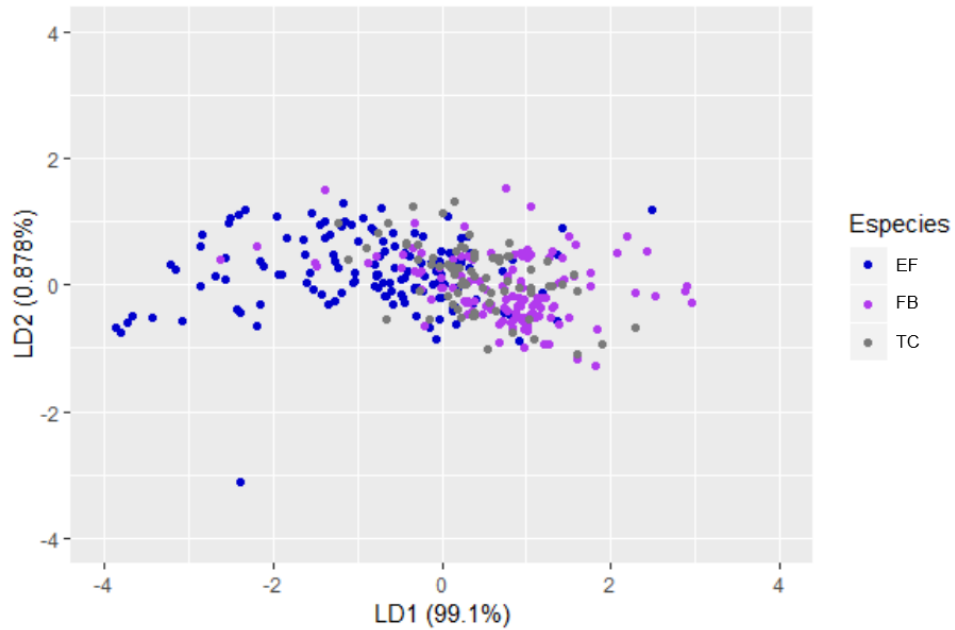


**Figura 7-1:** Análisis Discriminante Lineal con las 10 especies considerando todos los sitios de muestreo y datos transformados con MSC.

Discriminación de las especies *E. fusca*, *F. benjamina* y *T. catappa* utilizando los datos transformados con SNV:

**Tabla 7-2:** Matriz de confusión obtenida a partir del LDA con las especies *E. fusca*, *F. benjamina* y *T. catappa* considerando todos los sitios de muestreo y datos transformados con SNV.

Pronóstico	EF	FB	TC	Total	Precisión (%)
EF	110	33	0	143	76,92
FB	23	103	0	126	81,75
TC	17	56	0	73	0,00
Total	150	192	0	342	62,28

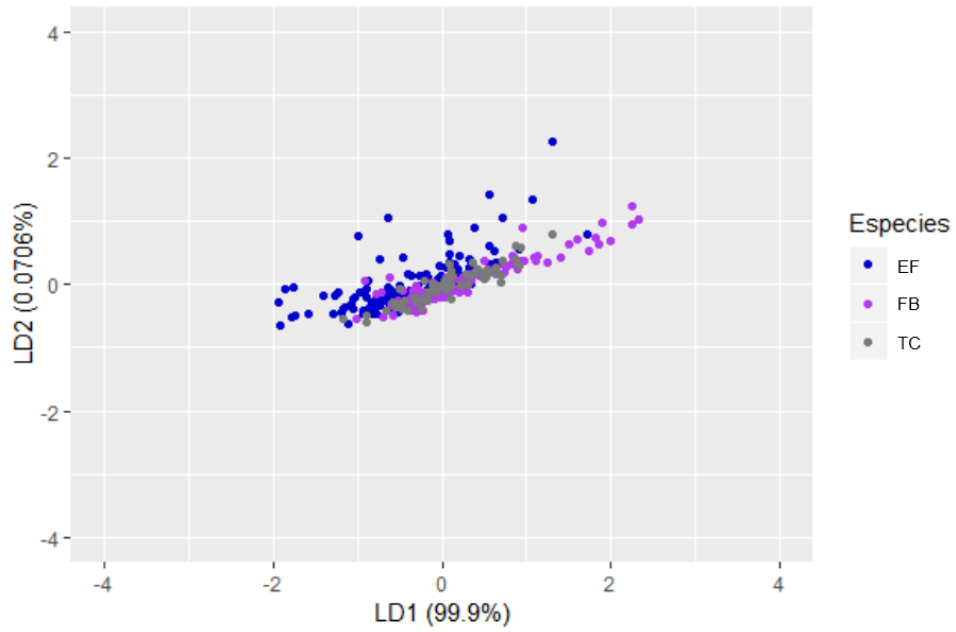


**Figura 7-2:** Análisis Discriminante Lineal con las especies *E. fusca*, *F. benjamina* y *T. catappa* considerando todos los sitios de muestreo y datos transformados con SNV.

Discriminación de las especies *E. fusca*, *F. benjamina* y *T. catappa* utilizando los datos transformados con MSC:

**Tabla 7-3:** Matriz de confusión obtenida a partir del LDA con las especies *E. fusca*, *F. benjamina* y *T. catappa* considerando todos los sitios de muestreo y datos transformados con MSC.

Pronóstico	EF	FB	TC	Total	Precisión (%)
EF	123	20	0	143	86,01
FB	95	31	0	126	24,60
TC	50	23	0	73	0,00
<b>Total</b>	268	74	0	342	45,03



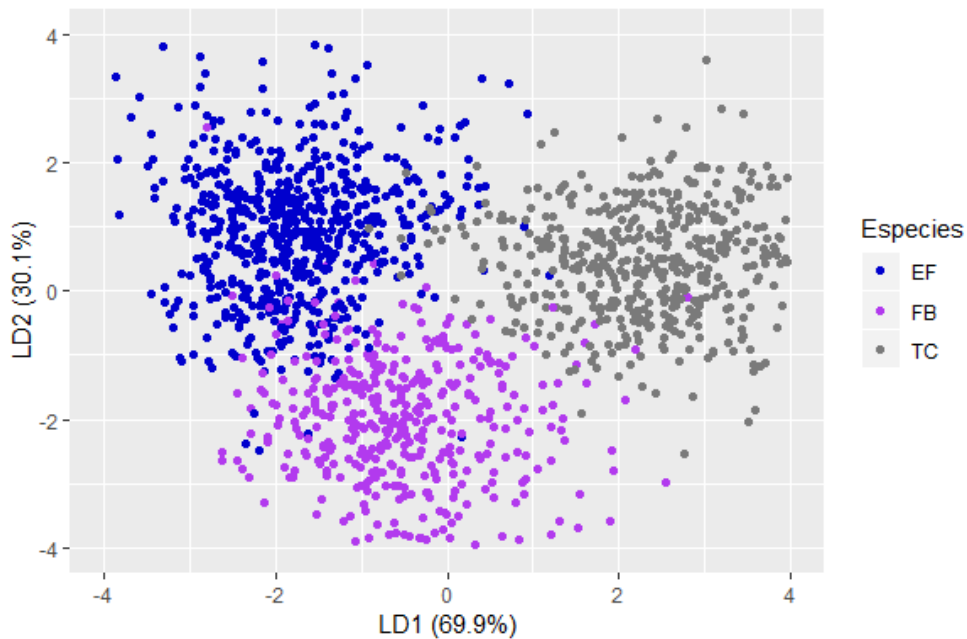
**Figura 7-3:** Análisis Discriminante Lineal con las especies *E. fusca*, *F. benjamina* y *T. catappa* considerando todos los sitios de muestreo y datos transformados con MSC.

