

UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS AGRONÓMICAS
ESCUELA DE PREGRADO

MEMORIA DE TÍTULO

**FIRMAS ESPECTRALES DE VEGETACIÓN URBANA EN SANTIAGO DE
CHILE: CARACTERIZACIÓN Y SEPARABILIDAD DE ESPECIES A ESCALA
FOLIAR.**

TOMÁS IGNACIO ACUÑA RUZ

SANTIAGO DE CHILE
2015

UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS AGRONÓMICAS
ESCUELA DE PREGRADO

MEMORIA DE TÍTULO

**FIRMAS ESPECTRALES DE VEGETACIÓN URBANA EN SANTIAGO DE
CHILE: CARACTERIZACIÓN Y SEPARABILIDAD DE ESPECIES A ESCALA
FOLIAR.**

**SPECTRAL SIGNATURES OF URBAN VEGETATION IN SANTIAGO DE
CHILE: CHARACTERIZATION AND SPECIES SEPARABILITY AT LEAF SCALE.**

TOMÁS IGNACIO ACUÑA RUZ

SANTIAGO DE CHILE
2015

UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS AGRONÓMICAS
ESCUELA DE PREGRADO

FIRMAS ESPECTRALES DE VEGETACIÓN URBANA EN SANTIAGO DE CHILE: CARACTERIZACIÓN Y SEPARABILIDAD DE ESPECIES A ESCALA FOLIAR.

Memoria para optar al Título Profesional de:
Ingeniero en Recursos Naturales Renovables

TOMÁS IGNACIO ACUÑA RUZ

PROFESORES GUÍAS	CALIFICACIONES
Cristián Mattar Bader Ingeniero en Recursos Naturales Renovables, Dr.	7,0
Jaime Hernández Palma Ingeniero Forestal, Dr.	7,0
PROFESORES EVALUADORES	
Mauricio Galleguillos Torres Ingeniero Agrónomo, M.S. Dr.	7,0
Marcos Mora González Ingeniero Agrónomo, Dr.	6,8

SANTIAGO DE CHILE
2015

Dedicado a mis padres

AGRADECIMIENTOS

Es imposible agradecer a todas las personas que han cambiado mi forma de ver el mundo.

Entre ellos, mis padres Ximena y Luis les debo todo lo que soy, porque me criaron con toda la dedicación que un hijo puede querer y también me dieron la libertad suficiente para aprender por mí mismo. Mis hermanos, Catalina y Felipe, son por supuesto otro pilar fundamental en mi vida, de ellos aprendí a caminar, a cocinar, a hablar, a reír, a llorar, y bueno por toda una infancia persiguiéndolos, agradezco sus innumerables consejos.

Por otro lado, mis amigos y amigas me enseñaron sobre otros aspectos de la vida, por ejemplo a disfrutar al máximo cada momento que pueda, pero también, a tranquilizar la situación siempre que sea necesario. Uno nunca sabe cómo las personas moldean tu percepción ante la vida, y eso me sucedió cuando conocí a mis amigos del colegio, entre viajes y experiencias, tengo semanas para reírme de acontecimientos inolvidables. En especial, “los watones apnea”, que con hambre me enseñaron el infinito mundo que hay debajo del agua.

Luego en una siguiente etapa, en la universidad me topé con una serie de personajes convencionales y otros no tanto... se podría decir que conocí una gran diversidad de estilos de vida y que las personas no se definen por el origen familiar, ni por los logros, ni siquiera de donde provienen, sino de la forma en cómo interactúan y relacionan con los otros. Es por esto, que guardo un recuerdo presente de cada uno, al Renzo al Tropie, al Lucho, Yayo, Refaso y al Badulaque. También al eterno gremio en las bancas de JGM, entre ellos al Álvaro, Yoyo y Edu. Últimamente, agradezco haber conocido la doble personalidad de carla o claudia, porque ya... ya... ya.. entendíiiiiiiii. quien es al fin y al cabo. Especialmente, quiero agradecer a la Luci, que juntos compartimos tantas enseñanzas que nunca olvidaré. Quiero reconocer también los momentos de aprendizaje en torno a la agricultura a pequeña escala, porque haciendo algo que históricamente ha sido el insumo esencial de la humanidad, aprendimos sobre la producción de comida y de medicina en el entorno, todos los chiquillos de Rengalentun y de Recicla Orgánico. Y definitivamente a Rayén, por como es, y por el infinito cariño que despertaste en mí.

Además, agradezco a los miembros del Laboratorio para el Análisis de la Biosfera (LAB) y miembros del Laboratorio de Geomática y Ecología del Paisaje (GEP) por la genial combinación de humor e interés científico que tienen y que con un alto grado de dedicación, resuelven creativamente los problemas de las ciencias ambientales.

Finalmente, quiero agradecer a los profesores que guiaron este trabajo, por señalarme lo que significa desarrollar un trabajo científico serio, concreto y con el rigor suficiente para hacer un aporte real al mundo.

ÍNDICE

1. RESUMEN	1
2. ABSTRACT	2
3. INTRODUCCIÓN	3
3.1. OBJETIVOS	6
3.1.1. Objetivo General	6
3.1.2. Objetivos Específicos	6
4. MATERIALES Y MÉTODO	6
4.1. ÁREA DE ESTUDIO	6
4.2. MATERIALES	8
4.2.1. Selección de muestras vegetales.....	8
4.2.2. Equipo de espectro radiometría de campo.....	8
4.3. METODOLOGÍA.....	9
4.3.1. Método para el protocolo de adquisición de firmas a escala foliar	9
4.3.2. Método para la construcción del catastro de firmas espectrales foliares	9
4.3.3. Método para caracterizar la separabilidad espectral.....	11
5. RESULTADOS Y DISCUSIÓN	17
5.1. PROTOCOLO FIRMAS ESPECTRALES A ESCALA FOLIAR	17
5.2. CATASTRO DE FIRMAS ESPECTRALES PARA ESPECIES SELECCIONADAS.....	20
5.2.1. Análisis de las firmas espectrales foliares	20
5.2.2. Análisis comparado de las firmas espectrales promedio	24
5.3. CARACTERIZACIÓN DE LA SEPARABILIDAD ESPECTRAL	25
5.3.1. Análisis de variabilidad de la amplitud espectral (D).....	25
5.3.2. Análisis de características espectrales diferenciables.....	27
5.4. DISCUSIÓN	32
6. CONCLUSIÓN	35
7. BIBLIOGRAFÍA	36
8. APÉNDICE	41
APÉNDICE 1: ÍNDICES HIPERESPECTRALES DE VEGETACIÓN	41
APÉNDICE 2: EJEMPLO <i>METADATA</i> DEL PROTOCOLO DE FIRMAS ESPECTRALES FOLIARES.....	42
APÉNDICE 3: FIRMAS ESPECTRALES FOLIARES DE REFLECTANCIA	43
APÉNDICE 4: EXTRACCIÓN DE VARIABLES POR ANÁLISIS DISCRIMINANTE LINEAR.....	45
APÉNDICE 5: SELECCIÓN DE VARIABLES POR RANDOM FOREST	48

Índice de Tablas

Tabla 1. Especies seleccionadas a partir del porcentaje total de árboles en la comuna.	8
Tabla 2. Resumen de variables de reflectancia original y variables derivadas.	13
Tabla 3. Matriz de confusión para evaluar la precisión de clasificación.....	15
Tabla 4. Protocolo de Adquisición de Firmas Espectrales Foliares	17
Tabla 5. Relación entre la altura del sensor sobre la muestra y el área leída	19
Tabla 6. Número de registros obtenidos por especie.....	20
Tabla 7. Desempeño de Overall Accuracy (OA) e Índice Kappa para la clasificación con Análisis Discriminante Linear. Las variables son: Reflectancia promedio cada 5 nm REF (5); Derivada cada 10 nm DER; Índices Vegetales VI; y todas las variables TV.	27
Tabla 8. Desempeño de Overall Accuracy (OA) e Índice Kappa para la clasificación con Random Forest. Las variables son: Reflectancia promedio cada 5 nm REF (5); Derivada cada 10 nm DER; Índices Vegetales VI; y todas las variables TV.....	27
Tabla 9. Importancia de las variables (1 = mayor importancia y 5 = menor importancia) para cada especie según LDA utilizando la reflectancia % promedio cada 5 nm.	29
Tabla 10. Importancia de las variables (1 = mayor importancia y 5 = menor importancia) para cada especie según LDA utilizando todas las variables TV.	30
Tabla 11. Índices de vegetación (VI).....	41
Tabla 12. Ejemplo de variables para las categorías de campaña de la metadata.....	42
Tabla 13. Ejemplo de variables para las categorías de campaña de la metadata.....	42
Tabla 14. Importancia de las variables (1 = mayor importancia y 5 = menor importancia) para cada especie según LDA utilizando la reflectancia % promedio cada 5 nm.	45
Tabla 15. Importancia de las variables (1 = mayor importancia y 5 = menor importancia) para cada especie según LDA utilizando todas las variables (Reflectancia promedio, índices vegetacionales y primera derivada)	46
Tabla 16. Continúa.	47

Índice de Figuras

Figura 1. Área de estudio en falsocolor estándar (Landsat 8 RGB 543, Febrero, 2014) y las especies dominantes: A) <i>Prunus ceracifera</i> var. <i>nigra</i> , B) <i>Robinia pseudoacacia</i> y.....	7
Figura 2. Diseño experimental para la obtención de las muestras.	10
Figura 3. Algoritmo utilizado para la modelación predictiva de clasificación.....	16
Figura 4. Esquema gráfico del sitio de experimentación del Protocolo FEF. La orientación N aplica solo para el hemisferio Sur.	19
Figura 5. Firmas espectrales foliares de reflectancia promedio \pm 1 S.D. (línea punteada). Código y número de registros por especie.	21
Figura 6. Firmas espectrales foliares de reflectancia promedio \pm 1 S.D. (línea punteada). Código y número de registros por especie.	22
Figura 7. Firmas espectrales foliares de reflectancia promedio \pm 1 S.D. (línea punteada). Código y número de registros por especie.	23
Figura 8. Firma espectral foliar de reflectancia promedio para las especies del catastro de árboles urbanos de Santiago.....	24
Figura 9. Comparación intraespecífica de amplitud espectral (D_{ws}) en el visible (VIS 400-699 nm en azul) e infrarrojo cercano (NIR 700-960 nm en rojo).....	25
Figura 10. Comparación interespecífica de amplitud espectral (D_{as}) para el visible VIS 400-699 (Esquina superior) e infrarrojo cercano NIR 700-960 (Esquina inferior). Se utilizó el color (de amarillo a rojo) para indicar la magnitud de D_{as}	26
Figura 11. Distribución de la Tasa de error con 1000 iteraciones para Análisis Discriminante Lineal (arriba) y Random Forest (abajo), en el cuadro se indica el subconjunto de variables utilizado en cada método.....	28
Figura 12. Gráficos de caja que muestran la variabilidad de una selección de métricas espectrales en las distintas especies. Las métricas son: pendiente entre bandas 473-483 (M483); pendiente entre bandas 553-563 (M563); Photochemical Reflectance Index (PRI); Water Band Index (WBI); Anthocyanin Reflectance Index 2 (ARI2); y Carotenoid Reflectance Index 2 (CRI2).	31
Figura 13. Firmas espectrales foliares de reflectancia promedio \pm 1 S.D. (línea punteada). Código y número de registros por especie	43
Figura 14. Firmas espectrales foliares de reflectancia promedio \pm 1 S.D. (línea punteada). Código y número de registros por especie	44
Figura 15. Importancia de las primeras 40 variables según el método de clasificación Random Forest para métricas espectrales del subconjunto combinado de índices de vegetación y derivadas cada 10 nm.	48

1. RESUMEN

La vegetación arbórea es un componente valioso para el desarrollo del ambiente urbano y su monitoreo es inevitable para cuantificar el impacto ambiental sobre la sociedad. Los árboles, entre otras funciones, entregan valor paisajístico, regulan la temperatura y mejoran la calidad del aire. En este sentido, la información espectral de las hojas de los árboles podría facilitar el monitoreo de la vegetación urbana, conociendo los rasgos foliares y fisiológicos que gobiernan la cantidad de luz reflejada por las hojas en las bandas del visible VIS (400-700 nm) e infrarrojo cercano NIR (700-970 nm). Aquí se presenta un protocolo de adquisición de firmas espectrales foliares (FEF) con un espectro radiómetro portátil *ASD HandHeld* para determinar el grado de separabilidad espectral entre diferentes especies de árboles en Santiago de Chile. La separabilidad espectral se analizó de dos formas: la primera, comparando la variabilidad por especie de la amplitud espectral de la reflectancia en el VIS y en el NIR (VNIR); la segunda, utilizando variables de reflectancia para evaluar dos modelos de clasificación de especies, Análisis Discriminante Linear (LDA) y *Random Forest* (RF), ambos modelos permitieron analizar el subconjunto de variables para estudiar los patrones espectrales intraespecíficos e interespecíficos. Las variables predictivas utilizadas en la modelación son: (1) el promedio de la reflectancia cada 5 nm (N° de variables 115); (2) la derivada estática cada 10 nm (N° de var. 57); y (3) índices espectrales de vegetación (N° de var. 15). El resultado del protocolo FEF, es una pauta de cinco etapas que detalla el procedimiento para recolectar, preparar y registrar la firma espectral junto a un archivo que describe parámetros de instrumentación (*metadata*). El catastro espectral contiene un total de 130 firmas foliares de 18 especies arbóreas, las cuales presentan la huella espectral característica de la vegetación, una baja reflectancia porcentual en el VIS y alta en el NIR. En relación a la clasificación de especies, se tuvo una precisión global mayor al 90% con LDA y en general, su precisión fue mejor que RF (74%). Sin embargo, al usar la primera derivada cada 10 nm del espectro (pendiente), se logra predecir satisfactoriamente en ambos métodos. De hecho, la especie *Quillaja saponaria* muestra patrones espectrales que la hacen separable del resto de las especies en el 100% de las veces (13 registros). Tiene una variabilidad moderada de la reflectancia, y muestra dos rasgos distinguibles, una pendiente elevada en el intervalo 470 - 485 nm de reflectancia y, en promedio, posee un contenido de agua (*Water Band Index*) mayor al resto de las especies. Este mismo análisis se realizó para el resto de las especies, encontrando similares patrones espectrales en *Schinus molle* y *Prunus ceracifera var.nigra*.

Finalmente, para desarrollar un programa de monitoreo de árboles es necesario aumentar el número de registros para consolidar el protocolo FEF, siendo necesario una adquisición periódica y en distintas épocas del año, para así, trabajar en un catálogo de firmas espectrales anuales para diferentes especies y luego, utilizar esta información como dato de campo en estudios de percepción remota.

2. ABSTRACT

Trees are an important component for urban development and their monitoring it is necessary in order to quantify the environmental impact over society. The primary function of trees is related with the regulation of air temperature, improving air quality and also increasing the landscapes value. In this sense, the spectral information may facilitate the monitoring of urban trees, knowing the leaf traits and their physiological aspect that govern the amount of light that is reflected by leaves in the optic region. In this paper is presented a foliar spectral signatures protocol with a portable spectroradiometer, for determine the spectral separability degree between different tree species in Santiago, Chile. The spectral separability was analyzed in two ways, first, by the spectral amplitude of reflectance in the visible VIS (400-700 nm) and in the near infrared NIR (700-970 nm). Second, with two methods of features extraction classification, Linear Discriminant Analysis (LDA) and Random Forest (RF), they were used to characterize the variables that extract intraspecific and interspecific spectral patterns. The predictive variables used to modeling are: (1) the average reflectance every 5 nm (N° of variables 115); (2) static derivate every 10 nm (N° of var. 57); and (3) spectral vegetation indices (N° of var. 15). The result of the protocol is a template of five steps that details the procedure for collecting, preparing and recording the spectral signature at leaf scale with a descriptive document of instrumental parameters (metadata). The spectral library contain a total of 130 leaf spectral signatures of 18 species of trees, which have the characteristically spectral fingerprint of vegetation, a low percentage of reflectance in VIS and relatively high in NIR. Regarding to the predictive classification modeling of species, LDA has a higher overall accuracy (90%) than RF (74%), each with different types of predictive variables. However, when using the first derivate of the spectrum (static slope), both achieved good performance. In fact, the native tree *Quillaja saponaria* has unique spectral patterns that make it separable from other species, showing two remarkable characteristics, an increase in the reflectance between the ranges of 470-485 nm and also, has a greater water content than rest of species (Water Band Index). A similar analysis could be done for the rest of the species, each one showing particular spectral pattern in *Schinus areira* and *Prunus ceracifera var.nigra*. Finally, for developed a monitor program of urban trees it should be increased the number of records with the foliar protocol for consolidation, being necessary a periodically acquisition and in different seasons, then will be able to generate annuals spectral signatures for the studied species, this information could serve as a ground data in remote sensing studies.

3. INTRODUCCIÓN

En el ámbito urbano, la vegetación arbórea se ha vuelto un componente estratégico para el desarrollo sostenible de la ciudad, ya que provee beneficios positivos tales como reducir la polución en el aire, enriquecer el valor estético y regular la oscilación térmica (Hassan & Lee, 2015). Por esta razón, el arbolado urbano debe ser monitoreado periódicamente para evaluar su estado y condición fitosanitaria, pero esto implica una demanda de tiempo y recursos. En este contexto, la percepción remota provee una visión retrospectiva y sinóptica para extensas áreas (Song, 2005), ofreciendo la oportunidad de monitorear la dinámica espaciotemporal del paisaje. De hecho, con sensores de alta resolución espectral (hiperespectral) se estudia la química del dosel arbóreo, para reflejar y predecir el estado funcional de la vegetación y del ecosistema (Asner *et al.*, 2011; Ustin *et al.*, 2009). Sin embargo, para lograr este desafío, primero se deben adquirir datos espectrales en terreno para validar la información remota, principalmente por la frecuencia de muestreo requerida (Pfitzner *et al.*, 2011). Así, la espectroscopia es la técnica que utiliza la interacción entre la radiación electromagnética y la materia, generando la denominada firma espectral, que sirve como dato de referencia para prácticas de calibración y validación remota (Milton *et al.*, 2009).

La firma espectral es un registro fotoeléctrico que representa la propiedad de un material para absorber, reflejar y transmitir el flujo de energía provisto por una fuente emisora. Se utiliza para estudiar la relación entre la reflectancia espectral (ρ_λ) y la estructura biológica, física, química y atómica de los distintos materiales de la Tierra (Bojinski *et al.*, 2003; Pfitzner *et al.*, 2006). En efecto, los organismos fotosintéticos absorben la energía del sol y brindan la condición base para la vida en la Tierra por liberar oxígeno gaseoso, a partir del secuestro de dióxido de carbono (Kiang *et al.*, 2007). Esta absorción puede ocurrir en la región fotosintéticamente activa (400-700 nm), la cual se ubica en la ventana atmosférica de máxima radiación solar (Kiang *et al.*, 2007; Ustin *et al.*, 2009). La vegetación logra absorber la energía solar por un complejo sistema (fotosistemas) formado por distintos pigmentos moleculares (*e.g.* clorofila a-b, carotenos) y selectivamente utilizan una porción específica de la radiación solar dependiendo de la composición estructural y bioquímica de cada grupo taxonómico, lo que resulta en el patrón característico de la vegetación. Este registro o huella espectral es una referencia para cada tipo de coberturas del paisaje, los cuales se integran como componentes puros (*endmember*) en librerías espectrales para analizar la separabilidad espectral de firmas o bien, para estudiar la relación espectro-píxel en una imagen remota (Dennison & Roberts, 2003).

