

UNIVERSIDAD NACIONAL DE INGENIERÍA  
FACULTAD DE CIENCIAS



TESIS

**“ANÁLISIS E IDENTIFICACIÓN DE SUELOS A  
PARTIR DE IMÁGENES HIPERESPECTRALES  
USANDO PPI, MNF Y PCA”**

PARA OBTENER EL GRADO ACADÉMICO DE MAESTRO EN  
CIENCIAS EN FÍSICA

ELABORADO POR:

**DANTE CASTILLO CCORAHUA**

**ASESOR:**

Dr. HÉCTOR RAÚL LORO RAMÍREZ

**LIMA – PERÚ**

**2021**

A Dios y a la Virgen María por sus bendiciones. A mis padres; Alejandro y Luisa, también a mis queridos hermanos, sobrinos y familiares.

## AGRADECIMIENTOS

A mis familiares, por ser parte fundamental de mi formación en principios y valores, por confiar siempre en mí y enseñarme que los sueños son para hacerlos reales con la ayuda de Dios.

A mis compañeros de la maestría, con quienes comparto horas de estudio e investigación, y en especial, al grupo del laboratorio de imágenes hiperespectrales; Ricardo, Walter, Ronald y Leonardo.

Al programa de CONCYTEC “Consejo Nacional de Ciencia, Tecnología e Innovación Tecnológica” por impulsar la investigación en jóvenes profesionales y el desarrollo tecnológico en el Perú. También al Dr. Héctor Raúl Loro Ramírez, por asesorarme y compartirme su experiencia y sabiduría. Al personal de laboratorio y taller, Rubén, el maestro Quinde y Efraín.

## Índice

Nomenclatura .....	viii
Índice de figuras .....	ix
Índice de tablas .....	xii
Resumen .....	xiii
Abstract .....	xiv

## INTRODUCCIÓN .....1

### CAPITULO 1

#### 1. FUNDAMENTO TEÓRICO .....6

##### 1.1. Fundamentos físicos .....6

###### 1.1.1. Energía electromagnética .....6

##### 1.2. Espectro electromagnético .....7

##### 1.3. Radiación .....9

###### 1.3.1. Leyes de radiación .....9

##### 1.4. Interacción con la superficie .....11

###### 1.4.1. Reflexión especular .....11

###### 1.4.2. Reflexión difusa.....12

##### 1.5. Proceso de interacción entre energía y materia .....12

###### 1.5.1. Reflectancia .....13

###### 1.5.2. Transmitancia .....13

###### 1.5.3. Absorbancia.....13

### CAPITULO 2

#### 2. VISIÓN HIPERESPECTRAL.....14

##### 2.1. Visión artificial .....14

###### 2.1.1. ¿Qué es una imagen hiperespectral? .....14

##### 2.2. Espectroscopía de imágenes hiperespectrales .....15

##### 2.3. Resolución de una imagen hiperespectral .....16

###### 2.3.1. Resolución espacial .....16

###### 2.3.2. Resolución espectral .....17

###### 2.3.3. Resolución radiométrica .....18

###### 2.3.4. Resolución Temporal.....19

##### 2.4. Tipo de sensores hiperespectrales .....19

2.4.1. Modo de detección espacial .....	20
<b>2.5. Curvas espectrales.....</b>	<b>21</b>
2.5.1. Reflectancia espectral de la vegetación, suelo y agua .....	22
2.5.2. Librerías espectrales .....	23
<b>2.6. Ventajas y desventajas de la información hiperespectral .....</b>	<b>24</b>
<b>2.7. Campo de aplicación de las imágenes hiperespectrales .....</b>	<b>25</b>
2.7.1. Biomedicina .....	25
2.7.2. Agricultura de precisión .....	26
2.7.3. Análisis de minerales .....	27
2.7.4. Monitoreo ambiental .....	28
2.7.5. Visión de máquina para productos alimenticios e industriales .....	28

### **CAPITULO 3**

<b>3. SUELOS DE CUSCO Y MADRE DE DIOS .....</b>	<b>29</b>
<b>3.1. Aspectos generales de suelos .....</b>	<b>29</b>
3.1.1. ¿Qué es el suelo? .....	29
3.1.2. Principales características físicas de suelo .....	29
<b>3.2. Clasificación de suelos .....</b>	<b>35</b>
3.2.1. La clasificación según SOIL TAXONOMY, USDA 1999 .....	35
3.2.2. La clasificación según la FAO, 1971 .....	35
3.2.3. Relaciones aproximadas entre los mayores órdenes y grupos de suelos .....	38
<b>3.3. Comportamiento espectral de los suelos y sus constituyentes .....</b>	<b>38</b>
<b>3.4. Suelos de Cusco .....</b>	<b>40</b>
<b>3.5. Suelos de Madre de Dios .....</b>	<b>41</b>

### **CAPITULO 4**

<b>4. ANÁLISIS Y PROCESAMIENTO DE UNA IMAGEN HIPERESPECTRAL .....</b>	<b>43</b>
<b>4.1. Datos de una imagen hiperespectral .....</b>	<b>43</b>
4.1.1. Diseño y cadena de estudio hiperespectral.....	45
<b>4.2. Pre – procesamiento de una HSI.....</b>	<b>46</b>

4.2.1. Normalización .....	46
4.2.2. Suavizamiento espectral con Savitzky – Golay (SG) .....	48
4.2.3. Corrección de escala multiplicativa con la Standard Normal Variate SNV ...	50
4.2.4. Reducción de dimensionalidad de la imagen hiperespectral .....	51
4.2.4.1. Análisis de Componentes Principales - PCA .....	51
4.2.4.2. Transformación Mínima Fracción de Ruido - MNF.....	53
<b>4.3. Procesamiento hiperespectral .....</b>	<b>55</b>
4.3.1. Clasificación y detección de objetos.....	55
4.3.2. Desmezclado espectral.....	57
<b>4.4. Técnica de extracción de endmembers .....</b>	<b>58</b>
4.4.1. Algoritmo Pixel Purity Index - PPI .....	58
<b>4.5. Otra técnica de clasificación .....</b>	<b>60</b>
4.5.1. K - means .....	60
 <b>CAPITULO 5</b>	
<b>5. RESULTADOS Y DISCUSIONES .....</b>	<b>62</b>
<b>5.1. Proceso de preparación de muestras de suelo .....</b>	<b>62</b>
5.1.1. Secado de muestra .....	63
5.1.2. Empastillado de muestra .....	64
5.1.3. Procedimiento y resultados .....	65
<b>5.2. Protocolo de adquisición de imágenes .....</b>	<b>67</b>
5.2.1. Sistema de cámara hiperespectral.....	67
5.2.2. Hardware .....	68
5.2.3. Software .....	69
5.2.4. Calibración del sistema .....	70
5.2.5. Calibración de la imagen.....	72
5.2.6. Adquisición de imágenes HSI .....	74
<b>5.3. Pre procesamiento HSI .....</b>	<b>77</b>
<b>5.4. Estandarización de los Espectro de reflectancia de los suelos .....</b>	<b>81</b>
5.4.1. Librería espectral .....	84
<b>5.5. Procesamiento HSI.....</b>	<b>85</b>

5.5.1. Análisis de los suelos usando el algoritmo PPI.....	85
5.5.2. Análisis de componentes principales (PCA) en suelos.....	93
5.5.3. Comparación de resultados de cada modelo .....	102
5.5.4. Correlación .....	103

<b>5.6. Discusión de resultados .....</b>	<b>104</b>
---	------------

## **CAPITULO 6**

<b>CONCLUSIONES.....</b>	<b>105</b>
--------------------------	------------

<b>BIBLIOGRAFIA .....</b>	<b>106</b>
---------------------------	------------

## ABREVIATURAS

- ❖ BIL: Banda intercalada por línea (Band interleaved by line)
- ❖ BIP: Banda intercalada por pixel (Band interleaved by pixel)
- ❖ BSQ: Banda secuenciada (Band Sequential)
- ❖ CCA: Análisis de correlación canónica (Canonical correlation analysis)
- ❖ CCD: Dispositivos de carga acoplada
- ❖ CE: Conjunto de espectros
- ❖ CEna: Conjunto de espectros no anómalos
- ❖ EEA: Espectroscopía de absorción atómica
- ❖ EVI: Índice vegetativo mejorado (Enhanced vegetative index)
- ❖ ENVI: Environment for visualizing images
- ❖ FAO: Organización para la agricultura y los alimentos (Food and agricultural organization)
- ❖ FOV: Campo de visión (Field of view)
- ❖ FPS: Fotogramas por segundo (Frames per second)
- ❖ HDR: Reducción de alta dimensionalidad (High dimensionality reduction)
- ❖ HSI: Imágenes hiperespectrales (Hyperspectral image)
- ❖ IFOV: Campo de visión instantáneo (instantaneous field of view)
- ❖ IR: Rango de onda infrarrojo (1700 – 2700 nm)
- ❖ MNF: Mínima fracción de ruido (Minimum noise fraction)
- ❖ NDVI: Diferencia vegetativa normalizada (Normalized difference vegetative index)
- ❖ NIR: Infrarrojo cercano (Near infrared)
- ❖ PCA: Análisis de componentes principales (Principal component analysis)
- ❖ PPI: Índice de pixel puro (Pixel purity index)
- ❖ RGB: Formato de imagen tricolor: Rojo, Verde y Azul (Red, green and blue format)
- ❖ ROI: Región de interés (Region of interest)
- ❖ SWIR: Infrarrojo de onda corta
- ❖ USDA: Departamento de agricultura de los EE.UU (United States Geological Survey Spectroscopy)

## Índice de Figuras

<b>Figura 1. 1.</b>	Se muestran las características de longitud de onda y amplitud de las ondas .....	6
<b>Figura 1. 2.</b>	Espectro de ondas electromagnéticas en términos de su longitud de onda [38] .....	8
<b>Figura 1. 3.</b>	Las componentes del campo eléctrico (E) y el campo magnético (H) de la radiación electromagnética, están orientados perpendicularmente entre si y en ángulo recto con respecto a la dirección de propagación [36] .....	9
<b>Figura 1. 4.</b>	a) Ley de la reflexión, b) Esquema que muestra la reflexión especular y c) muestra la reflexión difusa [40] .....	12
<b>Figura 2.1.</b>	Estructura de una imagen hiperespectral [46] .....	15
<b>Figura 2.2.</b>	Diagrama esquemático del espectrógrafo empleado, arquitectura de un sensor de barrido de línea (pushbroom) [49] .....	16
<b>Figura 2.3.</b>	Campo de visión instantáneo (IFOV) .....	17
<b>Figura 2.4.</b>	Comparación de resolución espectral de una imagen multiespectral e hiperespectral [50] .....	18
<b>Figura 2.5.</b>	Tipos de orientación de los sensores .....	20
<b>Figura 2.6.</b>	Curvas de reflectancia espectral para la vegetación, suelo y agua [53] .....	22
<b>Figura 2.7.</b>	Predicción del diagnóstico de tejido cerebral. Imagen izquierda: diagnóstico según la vista del neurocirujano de tejido canceroso y sano. Imagen derecha: detección correcta de un tumor de grado IV (glioblastoma)[58] .....	26
<b>Figura 2.8.</b>	Predicción espacial para el rendimiento de grano basado en el modelo MLR [60] .....	27
<b>Figura 2.9.</b>	A) Fotografía de corte pulido del espécimen CTP-1. B) Fotografía de la imagen hiperespectral etiquetada del espécimen CTP-1 [31] .....	27
<b>Figura 2.10.</b>	Imagen hiperespectral de agua de río contaminada [61] .....	28
<b>Figura 3.1.</b>	Partículas de suelo [67] .....	30
<b>Figura 3.2.</b>	La figura representa al triangulo textural del suelo de acuerdo a la FAO y al Departamento de Agricultura de USA .....	31
<b>Figura 3.3.</b>	Formas de estructuras de suelos comunes (imagen derecha). Disposición de partículas de cuarzo, arcillas y de coloides orgánicos en un agregado de suelo (imagen izquierda). [68] .....	32
<b>Figura 3.4.</b>	Escala de pH de suelos [69] .....	33
<b>Figura 3.5.</b>	Se muestra las variaciones de la reflectancia debido a diversos factores en el suelo [74] a) Relación entre el contenido de humedad del suelo y la reflectancia espectral. b) Relación entre el contenido de materia orgánica de los suelos minerales y su reflectancia espectral .....	39
<b>Figura 3.6.</b>	Variaciones de la reflectancia debido al contenido orgánico y de óxido de hierro [73]. c) Relación entre el estado de descomposición de un suelo orgánico y su reflectancia espectral. d) Relación entre el contenido de óxido de hierro y la reflectancia espectral del suelo .....	40
<b>Figura 3.7.</b>	Parcelas de valle interandino en Cusco [75] .....	41
<b>Figura 3.8.</b>	Suelos arcillosos rojizos en Madre de Dios [75] .....	42
<b>Figura 4.1.</b>	Reconocimiento y diferenciación de los pixeles puro y mezclas y su posterior desmezclado para encontrar los miembros finales (endmembers) .....	44
<b>Figura 4.2.</b>	Curvas espectrales de muestras de cuatro especies diferentes de hierba [77] .....	48
<b>Figura 4.3.</b>	Corrección de espectros utilizando el filtro de suavizamiento de SG .....	49

<b>Figura 4.4.</b>	(a) Espectros originales de un fruto seco (b) Espectros tratados con el SNV...	50
<b>Figura 4.5.</b>	Gráfico del PCA [87]. a) Datos iniciales, representados como puntos tridimensionales, b) Componentes principales (PC <sub>i</sub> ) de los datos, ordenados de manera descendentes según varianza, c) Proyección de los datos originales sobre PC1 y PC2, descartando PC3 .....	52
<b>Figura 4.6.</b>	Técnicas de aprendizaje automático, clasificación supervisada y no supervisada .....	56
<b>Figura 4.7.</b>	Flujo de procesamiento de reloj de arena para la detección de los miembros finales .....	58
<b>Figura 4.8.</b>	Ejemplo básico que ilustra el funcionamiento del algoritmo de extracción de endmembers PPI en un espacio de 2 dimensiones [52] .....	59
<b>Figura 5.1.</b>	Muestras Cus1, Cus2, Cus3, Cus4 de suelos de Cusco y muestras MD1, MD2, MD3 de Madre de Dios .....	63
<b>Figura 5.2.</b>	Puesta de material en el horno por 24 horas, a $110^{\circ} \pm 5^{\circ}C$ .....	64
<b>Figura 5.3.</b>	Materiales usados para el empastillamiento de muestras .....	64
<b>Figura 5.4.</b>	Muestras secadas durante 24 horas a una temperatura de $110^{\circ} \pm 5^{\circ}C$ .....	65
<b>Figura 5.5.</b>	Muestras empastilladas sometidas a diferentes presiones .....	66
<b>Figura 5.6.</b>	Sistema de adquisición de imágenes hiperespectrales .....	68
<b>Figura 5.7.</b>	Primera visualización del sistema (imagen borrosa), donde se hace el reconocimiento de la luz y objetos .....	71
<b>Figura 5.8.</b>	Resultado del correcto ajuste en el enfoque de luz entrante hacia los sensores .....	71
<b>Figura 5.9.</b>	Resultado del ajuste de controles de la cámara. Las muestras de suelos fueron tomadas con un Framerate de 47.5 Hz e Integration Time de 21.35 ms .....	72
<b>Figura 5.10.</b>	Cubierta de lente para la eliminación de corriente oscura .....	72
<b>Figura 5.11.</b>	Referencia de reflectancia usando un teflón como patrón de calibración ....	73
<b>Figura 5.12.</b>	Visualización de 6 bandas (9-14) de las 164 que hay en la imagen hiperespectral de los suelos en estado natural .....	74
<b>Figura 5.13.</b>	Visualización de 6 bandas (9-14) de las 168 que hay en la imagen hiperespectral de los suelos molidos .....	75
<b>Figura 5.14.</b>	Visualización de 6 bandas (10,11,12,20,80,150) de las 168 que hay en la imagen hiperespectral de los suelos molidos .....	76
<b>Figura 5.15.</b>	(a-c) Espectros con influencia de ruido en las bandas iniciales y finales. (b-d) Eliminación de bandas extremas, considerando como rango de espectro (980.196 - 1642.114 nm) y longitud de espectro 4.762 nm .....	77
<b>Figura 5.16.</b>	La columna izquierda representa el preprocesamiento de las firmas espectrales con la normalización de máximos y mínimos, la columna derecha es el suavizamiento de firmas espectrales con Savitzky – Golay .....	78
<b>Figura 5.17.</b>	a) Espectros de los 100 primeros píxeles del suelo natural Cus3, al cual se aplicó la normalización máx – mín, SG y la corrección de dispersión SNV. b) Imagen umbralizada del suelo natural de Cus3 .....	79
<b>Figura 5.18.</b>	c - e) Espectros de los 100 primeros píxeles del suelo molido y empastillado Cus3, respectivamente, al cual se aplicó la normalización máx – mín, SG y la corrección de dispersión SNV. d - f) Imagen umbralizada del suelo molido y empastillado de Cus3 respectivamente .....	80
<b>Figura 5.19.</b>	Espectros de reflectancia de las muestras Cus1, Cus2, Cus3 y Cus4 en diferentes condiciones .....	81
<b>Figura 5.20.</b>	Espectros de reflectancia de las muestras MD1, MD2 y MD3 en diferentes condiciones .....	82

<b>Figura 5.21.</b> Firmas espectrales de distintos materiales encontrados en cada muestra de los suelos estudiados .....	83
<b>Figura 5.22.</b> Espectros de reflectancia estandarizados de las muestras estudiadas de Cusco. El gráfico insertado muestra la intensidad relativa de las absorciones de las muestras para $\lambda = 1415 \text{ nm}$ .....	84
<b>Figura 5.23.</b> Espectros de reflectancia estandarizados de las muestras estudiadas de Madre de Dios. El gráfico muestra la intensidad relativa de las absorciones de las muestras para $\lambda = 1415 \text{ nm}$ .....	84
<b>Figura 5.24.</b> a) Imagen hiperespectral en falso color entre las bandas iniciales 15, 16 y 17, de los 4 suelos de Cusco. b) Imagen hiperespectral en falso color entre las bandas iniciales 15, 35 y 79, de los 3 suelos de Madre de Dios .....	85
<b>Figura 5.25.</b> a) y b) Blanqueamiento de ruido de la imagen de los suelos de Cusco y Madre de Dios respectivamente, luego de aplicar el MNF .....	87
<b>Figura 5.26.</b> Visualización de 12 bandas MNF, de la imagen hiperespectral del Cusco ...	88
<b>Figura 5.27.</b> Visualización de 16 bandas MNF, de la imagen hiperespectral de Madre de Dios .....	88
<b>Figura 5.28.</b> Gráfico de índices de pureza de píxeles, para los suelos de a) Cusco y b) Madre de Dios .....	89
<b>Figura 5.29.</b> Visualización n-D y distribución de la nube de píxeles en los extremos para: a) Suelos de Cusco y b) Suelos de Madre de Dios .....	90
<b>Figura 5.30.</b> Clasificación e identificación de miembros finales según la abundancia de píxeles para: Suelos de Cusco y b) Suelos de Madre de Dios .....	90
<b>Figura 5.31.</b> Firmas espectrales de los endmembers extraídos de las imágenes hiperespectrales de: a) Suelos de Cusco y b) Suelos de Madre de Dios .....	91
<b>Figura 5.32.</b> Pre procesamiento de la imagen hiperespectral: PC1 de Cusco y Madre de Dios .....	93
<b>Figura 5.33.</b> Muestra la carga de los primeros cinco componentes principales de los suelos de Cusco .....	94
<b>Figura 5.34.</b> Muestra la carga de los primeros cuatro componentes principales de los suelos de Madre de Dios .....	94
<b>Figura 5.35.</b> Porcentaje de variabilidad explicados por los primeros cinco componentes principales dentro de las bandas de la imagen de Cusco .....	95
<b>Figura 5.36.</b> Porcentaje de variabilidad explicados por los primeros cuatro componentes principales dentro de las bandas de la imagen de Madre de Dios .....	95
<b>Figura 5.37.</b> Visualización de izquierda a derecha de las 12 PCs de los suelos de Cusco ..	96
<b>Figura 5.38.</b> Visualización de izquierda a derecha de las 12 PCs de los suelos de Madre de Dios .....	96
<b>Figura 5.39.</b> Histogramas de frecuencia de los scores de los suelos de Cusco y Madre de Dios .....	98
<b>Figura 5.40.</b> Imágenes hiperespectrales de los suelos de Cusco y Madre de Dios sin fondo .....	98
<b>Figura 5.41.</b> a) y c) Selección de las regiones de interés (ROI's) mediante clasificación supervisada de cada tipo de suelo de Cusco y Madre de Dios. b) y e) Histogramas de cada imagen hiperespectral sin fondo .....	99
<b>Figura 5.42.</b> a) y c) Grupo de firmas espectrales de cada región de interés de cada suelo. b) y d) Firmas espectrales representativas de cada tipo de suelo .....	100

## Índice de tablas

<b>Tabla 1.1.</b> Principales divisiones del Espectro Electromagnético .....	8
<b>Tabla 2.1.</b> Modo de detección y escaneo de imágenes hiperespectrales [26] .....	20
<b>Tabla 2.2.</b> Identificación de materiales en distintos rangos de longitud de onda .....	23
<b>Tabla 2.3.</b> Principales diferencias entre métodos convencionales espectroscopia e imagen hiperespectral [54] .....	24
<b>Tabla 3.1.</b> Límites de tamaño de partícula .....	30
<b>Tabla 3.2.</b> Muestra los colores más comunes y su composición .....	34
<b>Tabla 3.3.</b> Clasificación natural del soil taxonomy - USDA [72] .....	36
<b>Tabla 3.4.</b> Clasificación según la FAO [65] .....	37
<b>Tabla 3.5.</b> Comparativa de ambas clasificaciones .....	38
<b>Tabla 3.6.</b> Relación aproximada de suelos del Cusco con las principales clasificaciones del Soil Survey Staff (USA) y de la FAO/UNESCO .....	41
<b>Tabla 3.7.</b> Relación aproximada de suelos del Madre de Dios con las principales clasificaciones del Soil Survey Staff (USDA) y de la FAO/UNESCO .....	42
<b>Tabla 5.1.</b> Cantidad y descripción de muestras a tratar .....	62
<b>Tabla 5.2.</b> Peso de cada muestra húmeda (MH), sin y con recipiente (S/R y C/R) .....	63
<b>Tabla 5.3.</b> Resumen de resultados del empastillamiento de las muestras .....	65
<b>Tabla 5.4.</b> Calculo del contenido de humedad (%) en las muestras, S/R (sin recipiente), C/R (con recipiente) .....	67
<b>Tabla 5.5.</b> Especificaciones de la cámara Pika NIR – 320 .....	68
<b>Tabla 5.6.</b> Softwares empleados para el procesamiento de imágenes hiperespectrales ...	70
<b>Tabla 5.7.</b> Abundancia de pixeles de los suelos de Cusco y el área en cm <sup>2</sup> que representan en la imagen .....	86
<b>Tabla 5.8.</b> Abundancia de pixeles de los suelos de Madre de Dios y el área en cm <sup>2</sup> que representan en la imagen .....	86
<b>Tabla 5.9.</b> Cuadro de estimación de abundancias obtenidos con el algoritmo PPI para los suelos de Cusco .....	92
<b>Tabla 5.10.</b> Cuadro de estimación de abundancias obtenidos con el algoritmo PPI para suelos de Madre de Dios .....	92
<b>Tabla 5.11.</b> Valores de Porcentaje de variabilidad PC para suelos de Cusco y su valor acumulado .....	97
<b>Tabla 5.12.</b> Valores de Porcentaje de variabilidad PC para los suelos de Madre de Dios y su valor acumulado .....	97
<b>Tabla 5.13.</b> Cuadro de estimación de abundancias obtenidos con el PC1 para los suelos de Cusco .....	101
<b>Tabla 5.14.</b> Cuadro de estimación de abundancias obtenidos con el PC1 para suelos de Madre de Dios .....	101
<b>Tabla 5.15.</b> Cuadro comparativo de estimación de abundancias entre los diferentes métodos de clasificación .....	102
<b>Tabla 5.16.</b> Cuadro comparativo de estimación de abundancias entre valores estandarizados versus el algoritmo de PPI .....	103

## RESUMEN

El uso de imágenes hiperespectrales es una importante herramienta tecnológica actual, la cual se aplica también al estudio de suelos y permite entre otras cosas la elaboración de mapas litológicos y una biblioteca espectral. En esta investigación, se trabajó con 7 muestras de suelos de 2 regiones del Perú (Cusco - Región kastanosólica y Madre de Dios – Región acrisólica) con el fin de identificar y analizar su firma espectral.

En el procesamiento de las imágenes hiperespectrales obtenidas se usó la técnica del algoritmo PPI (Pixel Purity Index) usando el software ENVI. En esta técnica se aborda el problema de la identificación de píxeles puros (píxeles con distintas firmas espectrales) y busca distinguir los distintos elementos (suelos) que conforman una imagen hiperespectral. A este proceso se le denomina extracción de Endmembers (miembros finales). Paralelamente a esta técnica se hace un análisis de componentes principales (PCA) con el fin de contrastar resultados.

Los resultados muestran que es posible identificar y construir una librería de firmas espectrales de distintos suelos en base a la región de procedencia usando la cámara hiperespectral.

**Palabras Clave:** imágenes hiperespectrales, firmas espectrales, PPI, MNF y PCA.

## ABSTRACT

The use of hyperspectral images is an important current technological tool, which is also applied to the study of soils and allows, among other things, the elaboration of lithological maps and a spectral library. In this research, we worked with 7 soil samples from 2 regions of Peru (Cusco - Kastanosolic Region and Madre de Dios - Acrisolic Region) in order to identify and analyze their spectral signature.

In the processing of the obtained hyperspectral images, the PPI algorithm (Pixel Purity Index) was used using the ENVI software. This technique addresses the problem of the identification of pure pixels (pixels with different spectral signatures) and seeks to distinguish the different elements (soils) that make up a hyperspectral image. This process is called extraction of Endmembers (final members). Parallel to this technique a principal component analysis (PCA) is made in order to contrast results.

The results show that it is possible to identify and build a library of spectral signatures of different soils based on the region of origin using the hyperspectral camera.

**Keywords:** hyperspectral images, spectral signatures, PPI, MNF y PCA.

## INTRODUCCIÓN

El uso de cámaras hiperespectrales en el análisis de suelos, minerales y productos orgánicos es un importante avance en los estudios de suelos. La coloración de la superficie terrestre representa una característica básica para su identificación y puede hacerse con suficiente precisión a partir de mediciones realizadas incluso con solo cinco longitudes de onda [1][2]. A partir de la reflectancia se puede obtener firmas espectrales que ayudan a determinar las concentraciones de carbonatos, óxidos de hierro y aluminio en suelos áridos y semiáridos [3][4][5][6].

El suelo es un componente crítico del ecosistema de la Tierra y juega un papel importante en roles de la producción agrícola de alimentos, cambio climático global, ciclo hidrológico, diversidad biológica y protección ambiental [7]. En las últimas décadas, debido a la disminución en la cantidad y calidad del suelo [8], la necesidad de monitorear y evaluar el suelo a diferentes escalas temporales y espaciales ha aumentado.

La información espectral de un cuerpo es el resultado de la interacción de la radiación incidente con los objetos en estudio. La información obtenida es consecuencia principalmente de cuatro factores: i) contenido de materia orgánica; ii) humedad del suelo; iii) contenido de minerales, y iv) textura [9].

La reflectancia del suelo, es uno de los atributos más útiles para delinear las diferencias entre los suelos y describir las características de un perfil de suelo [10], sin embargo su medida se dificulta debido a la complejidad de los factores ambientales [11], especialmente en regiones montañosas cubiertas de vegetación densa. Así mismo para apoyar este trabajo, es necesario un análisis en el laboratorio de estos suelos con el fin de identificarlos en escenarios de grandes áreas.

Dentro de las cámaras hiperespectrales disponibles, resulta de particular interés las cámaras con capacidad de detección en el infrarrojo cercano (NIR, por sus siglas en ingles), lo cual resulta importante al estudiar elementos orgánicos como los que se encuentran en los suelos ([12][13][14][15][16][17][18][19][20][21]).

El objetivo de este trabajo es obtener perfiles espectrales de suelos de interés en condiciones controladas en el laboratorio que resulta útil para una posterior identificación y clasificación de suelos. la información espectral de los suelos fue obtenida con la cámara hiperespectral Pika NIR 320, entre (900 – 1700 nm) con 168 bandas y software Spectronon Pro, analizando 7 muestras de suelos representativos de dos departamentos; Cusco (4) y Madre de Dios (3)

con un patrón de calibración (teflón) al 100% de reflectancia. Para representar los perfiles espectrales de cada suelo, se utilizó la técnica del algoritmo PPI (Pixel Purity Index) que da solución al problema habitual que se tiene cuando se procesa y analiza una imagen hiperespectral, “Mezcla de píxeles” [22]. Al ejecutar este algoritmo sobre la serie de datos reales, obtenidos de las imágenes hiperespectrales de cada muestra se pueden identificar las firmas espectrales o miembros finales (endmembers) extraídos después de la ejecución del algoritmo. Estos resultados fueron contrastados con otros métodos de análisis y guarda una estimación confiable entre ellos.

### **Planteamiento del problema**

En la actualidad el Perú, según información del IV Censo Nacional Agropecuario, se observó que el total de superficie agropecuaria asciende a 38 millones 742 mil ha, siendo la mayor proporción en superficie no agrícola: pastos naturales con 18 millones 19 mil ha (46,5%), montes y bosques con 10 millones 939 mil ha (28,2%) y otros usos con 2 millones 659 mil ha (6,9%). La superficie agrícola del Perú, según resultados del último censo agropecuario arrojan un total de 7 millones 125 mil ha; que representan el 18,4% del total de la superficie agropecuaria [23]. Según estas cifras, la superficie agropecuaria va aumentando año tras año, siendo una de las actividades alternas frente al auge minero que vive el país. No obstante, el mal manejo de suelos y el uso indebido de fertilizantes, insumos químicos degradan y contaminan no solo los suelos, sino los productos cultivados allí; por ello se busca alternativas tecnológicas para el manejo, supervisión y control adecuado de suelos.

Existe una gran diversidad de técnicas analíticas en el estudio de suelos, siendo las técnicas espectroscópicas las que mayor auge han tenido recientemente ante la necesidad de buscar metodologías que permitan estudiar adecuadamente grandes extensiones de tierra y que sean amigables con el ambiente [24].

Una alternativa de estudio y tratamiento de suelos en esta investigación es incorporar el uso de técnicas de análisis de imágenes hiperespectrales, debido a la gran capacidad de obtener abundante información espectral de todo tipo de suelos en el rango de espectral de (900 a 1700 nm) [25], ya que dentro de esta región existen bandas características que hacen fácil identificar si el suelo está o no contaminado.

## **Justificación**

Originalmente, las imágenes hiperespectrales fueron desarrollados para aplicaciones en control/análisis remoto, las cuales utilizan imágenes satelitales de la tierra, luna y planetas, pero han encontrado un gran potencial en otros campos como: la agricultura, industria farmacéutica y alimentaria [26].

En el análisis de suelos tradicional se busca caracterizar física y químicamente el suelo, cuya información obtenida permite posteriormente incrementar y llevar un mejor control en la productividad, buscando prácticas agrícolas adecuadas para lograr una alta calidad de suelo[27], estos análisis han permitido potenciar las prácticas agrícolas tales como riego, fertilización y encalado de suelos bajo bosques inicialmente y convertidos en suelos de uso agrícola, esto implica identificar componentes en los suelos e insumos compatibles con los que se puede trabajar para determinado tipo de suelo [28].

En el Perú, el análisis físico y químico está basado en técnicas volumétricas, gravimétricas de espectrometría visible, difracción de rayos X, espectroscopia de absorción atómica (EAA) etc... [29].

Los estudios basados en sistemas hiperespectrales brindan amplia información espectral para identificar componentes diferentes dentro de una escena o imagen de áreas específicas, ventaja mayor a las técnicas tradicionales, cuyos resultados son óptimos y se muestran en tiempos relativamente menores a las otras técnicas.

Por lo expuesto, esta investigación incorpora como primera etapa el análisis e identificación de suelos en laboratorio bajo condiciones controladas a partir de imágenes hiperespectrales, buscando obtener firmas espectrales característicos de suelos de la región de Cusco y Madre de Dios y así poder elaborar una librería local de firmas espectrales para el uso agrícola u otros fines.

## **Objetivos**

### **Objetivo general**

Analizar e identificar suelos a partir de sus imágenes hiperespectrales en el rango infrarrojo cercano (900 – 1700 nm) utilizando métodos de clasificación.

### **Objetivos Específicos**

- Calibración de la cámara hiperespectral y familiarización con su fundamento físico y funcionamiento.
- Etiquetar y manipular (secado, molido y empastillado) las muestras de suelo.
- Desarrollar un protocolo de adquisición de imágenes hiperespectrales(HSI) de suelos.
- Adquirir y hacer el pre – procesamiento de las HSI de los suelos en estudio.
- Procesar las imágenes mediante algoritmos de clasificación no supervisada: PPI y K-means.
- Extraer firmas espectrales para su análisis cuantitativo y cualitativo.

### **Hipótesis**

A partir de la adquisición, procesamiento y análisis de imágenes hiperespectrales, es posible identificar y clasificar suelos procedentes de diferentes regiones.

### **Metodología**

Se desarrollará una metodología en seis etapas acordes a los objetivos específicos, donde se abarca aspectos teóricos y prácticos.

- a) Calibración de la cámara hiperespectral.

En esta primera metodología se debe asegurar las condiciones mínimas del funcionamiento de la cámara; ubicación, velocidad de la plataforma, iluminación, calibración de la lente, del blanco y negro, pues de ello depende el desarrollo óptimo de toda esta investigación [25][30].

- b) Etiquetado y manipulación de muestras de suelo.

De acuerdo a las referencias, los suelos deben ser preparados para su posterior estudio, por ello primero se le asignará una etiqueta a cada muestra de suelo para su

fácil reconocimiento [31], luego se procederá hacer un tratamiento y estudio bajo tres tipos de condiciones: suelo húmedo, suelo secado y finalmente suelo empastillado.

- c) Desarrollo de un protocolo de adquisición de imágenes hiperespectrales(HSI) de suelos[26].

Se debe seguir un protocolo de seguridad y de adquisición de imágenes hiperespectrales de suelos, para lo cual se implementa una normativa de manejo del equipo y de la toma de imágenes de suelos, con el fin de adquirir las posteriores imágenes de manera correcta.

- d) Adquisición y pre – procesamiento de las HSI de los suelos en estudio.

Una vez seguido paso a paso el protocolo de adquisición de HSI, se procede a tomar las imágenes hiperespectrales de las 7 muestras de suelo. Dichas imágenes son guardadas en un archivo (.bil). El pre – procesamiento consiste en; detectar bordes, realzar bordes, eliminar ruidos y suavizar la imagen, para dichos procedimientos se utiliza métodos de reducción de dimensionalidad [32] que optimizan el procesado y disminuyen el coste computacional, dicho trabajo se implementa en las plataformas de Envi y RStudio.

- e) Procesamiento de las imágenes mediante algoritmos de clasificación [33].

Se entrena algoritmos de clasificación no supervisada en plataformas de Envi y RStudio, con la finalidad de identificar y clasificar suelos a partir de sus respuestas espectrales, y encontrar sus abundancias.

- f) Extracción de firmas espectrales.

Cada suelo muestra una gran cantidad de espectros, debido a su composición y estado, por ello a partir del procesado se extraerá las firmas características de cada tipo de suelo para su posterior estandarización y comparación con una librería espectral [34] .