## B. Anexo: Análisis Discriminante Lineal a partir de datos hiperespectrales

Discriminación de las especies *E. fusca*, *F. benjamina* y *T. catappa* utilizando los datos transformados con MSC:

**Tabla 7-4:** Matriz de confusión obtenida a partir del LDA con los espectros transformados MSC de las 3 especies seleccionadas considerando todos los sitios de muestreo.

Pronóstico	EF	FB	TC	Total	Precisión (%)
EF	548	73	2	623	87,96
FB	168	217	31	416	52,16
TC	1	41	486	528	92,05
Total	717	331	519	1567	79,83

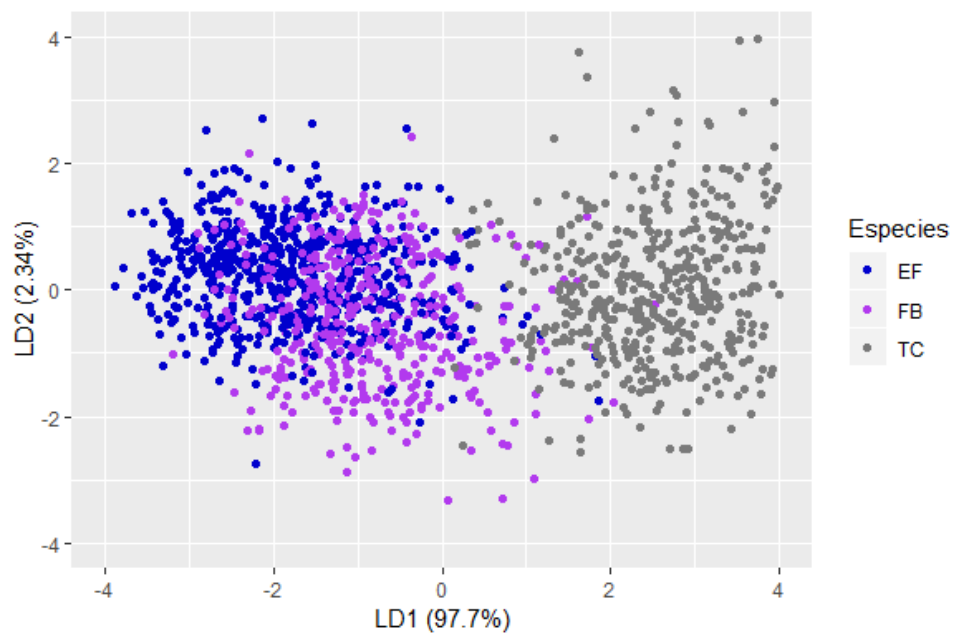


**Figura 7-4:** Análisis Discriminante Lineal con los espectros transformados MSC de las 3 especies seleccionadas considerando todos los sitios de muestreo.

Discriminación de las especies *E. fusca*, *F. benjamina* y *T. catappa* utilizando los datos sin transformar:

**Tabla 7-5:** Matriz de confusión obtenida a partir del LDA con los espectros sin transformar de las 3 especies seleccionadas considerando todos los sitios de muestreo.

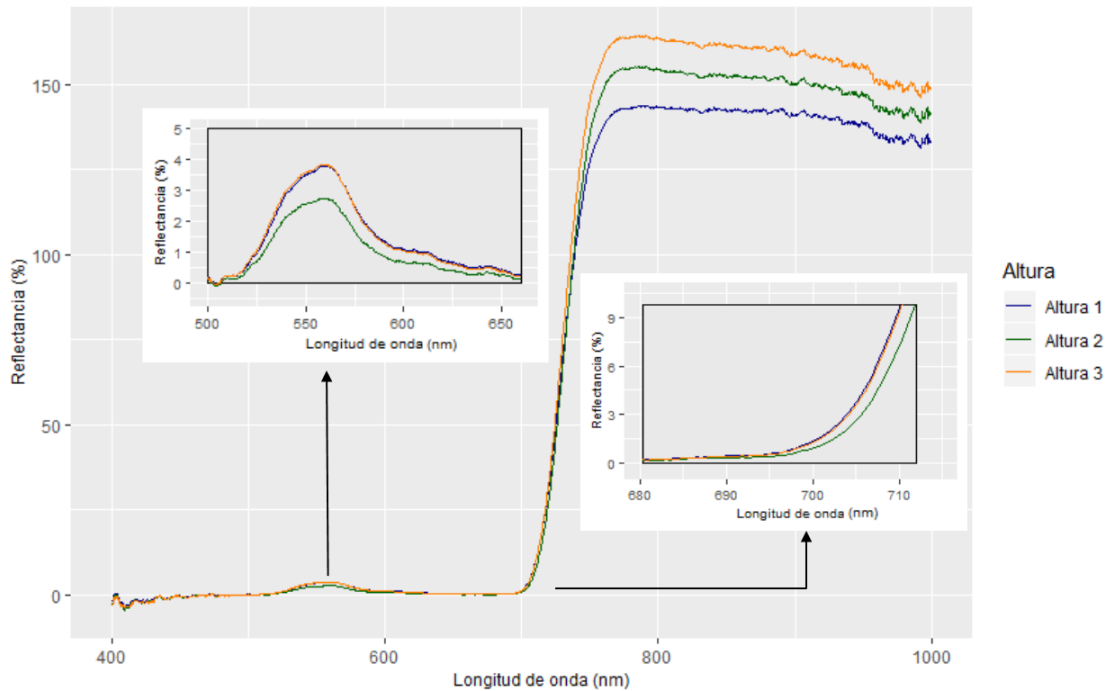
Pronóstico	EF	FB	TC	Total	Precisión (%)
EF	515	103	5	623	82,66
FB	214	184	18	416	44,23
TC	0	13	515	528	97,54
Total	729	300	538	1567	77,47



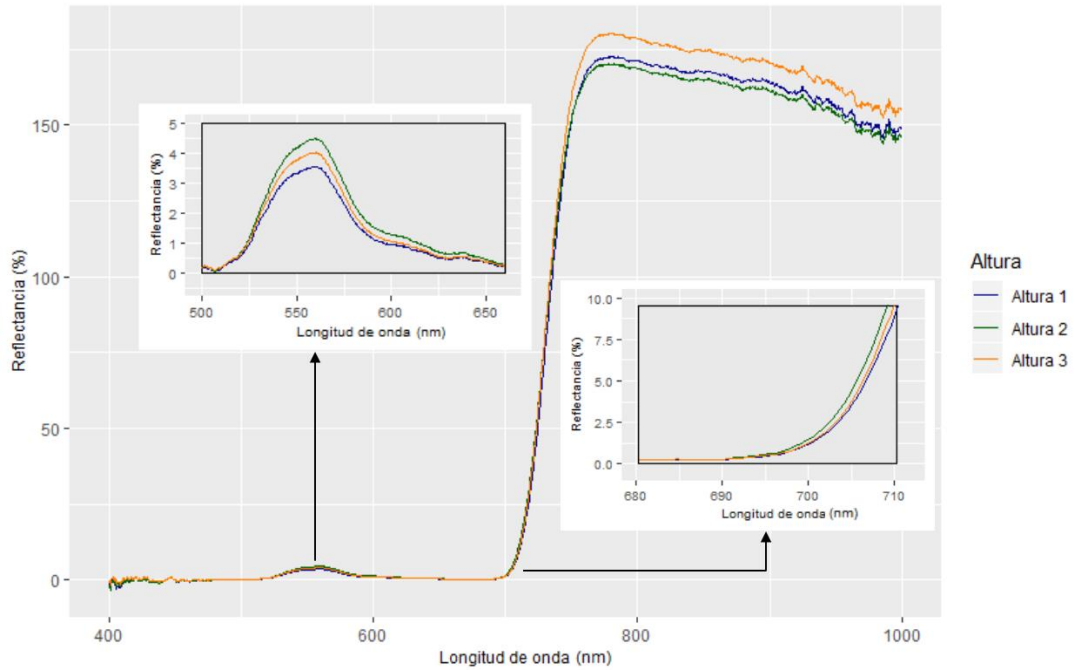
**Figura 7-5:** Análisis Discriminante Lineal con los espectros sin transformar de las 3 especies seleccionadas considerando todos los sitios de muestreo.