La espectroscopia de campo se ejecuta generalmente bajo iluminación solar, con un radiómetro portátil que mide durante un tiempo dado, a una distancia fija y área reducida (Milton *et al.*, 2009). Sin embargo, poco esfuerzo se ha dedicado para estandarizar el proceso de medición en sí mismo (Hueni *et al.*, 2009), acotando la utilidad de la firma espectral e impidiendo su comparación con otras librerías (Rasaiah *et al.*, 2014). La información de la

firma espectral se puede maximizar incluyendo un archivo descriptivo (*metadata*) que asegure su reproducción en otros sitios (Milton *et al.*, 2009). La *metadata* explicita los parámetros del protocolo de adquisición, aspectos como el lugar, la hora, la configuración geométrica de observación e iluminación, los datos instrumentales (*e.g.* equipo, estándar de calibración), y la técnica de campo (*e.g.* método de muestreo, tiempo de medición). Entre los esfuerzos internacionales para mejorar esta técnica, se puede consultar Bojinski *et al.* (2003), Milton *et al.* (2009) y Pfitzner *et al.* (2011, 2006); la encuesta internacional de Rassaiah *et al.* (2014) sobre factores críticos de la *metadata*; y el protocolo de la plataforma NEON (*National Ecological Observatory Network*) en Kampe *et al.* (2010, 2013).

La respuesta espectral de la vegetación depende de factores biológicos, instrumentales y ambientales que provocan que la firma espectral de la especie sea variable y, a veces, indistinguible entre diferentes especies (Cochrane, 2000; Price, 1994). Sin embargo, si las condiciones externas son similares, cada medición será indicativa para un tiempo y espacio, de modo que sería posible desarrollar un perfil de la especie a través del tiempo si se registran los distintos parámetros de la *metadata* (Blackburn & Milton, 1995; Pfitzner *et al.*, 2011). De este modo, mediante la repetición sistemática del protocolo para diferentes especies, se proporcionaría la información espectral necesaria para evaluar el potencial de discriminación de especies en determinadas épocas y escalas de análisis (Castro-Esau *et al.*, 2006b; Pfitzner *et al.*, 2011). Esto facilitaría la evaluación espectral de imágenes, debido a que la propiedad espectral a nivel de copa se relaciona directamente con el follaje y el fondo del suelo y, técnicamente, su interpretación se limita al detalle a escala foliar (Jacquemoud & Baret, 1990; Knipling, 1970). Luego, esta información se utiliza en el proceso de escalamiento (*scaling-up*) de firmas espectrales de coberturas (Gamon *et al.*, 2006), integrando la contribución energética de cada elemento y analizando escalas mayores. Por otro lado, la variabilidad biológica de la vegetación se ha estudiado en profundidad por Asner (1998), Blackburn (1998, 1999), Feret *et al.* (2008), Ollinger (2011) y Ustin *et al.* (2009), y se establecen factores biológicos como la proporción de pigmentos durante la fenología, el desarrollo del dosel y el tipo de estrés ambiental que modifican la firma espectral, tales como el bloqueo de la fotosíntesis por polución o déficit hídrico (Feret *et al.*, 2008; Gitelson *et al.*, 2009).

La vegetación terrestre tiene una huella espectral característica: el denominado *green bump* o incremento del verde en el visible (VIS), producido por una mayor reflectancia en este color en relación a la absorbida en el azul y el rojo por la clorofila (Kiang *et al.*, 2007); Por otra parte, el *red-edge* (680-730 nm) muestra el contraste de una baja reflectancia en el rojo y una alta reflectancia en el infrarrojo cercano (NIR) producto de la dispersión de fotones entre las paredes celulares y el espacio de aire en el mesófilo. Estas características de la vegetación son potencialmente diferenciables entre las especies. De hecho, la primera derivada del *red-edge* sería capaz de detectar estas diferencias (Portigal *et al.*, 1997) y ser utilizado para la discriminación espectral entre especies (Bruce & Li, 2001). Otra métrica hiperespectral son los índices vegetales (VI), útiles porque caracterizan la concentración de pigmentos, el verdor o la eficiencia en el uso de la luz (Clark & Roberts, 2012; Gitelson *et al.*, 2009; Ollinger, 2011; Zarco-Tejada *et al.*, 2001).

La variabilidad espectral de la vegetación es fundamental para su discriminación. Price (1994) elaboró un método para comparar la amplitud espectral entre dos espectros y evaluar su grado de parentesco. Este método ha sido aplicado en varios estudios para discriminar especies de árboles tropicales (Castro-Esau *et al.*, 2006a y 2006b; Cochrane, 2000; Féret & Asner, 2011) y de estos, Castro-Esau *et al.* (2006a) reportó diferencias entre los espectros foliares adquiridos por tres radiómetros distintos. Por otro lado, la discriminación de especies con métricas hiperespectrales está sujeta, en gran medida, a la habilidad de los métodos estadísticos para generar un modelo que capte las diferencias entre y dentro de las especies (Féret & Asner, 2011). Por ejemplo, algunos estudios utilizan métodos lineales de clasificación como análisis discriminante (Clark *et al.*, 2005; Fung *et al.*, 2003). Éste método se basa en encontrar límites de decisión entre clases, pero como prerequisite, deben tener una distribución gaussiana y la misma matriz de covarianza (Hastie *et al.*, 2009). Otros estudios prefieren métodos no paramétricos, porque evitan los supuestos de distribución normal y se pueden aplicar a una variedad mayor de problemas (Immitzer *et al.*, 2012). Entre este último tipo de métodos, *Random forest* es un buen ejemplo de clasificador para discriminar especies a escala foliar y de copa (Clark & Roberts, 2012; Immitzer *et al.*, 2012; Pal, 2005).

En este sentido, los métodos estadísticos permiten discriminar el conjunto de variables espectrales para cada especie, lo que permitiría establecer un esquema para monitorear su condición fitosanitaria y analizar desde un punto de vista funcional, cómo la actividad biológica de la vegetación beneficia el medioambiente urbano. Por ejemplo, la cobertura vegetal amortigua la concentración de calor y mitiga la emisión de CO₂ antrópico (McPherson & Simpson, 1999). Se ha demostrado que la vegetación urbana mejora la calidad del aire a través de la remoción de contaminantes atmosféricos, la modificación del clima local, modificando los patrones de flujo de viento y alterando la forma en cómo las construcciones usan la energía (Nowak & Crane, 2002; Nowak *et al.*, 2013). Dado este contexto, los estudios del entorno urbano con percepción remota se focalizan en materiales impermeables (Deng & Wu, 2013; Kotthaus *et al.*, 2014) o bien para categorizar sus componentes (Herold *et al.*, 2003, 2004; Tooke *et al.*, 2009), pero no para evaluar el estado fitosanitario ni para la discriminación de especies.

Dada la importancia del arbolado urbano, en este trabajo se desarrolla un método práctico y replicable para adquirir firmas espectrales foliares, para ser usado en la construcción del primer catastro de especies de Santiago, facilitando así su monitoreo de forma remota. El estudio se focaliza en caracterizar y distinguir el conjunto de métricas hiperespectrales con potencial de separabilidad de especies.

3.1. Objetivos

3.1.1. Objetivo General

Construir un catastro de firmas espectrales foliares de especies de árboles urbanos y analizar su separabilidad espectral.

3.1.2. Objetivos Específicos

1. Elaborar un protocolo de adquisición de firmas espectrales foliares.
2. Elaborar un catastro de firmas espectrales para una selección de especies de árboles urbanos.
3. Caracterizar la separabilidad espectral de las distintas especies seleccionadas.

4. MATERIALES Y MÉTODO

4.1. Área de estudio

Para estudiar la vegetación arbórea en Santiago de Chile, se acotó el área de estudio al municipio de La Reina, ubicada en el sector oriente de la ciudad, sobre el piedemonte precordillerano. Se escogió esta comuna porque dispone de información pública actualizada (Hernández *et al.*, 2014) sobre la composición y el estado fitosanitario del arbolado público (*i.e.* árboles presentes en avenidas, calles, pasajes y áreas verdes).

La Reina contiene cerca de 50.000 árboles pertenecientes a más de 200 especies (Hernández *et al.*, 2014). La especie más abundante es *Acer negundo* con un 16,3 % de participación (7.410 individuos). Le siguen en proporción de individuos, *Prunus cerasifera var. nigra* con un 12,4 % y *Robinia pseudoacacia* con un 9,8 %. En la Figura 1 se presenta el área de estudio y un esquema de la forma de selección de los ejemplares a considerar (árboles selectos).

El clima predominante es del tipo templado, con un verano (Diciembre-Febrero) de condición atmosférica seca y de cielo despejado. Además, posee un promedio de temperatura anual de 15 °C. Durante el mes de Enero promedia una temperatura máxima diaria de 29,7 °C y una precipitación total de 0,4 mm.

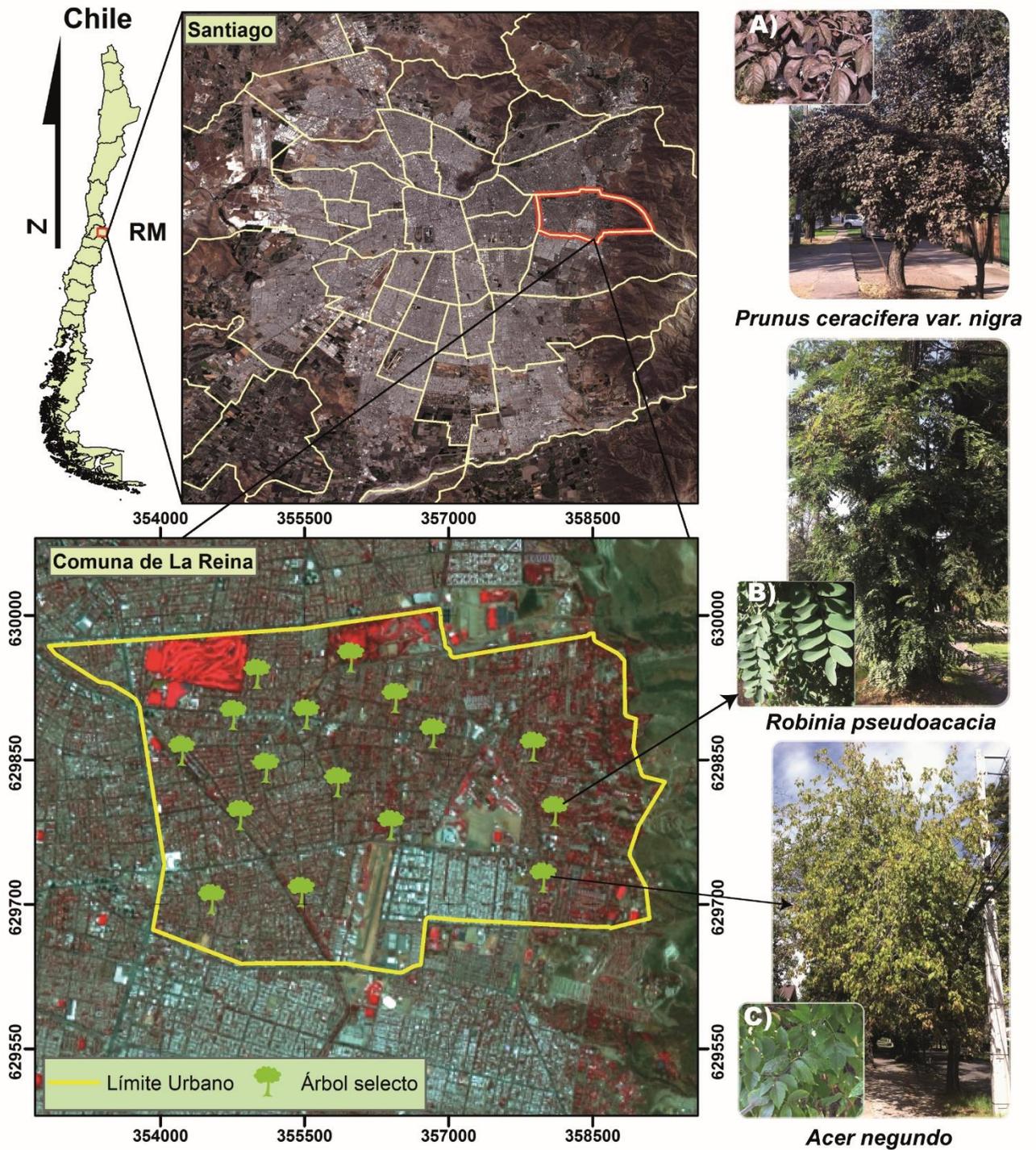


Figura 1. Área de estudio en falsocolor estándar (Landsat 8 RGB 543, Febrero, 2014) y las especies dominantes: A) *Prunus ceracifera var. nigra*, B) *Robinia pseudoacacia* y C) *Acer negundo*.

4.2. Materiales

Los materiales utilizados en esta investigación se clasifican en dos grupos, el primero es el inventario de árboles públicos de la comuna de La Reina y, el segundo, se refiere al equipo instrumental para la adquisición de firmas espectrales a escala foliar.

4.2.1. Selección de muestras vegetales

Se seleccionaron las 18 especies más abundantes de acuerdo a los datos disponibles en el Plan Maestro de Arbolado Público La Reina (Hernández *et al.*, 2014). En la Tabla 1 se presentan las especies seleccionadas, que incluyen el 70 % aproximado del total de individuos arbóreos presentes en la comuna.

Tabla 1. Especies seleccionadas a partir del porcentaje total de árboles en la comuna.

N°	Nombre Común	Nombre Científico	Familia	Estacionalidad del follaje	N° de individuos	%
1	Arce	<i>Acer negundo</i>	Aceraceae	Caducifolio	7410	16,3
2	Ciruelo rojo	<i>Prunus ceracifera var. nigra</i>	Rosaceae	Caducifolio	5615	12,3
3	Falso acacio	<i>Robinia pseudoacacia</i>	Fabaceae	Caducifolio	4467	9,8
4	Quillay	<i>Quillaja saponaria</i>	Quillajaceae	Siempre verde	2879	6,3
5	Liquidambar	<i>Liquidambar styraciflua</i>	Hamamelidaceae	Caducifolio	2573	5,09
6	Ligustro	<i>Ligustrum japonicum</i>	Oleaceae	Siempre verde	2280	5,0
7	Jacaranda	<i>Jacaranda mimosifolia</i>	Bignoniaceae	Caducifolio	1501	3,3
8	Grevillea	<i>Grevillea robusta</i>	Proteaceae	Siempre verde	1025	2,3
9	Pimiento	<i>Schinus areira</i>	Anacardiaceae	Siempre verde	789	1,7
10	Ailanto	<i>Ailanthus altissima</i>	Simaroubaceae	Caducifolio	555	1,2
11	Peumo	<i>Cyptocarya alba</i>	Lauraceae	Siempre verde	513	1,1
12	Melia	<i>Melia azedarach</i>	Meliaceae	Caducifolio	506	1,1
13	Olmo	<i>Ulmus campestris</i>	Ulmaceae	Caducifolio	448	0,88
14	Plátano oriental	<i>Platanus acerifolia</i>	Platanaceae	Caducifolio	336	0,66
15	Sophora japonesa	<i>Styphnolobium japonicum</i>	Fabaceae	Caducifolio	298	0,59
16	Catalpa	<i>Catalpa bignonioides</i>	Bignoniaceae	Caducifolio	261	0,52
17	Tulípero	<i>Liriodendron tulipifera</i>	Magnoliaceae	Caducifolio	224	0,44
18	Alamo musolini	<i>Populus deltoides</i>	Salicaceae	Caducifolio	186	0,37

Fuente: Elaborado a partir de Hernández *et al.*, (2014)

4.2.2. Equipo de espectro radiometría de campo

Para la adquisición de las firmas espectrales de reflectancia porcentual se usó el espectro radiómetro portátil *Analitycal Spectral Devices Hand Held* (ASD; ASD_{INC.}). El ASD es un instrumento de 512 fotodiodos que registra el valor de radiancia espectral ($W \cdot sr^{-1} \cdot m^{-2} \cdot nm^{-1}$) para el rango 325-1075 nm de longitud de onda, con un ancho de banda de 1,5 nm, un tiempo mínimo de escaneo de 17 ms y un campo de visión de 25° (*Field of view*). La muestra objetivo se calibró con un registro de referencia usando un panel *Spectralon* (*Labsphere Inc.*), bajo las mismas condiciones de iluminación y observación.

4.3. Metodología

La metodología se subdividió en tres partes, la primera se refiere a los criterios para elaborar el protocolo de adquisición de firmas espectrales foliares. La segunda describe el proceso de adquisición de firmas para elaborar un catastro espectral. Por último, la tercera parte aborda la caracterización y separabilidad espectral de las 18 especies seleccionadas.

4.3.1. Método para el protocolo de adquisición de firmas a escala foliar

El protocolo de adquisición se elaboró a partir de las experiencias internacionales de Féret & Asner (2011) y Pfitzner *et al.* (2006). Se dividió en dos secciones: (1) Proceso de extracción de la muestra foliar y preparación del equipo experimental; y (2) Procedimiento de adquisición del espectro objetivo y sus parámetros instrumentales asociados (*metadata*).

1.a) Extracción de muestra foliar y preparación del equipo experimental

Esta sección del protocolo incluye el criterio de selección y metodología de colecta de las muestras vegetales, las condiciones de traslado y manipulación para el registro espectral. Los siguientes aspectos fueron considerados: estado fitosanitario, exposición y altura de la muestra desde el suelo; el medio y forma de traslado de la muestra; la manipulación y arreglo de la muestra para el registro espectral. Por último se describe la disposición del equipo de experimentación, similares a las señaladas por Pfitzner *et al.* (2006, 2011) y Pu *et al.* (2003).

1.b) Método de adquisición firma espectral foliar

En esta sección, el protocolo incluye la descripción de los factores instrumentales a controlar para la adquisición del registro espectral. Estos criterios se incorporan en la *metadata* asociada a cada firma espectral. Para este efecto, se plantean doce categorías (campos) que permiten su comparación e interoperabilidad en el tiempo como recomiendan Milton *et al.* (2009) y Rasaiah *et al.* (2014). Existen dos clases de categorías, aquellas relacionadas con el contexto general de la campaña (*e.g.* información general y aplicación específica, preparación muestra vegetal), y aquellas relacionadas con el registro espectral (*e.g.* descripción muestra vegetal, ubicación geográfica, técnica de campo, intervalo de tiempo de integración, campo de visión).

4.3.2. Método para la construcción del catastro de firmas espectrales foliares

2.a) Adquisición de registros espectrales

Con el protocolo descrito, se adquirieron las firmas espectrales para las especies seleccionadas (Tabla 1). La variable medida es la reflectancia porcentual y representa la proporción de radiancia espectral ($W \cdot sr^{-1} \cdot m^{-2} \cdot nm^{-1}$) reflejada por una superficie objetivo, frente a la radiancia espectral reflejada por una superficie lambertiana (*Spectralon*), a igual geometría de iluminación y de observación. Se asume que no existe una variación significativa en la condición de la atmósfera entre el instante de escaneo de la muestra y la

referencia (pocos minutos después). Por consiguiente, la firma espectral representa la reflectancia porcentual como variable respuesta a la señal medida para cada longitud de onda captada por el ASD en el visible e infrarrojo cercano (VNIR 400-1000 nm).

Se define como unidad experimental la muestra vegetal recolectada (conjunto de hojas). En la literatura se reporta que entre 5 y 15 muestras son suficientes para estudiar la variabilidad espectral de una especie (Féret & Asner, 2011). Para minimizar el error aleatorio en cada medición, se programó el ASD para escanear 10 registros consecutivos por muestra, y usar el promedio resultante. Para la recolección de las muestras vegetales, se extrajo un número de hojas suficiente por especie para cubrir el área medida por el ASD, dado que la dimensión de las hojas varía según especie (*e.g.* una o dos para *Catalpa bignonioides* o 20 a 25 para *Robinia pseudoacacia*).

