



Caracterización espectral de arvenses que afectan los cultivos de *Saccharum* spp., métodos de obtención y procesamiento de datos: Una revisión

Spectral characterization of weeds affecting *Saccharum* spp., data collection and processing methods: A review

Rubén Orozco-Morales^{1*} , Katia Ojito-Ramos^{2*} , Osmany de la C. Aday Díaz³ ,
Luis Hernández Santana⁴ 

¹ Centro de Investigaciones de Métodos Computacionales y Numéricos en la Ingeniería, Universidad Central "Marta Abreu" de Las Villas, carretera a Camajuaní km 5,5, Santa Clara 54830, Cuba

² Departamento de Biología, Facultad de Ciencias Agropecuarias, Universidad Central "Marta Abreu" de Las Villas, carretera a Camajuaní km 5,5, Santa Clara 54830, Cuba

³ Instituto de Investigaciones de la Caña de Azúcar Villa Clara (INICA Villa Clara), Cuba

⁴ Grupo de Automatización, Robótica y Percepción, Universidad Central "Marta Abreu" de Las Villas, carretera a Camajuaní km 5,5, Santa Clara 54830, Cuba

INFORMACIÓN DEL ARTÍCULO

Recibido: 16/05/2025

Aceptado: 23/07/2025

CONFLICTO DE INTERESES

Los autores declaran no existir conflicto de intereses

CORRESPONDENCIA

Rubén Orozco Morales

rorozco@uclv.edu.cu

Katia Ojito Ramos

kojito@uclv.edu.cu



CF: cag101252458

RESUMEN

Contexto: La caña de azúcar (*Saccharum* spp.) es uno de los principales cultivos de muchos países, en el cual el manejo de las arvenses es esencial. La caracterización de las propiedades de reflectancia espectral de la vegetación es la base de diversos métodos para discriminar entre arvenses y *Saccharum* spp., utilizados en vehículos terrestres o aéreos no tripulados. **Objetivos:** Realizar una actualización sobre elementos relacionados con la caracterización espectral de arvenses que afectan el cultivo de *Saccharum* spp., así como los métodos de obtención y el procesamiento de datos espectrales. **Métodos:** Se realizó una revisión sistemática narrativa de la bibliografía científica con un enfoque cualitativo para evaluar tendencias y consensos, así como las brechas de conocimiento y futuras líneas de investigación. **Resultados:** Uno de los inconvenientes en la identificación de especies de plantas a partir de sus propiedades espectrales es la variabilidad de las características de reflectancia que se observa a causa de múltiples factores, por lo que trabajos recientes toman más en cuenta el *bundle* de características medias de diversas muestras que el tradicional concepto de firma espectral como una característica única. Se discute la necesidad del análisis de la normalidad de los datos, muestra la utilidad de las medidas

de correspondencia espectral e ilustra algunos enfoques útiles del análisis de componentes principales y el análisis de varianza. **Conclusiones:** Los resultados pueden constituir la base de futuras investigaciones que caractericen espectralmente arvenses y cultivares de *Saccharum* spp.

Palabras clave: agricultura de precisión, caña de azúcar, control de malezas, firma espectral, reflectancia foliar

ABSTRACT

Context: The Sugarcane (*Saccharum* spp.) is one of the main crops in many countries, where weed management is essential. The characterization of the spectral reflectance properties of vegetation is the basis for various methods used to discriminate between weeds and *Saccharum* spp., employed in ground vehicles or unmanned aerial vehicles. **Objectives:** To provide an update on elements related to the spectral characterization of weeds that affect the cultivation of *Saccharum* spp., as well as methods for obtaining and processing spectral data. **Methods:** A systematic narrative review of the scientific literature was conducted using a qualitative approach to assess trends and consensus, as well as knowledge gaps and future lines of research. **Results:** One of the drawbacks in identifying plant species based on their spectral properties is the variability in reflectance characteristics observed due to multiple factors. Therefore, recent studies take into account the *bundle* of average characteristics of various samples rather than the traditional concept of spectral signature as a single characteristic. The need for data normality analysis is discussed, the usefulness of spectral correspondence measures is demonstrated, and some useful approaches to principal component analysis and analysis of variance are illustrated. **Conclusions:** The results may form the basis for future research characterizing weeds and cultivars of *Saccharum* spp. spectrally.

Keywords: leaf reflectance, precision agriculture, spectral signature, sugarcane, weed control

INTRODUCCIÓN

Después de la pandemia de COVID-19 y en medio de varios conflictos regionales, el costo de los alimentos se ha incrementado, siendo difícil para muchos países no solo adquirirlos, sino también acceder a insumos para producirlos; entre ellos, los herbicidas. La geomática y la agricultura de precisión desarrollan un conjunto de tecnologías que tienen como propósito optimizar el uso de recursos e insumos, garantizar la sustentabilidad de la actividad agrícola e incrementar sus rendimientos. Las pérdidas en la agricultura a causa de las arvenses pueden ser considerables si no se les maneja adecuadamente, principalmente durante la primera fase del crecimiento del cultivo. Entre los métodos de control de arvenses, el control químico es el más utilizado debido a que la aplicación de herbicidas es un método rápido y eficaz, pero el uso indiscriminado de estos no solo incrementa los costos de producción, sino que también es dañino para el medio ambiente, siendo cada vez mayor el interés por los alimentos ecológicos.

La caña de azúcar (*Saccharum* spp.) constituye uno de los principales cultivos a nivel mundial pero sus rendimientos agrícolas se ven seriamente afectados por la presencia de las arvenses que influyen de forma negativa en el desarrollo de las plantas (FAOSTAT, 2023), las cuales varían de una región del planeta a otra (Ronay *et al.*, 2022; Galon *et al.*, 2023). La distribución espacial de las arvenses, la topografía del terreno, la variabilidad espacial del tipo de suelo, el drenaje y el microclima, pueden conducir a un uso eficiente de los herbicidas (Su, 2020). Para ello, la agricultura de precisión puede reducir de forma inteligente las pérdidas en rendimiento producidas por las arvenses, así como los daños al medio ambiente producido por los herbicidas.

Utilizando como referencia la reflectancia espectral de arvenses y el cultivo se han desarrollado múltiples estrategias no invasivas para reconocer entre unas u otras de una forma eficaz y mucho más eficiente que con los métodos tradicionales basados en la observación directa (Che'Ya *et al.*, 2022). Las características

de reflectancia espectral son la base de los métodos que se utilizan para identificar arvenses basados en imágenes hiperespectrales o multiespectrales, ya sea obtenidas mediante teledetección desde satélites, aviones, o vehículos aéreos no tripulados, así como directamente en el terreno por medio de vehículos terrestres no tripulados, a escala de laboratorios y hasta aplicaciones para teléfonos móviles (Su, 2020). Para ello, se han desarrollado bibliotecas de reflectancias espectrales de diversos cultivos y arvenses, al igual que para minerales, suelos, materiales desarrollados por el hombre, etc. Estas bibliotecas se almacenan en bases de datos digitales para su utilización, ya sea en tareas de teledetección mediante imágenes hiper- o multiespectrales, en detección directamente en el terreno, o para fines científicos por parte de investigadores. Todo ello ha dado lugar a lo que se denomina caracterización espectral de especies, que proporcionan muy buenos resultados para la identificación de plantas, y representan una técnica de muestreo rápida y de bajo costo (Soca-Muñoz, 2020; Kumar *et al.*, 2021; Roslin *et al.*, 2021; Johnson *et al.*, 2023).

El objetivo del presente trabajo fue realizar una revisión exploratoria sobre elementos relacionados con la caracterización espectral de arvenses que afectan el cultivo de *Saccharum* spp., así como los métodos de obtención y el procesamiento de datos espectrales, para proveer de una actualización para los investigadores que de introducen en este tema, constituyendo una base para futuras investigaciones.