## C. Anexo: Gráficas de la variabilidad en la respuesta espectral considerando la posición del dosel

### *E. fusca*

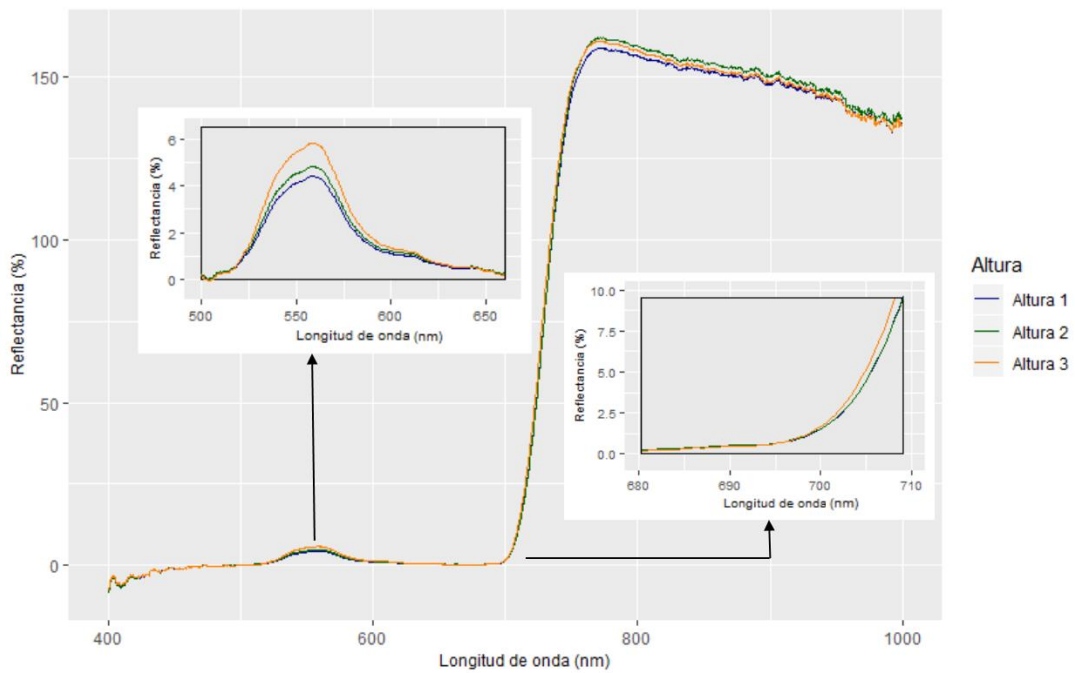


**Figura 7-6:** Respuesta espectral media en cada posición del dosel para la especie *E. fusca* en el sitio de muestreo UNAL.



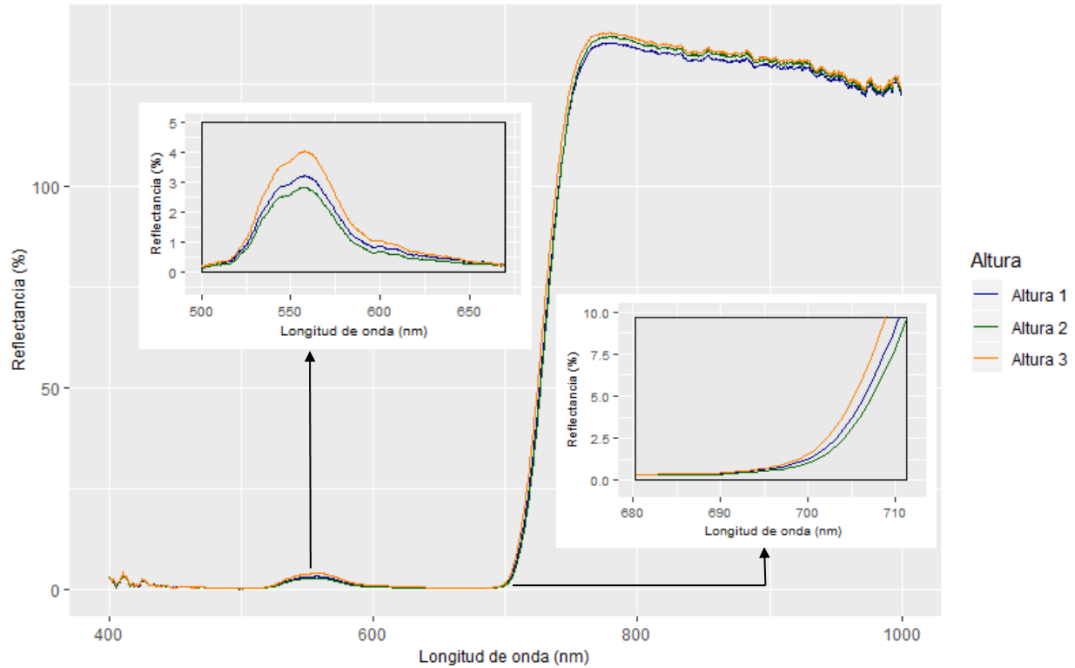
**Figura 7-7:** Respuesta espectral media en cada posición del dosel para la especie *E. fusca* en el sitio de muestreo EAFIT.