La variabilidad espectral de las muestras adquiridas se organizó en dos niveles según el esquema propuesto por Hueni & Tuohy (2006): (1) Reflectancia espectral de distintos individuos de la misma especie; y (2) Reflectancia espectral de distintas especies. La Figura 2 resume este esquema. Para la especie N°1 se extrae una muestra del sitio 1 y 2, el registro espectral se compone como el promedio de los diez registros obtenidos por el ASD. Con ello, se puede estudiar la variabilidad intraespecífica e interespecífica.



Figura 2. Diseño experimental para la obtención de las muestras.

2.b) Pre procesamiento de las firmas espectrales para el catastro

Las firmas espectrales adquiridas se pre procesaron para remover registros erróneos y promediar los registros consecutivos (1-10). En primer lugar, se acotó el intervalo analizado entre los 400 y 970 nm para descartar los ruidos electrónicos extremos (propios del ASD). Luego, se utilizó la distribución estadística de *T-Student* para encontrar los valores máximos y mínimos aceptables del intervalo de confianza para cada banda de reflectancia, utilizando un coeficiente de confianza de 95%. Este procedimiento se realizó a nivel de muestra espectral y a nivel de especie. En el primer caso, se promedió los registros dentro del rango no rechazable según la ecuación 1, y obtener la firma espectral de la muestra.

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} \quad (1)$$

Donde, \bar{x} es la media de la banda, μ el valor de la banda del registro a analizar, s la desviación estándar y n el número de registros de la banda. Si t se encuentra fuera de la distribución teórica se excluye el valor del registro. En el segundo caso, se removió aquellas firmas espectrales que mostraron en una o varias bandas un valor fuera del rango establecido por el estadístico t (*outliers*).

4.3.3. Método para caracterizar la separabilidad espectral

Para este objetivo, se analizó la separabilidad espectral de las muestras incluidas en el catastro a partir de dos criterios: (1) variabilidad espectral de la especie a escala interespecífica e intraespecífica en distintas regiones del espectro; (2) analizando las características diferenciables de una especie frente al resto.

3.a) Análisis de Variabilidad en la Amplitud Espectral

La variación de las firmas espectrales se calculó con la diferencia en amplitud espectral (D) entre un par de espectros. Esta métrica, desarrollada por Price (1994), se ha utilizado para medir la diferencia de la magnitud de la reflectancia y así, obtener la raíz cuadrática media según la ecuación 2 (Castro-Esau *et al.*, 2006b; Cochrane, 2000).

$$D = \left[\frac{1}{\lambda_b - \lambda_a} \int_{\lambda_a}^{\lambda_b} [S1(\lambda) - S2(\lambda)]^2 d\lambda \right]^{1/2} \quad (2)$$

Donde $S1$ y $S2$ son los registros espectrales a comparar por banda (λ) y promediado sobre el intervalo observado ($\lambda_a - \lambda_b$). Luego, para estudiar la amplitud espectral de una misma especie (variación intraespecífica; D_{WS}), se calculó D para todas las posibles combinaciones entre cada par de muestras disponibles de la especie. La variación interespecífica (D_{AS}) se calculó entre el perfil promedio de cada especie, evaluando el grado de parentesco espectral.

3.b) Análisis de las características espectrales diferenciables

Además del análisis de amplitud espectral, se determinó si era posible discriminar entre especies usando espectros de referencia. La estrategia de análisis fue probar diferentes métricas hiperespectrales como variables predictivas de las diferentes especies (clases). El propósito es detectar un patrón o rasgo diferenciable de una especie frente al resto utilizando métodos multivariados. Para ello, se calcularon tres tipos variables predictivas utilizadas en la literatura y se resumen en la Tabla 2 (Cochrane, 2000; Curran *et al.*, 1995): (1) la reflectancia promedio; (2) la derivada de la reflectancia; (3) índices de vegetación.

Cálculo de métricas Hiperespectrales

i. Cálculo de la reflectancia promedio

Los datos espectroscópicos pueden tener información redundante por el elevado número de bandas adquiridas a alta resolución espectral (*e.g.* alta correlación entre bandas contiguas), lo cual podría disminuir el desempeño del modelo de clasificación (Féret & Asner, 2011; Fung *et al.*, 2003). Luego del pre procesamiento de las firmas, se obtuvieron 571 bandas de reflectancia (rango 400-970 nm), y para evitar la alta correlación de los datos, para cada registro, se obtuvo el promedio móvil entre bandas espectrales cada 5 nm para disminuir el ruido electrónico, finalmente se redujo a 115 bandas de reflectancia.

ii. Cálculo de derivada por rango espectral.

Se calculó la derivada estática cada 10 nm del espectro electromagnético. Para cada muestra se utilizó la ecuación 3, que obtiene la razón de cambio o pendiente de la reflectancia (p) cada 10 bandas espectrales (λ) obtenidas por el sensor.

$$\frac{dp}{d\lambda} = \frac{p_{i+1} - p_i}{\lambda_{i+1} - \lambda_i} \quad (3)$$

Esta métrica evalúa la tendencia al cambio de la reflectancia, analizando por especie en que región del espectro ocurren cambios diferenciables en la pendiente (positiva o negativa), producto de la interacción entre la radiación y la composición y estructura foliar. Este método se ha utilizado en Cochrane (2000) y Vogelmann *et al.* (1993), para estudiar las características diferenciables por especie en el *red-edge*, pero aquí se amplió a todo el VNIR.

iii. Cálculo de Índices de Vegetación (VI)

Los índices de vegetación (VI) son métricas altamente utilizadas en percepción remota por su directa relación con aspectos fisiológicos de la vegetación (Kiang *et al.*, 2007). En la Tabla 2 y Apéndice 1 se presentan los VI utilizados. Los VI seleccionados responden a aplicaciones específicas como: Verdor; Eficiencia en el uso de la luz; Sequedad; Pigmentos foliares; Contenido de agua (Ollinger, 2011).

Tabla 2. Resumen de variables de reflectancia original y variables derivadas.

Variables	Tipo de variable	N° de variables	Significado	Método	Referencia
Reflectancia	VNIR	115	Intervalo 400 – 970	Promedio cada 5 nm	-
Reflectancia Derivada	VI *	15	Verdor	RSVI; NDVI705; mSR705; mNDVI705; VOG1; TVI	Sims & Gamon, 2002; Broge & Leblanc, 2001
			Eficiencia en el uso de la luz	PRI; SIPI	Peñuelas <i>et al.</i> , 1995; Sims & Gamon, 2002
			Sequedad	PSRI;	Merzlyak <i>et al.</i> , 1999
			Pigmento foliar	CRI; ARI	Gitelson <i>et al.</i> , 2001, 2002
			Contenido de agua	WBI	Peñuelas <i>et al.</i> , 1995
	Primera derivada	57	Pendiente cada 10 nm	$\frac{dp}{d\lambda}$	Cochrane, 2000; Curran <i>et al.</i> , 1995

* En Apéndice 1 se presenta una descripción completa de cada índice vegetacional VI.

Modelación de la separabilidad espectral

Para comprobar la hipótesis de que las especies son separables espectralmente, se exploraron dos métodos de clasificación multivariados: Análisis Discriminante Linear y *Random Forest*.

i. Análisis Discriminante Linear (LDA)

LDA se ha utilizado para discriminar especies en varios ecosistemas y principalmente en bosques tropicales (Clark *et al.*, 2005; Féret & Asner, 2011; Fung *et al.*, 2003). El método genera funciones lineales que mejor discriminan las clases definidas (N° especies). La discriminación se consigue al combinar las variables predictivas, tal que la varianza entre-grupos sea maximizada en relación a la varianza dentro-del-grupo (Fung *et al.*, 2003). En otras palabras, busca la combinación de variables que tiene la mejor separación entre los centros de las clases, minimizando al mismo tiempo la variación interna de cada clase. Cada función discriminante otorga un peso específico a cada variable según la ecuación 4:

$$Y = W_1X_1 + W_2X_2 + \dots W_iX_i + \dots + W_nX_n \quad (4)$$

Donde Y es el puntaje de discriminación, X_i es la variable predictiva i y W_i es el peso asociado. El valor absoluto de W_i representa la contribución relativa de la variable en la función de discriminación, esto convierte a LDA en un método fácil de interpretar a través de los coeficientes específicos. LDA se entrenó con la técnica de re muestreo de validación cruzada (*k-fold cross-validation*) para asegurar una validación interna del modelo por agregación de los índices de precisión (Kuhn & Johnson, 2013).

ii. Random Forest (RF)

RF es un método de clasificación basado en árboles de decisión, propuesto por Breiman (2001) y utilizado para la clasificación de especies con métricas hiperespectrales (Clark & Roberts, 2012; Immitzer *et al.*, 2012; Pal, 2005). Este algoritmo utiliza un proceso llamado *bagging* (*Bootstrap + Aggregation*), donde un ensamble de árboles se entrena mediante un proceso de re muestreo de los datos ingresados, cada nodo del árbol, se separa al escoger variables predictivas de forma aleatoria. La clasificación de una muestra se determina por la proporción de votos dentro del ensamble de árboles. Este método entrena el algoritmo con 2/3 de los datos disponibles, y 1/3 queda fuera (*out of bag* o OOB) para ser predicho por el modelo. Esto produce una estimación interna e imparcial de la clasificación (Clark & Roberts, 2012). Por último, las predicciones se promedian y se calculan índices de precisión.

iii. Esquema para la modelación predictiva

Para abordar el esquema de modelación, primero se presentan los índices de precisión y luego el procedimiento para ajustar cada método. La modelación predictiva es el proceso de desarrollar un modelo matemático para obtener una predicción precisa con los datos disponibles (Kuhn & Johnson, 2013). En este sentido, es común que los modelos de clasificación se ajusten mediante técnicas de re-muestreo, utilizando diferentes subconjuntos de datos. Además, se suele subdividir en datos para entrenar y datos para testear la modelación, los segundos, se utilizan como validación externa para obtener una cuantificación real de la tasa de error en precisión de cada método (Castro-Esau *et al.*, 2006b). Los métodos LDA y RF evaluaron los siguientes conjuntos de variables: (1) reflectancia promedio cada 5 nm REF 5; (2) la derivada cada 10 nm DER; (3) índices de vegetación VI; (5) la combinación de DER+VI; (6) Todas las variables TV.

a. Índices de precisión

El desempeño de los métodos se midió a través de la matriz de error o de confusión (Tabla 3). Esta entrega una de las mediciones más efectivas sobre la clasificación, la denominada precisión global o *overall accuracy* (OA), definido como el número de clasificaciones correctas sobre el total (N) de muestras evaluadas (ecuación 5). Sin embargo, éste índice no considera la precisión obtenida para cada clase, útil para estudiar en cuál de ellas hubo clasificaciones erróneas. Para abordar este aspecto, se calculó el porcentaje de clasificaciones erróneas con la Tasa de error (ecuación 6) y por el Índice *Kappa* o coeficiente de acuerdo (ecuación 7), que incluye el factor marginal esperado (Pe) al remover el error obtenido por azar (Jones & Vaughan, 2010).

$$OA = \frac{\sum(VN + VN)}{N} \quad (5)$$

$$Tasa\ de\ error = \frac{\sum(FN + FP)}{N} \quad (6)$$

$$Kappa = \frac{OA + Pe}{1 - Pe} \quad (7)$$

Donde VN y VP corresponden a los valores que realmente fueron clasificados como negativos y positivos. FN y FP aquellos que erróneamente fueron clasificados (Tabla 3). Pe es la probabilidad esperada por acuerdo y se calcula con la ecuación 8.

$$Pe = \frac{\sum(NP \times NA + PP \times PA)}{N^2} \quad (8)$$

Donde NA , NP , PA y PP corresponde a los negativos actuales, negativos predichos, positivos actuales y positivos predichos respectivamente (Tabla 3).

Tabla 3. Matriz de confusión para evaluar la precisión de clasificación.

Matriz de Confusión		REFERENCIA		
		Actual Negativo	Actual Positivo	Total
CLASIFICACIÓN	Predicho Negativo	Verdadero Negativo (VN)	Falso Negativo (FN)	Negativos Predichos ($VN+FN=NP$)
	Predicho Positivo	Falso Positivo (FP)	Verdadero Positivo (VP)	Positivos Predichos ($FP+VP=PP$)
	Total	Negativos Actuales ($VN+FP=NA$)	Positivos Actuales ($FN+VP=PA$)	N

b. Esquema de entrenamiento del clasificador

Las técnicas de re muestreo estiman el desempeño de un modelo matemático de forma interna y proporcionan una medida imparcial de su precisión al evitar el sobreajuste o *overfitting* (Hastie *et al.*, 2009; Kuhn & Johnson, 2013). Sin embargo, la validación externa es una cuantificación más realista de su efectividad, porque evalúa datos que no fueron utilizados en su construcción. En consecuencia, se utilizó el 70% de los datos para entrenar cada modelo discriminante, mientras que el 30% restante, se predice con los modelos y son evaluados en la matriz de confusión. Así para cada set de datos, se realizó una partición aleatoria y proporcional al número de muestras por clase, en datos de entrenamiento y testeo.

Los métodos LDA y RF se ejecutaron con los datos de entrenamiento, cada uno con su procedimiento de ajuste y validación interna. Para LDA, se utilizó la metodología *k-fold cross-validation* ($k=10$), con $k-1$ subconjuntos para entrenar el modelo y el remanente, para la validación interna. Por otro lado, RF se entrenó por re muestreo aleatorio con 2/3 de los datos, utilizando en cada separación un total de $\sqrt[2]{p}$ variables (donde p es el número de variables predictoras en cada set de datos) y se validó internamente por OOB con 1/3 de los datos. Finalmente, se predicen los datos de testeo (30%) y se evalúa el desempeño de cada método. Dado que existe un grado de aleatoriedad en los procesos de validación interna, este procedimiento se iteró mil veces, considerado como una muestra representativa para estudiar la distribución de la tasa de error en la clasificación.

La modelación predictiva es una técnica computacional iterativa, por esta razón, en la Figura 3 se presenta un esquema del algoritmo utilizado en la modelación. El esquema muestra la secuencia de etapas y procesos de iteración o ciclos (*For...do*) realizados sobre cada set de datos, para facilitar su comprensión y replicación. Kuhn & Johnson (2013) utilizan esquemas similares para ejemplificar el procedimiento de los modelos disponibles en el paquete CARET (*Classification And REgression Training*), el cual fue utilizado para desarrollar este análisis utilizando el software estadístico *R-project* (*R Core Team, 2014*).

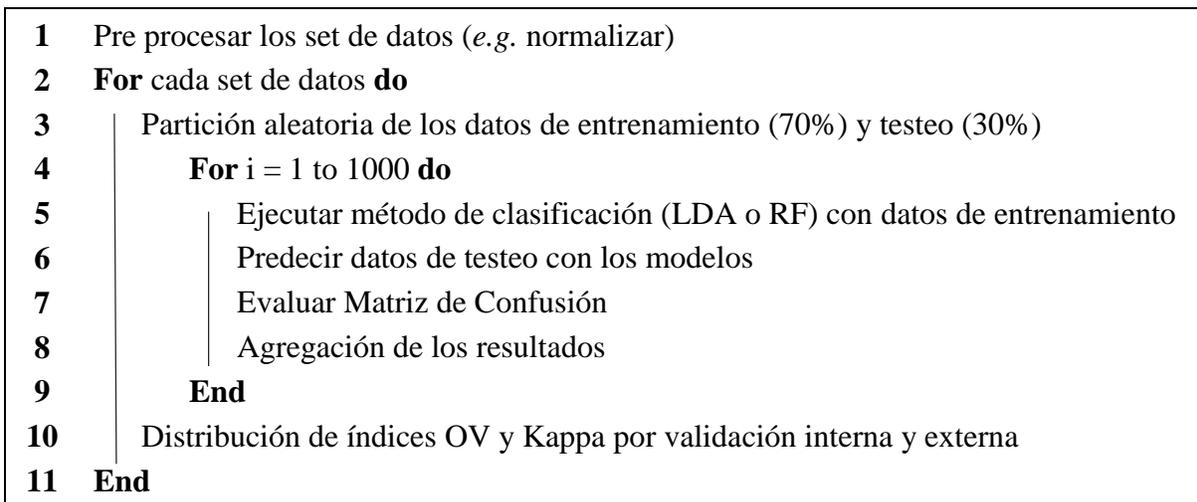


Figura 3. Algoritmo utilizado para la modelación predictiva de clasificación.

5. RESULTADOS Y DISCUSIÓN

5.1. Protocolo firmas espectrales a escala foliar

La Tabla 4 resume el protocolo de adquisición de Firmas Espectrales Foliar (FEF), en la página siguiente, se muestra paso a paso el protocolo. Este se dividió en cinco etapas consecutivas, cada una con una categoría y campos específicos de la *metadata* (Apéndice 2).

Tabla 4. Protocolo de Adquisición de Firmas Espectrales Foliar

Etapa	Categoría	Campo <i>metadata</i> [N° ; Tabla *]
I	Extracción y Traslado	[2; Tabla 12]
II	Preparación equipo experimental	[4; Tabla 12] ; [2-3; Tabla 13]
III	Preparación muestra objetivo	[3; Tabla 12] ; [1; Tabla 13]
IV	Adquisición firma espectral	[4-8; Tabla 13]
V	Elaboración <i>metadata</i>	-

* Los campos de la etapa V se encuentran en la Tabla 12 y Tabla 13, apéndice 2.

A continuación, se detallan los aspectos instrumentales para cada etapa. El límite de hora para recolectar muestras fue hasta las 11:00 AM, sólo muestras con exposición ecuatorial (Norte) y a una altura mínima de tres metros. Las muestras extraídas se guardan inmediatamente en bolsas grandes de polietileno, y luego en un contenedor hermético para su traslado (1-3 horas). Para la etapa II, el sitio de experimentación se ubicó en un sector abierto a la radiación solar directa, en un contexto urbano y con edificios presentes a 20 m de distancia. La técnica instrumental se ilustra en la Figura 4, constante para cada adquisición. La geometría de observación fue en ángulo nadir de 0° (vertical a la mesa) y se fijó un campo de visión de 15,4 cm² (altura de 10 cm entre el ASD y la muestra) según la Tabla 5.

Para la etapa III, la muestra vegetal se preparó bajo sombra, removiendo peciolas, ramas, polvo o cera adherida en la epidermis de la cara adaxial de las hojas. Se arregló un plano uniforme con las hojas (cara adaxial) y paralelo a la mesa de trabajo (Figura 4). Para la etapa IV, la firma de referencia del *Spectralon* se obtuvo cada 10 minutos. El horario de adquisición fue entre las 11:00 y 17:30 hrs, para obtener el máximo de radiación solar. La nubosidad del cielo se midió con el método de octavas de cobertura del cielo, no obstante, sólo se midió con cielo despejado (0 octas). La etapa V se ejemplifica para una muestra de *Quillaja saponaria* en Apéndice 2.

Protocolo de Firmas Espectrales Foliare (FEF)

Etapa 1: Extracción y traslado de la muestra vegetal

- Identificar el árbol objetivo y su ubicación geográfica.
- Extraer 1-2 ramas expuestas al sol con un gancho de poda (tijera) a una altura mínima de 3 metros.
- Dejar la muestra fuera de la luz solar directa e inmediatamente seleccionar 5 a 15 hojas maduras, con un estado homogéneo en cuanto a fenología y vigor.
- Guardar las hojas en una bolsa de polietileno grande con una toalla húmeda en su interior (para mantener hidratación). Seleccionar una cantidad de hojas suficiente para cubrir el área leída por el radiómetro (Tabla 5).
- Extraer el aire sobrante de la bolsa, rotular y guardar en un contenedor hermético (con botellas de agua congelada) para su traslado al sitio de experimentación.

Etapa 2: Preparación equipo experimental

- Ensamblar el equipo de experimentación previo a la etapa 1, según el esquema de la Figura 4.
- Fijar altura D1 (Figura 4) y la geometría de observación del sensor en ángulo nadir (0°).
- Configurar el número de registros (x10) y el tiempo de integración del sensor.
- Es recomendable precalentar el sensor durante una hora antes de la adquisición de registros.