# CAPÍTULO 1

## 1. FUNDAMENTO TEÓRICO

### 1.1. Fundamentos físicos

El término “Percepción remota” abarca un número de técnicas que utilizan información acerca de las propiedades de la superficie y los objetos a distancia; en general se define como la adquisición de datos físicos de manera no invasiva y sin contacto de objetos [35].

Estas técnicas permiten la obtención de información a partir de la medición de la energía electromagnética de los objetos distantes, que pueden ser diferentes materiales y de esta manera se pueden identificar y clasificar los objetos por clase o tipo, materiales que los componen y su distribución espacial. Estos resultados son obtenidos analizando la radiación electromagnética en una o más espectros electromagnéticos, ya sea reflejada, dispersada o emitida desde cualquier superficie terrestre [36].

#### 1.1.1. Energía electromagnética

La energía electromagnética esta relacionada con las variaciones del campo electromagnético que viaja en forma de una onda armónica. El concepto de onda explica cómo se propaga la energía electromagnética y sólo puede ser detectada cuando interactúa con alguna materia [37]. Su naturaleza de onda permite describirlo en términos de: longitud de onda, frecuencia y amplitud [36] tal como se muestra en la figura 1.1.

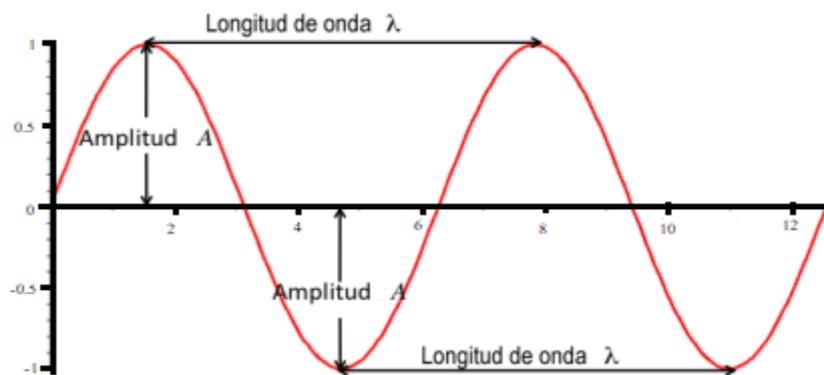


Figura 1.1. Se muestran las características de longitud de onda y amplitud de las ondas.

Donde:

- a) Longitud de onda ( $\lambda$ ). Es la distancia entre dos puntos en la misma fase de la onda, puede ser medida en cualquier unidad de longitud.

b) Frecuencia ( $\nu$ ). Es medida como el número de crestas que pasan por un punto fijo en un periodo de tiempo definido. Las unidades de medición son Hertz, el cual equivale a un ciclo por segundo.

c) Amplitud (A). Es el equivalente a la intensidad de cada pico, generalmente es medida como niveles de energía, expresada como Watts por metro cuadrado por micrómetro.

La energía electromagnética es generada por diversos mecanismos entre los cuales se incluyen los cambios en los niveles de energía de los electrones, la aceleración de las cargas eléctricas, el decaimiento radioactivo de sustancias y el movimiento térmico de los átomos y las moléculas. Las reacciones nucleares dentro del Sol producen un espectro completo de radiación electromagnética, que se transmite a través del espacio sin experimentar cambios importantes. Cuando esta radiación se acerca a la Tierra, atraviesa la atmósfera antes de llegar a la superficie de la Tierra. Algunos se reflejan hacia arriba desde la superficie de la Tierra; Es esta radiación la que forma la base para fotografías e imágenes similares [36].

Otra radiación solar se absorbe en la superficie de la Tierra y luego se irradia como energía térmica. Esta energía térmica también se puede utilizar para formar imágenes de detección remota, aunque difieren mucho de las fotografías aéreas formadas a partir de la energía reflejada. Finalmente, la radiación artificial, como la generada por los radares de imágenes, también se usa para la teledetección.

## **1.2. Espectro electromagnético**

Cada fotón del espectro electromagnético tiene una longitud de onda determinada por su nivel de energía. La luz y otras formas de radiación electromagnética se describen comúnmente en términos de sus longitudes de onda y frecuencia [38]. El espectro electromagnético se muestra en la Figura 1.3. Todas las divisiones son establecidas por conveniencia y tradición entre las diferentes disciplinas, ver la tabla 1.1 y se pueden encontrar diferentes definiciones en diversas fuentes de consulta [39].

Los aspectos que no se muestran en la tabla 1.1 son:

1) El espectro óptico que comprende desde 0.3 - 15  $\mu\text{m}$ , las cuales son las longitudes de onda que pueden ser reflejadas y refractadas con lentes y espejos.

2) El espectro de reflexión, el cual se extiende desde 0.38 – 3.0  $\mu\text{m}$  que corresponde a la porción del espectro solar utilizado directamente en percepción remota [36].

Tabla 1.1. Principales divisiones del Espectro Electromagnético. En esta tabla se presenta una subdivisión para el caso de la radiación infrarroja.

DIVISIÓN	LONGITUD DE ONDA
Rayos gamma	< 0,03 nm
Rayos X	0.03 - 300 nm
Radiación ultravioleta	0.30 – 0.38 $\mu\text{m}$
Luz visible	0.38 – 0.72 $\mu\text{m}$
Radiación infrarroja	
Infrarrojo cercano	0.72 – 1.30 $\mu\text{m}$
Infrarrojo medio	1.30 – 3.00 $\mu\text{m}$
Infrarrojo lejano	7.1 – 1,000 $\mu\text{m}$
Microondas	1 mm – 30 cm
Radio	$\geq$ 30 cm

La región visible del espectro electromagnético constituye una parte muy pequeña, sus límites están definidos por la sensibilidad visual de los humanos. Las propiedades ópticas de la radiación visible fueron investigadas por Isaac Newton, quien realizó experimentos entre los años de 1665-1666, los cuales revelaron que la luz visible se puede dividir en tres segmentos y son conocidos en función de los colores primarios aditivos, que se encuentran contenidos en las siguientes categorías: azul (0.4 – 0.5  $\mu\text{m}$ ), verde (0.5 – 0.6  $\mu\text{m}$ ) y rojo (0.6 – 0.7  $\mu\text{m}$ ).

Los colores primarios están definidos de forma tal, que no pueden ser formados a partir de ninguna combinación de los otros colores, se detalla en la figura 1.2.

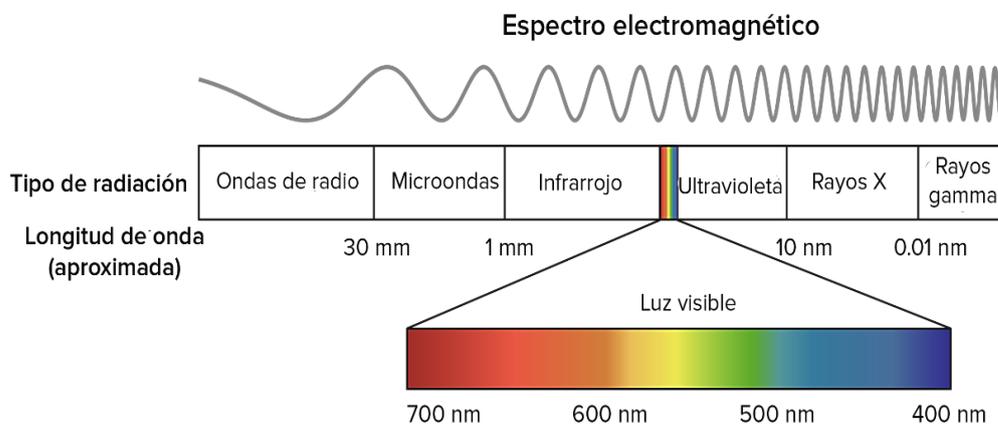


Figura 1.2. Espectro de ondas electromagnéticas en términos de su longitud de onda [38].

### 1.3. Radiación

La radiación electromagnética consiste en un campo eléctrico (E), que varía en magnitud en una dirección perpendicular a la dirección de propagación. Este campo eléctrico está acoplado a un campo magnético (H) orientado de forma perpendicular al campo eléctrico el cual se propaga en fase con el campo eléctrico, tal como se muestra en la figura 1.3.

Todas las ondas electromagnéticas viajan a la misma velocidad en el vacío ( $c= 299,792$  km/s), ésta es comúnmente mencionada como velocidad de la luz siendo la luz una forma de denominar a la energía electromagnética en la región del espectro electromagnético que corresponde al visible [37].

La frecuencia y la longitud de onda están relacionadas por la ecuación 1.1 y por lo tanto las características de la energía electromagnética pueden ser especificadas en términos de la frecuencia o la longitud de onda.

$$C = \lambda\nu \quad (1.1)$$

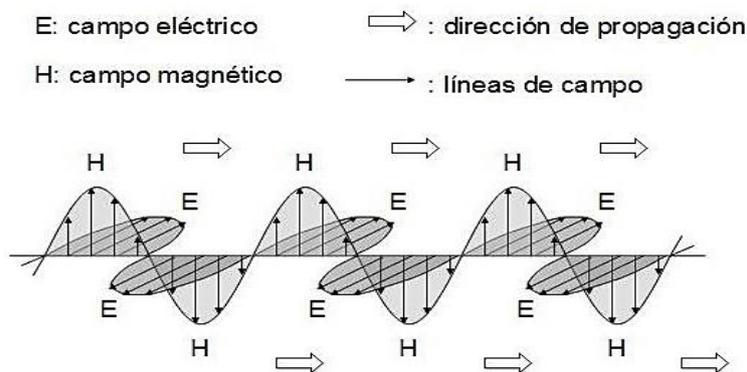


Figura 1.3. Las componentes del campo eléctrico (E) y el campo magnético (H) de la radiación electromagnética, están orientados perpendicularmente entre si y en ángulo recto con respecto a la dirección de propagación [36].

#### 1.3.1. Leyes de radiación

La propagación de la energía electromagnética obedece ciertas leyes físicas. En aras de un básico entendimiento y relación, algunas de estas leyes se describen en forma abreviada, como se describe a continuación [36].

1) Max Planck descubrió que la energía electromagnética es absorbida y emitida en unidades discretas llamadas cuantos o fotones. El tamaño de cada unidad es directamente proporcional a la frecuencia de la radiación de la energía. Planck definió una constante ( $h$ ) para relacionar la frecuencia ( $\nu$ ) con la energía radiante, ver ecuación 1.2. Este modelo explica el efecto foto eléctrico, la generación de corrientes por la exposición de ciertas sustancias a la luz, como el efecto del impacto de estas unidades discretas de energía (cuantos) sobre la superficie de ciertos metales, causando la emisión de electrones.

$$E = h \cdot \nu \quad \text{ó} \quad E = \frac{hc}{\lambda} \quad (1.2)$$

Donde: “ $h$ ” es la constante de Planck ( $6,626 \times 10^{-34}$  J.s) y “ $\nu$ ” es la frecuencia en Hertz.

2) Stefan Boltzmann indica que cualquier cuerpo en la naturaleza emite radiación y lo hace con diferentes longitudes de onda. La cantidad de energía que emite un cuerpo por radiación como la distribución de esta energía en diferentes longitudes de onda depende fundamentalmente de la temperatura del cuerpo (ley de Stefan-Boltzmann):

$$M = \sigma T_e^4 \quad (1.3)$$

Donde:  $T_e$  es la temperatura efectiva, es decir la temperatura absoluta de la superficie y “ $\sigma$ ” es la constante de Stefan-Boltzmann ( $\sigma = 5.67 \times 10^{-8} \text{ W/m}^2 \text{ K}^4$ )

3) La Ley de Wien nos permite calcular a que longitud de onda se produce el pico de máxima emisión de energía de un cuerpo a una temperatura  $T$ . Como puede observarse a mayor temperatura, menor longitud de onda, ver ecuación 1.4.

$$\lambda_{\max} = \frac{2898}{T} \quad (1.4)$$

La temperatura del Sol es de aproximadamente 6000 K mientras que la tierra es de 300 K. Esto significa que la radiación solar va a ser máxima para una longitud de onda de  $0.48 \mu$  (región visible) y la terrestre máxima en una longitud de onda de  $9.66 \mu$  (infrarrojo térmico).

Por tanto, puede concluirse que la radiación solar domina aquellas regiones del espectro electromagnéticos que corresponden a la radiación visible y al infrarrojo cercano. La radiación terrestre domina el infrarrojo lejano.

4) Ley de Kirchhoff establece que el cociente de radiación emitida entre el flujo radiante es la misma para todos los cuerpos negros a la misma temperatura.

Esta ley forma la base para la definición de Emisividad ( $\epsilon$ ), definida en la ecuación 1.5 como, el cociente entre la emitancia de un objeto cualquiera ( $M$ ) y la de un cuerpo negro a la misma temperatura ( $M_b$ ):

$$\epsilon = \frac{M}{M_b} \quad (1.5)$$

La emisividad de un cuerpo negro real sería 1, y la de un reflector perfecto (un cuerpo claro) sería 0. Los cuerpos negros y los cuerpos claros son conceptos hipotéticos, aproximados en los laboratorios bajo condiciones forzadas.

#### **1.4. Interacción con la superficie**

En teledetección, la radiación usada en los sensores debe pasar por la atmosfera de la Tierra y bajo estas condiciones tienen un impacto en la calidad de imagen debido a las partículas suspendidas en la atmósfera o por moléculas de gases atmosféricos. En este trabajo ese impacto se reduce al mínimo, ya que el análisis se hace en laboratorio bajo condiciones controladas.

La proporción de energía electromagnética incidida pasa procesos que dependerá de la naturaleza de la superficie, la longitud de onda de la energía y el ángulo de iluminación. Entonces ocurre el fenómeno de reflexión, es cuando un rayo de luz es redirigido al impactar con una superficie no transparente. La naturaleza de la reflexión dependerá del tamaño de las irregularidades de la superficie en comparación a la longitud de onda de la radiación considerada.

De acuerdo a esto puede ocurrir:

**1.4.1. Reflexión especular.** Si la superficie es relativamente suave o liza para la longitud de onda, ésta redirigirá la radiación incidente en una sola dirección y cumplirá la ley de reflexión, el ángulo de incidencia será igual al ángulo de reflexión ( $\theta_i = \theta_r$ ).

**1.4.2. Reflexión difusa.** Cuando su superficie es rugosa o isotrópica, donde la longitud de onda de la energía incidente será dispersada en todas las direcciones, como se ve en la figura 1.4.

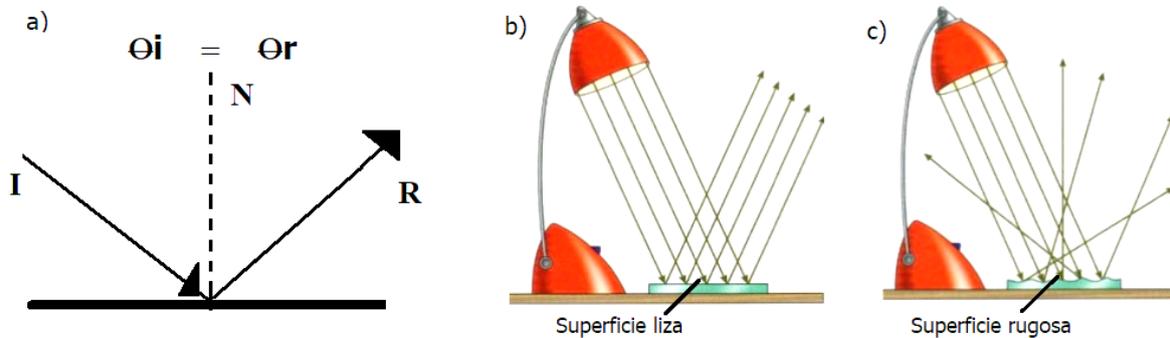


Figura 1.4. a) Ley de la reflexión, b) Esquema que muestra la reflexión especular y c) muestra la reflexión difusa [40].

### 1.5. Proceso de interacción entre energía y materia

Cuando la energía electromagnética alcanza la superficie del objeto de estudio, esta será: reflejada, absorbida, y transmitida.

Entonces cuando existe un flujo de energía radiante  $\phi_i$ , que alcanza la superficie de cualquier material, una parte de esta radiación es reflejada, originando un flujo reflejado,  $\phi_r$ . Otra parte es absorbida por la superficie, constituyendo un flujo absorbido  $\phi_a$ , por último, una fracción del flujo incidente será transmitida  $\phi_t$  normalmente en otras formas de energía.

A partir de la conservación de la energía, el flujo incidente puede expresarse en términos del flujo reflejado, absorbido y transmitido [36] [41] [42].

$$\phi_i = \phi_r + \phi_a + \phi_t \quad (1.6)$$

Sin embargo, es habitual expresar esta relación en términos de cantidades relativas, dividiendo la ecuación anterior por el flujo incidente obtenemos la siguiente relación:

$$1 = \frac{\phi_r}{\phi_i} + \frac{\phi_a}{\phi_i} + \frac{\phi_t}{\phi_i} = \rho + \alpha + \tau \quad (1.7)$$

Donde:

$$\rho = \frac{\phi_r}{\phi_i} = \text{reflectividad o reflectancia}$$

$$\alpha = \frac{\phi_a}{\phi_i} = \text{absortividad o absorbanza}$$

$$\tau = \frac{\phi_t}{\phi_i} = \text{transmisividad o transmitancia}$$

Hay que notar que la reflectividad, absorptividad y transmisividad, en términos más rigurosos depende de la longitud de onda [38].

### 1.5.1. Reflectancia

Es la fracción de radiación incidente que se refleja como: reflectancia especular, que no aporta información sobre la composición de la muestra y solo contribuye al ruido, por su parte la reflectancia difusa tiene lugar en todas las direcciones como consecuencia de los procesos de absorción y dispersión. Además, la reflectancia es una cantidad sin unidades y con valores en el rango [0 -1] en casos particulares son expresados en porcentajes. Existe una expresión empleada con mucha frecuencia para cuantificar su valor.

$$R = \frac{I_{refl}}{I_{in}} \quad (1.8)$$

Donde  $I_{refl}$  es la cantidad de energía reflejada e  $I_{in}$  es la cantidad de energía que incide sobre la muestra u objeto.

### 1.5.2. Transmitancia

Es una medida cuantitativa que resulta de la razón entre la cantidad de luz transmitida  $I_{out}$  por una muestra y la cantidad de energía incidida  $I_{in}$  sobre ella, tanto la energía radiante incidente como la transmitida deben ser medidas a la misma longitud de onda, tal como se muestra en la ecuación siguiente.

$$T = \frac{I_{out}}{I_{in}} \quad (1.9)$$

Donde  $I_{in}$  es la cantidad de energía que incide sobre la muestra,  $I_{out}$  es la intensidad transmitida.

### 1.5.3. Absorbancia

Se define como la cantidad de energía radiante absorbida por una sustancia o materia. Matemáticamente corresponde al logaritmo negativo de la transmitancia T, o al logaritmo de la inversa de la reflectancia 1/R, tal como se observa en la siguiente ecuación.

$$A = \frac{I_A}{I_{in}} \quad \text{ó para muestras opacas} \quad A = \log \frac{1}{R} \quad (1.10)$$

Donde: A es la absorbancia,  $I_A$  es intensidad absorbida y R la reflectancia.

## CAPÍTULO 2

### 2. VISIÓN HIPERESPECTRAL

En esta sección se describirá los conceptos fundamentales relacionados con las imágenes hiperespectrales, los tipos de sensores que se utilizan y modos de detección, la resolución de una imagen, variables que se obtienen, características de la cámara y su campo de aplicación.

#### 2.1. Visión artificial

En la actualidad, el análisis y procesamiento de imágenes se está volviendo cada vez más importante y generalizando porque se puede aplicar a diferentes campos de ciencia y tecnología. Por ejemplo, se ha utilizado para evaluar o cuantificar las características externas (es decir, color, tamaño, composición, forma y textura de la superficie) y estructuras internas (arquitectura y / o conectividad de los componentes del material) de las imágenes en estudio. El análisis de imágenes se basa en gran medida en la tecnología de visión artificial [43]. El crecimiento explosivo tanto en las plataformas de hardware como en los marcos de software ha llevado a avances significativos en el análisis de imágenes digitales.

##### 2.1.1. ¿Qué es una imagen hiperespectral?

Es una imagen espectral que proporciona datos digitales con mucha más información espectral (color) para cada píxel que las cámaras de color tradicionales [44]. Esto puede pensarse como una pila de decenas a cientos de imágenes con cada imagen sucesiva que representa su propio color específico (banda espectral), o de manera equivalente, como una curva espectral detallada para cada píxel.

En estas imágenes cada punto no viene descrito por un sólo valor de intensidad (como en una imagen "en blanco y negro") o por tres componentes de color (como en una imagen RGB de la pantalla del ordenador), sino por un vector de valores espectrales que se corresponden con la contribución de la luz detectada en ese punto a estrechas bandas del espectro. Se emplean el término "*imaging spectroscopy*" para referirse a los sensores de imagen hiperespectral, porque en cierto modo es como si estuviéramos midiendo simultáneamente con multitud de espectrómetros (que sólo miden la luz incidente en un punto) toda una rejilla de puntos en una gran zona geográfica [45].

La figura 2.4, muestra la salida de datos sin procesar y se visualiza a menudo como un "data cube", que es la estructura de una imagen hiperespectral; donde el eje X es el indicador de las líneas, el eje Y es el indicador de las muestras y representan la ubicación espacial de un

píxel, el eje Z representa la singularidad espectral de cada píxel y es el número de bandas, es decir, la longitud de onda.

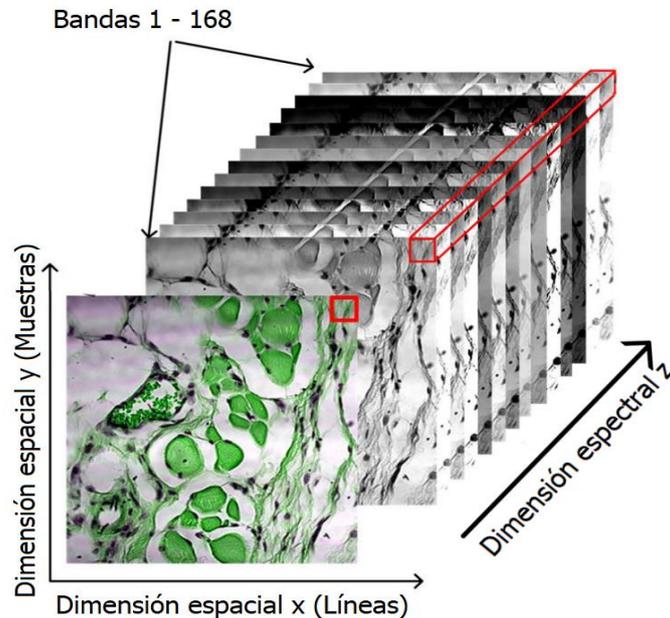


Figura 2.1. Estructura de una imagen hiperespectral [46].

## 2.2. Espectroscopía de imágenes hiperespectrales

Es una técnica espectroscópica producidas por instrumentos llamados espectrómetros de imágenes. Estos instrumentos son ópticos que sirve para medir las propiedades de la luz sobre una porción específica del espectro electromagnético.

Su principal función es identificar diferentes materiales que forman una escena, mediante la medida de la intensidad de la luz en función de la longitud de onda y son utilizados para realizar mediciones de campo o de laboratorio.

Para esta investigación se usa la espectrometría del infrarrojo cercano (NIR) donde el sistema de adquisición de imagen realiza un escaneo lineal o pushbroom, ya que registra línea por línea la imagen a lo largo de la muestra [44].

El funcionamiento radica en que el haz de luz reflejada, procedente de la línea de muestra escaneada, es dirigido a través de la lente hacia la abertura de entrada del espectrógrafo, ver figura 2.2. aquí se dispersa en las diferentes longitudes de onda del rango espectral medido y se proyecta en el detector de la cámara formando una imagen bidimensional; una dimensión representa el eje espectral (z) y la otra dimensión espacial (i) de la línea escaneada.

Moviendo la muestra línea tras línea a través de la zona de escaneo se consigue la segunda dimensión espacial (j) y tras juntar las imágenes bidimensionales generadas del escaneo completo de la muestra se obtiene la imagen hiperespectral tridimensional (i, j, z) [47].

Al usar cientos o miles de detectores, los espectrómetros pueden realizar mediciones espectrales de bandas tan próximas como  $0.01 \mu\text{m}$  sobre un rango amplio de longitudes de onda, típicamente al menos de  $0.4$  a  $2.4 \mu\text{m}$  (rangos de longitudes de onda del visible al infrarrojo medio)[48].

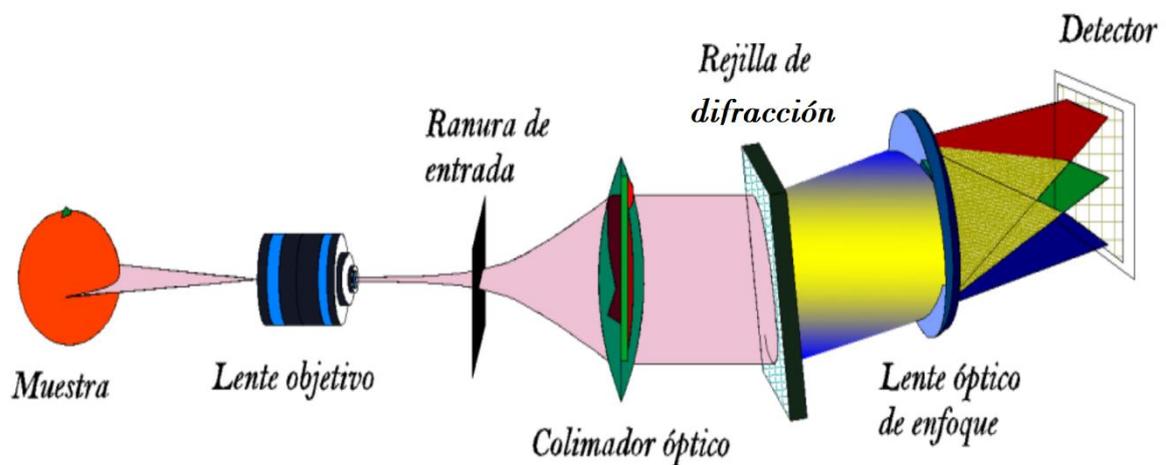


Figura 2.2. Diagrama esquemático del espectrógrafo empleado, arquitectura de un sensor de barrido de línea (pushbroom) [49].

### 2.3. Resolución de una imagen hiperespectral

En la adquisición de imágenes hiperespectrales se debe tener en cuenta algunos parámetros que son importantes como la resolución; que es la capacidad de distinguir objetos de un determinado tamaño en imágenes captadas a cierta distancia de una superficie. A mayor número de píxeles por unidad de superficie, mayor resolución del sensor, pero también mayor es el volumen del archivo informático generado.

Entre los principales tipos de resolución esta:

#### 2.3.1. Resolución Espacial:

Este concepto hace referencia al tamaño de cada pixel de la imagen y se refiere al tamaño de la característica más pequeña que puede ser detectada.

Dicha medición depende de la distancia del sensor con respecto a la superficie, del campo de visión instantáneo IFOV, del ángulo de visión FOV, la velocidad de escaneado y las características ópticas del sensor.

En algunos casos para determinar la resolución espacial de un sensor es necesario conocer el Campo de Visión Instantáneo IFOV, figura 2.2, que se define como la sección angular observada en un momento determinado.

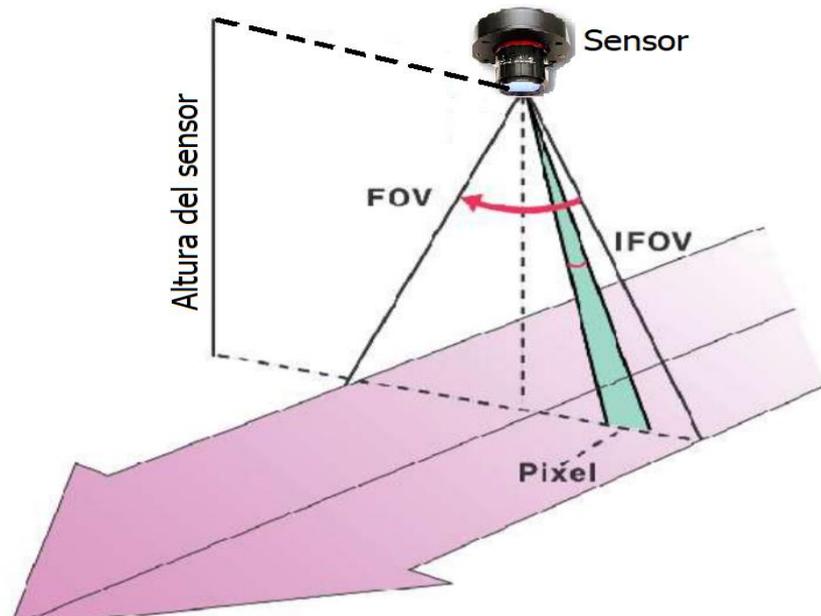


Figura 2.3. Campo de visión instantáneo (IFOV).

El IFOV está en función de la óptica del sensor, del tamaño del detector y de la altitud a la que se encuentra respecto a la superficie. La relación entre el IFOV y el tamaño de pixel ( $d$ ) viene dada por la siguiente ecuación:

$$d=2H\tan\left(\frac{IFOV}{2}\right) \quad (2.1)$$

Donde,

$d$  =Tamaño del pixel

$H$  = Distancia del sensor a la superficie terrestre

### 2.3.2. Resolución Espectral:

Es la medida del ancho de banda, mientras más números de canales espectrales tenga el sensor, el ancho de banda será muy estrecho y esto ayuda a separar de forma precisa distintos

objetos por su comportamiento espectral y los elementos de interés podrán ser distinguidos mediante la comparación de las curvas espectrales. Por ejemplo: El agua, la vegetación, los minerales, tienen una respuesta diferente a la energía electromagnética reflejada.

Dependiendo del tipo de sensor utilizado, se puede disponer de datos hiperespectrales formados por entre 25 y cientos de bandas por imagen. Actualmente existen sensores que son capaces de proporcionar más de 1000 bandas de información, a diferencia de los sensores multispectrales que solo tienen no más de 20 bandas, ver figura 2.3.

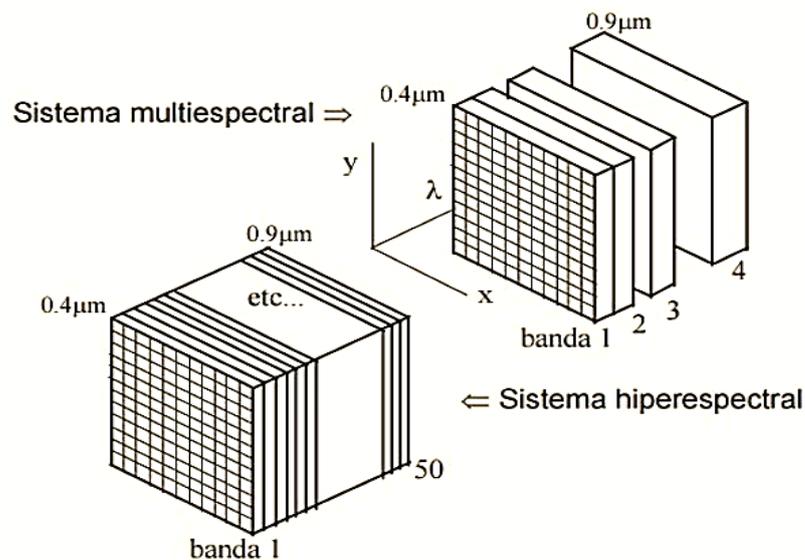


Figura 2.4. Comparación de resolución espectral de una imagen multispectral e hiperespectral [50].

### 2.3.3. Resolución Radiométrica:

Representa el número de bits (Unidad de medida de la velocidad de transmisión de los datos) necesarios para la representación del valor numérico de un pixel en una determinada banda de frecuencia, este valor numérico representa la cantidad de energía electromagnética reflejada por la superficie y captada por el sensor. Por lo tanto, indica la sensibilidad del sensor discriminando los niveles de radiancia espectral para un determinado rango de longitud de onda. Cuanto mayor sea la precisión radiométrica mayor número de detalles podrán captarse en la imagen.

Generalmente, a mayor resolución espacial disminuye la temporal y también la espectral. El problema es que cada aumento de resolución multiplica el volumen de datos que el sensor

debe procesar. La alta dimensionalidad de los datos obtenidos dificulta el análisis y almacenamiento de los mismos [50].

#### **2.3.4. Resolución Temporal:**

Hace referencia a la periodicidad con que el sensor adquiere información de un mismo sector de la superficie observada. Es el tiempo que transcurre entre dos imágenes. Resulta de gran utilidad cuando se pretende realizar un estudio evolutivo, por ejemplo: agricultura de precisión, meteorológico u oceanográfico.

#### **2.4. Tipo de sensores hiperspectrales**

Son dispositivos electrónicos que se caracterizan por su capacidad para medir la radiación reflejada en una amplia gama de longitudes de onda, y registran información en cientos de canales espectrales[51] [52].

Con respecto a los tipos de sensores, existen dos grandes tipos:

- Sensores activos: poseen fuentes internas que generan artificialmente la radiación (radares).
- Sensores pasivos, que detectan la radiación electromagnética emitida o reflejada de fuentes naturales o del medio.

En esta investigación nos centraremos en este último, los sensores pasivos. Existe una gran variedad ellos: sistemas fotográficos, espectrómetros de imagen, hiperspectrales, multiespectrales y radiómetros.

En el caso de las imágenes hiperspectrales la tecnología utilizada es la CCD (dispositivo de carga acoplada) y los sensores son de tipo pasivo ya que se han diseñado para registrar la radiancia reflejada o emitida por una gran superficie.

Cada sensor tiene una manera particular de captar la energía, ya que cada uno de ellos tiene una orientación o inclinación según su eje central, éstos se clasifican en tipos de orientación:

- Orientación vertical: Utilizado en satélites de resolución espacial baja o media (meteosat landsat).
- Orientación oblicua: Utilizados principalmente para sensores de radar.
- Orientación modificable: Utilizados principalmente en sensores de alta resolución espacial y alta resolución temporal.

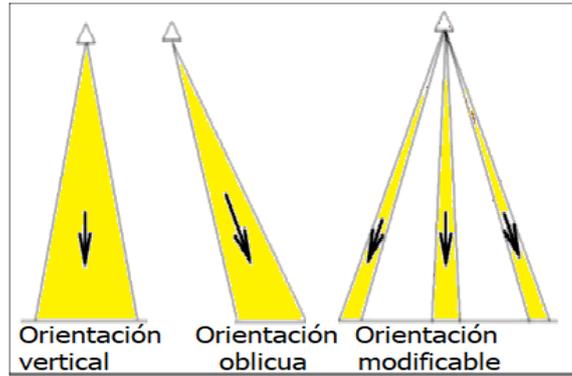


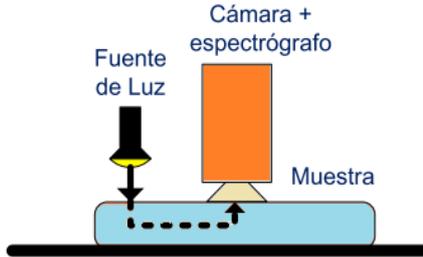
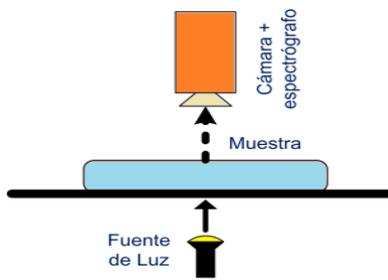
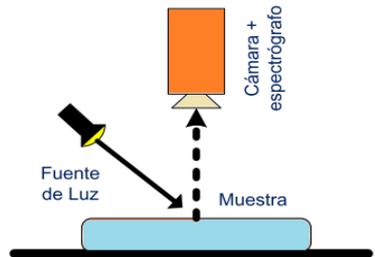
Figura 2.5. Tipos de orientación de los sensores.