### *F. benjamina*



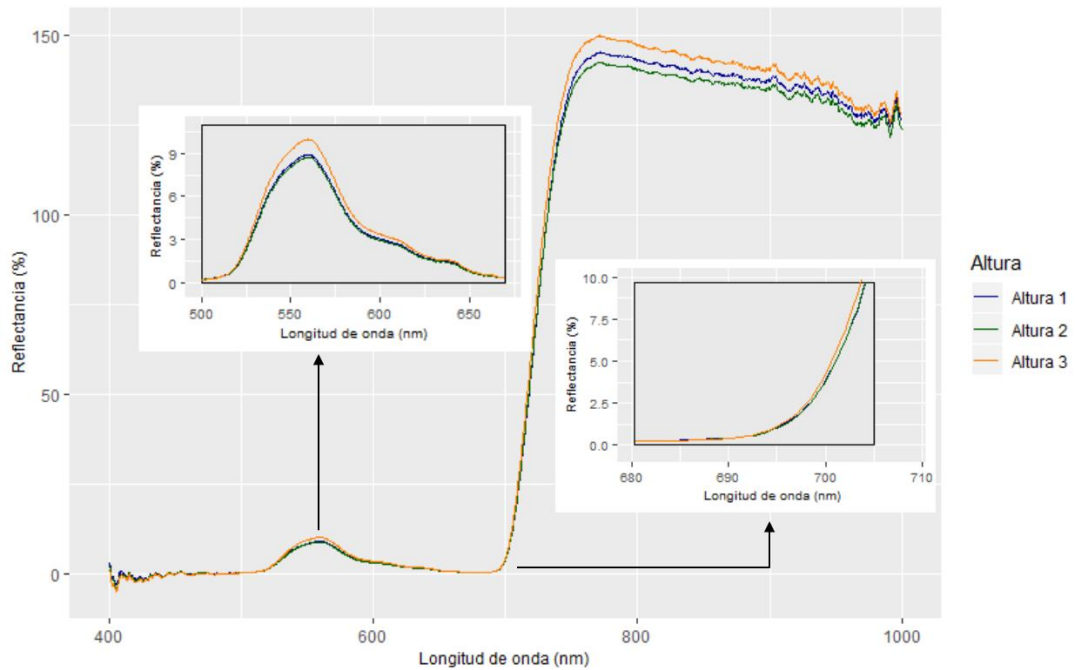
**Figura 7-8:** Respuesta espectral media en cada posición del dosel para la especie *F. benjamina* en el sitio de muestreo UNAL.



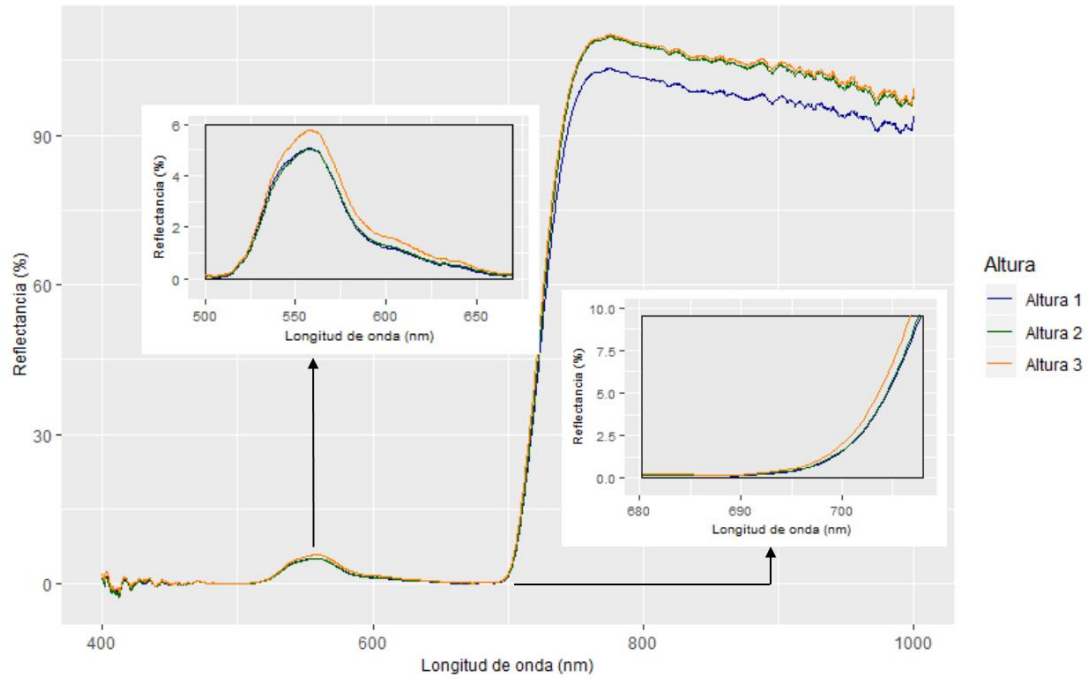


**Figura 7-9:** Respuesta espectral media en cada posición del dosel para la especie *F. benjamina* en el sitio de muestreo EAFIT.

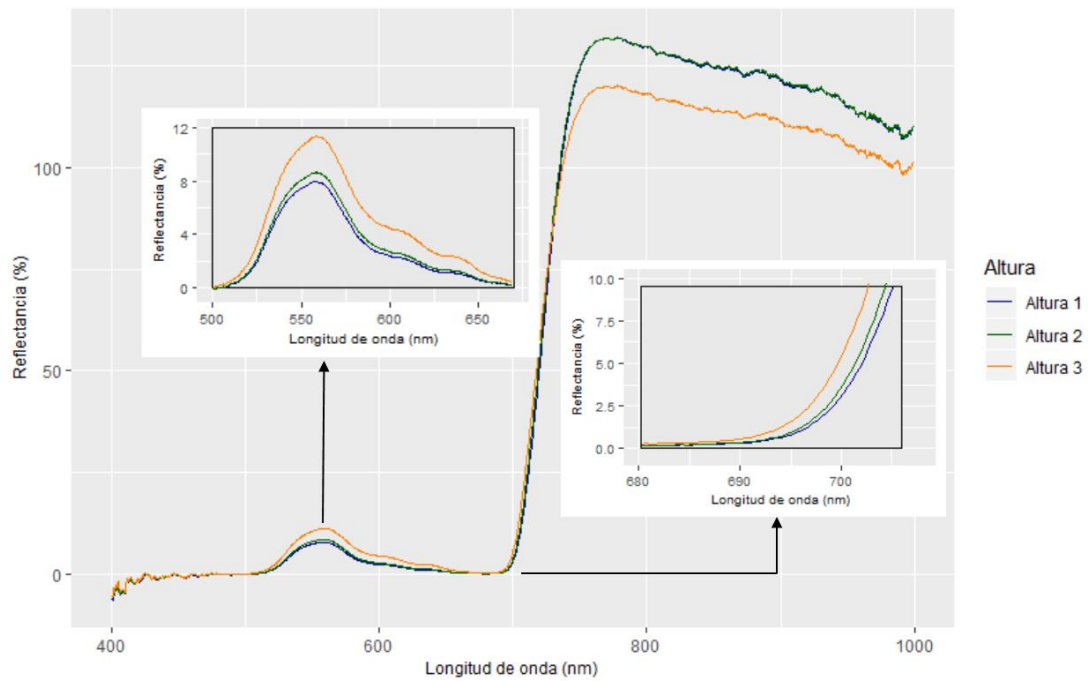
### *T. catappa*



**Figura 7-10:** Respuesta espectral media en cada posición del dosel para la especie *T. catappa* en el sitio de muestreo UNAL.



**Figura 7-11:** Respuesta espectral media en cada posición del dosel para la especie *T. catappa* en el sitio de muestreo EAFIT.



**Figura 7-12:** Respuesta espectral media en cada posición del dosel para la especie *T. catappa* en el sitio de muestreo UDEM.