Etapa 3: Preparación muestra objetivo

- Descartar peciolo y elementos adheridos en la cara adaxial de cada hoja de la muestra objetivo. Evitar exposición solar trabajando bajo sombra.
- Arreglar un plano homogéneo con las hojas de la muestra, evitando superposición y vacíos de vacíos entre ellas (Figura 4). Cubrir como mínimo el área leída por el sensor según la Tabla 5.

Etapa 4: Adquisición firma espectral

- Preparar la referencia estándar (*Spectralon*).
- Registrar la reflectancia porcentual de la referencia estándar.
- Registrar la reflectancia porcentual de la muestra objetivo.

Esquema gráfico del Protocolo FEF

Tabla 5. Relación entre la altura del sensor sobre la muestra y el área leída

D1: Altura (cm)	Área leída por al ASD en ángulo Nadir FOV 25° (cm ²)
10	15,4
20	61,8
50	386,0
100	1544,0

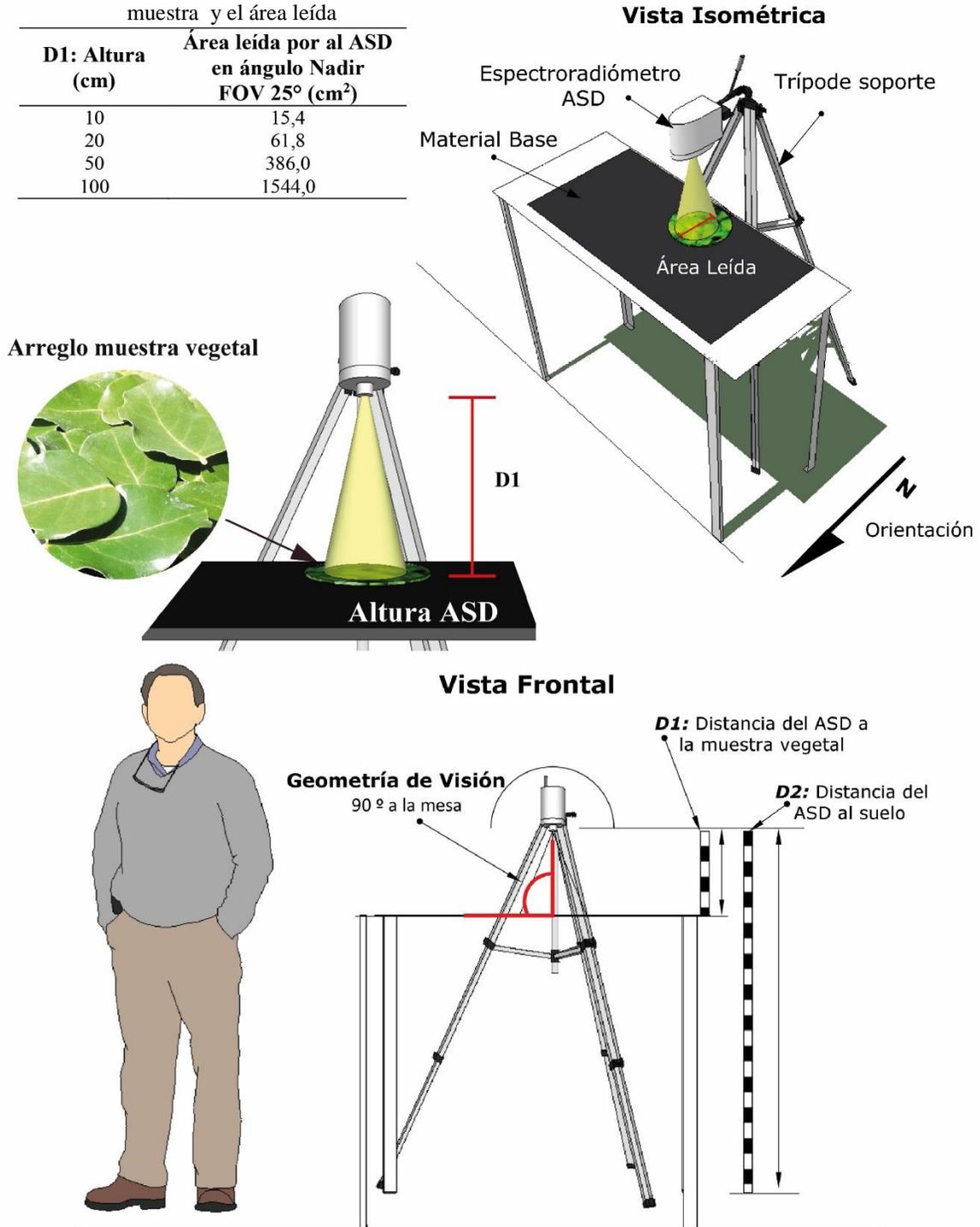


Figura 4. Esquema gráfico del sitio de experimentación del Protocolo FEF. La orientación N aplica solo para el hemisferio Sur.

5.2. Catastro de firmas espectrales para especies seleccionadas.

Aplicando el protocolo descrito anteriormente se obtuvo un total de 140 firmas espectrales y 10 de ellas, se descartan por presentar un estadístico *t* fuera del intervalo de confianza (*p*-valor > 0,05), y con variaciones evidentes en la reflectancia porcentual (desde ahora reflectancia). Lo anterior, se produjo posiblemente por errores metodológicos al momento de adquisición (e.g. por movimiento mecánico del ASD o de la muestra vegetal por el viento). La Tabla 6 presenta el número de registros obtenidos por especie y el código de identificación utilizado. El código USDA es la identidad otorgada por el Departamento de Agricultura de los Estados Unidos.

Tabla 6. Número de registros obtenidos por especie.

Nº	Nombre Científico	Código	Código USDA	Nº de Registros
1	<i>Acer negundo</i>	A_ne	ACNEN	10
2	<i>Prunus ceracifera var. nigra</i>	P_ce	PRCEP	7
3	<i>Robinia pseudoacacia</i>	R_ps	ROPS	6
4	<i>Quillaja saponaria</i>	Q_sa	QUSA	13
5	<i>Ligustrum japonicum</i>	L_ja	LJJA	10
6	<i>Jacaranda mimosifolia</i>	J_mi	JAMI	5
7	<i>Liquidambar styraciflua</i>	L_st	LIST2	8
8	<i>Grevillea robusta</i>	G_ro	GRRO	6
9	<i>Schinus areira</i>	S_ar	SCMO	11
10	<i>Ailanthus altissima</i>	A_al	AIAL	5
11	<i>Cryptocarya alba</i>	C_al	CRRU12	10
12	<i>Melia azedarach</i>	M_az	MEAZ	6
13	<i>Ulmus campestris</i>	U_ca	UMGL	5
14	<i>Platanus acerifolia</i>	P_ac	PLHY3	6
15	<i>Styphnolobium japonicum</i>	S_ja	STJA9	6
16	<i>Catalpa bignonioides</i>	C_bi	CABI8	5
17	<i>Liriodendron tulipifera</i>	L_tu	LITU	6
18	<i>Populus deltoides</i>	P_de	PODED	5

5.2.1. Análisis de las firmas espectrales foliares

Los registros obtenidos con el protocolo FEF revelan la huella espectral característica de la vegetación: una baja reflectancia en el VIS y relativamente alta en el NIR, esta transición ocurre en el sector denominado *red-edge* y señala el contraste entre ambas regiones del espectro electromagnético. Los sectores particulares que otorgan ésta característica son, en primer lugar, la subida en el verde o *green bump* del VIS (entre los 500 y 570 nm) por la presencia de pigmentos verdes (e.g. clorofila). En este sector, las firmas presentan un máximo cerca de la banda 550 nm, donde el rango promedio de reflectancia varía entre 6 % para *Prunus ceracifera var. nigra* (P_ce) y 20 % para *Liriodendron tulipifera* (L_tu). En segundo lugar, tanto la reflectancia en el azul (450-500 nm) como en el rojo (620-700 nm) presentan una menor magnitud, con valores que varían entre el 2 % y 11 % en los 470 nm y entre 3 % y 11 % en los 660 nm respectivamente. Esta menor reflectancia, se produce principalmente por un aumento de la absorbancia por los pigmentos fotosintéticos en dichos sectores. Por otro lado, la reflectancia en el NIR es relativamente alta (entre 50-70 %) y tiende a ser más variable que el VIS. Para ilustrar estas características, se graficó para cada especie la firma espectral promedio \pm 1 desviación estándar (S.D.), para esto, se utilizó el promedio cada cinco bandas espectrales dentro del intervalo acotado (400-970 nm).

La Figura 5 señala los espectros de: A) *Acer negundo* (A_ce); B) *Catalpa bignonioides* (C_bi); C) *Cryptocarya alba* (C_al); y D) *Ligustrum japonicum* (L_ja). Entre las 18 especies evaluadas, la especie C_bi (Figura 5.B), presenta la mayor variación de reflectancia en la banda 550 nm ($13 \pm 4\%$). Además, tiene una alta desviación en el *red-edge*. A partir de esto, es probable que las muestras analizadas tuvieran diferentes concentraciones de clorofila, por causas fenológicas o de estrés ambiental. Las demás especies, también muestran la firma singular de la vegetación, pero con una desviación mayor en el NIR que en el VIS. Aunque L_ja (Figura 5.D), varía a lo largo del VNIR y probablemente sea causado por la morfología foliar (lámina en “v”), que complicó el arreglo y disposición de las hojas en el experimento.

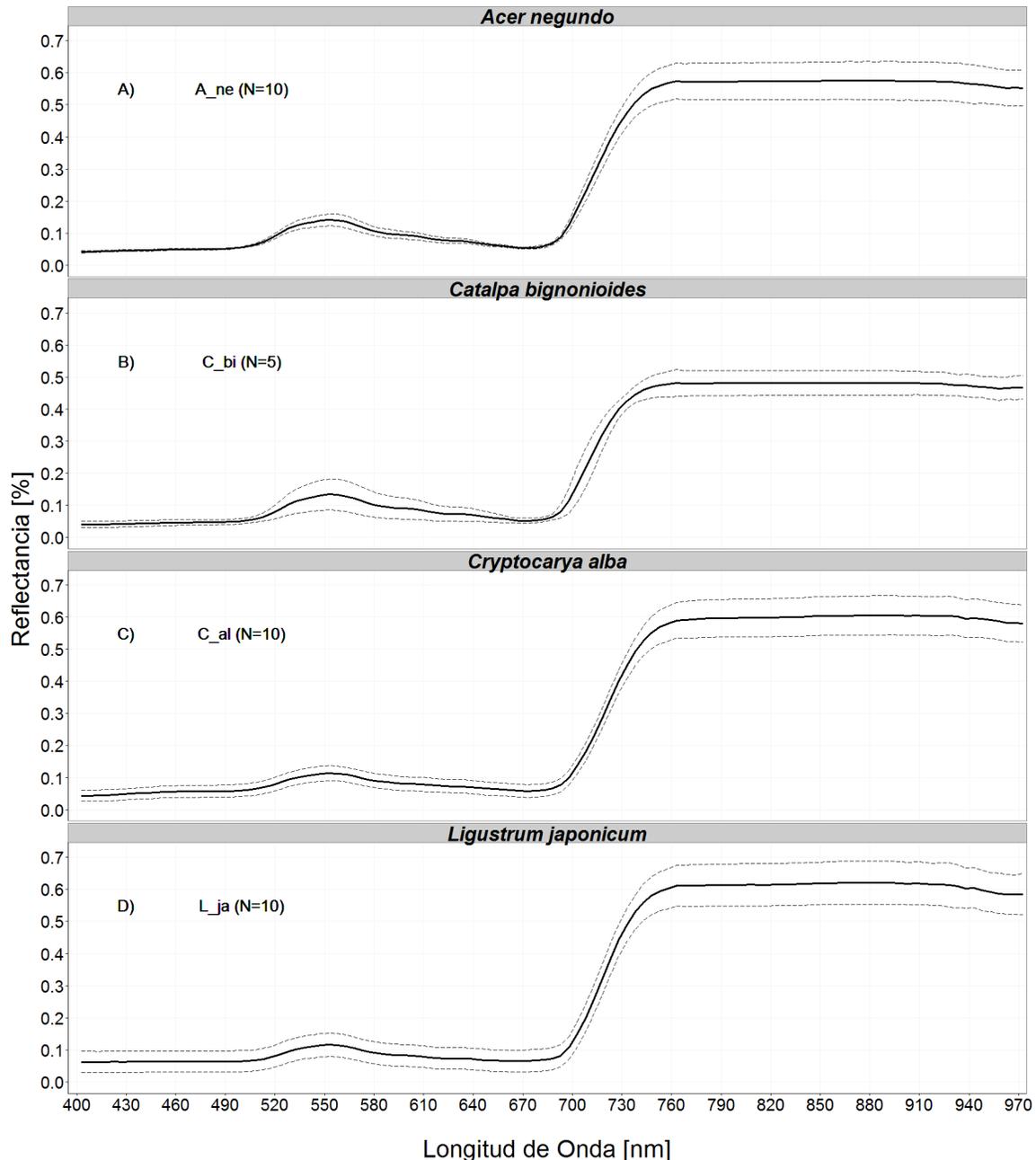


Figura 5. Firmas espectrales foliares de reflectancia promedio \pm 1 S.D. (línea punteada). Código y número de registros por especie.

La Figura 6.G, muestra la firma espectral promedio de *Platanus acerifolia* (P_ac). Entre todas las especies evaluadas, P_ac posee la menor desviación en el NIR (cerca del 1%), lo que indicaría que esta región es relativamente constante para la especie. Como era esperable, el ciruelo rojo (P_ce, Figura 6.H) fue la única especie que presenta una firma con una subida mínima en el verde, expresando una magnitud de reflectancia casi invariable en el azul, verde y rojo del VIS ($5 \pm 1\%$). Sin embargo, entre las especies evaluadas, presenta la mayor desviación en el NIR (15%), seguida por *Ailanthus altissima* (Figura 13.M en Apéndice 3) y *Schinus areira* (S_ar, Figura 8.K).

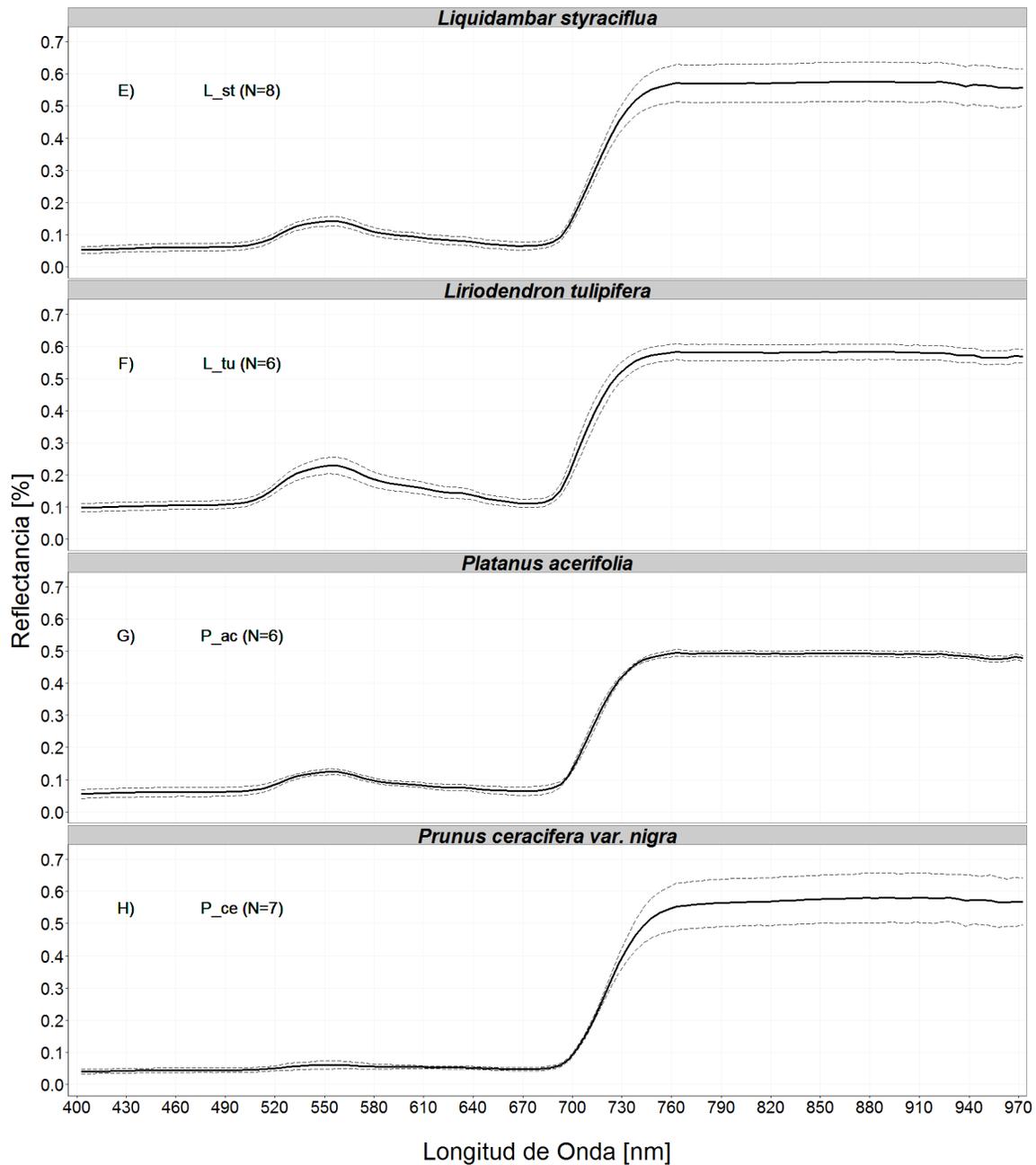


Figura 6. Firmas espectrales foliares de reflectancia promedio \pm 1 S.D. (línea punteada). Código y número de registros por especie.

En la Figura 7.I se presenta el promedio de *Quillaja saponaria* (Q_sa), con una desviación inferior a 2 % en el VIS y a 6 % en el NIR, además muestra una caída en la región de absorción del agua (entre los 950-970 nm). En la Figura 7.J, se observa la firma promedio de *Robinia pseudoacacia* (R_ps), destacando el parentesco espectral de las muestras analizadas en todo el VNIR. La Figura 7.L exhibe la firma de *Stiphlobium japonicum* (S_ja), con una variación aproximada de 10 % a lo largo del VNIR, excepto en el *red-edge*. El pimiento o S_ar (Figura 7.K) tiende a incrementarse gradualmente en el azul (desde los 430 nm).