### 2.4.1. Modo de detección espacial

Según estos tipos de captación de energía podemos identificar tres tipos básicos de sensores con su modo de detección y escaneo, tal como se muestra en la tabla 2.1.

Tabla 2.1. Modo de detección y escaneo de imágenes hiperespectrales[26].

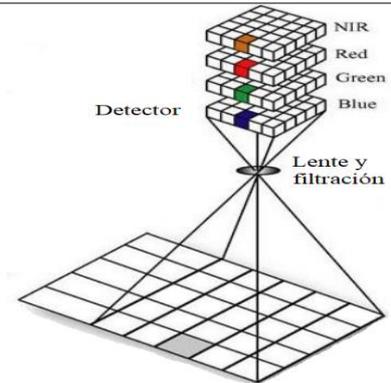
Modo de detección	Esquema
Reflectancia	<ul style="list-style-type: none"> <li>- El sensor captura la luz reflejada desde el mismo lado en que la muestra es iluminada.</li> <li>- Son utilizadas para la identificación externa de muestras.</li> </ul>
Transmitancia	<ul style="list-style-type: none"> <li>- El sensor capta la luz transmitida a través de la muestra que lleva información interna.</li> <li>- Su uso sirve para determinar la concentración de los componentes internos y defectos internos de objetos transparentes como pescado, frutas y verdura.</li> </ul>
Interactancia	<ul style="list-style-type: none"> <li>- La fuente de luz y el sensor están situados paralelamente y en el mismo lado.</li> <li>- Detecta información más profunda de la muestra y tiene menos efecto de superficie.</li> </ul>



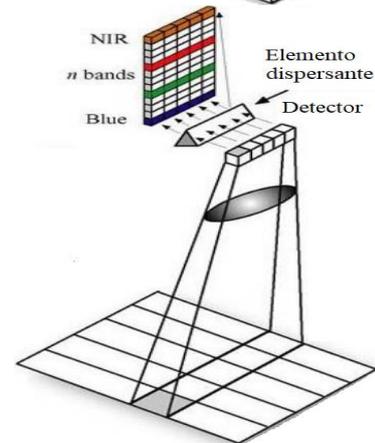
---

**Modo de escaneo****Esquema****Matricial o de Área**

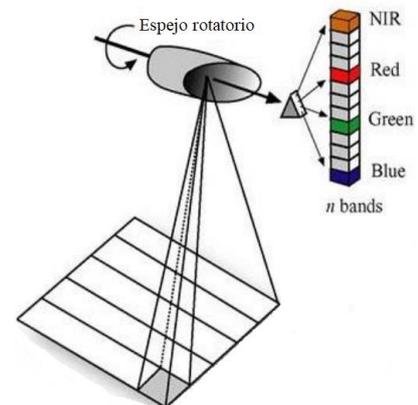
Este tipo de escaneo forma matrices de pixeles que componen una imagen. El uso de estos sensores se determina por sus características y sus aplicaciones, tales como la velocidad de obturación, el “Time delay integration” etc...

**Lineal o de empuje  
(Pushbroom  
Scanner)**

Como su propio nombre lo indica son sensores lineales de estado sólido, se encuentran alineados entre ellos, se encuentra perpendicular a la trayectoria del dispositivo.

**Barrido o de rotación  
(Whiskbroom  
Scanners)**

Son sensores electromecánicos, que utiliza la rotación de un prisma que hace mover el Campo de visión instantáneo IFOV perpendicularmente a la dirección del sensor, de esta forma se obtiene la imagen.



---

**2.5. Curvas espectrales**

Es la representación de las características que tienen los materiales que reaccionan con la radiación electromagnética incidente a cada banda de frecuencia. Cada curva espectral esta relacionando con la reflectancia espacial, que indica el comportamiento de la relación de la energía reflejada sobre la energía incidente y esto se muestra como una función de longitud de onda. A continuación, mostraremos las curvas espectrales de algunos elementos de nuestro entorno.

### 2.5.1. Reflectancia espectral de la vegetación, suelo y agua

Para mostrar las curvas espectrales de los elementos mencionados, ver figura 2.6, debemos destacar la importancia de cada una de las regiones del espectro que corresponden a un determinado tipo de transición atómica o molecular correspondientes a las diferentes energías, la parte visible (0.4  $\mu\text{m}$  – 0.7  $\mu\text{m}$ ) y la no visible (0.7  $\mu\text{m}$  – 2.4  $\mu\text{m}$ ) donde se ubican dichos elementos, estas regiones se dividen también en:

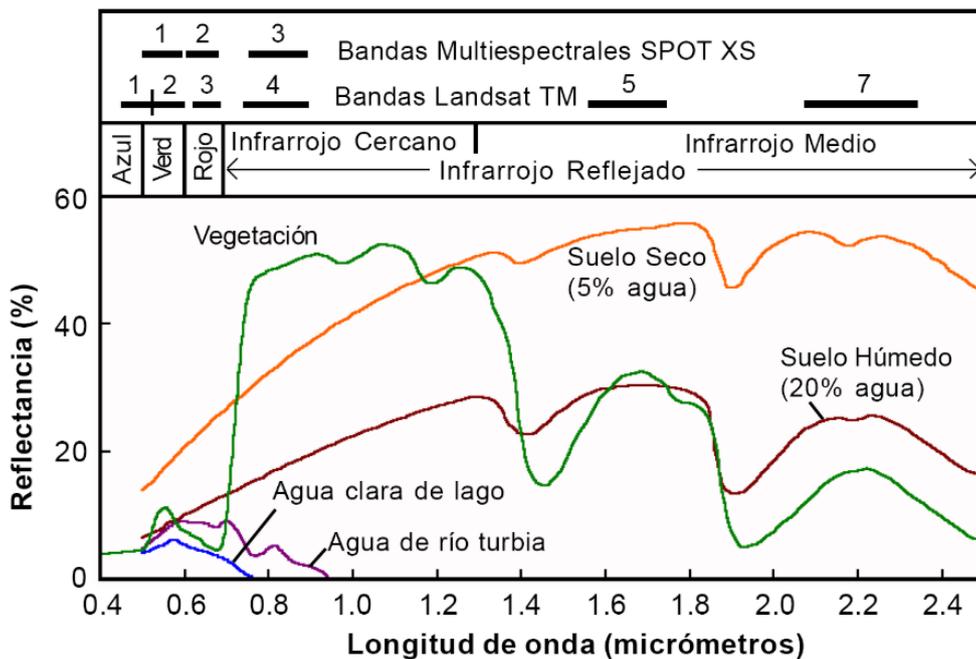


Figura 2.6. Curvas de reflectancia espectral para la vegetación, suelo y agua [53].

- **UV (ultra violeta):** Es la porción del espectro comprendida entre (0.1  $\mu\text{m}$  – 0.4  $\mu\text{m}$ ) donde la región del espectro tiene longitudes de onda corta, esto permite detectar en la superficie algunos materiales como rocas y minerales, incluso algunos de estos materiales emiten luz visible o fluorescente cuando son iluminados por la radiación UV.
- **IR (Infrarrojos):** Esta región comprende las longitudes de onda de radiación entre (0.7  $\mu\text{m}$  – 20  $\mu\text{m}$ ) basada en sus propiedades y a su amplia gama de longitudes, se subdivide en cinco regiones: infrarrojo cercano o NIR (0.7 $\mu\text{m}$  – 1.3 $\mu\text{m}$ ), infrarrojo de onda corta o SWIR (1.3 $\mu\text{m}$  – 2.5 $\mu\text{m}$ ), infrarrojo de onda media o MWIR (2.5 $\mu\text{m}$  – 8 $\mu\text{m}$ ), infrarrojo de onda larga LWIR (8 $\mu\text{m}$  – 14 $\mu\text{m}$ ) e infrarrojo lejano (14 $\mu\text{m}$  – 20 $\mu\text{m}$ ) de los cuales resaltan dos por la abundante información espectral en la parte geológica, ver la tabla 2.2:

Tabla 2.2. Identificación de materiales en distintos rangos de longitud de onda [31]

Subregión	Rango $\mu\text{m}$	Características: <b>Detectan elementos como:</b>
VNIR (Región visible e infrarrojo cercano)	<b>0.4 – 1.0</b>	<ul style="list-style-type: none"> <li>- Óxidos de hierro y minerales que contengan hierro.</li> <li>- Minerales raros de la tierra.</li> <li>- Vegetación.</li> <li>- Pirita</li> <li>- Oxido de Hematita</li> </ul>
SWIR (infrarrojo de onda corta)	<b>1.3 – 2.5</b>	<ul style="list-style-type: none"> <li>- Minerales que contienen OH (arcillas, micas, talco, etc...)</li> <li>- Sulfatos (alunita y yeso)</li> <li>- Filosilicatos (Micas, cloritas y arcillas)</li> <li>- Hidróxidos</li> <li>- Hidrocarburos</li> <li>- Carbonatos</li> </ul>

### 2.5.2. Librerías espectrales

Existen varias bibliotecas de espectros de reflectancia de materiales naturales y artificiales que están disponibles para uso público. Estas bibliotecas proporcionan una fuente de espectros de referencia que pueden ayudar a la interpretación de imágenes hiperespectrales y multispectrales [53], las más conocidas son:

- **Librería espectral ASTER.** Esta librería está accesible por la NASA, actualmente contiene cerca de 2000 espectrometrías (firmas espectrales), incluyendo minerales, suelos, rocas, agua, nieve e incluso materiales hechas por el hombre. La mayoría de las coberturas espectrales comprenden las longitudes de onda entre 0.4 y 1.4  $\mu\text{m}$ . Este registro se desarrolló como parte del programa instrumental Advanced Spaceborne Thermal Emission and Reflectance Radiometer (ASTER). Esta biblioteca incluye compilaciones espectrales del Jet Propulsion Laboratory de la NASA, Johns Hopkins University y United States Geological Survey. Esta información espectral se encuentra disponible en la web <http://speclib.jpl.nasa.gov>. La búsqueda de firmas espectrales se puede hacer por categorías, ver su gráfico e incluso descargar como archivo de texto.
- **Librería espectral USGS (The United States Geological Survey Spectroscopy).** Esta biblioteca fue elaborada en el laboratorio de espectrometría del United States

Geological Survey en Denver, Colorado, donde se ha compilado una biblioteca de cerca de 500 firmas espectrales de reflectancia de minerales y unas pocas plantas dentro de las longitudes de onda en el rango de 0.2 a 3.0  $\mu\text{m}$ . Esta biblioteca se encuentra disponible y es de acceso online en <http://speclab.cr.usgs.gov/spectral.lib04/spectral-lib04.html>. La búsqueda de una firma espectral es individual online y se puede descargar incluso la biblioteca completa.

## 2.6. Ventajas y desventajas de la información hiperespectral

Para hablar sobre las ventajas y desventajas [26][54] es necesario saber que existe métodos convencionales y técnicas espectroscópicas que guardan diferencias respecto a las imágenes hiperespectrales. Dichas principales diferencias se exponen en la tabla 2.3.

**Tabla 2.3. Principales diferencias entre métodos convencionales, espectroscopia e imagen hiperespectral [54].**

Características	Imagen 2D clásica	Espectroscopía	Hiperespectral
Información espacial	Sí	No	Sí
Información espectral	No/RGB	Sí	Sí
Composición química	No	Sí	Sí
Robustez	Sí	No	Sí

Entonces, a continuación, enumeraremos algunas ventajas de las imágenes hiperespectrales:

- Es un análisis no destructivo, no invasivo ya que es tomada a una distancia y por ende no altera a la muestra.
- No requiere preparación de las muestras, adquiriendo inmediatamente la imagen.
- Una vez realizada la calibración y puesta en marcha, se adquiere las imágenes hiperespectrales de manera rápida y sencilla.
- Gracias a la resolución espacial se puede analizar varios materiales en una misma escena.
- Se puede seleccionar regiones de interés descartando sombras o zonas de diferente composición.
- La información que se obtiene es más completa que la de un espectrógrafo al no

depender la medida de un solo punto y tener información redundante.

- Identifica compuestos químicos en una muestra, evitando comprar reactivos ni ningún material de análisis para realizar el estudio de la muestra.
- Se puede aplicar a diferentes campos de investigación.
- Contiene y se extrae información cuantitativo y cualitativo.

Ahora, por el contrario, mencionaremos las desventajas de las imágenes hiperespectrales.

- Necesidad de calibración precisa y de modelos robustos.
- Es necesario tener la iluminación controlada en el caso de trabajo en laboratorio, pero en un sistema aerotransportado se debe tomar factores del medio.
- Proporciona una gran cantidad de información redundante que conlleva al error y a un coste computacional a la hora del procesamiento.
- Es necesario un equipo de procesado alto con características acorde a su abundante información.
- Para muestras líquidas u homogéneas puede resultar más económico el uso de espectrografía.
- Es un método indirecto ya que requiere la utilización de modelos para su validación.
- Limitada penetración de la luz.

## **2.7. Campo de aplicación de las imágenes hiperespectrales**

Los sistemas hiperespectrales permiten distinguir de manera precisa y consistente entre colores o materiales similares, dichos sistemas son utilizados en plataformas que van portados en avionetas, drones o equipados en laboratorios para diferentes objetivos de estudio, por ejemplo:

### **2.7.1. Biomedicina**

Las imágenes hiperespectrales (HSI) ha emergido como una nueva técnica para identificar distintas afecciones de manera no invasiva y se basa en el muestreo de imágenes a diferentes longitudes de onda del espectro electromagnético, esta tecnología ha tenido un gran auge en el campo biomédico, como por ejemplo en la estimación de niveles de colesterol, para la detección de artritis en la piel, para la identificación de órganos, cáncer de piel [55] [56], cicatrización de la herida [57] y tumores cerebrales, mostrado en la figura 2.6 [58]. El enorme

potencial de las imágenes hiperespectrales reside en la gran cantidad de información que contienen, lo que permite distinguir materiales de forma detallada y no invasivos.

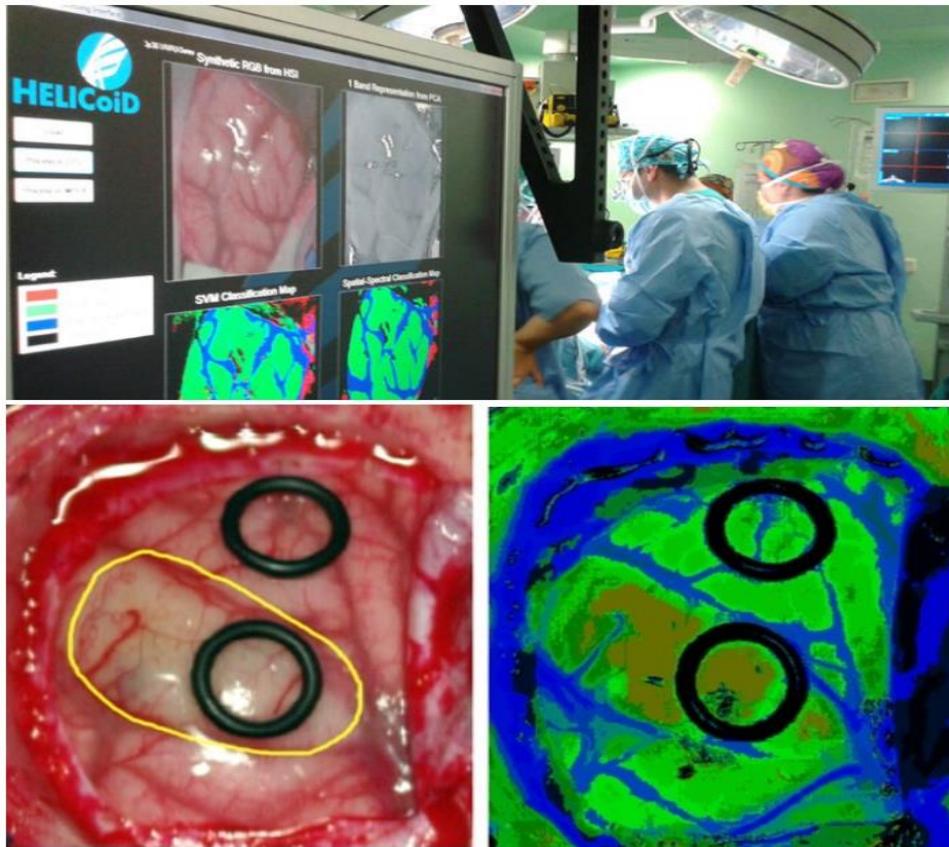


Figura 2.7. Predicción del diagnóstico de tejido cerebral. Imagen izquierda: diagnóstico según la vista del neurocirujano de tejido canceroso y sano. Imagen derecha: detección correcta de un tumor de grado IV (glioblastoma)[58].

### 2.7.2. Agricultura de precisión

Las imágenes hiperespectrales permite la identificación de malezas [59], el monitoreo de la salud de las plantas y la evaluación de la madurez. La detección temprana del estrés del cultivo es una aplicación común. Por ejemplo, se hizo estimaciones de LAI (índice de área foliar) y clorofila (CHL) a partir de trigo para predicción de rendimiento en función de la concentración de fertilizantes, ver figura 2.7. Usando una cámara hiperespectral en el rango de (400 – 1000 nm) [60].

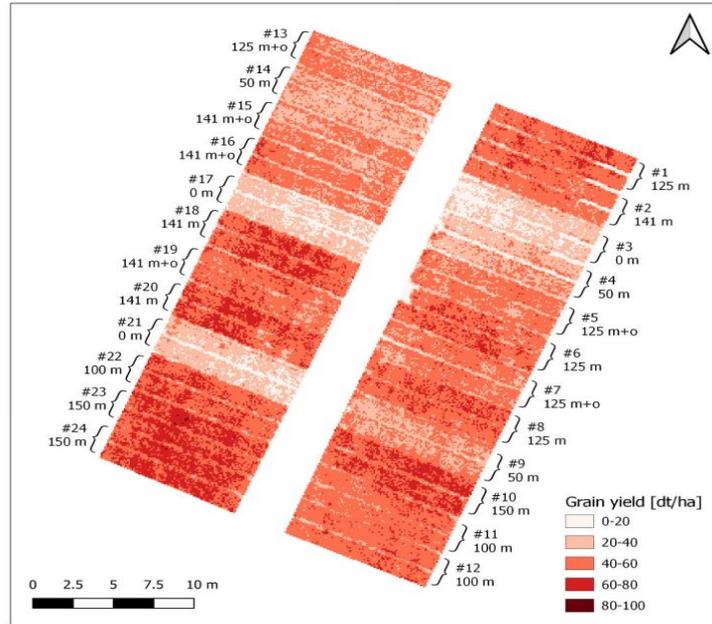


Figura 2.8. Predicción espacial para el rendimiento de grano basado en el modelo MLR [60].

### 2.7.3. Análisis de minerales

En el campo de minería, la caracterización del tipo de roca, tanto de la mena como la ganga, es de gran importancia no solo para definir el tipo de procesamiento para recuperar el metal de interés, sino para predecir su comportamiento en planta durante el proceso metalúrgico. A raíz de lo anterior diversos trabajos se enfocan en determinar un método que permita detectar y clasificar sulfuros y óxidos mediante el uso de tecnologías hiperespectrales [31].

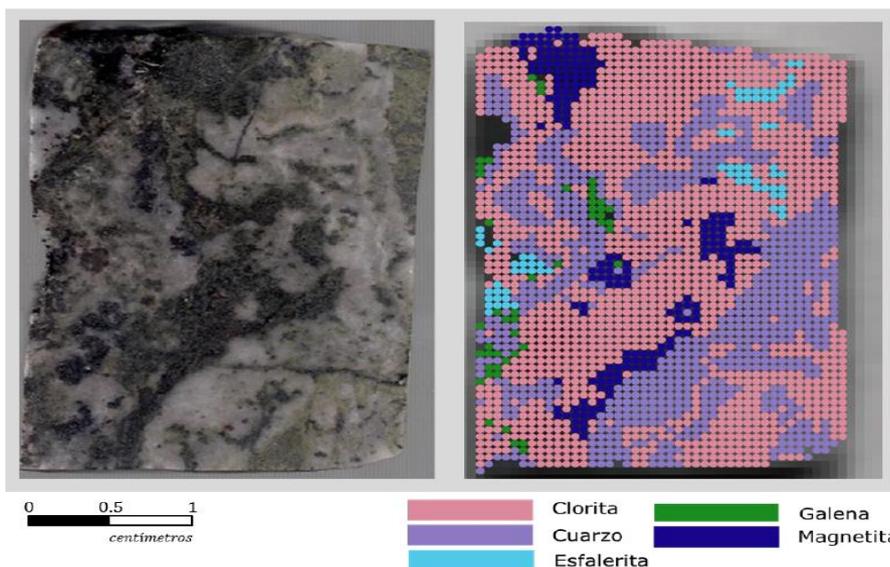


Figura 2.9. A) Fotografía de corte pulido del espécimen CTP-1. B) Fotografía de la imagen hiperespectral etiquetada del espécimen CTP-1 [31].

#### 2.7.4. Monitoreo ambiental

Las imágenes hiperespectrales también se utilizan para rastrear la salud del bosque, la calidad del agua y la contaminación de la superficie: suelos, ríos, mares y lagos: por ejemplo, se hicieron mediciones in situ de la cuenca del río Mississippi para detectar nitratos, fosfatos y sedimentos [61]. Por otro lado, el Instituto cooperativo de investigación de los Grandes Lagos de la Universidad de Michigan realizan campañas semanales de adquisición de imágenes hiperespectrales aéreas para rastrear y comprender la propagación de las floraciones de algas nocivas en el oeste del lago Erie [62].



Figura 2.10. imagen hiperespectral de agua de río contaminada [61].

#### 2.7.5. Visión de máquina para productos alimenticios e industriales

La visión artificial hiperespectral detecta pequeñas diferencias de color con mayor precisión e identifica diferentes materiales de manera más confiable que las imágenes convencionales. En el caso de alimentos caracteriza la calidad del producto, estado y localiza contaminantes. Por ejemplo esta tecnología se ha usado en la detección de podredumbre en cítricos y legumbres [30], detección de madurez en mango y melocotón [63], detección y adulteración de queso fresco con harina de maíz y detección de la calidad de carne de res [26]. Por otro lado, las imágenes hiperespectrales identifican falsificaciones, encontrando defectos y eliminando errores de prescripción en productos farmacéuticos [64].

## CAPITULO 3

### SUELOS DE CUSCO Y MADRE DE DIOS

En este apartado, describiéremos algunas características de los suelos y su comportamiento espectral en cada estado en el que se encuentra. También se enumerará algunos tipos de suelos presentes en los departamentos de Cusco y Madre de Dios, suelos con los que se trabajará.

#### 3.1. Aspectos generales de suelos

Entre los muchos factores que deben tomarse en consideración para obtener buenos resultados en el análisis de suelos a partir de su imagen hiperespectral, es saber la clasificación y sus constituyentes.

##### 3.1.1. ¿Qué es el suelo?

El suelo, según la FAO, es una mezcla compleja de organismos vivos, material orgánico, minerales, agua y aire. En general, se compone de partículas orgánicas (material descompuesto de vegetal y animal) y partículas de minerales (arena, arcilla, piedras o grava que provienen de rocas mayores) [65].

##### 3.1.2. Principales características físicas del suelo:

Sus características principales son la textura, la estructura, el pH, la salinidad del suelo y color.

❖ **Textura.** - Es la característica física que diferencia unos suelos de otros según el contenido de partículas de mayor o menor tamaño, se clasifican en tres grupos básicos: arena fina, limo y arcilla, los límites de tamaño para estos se dan en la tabla 3.1. Se han utilizado varios sistemas para su clasificación, uno de los más utilizados es el triángulo de texturas USDA-FAO que se muestra en la figura 3.1 dando los nombres de los suelos según diferentes proporciones de arena, limo y arcilla. Las determinaciones de textura se pueden hacer en un caso simple por “sensación” del suelo humedecido o por técnicas más precisas como: utilizando un hidrómetro, una pipeta o granulometría [66].

Tabla 3.1 Límites de tamaño de partícula

Suelo separado	Diámetro (mm)
Arena	
Muy grueso	2.0 – 1.0
Grueso	1.0 – 0.5
Medio	0.5 – 0.25
Fino	0.25 – 0.10
Muy bueno	0.10 – 0.05
Limo	0.05 – 0.002
Arcilla	< 0.002

Detallaremos de manera breve los conceptos de clasificación de los suelos según su textura, ver también la figura 3.1.

- **Arena:** Se le llama así al conjunto de partículas de rocas silíceas y de otro tipo, que se acumulan en las costas y los diversos ríos. Los suelos arenosos son de textura gruesa, ásperos y poco adhesivos.
- **Limo:** Estos tipos de suelo se componen de partículas más pequeñas, son una mezcla de arcilla, arena fina y es suave al tacto. Estos suelos retienen el agua por más tiempo, así como los nutrientes.
- **Arcilla:** Son grupos de minerales definidos, como caolinita, illita y montmorillonita, donde participan estructuras octaédricas y tetraédricas. La arcilla, como el humus, posee propiedades coloidales. Las arcillas, en el sentido mineralógico, son cristales microscópicos cuyos átomos están dispuestos en planos.

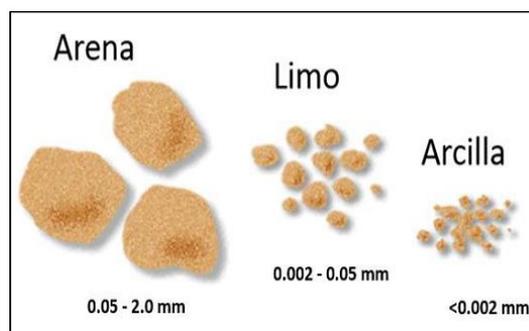


Figura 3.1. Partículas de suelo [67].

La figura 3.2 muestra el triángulo textural del suelo según la USDA y la FAO, que sirve para una correcta clasificación de suelo.

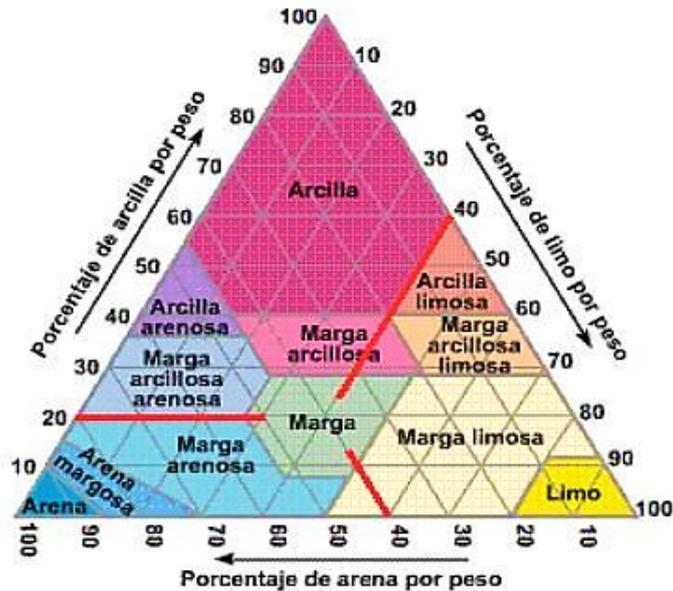


Figura 3.2. La figura representa al triangulo textural del suelo de acuerdo a la FAO y al Departamento de Agricultura de USA.

- ❖ **Estructura.** – En condiciones de campo, las propiedades de textura determinadas, pueden modificarse considerablemente debido a su estructura de suelo. Las partículas sólidas de los minerales tienen una disposición definida, y los espacios de poros entre ellos contienen en parte agua y en parte gases. La disposición de las partículas individuales en agregados o capas más grandes de varios tamaños y formas es la estructura del suelo. Aunque la "textura" y la "estructura" parecen usarse indistintamente en el uso de medios populares, existe una diferencia real en el significado. La arcilla y el limo no se extienden de manera uniforme por todo el suelo. Los finos cubren partículas de arena. Las unidades de arcilla individuales se unen en dominios de arcilla, como las hojas de un libro. Este es un tipo de orientación paralela, con los dominios de arcilla formando recubrimientos alrededor de partículas de arena o unidades de estructura del suelo. Otra distinción entre textura y estructura radica en el papel de la materia orgánica. Los coloides húmicos influyen considerablemente en las propiedades derivadas de la textura. La materia orgánica puede mejorar la capacidad de retención de agua de la arena y puede mejorar las propiedades de drenaje de la arcilla. Esto se logra promoviendo la formación de estructuras. Figura 3.3.

	<b>Sin estructura, un solo grano</b> Granos individuales; característico de arena suelta.
	<b>Sin estructura, mixto</b> No hay estructura visible; terrones poco desarrollados.
	<b>Migajosa</b> Más o menos redondeados, agregados finos (<2 mm).
	<b>Bloque</b> Capas más o menos equidimensionales, ya sean angulares o sub angulares en contorno.
	<b>Prismático</b> Capas alargados verticalmente con caras planas.
	<b>Columnar</b> Como arriba, pero con las puntas redondeadas debido a la reducción.
	<b>Placas</b> Placas alineadas horizontalmente.

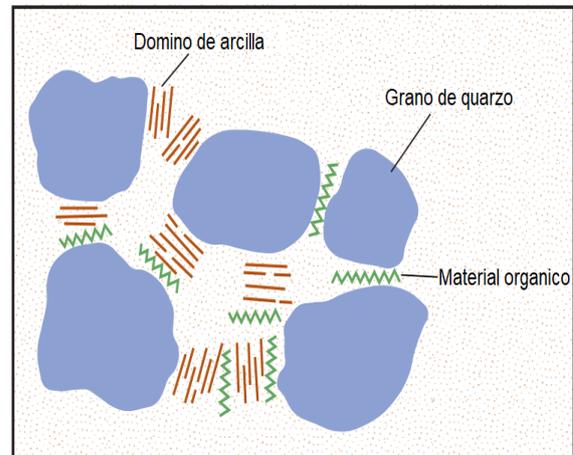


Figura 3.3. Formas de estructuras de suelos comunes (imagen derecha). Disposición de partículas de cuarzo, arcillas y de coloides orgánicos en un agregado de suelo (imagen izquierda). [68].

❖ **pH del suelo.** – Algunos suelos de regiones tropicales se han desarrollado bajo condiciones donde la precipitación excede la evapotranspiración durante periodos largos. Este hecho da lugar a procesos de lixiviación o lavado que han generado un agotamiento gradual de las bases del suelo (tales como Ca, Mg y K) y un desarrollo de la acidez del mismo. Las arcillas de los suelos contienen también capas de Fe y de Al y estos materiales afectan levemente la retención y la disponibilidad de los cationes y aniones en suelos ácidos. El pH del suelo es una medida de la acidez o de la alcalinidad de la solución del suelo [69]. Por definición, el pH es el negativo del logaritmo de la concentración del ión de hidrogeno [H\*], es decir:

$$\text{pH} = -\log[\text{H}^*] \quad (3.1)$$

Se dice que los suelos son ácidos, neutros, o alcalinos (básicos), dependiendo sus valores de pH, en una escala de 0 a 14 (ver figura 3.4).

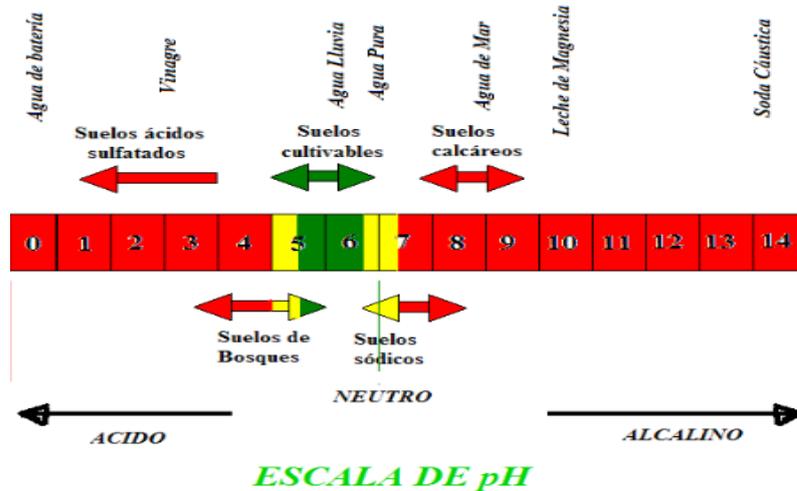


Figura 3.4. Escala de pH de suelos [69].

Los suelos minerales ácidos con pH inferiores a 5.0 contienen a menudo cantidades apreciables de Al y de Mn en la solución del suelo, que son perjudiciales para el crecimiento vegetal. El crecimiento óptimo y el uso eficiente de los nutrientes de los fertilizantes en suelos ácidos requieren de la adición de enmiendas para eliminar los efectos tóxicos del Al, H y Mn.

Los microorganismos del suelo no funcionan con eficacia en suelos ácidos. A medida que disminuye el pH del suelo disminuye también la actividad de los microorganismos que descomponen la materia orgánica y proveen nutrientes a las plantas. Aunque estos organismos funcionan mejor en niveles de pH del suelo de 8.0, su eficacia no cae rápidamente hasta que los niveles del pH están por debajo de 6.0. La descomposición de la materia orgánica no solo provee nutrientes, sino que también contribuye a la agregación (agrupamiento) de las partículas del suelo que genera buena estructura, aireación y drenaje del suelo [69].

❖ **Salinidad del suelo.** - La salinidad de los suelos puede ocurrir por efectos naturales, sin embargo, la principal causa es el mal manejo en la agricultura al momento del riego o drenaje, aplicación excesiva de estiércol de animal, fertilizantes y aguas residuales. La salinidad tiene efectos negativos en el desarrollo de los cultivos ya que en estas condiciones el potencial osmótico del suelo supera al del sistema de las plantas limitando así la entrada del agua en la raíz [70] además de ello la salinidad trae consigo otros problemas como:

- 1) Absorción limitada de nutrientes

- 2) Afecta la translocación y el reciclado de iones en la planta
- 3) Acumulación de Cl, Na y B en distintas partes de las plantas, como en las semillas, los tallos y las hojas.
- 4) El exceso de ciertos iones puede provocar toxicidad en las plantas.
- 5) La presencia en exceso de determinados elementos, provoca antagonismo entre nitratos-cloruro, potasio-sodio, calcio-sodio.

Ahora, para medir la salinidad de un suelo se hace a través de la Conductividad Eléctrica (CE) en el extracto de la saturación, de acuerdo a la metodología descrita por (Chavira y Castellanos 1987) o en una relación Suelo: Agua 1:2 o 1:5, la cual es la más práctica y sencilla, por lo que la medición de la salinidad en esta relación está tomando mucha importancia.

❖ **Color.** - Es una característica del suelo, fácil de observar y de uso cómodo para identificar un tipo de suelo, pero no puede ser evaluado con precisión al ser una apreciación subjetiva. Para ello se hacen comparaciones mediante una carta de colores, normalmente se utiliza la carta Munsell. El color del suelo generalmente está en relación con los procesos de edafogénesis o con uno de los factores de formación. Pero, por una parte, el proceso que colorea el suelo no es siempre fundamental, también la misma coloración, o matices vecinos bien pueden resultar de causas diferentes. Es así que esta característica debe ser utilizado con circunspección y sabiendo que raramente tiene valor como criterio de clasificación al nivel de los grandes tipos de suelos [71]. A continuación, veremos (tabla 3.2) las cuestiones diversas que se relacionan con el color del suelo: elementos y procesos que lo determinan, disposición en los suelos, apreciación e interpretación en el terreno.

Tabla 3.2. Muestra los colores más comunes y su composición.