MATERIALES Y MÉTODOS

Se realizó una revisión sistemática narrativa de la bibliografía científica para lo cual se consultaron las bases de datos Scopus y Web of Science. Se consideraron palabras clave específicas relacionadas con el tema como «VIS-NIR», «spectral signature», «*Saccharum* spp.», «weeds» y combinaciones de estos términos con métodos de obtención y procesamiento de datos espectrales. Se aplicaron filtros por idioma (inglés y español), tipo de documento («artículo» y «revisión»), y un período de 2000 a 2024 para garantizar la relevancia de los resultados. Ambas bases de datos se fusionaron, se revisaron las duplicaciones y se excluyeron las publicaciones no directamente relacionadas con el tema central para evitar sesgos en los análisis. Los registros obtenidos se exportaron en

formato CSV y BibText. Se adoptó un enfoque cualitativo para evaluar tendencias y consensos, así como las brechas de conocimiento y futuras líneas de investigación.

RESULTADOS Y DISCUSIÓN

Características espectrales de las plantas

Los resultados de diversos trabajos demuestran que los cultivos pueden distinguirse con precisión de las arvenses y entre varias arvenses basándose en la reflectancia espectral, dando lugar a la estrategia denominada manejo de arvenses específico del sitio. Esta estrategia es un enfoque preciso y eficiente en cuanto a recursos, adapta la gestión de las arvenses a las variaciones de su localización, densidad y composición dentro de cada campo, dando lugar a prácticas agrícolas más productivas y sostenibles.

De esta forma, la aplicación de herbicida se regula de acuerdo a las necesidades locales, reportándose una disminución entre un 40% y un 80% de su consumo sin afectar los rendimientos de los cultivos (Su, 2020). Para ello se necesitan los mapas de arvenses, los cuales se obtienen a diferentes escalas, ya sea mediante la teledetección, que puede proporcionar un mapeado rápido de grandes áreas de cultivos, o directamente en el terreno. En estos mapas georreferenciados, los píxeles segmentados como arvenses indican su localización exacta diferenciándose de los píxeles asociados al cultivo principal (Ronay *et al.*, 2022). De esta forma, las distintas especies de arvenses, que tienen características fisiológicas diferentes, recibirán estrategias diferentes para contrarrestar su crecimiento (Norasma *et al.*, 2020).

Dado que la característica espectral de una especie de planta no es única, sino que plantas de una misma especie pueden mostrar determinada variabilidad espectral, se han dedicado grandes esfuerzos al desarrollo de algoritmos que incorporan diferentes estrategias que permitan que la combinación de diversos componentes varíe dentro de una imagen hiperespectral, utilizando, por ejemplo, paquetes de firmas espectrales conocidas *a priori*. Así, las bibliotecas espectrales utilizadas por estos métodos se denominan a veces paquetes y, en principio, deberían dar cuenta de todas las variaciones posibles de cada especie de planta (Borsoi *et al.*, 2021).

En particular, se ha mostrado la posibilidad de diferenciar las plantas de *Saccharum* spp. de las ar-

venses por el comportamiento espectral de las hojas (de Souza *et al.*, 2020; Aday *et al.*, 2023). La firma espectral de *Saccharum* spp. depende de las mismas variables biofísicas que también determinan las propiedades ópticas de otros tipos de vegetación. Según Som-ard *et al.* (2021), estas pueden ser variables estructurales/morfológicas, coeficientes de absorción, dispersión y transmisión de la hoja, y la reflectancia de fondo del suelo.

Entre las variables estructurales/morfológicas, se encuentran el índice de área foliar, la inclinación media del ángulo foliar, la densidad y el agrupamiento de las plantas, etc. Los coeficientes de absorción, dispersión y transmisión de la hoja están en función de la pigmentación, el contenido de agua y la anatomía de la hoja. La reflectancia de fondo del suelo está en función del material parental, el contenido de materia orgánica, la humedad de la superficie y la rugosidad. Además, dado que estas variables varían con el tiempo, se hace necesario disponer de todas las pistas necesarias y rasgos espectrales de la dinámica temporal de los cultivos de *Saccharum* spp., útiles para la correcta identificación y mapeo de diferentes plantaciones de cultivos (Som-ard *et al.*, 2021).

El hecho de que existan diferencias regionales en el patrón del cultivo, también indica límites naturales para un mapeo y monitoreo perfectos, ya que los patrones espectro-temporales muestran una amplia variabilidad y, por lo tanto, un fuerte solapamiento con otros cultivares. Además, a pesar de que un conocimiento detallado de la firma espectral de un manto de *Saccharum* spp. es importante para el análisis preciso y el despliegue de datos de teledetección, al igual que otros cultivos, las características de distribución de reflectancia bidireccional de los cultivos de *Saccharum* spp. no son Lambertianas.

Por lo tanto, los ángulos de iluminación y observación, junto con la orientación y el espaciado de las hileras, tienen un profundo efecto en la reflectancia espectral medida remotamente. Todo ello da lugar a lo que se denomina variabilidad espectral, que es la variación que muestran las características de reflectancia espectral de una misma especie de planta, lo cual puede dar lugar a errores significativos en la mezcla hiperespectral. Es por ello, que las bases de datos de características espectrales, que abundan en el campo de la teledetección y son utilizadas sistemáticamente por los científicos como ayuda para sus

análisis, no pueden considerarse resultados perennes o absolutos. Las aplicaciones para la identificación de especies de plantas o la detección de insectos plaga y enfermedades deben basarse en la capacidad de adquirir el conjunto de características espectrales representativo de la planta y su estado fitosanitario (Som-ard *et al.*, 2021; Che'Ya *et al.*, 2022).

Recopilación de datos espectrales

Para poder tomar decisiones a partir de bibliotecas de firmas espectrales, o para que estas bibliotecas sean eficaces en los métodos hiperespectral supervisado o semi-supervisado, es necesario seguir prácticas correctas en la recolección de los datos (Katsoulas *et al.*, 2016). Con todo ello, los datos brindan siempre una representación muy específica de las verdaderas propiedades espectrales de cada planta, por lo que definir con exactitud la configuración seguida para su recopilación permite conocer el marco de validez de los mismos. Las firmas espectrales que se encuentran en las bases de datos representan una configuración de recogida de datos específica. En ello interviene la fuente de iluminación, ya sea artificial o natural, pues para obtener datos espectrales de un objeto, sus enlaces químicos deben ser estimulados por energía externa.

En condiciones de laboratorio, se suelen emplear fuentes de energía artificiales, mientras que las mediciones sobre el terreno dependen sobre todo del sol, aunque algunas tecnologías utilizan fuentes de luz artificial. Uno de los problemas de la iluminación natural es el rápido cambio de las condiciones de iluminación, incluso en días despejados (Jafarbiglu y Pourreza, 2023). Por lo tanto, la calibración periódica del instrumento con respecto a una referencia blanca es de gran importancia para obtener lecturas coherentes y libres de metamerismo.

En este sentido, Ronay *et al.* (2024) recomiendan que los metadatos que debería contener una biblioteca de firmas espectrales deben ser: ubicación del lugar, hora/fecha, condiciones del cielo, detalles del instrumento sensor, geometría de visualización, altura del sensor sobre el suelo e información sobre la banda de longitudes de onda. La falta de metadatos, así como la inadecuada organización de los datos espectrales, pueden hacer que la información espectral sea inútil, ya que se desconocen las circunstancias de su captura.

Es común recopilar los datos cuando los cultivos y las arvenses aún están verdes, tomando medidas de reflectancia espectral de al menos 10 hojas individuales en el laboratorio, utilizando un círculo integrador que mide la reflectancia de una pequeña área de la hoja frente a una lámpara estándar. Para cada especie, se recoge material vegetal de campos individuales, asegurando que se tome una parte similar de las plantas. Long *et al.* (2021) y Johnson *et al.* (2023) recomiendan que en el caso de las hojas muy estrechas se coloquen secciones de hojas individuales inmediatamente adyacentes (pero no superpuestas) para obtener una superficie de tamaño suficiente para permitir la evaluación mediante el círculo integrador.