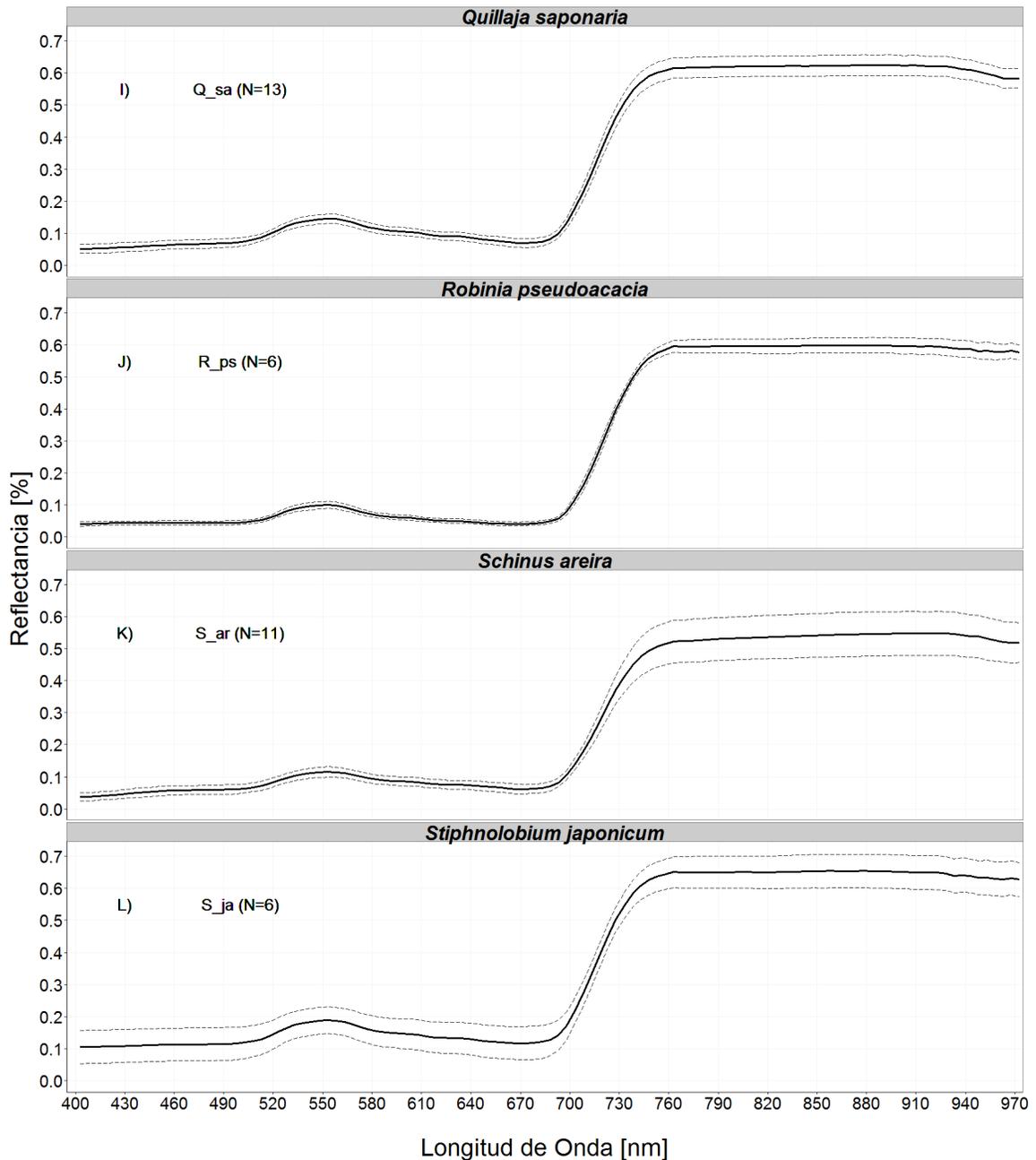


Figura 7. Firmas espectrales foliares de reflectancia promedio \pm 1 S.D. (línea punteada). Código y número de registros por especie.

A pesar de recolectar para cada especie muestras vegetales más o menos homogéneas (*e.g.* mismo estado de maduración y fitosanidad), se constata variación de la reflectancia para algunas especies en diferentes regiones del espectro electromagnético. Esto se produce principalmente por rasgos foliares, tales como el nivel de clorofila, el grosor del mesófilo y el porcentaje de espacio de aire en el tejido esponjoso (Castro-Esau *et al.*, 2006b). En la Figura 13 y Figura 14 (Apéndice 3) se presentan las firmas espectrales para el resto de las especies. A continuación, se compara el perfil espectral promedio de las especies del catastro.

5.2.2. Análisis comparado de las firmas espectrales promedio

El ciruelo rojo (línea roja) se discrimina fácilmente en la Figura 8 como ya fue indicado. El tulipero o *L_tu* (línea café clara) tiene el mayor porcentaje de reflectancia (> 20%) cerca de los 550 nm, luego le siguen *S_ja* (línea beige) y *Populus deltoides* (*P_de*, línea verde), éstas dos últimas se asemejan en todo el VNIR. En cambio, el resto de las especies están moderadamente agrupadas en el VIS. Entre ellas, *Jacaranda mimosifolia* (*J_mi*, línea gris) tiene la menor reflectancia. Además, se constata la existencia de un comportamiento distintivo en algunas especies en ciertas regiones del espectro. Por ejemplo, la reflectancia se mantiene casi constante a través del NIR en *Q_sa*, *P_de*, *C_bi* y *Liquidambar styraciflua* (*L_st*). Luego, en *S_ar*, *P_ce*, *Grevillea robusta* (*G_ro*) y *Melia azedarach* (*M_az*) tiende a incrementarse desde los 760 a los 940 nm. Del mismo modo, en *S_ar* (Figura 7.K), *Q_sa* (Figura 7.I) y *C_al* (Figura 4.C) la reflectancia se eleva progresivamente entre los 420 y 480 nm.

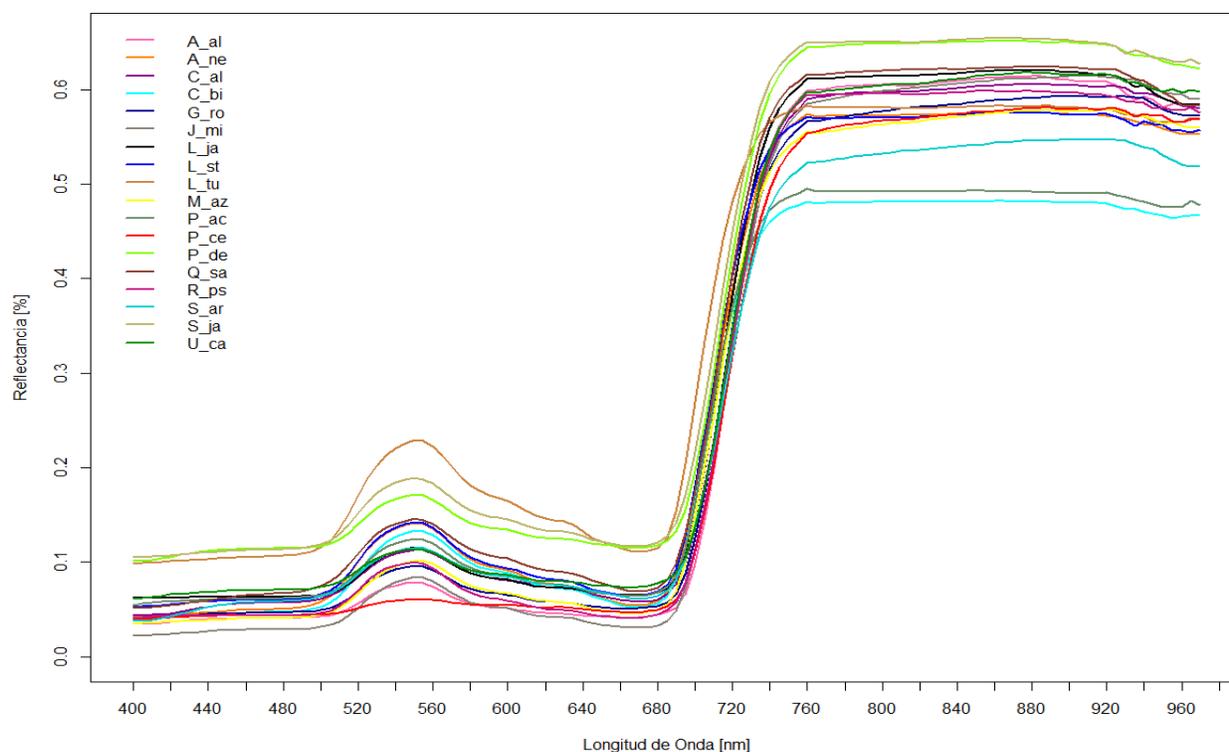


Figura 8. Firma espectral foliar de reflectancia promedio para las especies del catastro de árboles urbanos de Santiago.

5.3. Caracterización de la separabilidad espectral

5.3.1. Análisis de variabilidad de la amplitud espectral (D)

D intraespecífico

La Figura 9 presenta la variación de la amplitud de la reflectancia en el VIS y NIR para cada especie del catastro. Como primera observación, la D_{ws} (amplitud intraespecífica) tiene una variación mayor en el NIR que en el VIS. El D_{ws} promedio más bajo en el VIS es para J_mi (0,78), seguido por P_ce y R_ps. Al contrario el D_{ws} promedio más elevado es para L_ja y S_ja (4,1 y 5,9). Lo mismo ocurre en el NIR, donde P_ac y R_ps tienen el menor D_{ws} (1,2 y 2,7), y las especies P_ce y A_al el mayor (7,9 y 8,1). Así, esta métrica indicaría la región del espectro donde sería probable separar las especie dado el parentesco espectral de las muestras utilizadas. Un valor bajo significa que las especies tendrán un rasgo o característica potencialmente diferenciable (siendo no diferenciable un valor alto). Para las muestras del catastro, D_{ws} es más homogénea en el VIS para la mayoría de las especies. Por ejemplo, en la Figura 9 se observan ocho muestras que escapan de la distribución teórica (*outliers*), presentando un $D_{ws} > 15$ en algunos casos. Esto podría indicar aquellos registros por especie que tienen una variación abrupta en determinados sectores del espectro, lo que podría estar relacionado con una variación intrínseca de la especie o debido a un error experimental en la adquisición del registro con el Protocolo FEF. En cualquier caso, se debe obtener más registros para determinar la respuesta. Finalmente, se establece que el NIR tiene un D_{ws} promedio mayor que el VIS (y más variable), coincidiendo con la descripción de las firmas espectrales antes presentada, especialmente para las especies a la derecha de la Figura 9. S_ja escapa de este comportamiento. Además, tanto S_ja y L_ja tienen una variación interespecífica alta en todo el VNIR (D_{ws} promedio >4).

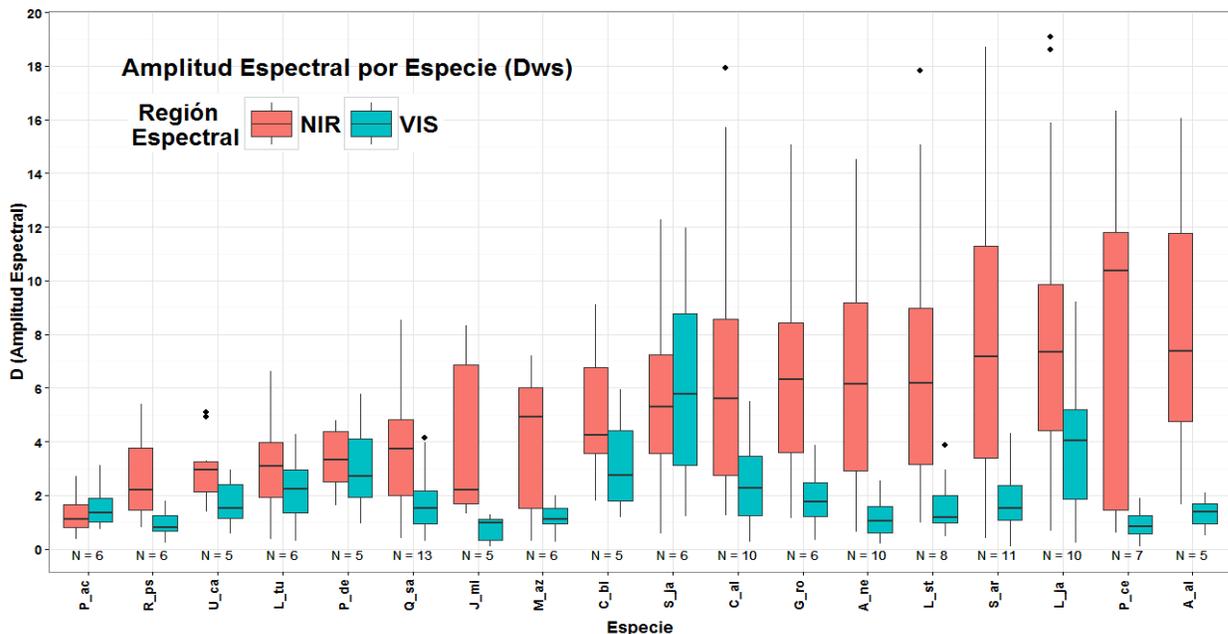


Figura 9. Comparación intraespecífica de amplitud espectral (D_{ws}) en el visible (VIS 400-699 nm en azul) e infrarrojo cercano (NIR 700-960 nm en rojo).

D interespecífico

La Figura 10 representa la diferencia de la amplitud espectral entre las firmas espectrales promedio de las 18 especies (*D* interespecífico o *D_{as}*). Esta métrica mide el grado de separabilidad entre especies, obteniendo una variación de 0 a 16. Se graficó en una matriz de color (de amarillo a rojo) para dos sectores del espectro, el VIS en la esquina superior y el NIR en la esquina inferior. La diagonal principal tiene un valor igual a 0 ya que es la misma especie evaluada, entonces a menor valor, nula separabilidad. En general, se tiene un grado mayor (*D_{ws}* rojo) en el sector inferior de la matriz, lo que indica que el NIR tiene mejor capacidad para diferenciar entre especies. Por ejemplo, entre P_ac y P_de, existe aproximadamente un valor 15 en el NIR y 4 en el VIS. Esto indicaría que la reflectancia en el NIR discrimina mejor (potencialmente) estas dos especies (Figura 8). De igual forma, Q_sa sería separable de C_bi y P_ac en el NIR y no en el VIS (*D_{as}* cercano a 2).

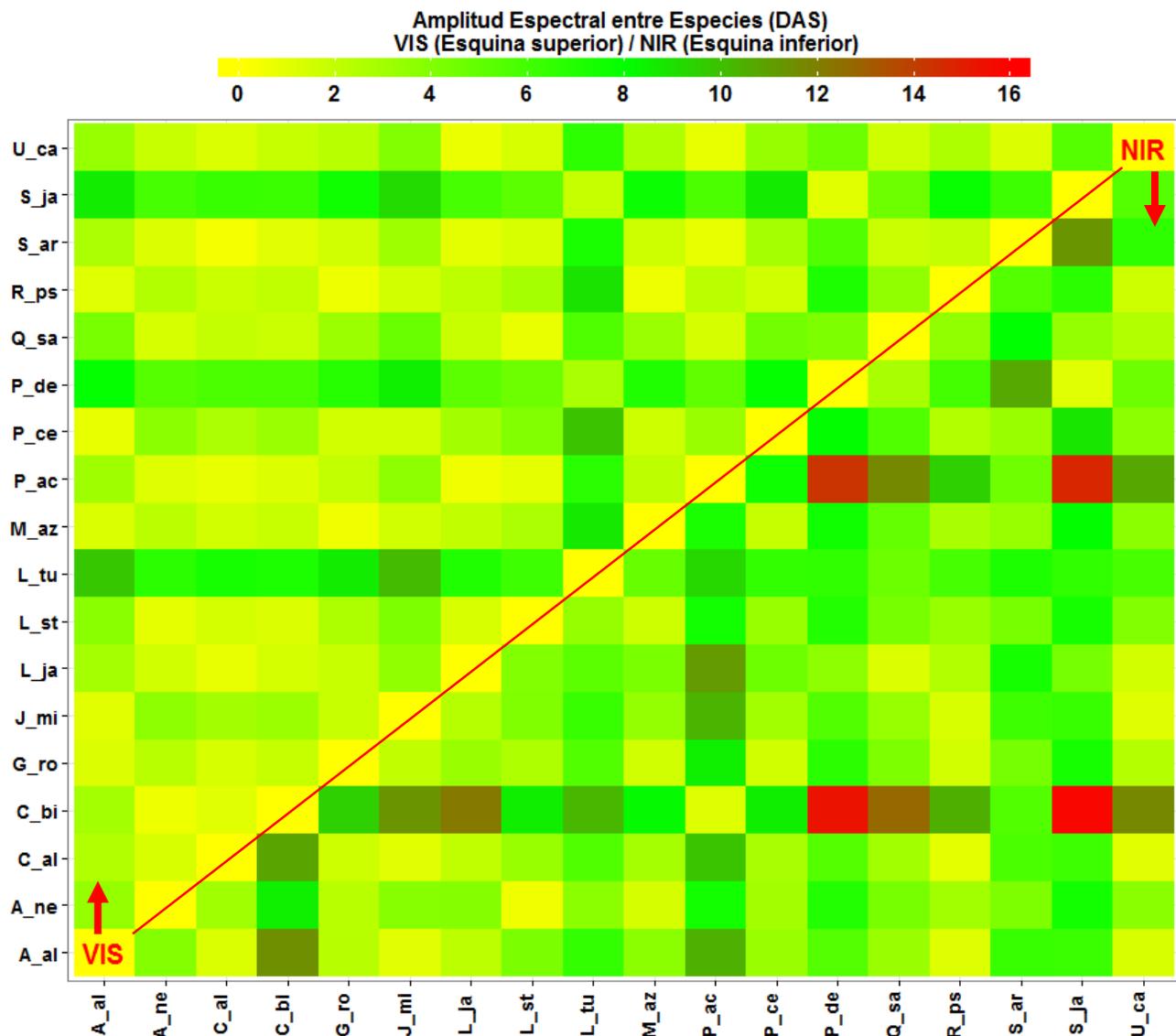


Figura 10. Comparación interespecífica de amplitud espectral (*D_{as}*) para el visible VIS 400-699 (Esquina superior) e infrarrojo cercano NIR 700-960 (Esquina inferior). Se utilizó el color (de amarillo a rojo) para indicar la magnitud de *D_{as}*.

5.3.2. Análisis de características espectrales diferenciables

A continuación se presenta en la Tabla 7 y Tabla 8 la predicción de la modelación o clasificación por LDA y RF, y luego en la Figura 11, se presenta la distribución de la tasa de error obtenido para cada método utilizando el subconjunto de variables con mejor predicción (en rojo en la Tabla 7 y Tabla 8). Las Tablas 7 y 8 se ordenan descendientemente según el promedio de la precisión global o *overall accuracy* (OA) obtenido para la validación externa (Matriz de confusión). En general, existe un mejor desempeño con LDA que con RF, porque tanto la reflectancia promedio cada 5 nm REF (5) como la derivada cada 10 nm (DER) tienen una predicción perfecta de los datos de testeo (30 %). Los índices de vegetación (VI) tuvieron una baja predicción, incluso disminuye la clasificación del subconjunto combinado VI+DER en LDA. Por otro lado, como la naturaleza de los métodos es diferente (lineal v/s árbol de decisión), el desempeño obtenido cambia según el set de variables. Por ejemplo, al utilizar REF (5) en RF, se obtiene un resultado contrario a LDA, con un desempeño pobre e inferior al 0,43 en OA e Índice Kappa. Del mismo modo, al utilizar VI+DER, se obtiene una precisión inversa en la validación interna (LDA inferior al 0,57 y RF superior al 0,73). A pesar de esto, el subconjunto DER, por sí solo, es sensible en ambos métodos y probablemente el desempeño de VI está condicionado por el bajo número de variables utilizadas (15).

Tabla 7. Desempeño de Overall Accuracy (OA) e Índice Kappa para la clasificación con Análisis Discriminante Lineal. Las variables son: Reflectancia promedio cada 5 nm REF (5); Derivada cada 10 nm DER; Índices Vegetales VI; y todas las variables TV.

Análisis Discriminante Lineal (LDA)					
Tipo de variable	N° de Variables	Validación Interna (70%)		Matriz de Confusión (30%)	
		OA	Kappa	OA	Kappa
REF (5)	115	0,91	0,90	1	1
DER	57	0,88	0,86	1	1
TV	187	0,83	0,81	0,97	0,96
VI+DER	72	0,57	0,53	0,87	0,86
VI	15	0,62	0,59	0,77	0,75

Tabla 8. Desempeño de Overall Accuracy (OA) e Índice Kappa para la clasificación con Random Forest. Las variables son: Reflectancia promedio cada 5 nm REF (5); Derivada cada 10 nm DER; Índices Vegetales VI; y todas las variables TV.