Colores de Suelo	Elementos y Composición
Grises, blancos o rojizos	Arcillas
Café oscuro, negro o azulado	Abundante contenido de material orgánico-Humus
Blancos o grises	Carbonatos, calizos y cuarzo
Tonos rojizos	Feldespatos
Amarillo	Limonita ( $\text{Fe}_2\text{O}_3 \cdot n\text{H}_2\text{O}$ )
Pardos amarillentos	Hidróxido de hierro férrico
Rojizo	Hidróxido férrico $\text{Fe}(\text{OH})_3$
Verde grisáceo	Hidróxido ferroso $\text{Fe}(\text{OH})_2$

### **3.2. Clasificación de suelos**

Para la clasificación natural de los suelos normalmente se usa el sistema del Soil Taxonomy USDA (Departamento de Agricultura de los Estados Unidos), [72] y su correspondiente correlación son el sistema FAO (Organización de las Naciones Unidas para la Agricultura y la Alimentación) [65]. La taxonomía del suelo ha sido desarrollada para agrupar suelos de características similares para fines de manejo y para facilitar un lenguaje universal entre los agrónomos y especialistas en suelos del mundo. El sistema utiliza categorías, una categoría es un conjunto de clases de suelos definidas aproximadamente al mismo nivel de abstracción y generalización. En orden decreciente las categorías son seis: Orden, Suborden, Gran Grupo, Subgrupo, Familia, Serie. Según [72], en la categoría más alta se separan todos los suelos en un número pequeño de clases lo que permite comprender y recordar las doce clases: Histosols, Spodosols, Oxisols, Vertisols, Aridisols, Ultisols, Mollisols, Alfisols, Inceptisols, Entisols, Andisols y Gelisols.

#### **3.2.1 La clasificación según Soil Taxonomy, USDA 1999**

Esta clasificación taxonómica (USDA) fue publicada por el Soil Survey Staff, elaborada para utilizarse en suelos de cualquier parte del mundo. Basada en la identificación de horizontes diagnóstico como pedones y polipedones para determinar: Ordenes, Subórdenes, Gran grupo, Sub grupo, Familia y Serie. Para la determinación de órdenes se tomó en cuenta la clasificación natural de [72], tal como se muestra la tabla 3.3.

#### **3.2.2 La clasificación según la FAO, 1971**

La clasificación de la FAO se publicó por primera vez en 1974 como la leyenda del Mapa del suelo 1: 5,000,000 del mundo, y desde entonces se ha utilizado cada vez más para la comunicación internacional, con revisiones en 1985 y 1988. Con algunas modificaciones, la clasificación se utilizó en 1985 para el 1: 1,000,000 Mapa de suelos de los países de la CEE. La clasificación de la FAO consiste en una amalgama de nombres tradicionales (Podzol, Chernozem), nombres recién acuñados (Lixisol, Alisol) y préstamos de la Taxonomía de Suelos de los Estados Unidos (Histosol, Vertisol). En la revisión de 1988 hay veintiocho grupos de suelo, subdivididos en el segundo nivel en 153 unidades de suelo. Los grupos de suelos se enumeran en la Tabla 3.4 junto con una breve descripción de sus características principales [66].

Tabla 3.3. Clasificación natural del soil taxonomy - USDA [72].

N°	Nombre	Derivación Y Significado	Descripción
1	Entisol	A partir de (reciente)	Dominan los materiales inorgánicos; ausencia de horizontes bien definidos; en llanuras de inundación y suelos rocosos.
2	Vertisol	Latin verto (invertido)	Suelos arcillosos y oscuros que muestran fracturas anchas y profundas al secarse.
3	Inceptisol	Latin inceptio (inici)	Textura más fina que la arena margosa; escasa translocación de las arcillas; poco espesor; desarrollo moderado de los horizontes.
4	Aridisol	Latin aridus(arido)	Seco durante periodos largos; pobre en humus; rico en bases; suele tener horizontes de carbonato, yeso y/o arcilla.
5	Mollisol	Latin mollis	Horizontes superiores de color desde pardo oscuro hasta negro, con una consistencia blanda; rico en bases; en regiones semihúmedas.
6	Espodosol	Griego spodus (ceniciento)	Horizonte E gris pálido, blanquecino, sobre un horizonte B negro y rojizo, rico en hierro y aluminio que pueden extraerse.
7	Alfisol	A partir de Al y Fe	Humus a poca profundidad; translocación de la arcilla; horizontes bien desarrollados.
8	Ultisol	Latín ultimus (último)	Con lavado intenso; intensa translocación de arcilla; pobre en bases; en climas húmedos y cálidos.
9	Oxisol	Francés oxidé (oxidado)	Suelos con una meteorización intensa; rojos, amarillos o grises; ricos en caolinita, óxidos de hierro y usualmente también humus; en zonas tropicales y subtropicales.
10	Gelisol		Suelos afectados por los permafrost o heladas.
11	Andisol	Japones an (negro) y do (suelo)	Desarrollados a partir de materiales volcánicos; con meteorización débil; capas superiores de color oscuro; densidad reducida.
12	Histosol	Griego histos (orgánico)	Muy rico en materia orgánica; en zonas de turbera.

Tabla 3.4. Clasificación según la FAO [65].

<b>Símbolo</b>	<b>Grupo de suelo</b>	<b>Principales características</b>
<b>AC</b>	Acrisols	Argilica Bt de baja actividad arcillosa
<b>AN</b>	Andosols	Suelos sobre ceniza volcánica
<b>AL</b>	Alisols	Argillica Bt con arcillas de alta actividad y aluminio
<b>AT</b>	Anthrosols	Suelos afectados por cultivos humanos y / o perturbaciones
<b>AR</b>	Arenosols	Suelos de textura gruesa sin horizontes Marrón
<b>CM</b>	Cambisols	Tierra marrón con un horizonte Bw degradado
<b>CL</b>	Calcisols	Suelos con un horizonte cálcico o petrocálcico
<b>CH</b>	Chernozem	Suelos con un horizonte negro ah húmico
<b>FR</b>	Ferralsols	Los suelos tienen un horizonte B de sesquióxidos
<b>FL</b>	Fluvisols	Suelos crudos en aluviones recientes
<b>GL</b>	Gleysols	Suelos húmedos con horizontes descoloridos
<b>GR</b>	Greyzems	Suelos con horizontes húmicos A y argílicos Bt
<b>GY</b>	Gypsisols	Suelos con horizonte gípsico o petrogípsico
<b>HS</b>	Histosols	Suelos dominados por material orgánico (turba)
<b>KS</b>	Kastanozems	Suelos con horizontes húmicos A y cálcicos o gípsicos B
<b>LP</b>	Leptosols	Suelos poco profundos sobre roca dura
<b>LX</b>	Lixisols	Suelos con horizonte Bt argílico de baja actividad arcillosa
<b>LV</b>	Luvisols	Suelos con Bt argílico de actividad arcillosa media / alta
<b>NT</b>	Nitisols	Otros suelos con un profundo horizonte argílico Bt
<b>PH</b>	Phaeozems	Humic A sobre horizontes variados Pesado
<b>PL</b>	Planosols	Textura pesada en el horizonte B
<b>PT</b>	Plinthosols	Suelos con plintita que causan humedad en la superficie
<b>PZ</b>	Podzols	Suelos con horizonte b podzólico
<b>PD</b>	Podzoluvisols	Argillic Bt con parte superior irregular
<b>RG</b>	Regosols	Suelos de materiales no consolidados Enriquecidos con sal
<b>SC</b>	Solonetz	Suelos enriquecidos con sal
<b>SN</b>	Solonetz	Suelos con horizonte alcalino B
<b>VR</b>	Vertisols	Arcillas negras profundamente agrietadas

### 3.2.3 Relaciones aproximadas entre los mayores órdenes y grupos de suelos

Tabla 3.5. Comparativa de ambas clasificaciones.

Clasificación USDA	Clasificación FAO/UNESCO
Alfisol	Luvisol
Andisol	Andosol
Ardisol	Calcisol, gypsisol, solonchank, solonetz
Entisol	Arenosol, fluviosol, leptosol, regosol
Histosol	Histosol
Inceptisol	Cambisol
Mollisol	Chernozem, greyzems, kastanozems, phaenzems
Oxisol	Alisol, ferrasol, nitosol, plinthosol
Spodosol	Podzol
Ultisol	Acrisol, lixisol
Vertisol	Vertisol

### 3.3. Comportamiento espectral de los suelos y sus constituyentes

El comportamiento espectral de los suelos es la interacción física entre la materia y la energía, en este caso veremos 6 variaciones de la reflectancia espectral del suelo, por ejemplo, con el contenido de humedad, contenido de materia orgánica, distribución del tamaño de las partículas, contenido de óxido de hierro, mineralogía del suelo y estructura del suelo [73], a continuación detallaremos algunas de las variables más importantes .

#### ❖ Contenido de humedad

De las variables mencionadas en el párrafo anterior, el contenido de humedad es la más importante debido a su dinámica temporal y su gran impacto en la reflectancia de todo el suelo. Como se ve en la figura 3.5 (a), existe una relación inversamente proporcional entre el contenido de humedad edáfica y la reflectancia espectral del suelo, teniendo dos bandas de absorción de consideración en (1450 y 1920 nm). También es notable la fuerte absorción de la banda de hidroxilos a los 2200 nm, en la cual se muestran suelos ricos en arcillas.

### ❖ Contenido de materia orgánica

Son suelos con 20 % de contenido de carbón orgánico, a diferencia de los suelos minerales. Se muestra en la figura 3.5 (b), que la reflectancia en los suelos minerales con contenido orgánico disminuye a medida que aumenta el material orgánico y viceversa. La reflectancia de los suelos orgánicos, por otro lado, es controlada por el estado de descomposición del material en el suelo, ver fig. 3.6 (c) como la turba (materiales fibrosos) que está compuesta por restos de plantas con una descomposición mínima. Los materiales altamente descompuestos sapric (estiércol) usualmente son de color negro.

### ❖ Contenido de óxido de hierro

Las causas principales de la coloración rojiza de los suelos son debido al contenido del óxido de hierro ( $Fe_2O_3$ ). El contenido de óxido de hierro y materia orgánica, son los dos factores más importantes que afectan la reflectancia espectral de los suelos erosionados, especialmente en las bandas de 500 a 1200 nm, como se muestra en la figura 3.6 (d) [10]. Algunas pruebas muestran el incremento de la reflectancia en las longitudes de onda menores a los 1100 nm, al reducir el contenido de hierro químicamente.

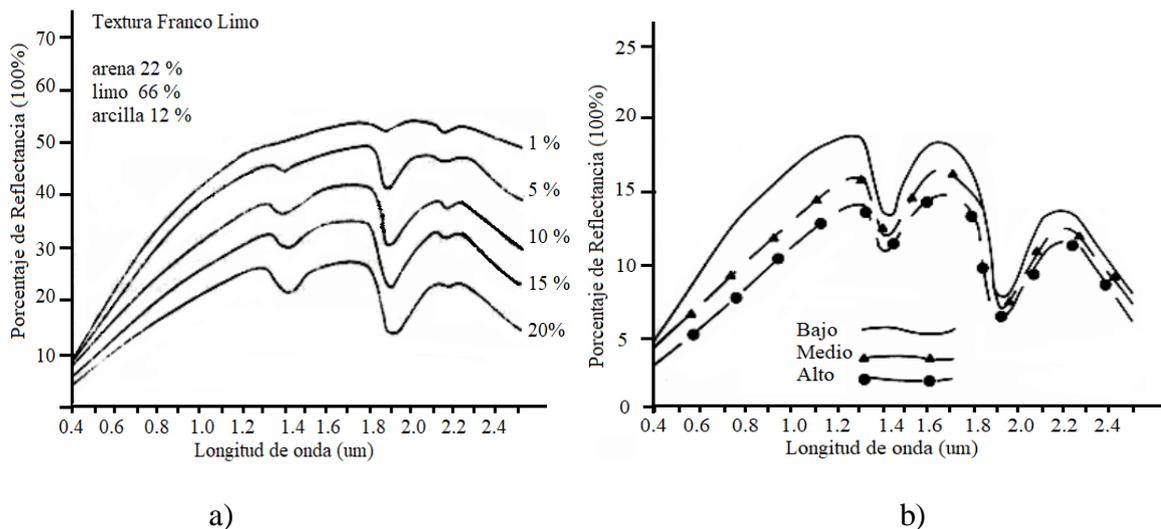


Figura. 3.5. Se muestra las variaciones de la reflectancia debido a diversos factores en el suelo [74]. a) Relación entre el contenido de humedad del suelo y la reflectancia espectral. b) Relación entre el contenido de materia orgánica de los suelos minerales y su reflectancia espectral.

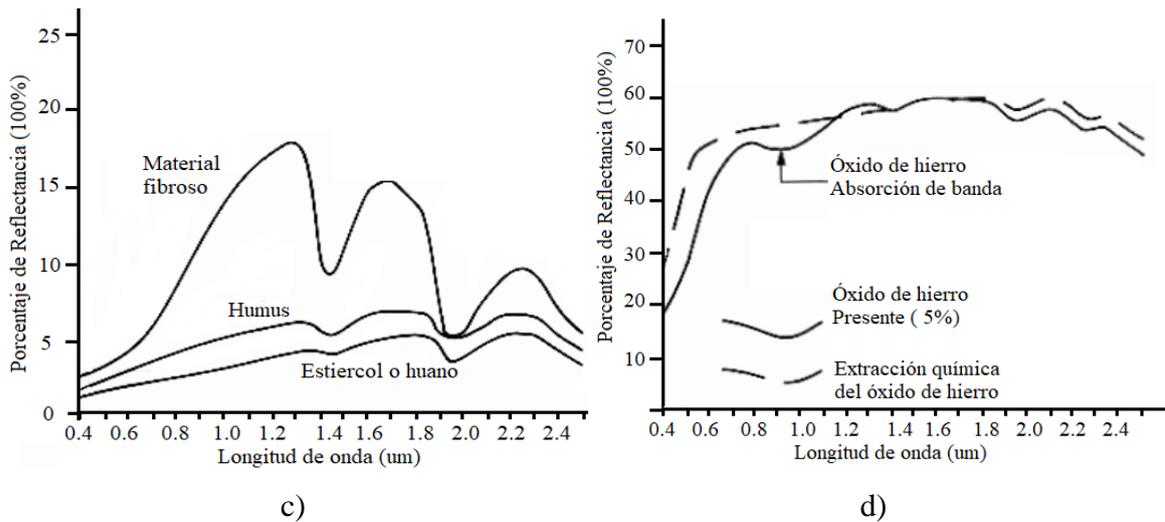


Figura. 3.6. Variaciones de la reflectancia debido al contenido orgánico y de óxido de hierro [73]. c) Relación entre el estado de descomposición de un suelo orgánico y su reflectancia espectral. d) Relación entre el contenido de óxido de hierro y la reflectancia espectral del suelo.

### 3.4. Suelos de Cusco

La región Kastanosólica, referida a los valles interandinos altos y zonas intermedias, está ubicado entre los 2200 y 4000 msnm y existen diversos tipos de suelos, principalmente los “kastanozems cálcicos”, de textura media, alcalinos y de color rojizo ó pardo rojizo. Los “kastanozems lúvicos”, similares pero arcillosos; así como suelos profundos y de textura fina (phaeozems). En las zonas de alta pendiente, predominan los suelos rocosos y calcáreos. En esta región se ubica el departamento del Cusco, a 3 399msnm, latitud sur 13°30'45”, con una superficie de 71 986 50 km<sup>2</sup> y una población de 1 205 527 habitantes. Las muestras de suelos para esta investigación fueron extraídas de la provincia de Anta, distrito de Ancahuasi a 3457 msnm, y coordenadas geográficas 13°27'21”S 72°18'02”O.

Esta región es un área agrícola tradicional, con uso intensivo de hace miles de años, cultivándose principalmente cereales, tubérculos, leguminosas y algunas hortalizas. Las partes altas son pastizales, usadas con fines pecuarios y las partes bajas usadas para cultivos permanentes como frutos y otros.

En el siguiente cuadro; Tabla 3.6, ubicaremos los suelos de Cusco en las dos principales clasificaciones:

Tabla 3.6. Relación aproximada de suelos del Cusco con las principales clasificaciones del Soil Survey Staff (USA) y de la FAO/UNESCO.

Ubicación de Suelo	Clasificación USDA	Clasificación FAO/UNESCO	Tipos de textura
Región del Cusco	Andisol	Andosol	Arena Limo y arcilla
	Entisol	Fluvisol, Gleysol, Leptosol y Regosol	
	Inceptisol	Cambisol	
	Mollisol	Kastanozems	
	Oxisol	Alisol	
	Ultisol	Acrisol y Lixisol	

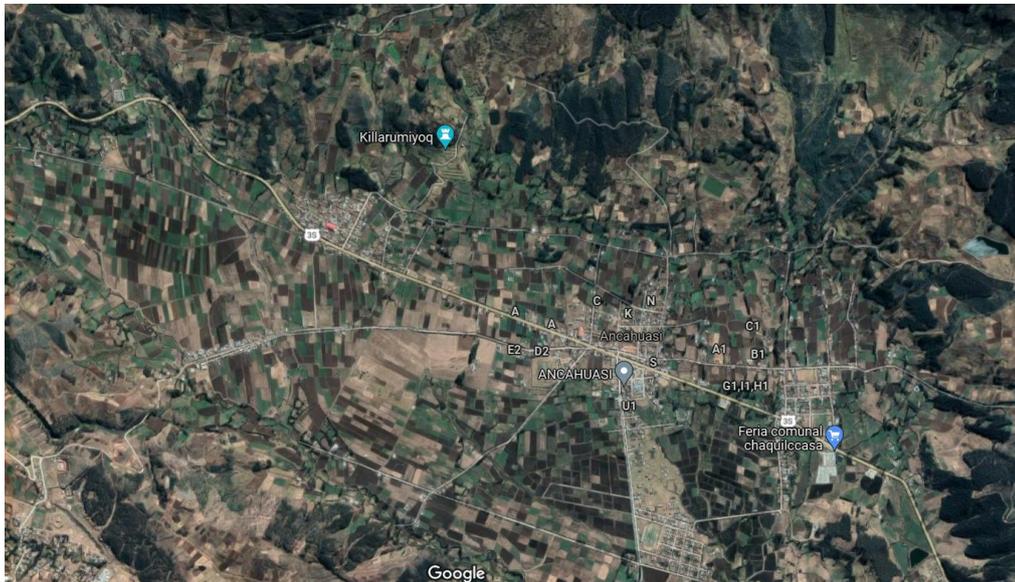


Figura 3.7. Parcelas de valle interandino en Cusco.

### 3.5. Suelos de Madre de Dios

Se ubican en la selva alta, entre los 500 y 2200 msnm, con un relieve escarpado, pero con ciertos valles. Los suelos provienen de la región lito-cambisólica, pero son más profundos. Es una zona con una fuerte meteorización o descomposición del material parental y de reacción ácida. Predominan los suelos profundos, de tonalidad amarilla y rojiza, ácidos y de buen drenaje, llamados “acrisoles” ó “rojo amarillo podsólicos”; arcillosos profundos. En la zona cercana a la selva baja se encuentran suelos arcillosos con hierro (acrisoles plínticos).

El departamento de Madre de Dios, está a 469 msnm, latitud oeste 11°59'S 70°35', con una superficie de 8530054 km<sup>2</sup> y una población de 141070 habitantes. Las muestras de suelos fueron extraídas de la provincia de Puerto Maldonado, distrito de Tambopata. Entre las actividades que desarrollan es el cultivo de arroz, caucho, maíz, yuca, plátano y coco. La madera es también fuente de ingresos, donde destacan la caoba y el cedro. Se practica la caza y la ganadería vacuna. Además, los lavaderos de oro en sus ríos le proporcionan importancia en la minería informal.

En el siguiente cuadro; Tabla 3.7, ubicaremos los suelos de Madre de Dios en las dos principales clasificaciones:

Tabla 3.7. Relación aproximada de suelos del Madre de Dios con las principales clasificaciones del Soil Survey Staff (USDA) y de la FAO/UNESCO.

Ubicación de Suelo	Clasificación USDA	Clasificación FAO/UNESCO	Tipos de textura
Región de Madre de Dios	Alfisol	Luvisol	Arena Limo y arcilla
	Entisol	Fluvisol, Gleysol, Leptosol y Regosol	
	Inceptisol	Cambisol	
	Oxisol	Alisol	
	Ultisol	Acrisol y Lixisol	



Figura 3.8. Suelos arcillosos rojizos en Madre de Dios [75].

## CAPITULO 4

### 4. ANÁLISIS Y PROCESAMIENTO DE UNA IMAGEN HIPERESPECTRAL (HSI)

#### 4.1. Datos de una imagen hiperespectral

En este apartado detallaremos la información espectral de cada pixel de la escena que se obtiene de una imagen hiperespectral. Dichos pixeles son utilizados para identificar materiales y/o propiedades en la escena. Nombraremos algunos componentes relevantes de una imagen hiperespectral, como:

- **Ruido:** Es un componente de error que se debe considerar al momento de procesar una imagen hiperespectral, ya que presenta bandas que contienen información errónea que no pueden ser explicadas como combinación de los propios materiales, en algunos casos esta constituido por sombras o puntos ciegos en superficies de relieve.
- **Relación Señal – Ruido (RSR):** Típicamente se da el valor máximo para alguna longitud de onda que depende del espectro utilizado para cuantificarlo y del tiempo de integración de la señal. Un valor máximo (“peak value”) no es suficiente para predecir el comportamiento del sensor.
- **Pixel Puro:** Es aquel que tiene solo un componente espectral, es decir, que en la representación de su firma espectral solo se ha identificado un elemento o componente en esa pequeña parte de la imagen.
- **Píxeles Mezcla:** Comprende una mezcla de componentes o de diferentes tipos de cubiertas de una superficie. Esta situación, a menudo denominada como “problema mixto” en terminología de análisis hiperespectrales, es una de las más cruciales y con propiedades de distinguir análisis de imágenes hiperespectrales. Los píxeles mixtos existen por varias razones. Primero, si la resolución espacial del sensor no es suficientemente alta para separar diferentes materiales, este puede ocupar conjuntamente un píxel sencillo, y la medida espectral resultante será un compuesto de los espectros individuales. Segundo, los píxeles mixtos pueden ser consecuencia también de la combinación de dos materiales distintos en una mezcla heterogénea.
- **Endmembers (miembros finales):** Son los componentes espectralmente puros de la imagen. Podemos decir que un Endmember es un pixel, cuyas características son las

mismas o similares a las tomadas en un laboratorio, sin ningún tipo de interferencias que altere o modifique el resultado final del material de estudio. En este punto llegamos a una conclusión sobre lo que es o no útil en una imagen hiperespectral, ya que al analizar los endmembers de una imagen nos damos cuenta de la veracidad de la misma y el aporte que ésta nos da al estudio, y lo que está con interferencias o lo que no está claro (ruido) se descarta [76]. Para conocer un endmember en una imagen primero hay que extraerlo y para ello existen diversas técnicas de identificación y software capaces de tratarlos.

En el siguiente gráfico, figura 4.1, se puede reconocer estos componentes presentes en una imagen hiperespectral.

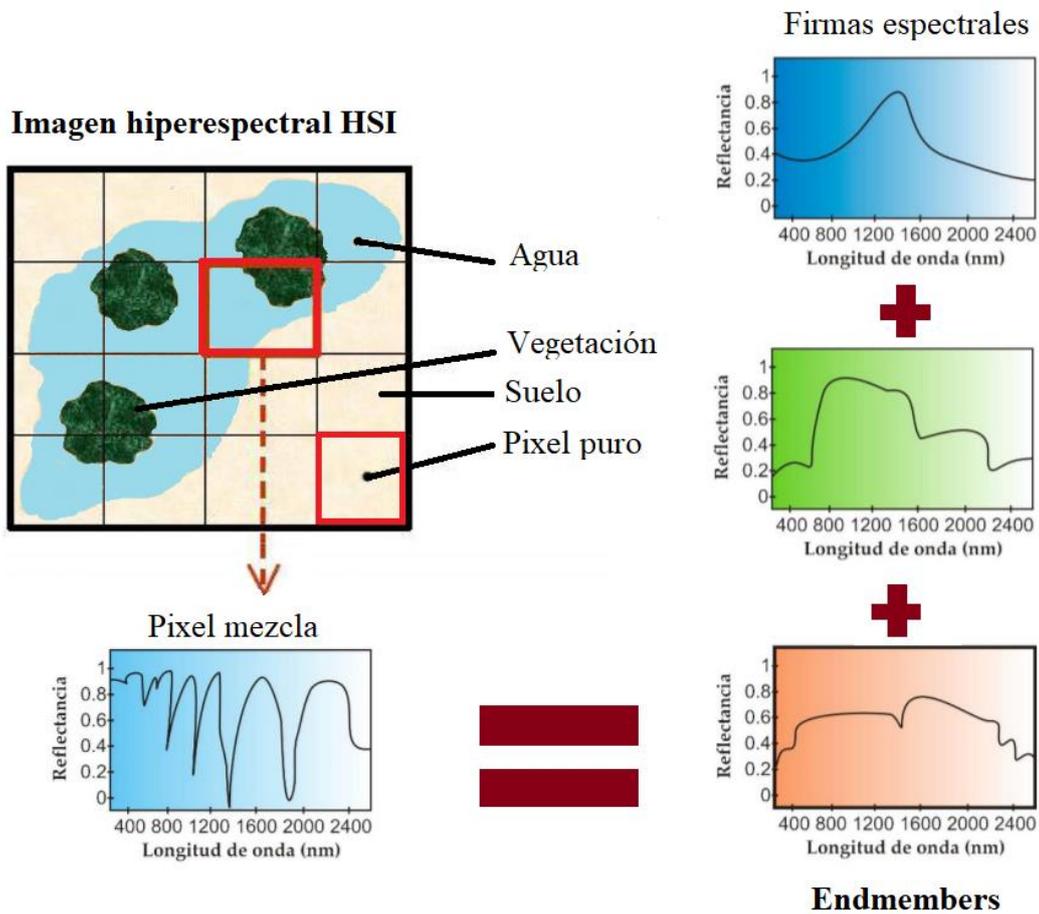
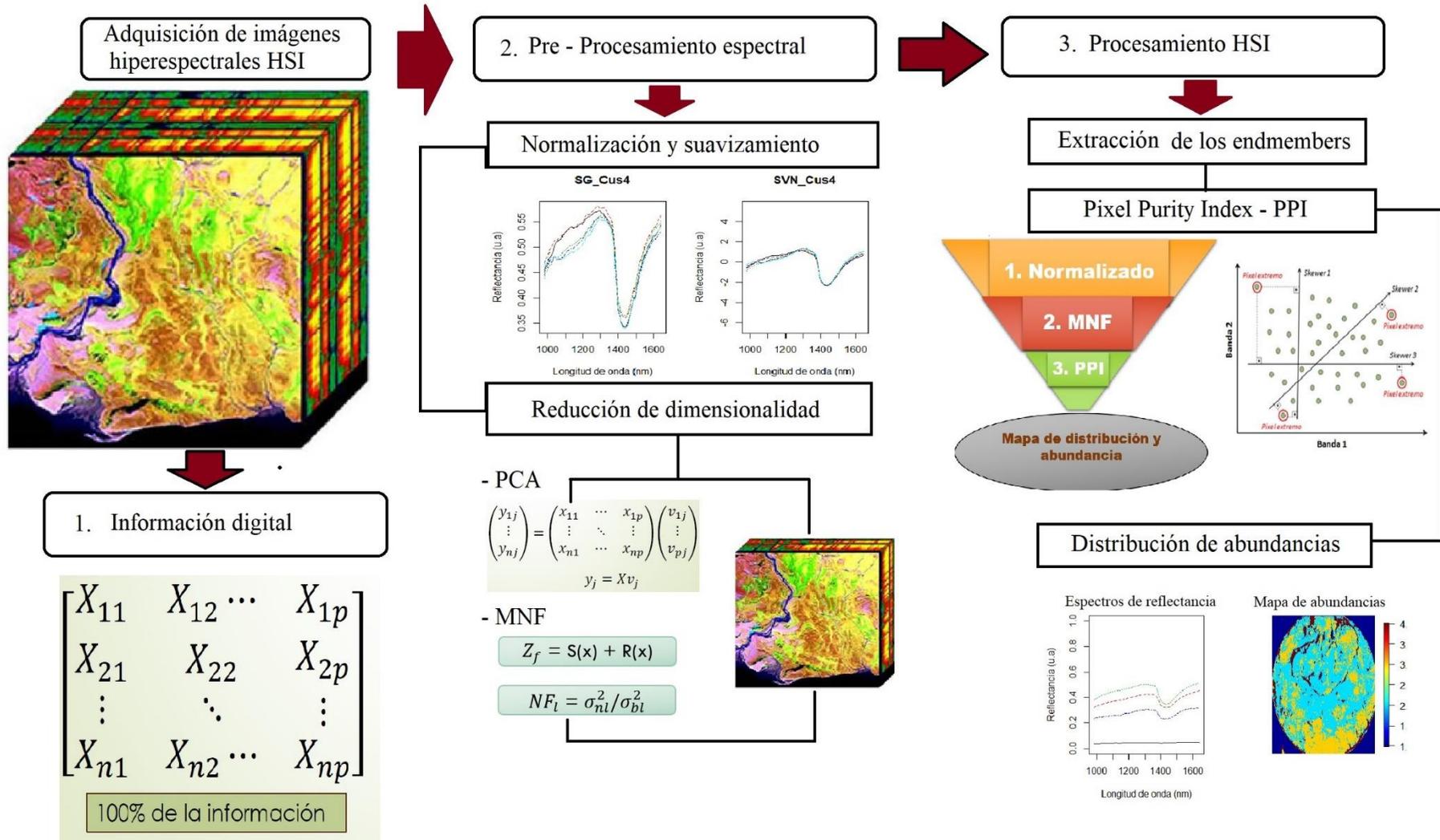


Figura 4.1. Reconocimiento y diferenciación de los pixeles puro y mezclas y su posterior desmezclado para encontrar los miembros finales (endmembers).

### 4.1.1. Diseño y cadena de estudio hiperespectral



## **4.2. Pre – procesamiento de una HSI**

Existen varios factores que alteran o producen ligeras variaciones espaciales durante la adquisición, uno de estos factores es la iluminación sobre el plano de la escena. Además la textura del suelo no es plana, su relieve es granular y esto genera sombras o puntos ciegos para el detector produciendo cambios en la reflectancia. Estos fenómenos fueron identificados como ruidos, en algunos casos como mezcla de píxeles de los bordes con el fondo. Para dar solución a estos problemas en primer lugar se ha abordado la corrección de la iluminación por medio de una calibración del sistema de adquisición tomando en cuenta un blanco de referencia [30], así mismo verificar que la plataforma genere la menor vibración posible. Sin embargo, no es suficiente, por ello se recurre al tratamiento previo de la información adquirida segmentando las imágenes y ubicando las regiones de interés (ROIs) a esto se le conoce como preprocesamiento espacial, inmediatamente se continúa con el preprocesamiento espectral utilizando técnicas de normalización, suavizamiento, corrección de escala multiplicativa y reduciendo la dimensionalidad.

A continuación, detallaremos algunas técnicas de preprocesamiento espectral:

### **4.2.1. Normalización**

La normalización de los valores de la imagen reduce el efecto de las variaciones en los valores de reflectancia absoluta y resalta la información de la forma espectral [77]. Existen varios métodos de normalización que tienen diferentes rendimientos al momento de particionar algoritmos de agrupamiento [78], por ejemplo K-means, es el algoritmo de agrupamiento más utilizado en teledetección que identifica grupos dentro de datos multidimensionales basados en algunas medidas de similitud (SM). Las SM asignan más pesos a las funciones con rangos grandes que aquellas con rangos pequeños. De esta manera, las características de rango pequeño son suprimidas por las características de rango grande para que no puedan tener ningún efecto durante el procedimiento de agrupamiento. Este problema se deteriora para los datos de alta dimensión, como las imágenes hiperespectrales de detección remota. Para abordar este problema, se recurre a la normalización de funciones (FN). En algunos casos los resultados de normalización muestran una mejora en los conjuntos de datos reales, y por otro lado se reduce la carga computacional [79]. A continuación, formularemos algunos métodos de normalización con el que se trabajará.

#### 4.2.1.1. Método basado en estiramiento (SBM).

- Mapeo entre [0,1]: Esta normalización se da cuando hay valores de límite inferior “min” y superior “max” para un componente de características  $x_i$ , esto se puede escribir de la siguiente manera:

$$x\eta_i = \frac{x_i - \min}{\max - \min} \quad (4.1)$$

Donde "i" es el número de componentes de la función y  $x\eta_i$  es la función normalizada de 0 a 1.

- Recorte: Es la normalización máxima donde el valor de cada componente de una banda es dividido por el máximo valor de dicha banda.

$$x\eta_i = \frac{x_i}{\max} \quad (4.2)$$

#### 4.2.1.2. Método de normalización estadística.

En este método de normalización, un componente de característica  $x_i$ , se transforma en una variable aleatoria con cero media, mediana o varianza de modo y unidad como:

$$x\eta_i = \frac{x_i - \mu}{\sigma} \quad (4.3)$$

Donde  $\mu$  es media, mediana o moda y  $\sigma$  es la desviación estándar muestral de esa característica.

#### 4.2.1.3. Normalización de rango.

Con la normalización de rango, cada componente de característica  $x_i$  en una matriz se reemplaza por su posición (rango) en la matriz ordenada contada desde el valor más pequeño dividido por el número total de componentes, pero si cada característica contiene un valor repetitivo, cada valor se divide por el máximo de la función. Denotamos  $r_i$  como el rango de  $x_i$  en la matriz a la que pertenece, las expresiones componentes normalizadas son:

$$x\eta_i = \frac{r_i}{m} \quad (4.4)$$

Donde “m” es el número total de componentes o el máximo de características.

#### 4.2.1.4. Normalización de banda máxima.

La normalización de banda se da cuando el valor máximo de todas las bandas de los píxeles  $\text{bandmax}(x_i)$  divide a cada valor de pixel  $x_i$  por banda [80].

$$x_i^* = \frac{x_i}{\text{bandmax}(x_i)} \quad (4.5)$$

Mostramos un resultado de normalización en el grafico siguiente.

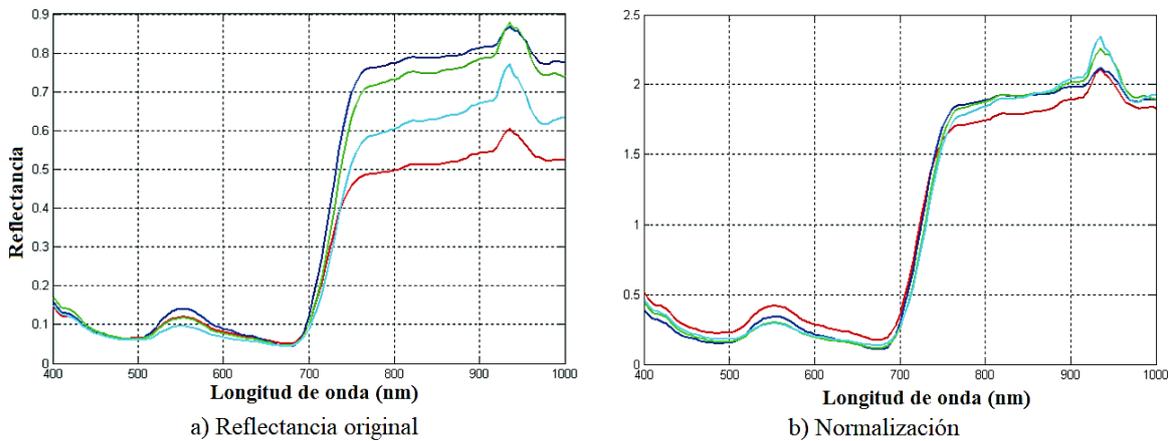


Figura 4.2. Curvas espectrales de muestras de cuatro especies diferentes de hierba [77].