## Bibliografía

- Adam, E., Mutanga, O., & Rugege, D. (2010). Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation: a review. *Wetlands Ecology and Management*, 18(3), 281–296. <https://doi.org/10.1007/s11273-009-9169-z>
- Alcaldía de Medellín. (2011). *Árboles nativos y ciudad, aportes a la silvicultura urbana de Medellín* (Primera ed). Medellín: Fondo Editorial Jardín Botánico de Medellín.
- Alonzo, M., Bookhagen, B., & Roberts, D. A. (2014). Urban tree species mapping using hyperspectral and lidar data fusion. *Remote Sensing of Environment*, 148(Supplement C), 70–83. <https://doi.org/https://doi.org/10.1016/j.rse.2014.03.018>
- Alonzo, M., McFadden, J. P., Nowak, D. J., & Roberts, D. A. (2016). Mapping urban forest structure and function using hyperspectral imagery and lidar data. *Urban Forestry & Urban Greening*, 17, 135–147. <https://doi.org/https://doi.org/10.1016/j.ufug.2016.04.003>
- Araque, L; Jiménez, A. (2009). Caracterización de firma espectral a partir de sensores remotos para el manejo de sanidad vegetal en el cultivo de palma de aceite | Revista Palmas. *PALMAS*, 30(3), 63–79. Retrieved from <https://publicaciones.fedepalma.org/index.php/palmas/article/view/1455>
- Área Metropolitana del Valle de Aburrá, & Universidad Nacional de Colombia. (2015). *Guía para el manejo del arbolado urbano en el Valle de Aburrá*. (F. Moreno & C. H. Hoyos, Eds.). Medellín.
- Asner, G. P., Martin, R. E., Carranza-Jiménez, L., Sinca, F., Tupayachi, R., Anderson, C. B., & Martinez, P. (2014). Functional and biological diversity of foliar spectra in tree canopies throughout the Andes to Amazon region. *New Phytologist*, 204(1), 127–139. <https://doi.org/10.1111/nph.12895>
- Avtar, R., Kumar, P., Oono, A., Saraswat, C., Dorji, S., & Hlaing, Z. (2017). Potential application of remote sensing in monitoring ecosystem services of forests, mangroves and urban areas. *Geocarto International*, 32(8), 874–885. <https://doi.org/10.1080/10106049.2016.1206974>
- Barnes, R. J., Dhanoa, M. S., & Lister, S. J. (1989). Standard Normal Variate Transformation and De-Trending of Near-Infrared Diffuse Reflectance Spectra. *Applied Spectroscopy*, 43(5), 772–777. <https://doi.org/10.1366/0003702894202201>
- Baró, F., Chaparro, L., Gómez-Baggethun, E., Langemeyer, J., Nowak, D. J., & Terradas,

- J. (2014). Contribution of Ecosystem Services to Air Quality and Climate Change Mitigation Policies: The Case of Urban Forests in Barcelona, Spain. *AMBIO*, 43(4), 466–479. <https://doi.org/10.1007/s13280-014-0507-x>
- Bergo, M. C. J., Pastore, T. C. M., Coradin, V. T. R., Wiedenhoef, A. C., & Braga, J. W. B. (2016). NIRS IDENTIFICATION OF SWIETENIA MACROPHYLLA IS ROBUST ACROSS SPECIMENS FROM 27 COUNTRIES. *IAWA Journal*, 37(3), 420–430. <https://doi.org/10.1163/22941932-20160144>
- Botelho, I. M. S. (2017). *Uso da espectroscopia no infravermelho próximo (NIR) como ferramenta no reconhecimento de espécies de Burseraceae de diferentes regiões geográficas da Amazônia*. Dissertação de mestrado, Instituto Nacional de Pesquisas da Amazônia/ Universidade Federal do Amazonas.
- Botero, J. M. (2009). *Determinación del nivel foliar de nutrientes mediante espectroscopía de reflectancia*. Universidad Nacional de Colombia.
- Brackx, M., Verhelst, J., Scheunders, P., & Samson, R. (2017). On the use of dorsiventral reflectance asymmetry of hornbeam (*Carpinus betulus* L.) leaves in air pollution estimation. *Environmental Monitoring and Assessment*, 189(9), 472. <https://doi.org/10.1007/s10661-017-6168-z>
- Braga, J. W. B., Pastore, T. C. M., Coradin, V. T. R., Camargos, J. A. A., & da Silva, A. R. (2011). The use of near Infrared Spectroscopy to Identify solid wood Specimens of *Swietenia Macrophylla* (Cites Appendix II). *IAWA Journal*, 32(2), 285–296. <https://doi.org/10.1163/22941932-90000058>
- Camacho, M. (2014). *Metodología para identificar frailejón (*Espeletia* sp) usando las características pictórico-morfológicas de imágenes de satélite para su patronamiento espectral*. Universidad Nacional de Colombia.
- Carvalho, S., Schlerf, M., van der Putten, W. H., & Skidmore, A. K. (2013). Hyperspectral reflectance of leaves and flowers of an outbreak species discriminates season and successional stage of vegetation. *International Journal of Applied Earth Observation and Geoinformation*, 24, 32–41. <https://doi.org/10.1016/J.JAG.2013.01.005>
- Castro-Esau, K. L., Sánchez-Azofeifa, G. A., Rivard, B., Wright, S. J., & Quesada, M. (2006). Variability in leaf optical properties of mesoamerican trees and the potential for species classification. *American Journal of Botany*, 93(4), 517–530. <https://doi.org/10.3732/ajb.93.4.517>
- Chen, Q. (2011). Comparison of Worldview-2 and IKONOS-2 Imagery for Identifying Tree Species in the Habitat of an Endangered Bird Species in Hawaii; 8-Band Research Challenge. *DigitalGlobe: Longmont, CO, USA*.
- Chuvieco, E. (2000). *Fundamentos de teledetección espacial* (3a edición). Madrid: RIALP.
- Chuvieco, E. (2008). *TELEDETECCIÓN AMBIENTAL. La observación de la tierra desde el espacio*. (Ariel, Ed.) (3a edición). Barcelona.

- Clark, M. L., Roberts, D. A., & Clark, D. B. (2005). Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales. *Remote Sensing of Environment*, 96(3–4), 375–398. <https://doi.org/10.1016/J.RSE.2005.03.009>
- Clark, M. L., Roberts, D. A., Clark, M. L., & Roberts, D. A. (2012). Species-Level Differences in Hyperspectral Metrics among Tropical Rainforest Trees as Determined by a Tree-Based Classifier. *Remote Sensing*, 4(6), 1820–1855. <https://doi.org/10.3390/rs4061820>
- Coble, A. P., VanderWall, B., Mau, A., & Cavaleri, M. A. (2016). How vertical patterns in leaf traits shift seasonally and the implications for modeling canopy photosynthesis in a temperate deciduous forest. *Tree Physiology*, 36(9), 1077–1091.
- Cortés, V., Cubero, S., Blasco, J., Aleixos, N., & Talens, P. (2019). In-line Application of Visible and Near-Infrared Diffuse Reflectance Spectroscopy to Identify Apple Varieties. *Food and Bioprocess Technology*, 12(6), 1021–1030. <https://doi.org/10.1007/s11947-019-02268-0>
- Dittmann, S., Thiessen, E., & Hartung, E. (2017). Applicability of different non-invasive methods for tree mass estimation: A review. *Forest Ecology and Management*, 398, 208–215. <https://doi.org/http://dx.doi.org/10.1016/j.foreco.2017.05.013>
- Dudley, K. L., Dennison, P. E., Roth, K. L., Roberts, D. A., & Coates, A. R. (2015). A multi-temporal spectral library approach for mapping vegetation species across spatial and temporal phenological gradients. *Remote Sensing of Environment*, 167, 121–134. <https://doi.org/10.1016/j.rse.2015.05.004>
- Durgante, F. M., Higuchi, N., Almeida, A., & Vicentini, A. (2013). Species Spectral Signature: Discriminating closely related plant species in the Amazon with Near-Infrared Leaf-Spectroscopy. *Forest Ecology and Management*, 291, 240–248. <https://doi.org/10.1016/j.foreco.2012.10.045>
- EIA, U. (2019). *Catálogo Virtual de Flora del Valle de Aburrá*. Medellín. Retrieved from <https://catalogofloravalleaburra.eia.edu.co/>
- Encina-Zelada, C., Cadavez, V., Pereda, J., Gómez-Pando, L., Salvá-Ruiz, B., Teixeira, J. A., ... Gonzales-Barron, U. (2017). Estimation of composition of quinoa (*Chenopodium quinoa* Willd.) grains by Near-Infrared Transmission spectroscopy. *LWT - Food Science and Technology*, 79, 126–134. <https://doi.org/https://doi.org/10.1016/j.lwt.2017.01.026>
- Fassnacht, F. E., Latifi, H., Stereńczak, K., Modzelewska, A., Lefsky, M., Waser, L. T., ... Ghosh, A. (2016). Review of studies on tree species classification from remotely sensed data. *Remote Sensing of Environment*, 186, 64–87. <https://doi.org/http://dx.doi.org/10.1016/j.rse.2016.08.013>
- Fassnacht, F. E., Neumann, C., Forster, M., Buddenbaum, H., Ghosh, A., Clasen, A., ... Koch, B. (2014). Comparison of feature reduction algorithms for classifying tree species with hyperspectral data on three central european test sites. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(6), 2547–2561. <https://doi.org/10.1109/JSTARS.2014.2329390>