Random Forest (RF)					
Tipo de variable	N° de Variables	Validación Interna (70%)		Matriz de Confusión (30%)	
		OA	Kappa	OA	Kappa
VI+DER	72	0,74	0,73	0,93	0,92
DER	57	0,72	0,70	0,90	0,90
TV	187	0,68	0,66	0,89	0,88
VI	15	0,56	0,53	0,81	0,80
REF (5)	115	0,32	0,27	0,43	0,40

Distribución del error

A continuación se presenta en la Figura 11 la distribución de la tasa de error obtenida para LDA (arriba) con el subconjunto REF (5) y para RF (abajo) con el subconjunto VI+DER. Como se explicó anteriormente, los métodos utilizan procesos aleatorios en la partición de los datos de predicción, esto significa que para las mil iteraciones, un grupo distinto de muestras/variables se subdivide para la validación interna del error, generando entonces, diferentes resultados en la predicción de los datos. El promedio de la tasa de error para LDA es igual a 9 % y para RF igual a 25 %, en ambos casos, las distribuciones siguen una función gaussiana (normal), aunque LDA es más achatada porque tiene una varianza menor que RF. Esto indicaría que LDA tiene un error menor que RF en la predicción de las firmas espectrales.

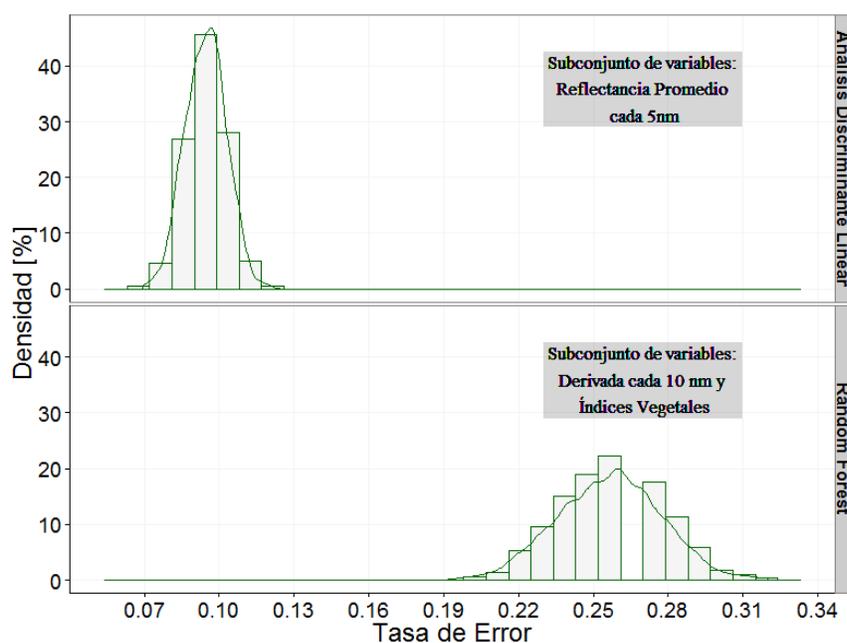


Figura 11. Distribución de la Tasa de Error con mil iteraciones para Análisis Discriminante Lineal (arriba) y *Random Forest* (abajo), en el cuadro gris se indica el subconjunto de variables utilizado en cada método.

Características diferenciales por especie

A continuación se describe la importancia de las primeras cinco variables seleccionadas (variables extraídas) por LDA especies (Tabla 9). Éstas se obtuvieron a partir del peso relativo (valor absoluto) de las diferentes variables del modelo construido (promedio de reflectancia cada 5 nm), esto significa que la categoría 1 es la variable de mayor peso en la función discriminante y la categoría 5 representa el quinto valor de importancia para la clase. Por ejemplo, para P_ce las bandas de reflectancia (promedio) que más contribuyen para su clasificación son 553, 548, 558, 543 y 538 nm (Tabla 9). Estas bandas se encuentran en la transición de reflectancia del verde al rojo y era esperable, debido a que la especie tiene una coloración roja (Figura 1 y Figura 6.H) y LDA detectó éste patrón espectral para reconocer

la especie frente al resto. Un análisis similar se puede realizar para el resto de las especies. Un aspecto interesante del método, es que asocia regiones específicas del espectro para discriminar cada especie, en otras palabras, las especies se diferencian por determinados sectores de la firma espectral. A pesar de esto, Q_sa y S_ar tienen las mismas variables (733, 738, 743, 728 y 748, en negrita), y de hecho no es hasta la octava variable de la función discriminante donde se diferencian (723 para Q_sa y 763 para S_ar). Cabe destacar, que si bien existe un mismo orden de importancia de las variables predictivas, no se señala ni el signo ni el peso relativo en cada función, lo que podría explicar este parentesco sin ningún tipo de error en la clasificación de los datos de entrenamiento y de testeo por el método al (Tabla 7). Para el resto de las especies analizadas revisar la Tabla 14 (Apéndice 5).

Tabla 9. Importancia de las variables (1 = mayor importancia y 5 = menor importancia) para cada especie según LDA utilizando la reflectancia % promedio cada 5 nm.

Rango Longitud de Onda	REF % cada 5 nm	P_ce	Q_sa	S_ar
Visible (VIS)	538	5		
	543	4		
	548	2		
	553	1		
	558	3		
Infrarrojo cercano (NIR)	728		4	4
	733		1	1
	738		2	2
	743		3	3
	748		5	5

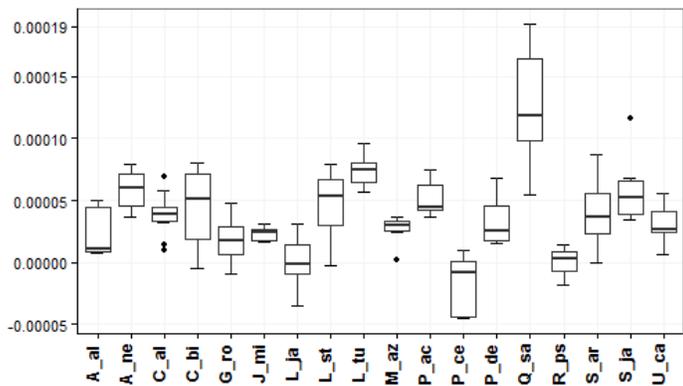
Para ampliar el análisis de las variables diferenciables, se investigó las variables seleccionadas por LDA usando el conjunto de datos que considera todas las variables (TV), para evaluar la contribución relativa de cada variable en la función discriminante de cada especie. La Tabla 10 resume las primeras cinco variables para las especies P_ce, Q_sa y S_ar, y en la Tabla 15 y Tabla 16 (Apéndice 5), para el resto de las especies. Se observa que LDA extrajo nuevas variables para discriminar las clases. Aunque este conjunto de variables tiene un menor desempeño de clasificación (Tabla 7), se ve que los índices que evalúan el contenido de antocianinas ARI2 y ARI1 se vuelven predeterminantes en P_ce (primero y segundo lugar de importancia). Del mismo modo, la pendiente M483 aparece como la variable que mejor discrimina a la especie Q_sa (Tabla 10), seguido por la variable WBI que mide el contenido de agua.

Tabla 10. Importancia de las variables (1 = mayor importancia y 5 = menor importancia) para cada especie según LDA utilizando todas las variables TV.

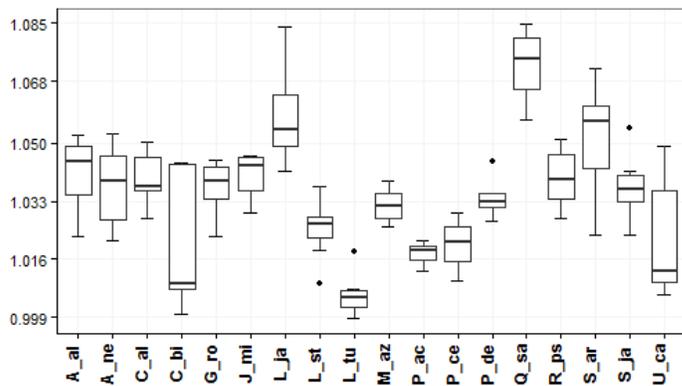
Tipo de variable	Variable	P_ce	Q_sa	S_ar
<i>Pendiente en el VIS</i>	M423			3
	M433			2
	M443			1
	M453			4
	M463		4	
	M473		3	
	M483		1	
	M563	3		
	M573	4		
	M583	5		
<i>Pendiente en el NIR</i>	M713			5
	M963		5	
<i>Índices de vegetación</i>	ARI1	2		
	ARI2	1		
	WBI		2	

Con el propósito de evidenciar la capacidad de éstas métricas para discriminar las especies, la Figura 12 presenta gráficos de cajas para una selección de variables. Como se puede observar, el índice ARI2 separa a P_ce del resto de las especies, pero tiene una gran dispersión entre las muestras evaluadas. Del mismo modo, la pendiente M563 separa las muestras de esta especie frente al resto. Por otro lado, el índice de carotenos CRI1 es la variable más importante para separar a J_mi (Tabla 16), sin embargo, al compararla frente al resto de las especies en la Figura 12, se aprecia que posee un traslape con la distribución de la especie L_ja, cual tienen una variabilidad mayor y que explicaría porque CRI1 no fue seleccionada por LDA en esta última especie.

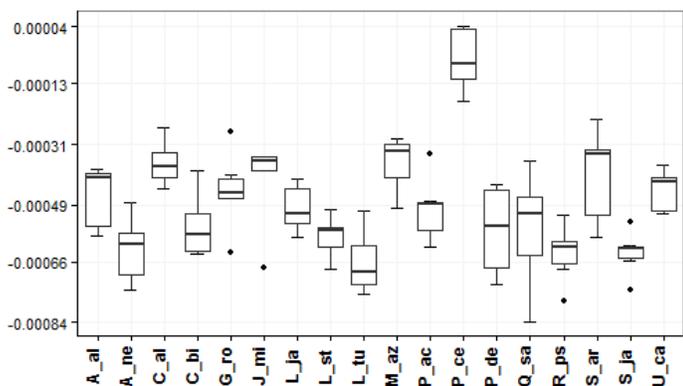
Por otra parte, la métrica que mide la eficiencia en el uso de la luz PRI (Sims & Gamon, 2002) obtuvo el mayor peso relativo en la función discriminante de R_ps (Tabla 16), pero en la Figura 12, se ve que la distribución de cuartiles se traslapa con A_al, A_ne y P_ac. Por último, Q_sa muestra valores elevados con las métricas M483 (>0.0001) y WBI (>1.068), y por esta razón, LDA otorga el primer y segundo peso de importancia en su función discriminante. Finalmente, al incorporar las variables derivadas de reflectancia (índices de vegetación y primera derivada), se obtiene una ordenación completamente diferente para S_ar y Q_sa (Tabla 10), sin repetirse ningún índice hiperespectral.



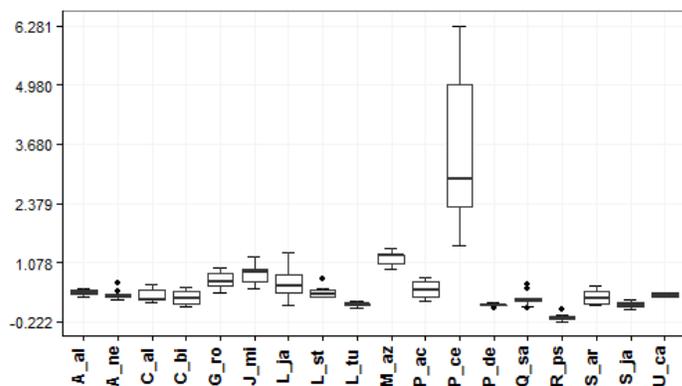
M483



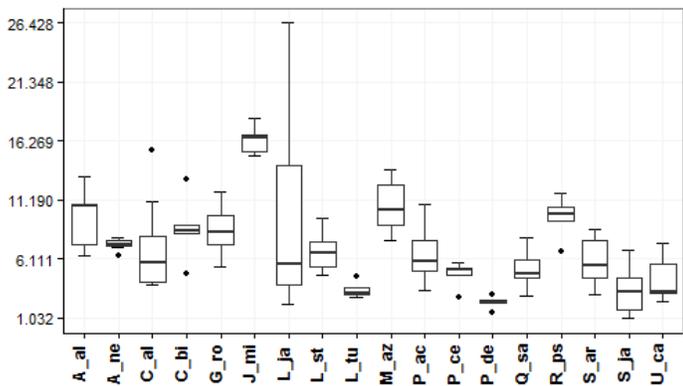
WBI



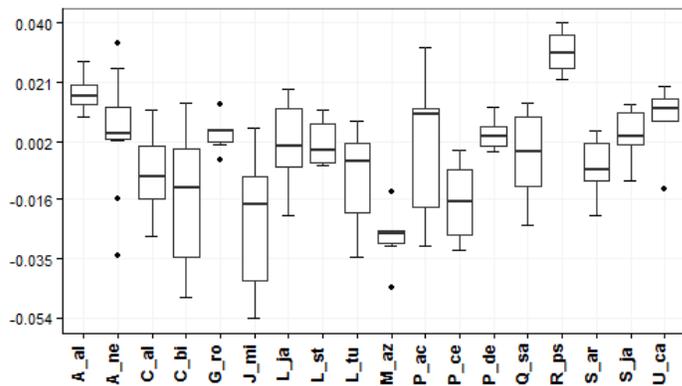
M563



ARI2



CRI1



PRI

Figura 12. Gráfico de caja de los valores medios para una selección de métricas espectrales para las especies evaluadas. Las métricas son: pendiente entre bandas 473-483 (M483); pendiente entre bandas 553-563 (M563); Photochemical Reflectance Index (PRI); Water Band Index (WBI); Anthocyanin Reflectance Index 2 (ARI2); y Carotenoid Reflectance Index 2 (CRI2).

5.4. Discusión

El protocolo de firmas espectrales foliares (FEF) se diseñó con el propósito de ser usado en campo bajo iluminación solar, de modo de garantizar, una fuente de radiación relativamente homogénea a lo largo del año (según fecha y hora de adquisición). Esta estandarización facilita la adquisición periódica y consecutiva de espectros foliares, condición necesaria según Pfitzner *et al.* (2011), para desarrollar un seguimiento espectral de las diferentes especies para su detección remota. El protocolo describe una secuencia de pasos para preparar las muestras vegetales y ubicar el equipo de experimentación (Figura 4), siendo replicable en cualquier parte del mundo por incorporar un archivo descriptivo de variables (*metadata*) estandarizadas en la literatura (Milton *et al.*, 2009; Rasaiah *et al.*, 2014).

Una de las desventajas del protocolo FEF, es la extracción y manipulación de las muestras foliares, primero porque la muestra vegetal se extrae del árbol objetivo y segundo, porque las hojas deben ser meticulosamente preparadas para la adquisición del registro (dentro del menor tiempo posible). Una solución apropiada es utilizar accesorios acoplados al radiómetro de campo, por ejemplo, la compañía ASD_{INC} (*Analytical Spectral Device*) ofrece dispositivos específicos para la adquisición de registros foliares, como la sonda de vegetación, que mide directamente la reflectancia del individuo sin extraer la muestra, sin embargo, el área medida es considerablemente inferior (< a 1 cm) y pueden existir variaciones a lo largo de la lámina del mesófilo (*e.g.* venas), como fue reportado en Castro-Esau *et al.* (2006b). Además, el sentido final del protocolo FEF, es integrar la reflectancia aportada por cada estructura de la lámina foliar, para utilizarla en función o en combinación con otros componentes (*e.g.* suelo), para comparar la reflectancia del árbol a escalas de mayor análisis espacial (Gamon *et al.*, 2006). En este sentido, los estudios de Clark & Roberts (2012) y Castro-Esau *et al.* (2004) analizan el aporte espectral de estructuras como cortezas de árboles y de lianas respectivamente, con el sentido de discriminar especies de árboles tropicales. Estos esfuerzos son, básicos para comprender la reflectividad total del dosel, pero como fue señalado por Knipling (1970), la medición espectral de una hoja no se puede obtener sin aplicar una modificación directa en su hábito de vida, lo que provoca que el porcentaje de reflectancia del dosel, tiene en general, una atenuación considerablemente inferior por variaciones en el ángulo de iluminación, la orientación de las hojas y por el trasfondo no-foliar. Este catastro espectral, es una base de referencia para comparar firmas de árboles en Santiago de Chile y, al compartir la información contenida en él, eventualmente se podrá derivar nuevos productos, como ha sucedido con el desarrollo de otras librerías espectrales (Hueni *et al.*, 2011).

En relación a las firmas espectrales del catastro, se pudo observar en la Figura 8 que responden al patrón típico de la vegetación, con una baja reflectancia en el visible y relativamente alta en el infrarrojo cercano. En relación a la variabilidad de las firmas espectrales por especie, se ve que algunas especies muestran mayor consistencia que otras, como plátano oriental (P_ac) y falso acacio (R_ps) y otras, como la catalpa (C_bi), la falsa acacia de Japón (S_ja), el ciruelo (P_ce) y el tulípero (L_tu) mostraron una amplia variación en alguna región del espectro. En relación al número de muestras por especie, es

recomendable aumentar el número de registros para mejorar el perfil espectral de la especie y conocer el grado máximo de variación, esto también fue señalado por Clark *et al.* (2005) y Cochrane (2000). En este sentido, en futuros estudios se debe seguir la propuesta de Pfitzner *et al.* (2011) de replicar la metodología o protocolo en distintas épocas del año y en diferentes condiciones ambientales, para saber si el comportamiento espectral del individuo está bajo la influencia de estrés o se debe a su variación natural (Feret *et al.*, 2008).

En relación con la separabilidad espectral de especies, la métrica de amplitud espectral D (Price, 1994) demuestra que la variación de reflectancia empieza a escala foliar y que para cada especie, existen regiones del espectro más o menos homogéneas (Figura 9). Esto coincide con los resultados obtenidos por Cochrane (2000) y Castro-Esau *et al.* (2006b). En este último, la amplitud espectral entre especies dio un valor promedio de $4 \pm 2,5$ de desviación estándar (S.D), similar al valor promedio obtenido aquí de $5 \pm 2,9$ S.D, ahora, si bien se utilizó el mismo sensor (ASD *FielSpec HH*), el método de adquisición, el intervalo analizado y el número de muestras es diferente, lo que podría explicar esta diferencia. En esta misma línea, interesante es el resultados de Féret & Asner (2011), porque obtienen un D promedio interespecífico de 5 ± 3 S.D, pero para un número mayor de especies (>100) y utilizando un dominio espectral superior (400-2500 nm) en el infrarrojo de onda corta (SWIR), con la intención de agregar información espectral con relación al contenido de agua. A pesar de éstas diferencias metodológicas, se puede decir que esta métrica es capaz de determinar las variaciones espectrales entre especies, manteniendo una magnitud similar entre distintos set de datos. Para estudios posteriores, sería novedoso acotar el D a intervalos específicos del espectro (*e.g.* azul, verde, bandas de absorción de agua en el infrarrojo), porque podría simplificar significativamente el análisis de variabilidad intraespecífica (Figura 9) y del mismo modo, considerar el dominio espectral en el infrarrojo de onda corta.

En relación a la separabilidad espectral de los registros foliares, las métricas hiperespectral de reflectancia arrojaron una precisión global mayor al 90% usando el análisis discriminante lineal (LDA), este resultado es similar a lo reportado por Castro-Esau *et al.* (2006), Féret & Anger (2011) y Fung *et al.* (2003), quienes obtuvieron precisiones de 80-85 %, 82-88 % y 87-89 % respectivamente. De hecho, en Castro-Esau *et al.* (2006b) se planteó una relación lineal entre la tasa de error y el número de especies, reportando que la tasa de error aumenta un 0,189 % cada 5 especies incluidas en el set de datos (intercepto igual a 12,348), luego si evaluamos esta función para las 18 especies aquí estudiadas, da un valor igual a 15,75 %, valor relativamente cercano al obtenido por LDA (9 %) y *Random Forest* (RF, 25 %) presentados en la Figura 11. Del mismo modo, Féret & Asner (2011) reportan un error del 12-18 % cuando evaluaron 20 especies con LDA. En tanto, con RF, se obtuvo una precisión global mayor al 70 %, inferior al resultado obtenido por Clark & Roberts (2012) de 86 % al clasificar 7 especies de árboles tropicales.