#### 4.2.2. Suavizamiento espectral con Savitzky – Golay (SG)

El filtro Savitzky – Golay se usa ampliamente para suavizar o mejorar la relación señal/ruido y diferenciar series temporales, especialmente en datos biomédicos [81]. El filtro de suavizado de SG propone una convolución simplificada de mínimos cuadrados para suavizar y calcular las derivadas de un conjunto de valores consecutivos (un espectro). La convolución se entiende como un filtro de medida móvil ponderada con ponderación dada como un polinomio de cierto grado [82]. Los datos que proporcionan las imágenes hiperespectrales satisfacen las condiciones para realizar un suavizado con la ecuación 4.6, donde los coeficientes de peso aplicados a una señal, realizan un ajuste de mínimos cuadrados polinomiales dentro de la ventana de filtro. Este polinomio está diseñado para preservar los momentos más altos dentro de los datos y para reducir el sesgo introducido por el filtro. Este filtro se puede aplicar a cualquier dato consecutivo cuando los puntos de los datos están en un intervalo fijo y uniforme a lo largo de la abscisa elegida, y las curvas formadas al representar gráficamente los puntos deben ser continuas y más o menos suaves.

Entonces la ecuación general de la convolución de mínimos cuadrados simplificada para el suavizado de la información de las HSI está dada por:

$$Y_i^* = \frac{\sum_{i=-m}^{i=m} C_i Y_{j+i}}{N} \quad (4.6)$$

Donde Y es el valor original de entrada, Y\* es el valor resultante de suavizado,  $C_i$  es el coeficiente para el i-ésimo valor de entrada del filtro (ventana de suavizado) y N es el número de enteros convolutantes y es igual al tamaño de la ventana de suavizado ( $2m + 1$ ). El índice j es el índice de ejecución de la tabla de datos de ordenadas originales. La matriz de suavizado (tamaño del filtro) consta de  $2m + 1$  puntos, donde m es el ancho medio de la ventana de suavizado y por lo general, un valor mayor de m produce un resultado más suave a expensas de aislar picos afilados. Los coeficientes de un filtro Savitzky-Golay ( $C_i$ ) se pueden obtener directamente del artículo [83] como una versión corregida del trabajo de SG [84]. El grado del polinomio de suavizado generalmente se establece en un rango de 2 a 4, ya que un valor menor produce un resultado más suave, pero puede introducir parcialidad; un valor más alto reducirá el sesgo del filtro, pero puede “ajustar demasiado” los datos y dar resultados más ruidosos. La figura 4.3, muestra un suavizado del espectro de reflectancia de un material respecto a su longitud de onda.

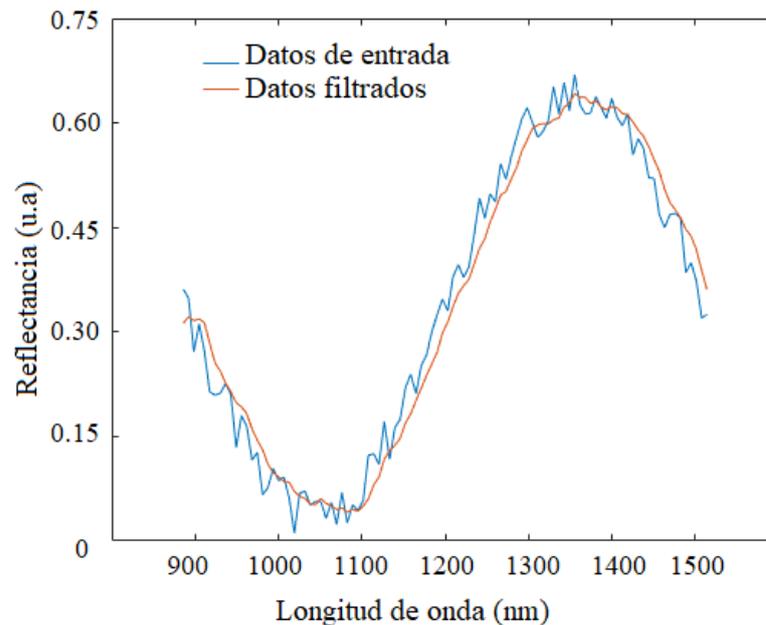


Figura 4.3. Corrección de espectros utilizando el filtro de suavizamiento de SG.

### 4.2.3. Corrección de escala multiplicativa con la *Standard Normal Variate* SNV

En la espectrometría de reflectancia difusa, los problemas habituales que se encuentran son el tamaño de partícula, la dispersión y la multicolinealidad y requieren un cierto cuidado al momento de analizar la data. En el estudio de suelos, el tamaño de partículas de la muestra representa la mayor parte de la varianza, mientras su composición química la menor varianza [85].

Para dar solución a estos problemas, se presenta las transformaciones matemáticas como: *Standard Normal Variate* SNV, *DT-trending* (DT), aplicados a los espectros de reflectancia difusa NIR individuales y a diferencia de otras correcciones no necesita un espectro de referencia porque estas necesitan un conjunto de calibración para calcular un espectro de referencia (media).

La transformación matemática de los espectros  $\log 1/R$  mediante el cálculo de la variación normal estándar en cada longitud de onda elimina la variación de la pendiente en una muestra individual mediante el uso del siguiente cálculo:

$$SNV: Y_{(1-700)}^* = \frac{(y_{(1-700)} - \bar{y})}{\sqrt{\frac{\sum (y_{(1-700)} - \bar{y})^2}{n - 1}}} \quad (4.7)$$

Donde SNV:  $Y_{(1-700)}^*$  son las variaciones normales estándar individuales para 700 longitudes de onda,  $y_{(1-700)}$  son los valores de  $\log 1/R$  vs. 700 longitudes de onda, y  $\bar{y}$  es la media de los valores de  $\log 1/R$  vs. 700 longitudes de onda. La figura 4.4, demuestra la corrección de los efectos de la dispersión de la luz y los del tamaño de las partículas de un fruto seco.

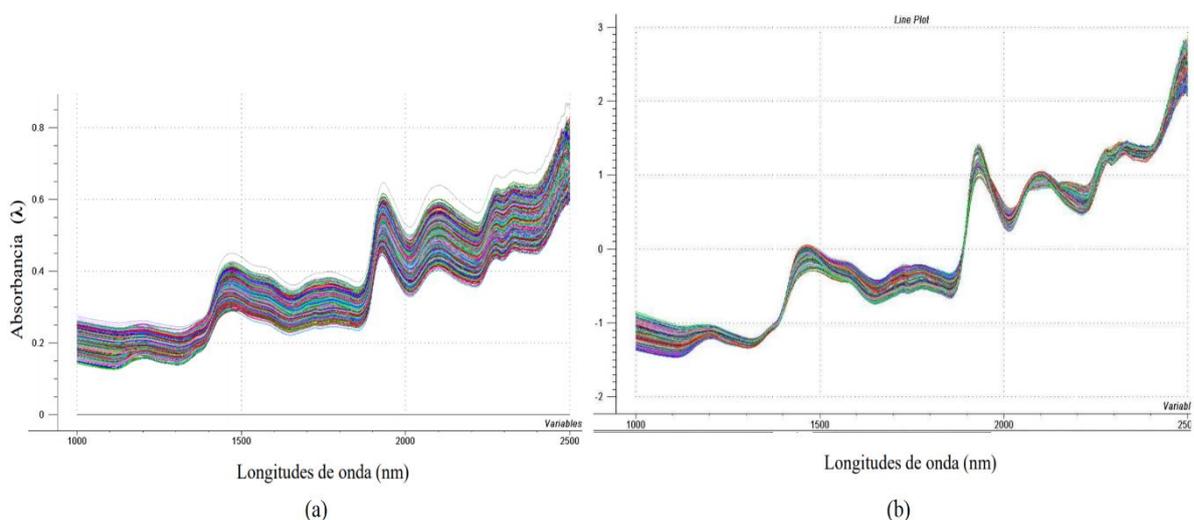


Figura 4.4. (a) Espectros originales de un fruto seco, (b) Espectros tratados con el SNV [85].

#### 4.2.4. Reducción de dimensionalidad de la imagen hiperespectral

La reducción de dimensionalidad (DR) es utilizado en la explotación de datos hiperespectrales para varios propósitos. Habitualmente es utilizado como una técnica de preprocesamiento que reduce un espacio de datos de muy alta dimensión a un espacio manejable, sin que este pierda la información relevante y donde el análisis de datos se realice de manera más efectiva. En particular, se utilizan dos enfoques comunes para DR, uno por transformación (DRT) y otro por selección de bandas (DRBS) [86]. El primer enfoque DRT, utiliza una transformación para compactar datos en un sentido óptimo, el segundo DRBS encuentra un subconjunto de bandas apropiado para representar los datos a través de un criterio óptimo. En este trabajo abarcaremos la reducción de dimensionalidad por transformación DRT, se desarrollan dos tipos; análisis de componentes (CA) y extracción de características (FE). Una transformación CA generalmente se considera como una transformación que utiliza estadísticas como criterio para descorrelacionar y convertir datos en un conjunto de componentes de datos no correlacionados para el análisis. Las transformaciones de este tipo incluyen dos transformaciones de componentes estadísticos de segundo orden de uso común, análisis de componentes principales (PCA) basado en varianza de datos y transformación de fracción de ruido máximo (MNF) basada en la relación señal / ruido (SNR), el CA basado en estadísticas de cuarto orden, más un análisis de componentes independientes (ICA) basado en independencia estadística que requiere un orden infinito de estadísticas. Una transformación FE utiliza un criterio basado en la extracción de características para producir un conjunto de vectores de características para que los datos puedan ser representados por estos vectores de características.

##### 4.2.4.1. Análisis de Componentes Principales – PCA

El análisis de componentes principales (PCA) o transformación de componentes principales, es una transformación óptima para representar datos en el sentido de varianza de los datos [33]. El PCA es un método estadístico que permite reducir la complejidad de espacios muestrales con muchas dimensiones conservando su información. Para una muestra con  $n$  elementos (píxeles) cada uno con  $p$  variables (longitud de onda o bandas) espacio muestral  $p$  – dimensional, PCA permite reducir las dimensiones a  $z$  ( $z < p$ ) reproduciendo aproximadamente lo mismo que las  $p$  variables originales. Donde antes se necesitaban  $n$  valores para caracterizar a cada muestra, ahora bastan  $z$  valores. Cada una de estas  $z$  nuevas

variables recibe el nombre de componentes principal. PCA pertenece a la familia de técnicas conocida como unsupervised learning.

Cada componente principal ( $z_i$ ) se obtiene por combinación lineal de las variables originales, buscando maximizar la varianza para la componente  $z_i$ :

$$z_i = \phi_{11}x_1 + \phi_{21}x_2 + \dots + \phi_{p1}x_p \quad (4.8)$$

los términos  $\phi_{11}, \dots, \phi_{p1}$  reciben el nombre de loadings y son los que definen a la componente principal. Los loadings pueden interpretarse como el peso/importancia que tienen cada variable en cada componente y, por lo tanto, ayudan a conocer que tipo de información recoge cada una de las componentes.

En la figura 4.5, se ilustra el proceso de PCA mediante un ejemplo gráfico.

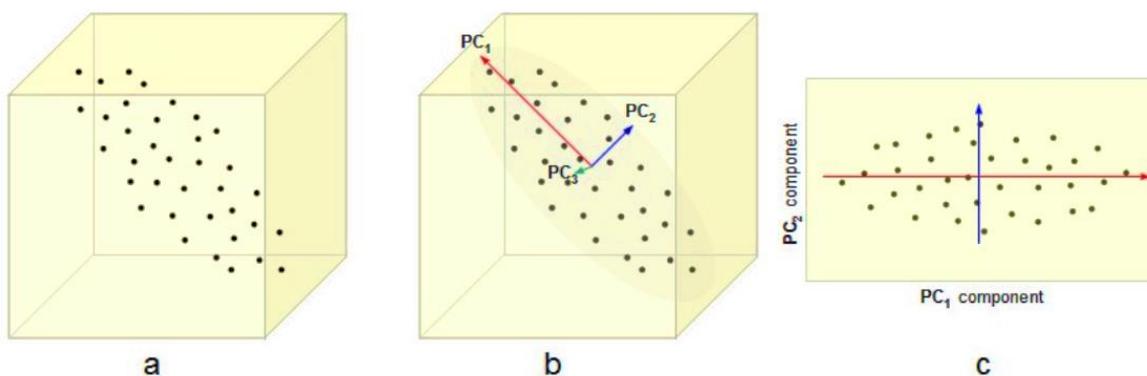


Figura 4.5. Gráfico del PCA [87]. a) Datos iniciales, representados como puntos tridimensionales, b) Componentes principales (PCi) de los datos, ordenados de manera descendentes según varianza, c) Proyección de los datos originales sobre PC1 y PC2, descartando PC3.

De esta manera, analizando el gráfico anterior, podemos destacar dos aspectos resaltantes en el PCA.

- Reducción de dimensionalidad: Es notorio la disminución de información, esto permite trabajar con un tamaño de datos inferior, pero conservando sus características principales.
- Incorrelación: Indica que no existe redundancia en la información que se transporta. Por lo tanto, las componentes principales representan los patrones de variación del conjunto inicial de datos.

El proceso que se sigue para encontrar las componentes principales es:

1.- Centrado de las variables: se resta a cada valor la media de la variable a la que pertenece. Con esto se consigue que todas las variables tengan media cero.

2.- Se encuentran los loadings de  $(z_i)$  diagonalizando la matriz de covarianza.

3.- Una vez hallada la primera componente  $z_1$  se calcula la segunda  $z_2$  repitiendo el mismo proceso, pero añadiendo la condición de que la combinación lineal no puede estar correlacionada con la primera componente. Esto equivale a decir que  $z_1$  y  $z_2$  tienen que ser perpendiculares. El proceso se repite de forma iterativa hasta calcular todos los posibles autovectores. El orden de importancia de las componentes se relaciona con la magnitud de los autovalores.

4.- Finalmente si se desea obtener una imagen remarcando los píxeles donde prevalece la primera componente principal se proyecta el vector intensidad:

$$I = a_1x_1 + a_2x_2 + a_3x_3 + \dots + a_px_p = a'_i x \text{ de cada píxel sobre el vector } z_1.$$

Donde:  $a'_i = (a_1, a_2, \dots, a_i)$  representa a los eigenvectores.

Las ventajas y desventajas del PCA se resumen en los siguientes puntos:

- Muestra la relación existente entre los datos de entrada con la información conseguida, identificando el comportamiento general de los PCA's
- Reducción del tamaño de datos, lo que implica menor coste computacional.
- Su proceso está limitado al uso estadístico de segundo orden, dejando excluidos a estadísticos de orden superior que podrían aportar información útil adicional sobre la estructura interna de los datos [87].
- No es posible determinar mediante PCA unas componentes principales que sean independientes, únicamente se puede afirmar que no existe correlación alguna entre ellas.

#### **4.2.4.2. Transformación Mínima Fracción de Ruido – MNF**

Una transformación por componentes principal (PCA) no siempre produce imágenes que muestran una cantidad de imagen sostenidamente decreciente al incrementarse el número de la componente. lo que se busca entonces es una transformación lineal que haga esta tarea. Una de las formas más comunes de medir la calidad de una imagen es la relación ruido-síñal. Así en vez de escoger nuevas componentes para maximizar la varianza, como en la transformación por componentes principales, ahora se escogen de tal manera que se

maximice la relación ruido-síñal, logrando de este modo el deseado ordenamiento óptimo en términos de la calidad de imagen [88].

Para un conjunto de datos multivariado de  $p$  –bandas con niveles de intensidad  $Z_1(x)$ , con  $i = 1, \dots, p$  donde  $x$  da las coordenadas de la muestra. Asumimos que:

$$Z(x) = S(x) + N(x) \quad (4.9)$$

donde  $Z^T(x) = Z_1(x), \dots, Z_p(x)$  y  $S(x), N(x)$  son las componentes de señal y ruido no correlacionados.

Luego:

$$Cov(Z(x)) = \Sigma = \Sigma_S + \Sigma_N \quad (4.10)$$

donde:  $\Sigma_S$  y  $\Sigma_N$  son las matrices de covarianza de  $S(x)$  y  $N(x)$  respectivamente.

Se define la fracción de ruido de la  $i$ -ésima banda como la razón de la varianza total de la banda, a la varianza del ruido, es decir como:

$$\frac{Var(Z_i(x))}{Var(N_i(x))} \quad (4.11)$$

Los autovectores obtenidos usando la fracción de ruido son los mismos que los obtenidos por el procedimiento de minimizar la razón señal-ruido.

Sea  $Y_i(x) = a_i^T Z(x)$ ,  $i = 1, \dots, p$  la transformación de fracción de ruido (MNF) con  $Y_i$  máxima entre toda las transformaciones ortogonales a  $Y_j, j = 1, \dots, i$ .

Los  $a_i$  son los autovectores de  $\Sigma_N \Sigma^{-1}$  y si  $\mu_i$  son los autovalores correspondientes a  $a_i$ , es igual a la fracción de ruido en  $Y_i(x)$ . Por lo tanto, a partir de la definición de la transformada MNF, vemos que  $\mu_1 \geq \mu_2 \geq \dots \geq \mu_p$ , por lo que los componentes MNF mostrarán una calidad de imagen en constante aumento (a diferencia del orden habitual de componentes principales).

Con fines prácticos se asumirá que los autovectores  $a_i$ , están normalizados de manera que:

$$a_i^T \Sigma a_i = 1, \quad i = 1, \dots, p. \quad (4.12)$$

Esta norma es solo por conveniencia matemática y no tiene ningún efecto sobre los procedimientos de eliminación de ruido, entonces podemos expresar la transformada MNF en forma matricial.

$$Y_i(x) = A^T Z(x) \quad (4.13)$$

Donde:  $Y^T(x) = (Y_1(x), \dots, Y_p(x))$  y  $A = (a_1, \dots, a_p)$

Generalmente,  $\Sigma$  se estima utilizando la matriz de covarianza muestral de  $\mathbf{Z}(x)$  mientras para estimar  $\Sigma_N$  se usa el método MAF (mínimum/máximum autocorrelation factors). También se puede estimar  $\Sigma_N$ , primero identificando las fuentes de ruido, que puede ser desde un fondo conocido hasta las sombras presentes en las muestras estudiadas y relacionando el ruido con datos evaluados previamente obtenidos a partir de las mismas imágenes estudiadas.

### 4.3. Procesamiento hiperespectral

En la actualidad, el procesamiento hiperespectral se refiere al uso de varios modelos matemáticos o algoritmos que extraen, manipulan y transforman información. Son modelos esquemáticos que permite el desarrollo de varias técnicas para el procesamiento, corrección y detección de endmembers o simplemente para el proceso de desmezclado de una imagen hiperespectral.

#### 4.3.1. Clasificación y detección de objetos

La finalidad de estas técnicas es buscar un material u objeto específico o de interés, en algunos casos no se necesita la aplicación de estas técnicas ya que basta con diferenciar las clases de material al inicio.

La información hiperespectral está conformada por un conjunto de firmas espectrales (puras o no) que representan los diferentes objetos que forman parte de la imagen procesada, los cuales tienen características espectrales diferentes entre sí.

Para la detección de las firmas espectrales de los píxeles de una imagen de manera automática, se recurre a dos tipos de detección o clasificación:

- **Clasificación Supervisada:** Es uno de los métodos asociados con el aprendizaje automático que implica la asignación de datos etiquetados para que se pueda deducir un determinado patrón o función a partir de estos datos. Vale la pena señalar que el aprendizaje supervisado implica la asignación de un objeto de entrada, un vector, mientras que, al mismo tiempo, anticipa el valor de salida más deseado, que se conoce principalmente como la señal de supervisión. La propiedad final del aprendizaje supervisado es que los datos de entrada son conocidos y etiquetados adecuadamente.

- Clasificación No supervisada:** Es el segundo método de algoritmo de aprendizaje automático donde las inferencias se obtienen de datos de entrada sin etiqueta. El objetivo del aprendizaje no supervisado es determinar los patrones ocultos o la agrupación en datos de datos no etiquetados. Se usa principalmente en el análisis de datos exploratorios. Uno de los caracteres definitorios del aprendizaje no supervisado es que tanto la entrada como la salida no se conocen.

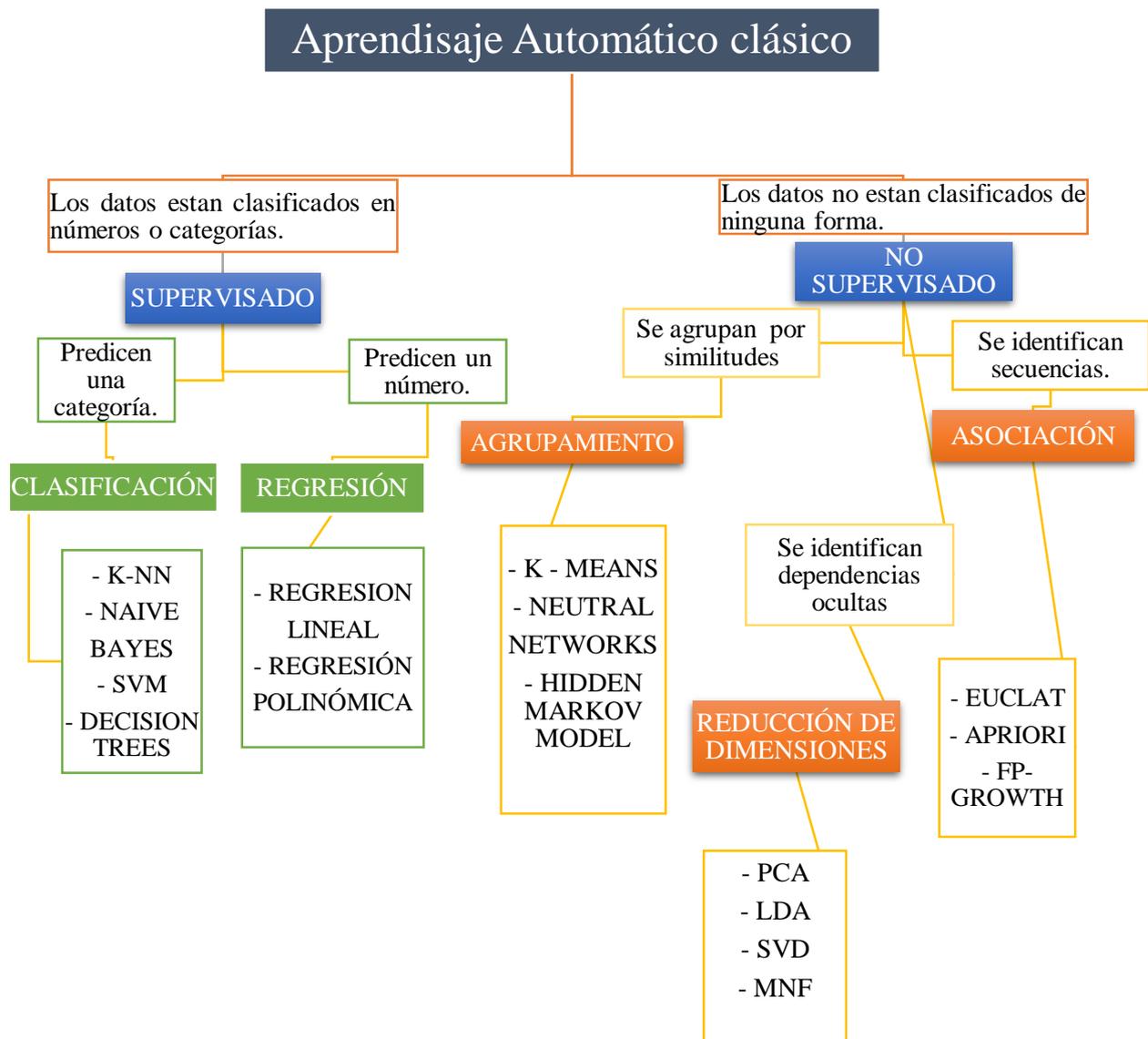


Figura 4.6. Técnicas de aprendizaje automático, clasificación supervisada y no supervisada.

### 4.3.2. Desmezclado espectral

Las imágenes en general, están conformadas por diversos elementos o componentes, las imágenes hiperespectrales que procesaremos no escapa de esta realidad, por ello recurriremos a diversas técnicas de desmezclado espectral. Un modelo básico es el modelo lineal de mezcla, que expresa los píxeles mezcla como una combinación lineal de las firmas relacionadas a componentes espectrales puros llamados endmembers [89]. En una primera aproximación, las intensidades  $X(x_{i\lambda})$  medidas en cada banda espectral  $\lambda = 1, \dots, L$  para un píxel dato  $i = 1, \dots, N$ , son combinaciones lineales de las características de reflectancia  $S(s_{m\lambda})$  de los materiales  $m = 1, \dots, M$  presentes en el área. Los posibles errores de esta aproximación y el ruido del sensor se tienen agregando un término de ruido  $N(n_{i\lambda})$ . En forma de matriz, esto se puede escribir como:

$$X = AS + N, \text{ teniendo en cuenta que: } A \geq 0$$

Donde la matriz  $A(a_{im})$  representa la abundancia de material  $m$  en el área correspondiente al píxel  $i$ , con restricciones de positividad y normalización.

El desmezclado espectral o análisis de mezclas espectrales son técnicas más utilizadas en el procesamiento de imágenes hiperespectrales. Estas técnicas o métodos comprenden dos etapas:

- Identificación de Endmember
- Estimación de abundancias

El primer proceso consiste en la extracción de las distintas firmas espectrales para identificar los componentes puros de los píxeles mezcla.

En la estimación de las abundancias se toma en cuenta para cada píxel de la imagen el grado de coincidencia con las firmas puras o endmembers extraídos.

## 4.4. Técnicas de identificación de endmembers

### 4.4.1 Algoritmo Pixel Purity Index (PPI)

Describiremos en esta sección el algoritmo PPI, el cual ha sido utilizado ampliamente en el análisis de imágenes hiperespectrales para la extracción de miembros finales o endmembers. Este algoritmo se encuentra presente en el software comercial Environment for Visualizing Images (ENVI) de ITT Visual Information Solutions (ITTVIS) desarrollado originalmente por Analytical Imaging and Geophysics (AIGSS) [22]. Debido a los derechos de propiedad y a los limitados resultados publicados, los detalles de su implementación hasta ahora son desconocidos. Por ello se recurre al software ENVI o se implementa versiones propias basados en la información disponible en la literatura.

El procedimiento general del algoritmo PPI, implementado por ENVI sugiere una metodología para la detección de los endmembers utilizando el flujo de procesamiento de reloj de arena, que consiste en las siguientes etapas: reducción de la dimensión espectral con (MNF), reducción de la dimensión espacial (índice de pureza del píxel - PPI), identificación (utilizando un visualizador n-D) y un mapa de distribución y abundancia [90], ver figura 4.7.



Figura 4.7. Flujo de procesamiento de reloj de arena para la detección de los miembros finales.

Un procedimiento para hallar los endmembers o elementos puros de las imágenes se detalla en [86]. La figura 4.8 muestra el proceso de identificación de los elementos extremos de un conjunto de píxeles en algunas direcciones denominadas skewers.

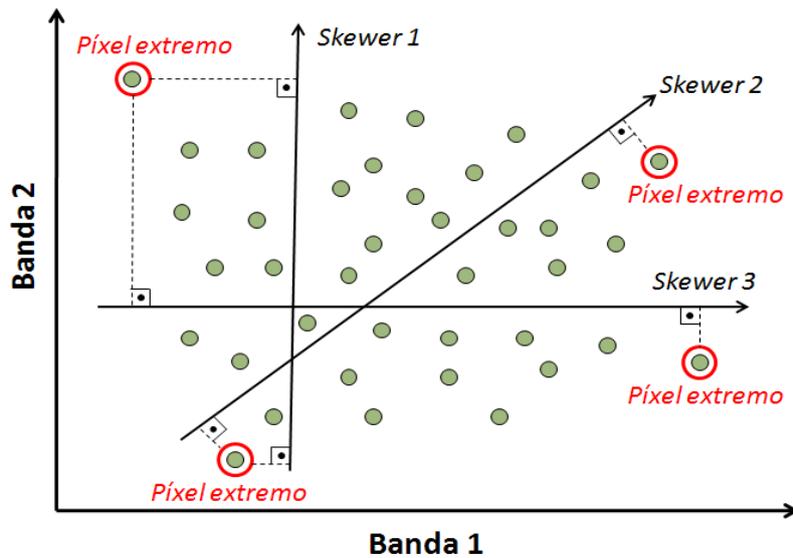


Figura 4.8 Ejemplo básico que ilustra el funcionamiento del algoritmo de extracción de endmembers PPI en un espacio de 2 dimensiones [52].

#### 4.4.1.1. Descripción del algoritmo

Debido a que los detalles mínimos de la secuencia de pasos específicos para implementar el algoritmo PPI del software ENVI son desconocidos, el algoritmo PPI descrito a continuación se basa en los limitados resultados publicados en [86][52] y tomando una propia interpretación. Sin embargo, excepto un paso final de supervisión manual (incluido en el PPI de ENVI) que es reemplazado por el paso 4, tanto la aproximación como el PPI en ENVI 4.0 producen resultados muy similares. Las entradas al algoritmo son un cubo de datos hiperespectrales  $F$  con  $N$  dimensiones; el número de skewers al azar que se generarán durante el proceso,  $K$ ; y un valor umbral de corte,  $t_v$ , utilizado para seleccionar como miembros finales solo aquellos píxeles que se han seleccionado como píxeles extremos al menos  $t_v$  veces durante todo el proceso de PPI [91].

El algoritmo consta de los siguientes pasos:

1. Generación de skewers. Produce un conjunto de  $K$  vectores unitarios generados aleatoriamente  $\{skewer_j\}_{j=1}^K$ .
2. Cálculo de proyecciones extremas. Para cada  $skewer_j$ ,  $j = \{1, 2, \dots, K\}$ , todos los píxeles  $f_i$  del conjunto original de datos  $F$  se proyectan en el  $skewer_j$  mediante un dot-products (productos escalares) de  $|f_i \cdot skewer_j|$  para encontrar los píxeles de la muestra que son

extremos (máximo y mínimo), dando como resultado un conjunto de extremos para el  $skewer_j$  que es representado por  $S_{extremos}(skewer_j)$ . Debido a la probabilidad de que exista algunos pixeles de la muestra que aparezcan en más de un conjunto de extremos, se debe definir la función  $I_S(x)$ , que indica la pertenencia de un elemento  $x$  a ese conjunto particular de la siguiente manera:

$$I_S(x) = \begin{cases} 1 & \text{si } x \in S \\ 0 & \text{si } x \notin S \end{cases} \quad (4.14)$$

3. Cálculo de la puntuación PPI. Usando la función de indicador anterior, calculamos la puntuación PPI asociada a cada píxel  $f_i$  (es decir, el número de veces que un píxel dado ha sido seleccionado como extremo en el paso 2) usando la siguiente ecuación:

$$N_{PPI}(f_i) = \sum_{j=1}^k I_{S_{extremos}(skewer_j)}(f_i) \quad (4.15)$$

4. Selección de endmembers (miembros finales). Se encuentra los pixeles que tengan una puntuación  $N_{PPI}(f_i)$  que esté por encima de  $t_v$  y se etiqueta como endmembers.

Como una desventaja se observó que el algoritmo PPI en la parte 2 (cálculo de proyecciones) consume más tiempo. Afortunadamente, el algoritmo PPI es muy adecuado para ser paralelizado. Los cálculos de las proyecciones de los píxeles en los skewers son independientes y pueden ser ejecutados simultáneamente, dando lugar a diversas formas de paralelización.

## 4.5. Otra técnica de clasificación

### 4.5.1 K – MEANS

K-Means (traducido como K-Medias en español), es un método de agrupamiento o clustering no supervisado. La idea principal del algoritmo se basa en la selección de un centroide para cada Cluster del conjunto de datos dado. Esta selección es inicialmente aleatoria. Para saber si los datos son parecidos o diferentes el algoritmo K-means utiliza la distancia entre los datos. Las observaciones que se parecen tendrán una menor distancia entre ellas[92].

En general, como medida se utiliza la distancia euclidiana, aunque también se pueden utilizar otras funciones. Una vez agrupados los puntos se calcula la media en cada grupo los cuales definen los nuevos centroides. Este proceso se repite hasta que no hay cambios en los

centroides. Este algoritmo tiene como objetivo minimizar el error cuadrático medio, con vectores reales de  $d$  dimensiones  $(x_1, x_2, \dots, x_j)$ , dentro de cada grupo  $S = \{S_1, S_2, \dots, S_k\}$ , a su centroide. Entonces, se puede formular de la siguiente forma:

$$\min_S E(U_i) = \min_S \sum_{i=1}^k \sum_{x_j \in S_i} \|x_j - \mu_i\|^2 \quad (4.16)$$

Donde  $S$  es el conjunto de datos cuyos elementos son los objetos  $x_j$  representados por vectores donde cada uno de sus elementos representa una característica o atributo. Tendremos  $k$  grupos o clusters con su correspondiente centroide  $\mu_i$ .

Así, los elementos que comparten características semejantes estarán juntos en un mismo grupo, separados de los otros grupos con los que no comparten características.

El algoritmo consta de tres pasos:

1. Inicialización: una vez escogido el número de grupos,  $k$ , se establecen  $k$  centroides en el espacio de los datos, por ejemplo, escogiéndolos aleatoriamente.
2. Asignación objetos a los centroides: cada objeto de los datos es asignado a su centroide más cercano.
3. Actualización centroides: se actualiza la posición del centroide de cada grupo tomando como nuevo centroide la posición del promedio de los objetos pertenecientes a dicho grupo.

## CAPITULO 5

### 5. RESULTADOS

Los resultados obtenidos en esta investigación siguen la cadena de procesado que se encuentra en el apartado (4.1.1).

#### 5.1. Proceso de preparación de muestras de suelo

En principio, solo se analizó siete muestras de suelo, las más representativas y predominantes de una zona. Cuatro clases del departamento de Cusco y tres clases del departamento de Madre de Dios.

Para una mejor identificación de las muestras estudiadas, se implementó una etiqueta para nombrar a las muestras de suelo, se asignará de la siguiente manera: Cus1, Cus2, Cus3 y Cus4 a los suelos de Cusco; MD1, MD2 y MD3 a los suelos de Madre de Dios.

Los suelos han sido tomados según la clasificación de regiones geodáficas identificadas por la FAO [65], la cual también es usada en el Perú por el Minagri (Ministerio de Agricultura) [93], La tabla 5.1 muestra las clasificaciones de las muestras de suelo estudiadas.

**Tabla 5.1.** Cantidad y descripción de muestras a tratar.

Nº MUESTRA	Tipo de suelo	Característica	Peso (kg)
<b>Cus1</b>	Mollisol-Kastanozems (cálcicos)	Suelos limosos calcáreos de color pardo rojizo.	3kg
<b>Cus2</b>	Mollisol-Kastanozems (húmicos)	Suelos limosos ricos en materia orgánica y ácidos de color negro humo.	3kg
<b>Cus3</b>	Oxisol-Alisol (argilica)	Suelos arcillosos de color rojo pardo.	3kg
<b>Cus4</b>	Ultisol-Acrisol (argilica)	Suelos arcillosos de color negro oscuro.	3kg
<b>MD1</b>	Inceptisol-Cambisol	Suelos de textura más fina que la arena con abundante material orgánico de color grisáceo.	2kg
<b>MD2</b>	Entisol-Arenosol (fluviosol)	Suelos muy infértiles, con textura arenosa color blanco arena.	2kg
<b>MD3</b>	Ultisol-Acrisol (plínticos)	Suelos arcillosos podsólicos de color rojo amarillo.	2kg

A continuación, se visualizará las imágenes de los suelos en su estado natural después de haberla clasificado.