- Filzmoser, P., & Gschwandtner, M. (2017). mvoutlier: Multivariate Outlier Detection Based on Robust Methods.
- Galenieks, A. (2017). Importance of urban street tree policies: A Comparison of neighbouring Southern California cities. *Urban Forestry & Urban Greening*, 22, 105–110. <https://doi.org/http://doi.org/10.1016/j.ufug.2017.02.004>
- Gara, T., Darvishzadeh, R., Skidmore, A., Wang, T., Gara, T. W., Darvishzadeh, R., ... Wang, T. (2018). Impact of Vertical Canopy Position on Leaf Spectral Properties and Traits across Multiple Species. *Remote Sensing*, 10(2), 346. <https://doi.org/10.3390/rs10020346>
- Ghosh, A., Fassnacht, F. E., Joshi, P. K., & Kochb, B. (2014). A framework for mapping tree species combining hyperspectral and LiDAR data: Role of selected classifiers and sensor across three spatial scales. *International Journal of Applied Earth Observation and Geoinformation*, 26(1), 49–63. <https://doi.org/10.1016/j.jag.2013.05.017>
- Gini, R., Sona, G., Ronchetti, G., Passoni, D., & Pinto, L. (2018). Improving Tree Species Classification Using UAS Multispectral Images and Texture Measures. *ISPRS International Journal of Geo-Information*. <https://doi.org/10.3390/ijgi7080315>
- Goetz, A. F. H., Vane, G., Solomon, J. E., & Rock, B. N. (1985). Imaging Spectrometry for Earth Remote Sensing. *Science*, 228(4704), 1147–1153. <https://doi.org/10.1126/science.228.4704.1147>
- Guindon, L., Bernier, P. Y., Beaudoin, A., Pouliot, D., Villemaire, P., Hall, R. J., ... St-Amant, R. (2014). Annual mapping of large forest disturbances across Canada's forests using 250 m MODIS imagery from 2000 to 2011. *Canadian Journal of Forest Research*, 44(12), 1545–1554. <https://doi.org/10.1139/cjfr-2014-0229>
- Hadlich, H. L., Durgante, F. M., dos Santos, J., Higuchi, N., Chambers, J. Q., & Vicentini, A. (2018). Recognizing Amazonian tree species in the field using bark tissues spectra. *Forest Ecology and Management*, 427, 296–304. <https://doi.org/10.1016/J.FORECO.2018.06.002>
- Harrison, D., Rivard, B., & Sánchez-Azofeifa, A. (2018). Classification of tree species based on longwave hyperspectral data from leaves, a case study for a tropical dry forest. *International Journal of Applied Earth Observation and Geoinformation*, 66, 93–105. <https://doi.org/10.1016/J.JAG.2017.11.009>
- Hawkins, D. M. (1980). *Identification of outliers* (Vol. 11). Springer.
- Heikkinen, V., Tokola, T., Parkkinen, J., Korpela, I., & Jääskeläinen, T. (2010). Simulated multispectral imagery for tree species classification using support vector machines. *IEEE Transactions on Geoscience and Remote Sensing*, 48(3 PART2), 1355–1364. <https://doi.org/10.1109/TGRS.2009.2032239>
- Hovi, A., Raitio, P., & Rautiainen, M. (2017). A spectral analysis of 25 boreal tree species. *Silva Fennica*, 51(4). <https://doi.org/10.14214/sf.7753>

- Immitzer, M., Atzberger, C., & Koukal, T. (2012). Tree Species Classification with Random Forest Using Very High Spatial Resolution 8-Band WorldView-2 Satellite Data. *Remote Sensing*. <https://doi.org/10.3390/rs4092661>
- Jensen, J. R. (2007). *Remote Sensing of the Environment An Earth Resource Perspective* (Second Edi). New Jersey: Prentice Hall. Retrieved from [https://s3.amazonaws.com/academia.edu.documents/7106414/306542412.pdf?response-content-disposition=inline%3Bfilename%3DRemote\\_sensing\\_of\\_the\\_environment.pdf&X-Amz-Algorithm=AWS4-HMAC-SHA256&X-Amz-Credential=AKIAIWOWYYGZ2Y53UL3A%2F20190703%2Fus-east-1%2Fs](https://s3.amazonaws.com/academia.edu.documents/7106414/306542412.pdf?response-content-disposition=inline%3Bfilename%3DRemote_sensing_of_the_environment.pdf&X-Amz-Algorithm=AWS4-HMAC-SHA256&X-Amz-Credential=AKIAIWOWYYGZ2Y53UL3A%2F20190703%2Fus-east-1%2Fs)
- Jensen, R. R., Hardin, P. J., & Hardin, A. J. (2012). Classification of urban tree species using hyperspectral imagery. *Geocarto International*, 27(July 2014), 443–458. <https://doi.org/10.1080/10106049.2011.638989>
- Ke, Y., Quackenbush, L. J., & Im, J. (2010). Synergistic use of QuickBird multispectral imagery and LIDAR data for object-based forest species classification. *Remote Sensing of Environment*, 114(6), 1141–1154. <https://doi.org/10.1016/j.rse.2010.01.002>
- Khavanin Zadeh, A. R., Veroustraete, F., Buytaert, J. A. N., Dirckx, J., & Samson, R. (2013). Assessing urban habitat quality using spectral characteristics of Tilia leaves. *Environmental Pollution*, 178, 7–14. <https://doi.org/https://doi.org/10.1016/j.envpol.2013.02.021>
- Khavaninzadeh, A. R., Veroustraete, F., Van Wittenberghe, S., Verrelst, J., & Samson, R. (2015). Leaf reflectance variation along a vertical crown gradient of two deciduous tree species in a Belgian industrial habitat. *Environmental Pollution*, 204, 324–332. <https://doi.org/10.1016/j.envpol.2015.05.028>
- Khorrarn, S., Koch, F., Wiele, C. van der, & Nelson, S. (2012). Remote sensing. *Springer Science & Business Media*. Retrieved from <https://books.google.es/books?hl=es&lr=&id=QpdQut59MPsC&oi=fnd&pg=PR5&ots=iX3hqGi2al&sig=hFmx6Q0wgK22eFky8xm9Qsk8lmE>
- Kira, K., Aaii, L. R., & 1992, undefined. (n.d.). The feature selection problem: Traditional methods and a new algorithm. *New.aaai.org*. Retrieved from <http://new.aaai.org/Library/AAAI/1992/aaai92-020.php>
- Korpela, I., Heikkinen, V., Honkavaara, E., Rohrbach, F., & Tokola, T. (2011). Variation and directional anisotropy of reflectance at the crown scale — Implications for tree species classification in digital aerial images. *Remote Sensing of Environment*, 115(8), 2062–2074. <https://doi.org/https://doi.org/10.1016/j.rse.2011.04.008>
- Kumar, L., Schmidt, K., Dury, S., & Skidmore, A. (2002). Imaging Spectrometry and Vegetation Science. In *Imaging spectrometry* (pp. 111–155). Dordrecht: Springer, Dordrecht. [https://doi.org/10.1007/978-0-306-47578-8\\_5](https://doi.org/10.1007/978-0-306-47578-8_5)
- Lang, C., Almeida, D. R. A., & Costa, F. R. C. (2017). Discrimination of taxonomic identity at species, genus and family levels using Fourier Transformed Near-Infrared