La espectroscopia de las plantas en el rango de longitudes de onda del espectro visible (Vis), entre 380 y 780 nm, revela la información del color. No obstante, las plantas producen la mayor reflectancia en el espectro infrarrojo. Este se divide en tres regiones: el infrarrojo cercano (NIR), también denominado infrarrojo de onda corta (SWIR), con longitudes de onda entre 780 y 2 500 nm, el infrarrojo medio (MIR), entre 2 500 y 25 000 nm, dentro del cual se encuentra el infrarrojo térmico (TIR), con longitudes de onda entre 10 600 y 12 500 nm, y el infrarrojo lejano, también denominado infrarrojo de onda larga (LWIR), entre 25 000 y 300 000 nm. El valor de la reflectancia en la zona del espectro NIR puede utilizarse para analizar el estiramiento y la flexión de los enlaces químicos, incluidos S-H, C-H, hidroxilo (O-H) y N-H, mientras que la zona del espectro MIR proporciona información característica relacionada con los grupos funcionales químicos. Ello está asociado a que varios compuestos de las plantas absorben la energía de la radiación electromagnética en mayor cantidad a determinadas longitudes de onda, por ejemplo, α -carotenoide (420, 440 y 470 nm), antocianina (400 - 550 nm), β -carotenoide (425, 450 y 480 nm), clorofila a (435, 670 - 680 y 740 nm), clorofila b (480 y 650 nm), luteína (425, 445 y 475 nm), humedad (970, 1450 y 1944 nm) y violaxantina (425, 450 y 475 nm). Estos compuestos, aunque presentes en las hojas de todas las plantas, varían su concentración de una especie a otra, permitiendo diferenciar entre cultivos y arvenses. Además, la reflectancia espectral también se ve afectada

por la estructura celular y la estructura física de la superficie de la planta (Su, 2020; Long *et al.*, 2021). Aunque las tres regiones del espectro mencionadas proporcionan una cantidad sustancial de información, el borde rojo (RE), con longitudes de onda entre 700 y 745 nm, se considera la principal fuente de información para los estudios de vegetación. Esto se debe a que el RE es el rasgo espectral más ordinario que se encuentra en la vegetación, persistiendo durante todo el desarrollo de la planta hasta las etapas iniciales de la senescencia. No obstante, el uso del RE se considera limitado para determinar el contenido de clorofila de las hojas, el índice de área foliar, la biomasa o la discriminación de especies y el estado de salud de la vegetación (Almeida y Filho, 2004). Por otra parte, Long *et al.* (2021) plantean que, aunque la región LWIR ha sido poco utilizada para la discriminación entre especies, hay varios estudios que demuestran la aplicabilidad y el potencial de esta a partir de la posible relación de los rasgos de la reflectancia de esta región con los compuestos químicos de la superficie de la hoja, como los compuestos cerosos que forman la cutícula de la hoja o la celulosa de las paredes celulares de la epidermis superior. Estos autores notifican también que hay investigaciones que demuestran que las hojas presentan rasgos espectrales distintivos en la región del infrarrojo térmico, y es posible discriminar los tipos de vegetación a nivel de especie. Por el contrario, también sostienen que el uso de todas las bandas desde el SWIR hasta el LWIR puede sesgar el análisis espectral, ya que los rasgos impulsados por compuestos de contraste relativamente bajo pueden quedar oscurecidos por los abrumadores rasgos del agua. Por lo antes expuesto es importante que se elijan adecuadamente los rangos de longitudes de ondas en la zona del espectro útil.

Preprocesamiento de los datos espectrales

El preprocesamiento de los datos tiene varios propósitos y sus consecuencias influyen significativamente en la posterior utilidad de los datos preprocesados, principalmente en tareas de análisis o clasificación (Sonobe y Hirono, 2022).

Entre los procedimientos utilizados en el preprocesamiento de los datos está la eliminación de bandas ruidosas o no calibradas, las que no aportan en la caracterización espectral de las especies, o en la

comparación entre ellas, con el fin de lograr la mayor relación señal-a-ruido (Piron *et al.*, 2008). Como se mostró en la sección anterior, lo más usual es desear longitudes de onda por debajo de 400 nm y por encima de 1 000 nm, aunque cada estudio limita el ancho de banda solo al necesario, de acuerdo a sus propósitos. De igual manera, el preprocesamiento también incluye eliminar las características atípicas (*outliers*), en función de los valores de reflectancia, producidas generalmente por errores de medición. También pueden obtenerse bandas ruidosas o que no proporcionan ninguna señal útil debido a una baja relación señal-a-ruido. Ello puede deberse a defectos del sensor o a interferencias durante la medición. Por ejemplo, en mediciones de campo las relaciones señal-a-ruido bajas se encuentran de forma natural en algunos rangos de longitud de onda debido a las interferencias atmosféricas, por ejemplo, el vapor de agua a 1 350 - 1 440 nm, 1 790 - 1 990 nm y 2 360 - 2 500 nm. El vapor de agua es la causa del mayor ruido encontrado en los datos de espectroscopia de campo. Sólo cuando la distancia entre el sensor y el objeto detectado es mínima la influencia de la atmósfera será prácticamente inexistente (Ronay *et al.*, 2024).

Otro de los propósitos del preprocesamiento de los datos es atenuar el ruido de reflectancia, denominado en ocasiones como perturbación en la característica espectral, proceso que debe realizarse

preservando al máximo los rasgos esenciales de la característica espectral, tales como pequeñas o grandes transiciones de interés al tiempo que se atenúan las fluctuaciones de alta frecuencia. Una de las causas de este ruido radica en la alta resolución espectral, lo que redundaría en un estrecho ancho de banda de cada una de las bandas del espectrómetro, el cual es muchas veces de tan solo 1 nm. Al disminuir el ancho de banda de cada banda se reduce la relación señal-a-ruido. El ruido (alta frecuencia) o el fondo (baja frecuencia) pueden eliminarse mediante métodos de preprocesamiento, mejorando las predicciones y la separación de grupos (especies) en el espacio espectral. A tales efectos, Laroche-Pinel *et al.* (2021) recomiendan promediar los espectros.

La figura 1 muestra el suavizado de las 50 características de reflectancia originales de muestras de *Saccharum* spp. y de *Albizia procera* (Roxb.) Benth. El suavizado obtenido es un compromiso entre la atenuación de las fluctuaciones de alta frecuencia y la preservación de la pendiente de las principales variaciones de reflectancia, como la producida en la frontera entre las regiones Vis y NIR del espectro (Aday *et al.*, 2023).

El suavizado mediante filtro de Savitzky-Golay es capaz de resaltar transiciones y picos al tiempo que suaviza los datos, suprimiendo el ruido aleatorio (Chandra Pandey *et al.*, 2020). A pesar de ello, la pendiente de la transición Vis-NIR puede redu-

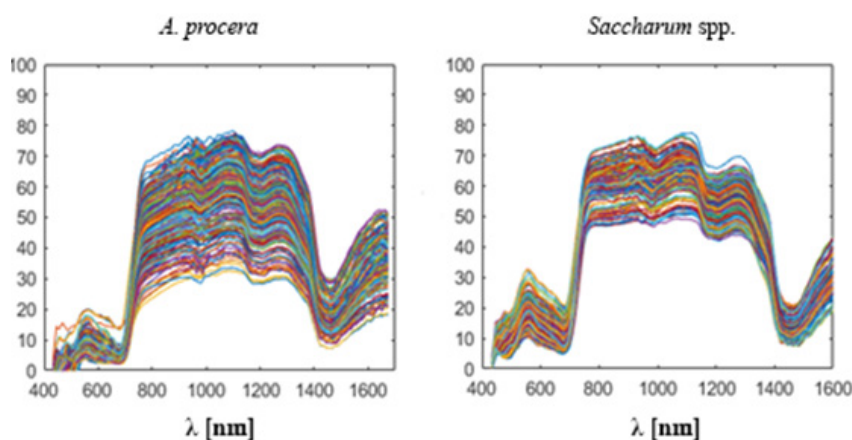


Figura 1. Características de reflectancia espectral suavizadas con filtro FIR de promedio móvil: 502 de *Albizia procera* y 524 de *Saccharum* spp. (Aday *et al.*, 2023).

cirse más después del suavizado con el filtro de Savitzky-Golay, que con el de promedio móvil, mientras que las fluctuaciones entre 900 y 1 000 nm son fuertemente atenuadas. Otra desventaja del filtro de Savitzky-Golay es que depende de dos parámetros de control (orden polinomial y longitud del marco), cuya sintonía puede resultar ardua. Por su parte, el filtrado *wavelet*, al ser localmente adaptable, suaviza las regiones con ruido de pequeña amplitud afectando también la pendiente de la transición Vis-NIR, mientras que preserva las fluctuaciones entre 900 y 1 000 nm. Dado que se ha desarrollado una gran diversidad de *kernels wavelet*, además del tipo *symlet*, también puede resultar arduo encontrar una combinación de orden y tipo de *kernel* satisfactoria para cada situación específica.