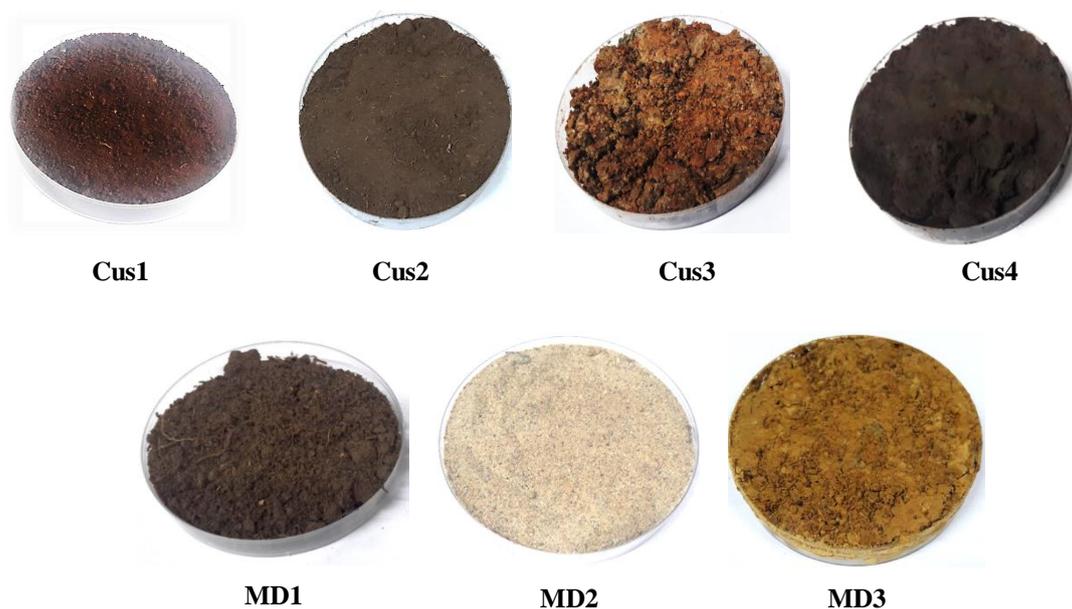


Figura 5.1. Muestras Cus1, Cus2, Cus3, Cus4 de suelos de Cusco y muestras MD1, MD2, MD3 de Madre de Dios.

### 5.1.1 Secado de muestra

Se toma muestras representativas de suelo, y a continuación se coloca la muestra húmeda en una capsula de cerámica previamente tarado. Así como se muestra en la siguiente tabla 5.2.

**Tabla 5.2.** Peso de cada muestra húmeda (MH), sin y con recipiente (S/R y C/R).

TIPO DE SUELO		PESO:	
		MUESTRA NATURAL(gr)	
<b>Cus1</b>	Pardo rojizo (75ml)	S/R	60.077
		C/R	102.238
<b>Cus2</b>	Negro humo (60ml)	S/R	60.108
		C/R	99.917
<b>Cus3</b>	Rojo pardo (75ml)	S/R	30.001
		C/R	69.997
<b>Cus4</b>	Negro oscuro (60ml)	S/R	60.01
		C/R	102.212
<b>MD1</b>	Grisáceo (60ml)	S/R	25.254
		C/R	66.310
<b>MD2</b>	Blanco arena (60ml)	S/R	60.096
		C/R	101.303
<b>MD3</b>	Rojo Amarillo (200ml)	S/R	100.091
		C/R	188.002

Luego, se coloca el conjunto de muestras en el horno por 24 horas, a una temperatura de  $110^{\circ} \pm 5^{\circ}\text{C}$ . Trascurrido dicho tiempo se obtiene lo siguiente:



Figura 5.2. Puesta de material en el horno por 24 horas, a  $110^{\circ} \pm 5^{\circ}\text{C}$ .

### 5.1.2 Empastillado de muestra

➤ **Objetivos:**

- ✓ Generar una muestra homogénea, y superficie uniforme.
- ✓ Identificar el valor de la relación de la reflectancia que se asigna a un pixel en una imagen hiperespectral.

➤ **Materiales:**

- ✓ Matriz de acero (base mayor con diámetro hueco de 4cm, diámetro mayor 6cm y 10cm de altura, eje de 12cm y base menor de 2cm. Peso total de la matriz 1.5kg.)
- ✓ Prensa de 100MP
- ✓ Balanza de precisión
- ✓ Mortero



Matriz de acero

Prensa en Mega Pascal (MP)

B. de precisión  
 $\Delta = 0.01$

Mortero

Figura 5.3. Materiales usados para el empastillamiento de muestras.

### 5.1.3 Procedimiento y resultados:

Una vez obtenido las muestras secadas, se prosede a triturarlas con el mortero, reduciendo el tamaño de grano. En la figura 5.4. observamos los resultados del secado de las muestras



Figura 5.4. Muestras secadas durante 24 horas a una temperatura de  $110^{\circ} \pm 5^{\circ}C$ .

Se hace el pesado respectivo de los 7 materiales, después de varios ensayos, se encontró una presión óptima para cada muestra, también un peso adecuado para obtener la pastilla con una buena consistencia.

**Tabla 5.3.** Resumen de resultados del empastillamiento de las muestras.

	Peso prueba (gr)	Peso de pastilla (gr)	Presión (MP)	Dimensiones de cada pastilla
<b>Cus1</b>	20	19.756	5	4cm de diámetro x 0.8cm de altura
<b>Cus2</b>	20	19.856	6	
<b>Cus3</b>	20	19.895	9	
<b>Cus4</b>	20	19.985	9	
<b>MD1</b>	20	19.657	9	
<b>MD2</b>	20	19.925	10	
<b>MD3</b>	20	19.872	9	

En algunas muestras la presión aplicada no fue la misma, ya que debido al tamaño de grano se requería una mayor presión. Cada pastilla obtenida abarcará un área superficial de  $12.57 \text{ cm}^2$  y volumen de  $10.05 \text{ cm}^3$ , dichas medidas se obtuvieron previo a la adquisición de las imágenes hiperespectrales.

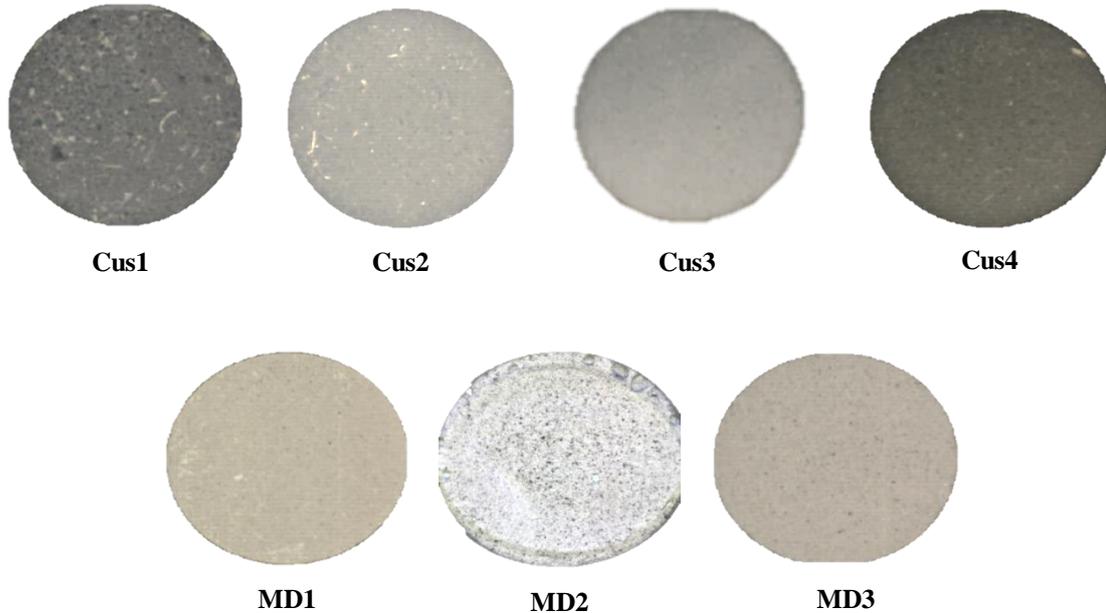


Figura 5.5. Muestras empastilladas sometidas a diferentes presiones.

**Nota:** En el proceso de empastillamiento, se pudo verificar que es dificultoso empastillar la muestra MD2, Suelo Blanco arena, debido a su estructura granular de sedimentos de roca y cal, que impidieron el prensado o agrupamiento de moléculas, razón por la cual no se pudo obtener la pastilla.

Ahora, se procede a tomar el peso de cada muestra seca ( $M_s$ ) con y sin el recipiente luego, se usará la ecuación (5.1) para el cálculo de porcentaje de humedad que existen en las muestras.

El objetivo de calcular el porcentaje de humedad es verificar la influencia que tiene en los espectros de cada suelo y así determinar el espectro característico de los mismos. Los cálculos del porcentaje de humedad se registran en la siguiente tabla.

$$W = \left( \frac{W_h - W_s}{W_s - W_r} \right) 100(\%) \quad (5.1)$$

Donde:

W = contenido de humedad expresado en %

Wh = peso del recipiente más la muestra del suelo húmedo en (gr.)

Ws = peso del recipiente más la muestra seca del suelo en (gr.)

Wr = peso del recipiente (gr.)

**Tabla 5.4.** Calculo del contenido de humedad (%) en las muestras, S/R (sin recipiente), C/R (con recipiente).

	Tipo de suelo		Peso: muestra natural(gr)	Peso: muestra secada(gr)	Peso: recipiente(gr)	Peso %w de humedad	Peso perdido
<b>1</b>	Cus1	S/R	60.077	50.588	43.012g	19.39%	9.489
		C/R	102.238	92.618			9.62
<b>2</b>	Cus2	S/R	60.108	49.62	40.059g	21.75%	10.488
		C/R	99.917	89.223			10.694
<b>3</b>	Cus3	S/R	30.001	23.682	43.012g	30.77%	6.319
		C/R	69.997	63.647			6.35
<b>4</b>	Cus4	S/R	60.01	42.459	41.386g	40.76%	17.551
		C/R	102.212	84.6			17.612
<b>5</b>	MD1	S/R	25.254	18.457	41.386g	35.41%	6.797
		C/R	66.31	59.793			6.517
<b>6</b>	MD2	S/R	60.096	56.336	41.386g	6.34%	3.76
		C/R	101.303	97.73			3.573
<b>7</b>	MD3	S/R	100.091	78.079	88.305g	29.09%	22.012
		C/R	188.002	165.534			22.468

En este cuadro, notamos que los suelos arcillosos (Cus4 y Cus3) contienen mayor humedad, teniendo el valor más alto el suelo Cus4 con un 40.76%, seguida del suelo Cus3 con un 30.77%, también notamos que el suelo de característica arenosa MD2, contiene una menor humedad de 6.341%, esto debido a que las partículas que la componen no retienen mucha agua.

El suelo MD1 tiene una humedad considerable de 35.41%, esto a que es un suelo limoso y con abundante material orgánico.

## 5.2. Protocolo de adquisición de imágenes

### 5.2.1. Sistema de cámara hiperespectral

➤ **Cámara Hiperespectral Pika Nir – 320:**

La cámara de Resonon, es un sistema integral de hardware y software para la adquisición y análisis de datos hiperespectrales.

Las características técnicas de la cámara se muestran en la Tabla 5.5:

**Tabla 5.5.** Especificaciones de la cámara Pika NIR - 320

Rango espectral	900 – 1700 nm
Resolución espectral	4.9 nm
Canales espectrales	168 bandas
Píxeles espaciales	320
Máximos cuadros por segundo	520 fps
Profundidad de Bits	14
Peso con lente	3.21kg
Dimensiones	11.4x27.0x8.9 cm
Tamaño pixel	≈30 μm
Requerimiento de potencia	12V, 5A

**5.2.2. Hardware**

Los avances logrados en la aplicación de esta tecnología han conllevado a mejoras en la velocidad y coste del hardware; sin embargo, los elementos típicos, esquematizados en la Figura 2.6, siguen siendo los mismos por los últimos veinte años.

➤ **Montaje experimental**

Se observa (figura 5.6), el sistema de adquisición de imágenes hiperespectrales en el laboratorio, usando el software Spectronon Pro.

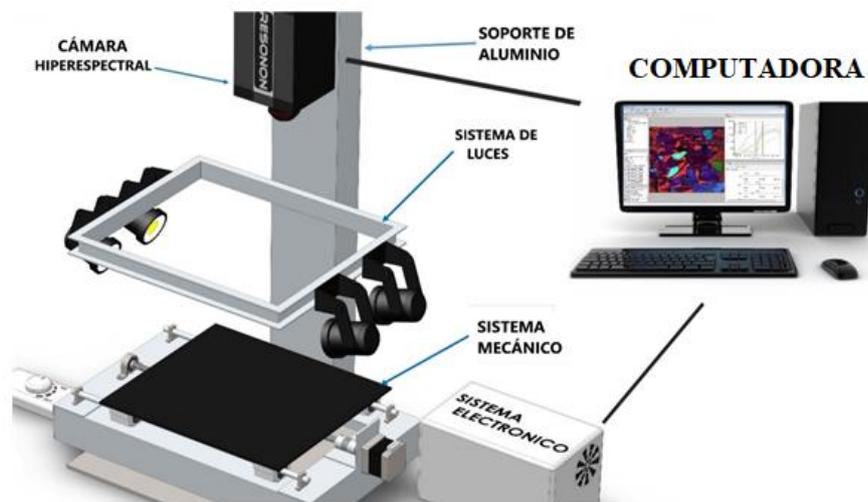


Figura 5.6. Sistema de adquisición de imágenes hiperespectrales.

**i) Mecanismo de soporte y componentes:**

Este sistema consta de un soporte de aluminio, de 120cm y soporta el peso de los equipamientos de la cámara. Cuenta con una plataforma desplazable que mide (18x30  $cm^2$ ), se desplaza una velocidad de 2,5 cm/s, dicho movimiento es generado por un motor a paso que es controlado por un arduino.

**ii) Sistema de luces:**

Está ubicada debajo del lente óptico y sujeta a la torre metálica, tiene cuatro posiciones, esto va depender de la muestra a estudiar. Este sistema de iluminación cuenta con 4 unidades de lámparas halógenas de corriente directa con voltaje de 220 V y de 50 W.

**iii) Cámara:**

Este dispositivo está ubicado a 34.5 cm de la plataforma, y su resolución espacial es de 11,53 cm. La cámara permite la creación de una imagen digital que comprende una resolución espectral muy alta. Además, un píxel de la imagen hiperespectral representa una curva continua de la intensidad de la luz entrante contra longitud de onda.

**iv) Computadora:**

Este sistema registra, controla y permite visualizar la adquisición de las imágenes mediante el software Spectronon – Pro. Además controla la plataforma a través del circuito electrónico arduino. La computadora es de procesador Intel (R) Corel (TM) i7-6700 con una memoria RAM de 16 GB y disco duro de 1 TB.

### **5.2.3. Software**

El software para estos sistemas hiperespectral incluye trabajos de procesamiento digital de imágenes hiperespectrales, las cuales incluyen: Corrección (espectral, espacial), selección de ROIs, reducción de dimensionalidad y modelización de algoritmos para su correcta interpretación [26].

En la Tabla 5.6 se enumeran algunos programas que, individualmente o combinados, son empleados como herramientas informáticas para el procesamiento y análisis de imágenes hiperespectrales.

En este trabajo utilizaremos los siguientes softwares: Spectronon Pro, RStudio y Envi.

Tabla 5.6. Softwares empleados para el procesamiento de imágenes hiperespectrales

SOFTWARE	APLICACIÓN PRINCIPAL	MÉTODOS Y ALGORITMOS
R	<ul style="list-style-type: none"> <li>- Reducción de dimensionalidad</li> <li>- Desarrollo de modelos</li> <li>- Plataforma estadística</li> <li>- Multiplataforma</li> <li>- Software libre</li> </ul>	<ul style="list-style-type: none"> <li>- PCA, MNF, IPC: reducción de dimensionalidad.</li> </ul>
Python	<ul style="list-style-type: none"> <li>- Desarrollo de modelos</li> <li>- Implementación de plataformas</li> <li>- Análisis estadístico</li> </ul>	<ul style="list-style-type: none"> <li>- PPI, N-FINDR y AMEE: Algoritmos clásicos de extracción de endmembers.</li> </ul>
Matlab	<ul style="list-style-type: none"> <li>- Desarrollo de programas de análisis</li> <li>- Desarrollo de modelos</li> <li>- Visualización de imágenes</li> <li>- Implementación de plataformas</li> </ul>	<ul style="list-style-type: none"> <li>- K-means (K-M), Support Vector Machines (SVM), K- nearest neighbors (K-NN), Decision Trees (DT), Artificial Neural Network (ANN), Partial Least Squares (PLS), etc...:</li> </ul>
Big data	<ul style="list-style-type: none"> <li>- Control de información</li> </ul>	<p>Son técnicas de aprendizaje de máquinas para la minería de datos o big data, cuyo procesamiento puede ser supervisado y no supervisado, mediante el cual se obtiene resultados cuantitativos, cualitativos y estadísticos que nos llevara a la toma de decisiones.</p>
Envi	<ul style="list-style-type: none"> <li>- Análisis de clasificación</li> <li>- Correcciones atmosféricas</li> <li>- Correlación de bandas</li> </ul>	
QGIS	<ul style="list-style-type: none"> <li>- Calculo de estadísticas e histogramas</li> <li>- Detección de cambios climáticos</li> </ul>	

#### 5.2.4. Calibración del sistema

A continuación, se enumerará los pasos para una correcta calibración del sistema y así obtener imágenes que contengan información relevante.

- i) Se fija la distancia del lente a la plataforma, para el estudio de las muestras de suelo se ubica la cámara a 34.5 cm del objeto.
- ii) Se enciende el sistema de iluminación, el tiempo estimado para que se estabilicen es de 15 a 20 minutos.

- iii) Se hace el encendido del sistema eléctrico y se abre el software SpectronPro, verificando el encendido correcto de los controladores de la cámara.
- iv) Se hace el reconocimiento de la fuente de luz a través de la cámara en tiempo real, esto se logra encendiendo el botón Focus (F) en la ventana del SpectronPro, tal como se visualiza en la figura 5.7.



Figura 5.7. Primera visualización del sistema (imagen borrosa), donde se hace el reconocimiento de la luz y objetos.

La figura anterior muestra una imagen distorsionada, esto requiere que se haga una calibración del lente, para fines de esta investigación y la distancia de la cámara hacia la plataforma se usa el lente con una apertura del objetivo  $f/1.8$  para regular la cantidad de luz entrante. La figura 5.8 muestra el resultado de una correcta calibración del enfoque de luz.

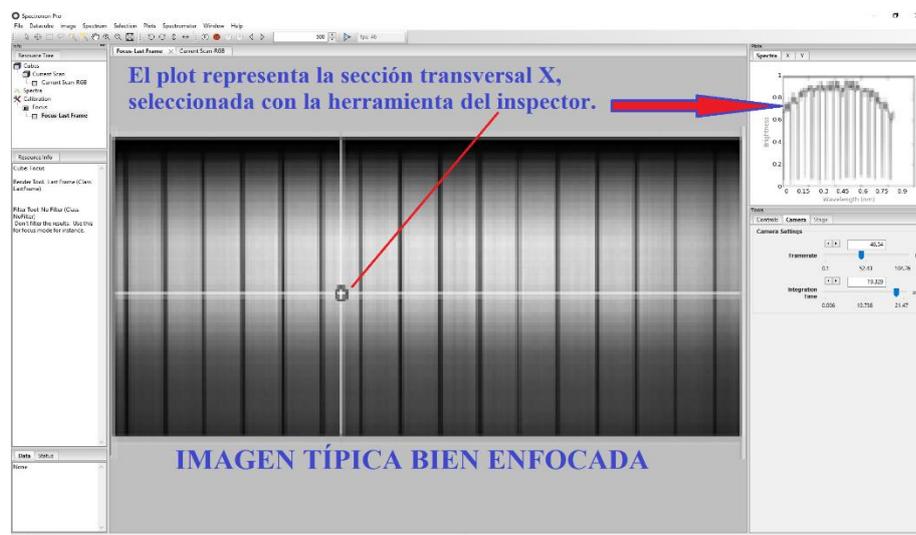


Figura 5.8. Resultado del correcto ajuste en el enfoque de luz entrante hacia los sensores.

v) Se debe hacer los ajustes geométricos del sistema, para lo cual se utiliza la hoja de calibración de la relación de aspecto (ver figura 5.9) para establecer la velocidad del escenario y la velocidad de fotogramas del generador de imágenes.

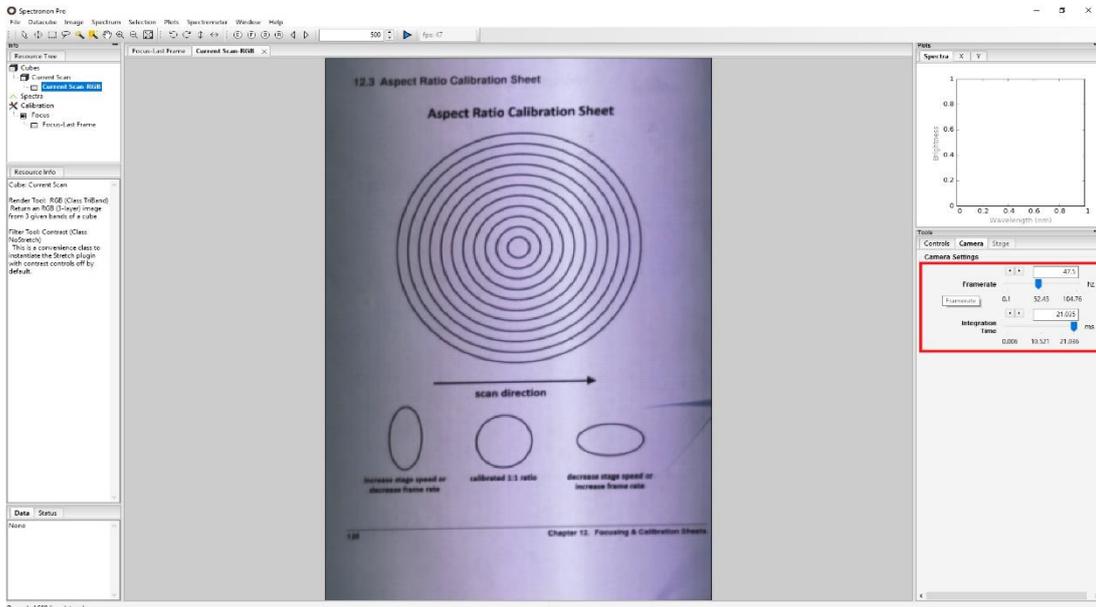


Figura 5.9. Resultado del ajuste de controles de la cámara. Las muestras de suelos fueron tomadas con un Framerate de 47.5 Hz e Integration Time de 21.35 ms.

### 5.2.5. Calibración de la imagen

En este apartado se describe cómo configurar el sistema para buscar reflectancia escalada a un objeto de referencia.

#### ❖ Eliminación de corriente oscura:

El software SpectronPro facilita la eliminación de ruido de corriente oscura promedio de sus escaneos. Para tal objeto, primero se bloquea toda la luz que ingresa al espectrómetro con su propia cubierta de lente (ver figura 5.10) luego se hace clic en el botón **Dark Current**, que se encuentra en la barra de herramienta de SpectronPro.



Figura 5.10. Cubierta de lente para la eliminación de corriente oscura.

❖ **Establecer referencia de reflectancia:**

La medición de la reflectancia absoluta de un objeto requiere una corrección para tener en cuenta los efectos de iluminación. En primer lugar, se coloca un material de referencia dentro del campo de visión del espectrómetro de imágenes, el material debe tener una superficie uniforme, puede ser una lámina de teflón blanco (ver figura 5.11) o el Spectralon® que viene con el sensor. Inmediatamente después se hace clic en el botón **Cubo de corrección de respuesta** ® en la barra de herramientas de SpectronPro. Por otro lado el software considera al blanco de referencia con un valor de reflectancia ~99.9%. La imagen espectral del teflón tiene que tener la reflectancia homogénea en todas sus bandas, tal como se muestra en la figura 5.10, esto asegura la reproducibilidad de los experimentos.

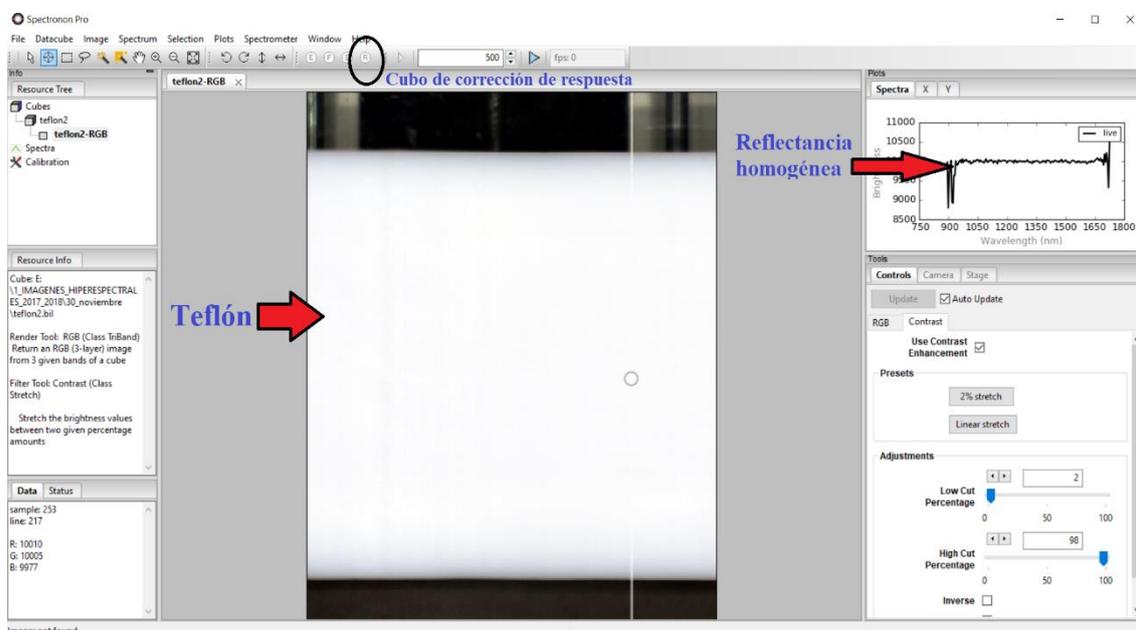


Figura 5.11. Referencia de reflectancia usando un teflón como patrón de calibración.

❖ **Escaneo de muestras y toma de imágenes:**

Después de realizar la calibración de todo el sistema, y que el sistema de luces este enfocado, se va a proceder a realizar el escaneo de las muestras de suelo, para la adquisición de las imágenes, dichas tomas se registraron con 300 líneas de escaneo. Estas imágenes hiperespectrales obtenidas en formato “.bil”, son guardadas en una carpeta creada en el disco duro del computador para luego ser procesadas.

### 5.2.6. Adquisición de las imágenes hiperespectrales de los suelos:

En este apartado, se mostrará las imágenes hiperespectrales adquiridas de los suelos en sus tres estados (natural, secado - molido y empastillado).

---

#### SUELOS EN ESTADO NATURAL

---

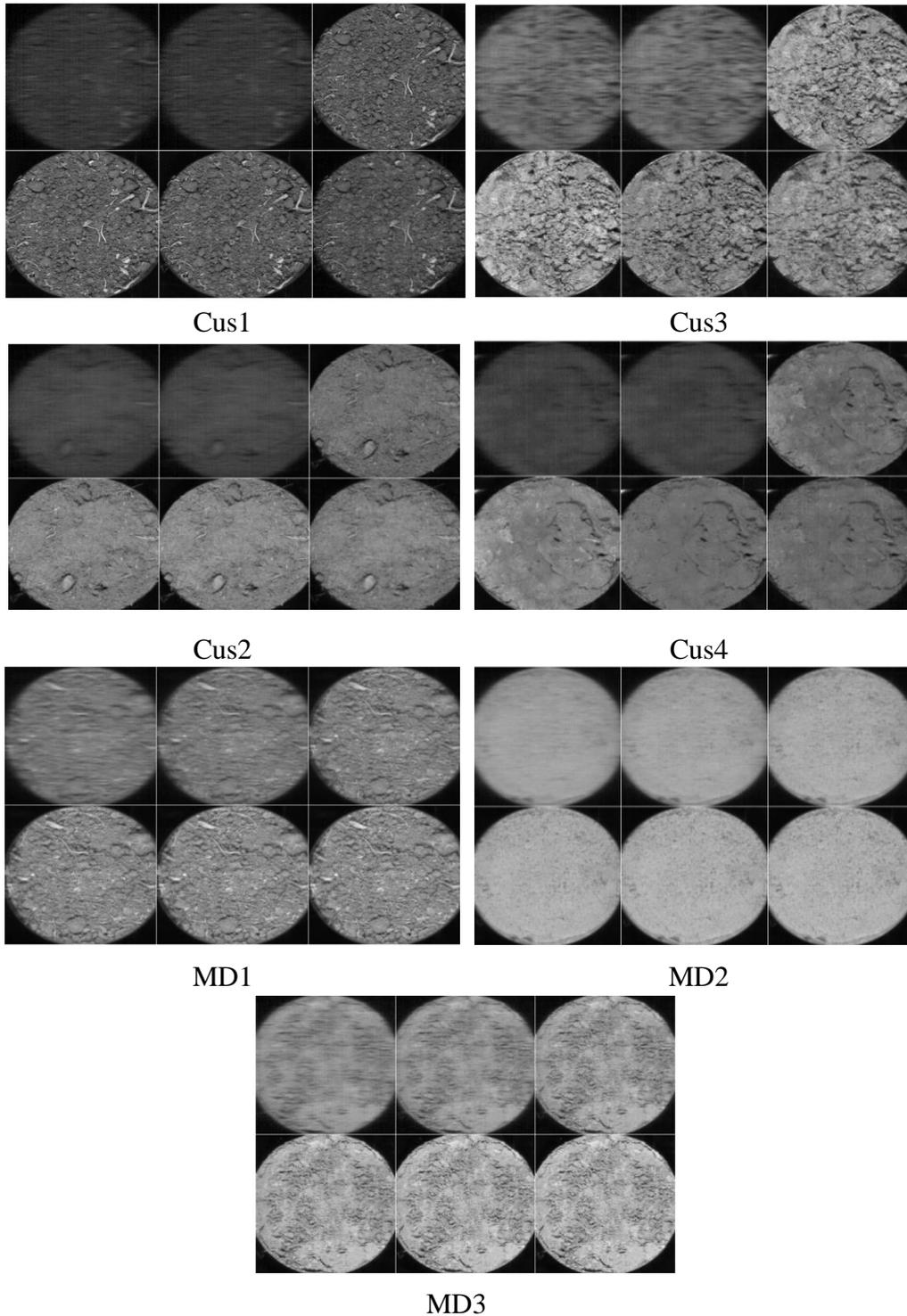


Figura 5.12. Visualización de 6 bandas (9-14) de las 168 que hay en la imagen hiperespectral de los suelos en estado natural.

En la figura anterior, se observa que las bandas del (1 – 10) y del (160 – 168), muestran imágenes borrosas o con ruido, que no aportan nada de información, además las muestras presentan otros componentes, dichas bandas son eliminadas en el pre procesamiento de la HSI.

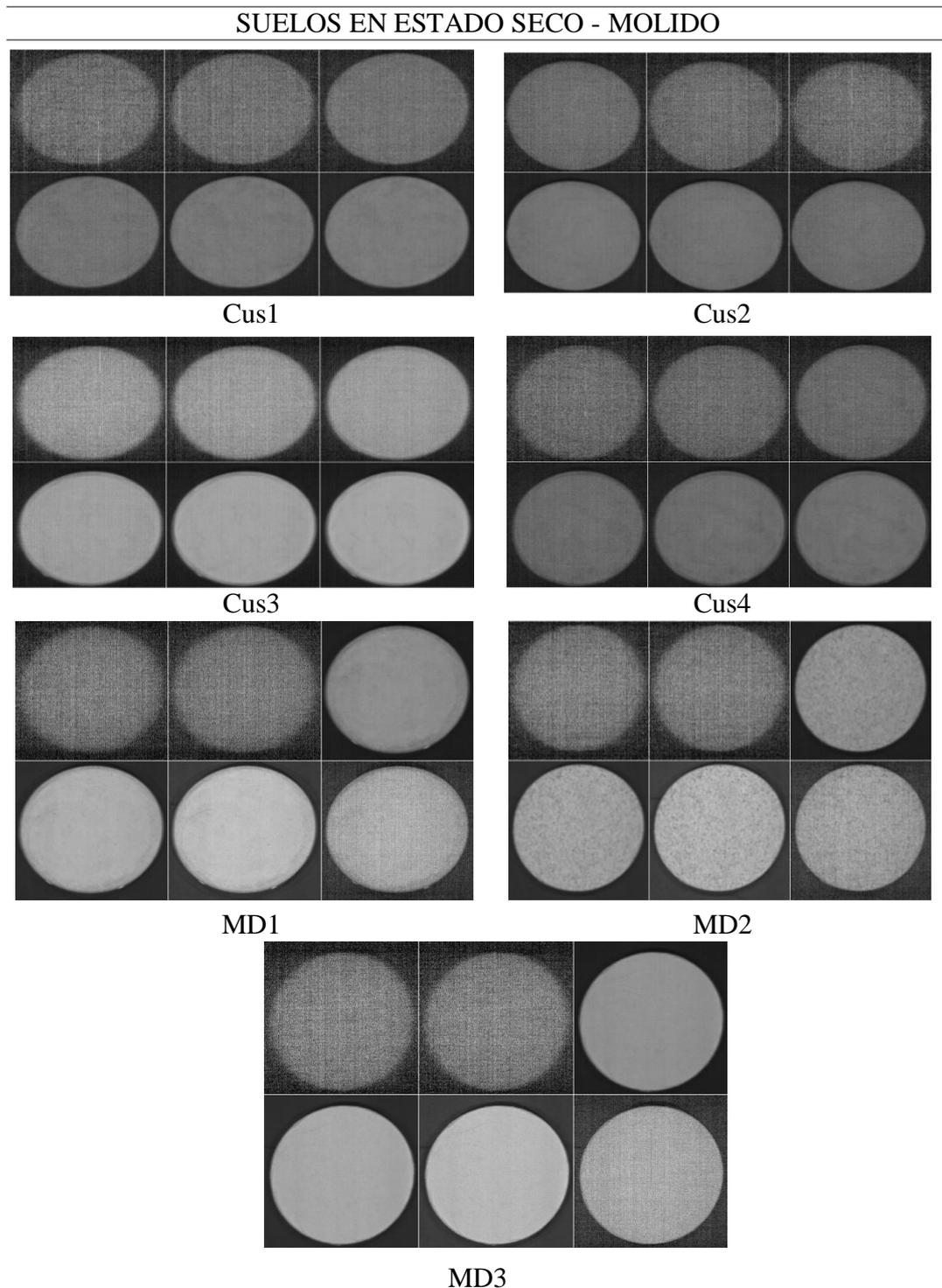


Figura 5.13. Visualización de 6 bandas (9-14) de las 164 que hay en la imagen hiperespectral de los suelos molidos.

Se acaba de visualizar las imágenes de los suelos molidos, se puede apreciar la homogeneidad de su composición y superficie, esto nos permitirá evaluar la variación del espectro de cada muestra.