- Spectroscopy (FT-NIR). *Forest Ecology and Management*, 406, 219–227.  
<https://doi.org/10.1016/J.FORECO.2017.09.003>
- Lang, C., Costa, F. R. C., Camargo, J. L. C., Durgante, F. M., & Vicentini, A. (2015). Near Infrared Spectroscopy Facilitates Rapid Identification of Both Young and Mature Amazonian Tree Species. *PLOS ONE*, 10(8), e0134521.  
<https://doi.org/10.1371/journal.pone.0134521>
- Larrahondo Aguilar, J. E., & Palau, F. J. (2001). *Espectroscopia de infrarrojo cercano - nir principios generales, determinación de nitrógeno en tejido foliar de caña de azúcar, calibración y validación del equipo Bran y Luebbe, modelo infraalyzer 500*. CENICAÑA. Retrieved from <http://ceiba.agro.uba.ar/cgi-bin/koha/opac-detail.pl?biblionumber=10542>
- Leandro, J. G. R., Gonzaga, F. B., & Latorraca, J. V. de F. (2019). Discrimination of wood species using laser-induced breakdown spectroscopy and near-infrared reflectance spectroscopy. *Wood Science and Technology*. <https://doi.org/10.1007/s00226-019-01119-7>
- Li, H., Zhao, C., Huang, W., & Yang, G. (2013). Non-uniform vertical nitrogen distribution within plant canopy and its estimation by remote sensing: A review. *Field Crops Research*, 142, 75–84. <https://doi.org/https://doi.org/10.1016/j.fcr.2012.11.017>
- Li, P., & Wang, Q. (2013). Developing and validating novel hyperspectral indices for leaf area index estimation: Effect of canopy vertical heterogeneity. *Ecological Indicators*, 32, 123–130. <https://doi.org/10.1016/j.ecolind.2013.03.025>
- LI, Z.-F., XU, G.-J., WANG, J.-J., DU, G.-R., CAI, W.-S., & SHAO, X.-G. (2016). Outlier Detection for Multivariate Calibration in Near Infrared Spectroscopic Analysis by Model Diagnostics. *Chinese Journal of Analytical Chemistry*, 44(2), 305–309.  
[https://doi.org/https://doi.org/10.1016/S1872-2040\(16\)60907-6](https://doi.org/https://doi.org/10.1016/S1872-2040(16)60907-6)
- Lin, L., Meng, Y., Yue, A., Yuan, Y., Liu, X., Chen, J., ... Chen, J. (2016). A spatio-temporal model for forest fire detection using HJ-IRS satellite data. *Remote Sensing*, 8(5). <https://doi.org/10.3390/rs8050403>
- Liu, S., Peng, Y., Du, W., Le, Y., & Li, L. (2015). Remote estimation of leaf and canopy water content in winter wheat with different vertical distribution of water-related properties. *Remote Sensing*, 7(4), 4626–4650. <https://doi.org/10.3390/rs70404626>
- Liu, Y. J., Tran, T., Postma, G., Buydens, L. M. C., & Jansen, J. (2018). Estimating the number of components and detecting outliers using Angle Distribution of Loading Subspaces (ADLS) in PCA analysis. *Analytica Chimica Acta*, 1020, 17–29.  
<https://doi.org/https://doi.org/10.1016/j.aca.2018.03.044>
- Lu, D., & Weng, Q. (2007). A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing*, 28(5), 823–870.
- Martens, H., Jensen, S. A., & Geladi, P. (1983). Multivariate linearity transformation for



- near-infrared reflectance spectrometry. In *Proceedings of the Nordic symposium on applied statistics* (pp. 205–234). Stokkand Forlag Publishers Stavanger, Norway.
- Masek, J. G., Hayes, D. J., Joseph Hughes, M., Healey, S. P., & Turner, D. P. (2015). The role of remote sensing in process-scaling studies of managed forest ecosystems. *Forest Ecology and Management*, 355, 109–123. <https://doi.org/http://dx.doi.org/10.1016/j.foreco.2015.05.032>
- Meerdink, S. K., Roberts, D. A., King, J. Y., Roth, K. L., Dennison, P. E., Amaral, C. H., & Hook, S. J. (2016). Linking seasonal foliar traits to VSWIR-TIR spectroscopy across California ecosystems. *Remote Sensing of Environment*, 186, 322–338. <https://doi.org/10.1016/j.rse.2016.08.003>
- MILTON, E. J. (1987). Review Article Principles of field spectroscopy. *International Journal of Remote Sensing*, 8(12), 1807–1827. <https://doi.org/10.1080/01431168708954818>
- Morales, L., & Varón, T. (2006). *Árboles ornamentales en el Valle de Aburrá. Elementos de Manejo* (Primera ed). Medellín.
- Östberg, J., Delshammar, T., Wiström, B., & Nielsen, A. B. (2013). Grading of Parameters for Urban Tree Inventories by City Officials, Arborists, and Academics Using the Delphi Method. *Environmental Management*, 51(3), 694–708. <https://doi.org/10.1007/s00267-012-9973-8>
- Parazoo, N. C., Bowman, K., Fisher, J. B., Frankenberg, C., Jones, D. B. A., Cescatti, A., ... Montagnani, L. (2014). Terrestrial gross primary production inferred from satellite fluorescence and vegetation models. *Global Change Biology*, 20(10), 3103–3121. <https://doi.org/10.1111/gcb.12652>
- Paz-Kagan, T., Silver, M., Panov, N., & Karnieli, A. (2019). Multispectral approach for identifying invasive plant species based on flowering phenology characteristics. *Remote Sensing*, 11(8). <https://doi.org/10.3390/rs11080930>
- Peñuelas, J., & Filella, I. (1998). Visible and near-infrared reflectance techniques for diagnosing plant physiological status. *Trends in Plant Science*, 3(4), 151–156. [https://doi.org/10.1016/S1360-1385\(98\)01213-8](https://doi.org/10.1016/S1360-1385(98)01213-8)
- Pflugmacher, D., Cohen, W. B., & E. Kennedy, R. (2012). Using Landsat-derived disturbance history (1972–2010) to predict current forest structure. *Remote Sensing of Environment*, 122, 146–165. <https://doi.org/http://dx.doi.org/10.1016/j.rse.2011.09.025>
- Ploton, P., Pélissier, R., Proisy, C., Flavenot, T., Barbier, N., Rai, S. N., & Coutron, P. (2012). Assessing aboveground tropical forest biomass using Google Earth canopy images. *Ecological Applications*, 22(3), 993–1003. <https://doi.org/10.1890/11-1606.1>
- Pu, R., & Landry, S. (2012). A comparative analysis of high spatial resolution IKONOS and WorldView-2 imagery for mapping urban tree species. *Remote Sensing of Environment*, 124(Supplement C), 516–533. <https://doi.org/https://doi.org/10.1016/j.rse.2012.06.011>