Reducción de la dimensionalidad de los datos

La alta resolución espectral da lugar a una alta redundancia de datos, ya que a lo largo de determinadas longitudes de onda la información es similar. Es decir, las longitudes de onda vecinas tienen un alto grado de correlación. Una práctica común en el preprocesamiento de los datos de reflectancia espectral es reducir su dimensionalidad con fines diversos, entre los que se encuentran reducir su volumen en bases de datos, disminuir la carga computacional que demanda su procesamiento, o investigar las bandas ideales necesarias para determinadas aplicaciones (Li *et al.*, 2022).

Para ello, se han aplicado varias estrategias como tomar los datos espectrales cada cierta cantidad de longitudes de onda. Esto es, un diezmado del espectro, por ejemplo, se dividen todos los espectros en intervalos de 5 nm (Che'Ya *et al.*, 2013) o 10 nm (Norasma *et al.*, 2020; Roslin *et al.*, 2021; Che'Ya *et al.*, 2022), en comparación con el valor original de 1 nm. De esta forma, se reduce el volumen de datos en un factor de 5 o 10. No obstante, según Ronay *et al.* (2024) este enfoque debe tratarse con precaución, ya que puede violar el teorema del muestreo de Shannon, el cual establece que las muestras discretas son una representación completa de la señal si el ancho de banda es inferior a la mitad de la frecuencia de muestreo.

La teoría del muestreo de Shannon es aplicable siempre que la función de entrada sea de banda limitada, permitiendo obtener su transformada discreta de

Fourier, siempre que el número de puntos espectrales de la transformada sea igual a superior a la cantidad de muestras que tenga la secuencia obtenida de la señal muestreada. Cuando se produce el diezmado del espectro (equivalente a un muestreo en el dominio de la frecuencia), la secuencia asociada a la función de entrada sufre tanto que se distorsiona, como del fenómeno Gibbs que le genera ondulaciones (Ronay *et al.*, 2024).

Se han utilizado otros enfoques para reducir la dimensionalidad de los datos. Uno de ellos es la transformación de componentes principales (*principal component transformation*, PCT), la cual realiza una descomposición en valores propios y vectores propios que se utilizan para construir una matriz de transformación que se aplica a los datos originales. Los componentes se ordenan en función de su capacidad para describir la variación de los datos, siendo los primeros los que explican la mayor parte de la variación. Otro método utilizado ha sido la fracción de ruido máximo (*maximum noise fraction*, MNF), similar a la PCT, pero resuelve el punto débil de esta última cuando la varianza del ruido no es uniforme en todas las bandas del conjunto de datos. La MNF es una transformación lineal y ordena los componentes resultantes por su relación señal-a-ruido, por lo que también es conocida como componentes principales ajustados al ruido (*noise-adjusted principal components*), siendo una técnica útil para reducir la dimensionalidad de un conjunto de datos conservando la mayor parte de la información y minimizando el ruido al mismo tiempo.

Selección de bandas de interés

Asociado a la reducción de la dimensionalidad están los métodos de selección de bandas de interés. Las mejores bandas son un subconjunto de las bandas originales que maximizan la separabilidad de las clases, por lo que estos métodos buscan las bandas no correlacionadas, desechando aquellas en las cuales existe una alta correlación en la reflectancia entre sus longitudes de onda, las cuales no aportan o aportan poco en la discriminación entre especies.

Entre los métodos utilizados para seleccionar bandas está el ángulo espectral sólido N-dimensional (*N-dimensional solid spectral angle*), basado en la diferencia geométrica de la forma espectral entre un conjunto de espectros. Utiliza el ángulo espectral en

un espacio N-dimensional para cuantificar la similitud entre múltiples objetivos y encontrar bandas que definan el mayor contraste espectral entre espectros (Long *et al.*, 2021; Marañón *et al.*, 2023).

Además, se encuentra el método variable de selección con bosque aleatorio (*selection with random forest*), que se basa en la precisión de la clasificación mediante bosques aleatorios en los que las muestras dadas se dividen aleatoriamente en datos de entrenamiento y de prueba. Este método ha demostrado un sólido rendimiento en el tratamiento de los datos espectrales de la vegetación en las bandas Vis y NIR. Además, el método de mínima redundancia máxima relevancia (*minimum redundancy maximum relevance*), que se basa en la teoría de la información, ha demostrado un excelente rendimiento en diversas aplicaciones de explotación de datos en los campos de la biología y la informática (Long *et al.*, 2021; Marañón *et al.*, 2023).

Prueba de normalidad de los datos

Dado que muchos de los análisis que se realizan con los datos de las características de reflectancia espectral son estadísticos, es necesario, ante todo, tener certeza de la normalidad o no de tales datos, o cuánto difiere la distribución de los datos observados respecto a lo esperado si procediesen de una distribución normal. Se sabe que los modelos, tales como los procesos Gaussianos (GP), han sido muy exitosos para modelar parámetros de la vegetación a partir de sus espectros de reflectancia. Según Gewali *et al.* (2019), los GP permiten modelar la distribución de probabilidad de estimación de la reflectancia, lo cual es más práctico que sus valores, proveyendo además la posibilidad de estimar valores de confianza y establecer un nivel de incertidumbre.

Métodos de análisis de las características espectrales

El análisis de los datos es la tarea principal a que se someten las características de reflectancia espectral obtenidas tras ser preprocesadas. Ello permite extraer toda la información posible del comportamiento espectral de cada especie y encontrar la mayor utilidad con vistas a la clasificación de las diversas especies, estimar el contenido de sus constituyentes, evaluar el estado de salud, predecir rendimiento, etc. Entre los diversos métodos de análisis de datos hípe-

respectrales utilizados se encuentran los análisis de gradientes, las medidas de correspondencia espectral (*spectral matching measures*, SMM), el análisis discriminante lineal (*linear discriminant analysis*, LDA) y el análisis discriminante lineal por pasos (*stepwise LDA*), la transformación de Fourier en el infrarrojo cercano (*Fourier-transform near-infrared*), el análisis discriminante cuadrático (*quadratic discriminant analysis*), el análisis de componentes principales (*principal component analysis*, PCA), la regresión logística (*logistic regression*), la regresión de mínimos cuadrados parciales (*partial least square regression*), la máquina de vectores de soporte (*support vector machine*, SVM), el bosque aleatorio (*random forest*), la red neuronal artificial (ANN) y la red neuronal convolucional. Además, otras técnicas no paramétricas como el análisis de conglomerados y los mínimos cuadrados parciales. También, se reporta la utilización del modelado independiente por analogía de clase (*soft independent modelling by class analogy*) y el análisis de varianza (ANOVA). Las SVM y las ANN son ejemplos de métodos de clasificación supervisada comúnmente utilizados (de Souza *et al.*, 2020; Su, 2020; Yuan *et al.*, 2020; Long *et al.*, 2021; Che'Ya *et al.*, 2022; Ktenioudaki *et al.*, 2022; Sulaiman *et al.*, 2022; Pandey *et al.*, 2023).

El proceso básico consiste en determinar primero los criterios discriminantes basándose en la categoría conocida de la especie y en el conocimiento previo y, a continuación, calcular la función discriminante. Pueden mencionarse también los métodos de regresión dispersa, bajo el supuesto de que los espectros observados pueden representarse como combinaciones lineales de firmas espectrales, por lo que estos métodos utilizan principalmente algunas bibliotecas espectrales conocidas como colecciones de endmembers (Yuan *et al.*, 2020; Su, 2020; de Souza *et al.*, 2020; Long *et al.*, 2021; Che'Ya *et al.*, 2022; Ktenioudaki *et al.*, 2022; Sulaiman *et al.*, 2022; Pandey *et al.*, 2023). A continuación, se profundiza en la aplicación de algunas de estas medidas de análisis espectrales.