Esta caracterización espectral de los métodos, indica que existen métricas que gobiernan patrones por especie a escala foliar, sin embargo, estos no pueden explicarse sin un análisis histológico acabado, de hecho, en Castro-Esau *et al.* (2006b) se analizó el tejido interno de tres especies y se llegó a la conclusión que la reflectancia está directamente relacionada a las características foliares, tales como, el nivel de clorofila por unidad de área, el grosor de la

hoja y/o el porcentaje de aire en el mesófilo. Sin embargo, a partir de los resultados de Sims & Gamon (2003) en relación a la métrica *Water Band Index* (WBI), se encontró un importante resultado para la especie siempre verde *Quillaja saponaria*, ya que en este estudio se discriminó perfectamente la especie al utilizar esta métrica (Tabla 10) y directamente relacionado con la pendiente en el intervalo 953-963 nm (M963), lo que indicaría que esta especie tiene un alto contenido de agua en relación a las otras especies. Este hallazgo es un aporte a la región central del país, ya que la especie tiene una gran capacidad de regeneración natural ante desastres naturales y es dominante en la zona mediterránea del país. En el caso particular de La Reina, la especie tuvo un incremento significativo de 500 a casi 3000 individuos plantados en los últimos diez años (Hernández *et al.*, 2014).

En relación a los patrones espectrales detectados por especie (Tabla 14 y Tabla 15-16), se puede constatar la propiedad de los datos hiperespectrales al analizar bandas contiguas (Clark & Robert, 2012), lo que significa que para cada especie existen zonas determinadas del espectro que permiten su caracterización. Esto se evidenció en el ciruelo rojo, tanto al utilizar la reflectancia porcentual promedio como en la tasa de cambio de la reflectancia (primera derivada) en la transición del verde al rojo. Es pertinente señalar, que esta selección de variables buscó determinar las variables hiperespectrales que optimizan la discriminación de las especies, evitando la construcción de un modelo estadístico, ya que podrían existir problemas producidos por el sobreajuste de variables altamente correlacionadas y que en otro contexto, podría provocar una mala predicción por el uso de éstas métricas. Sin embargo, Tan & Li (2008) señalan que utilizar bandas vecinas con alta correlación sería más consistente que seleccionar bandas individuales para analizar la naturaleza espectral en el NIR, siendo más apropiado para la modelación e interpretación química por espectroscopia.

Con lo expuesto anteriormente, se establece que a pesar de existir variabilidad espectral intraespecífica, existe una precisión global suficiente para discriminar firmas espectrales foliares, lo que se podría deber al esquema de validación cruzada para entrenar el modelo de LDA, como fue señalado por Castro-Esau *et al.* (2006b). En consecuencia, el desempeño de los métodos de clasificación depende de la capacidad de las variables para detectar patrones intra e interespecífico y además, según cómo los métodos las seleccionan para construir los modelos predictivos. Por ejemplo, los métodos no paramétricos de árboles de decisión, generalmente tienen un desempeño pobre (Mitchell, 1997), probablemente por la complejidad del árbol y la tendencia al sobre ajuste al utilizar métricas con una gran variabilidad. En cambio, Immitzer *et al* (2012) señalan que éstos dos métodos pueden ser considerados clasificadores efectivos de especies, cada uno con ventajas frente el otro, por un lado LDA es más interpretable, pero por otro lado, RF es más robusto y teóricamente más flexible en los supuestos de distribución de las clases, además de evaluar cuales son las variables que mejor explican el conjunto de datos (Figura 15, Apéndice 5).

6. CONCLUSIÓN

Las librerías espectrales son una fuente de información valiosa para múltiples áreas de la percepción remota, ya que albergan información de referencia para analizar cambios territoriales relacionados con la ecología del paisaje. En este estudio se desarrolló un protocolo metodológico para adquirir firmas espectrales foliares, y con éste, se levantó un catastro de 130 muestras para 18 especies del arbolado urbano de la comuna de La Reina, Santiago, Chile. Este conjunto de datos es el primer acercamiento de firmas espectrales de vegetación urbana en el país, y el protocolo se podría extender a otras aplicaciones en el área forestal, en el área de monitoreo de plantas invasoras en zonas mineras y/o en agricultura de precisión.

El protocolo propuesto, mostró resultados coherentes con las firmas espectrales de la vegetación, ya que fue apto para distinguir el comportamiento espectral característico que distinguen a este grupo ambiental (componente) en la región del visible e infrarrojo cercano del espectro electromagnético. En consecuencia, se logró un catastro espectral de las principales especies analizadas en la comuna de La Reina.

A partir de este catastro, fue posible caracterizar la separabilidad espectral de la mayoría de las especies. Se alcanzó precisiones sobre el 90%, lo que indicaría que el esquema de clasificación permitió construir un eficiente modelo predictivo de las especies a escala foliar. Estos resultados se deben atribuir a la selección de variables altamente discriminantes (bandas espectrales, derivadas de primer orden e índices de vegetación) que se convierten en una herramienta poderosa para el reconocimiento de patrones en la vegetación.

Esta caracterización de patrones es clave para monitorear remotamente la vegetación urbana, considerada como un componente estratégico para el desarrollo sustentable de las ciudades. Para conseguir este objetivo, habría que continuar los esfuerzos de muestreo, aumentando el número muestras y la diversidad del catastro, realizando campañas de adquisición de registros en diferentes épocas del año, para comenzar a generar un banco nacional de firmas espectrales de vegetación y avanzar en el monitoreo fitosanitario.

7. BIBLIOGRAFÍA

- Asner, G. P. 1998. Biophysical and biochemical sources of variability in canopy reflectance. *Remote Sensing of Environment*, 64(February), 234. Retrieved from WOS:000074765100010
- Asner, G. P.; Martin, R. E.; Knapp, D. E.; Tupayachi, R.; Anderson, C.; Carranza, L.; Weiss, P. 2011. Spectroscopy of canopy chemicals in humid tropical forests. *Remote Sensing of Environment*, 115(12), 3587–3598.
- Blackburn, G. 1998. Quantifying Chlorophylls and Carotenoids at Leaf and Canopy Scales. *Remote Sensing of Environment*, 66(98), 273–285.
- Blackburn, G. 1999. Relationships between spectral reflectance and pigment concentrations in stacks of deciduous broadleaves. *Remote Sensing of Environment*, 70(99), 224–237.
- Blackburn, G. & Milton, E. J. 1995. Seasonal variations in the spectral reflectance of deciduous tree canopies., (February 2015), 37–41.
- Bojinski, S.; Schaepman, M. E.; Schläpfer, D.; & Itten, K. 2003. SPECCHIO: A spectrum database for remote sensing applications. *Computers and Geosciences*, 29, 27–38.
- Breiman, L. 2001. Random forests. *Machine Learning*, 45, 5–32.
- Broge, N. H. & Leblanc, E. 2001. Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote Sensing of Environment*, 76(2000), 156–172.
- Bruce, L. M. & Li, J. 2001. Wavelets for computationally efficient hyperspectral derivative analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 39(7), 1540–1546.
- Castro-Esau, K. L.; Sánchez-Azofeifa, G. a.; & Caelli, T. 2004. Discrimination of lianas and trees with leaf-level hyperspectral data. *Remote Sensing of Environment*, 90(3), 353–372.
- Castro-Esau, K. L.; Sánchez-Azofeifa, G. A.; & Rivard, B. 2006. Comparison of spectral indices obtained using multiple spectroradiometers. *Remote Sensing of Environment*, 103, 276–288.
- Castro-Esau, K. L.; Sanchez-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, 93(4), 517–530.

- 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, 1820–1855.
- Clark, M. L.; Roberts, D.; & 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.
- Cochrane, M. A. 2000. Using vegetation reflectance variability for species level classification of hyperspectral data. *Remote Sensing of Environment*, 21(10), 2075–2087.
- Curran, P.; Dungan, J. L.; & Gholz, H. L. 1995. Exploring the Relationship Between Reflectance Red Edge and Chlorophyll Content in Slash Pine. *Tre Physiology*, 15, 203–206.
- Deng, C. & Wu, C. 2013. A spatially adaptive spectral mixture analysis for mapping subpixel urban impervious surface distribution. *Remote Sensing of Environment*, 133, 62–70.
- Dennison, P. E. & Roberts, D. 2003. Endmember selection for multiple endmember spectral mixture analysis using endmember average RMSE. *Remote Sensing of Environment*, 87, 123–135.
- Feret, J. B.; François, C.; Asner, G. P.; Gitelson, A.; Martin, R. E.; Bidel, L. P. R.; Jacquemoud, S. 2008. PROSPECT-4 and 5: Advances in the leaf optical properties model separating photosynthetic pigments. *Remote Sensing of Environment*, 112(6), 3030–3043.
- Féret, J.-B. & Asner, G. P. 2011. Spectroscopic classification of tropical forest species using radiative transfer modeling. *Remote Sensing of Environment*, 115(9), 2415–2422.
- Fung, T.; Yan Ma, H. F.; & Siu, W. L. 2003. Band Selection Using Hyperspectral Data of Subtropical Tree Species. *Geocarto International*, 18(4), 3–11.
- Gamon, J. a.; Rahman, a. F.; Dungan, J. L.; Schildhauer, M.; & Huemmrich, K. F. 2006. Spectral Network (SpecNet)-What is it and why do we need it? *Remote Sensing of Environment*, 103, 227–235.
- Gitelson, A.; Chivkunova, O. B.; & Merzlyak, M. N. 2009. Nondestructive estimation of anthocyanins and chlorophylls in anthocyanic leaves. *American Journal of Botany*, 96(10), 1861–1868.
- Gitelson, A.; Merzlyak, M. N.; & Chivkunova, O. B. 2001. Optical properties and nondestructive estimation of anthocyanin content in plant leaves. *Photochemistry and Photobiology*, 74(1), 38–45.

- Hassan, A. M. & Lee, H. 2015. Land Use Policy Toward the sustainable development of urban areas : An overview of global trends in trials and policies. *Land Use Policy*, 48, 199–212.
- Hastie, T.; Tibshirani, R.; & Friedman, J. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer (2nd ed.). New York, NY, USA.
- Hernández, J.; Grez, I.; Magni, C.; & González, J. 2014. Plan Maestro de Conservación, Mantenimiento y Recuperación del Arbolado Comunal de La Reina. *Plan Maestro de Arbolado Público La Reina*. Universidad de Chile. Facultad de Ciencias Forestales y Conservación de la Naturaleza. Santiago, Chile.
- Herold, M.; Gardner, M. E.; & Roberts, D. 2003. Spectral resolution requirements for mapping urban areas. *IEEE Transactions on Geoscience and Remote Sensing*, 41(9), 1907–1919.
- Herold, M.; Roberts, D.; Gardner, M. E.; & Dennison, P. E. 2004. Spectrometry for urban area remote sensing - Development and analysis of a spectral library from 350 to 2400 nm. *Remote Sensing of Environment*, 91, 304–319.
- Hueni, A.; Malthus, T.; Kneubuehler, M.; & Schaepman, M. E. 2011. Data exchange between distributed spectral databases. *Computers and Geosciences*, 37, 861–873.
- Hueni, A.; Nieke, J.; Schopfer, J.; Kneubühler, M.; & Itten, K. I. 2009. The spectral database SPECCHIO for improved long-term usability and data sharing. *Computers and Geosciences*, 35, 557–565.
- Hueni, A. & Tuohy, M. 2006. Spectroradiometer data structuring, pre-processing and analysis – an IT based approach. *Journal of Spatial Science*, 51(2), 93–102.
- 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*, 4(9), 2661–2693.
- Jacquemoud, S. & Baret, F. 1990. PROSPECT: A model of leaf optical properties spectra. *Remote Sensing of Environment*, 34, 75–91.
- Jones, H. & Vaughan, R. 2010. Sampling, error, and scaling: In *Remote Sensing of Vegetation. Principles, techniques, and applications* (First, p. 384). Oxford University Press. New York, NY, USA.
- Kampe, T.; Johnson, B. R.; Kuester, M.; & Keller, M. 2010. NEON: the first continental-scale ecological observatory with airborne remote sensing of vegetation canopy biochemistry and structure. *Journal of Applied Remote Sensing*, 4(1), 043510.

- Kampe, T.; Leisso, N.; Musinsky, J.; Petroy, S.; & Karpowicz, B. 2013. *The NEON 2013 Airborne Campaign at Domain 17 Terrestrial and Aquatic Sites in California*.
- Kiang, N. Y.; Siefert, J.; Govindjee; & Blankenship, R. E. 2007. Spectral signatures of photosynthesis. I. Review of Earth organisms. *Astrobiology*, 7(1), 222–251.
- Knipling, E. B. 1970. Physical and physiological basis for the reflectance of visible and near-infrared radiation from vegetation. *Remote Sensing of Environment*, 1(3), 155–159.
- Kotthaus, S.; Smith, T. E. L.; Wooster, M. J.; & Grimmond, C. S. B. 2014. Derivation of an urban materials spectral library through emittance and reflectance spectroscopy. *ISPRS Journal of Photogrammetry and Remote Sensing*, 94, 194–212.
- Kuhn, M. & Johnson, K. 2013. *Applied Predictive Modeling*. Springer (2nd ed.). New York, NY, USA.
- Max Kuhn (2015). Contributions from Jed Wing, Steve Weston, Andre Williams, Chris Keefer, Allan Engelhardt, Tony Cooper, Zachary Mayer, Brenton Kenkel, the R Core Team, Michael Benesty, Reynald Lescarbeau, Andrew Ziem and Luca Scrucca. caret: Classification and Regression Training. R package version 6.0-41. Disponible en: <http://CRAN.R-project.org/package=caret>
- McPherson, E. G. & Simpson, J. R. 1999. Carbon dioxide reduction through urban forestry: Guidelines for Professional and Volunteer Tree Planters. *Forest Research*, 237p.
- Merzlyak, M. N.; Gitelson, A. a; Chivkunova, O. B.; & Rakitin, V. Y. U. 1999. Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening. *Physiologia Plantarum*, 106(1), 135–141.
- Milton, E. J.; Fox, N. P.; & Schaepman, M. E. 2009. Progress in field spectroscopy. *International Geoscience and Remote Sensing Symposium (IGARSS)*, 113, 1966–1968.
- Nowak, D. J. & Crane, D. E. 2002. Carbon storage and sequestration by urban trees in the USA. *Environmental Pollution*, 116, 381–389.
- Nowak, D. J.; Greenfield, E. J.; Hoehn, R. E.; & Lapoint, E. 2013. Carbon storage and sequestration by trees in urban and community areas of the United States. *Environmental Pollution*, 178, 229–236.
- Ollinger, S. V. 2011. Sources of variability in canopy reflectance and the convergent properties of plants. *New Phytologist*, 189, 375–394.
- Pal, M. 2005. Random forest classifier for remote sensing classification. *International Journal of Remote Sensing*, 26(1), 217–222.

- Peñuelas, J.; Baret, F.; & Filella I. 1995. Semi empirical indices to asses carotenoids chlorophyll a ratio from leaf spectral reflectance. *Photosynthetica*, 31(2), 221–230.
- Pfritzer, K.; Bartolo, R.; Carr, G.; Esparon, A.; & Bollhöfer, A. 2011. *Standards for reflectance spectral measurement of temporal vegetation plots*.
- Pfritzer, K.; Bollhöfer, a.; & Carr, G. 2006. A standard design for collecting vegetation reference spectra: Implementation and implications for data sharing. *Journal of Spatial Science*, 51(2), 79–92.
- Price, J. C. 1994. How Unique Are Spectral Signatures ?, *186*, 181–186.
- R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- Rasaiah, B. a.; Jones, S. D.; Bellman, C.; & Malthus, T. J. 2014. Critical metadata for spectroscopy field campaigns. *Remote Sensing*, 6, 3662–3680.
- Sims, D. A. & Gamon, J. A. 2002. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages 2002.pdf, *81*, 337–354.
- Sims, D. & Gamon, J. 2003. Estimation of vegetation water content and photosynthetic tissue area from spectral reflectance: a comparison of indices based on liquid water and chlorophyll absorption features. *Remote Sensing of Environment*, 84, 526–537.
- Tan, C. & Li, M. 2008. Mutual information-induced interval selection combined with kernel partial least squares for near-infrared spectral calibration. *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, 71(4), 1266–1273.
- Tooke, T. R.; Coops, N. C.; Goodwin, N. R.; & Voogt, J. a. 2009. Extracting urban vegetation characteristics using spectral mixture analysis and decision tree classifications. *Remote Sensing of Environment*, 113, 398–407.
- Ustin, S. L.; Gitelson, A.; Jacquemoud, S.; Schaepman, M. E.; Asner, G. P.; Gamon, J. a.; & Zarco-Tejada, P. J. 2009. Retrieval of foliar information about plant pigment systems from high resolution spectroscopy. *Remote Sensing of Environment*, 113, S67–S77.
- Vogelmann, J. E.; Rock, B. N.; & Moss, D. M. 1993. Red edge spectral measurements from sugar maple leaves, (January 2015), 37–41.
- Zarco-Tejada, P. J.; Miller, J. R.; Noland, T. L.; Mohammed, G. H.; & Sampson, P. H. 2001. Scaling-up and model inversion methods with narrowband optical indices for chlorophyll content estimation in closed forest canopies with hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 39(7), 1491–1507.

8. APÉNDICE

Apéndice 1: Índices Hiperespectrales de Vegetación

Tabla 11. Índices de vegetación (VI)

Categoría	VI	Aplicación	Ecuación	Autor
Verdor	<i>Red Edge Normalized Difference Vegetation Index (NDVI705)</i>	Contenido de clorofila	$\frac{\rho_{750} - \rho_{705}}{\rho_{750} + \rho_{705}}$	Sims & Gamon, 2002
	<i>Modified Red Edge Simple ratio Index (mSR705)</i>	Contenido de clorofila	$\frac{\rho_{750} - \rho_{445}}{\rho_{705} + \rho_{445}}$	Sims & Gamon, 2002
	<i>Modified Red Edge Normalized Difference Vegetation Index (mNDVI705)</i>	Contenido de clorofila y corrige reflexión especular	$\frac{\rho_{750} - \rho_{705}}{\rho_{750} + \rho_{705} - 2\rho_{445}}$	Sims & Gamon, 2002
	<i>Vogelman Red Edge Index 1 (VOG1)</i>	Contenido de clorofila y agua por área foliar	$\frac{\rho_{740}}{\rho_{720}}$	Vogelmann <i>et al.</i> , 1993
	<i>Vogelman Red Edge Index 2 (VOG2)</i>	Contenido de clorofila y agua por área foliar	$\frac{\rho_{734} - \rho_{747}}{\rho_{715} + \rho_{726}}$	Vogelmann <i>et al.</i> , 1993
	<i>Triangular Vegetation Index (TVI)</i>	Relaciona el pick del verde, el mínimo de clorofila y NIR	$0.5[120(\rho_{750} - \rho_{550}) - 200(\rho_{670} - \rho_{550})]$	Broge & Leblanc, 2002
	Eficiencia uso de luz	<i>Photochemical Reflectance Index (PRI)</i>	Eficiencia de xantofilas y relación carotenos/clorofila	$\frac{\rho_{531} - \rho_{570}}{\rho_{531} + \rho_{570}}$
<i>Structure Insensitive Pigment Index (SIPI)</i>		Maximiza la razón entre fibras de carotenos y clorofila	$\frac{\rho_{800} - \rho_{445}}{\rho_{800} - \rho_{680}}$	Peñuelas <i>et al.</i> , 1995
Sequedad o senescencia de carbón	<i>Plant Senescence Reflectance Index (PSRI)</i>	Maximiza la relación entre carotenos y clorofila	$\frac{\rho_{680} - \rho_{500}}{\rho_{750}}$	Merzlyak <i>et al.</i> , 1999
	<i>Red-edge Vegetation Stress Index (RSVI)</i>	Cambio del <i>red-edge</i> por estrés y contenido de clorofila.	$\frac{\rho_{714} + \rho_{752}}{2} - \rho_{733}$	Merton & Huntington 1999
Pigmento foliar	<i>Carotenoid Reflectance Index 1 (CRI1)</i>	Contenido de carotenos	$\left(\frac{1}{\rho_{510}}\right) - \left(\frac{1}{\rho_{550}}\right)$	Gitelson <i>et al.</i> , 2002
	<i>Carotenoid Reflectance Index 2 (CRI2)</i>	Contenido de carotenos	$\left(\frac{1}{\rho_{510}}\right) - \left(\frac{1}{\rho_{700}}\right)$	Gitelson <i>et al.</i> , 2002
	<i>Anthocyanin Reflectance Index 1 (ARI1)</i>	Contenido de antocianinas	$\left(\frac{1}{\rho_{550}}\right) - \left(\frac{1}{\rho_{700}}\right)$	Gitelson <i>et al.</i> , 2001
	<i>Anthocyanin Reflectance Index 2 (ARI2)</i>	Contenido de antocianinas	$\rho_{800} \left[\left(\frac{1}{\rho_{550}}\right) - \left(\frac{1}{\rho_{700}}\right) \right]$	Gitelson <i>et al.</i> , 2001
Índice de agua	<i>Water Band Index (WBI)</i>	Contenido de agua del dosel	$\frac{\rho_{900}}{\rho_{970}}$	Peñuelas <i>et al.</i> , 1995

Fuente: Elaborado a partir de Clark & Roberts (2012), Ollinger (2011), Sims & Gamon (2002) y Vogelmann *et al.* (1993).