- Rinnan, Å., Berg, F. van den, & Engelsen, S. B. (2009). Review of the most common pre-processing techniques for near-infrared spectra. *TrAC Trends in Analytical Chemistry*, 28(10), 1201–1222. <https://doi.org/https://doi.org/10.1016/j.trac.2009.07.007>
- Sánchez-Azofeifa, G. A., Castro, K., Wright, S. J., Gamon, J., Kalacska, M., Rivard, B., ... Feng, J. L. (2009). Differences in leaf traits, leaf internal structure, and spectral reflectance between two communities of lianas and trees: Implications for remote sensing in tropical environments. *Remote Sensing of Environment*, 113(10), 2076–2088. <https://doi.org/https://doi.org/10.1016/j.rse.2009.05.013>
- Secretaría de Medio Ambiente de Medellín. (2015). *Manual de silvicultura urbana para Medellín– Gestión, Planeación y manejo de la infraestructura verde* (Segunda ed). Medellín: Fondo Editorial Jardín Botánico de Medellín.
- Secretaría de Movilidad. Alcaldía de Medellín. Resolución 2585 de 2016 (2016).
- Serbin, S. P., Dillaway, D. N., Kruger, E. L., & Townsend, P. A. (2012). Leaf optical properties reflect variation in photosynthetic metabolism and its sensitivity to temperature. *Journal of Experimental Botany*, 63(1), 489–502. <https://doi.org/10.1093/jxb/err294>
- Silva, D. C., Pastore, T. C. M., Soares, L. F., de Barros, F. A. S., Bergo, M. C. J., Coradin, V. T. H., ... Braga, J. W. B. (2018). Determination of the country of origin of true mahogany (*Swietenia macrophylla* King) wood in five Latin American countries using handheld NIR devices and multivariate data analysis. *Holzforschung*, 72(7), 521–530. <https://doi.org/10.1515/hf-2017-0160>
- Strunk, J. L., Mills, J. R., Ries, P., Temesgen, H., & Jeroue, L. (2016). An urban forest-inventory-and-analysis investigation in Oregon and Washington. *Urban Forestry & Urban Greening*, 18, 100–109. <https://doi.org/http://doi.org/10.1016/j.ufug.2016.04.006>
- Suarez, L., Restrepo-Coupe, N., Hueni, A., Chisholm, L. A. (2015). Vegetation spectroscopy. In & S. J. Held, S. Phinn, M. Soto-Berelov (Ed.), *AusCover Good Practice Guidelines: A technical handbook supporting calibration and validation activities of remotely sensed data product* (pp. 221–233).
- Thennadil, S. N., Dewar, M., Herdsman, C., Nordon, A., & Becker, E. (2018). Automated weighted outlier detection technique for multivariate data. *Control Engineering Practice*, 70, 40–49. <https://doi.org/https://doi.org/10.1016/j.conengprac.2017.09.018>
- Tovar-Corzo, G. (2013). An approximation urbanistic silviculture in Colombia. *Bitacora Urbano Territorial*, 22(1), 119–136. Retrieved from <https://www.scopus.com/inward/record.uri?eid=2-s2.0-84890545275&partnerID=40&md5=8ce5825ba729e546ab655c206dbc1202>
- Ullah, S., Skidmore, A. K., Ramoelo, A., Groen, T. A., Naeem, M., & Ali, A. (2014). Retrieval of leaf water content spanning the visible to thermal infrared spectra. *ISPRS Journal of Photogrammetry and Remote Sensing*, 93, 56–64.

<https://doi.org/10.1016/j.isprsjprs.2014.04.005>

- Van Deventer, H., Azong Cho, M., Mutanga, O., Naidoo, L., & Dudeni-Tlhone, N. (2015). Reducing Leaf-Level Hyperspectral Data to 22 Components of Biochemical and Biophysical Bands Optimizes Tree Species Discrimination. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(6), 3161–3171. <https://doi.org/10.1109/JSTARS.2015.2424594>
- Waser, T. L., Küchler, M., Jütte, K., & Stampfer, T. (2014). Evaluating the Potential of WorldView-2 Data to Classify Tree Species and Different Levels of Ash Mortality. *Remote Sensing*. <https://doi.org/10.3390/rs6054515>
- Wenz, J. J. (2018). Examining water in model membranes by near infrared spectroscopy and multivariate analysis. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1860(3), 673–682. <https://doi.org/10.1016/J.BBAMEM.2017.12.007>
- Wilkes, P., Jones, S. D., Suarez, L., Haywood, A., Mellor, A., Woodgate, W., ... Skidmore, A. K. (2016). Using discrete-return airborne laser scanning to quantify number of canopy strata across diverse forest types. *Methods in Ecology and Evolution*, 7(6), 700–712. <https://doi.org/10.1111/2041-210X.12510>
- Xu, J.-L., Esquerre, C., & Sun, D.-W. (2018). Methods for performing dimensionality reduction in hyperspectral image classification. *Journal of Near Infrared Spectroscopy*, 26(1), 61–75. <https://doi.org/10.1177/0967033518756175>
- Zhang, C., Jiang, H., Liu, F., & He, Y. (2017). Application of Near-Infrared Hyperspectral Imaging with Variable Selection Methods to Determine and Visualize Caffeine Content of Coffee Beans. *Food and Bioprocess Technology*, 10(1), 213–221. <https://doi.org/10.1007/s11947-016-1809-8>
- Zhang, J., Rivard, B., Sánchez-Azofeifa, A., & Castro-Esau, K. (2006). Intra- and inter-class spectral variability of tropical tree species at La Selva, Costa Rica: Implications for species identification using HYDICE imagery. *Remote Sensing of Environment*, 105(2), 129–141. <https://doi.org/10.1016/J.RSE.2006.06.010>
- Zhang, K., & Hu, B. (2012). Individual urban tree species classification using very high spatial resolution airborne multi-spectral imagery using longitudinal profiles. *Remote Sensing*, 4(6), 1741–1757. <https://doi.org/10.3390/rs4061741>
- Zolkos, S. G., Goetz, S. J., & Dubayah, R. (2013). A meta-analysis of terrestrial aboveground biomass estimation using lidar remote sensing. *Remote Sensing of Environment*, 128, 289–298. <https://doi.org/http://dx.doi.org/10.1016/j.rse.2012.10.017>
- Zomer, R. J., Trabucco, A., & Ustin, S. L. (2009). Building spectral libraries for wetlands land cover classification and hyperspectral remote sensing. *Journal of Environmental Management*, 90(7), 2170–2177. <https://doi.org/10.1016/j.jenvman.2007.06.028>