Análisis de gradientes

El gradiente (primera derivada) ofrece información sobre los cambios de reflectancia en cada longitud de onda a partir de resaltar las variaciones de reflectancia entre bandas adyacentes, como los que

ocurren en la región de inflexión del borde rojo (Sonobe y Hirono, 2022). Como resultado, cambios graduales de reflectancia entre longitudes de onda adyacentes son eliminados, ya que por lo general no están asociados a rasgos característicos del espectro. Contrariamente, cambios abruptos de los valores de reflectancia entre longitudes de onda adyacentes son convertidos a un valor numérico asociado a su gradiente. Esos cambios de alto gradiente son los que generalmente están asociados a picos de reflectancia o de absorción característicos de los elementos constituyentes de la especie en estudio. No obstante, dado el carácter pasoalto de la primera derivada, es necesario atenuar previamente el ruido de alta frecuencia en la característica espectral mediante un suavizado que no afecte la pendiente de las transiciones significativas. Por ejemplo, Roslin *et al.* (2021), utilizando el gráfico de la primera derivada para cada especie de arvenses, mostraron que la región entre 710 nm y 750 nm puede ser un indicador para identificar las especies de arvenses. Esta región del infrarrojo cercano está relacionada con el proceso de fotosíntesis, así como con los pigmentos biogeoquímicos y foliares de las especies de arvenses. También, Che'Ya *et al.* (2013) utilizaron la

primera derivada para discriminar entre seis tipos de arvenses y el sorgo a partir del pico a 720 nm.

A modo de ejemplo, la figura 2 contiene las primeras derivadas de las características medias de reflectancia espectral de *Saccharum* spp. y *A. procera* (Aday *et al.*, 2023). Se puede observar, en la parte inferior de la figura, que se ha trazado el gradiente medio de cada especie elevado al cuadrado y suavizado con filtro de promedio móvil. En este sentido, se hace notar que las diferencias de gradiente pueden ser rasgos característicos útiles para discriminar entre especies de plantas, como sucede a $\lambda = 730,48$ nm y a $\lambda = 1395,91$ nm.

Medidas de correspondencia espectral

Las SMM comparan las características de los espectros de reflectancia de dos especies dentro de un intervalo seleccionado de longitudes de onda y devuelven un valor numérico de acuerdo al grado de correspondencia entre los espectros. El valor del grado de correspondencia entre el espectro incógnito y el de referencia compilado en bibliotecas espectrales permite aceptar o rechazar la pertenencia del espectro incógnito a la clase del de referencia (Pathak *et al.*, 2023). En ocasiones, la comparación solo en algunas bandas seleccionadas, teniendo en cuenta

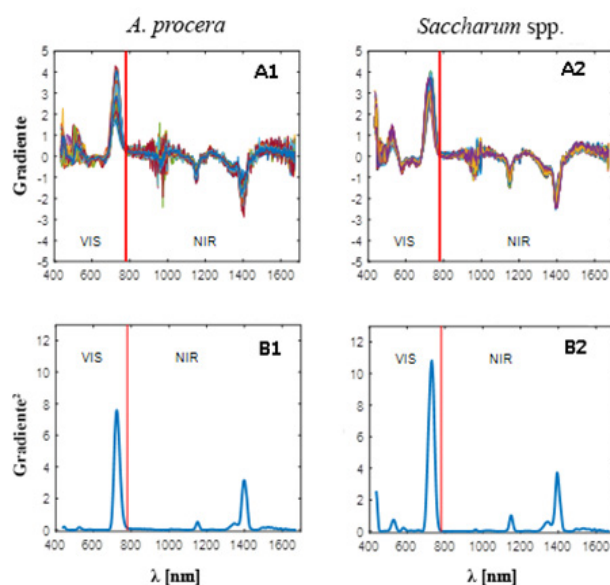


Figura 2. Primeras derivadas de las características medias de reflectancia espectral de *Albizia procera* y *Saccharum* spp. A1 y A2) Promedio de los gradientes de las muestras y suavizado. B1 y B2) Gradientes de las muestras elevado al cuadrado, de *Albizia procera* y *Saccharum* spp. La línea roja separa la región del espectro visible (VIS) de la del infrarrojo cercano (NIR) (Aday *et al.*, 2023).

ta que la forma de un espectro de reflectancia puede dividirse en regiones amplias que cambian suavemente y que definen la forma general del espectro y regiones estrechas en forma de depresión asociadas a rasgos de absorción (Padma y Sanjeevi, 2014). Así, una de las estrategias más comunes consiste en comparar únicamente los rasgos de absorción de cada espectro e ignorar las demás partes.

Es amplia la diversidad de métodos desarrollados para medir la correspondencia entre dos características espectrales. No obstante, entre los métodos clásicos de SMM más reportados para comparar características de reflectancia espectral están la distancia Euclidiana espectral (*spectral Euclidean distance*, SED), el ángulo espectral medio (*spectral angle mapper*, SAM), la medida de la correlación espectral (*spectral correlation measure*, SCM), y la medida de la divergencia de la información espectral (*spectral information divergence*, SID), así como métodos híbridos, como entre SID y SAM (SIDSAM). Se han utilizado otros métodos clásicos como la distancia de Bhattacharaya (BhattD), la distancia de Jeffries-Matusita (JMD) o la similitud de Bray-Curtis (Borsoi *et al.*, 2021; Pathak *et al.*, 2023). Kumar *et al.* (2021) proponen el coeficiente Dice de similitud espectral (*Dice spectral similarity coefficient*, DSCC), el coeficiente de similitud espectral de Kumar-Johnson (*Kumar-Johnson spectral similarity coefficient*, KJSSC), y el DSCC-KJSSC híbrido (KJDSSC).

Un valor de SED pequeño indica gran similitud entre los dos espectros. Un inconveniente de la SED es que depende de la intensidad de los espectros más que de su forma. El SAM determina la separación angular entre dos espectros. Una ventaja de SAM es su insensibilidad a los cambios en la intensidad de los espectros, pero al mismo tiempo se afecta por la presencia de ruido de alta frecuencia en la característica espectral, aunque tenga baja amplitud. Un albedo mayor o menor de la característica espectral sólo modifica la longitud del vector de la señal, pero no el ángulo de su dirección. Un ángulo espectral menor indica una mayor similitud entre dos espectros de reflectancia. El SCM puede tomar valores positivos o negativos, siendo los valores positivos mayores indicativos de similitud entre espectros. La SID ve cada espectro de reflectancia como una variable aleatoria y mide la discrepancia del desempeño probabilístico entre los dos espectros. Para carac-

terísticas de reflectancia espectral idénticas, el valor de SID es cero, aumentando en la medida en que son diferentes (Ma *et al.*, 2016).

En un esfuerzo por combinar las fortalezas de diversas SMM se han propuesto varios métodos híbridos. El SAM calcula el ángulo que permite discriminar entre dos características de reflectancia espectral, mientras que la SID modela los datos como una distribución de probabilidad, de forma tal que la variabilidad espectral entre las bandas puede ser extraída como una medida estocástica. La opción híbrida entre los métodos SID y SAM, denominada SIDSAM, ha demostrado ser mejor discriminador que sus componentes por separado. Así, el SIDSAM combina la distancia estocástica SID, que tiene en cuenta el contenido de información de las características espectrales, con la medida determinística SAM, que es insensible a los cambios en la intensidad de los espectros. Su propósito es que dos características espectrales semejantes se consideren más similares, mientras que si son más diferentes se consideren más disímiles (Borsoi *et al.*, 2021).

La BhattD mide la divergencia entre dos distribuciones de probabilidad (normalizadas), por lo que se utiliza para determinar cuán similares o distintas son. Tiene en cuenta la característica completa de reflectancia y no solo algunos de sus parámetros como la media o la varianza, como en algunas otras SMM. Además, es robusta ante datos atípicos (*outliers*). La distancia JMD calcula la separabilidad entre dos características espectrales suponiendo normalidad de los datos. Es asintótica a 2, por lo que un valor de JMD = 2,0 equivaldría a una separabilidad de las dos clases del 100%, considerándose buena separabilidad cuando es superior a 1,9. Toma valor cero cuando las características espectrales son idénticas. Se ha aplicado como criterio de separabilidad para la selección óptima de bandas y la evaluación de los resultados de clasificación. El criterio de separabilidad JMD tiene en cuenta la distancia entre las medias de cada par de clases y la distribución de los valores de sus medias, lo que se consigue implicando las matrices de covarianza de las clases en la medida de separabilidad. La JMD permite evaluar la calidad de las muestras de clases seleccionadas en el espacio de rasgos disponible. En el caso de un espacio de rasgos de alta dimensión, el criterio de separabilidad JMD es un índice estadístico común

que permite seleccionar un subconjunto de rasgos adecuado asociado a un alto grado de separabilidad de clases (Hong *et al.*, 2019).