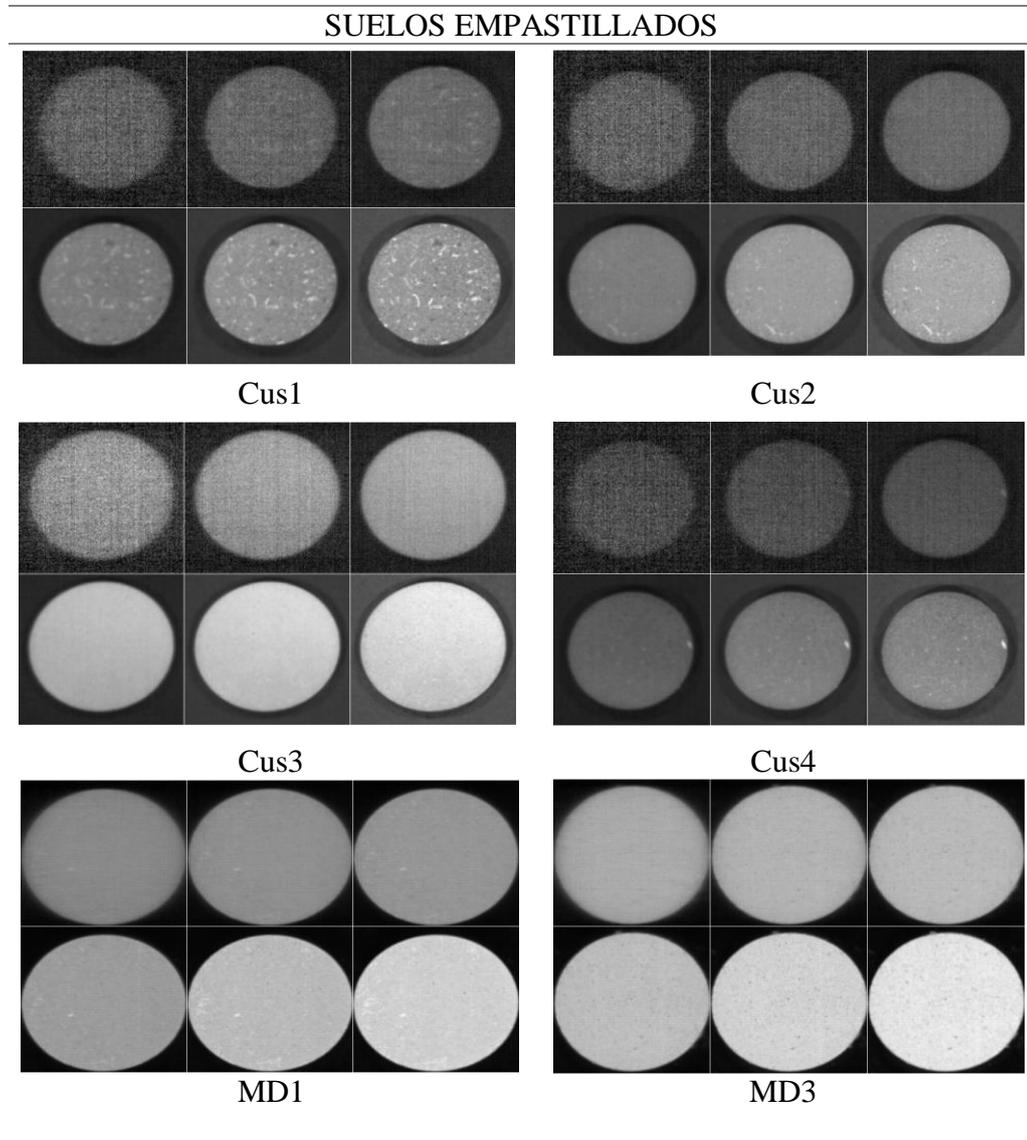


Figura 5.14. Visualización de 6 bandas (10,11,12,20,80,150) de las 168 que hay en la imagen hiperespectral de los suelos molidos.

Esta figura muestra los suelos empastillados, se escogió algunas bandas al azar para visualizar algunas bandas correctas y otras con ruido, una vez más queda demostrado que las bandas iniciales (1 – 12) y finales (160 – 168) presentan un ruido considerable. También, líneas más arriba mencionamos que el suelo MD2, no se pudo empastillar debido al tamaño de grano o textura.

### 5.3. Pre procesamiento HSI

#### i) Corrección de bordes:

Al visualizar las imágenes hiperespectrales, se identificó bandas con influencia de ruido, dichas bandas serán extraídas para un mejor tratamiento de datos y solo consideraremos 140 bandas en todo nuestro proceso, teniendo como banda inicial la imagen 19, y como banda final la imagen 158. En la figura 5.15, se muestra los espectros de una imagen con corrección de borde.

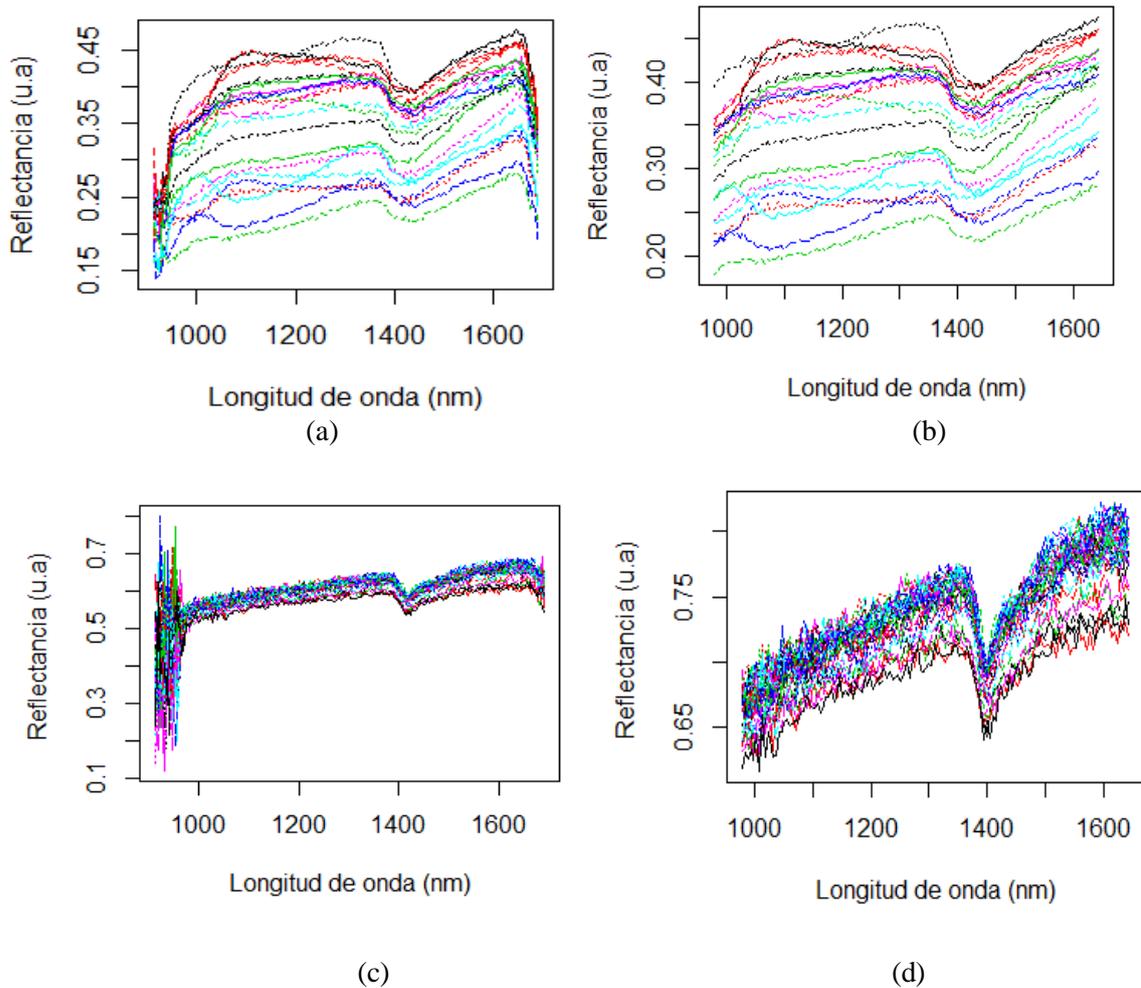


Figura 5.15. (a-c) Espectros con influencia de ruido en las bandas iniciales y finales. (b-d) Eliminación de bandas extremas, considerando como rango de espectro (980.196 - 1642.114 nm) y longitud de espectro 4.762 nm.

Este procedimiento se repetirá para todas las muestras de suelo a tratar, la corrección de bandas es necesaria, ya que permite descartar aquellas bandas que presentan ruido e información redundante al inicio y al final, es por eso que se hace un recorte de 168 a 140 bandas.

ii) Normalización con Máx - Min y Suavizamiento con SG:

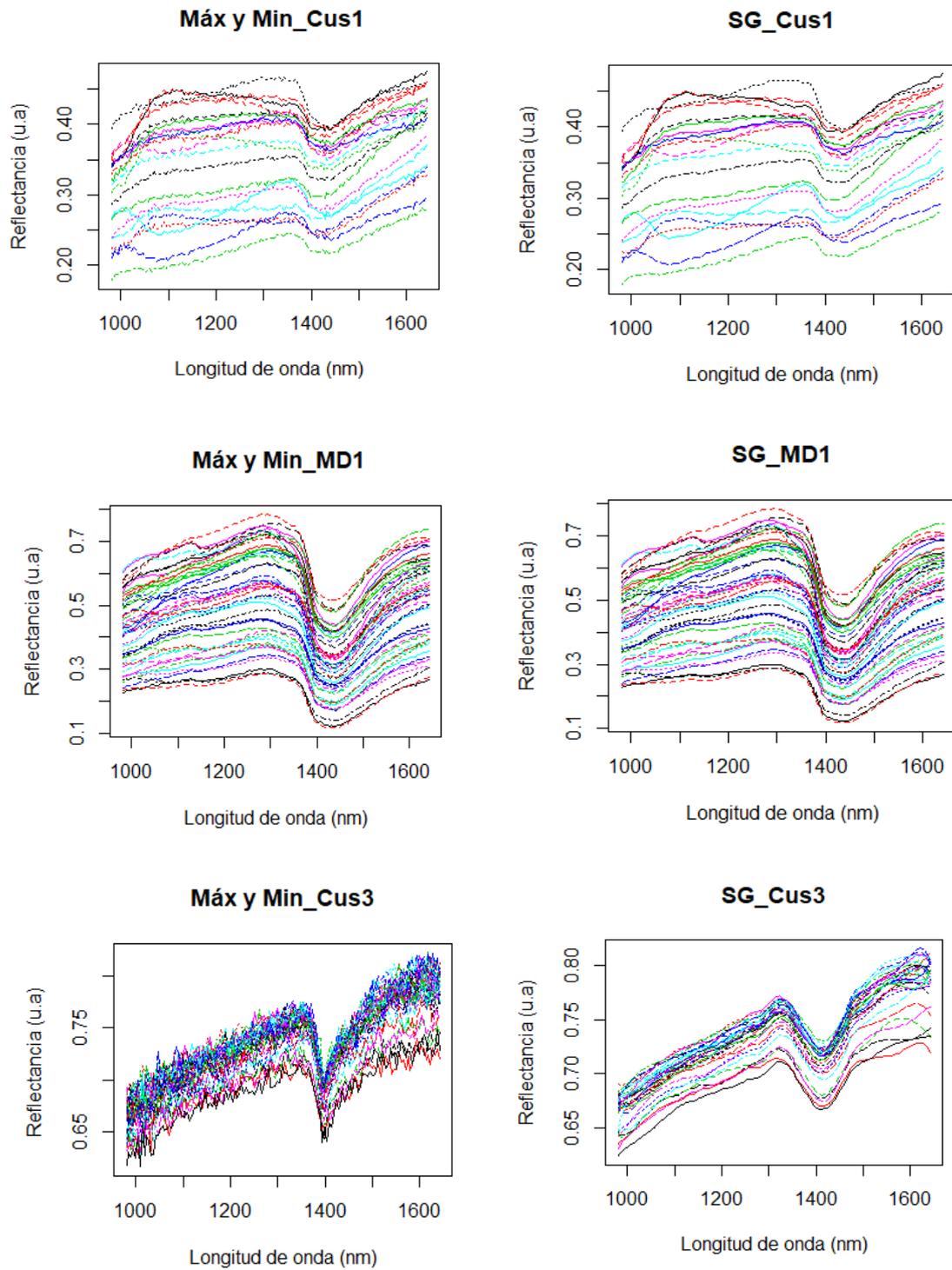


Figura 5.16. La columna izquierda representa el preprocesamiento de las firmas espectrales con la normalización de máximos y mínimos, la columna derecha es el suavizamiento de firmas espectrales con Savitzky - Golay.

La figura anterior (fig. 5.16) muestra la comparativa de la normalización de máximos y mínimos con el suavizamiento de Savitzky - Golay, donde usamos un polinomio de grado 5 con 11 puntos, verificamos que, para una mayor cantidad de puntos, los espectros se suavizan más y que las bandas de importancia tienen tendencia a perderse, es por eso que fijamos 11 puntos. Este procedimiento se repite para todas las muestras de suelos (natural, molido y empastillado).

### iii) Corrección de escala multiplicativa con la *Standard Normal Variate* SNV

Luego de la normalización por máximos - mínimos y el suavizado con Savitzky – Golay, notamos que, debido a la dispersión de la luz, existe las diferencias de escala multiplicativa, como se evidencia en la figura 5.16; por esta razón recurrimos a la *Standard Normal Variate* (SNV), dicho método corrige la dispersión, tal como se muestra en las siguientes figuras (fig. 5.17 y fig. 5.18).

Esta corrección de dispersión fue aplicada a todas las muestras, simultáneamente umbralizamos la imagen, quitamos el fondo restando los pixeles de menores scores, aquellos que no guardaban relación con los pixeles de mayor variabilidad.

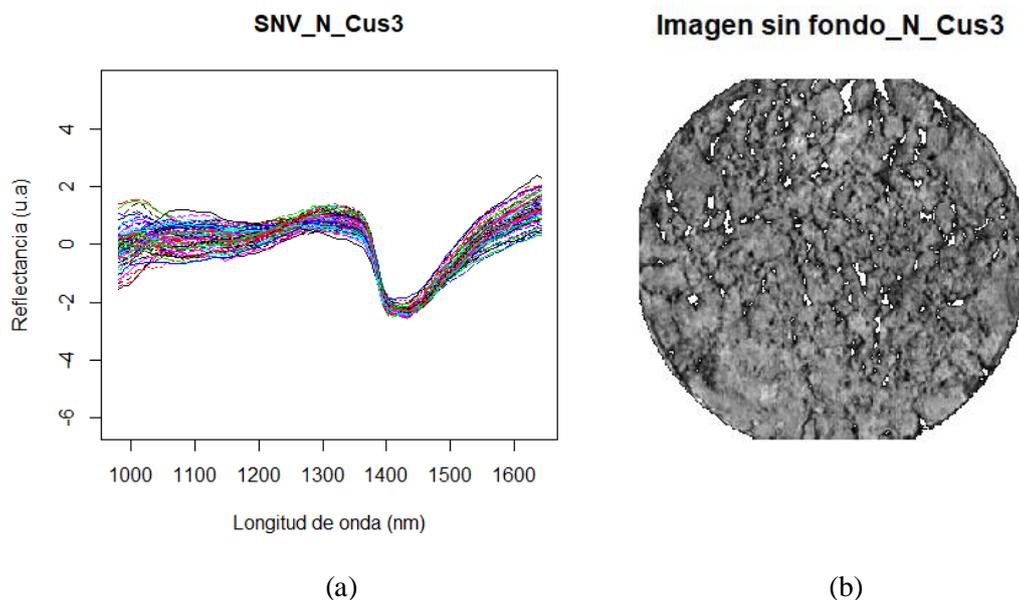


Figura 5.17. a) Espectros de los 100 primeros pixeles del suelo natural Cus3, al cual se aplicó la normalización máx – mín, SG y la corrección de dispersión SNV. b) Imagen umbralizada del suelo natural de Cus3.

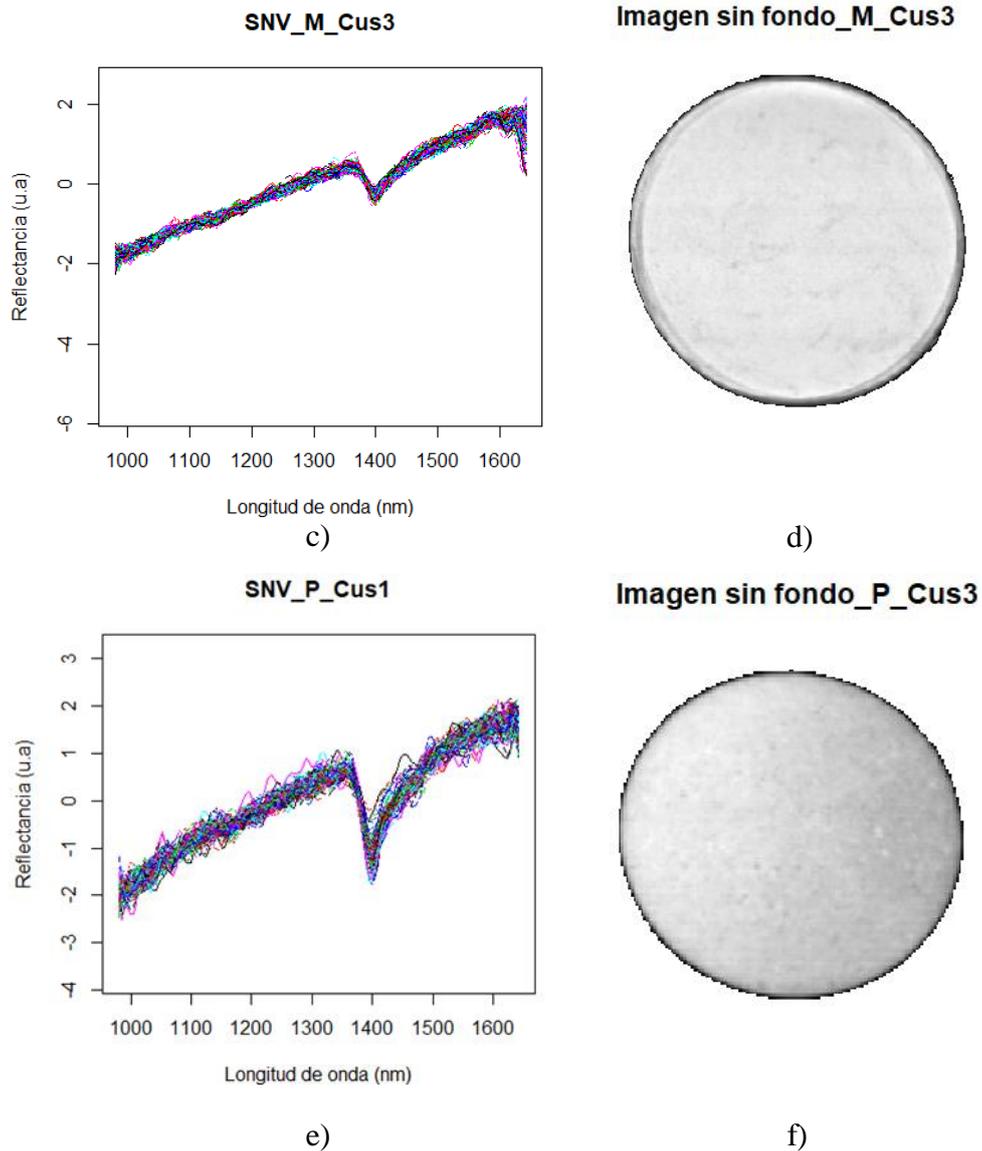


Figura 5.18. c - e) Espectros de los 100 primeros píxeles del suelo molido y empastillado Cus3, respectivamente, al cual se aplicó la normalización máx – mín, SG y la corrección de dispersión SNV. d - f) Imagen umbralizada del suelo molido y empastillado de Cus3 respectivamente.

Estas figuras muestran la corrección de dispersión usando el SNV, para la visualización de ello, usamos el suelo Cus3 en sus tres estados (natural, molido y empastillado) y así ver el comportamiento de los espectros. Los resultados de la figura 5.17 y 5.18 (a – c – e) muestran espectros más juntos, es decir, se eliminó la escala multiplicativa y la línea de base.

Por otro lado, identificamos la máscara de trabajo o región de interés (ROIS), análogo al procedimiento que se muestra en la figura 5-17 y 5.18 (b – d – f) se realizó para todas las muestras.

#### 5.4. Estandarización de los Espectro de reflectancia de los suelos

Se obtuvieron los espectros de reflectancia de las distintas muestras de suelos estudiadas. La figura 5.19 muestra los espectros obtenidos para los suelos de Cusco en sus tres estados.

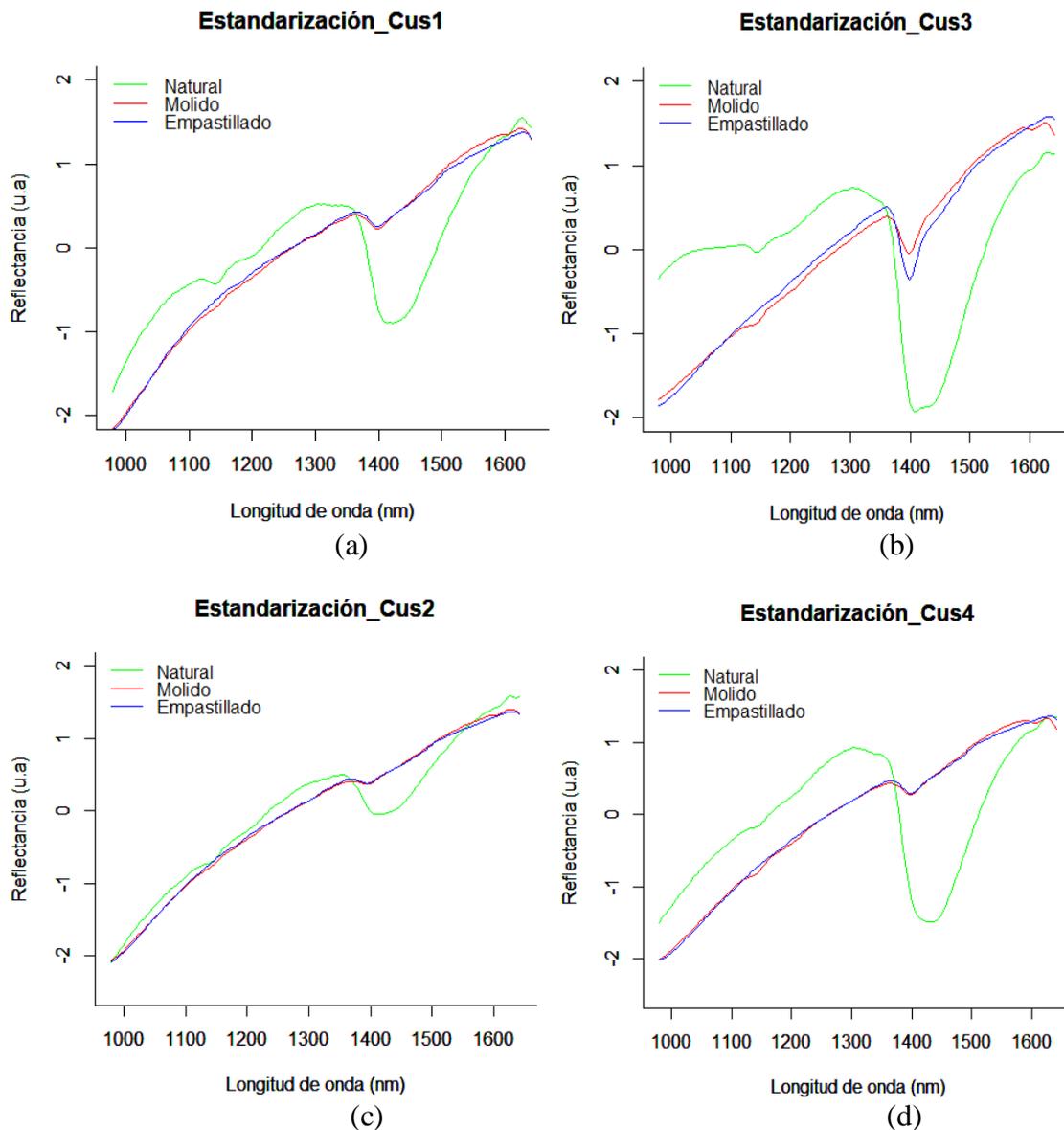


Figura 5.19. Espectros de reflectancia de las muestras Cus1, Cus2, Cus3 y Cus4 en diferentes condiciones.

La figura 5.20 muestra los espectros de reflectancia obtenidos para los suelos de Madre de Dios. Los espectros de color verde corresponden a espectros de reflectancias de suelos sin ningún tratamiento, estos espectros son el promedio de varios componentes que se encuentran en cada muestra natural. Los espectros de color rojo corresponden a las firmas

espectrales de los suelos molidos, previamente secadas. Los espectros de color azul representan a las firmas espectrales de los suelos empastillados.

En las muestras de Cusco los espectros sin ningún tratamiento térmico presentan una banda de absorción ancha típica del agua alrededor de 1450 nm, la cual se muestra más intensa en los espectros de reflectancia de la muestra Cus3.

Los espectros de las muestras secadas y empastilladas presentan la absorción típica del  $Fe_2O_3$  alrededor de 1415 nm. Esta banda es más intensa en los espectros de reflectancia de la muestra arcillosa Cus3 y la menos intensa se ve en la muestra Cus1 y Cus2.

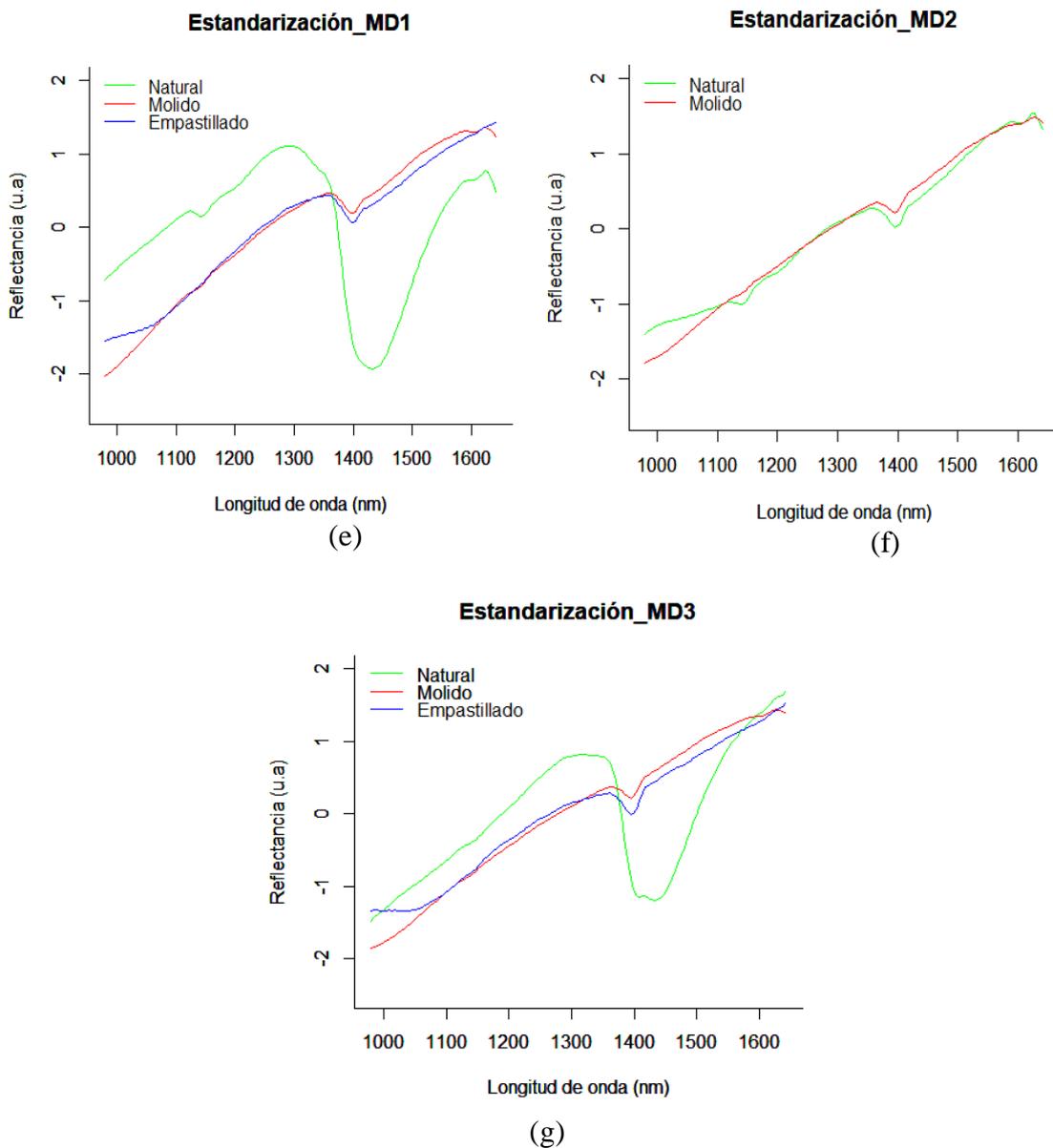


Figura 5.20. Espectros de reflectancia de las muestras MD1, MD2 y MD3 en diferentes condiciones.

Los espectros de las muestras secadas y empastilladas presentan la absorción típica del  $Fe_2O_3$  alrededor de 1415 nm. Esta banda es más intensa en los espectros de reflectancia de la muestra arcillosa Cus3 y la menos intensa se ve en la muestra Cus1 y Cus2.

Los espectros de las muestras secadas molidas muestran una reflectancias similar a las muestras empastilladas, aunque con una pendiente ascendente con la longitud de onda.

En los espectros de reflectancia de las muestras de Madre de Dios sin ningún tratamiento, la banda de absorción asociada al agua es muy intensa en el espectro de la muestra MD1, mientras que la muestra MD3 presenta esta banda con una intensidad similar a lo observado en las muestras de Cusco. La muestra MD2 correspondiente a arena de rio no presenta esta banda de absorción considerable. Los espectros de reflectancia de las muestras secadas y empastilladas presentan una banda de absorción asociada al  $Fe_2O_3$  con la misma intensidad a las muestras de Cusco, a pesar del tratamiento del secado, esta banda está relacionado al agua, pero también a la composición orgánica y de OH, la literatura [21] indica que estos suelos son arcillas, así como se describió en la tabla 5.1.

Por último, los espectros de reflectancia de pixeles que se encontraron en cada muestra que no guardaban relación con la mayoría de los pixeles, corresponde a espectros de diferentes materiales, como vegetación, raíces secas u otros materiales no identificados, tal como se observa en la figura 5.21.

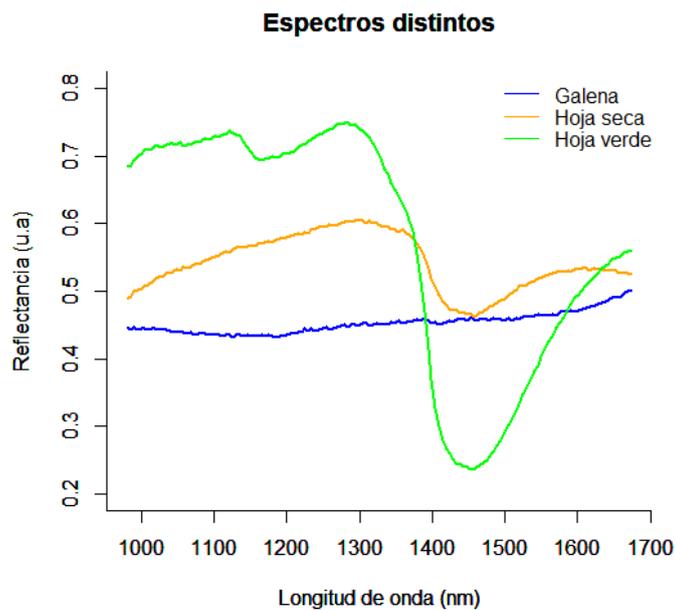


Figura 5.21. Firmas espectrales de distintos materiales encontrados en cada muestra de los suelos estudiados.

### 5.4.1. Librería espectral

Para las muestras secadas y empastilladas Cus1, Cus2, Cus3, Cus4, MD1, MD2 y MD3, se obtuvieron los promedios de los espectros de reflectancia de los pixeles de las imágenes de las muestras.

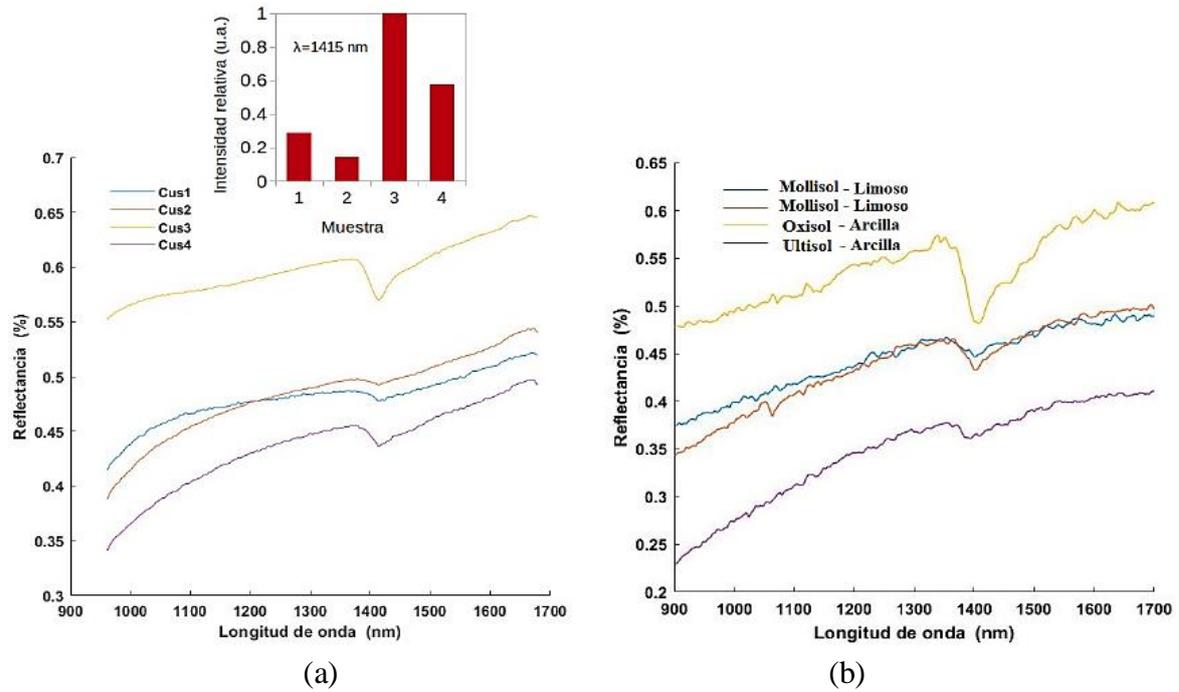


Figura 5.22. Espectros de reflectancia estandarizados de las muestras estudiadas de Cusco. El gráfico insertado muestra la intensidad relativa de las absorciones de las muestras para  $\lambda = 1415 \text{ nm}$ .