Apéndice 2: Ejemplo *metadata* del protocolo de firmas espectrales foliares

Tabla 12. Ejemplo de variables para las categorías de campaña de la *metadata*.

N°	Campos de la campaña	Valor
1	Información general de la campaña y aplicación específica	Firmas espectrales foliares de vegetación urbana en Santiago de Chile.
2	Método de colecta del material vegetal	Colecta de material vegetal a exposición ecuatorial, a más de 3 m y guardadas herméticamente.
3	Método de preparación de la muestra para su registro espectral	Manipulación de hojas: se removió polvo, peciolo y ramas.
4	Descripción del equipo experimental para la medir la firma espectral	El ASD se instala sobre un trípode, se ajustó altura y geometría de visión para controlar el área leída por el sensor. Se utilizó una goma negra opaca como material de trasfondo.

Tabla 13. Ejemplo de variables para las categorías de campaña de la *metadata*.

N°	Campos del registro espectral	Variable	Valor
1	Descripción taxonómica y condición de la muestra vegetal	Familia	Quillajaceae
		Especie	<i>Quillaja saponaria</i>
		Condición de la muestra (*)	B
2	Ubicación geográfica del sitio de adquisición	Latitud (°)	-33.438583°
		Longitud (°)	-70.565095°
		Altitud (m.s.n.m.)	625
3	Descripción espacial del sitio de adquisición	Exposición abierta al sol, espacio plano y con presencia de edificaciones a 20 m.	Imagen
4	Tiempo de medida (fecha y hora) al instante de adquisición	UTC - 04:00	15:25 hrs
		Fecha (dd/mm/aa)	24/04/ 2014
5	Nubosidad del cielo al instante de adquisición	Octas	0
6	Señal espectral medida por el instrumento	Reflectancia porcentual (nm)	325-1025
		Geometría de visión (°)	0
		Tipo de soporte	Trípode
		Altura del sensor sobre el suelo (m)	1.5
		Altura del sensor sobre la superficie de la muestra (m)	0.1
		Método de muestreo (N° de registros)	10
7	Técnica de campo experimental	Tiempo de medición de la muestras (s)	1.1
		Tiempo de calibración (s)	1
		Modelo espectro radiómetro	ASD <i>HandHeld</i> 1
		Campo de Visión Angular (°)	25
		Intervalo de muestreo espectral (nm)	1.5
8	Instrumentos y equipos utilizados	Panel de referencia de calibración	<i>Spectralon</i>

(*) La condición fitosanitaria refiere a la calidad de la muestra vegetal (e.g. B = Buena; M= Media; P= Pobre; S =Seco).

Apéndice 3: Firmas Espectrales Foliaras de Reflectancia

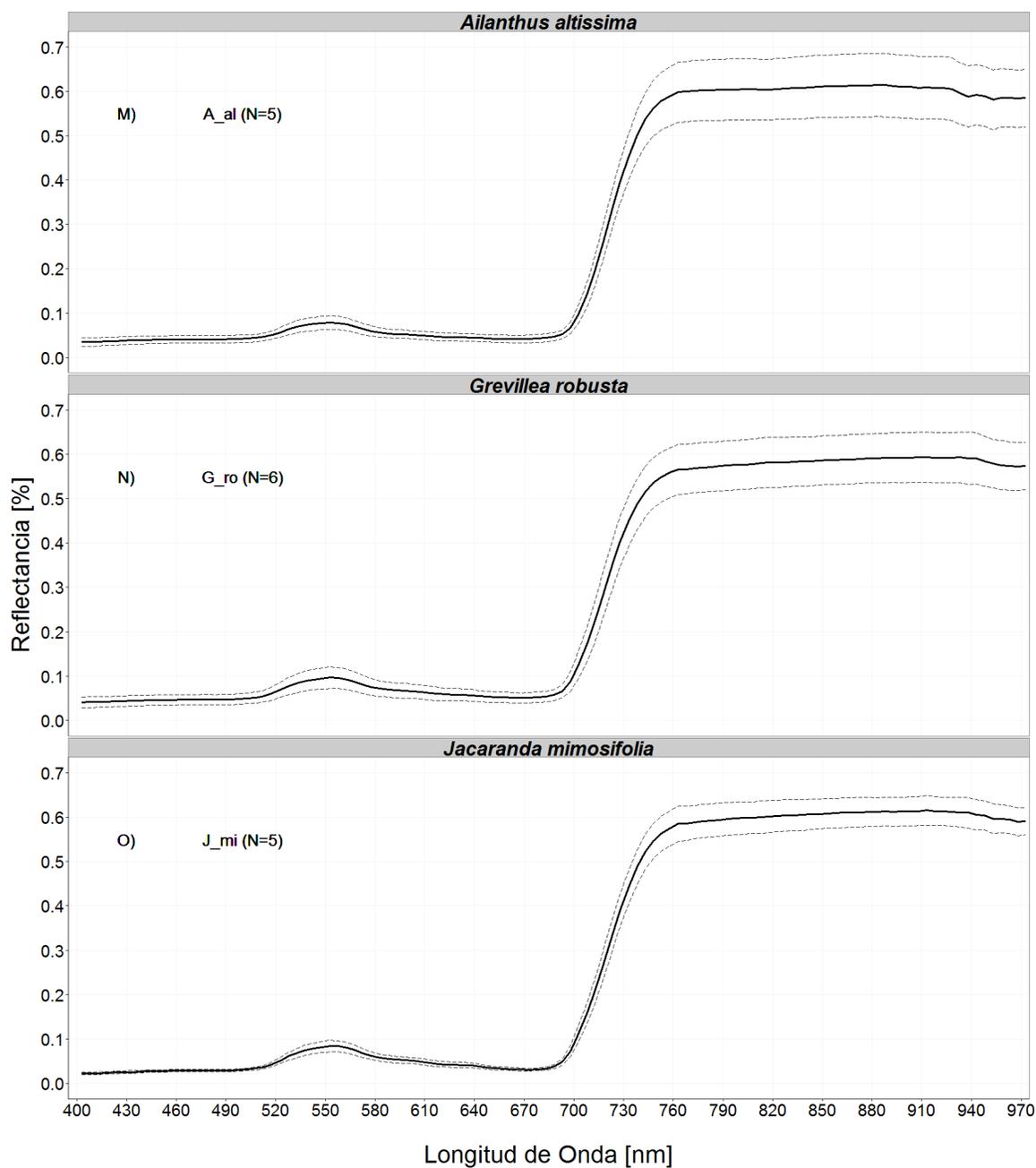


Figura 13. Firmas espectrales foliares de reflectancia promedio \pm 1 S.D. (línea punteada). Código y número de registros por especie

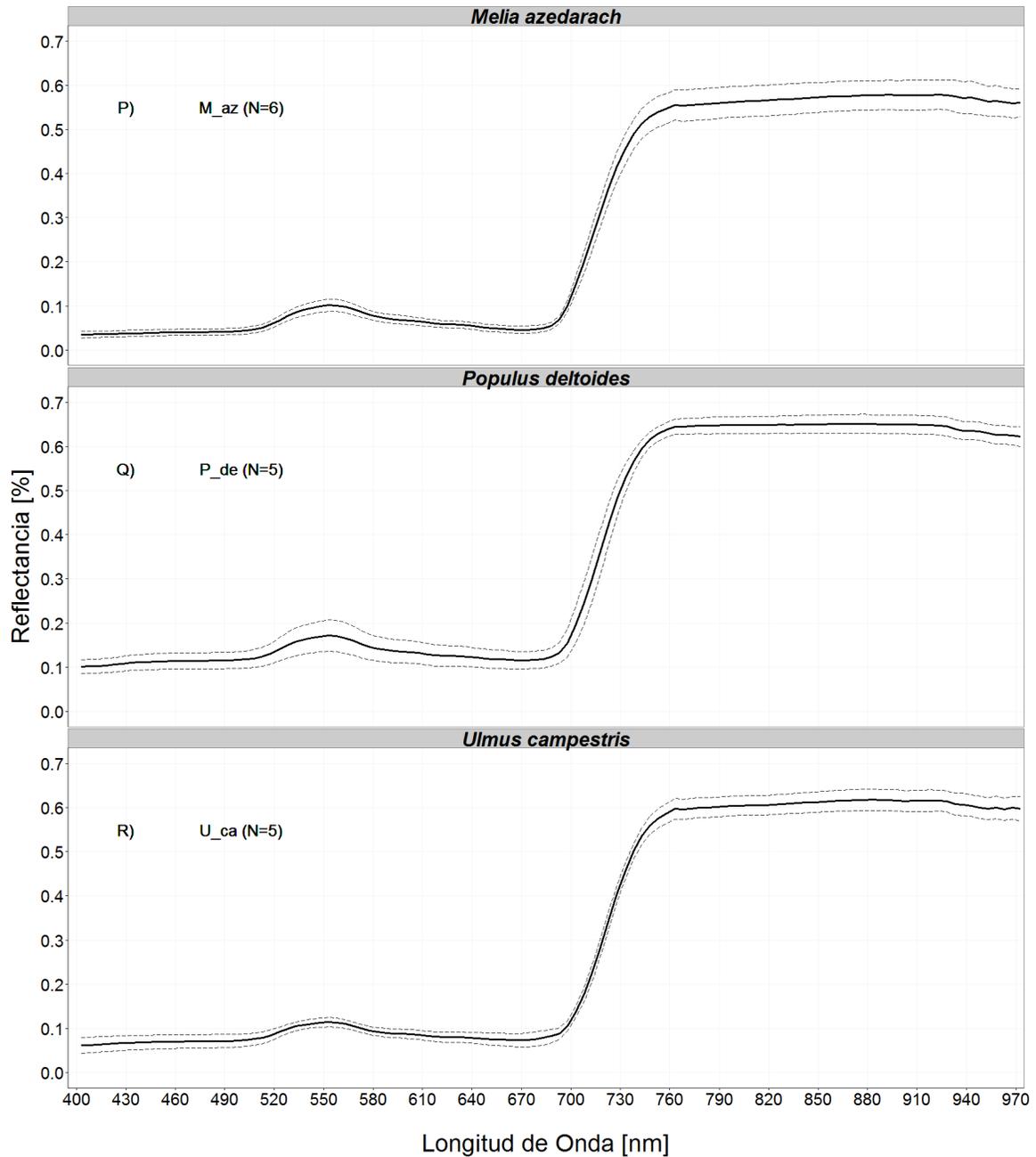


Figura 14. Firmas espectrales foliares de reflectancia promedio \pm 1 S.D. (línea punteada). Código y número de registros por especie

Apéndice 4: Extracción de variables por Análisis Discriminante Linear

Tabla 14. Importancia de las variables (1 = mayor importancia y 5 = menor importancia) para cada especie según LDA utilizando la reflectancia % promedio cada 5 nm.

Rango Longitud de Onda	REF % cada 5 nm	A_al	A_ne	C_al	C_bi	G_ro	J_mi	L_ja	L_st	L_tu	M_az	P_ac	P_ce	P_de	Q_sa	R_ps	S_ar	S_ja	U_ca
Visible (VIS x28)	403							1										3	
	408							2										2	
	413							4										1	
	418																	4	
	438													3					
	443													2					
	448		5											1					
	453													4					
	458													5					
	498											5							5
	503							5				3							
	508							2				2							
	513							1				1							
	518							4				4							
	523						5												
	533						3												
	538													5					
	543													4					
	548													2					
	553													1					
558										5			3						
563										4									
568						4													
573		5				1													
578		4				2													
688							3												
693																	4		
698		2								3							1		
Infrarrojo cercano (NIR x17)	703	1	3							1							2		
	708	3	1						4	2							3		3
	713		2	4					2								5		1
	718		4	2					1										2
	723			1					3										4
	728			3					5						4		4		
	733			5											1		1		
	738														2		2		
	743								3						3		3		
	748								5						5		5		
	928					3							5						
	933					1							4						
	938					5													
	943					4							2						
	948					2							1						
	953												3						

Tabla 15. Importancia de las variables (1 = mayor importancia y 5 = menor importancia) para cada especie según LDA utilizando todas las variables (Reflectancia promedio, índices vegetacionales y primera derivada)

Tipo de variable	Variable	A_al	A_ne	C_al	C_bi	G_ro	J_mi	L_ja	L_st	L_tu	M_az	P_ac	P_ce	P_de	Q_sa	R_ps	S_ar	S_ja	U_ca
<i>Visible</i> (VIS x 12)	403																		3
	408																		2
	413																		1
	418																		4
	438													3					
	443													2					
	448													1					
	453													4					
	458													5					
	498																	5	
563									5										
698									4										
<i>Infrarrojo cercano</i> (NIR x 8)	703									2									
	708									3									
	928				3							5							
	933				1							4							
	938				5														
	943				4							2							
	948				2							1							
	953											3							
<i>Pendiente en el VIS</i> (x 19)	M413							2											5
	M423										5							3	
	M433			3				4								5		2	
	M443			1														1	
	M453							3										4	
	M463													4		2			
	M473													3					
	M483													1					
	M503		1																
	M513		2																
	M523		3																
	M563												3						
	M573												4						
	M583												5						
	M593								2										
	M633								1										
	M643									1									
	M683		4	2															
	M693		5																

Tabla 16. Continúa.

Tipo de variable	Variable	A_al	A_ne	C_al	C_bi	G_ro	J_mi	L_ja	L_st	L_tu	M_az	P_ac	P_ce	P_de	Q_sa	R_ps	S_ar	S_ja	U_ca
<i>Pendiente en el NIR (x 16)</i>	M713																5		
	M723	5																	
	M753			5															
	M793					2	5												
	M803					4					1								2
	M813					1													
	M823						3												
	M833										3								1
	M853					3													3
	M883					5													4
	M893	3								5									
	M903								5										
	M923									4									
	M943									3									
	M953								1										
M963														5					
<i>Índices de vegetación (VI x 12)</i>	ARI1											2							
	ARI2											1							
	CRI1						1												
	CRI2						2												
	mSR705	1					4												
	NDVI705	2																	
	PRI										2					1			
	PSRI										4								
	RSVI			4															
	VOG1	4															4		
	VOG2																3		
	WBI														2				

Apéndice 5: Selección de variables por Random Forest

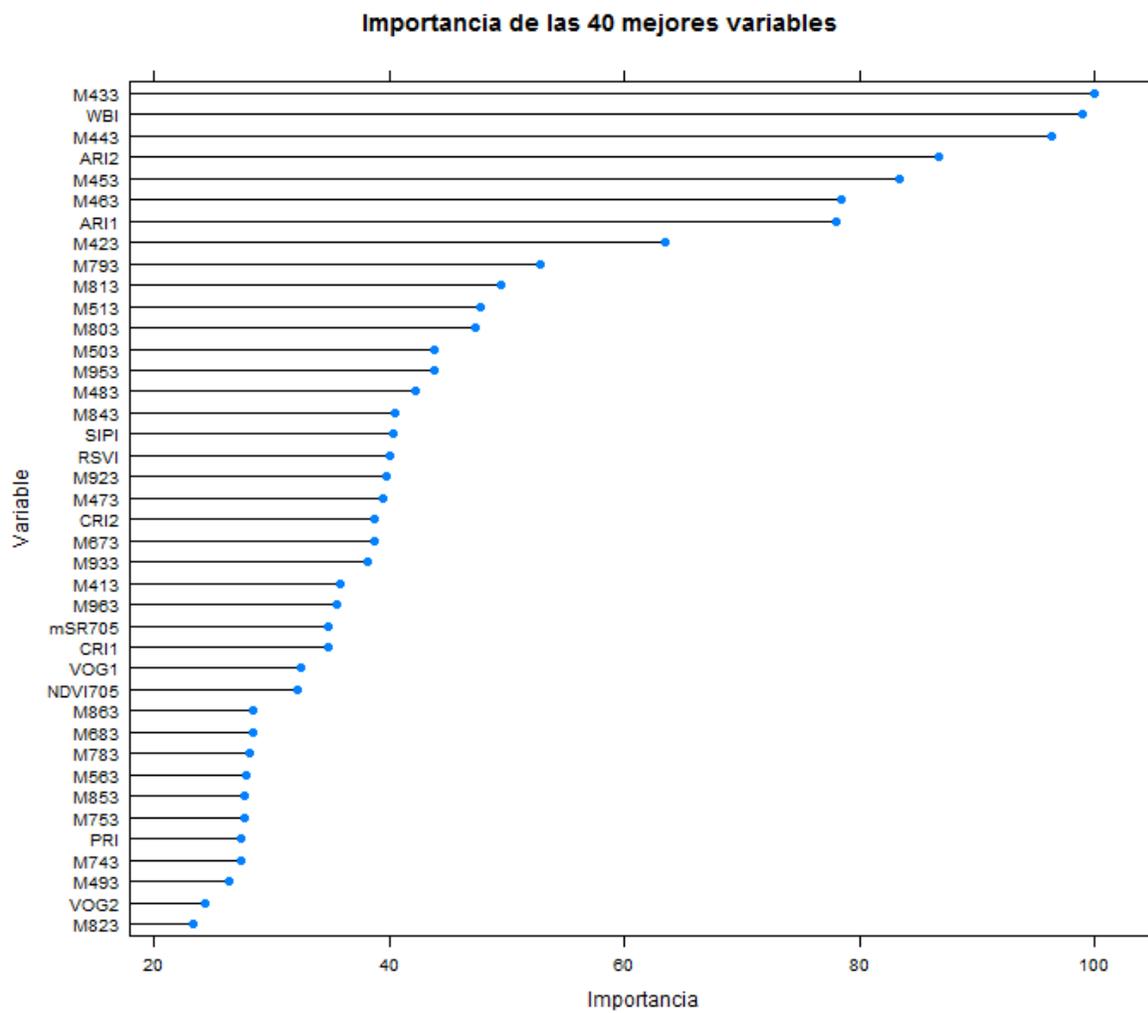


Figura 15. Importancia de las primeras 40 variables según el método de clasificación *Random Forest* para métricas espectrales del subconjunto combinado de índices de vegetación y derivadas cada 10 nm.