El DSSC entre dos características espectrales tiene como base el popular coeficiente Dice de similitud, con el cual se calcula el producto interno entre dos vectores para determinar su similitud. Tiene la propiedad de ser poco sensible al ruido y les da poco peso a los datos atípicos. El DSSC varía entre 0 y 1, indicando mayor similitud en la medida en que disminuye su valor y viceversa. El KJSSC entre dos características espectrales tiene como base la medida de la divergencia simétrica no paramétrica de Csizsar, la cual se utiliza en procesamiento de señales para medir la similitud entre distribuciones de probabilidad. El KJSSC puede variar entre 0 hasta infinito, donde un valor que tiende a 0 indica mayor similitud entre las dos características espectrales, mientras que un mayor valor de KJSSC denota mayor diferencia espectral.

Un aspecto importante de los métodos de SMM es que entre ellos no exista correlación, revelando cada uno un punto de vista diferente sobre la relación entre dos características de reflectancia espectral. Dado que los valores de SMM calculados mediante estos métodos no siguen un procedimiento equivalente o proporcional, entregando resultados en unidades de medida diferentes, no se pueden comparar directamente. En este sentido, Kumar *et al.* (2021) utilizan para ello la relative spectral discrimination power (RSDP). La RSDP cuantifica la habilidad de una SMM para discriminar entre dos características espectrales incógnitas con respecto a una característica espectral de referencia. Asumiendo que M (.,.) es una SMM calculada entre un espectro de referencia ρ_r y dos espectros incógnitos ρ_x y ρ_y , la magnitud de la RSDP es proporcional al grado de discriminación entre ρ_x y ρ_y con respecto a ρ_r , siendo mayor cuando se discrimina mejor entre ρ_x y ρ_r que entre ρ_y y ρ_r . Además, la RSDP es simétrica y acorada inferiormente en 1; esto es, $RSDP(\rho_x, \rho_y; \rho_r) \geq 1$, siendo 1 solo cuando $\rho_x = \rho_y$.

Espacio de rasgos

Además del habitual espacio espectral en que se muestran los datos como una característica de reflectancia contra longitud de onda, los datos espectrales pueden representarse también en un espacio

de características formado por vectores que definen puntos en un espacio de N dimensiones, creando grupos de cada una de las clases (Ma *et al.*, 2016). Cuando el espacio de rasgos es bidimensional ($N = 2$) los datos de la variabilidad de la reflectancia, tomados a dos longitudes de onda diferentes (λ_1 y λ_2) forman un plano sobre el cual se distribuyen los grupos, que se suponen Gaussianos, de cada especie a esas longitudes de onda. Un espacio de rasgos bidimensional solo permite estadísticos de primer orden, con el inconveniente adicional de que reduce todo el espacio de datos a solo dos dimensiones y con la consiguiente pérdida de información. Incrementar el número de rasgos a tres, permite representar los datos en un espacio tridimensional siendo aplicable estadísticos de segundo orden, con lo cual se incrementa la posibilidad de separación entre clases al tener en cuenta mayor información.

De esta forma, los grupos de especies diferentes deben tener centroides significativamente separados como para poder diferenciar las clases de las especies a partir del espacio de rasgos N -dimensional. La posición media de cada grupo en el espacio de rasgos está dada por el vector de su valor medio (centroide), mientras que la forma de la distribución la determina la matriz de covarianza. La correlación entre dimensiones puede hallarse interpretando la matriz de covarianza; así, si hay gran correlación entre los ejes de un espacio de rasgos, los elementos no diagonales de la matriz de covarianza se aproximan a 0. Por el contrario, si existe poca correlación, los elementos no diagonales serán grandes en comparación con los elementos diagonales. Cuando se combina los estadísticos media y covarianza se obtienen mejores resultados de clasificación (Ronay *et al.*, 2024).

La selección de rasgos puede realizarse con el fin de caracterizar una población de forma inequívoca, o su clasificación. Representar en un espacio de rasgos los datos de la familia de características de reflectancia espectral producidas por la variabilidad de una misma especie involucra estimar los rasgos que logran mayor separabilidad entre clases (especies). Por ejemplo, en el caso bidimensional, serían los valores de λ_1 y λ_2 que logran la mayor distancia entre los centroides de las clases. Para ello pueden utilizarse técnicas de Monte Carlo u otros tipos de análisis.

Uno de esos análisis es el discriminante, también llamado reconocimiento supervisado de patrones. Después de encontrar una función discriminante (lineal o cuadrática) esta se aplica para asignar nuevos datos desconocidos a una clase. También, la regresión mediante mínimos cuadrados parciales es otro método de análisis multivariante. Una de las utilidades del espacio de rasgos es que puede ser transformado a otro espacio de rasgos, como los índices derivados, los índices de dos bandas normalizadas y la PCT (Ronay *et al.*, 2024).

Algunos ejemplos de espacios de rasgos pueden verse en la figura 3, obtenidos a partir de los valores de los picos de gradiente elevados al cuadrado de cada una de las características espectrales de una muestra de *Saccharum* spp. y de *A. procera* (Figura 3A), así como de la relación entre los valores de algunas SMMs calculadas para todas características espectrales (Figura 3B).

Análisis de componentes principales

El PCA es una técnica no paramétrica multivariante que produce una transformación ortogonal que convierte un conjunto de observaciones de un conjunto de variables potencialmente correlacionadas en un conjunto de variables linealmente no correlacionadas denominadas componentes principales (PC). En una descomposición en PC la suma de la varianza de las relaciones de bandas originales es igual a la suma de la varianza de los PC (sus valores propios),

por lo que no hay pérdida ni adición de información o ruido durante el proceso de transformación, sino sólo una reorganización del conjunto de datos. Todos los PC son ortogonales entre sí. No obstante, los PC se ordenan de forma que el primero represente la mayor variabilidad de los datos y cada componente posterior contenga el resto de la varianza en orden descendente.

De esta forma, el PCA extrae características que ayudan a encontrar la mayor contribución de puntos, los cuales pueden identificarse a través de los mayores valores propios. A grandes rasgos, un conjunto de características espectrales necesitaría p componentes principales para reproducir la variabilidad total del sistema. No obstante, gran parte de esa variabilidad puede ser cuantificada por un número menor de los PC. Así, la utilidad práctica del PCA consiste en que las contribuciones más bajas pueden desecharse y sólo los puntos con la contribución más alta se seleccionan para su posterior procesamiento/análisis, con la ventaja de lograr una significativa reducción de los datos originales. Esta reducción, aunque tiene el costo de la pérdida de parte de la información inicial, puede ser regulada reduciendo la pérdida a límites aceptables. Esto es, si se elige bien el número de componentes, la compensación entre la pérdida de información y la simplificación del problema es casi siempre una ventaja (Butler *et al.*, 2021; Che'Ya *et al.*, 2022).

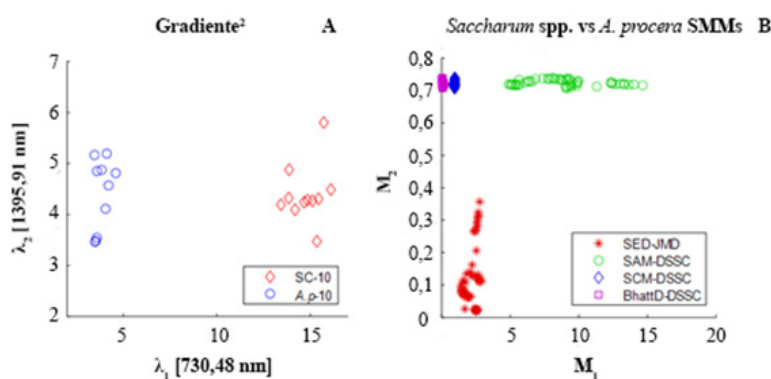


Figura 3. Espacios de rasgos de 10 lecturas de muestras de *Saccharum* spp. y de la arvense *Albizia procera*. A) Espacios de rasgos construidos a partir de los gradientes elevados al cuadrado a 730,48 nm y 1 395,91 nm de *Saccharum* spp. (SC) y de la arvense *Albizia procera* (A.p). B) Espacios de rango de los gradientes a partir de varias SMMs calculadas utilizando 50 lecturas de cinco muestras de ambas especies.

Análisis de varianza

Cuando se comparan simultáneamente tres o más características de reflectancia espectral, ANOVA es la prueba estadística paramétrica que permite determinar si la varianza de una de las características se produce dentro de los distintos grupos de población o entre ellos, en lugar de analizar las medias de las características, asumiendo para ello la normalidad de los datos y la independencia entre las lecturas. No obstante, el one-way ANOVA con valor $p < 0,05$ puede resultar de gran utilidad en el análisis de características espectrales utilizando los valores medios de las características espectrales. El propósito del one-way ANOVA es determinar si los datos de varios grupos de una variable independiente tienen una media común. Teniendo en cuenta que el tamaño de todas las muestras está balanceado con 10 lecturas de características espectrales (tamaño mínimo recomendado), así como que se probó la normalidad de los datos.

El análisis de varianza compara las medias de varios grupos para comprobar la hipótesis de que todos son iguales, frente a la alternativa general de que no todos son iguales. A veces, esta alternativa puede ser demasiado general, por lo que es necesario recabar información sobre qué pares de medias son significativamente diferentes y cuáles no. Una prueba de comparación múltiple puede proporcionar esta información. Por ejemplo, Estrada *et al.* (2023) prefieren la Prueba de comparación múltiple de Tukey (*Tukey's multiple comparison test*) ($p\text{-values} < 0,05$) en lugar del de Bonferroni, pero el *Tukey's test* es más apropiado para análisis de covarianza (ANCOVA) en problemas en que los grupos (lecturas espectrales) estén correlacionados.

Che'Ya *et al.* (2013) analizaron los datos mediante un análisis de varianza unidireccional, mostrando el patrón de discriminación único para cada tipo de arvense y el sorgo, con lo que se verificó la presencia de diferencias estadísticas entre especies en cada banda espectral. La hipótesis se comprobó para longitudes de onda de 450-1 000 nm, con límites de confianza del 95% ($\alpha = 0,05$). La prueba ANOVA demostró las diferencias individuales de longitud de onda entre siete especies (seis arvenses y el sorgo), permitiendo visualizar qué longitudes de onda tienen el potencial de diferenciar entre las siete especies.

Análisis discriminante

El DA es semejante a una clasificación, pero donde se conoce la clase a la que pertenece la entrada y se asume que cada clase genera datos basados en diferentes distribuciones Gaussianas. Los clasificadores pueden basarse en funciones discriminantes y, cuando utilizan una medida de distancia para seleccionar la clase cuya función discriminante produce la menor distancia entre la especie incógnita y la media de cada clase candidata, son los denominados de distancia mínima a la media. De esta forma, el análisis discriminante se utiliza para probar el clasificador y el poder discriminatorio del espacio de rasgos. Che'Ya *et al.* (2013) utilizaron un análisis unidireccional de la varianza y un LDA para identificar las longitudes de onda significativas para discriminar entre arvenses y cultivos. Los resultados indican que sería factible discriminar las arvenses de las plantas de cultivo dentro de un rango de longitudes de onda de 450-999 nm.

Relación entre bandas espectrales

La SBR es una de las técnicas de análisis espectral más sencillas y, sin embargo, más útiles ya que permite realzar diferencias espectrales, siendo invariante a las variaciones de iluminación. Entre sus aplicaciones están los índices de vegetación, tales como el $NR = NIR/Red$, el $BN = Blue/NIR$, o el $GN = Green/NIR$. Por ejemplo, Almeida y Filho (2004) propusieron índices espectrales capaces de distinguir constituyentes específicos de la vegetación, expresando así, las longitudes de onda en nm, para clorofila-a (461/422 y 807/638), clorofila-b (520/470 y 807/648), α -caroteno (520/442), carotenoides (539/490 y 807/490), antocianina (510/530), fitocromo P730 (845/730), fitocromo P660 (778/658), lignina (1 028/2 101), celulosa (2 211/2 400), nitrógeno (1 731/1 691), y contenido de agua de la hoja (1 066/1 452). Para ello, las relaciones se establecieron tomando como numerador las bandas en las que las absorciones espectrales eran insignificantes o inexistentes y como denominador las posiciones de las bandas en los picos de absorción de cada constituyente vegetal. Estos autores, además de proponer relaciones de bandas individuales, proponen también enfoque mucho más sólido consistente en utilizar un conjunto de varias relaciones de banda dirigidas a una serie de componentes de las plantas

que contribuyen simultáneamente a los espectros de reflectancia de la vegetación.

CONCLUSIONES

El manejo inteligente de arvenses en los cultivos, en base a las propiedades de reflectancia espectral, constituye una temática de gran interés por lo que se han creado bibliotecas de características espectrales como base de muchos métodos para el manejo de arvenses específico del sitio. No obstante, uno de los grandes inconvenientes en la identificación de especies de plantas a partir de sus propiedades espectrales es la variabilidad intra-muestra e inter-muestra de las características de reflectancia que se observa a causa de múltiples factores, por lo que trabajos recientes toman más en cuenta el bundle de características medias de diversas muestras que el tradicional concepto de firma espectral como una característica única.

Por esta razón, la caracterización de una especie requiere un profundo estudio para identificar los rasgos espectrales distintivos que permitan discriminar entre una especie y otra teniendo en cuenta tal variabilidad. La extensa literatura publicada sobre métodos de análisis de características espectrales de especies de plantas recoge el uso de muy diversos métodos para el procesamiento de los datos registrados, lo cual es consecuencia de lo complejo que resulta determinar una técnica para ello con eficacia. Consecuentemente, cada investigador debe realizar un análisis exhaustivo de la utilidad de cada método para el propósito de su investigación de manera que pueda solventar las limitaciones actuales.

AGRADECIMIENTOS

Este artículo ha sido posible gracias al proyecto de investigación titulado “Soluciones para la agricultura y el medio ambiente basados en vehículos no tripulados, articulados a una infraestructura de datos espaciales”, código PN223LH004, del Programa Nacional “Automática, Robótica e Inteligencia Artificial”, ARIA.

CONTRIBUCIÓN DE LOS AUTORES

Conceptualización: Katia Ojito-Ramos, Osmany de la C. Aday Díaz y Rubén Orozco Morales

Análisis formal: Rubén Orozco Morales

Investigación: Katia Ojito-Ramos, Osmany de la C. Aday Díaz, Rubén Orozco Morales

Metodología: Katia Ojito-Ramos, Osmany de la C. Aday Díaz, Rubén Orozco Morales

Redacción – borrador inicial: Rubén Orozco Morales, Katia Ojito-Ramos

Redacción - revisión y edición: Katia Ojito-Ramos, Osmany de la C. Aday Díaz, Rubén Orozco Morales

Adquisición de financiación, Administración de proyecto y Supervisión: Luís Hernández Santana

BIBLIOGRAFÍA

- ADAY DÍAZ, O. C., OJITO-RAMOS, K., OROZCO-MORALES, R. and HERNÁNDEZ SANTANA, L. 2023. Vis-NIR Spectral differentiation of *Albizia procera* (Roxb.) Benth. and *Saccharum* spp. *Agrica*, 12 (2): 244-252.
- ALMEIDA, T. I. R. and FILHO, D. S. 2004. Principal component analysis applied to feature-oriented band ratios of hyperspectral data: A tool for vegetation studies. *International Journal of Remote Sensing*, 25 (22): 5005-5023.
- BORSOI, R. A., IMBIRIBA, T., BERMUDEZ, J. C. M., RICHARD, C., CHANUSSOT, J., DRUMETZ, L., TOURNERET, J.-Y., ZARE, A. and JUTTEN, C. 2021. Spectral variability in hyperspectral data unmixing: A comprehensive review. *IEEE Geoscience and Remote Sensing Magazine*, 9 (4): 223-270.
- CHANDRA PANDEY, P., SRIVASTAVA, P. K., BALZTER, H., BHATTACHARYA, B. and PETROPOULOS, G. P. (Eds.) 2020. *Hyperspectral Remote Sensing. Theory and Application*. Primera Edición. Elsevier, 506 p.
- CHE'YA, N. N., GUPTA, M., GEORGE, D., LISLE, A., BASNET, B. and CAMPBELL, G. 2013. Spectral discrimination of weeds using hyperspectral radiometry. Proceedings of the 5th Asian Conference on Precision Agriculture (ACPA), June 25-28, Jeju, Korea.
- CHE'YA, N. N., MOHIDEM, N. A., ROSLIN, N. A., SABERIOON, M., TARMIDI, M. Z., ARIF SHAH, J., FAZLIL ILAHI, W. F. and MAN, N. 2022. Mobile computing for pest and disease management using spectral signature analysis: A Review. *Agronomy*, 12 (4): 967.
- DE SOUZA, M. F., RIOS DO AMARAL, L., ROBSON DE MEDEIROS OLIVEIRA, S., NERIS COUTINHO, M. A. and FERREIRA NETTO, C. 2020. Spectral differentiation of sugarcane

- from weeds, *Biosystems Engineering*, 190: 41-46.
- ESTRADA, F., FLEXAS, J., ARAUS, J. L., MORA-POBLETE, F., GONZALEZ-TALICE, J. and LOBOS, G. A. 2023. Exploring plant responses to abiotic stress by contrasting spectral signature changes. *Frontiers in Plant Sciences*, 13: 1026323.
- FAOSTAT. 2023. Organización de las Naciones Unidas para la Alimentación y Agricultura Disponible en: <http://www.faostat.fao.org>. Consultado: 21/04/2025.
- GALON, L., NIKPAY, A., MA, Y. L., FERREIRA, E. A., MUNSIF, F., ZIAEE, M., SHARAFIZADEH, P. and CONCENCO, G. 2023. Weeds Management in Sugarcane: Recent Developments and Future Perspectives. En: VERMA, K. K., SONG, XIU-PENG., RAJPUT, V. D., SOLOMON, S., LI, YANG-RUI. and GOVIND, P. R. (Eds.). *Agro-industrial Perspectives on Sugarcane Production under Environmental Stress*, Primera edición, Springer, Singapore, pp. 361-387.
- HONG, D., YOKOYA, N., CHANUSSOT, J. and ZHU, X. X. 2019. An augmented linear mixing model to address spectral variability for hyperspectral unmixing, *IEEE Transactions on Image Processing*, 28 (4): 1923-1938.
- JAFARBIGLU, H. and POURREZA, A. 2023. Impact of sun-view geometry on canopy spectral reflectance variability. *ISPRS Journal of Photogrammetry and Remote Sensing*, 196: 270-286.
- JOHNSON, R. M., ORGERON, A. J., SPAUNHORST, D. J., HUANG, I. S. and ZIMBA, P. V. 2023. Discrimination of weeds from sugarcane in Louisiana using hyperspectral leaf reflectance data and pigment analysis. *Weed Technology*, 37 (2): 123-131.
- KATSOULAS, N., ELVANIDI, A., FERENTINOS, K. P., KACIRA, M., BARTZANAS, T. and KIT-TAS, C. 2016. Crop reflectance monitoring as a tool for water stress detection in greenhouses: A review. *Biosystems Engineering*, 151: 374-398.
- KUMAR, C., CHATTERJEE, S., OOMMEN, T. and GUHA, A. 2021. New effective spectral matching measures for hyperspectral data analysis. *International Journal of Remote Sensing*, 42 (11): 4126-4156.
- LAROCHE-PINEL, E., ALBUGHDADI, M., DUTHOIT, S., CHÉRET, V., ROUSSEAU, J. and CLENET, H. 2021. Understanding vine hyperspectral signature through different irrigation plans: A first step to monitor vineyard water status. *Remote Sensing*, 13 (3): 536.
- LI, S., PENG, B., FANG, L. and LI, Q. 2022. Hyperspectral band selection via optimal combination strategy. *Remote Sensing*, 14 (12): 2858.
- LONG, Y., RIVARD, B., SANCHEZ-AZOFEIFA, A., GREINER, R., HARRISON, D. and JIA, S. 2021. Identification of spectral features in the longwave infrared (LWIR) spectra of leaves for the discrimination of tropical dry forest tree species. *International Journal of Applied Earth Observations and Geoinformation*, 97: 102286.
- MA, D., LIU, J., HUANG, J., LI, H., LIU, P., CHEN, H. and QIAN, J. 2016. Spectral similarity assessment based on a spectrum reflectance-absorption index and simplified curve patterns for hyperspectral remote sensing. *Sensors*, 16 (2): 152.
- NORASMA, C. Y. N., ALAHYADI, L. A. N., FAZILAH, F. F. W., ROSLAN, S. N. A. and TARMIDI, Z. 2020. Identification spectral signature of weed species in rice using Spectroradiometer Handheld Sensor. *IOP Conference Series: Earth and Environmental Science*, 540 (1): 012091.
- PADMA, S. and SANJEEVI, S. 2014. Jeffries Matusita based mixed-measure for improved spectral matching in hyperspectral image analysis. *International Journal of Applied Earth Observation and Geoinformation*, 32: 138-151.
- PATHAK, P., CHALOPIN, C., ZICK, L., KÖHLER, H., PFAHL, A., RAYES, N., GOCKEL, I., NEUMUTH, T., MELZER, A., JANSEN-WINKELN, B. and MAK TABI, M. 2023. Spectral similarity measures for *in vivo* human tissue discrimination based on hyperspectral imaging. *Diagnostics*, 13 (2): 195.
- PIRON, A., LEEMANS, V., KLEYNEN, O., LEBEAU, F. and DESTAIN, M.-F. 2008. Selection of the most efficient wavelength bands for discriminating weeds from crop. *Computers and Electronics in Agriculture*, 62: 141-148.
- RONAY, I., KIZEL, F. and LATI, R. 2022. The effect of spectral mixtures on weed species classification. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 3: 477-484.

- RONAY, I., LATI, R. N. and KIZEL, F. 2024. Weed species identification: Acquisition, feature analysis, and evaluation of a hyperspectral and RGB dataset with labeled data. *Remote Sensing*, 16 (15): 2808.
- ROSLIN, N. A., CHE'YA, N. N., SULAIMAN, N., NOR ALAHYADI, L. A. and ISMAIL, M. R. 2021. Mobile application development for spectral signature of weed species in rice farming. *Pertanika Journal of Science & Technology*, 29 (4): 2241-2259.
- SOCA-MUÑOZ, J. L., RODRÍGUEZ-MACHADO, E., ADAY-DÍAZ, O., HERNÁNDEZ-SANTANA, L. and OROZCO-MORALES, R. 2020. Spectral signature of brown rust and orange rust in sugarcane. *Revista Facultad de Ingeniería, Universidad de Antioquia*, REDIN, 96: 9-20.
- SOM-ARD, J., ATZBERGER, C., IZQUIERDO-VERDIGUIER, E., VUOLO, F. and IMMITZER, M. 2021. Remote sensing applications in sugarcane cultivation: A review. *Remote Sensing*, 13 (20): 4040.
- SONOBE, R. and HIRONO, Y. 2022. Applying variable selection methods and preprocessing techniques to hyperspectral reflectance data to estimate tea cultivar chlorophyll content. *Remote Sensing*, 15 (1): 19.
- SU, W. H. 2020. Advanced machine learning in point spectroscopy, RGB-and hyperspectral-imaging for automatic discriminations of crops and weeds: A review. *Smart Cities*, 3 (3): 767-792.

CITAR COMO:

OROZCO MORALES, R., OJITO RAMOS, K., ADAY DÍAZ, O.C., HERNÁNDEZ SANTANA, L. 2025. Caracterización espectral de arvenses que afectan los cultivos de *Saccharum* spp., métodos de obtención y procesamiento de datos: Una revisión, 52 (2025) e2458



Artículo de **libre Acceso** bajo los términos de una *Licencia Creative Commons Atribución-NoComercial-SinDerivar 4.0 Internacional*. Se permite, sin restricciones el uso, distribución, traducción y reproducción del documento, siempre que la obra sea debidamente citada.