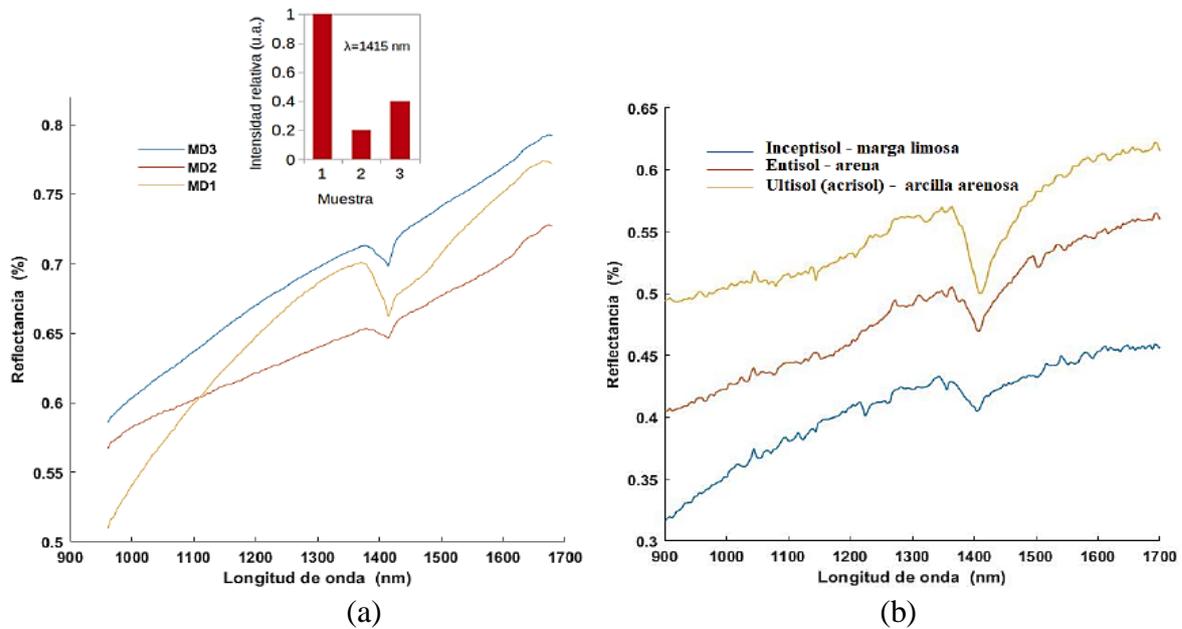


Figura 5.23. Espectros de reflectancia estandarizados de las muestras estudiadas de Madre de Dios. El gráfico muestra la intensidad relativa de las absorciones de las muestras para  $\lambda = 1415 \text{ nm}$ .

Estos espectros promedio denominados espectros estandarizados se muestran en la figura 5.22 y 5.23 parte a) y se compara con los espectros tomados de la librería ASTER [34]. Estos espectros aparecen en la figura 5.22 y 5.23 parte b). Para identificar estos espectros en la librería ASTER, se tomó en cuenta la clasificación según [72].

## 5.5. Procesamiento HSI

Para el procesamiento de imágenes de las muestras, utilizamos todos los procedimientos anteriores. En este apartado simularemos una mezcla en condiciones controladas de los suelos para probar la efectividad del algoritmo PPI y validar los espectros estandarizados.

### 5.5.1. Análisis de los suelos usando el algoritmo PPI

Esta técnica del PPI será utilizada para distinguir los tipos de suelos a partir de una imagen hiperespectral, para tal objeto las muestras fueron dispuestas en la plataforma de toma de imagen y ubicadas tal como se muestra en la figura 5.24. El conocimiento previo de la ubicación de las muestras de los suelos de Cusco y Madre de Dios, permite validar las clasificaciones obtenidas luego del análisis de las imágenes con el algoritmo PPI.

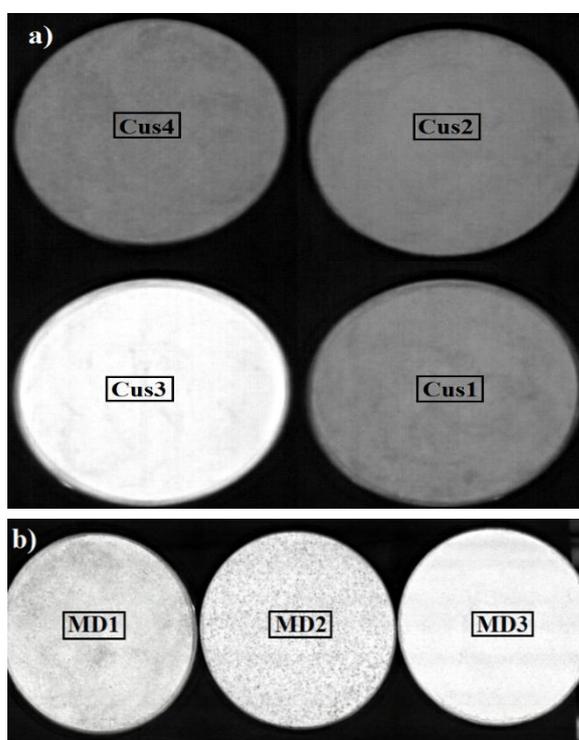


Figura 5.24. a) Imagen hiperespectral en falso color entre las bandas iniciales 15, 16 y 17, de los 4 suelos de Cusco. b) Imagen hiperespectral en falso color entre las bandas iniciales 15, 35 y 79, de los 3 suelos de Madre de Dios.

La imagen mostrada en la figura 5.24 se generó usando los datos iniciales de las imágenes hiperespectrales tomadas entre 900 y 1700 nm. Para generar la imagen de las muestras de suelos de Cusco se usaron la combinación de las bandas 15, 16 y 17, mientras que para generar las imágenes de las muestras de suelos de Madre de Dios se usaron las banas 15, 35 y 79.

Obtenida la imagen de los suelos, se procede a extraer datos iniciales de abundancia de pixeles y el área que representan en la imagen cada suelo presente en la figura 5.24. dichos valores serán resumidos en la tabla 5.7 para los suelos de Cusco y la tabla 5.8 para los suelos de Madre de Dios.

Tabla 5.7. Abundancia de pixeles de los suelos de Cusco y el área en cm<sup>2</sup> que representan en la imagen.

<b>Muestra</b>	<b>Pixeles</b>	<b>% Área</b>	<b>Área en cm<sup>2</sup></b>
<b>Imagen Total</b>	360 000	100 %	190.425
<b>Cus1</b>	57 520	15.98 %	30.426
<b>Cus2</b>	57 805	16.06%	30.576
<b>Cus3</b>	57 637	16.01 %	30.488
<b>Cus4</b>	57 224	15.90 %	30.269
<b>Fondo</b>	118 696	32.97 %	62.785
<b>Mezcla de pixel</b>	11 118	3.09 %	5.881
<b>Total relativo</b>	348 882	96.91 %	184.544
<b>%Error (mezcla)</b>			<b>3.09 %</b>

Tabla 5.8. Abundancia de pixeles de los suelos de Madre de Dios y el área en cm<sup>2</sup> que representan en la imagen.

<b>Muestra</b>	<b>Pixeles</b>	<b>% Área</b>	<b>Área en cm<sup>2</sup></b>
<b>Imagen Total</b>	270 000	100 %	142.82
<b>MD1</b>	57 092	21.15 %	30.19
<b>MD2</b>	57 053	21.13 %	30.18
<b>MD3</b>	57 829	21.49 %	30.59
<b>Fondo</b>	87 770	32.51 %	46.43
<b>Mezcla de pixel</b>	10 256	3.8 %	5.43
<b>Total relativo</b>	259 744	96.20 %	137.39
<b>%Error (mezcla)</b>			<b>3.8 %</b>

Las tablas 5.7 y 5.8 muestran la abundancia de píxeles de cada suelo que se muestra en la figura 5.24 con un error de 1.77 % y 2.33 %, respectivamente y a su vez el área en  $\text{cm}^2$  que representa cada muestra, para una resolución espacial de 0.025cm.

Estas dos tablas, nos muestran datos y resultados a priori que nos servirá para corroborar y contrastar los resultados obtenidos con el algoritmo PPI y otras técnicas de análisis, como el PCA, Least Squares (LS), Support Vector Machine (SVM) o K- Means.

Para el procesamiento con el algoritmo PPI, añadiremos un procedimiento que se describió en la figura 4.7 y que es fundamental para que dicho algoritmo funcione correctamente. Se trata de la reducción de la dimensionalidad usando el MNF.

#### 5.5.1.1 Reducción de datos espectrales MNF (Mínima fracción de ruido)

El método MNF, representa la segregación de ruido y eliminación de bandas con información redundante, en la figura 5.25 se visualiza el blanqueamiento de ruido de la imagen de los suelos, donde se distingue todos los componentes presentes en cada imagen.

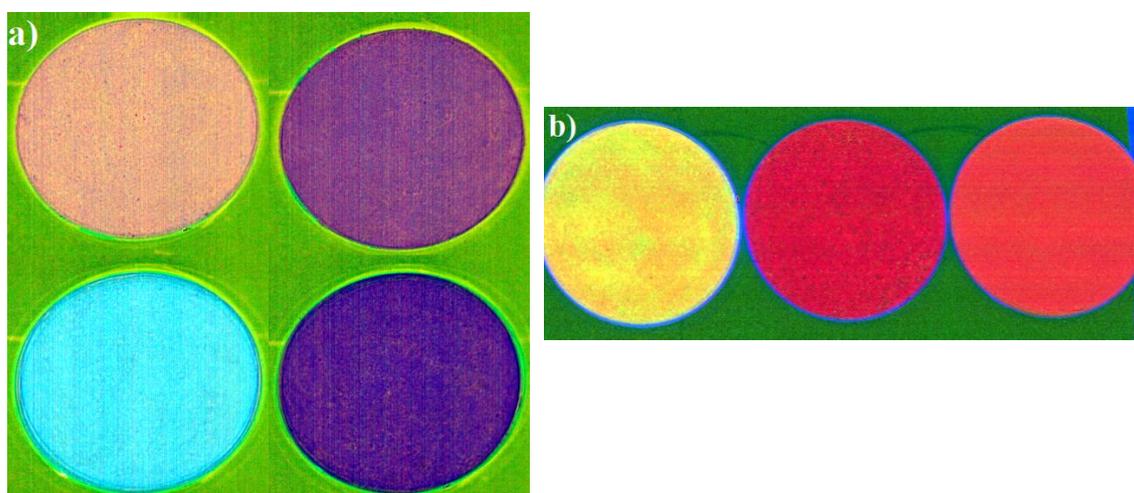


Figura 5.25. a) y b) Blanqueamiento de ruido de la imagen de los suelos de Cusco y Madre de Dios respectivamente, luego de aplicar el MNF.

La reducción de ruido de estas imágenes con MNF fue usando el software Environment for Visualizing Images (ENVI), adicional a ello se implementó un código en el software R, obteniendo resultados similares. Luego de la diagonalización de la matriz de covarianza, no se consideraron las bandas con valores propios por debajo de 1. El número final de bandas fue de 54 en ambos casos. A continuación, la figura 5.26 y 5.27 mostrará los MNF's

correspondientes a las muestras en estudio, los resultados están ordenados respecto al valor mínimo de ruido que poseen.

Se aprecia que a partir del MNF- 8 el resto de imágenes MNF ya no contienen información coherente, por ello se trabajó con las 7 primeras bandas MNF.

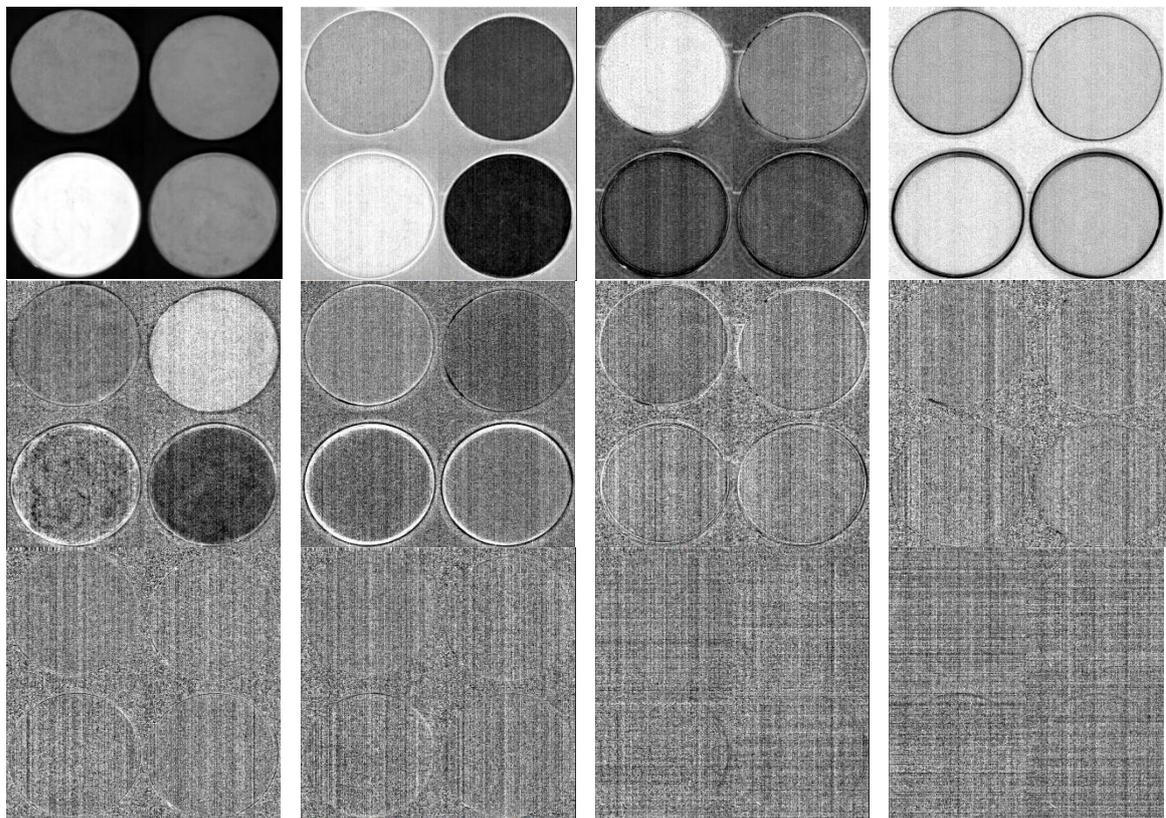


Figura 5.26. Visualización de 12 bandas MNF, de la imagen hiperespectral del Cusco.

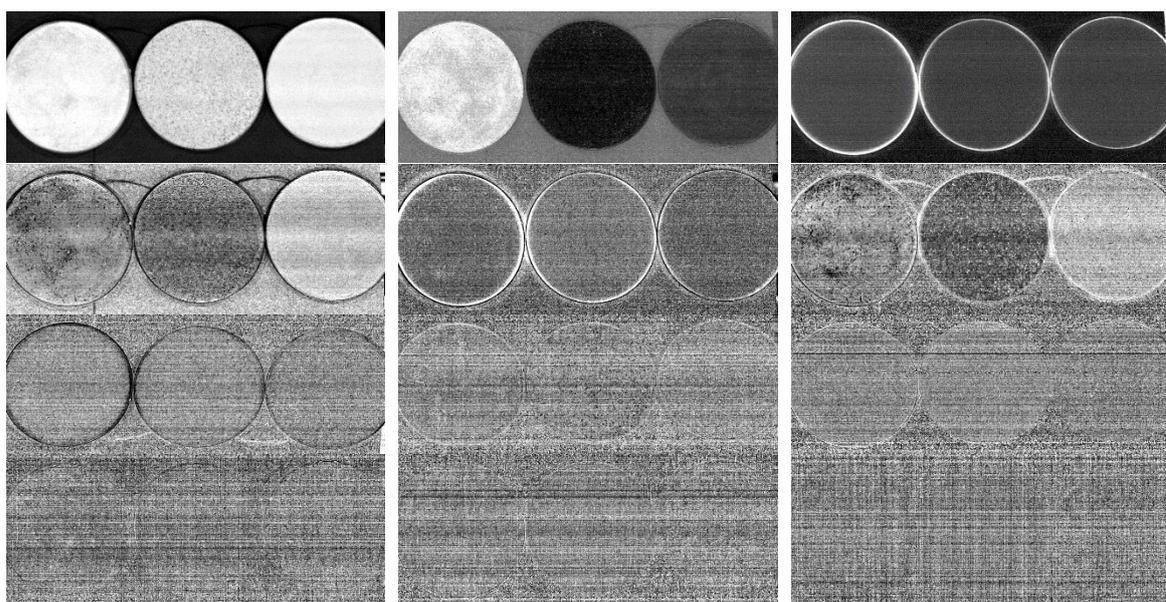


Figura 5.27. Visualización de 16 bandas MNF, de la imagen hiperespectral de Madre de Dios.

### 5.5.1.2 Reducción de datos espaciales del algoritmo PPI

A continuación, se hizo un análisis de las bandas MNF resultantes usando PPI, el cual permite identificar los endmembers o elementos principales en las imágenes analizadas. De manera específica, el algoritmo PPI asigna un índice de pureza a todos los píxeles de la imagen, se calcula proyectando repetidamente nD diagramas de dispersión en un vector unitario aleatorio. El contador de cada punto se inicializa al valor 0 y en la figura 5.28 está representado por las tonalidades oscuras.

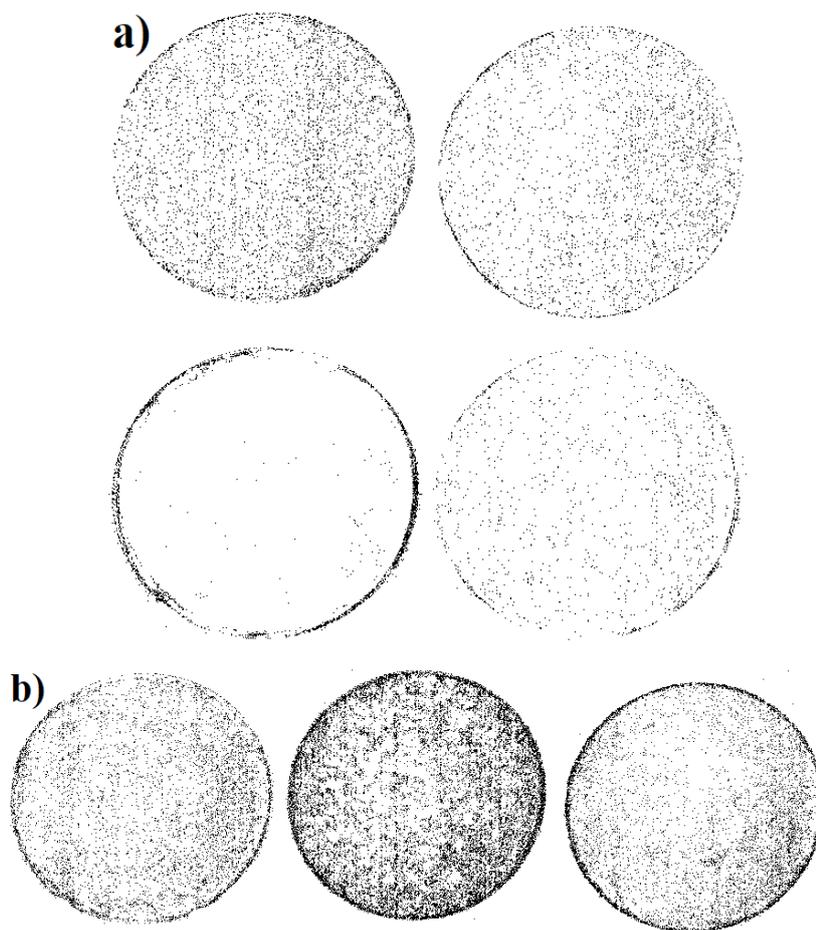


Figura 5.28. Gráfico de índices de pureza de píxeles, para los suelos de a) Cusco y b) Madre de Dios.

Esta figura 5.28, da como resultado una imagen de índice de pureza incrementada en la unidad, formada por los índices asociados a cada uno de los píxeles de la imagen.

Otra forma de divisar esos puntos de pureza, es haciendo un contraste inverso a la figura anterior, y obtener los PPI en escala de cero (negro).

### 5.5.1.3 Visualización n-D y Mapa de Distribución y Abundancia

Siguiendo con el flujo de procesamiento de reloj de arena (figura 4.7), se visualiza los datos en el espacio n-D (figura 5.29) y se agrupa los píxeles más puros en los miembros finales, opcionalmente proporciona sus propios miembros finales. Luego de este procedimiento se mapea la distribución y abundancia de los miembros finales, ver figura 5.30.

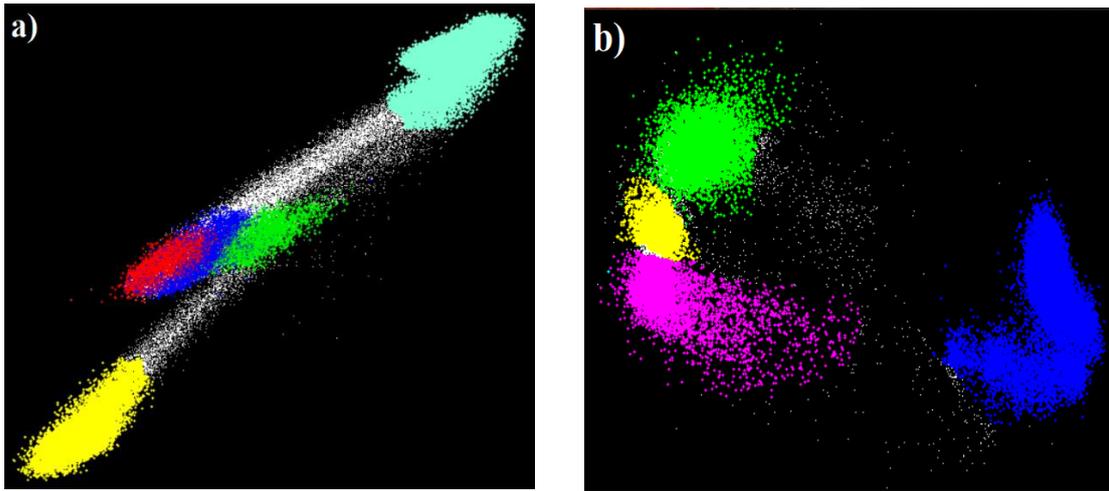


Figura 5.29. Visualización n-D y distribución de la nube de píxeles en los extremos para: a) Suelos de Cusco y b) Suelos de Madre de Dios.

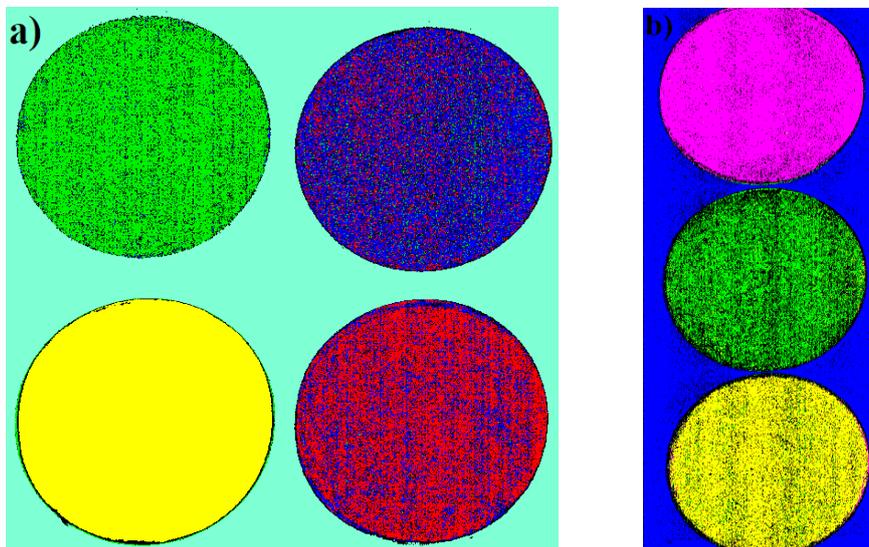


Figura 5.30. Clasificación e identificación de miembros finales según la abundancia de píxeles para: Suelos de Cusco y b) Suelos de Madre de Dios.

La figura 5.30 muestra una tonalidad de colores marcadamente diferentes para todas las muestras de los suelos, de Cusco y Madre de Dios, analizadas e incluso se logra distinguir

el fondo. Esto contrasta con las imágenes obtenidas usando los datos originales. Se puede observar la presencia de algunos píxeles de impurezas, como espectros de materiales distintos que están presentes en las muestras de suelos estudiadas.

Luego de obtener el mapa de distribución y abundancia, como resultado final se procede a la extracción y cuantificación de los espectros de los endmembers de los suelos.

#### 5.5.1.4 Extracción de espectros de los endmembers

Los resultados obtenidos hasta el apartado anterior, muestran que el método de análisis de imágenes PPI, es una herramienta muy eficiente para distinguir entre distintas muestras de suelos. Con el fin de comprobar la invariabilidad de los espectros de reflectancia obtenidas en la figura 5.22 y 5.23, se hace la extracción de los espectros de los endmembers (ver figura 5.31) luego de la distribución y abundancia de píxeles mostradas en la figura 5.29 y 5.30, a su vez la cuantificación en valores de área que se muestra en la tabla 5.9 y 5.10.

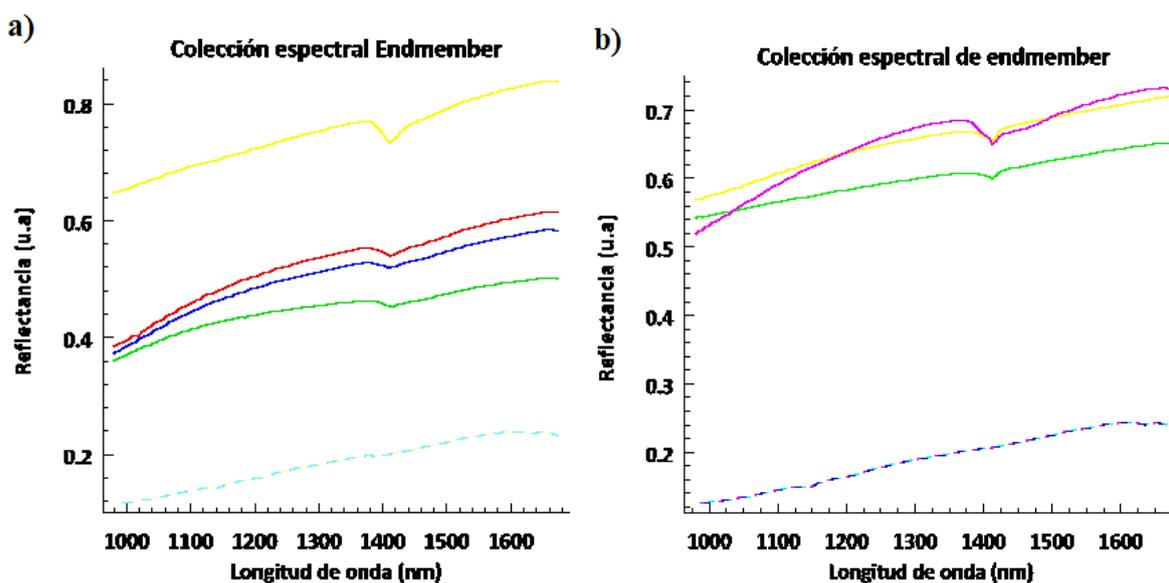


Figura 5.31. Firmas espectrales de los endmembers extraídos de las imágenes hiperespectrales de:  
a) Suelos de Cusco y b) Suelos de Madre de Dios.

Los resultados mostrados en la figura 5.31 muestran ser similares a los registrados para los datos originales o espectros estandarizados. En la parte a) y b) se distingue el espectro del fondo (líneas punteadas de color celeste y azul) respectivamente. A continuación, representaremos la abundancia de cada espectro por color en las siguientes tablas:

Tabla 5.9. Cuadro de estimación de abundancias obtenidos con el algoritmo PPI para los suelos de Cusco.

<b>Muestras</b>	<b>Color</b>	<b>Pixeles</b>	<b>% Área</b>
<b>Limite PPI</b>	Negro	359 796	100 %
<b>Cus1</b>	Rojo	49 978	13.89
<b>Cus2</b>	Azul	47 714	13.26
<b>Cus3</b>	Amarillo	57 527	15.99
<b>Cus4</b>	Verde	51 676	14.36
<b>Fondo</b>	Aguamarina	123 119	34.22
<b>Mezcla de Pixeles</b>	Varios	29 712	8.26
<b>Total pixel</b>		330 014	91.72
<b>Relativo</b>			

Tabla 5.10. Cuadro de estimación de abundancias obtenidos con el algoritmo PPI para suelos de Madre de Dios.

<b>Muestras</b>	<b>Color</b>	<b>Pixeles</b>	<b>% Área</b>
<b>Limite PPI</b>	Negro	264 187	100 %
<b>MD1</b>	Magenta	55 315	20.94
<b>MD2</b>	Verde	54 269	20.54
<b>MD3</b>	Amarillo	55 145	20.87
<b>Fondo</b>	Azul	89 337	33.81
<b>Mezcla Pixeles</b>	Varios	10 121	3.83
<b>Total pixel</b>		254066	96.17
<b>relativo</b>			

En estas dos tablas, la muestra Cus3 (amarillo) y MD1 (magenta) representan la mayor abundancia de tipo de suelo respecto al resultado obtenido mostradas en la figura 5.30, esto se debe a la pureza de dichos suelos. Por otro lado, pudo determinarse la abundancia del fondo, en ambas figuras representa la mayor parte de la imagen hiperespectral, puede ser relevante dicha información, si se considerara como una muestra o clase más.

El contraste entre las tablas iniciales 5.7 y 5.8 con las obtenidas, reflejan una similitud respecto a los valores obtenidos.

Por lo tanto, se puede afirmar o validar que dichos espectros corresponden a los suelos que inicialmente clasificamos.

## 5.5.2. Análisis de componentes principales (PCA) en suelos

### I) PCA \_ para suelos de Cusco y Madre de Dios

En primer lugar, se hará la reducción de dimensionalidad, para el cual usaremos el software R Studio, teniendo como resultado las siguientes imágenes del primer componente principal de ambos suelos.

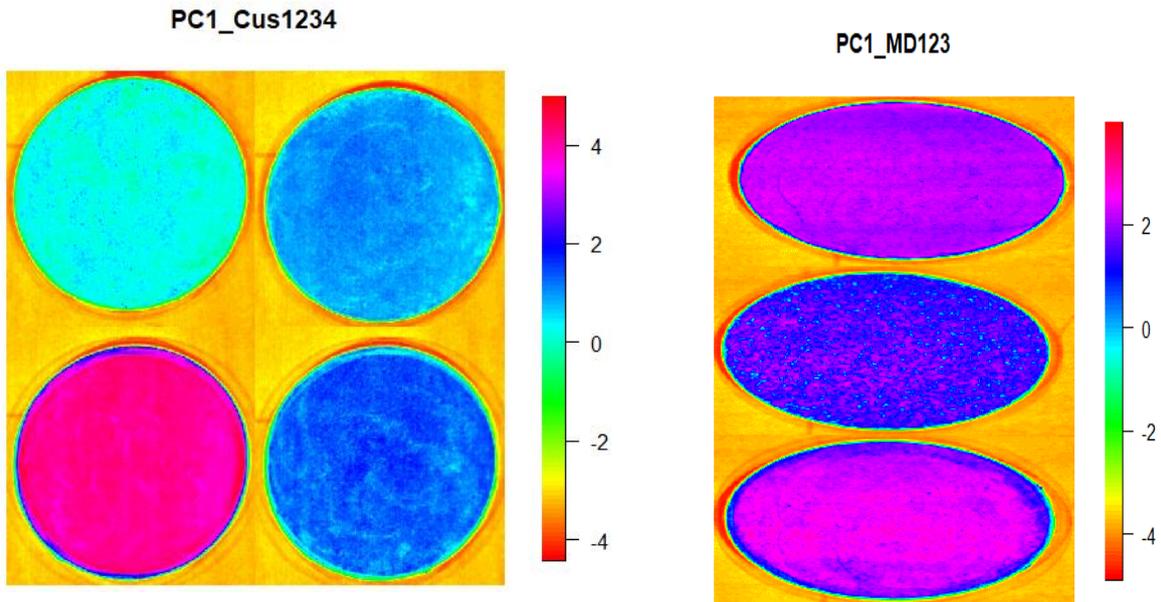


Figura 5.32. Pre procesamiento de la imagen hiperespectral de Cusco y Madre de Dios.

Se muestra los primeros PCA's de ambos suelos en falso color, donde se puede identificar visualmente los tipos de suelo presentes en ambas imágenes, tal como se aprecia en los dos gráficos. Esto haría suponer una clasificación anticipada de los suelos, sin embargo, no sabemos la variabilidad o el peso de cada tipo suelo en cada imagen.

Por ello, recurrimos al cálculo de los pesos o coeficientes de los primeros componentes principales ya que esto permite explicar una gran cantidad de variabilidad.

Ahora vamos a interpretar las cargas, es decir, los coeficientes de los componentes principales presentes en cada imagen y así determinar la escala de valores de cada PC. A continuación, se muestra la trama de esta.

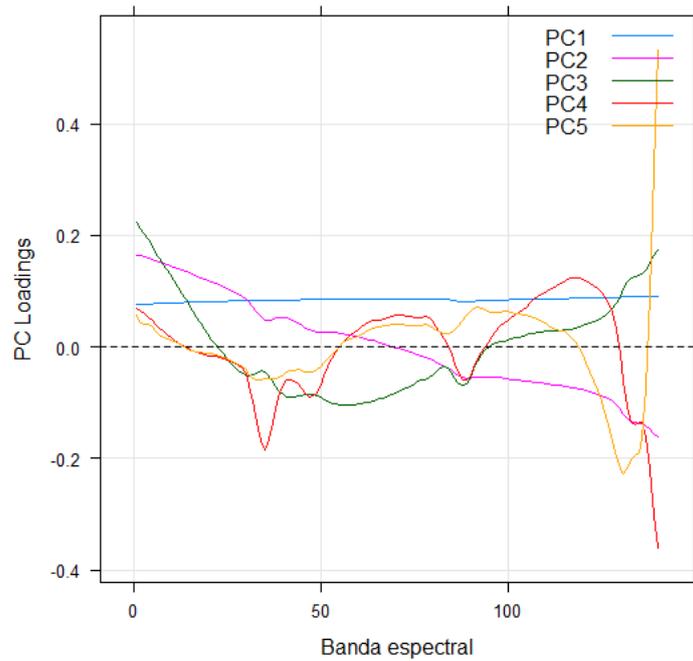


Figura 5.33. Muestra la carga de los primeros cinco componentes principales de los suelos de Cusco.

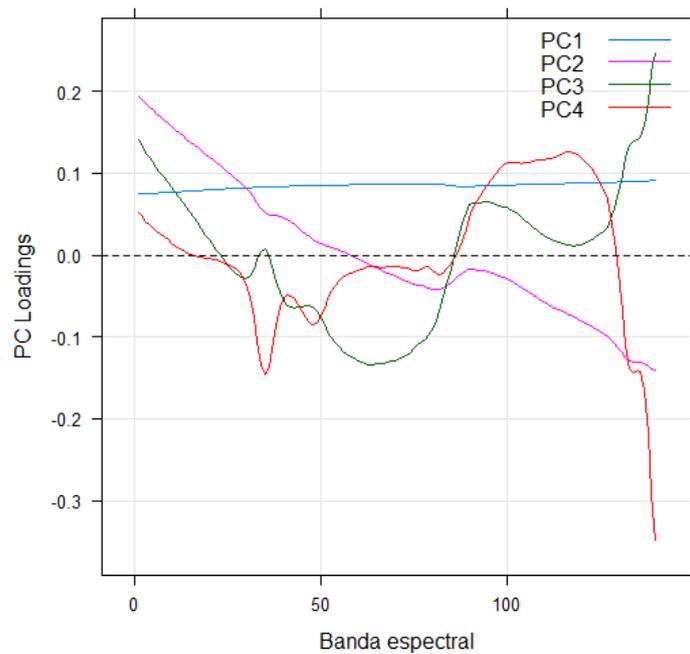


Figura 5.34. Muestra la carga de los primeros cuatro componentes principales de los suelos de Madre de Dios.

Los coeficientes del primer componente principal (PC1) en ambos caso son todos positivos y de casi la misma magnitud. Esto significa que el PC1 es (aproximadamente) proporcional al promedio de toda las bandas espectrales y representa una magnitud general de reflectancias en toda las bandas.

Ahora, introducimos la regla de “Porcentaje de variabilidad” que nos permitirá conocer la cantidad de bandas PC’s, con mayor variabilidad.

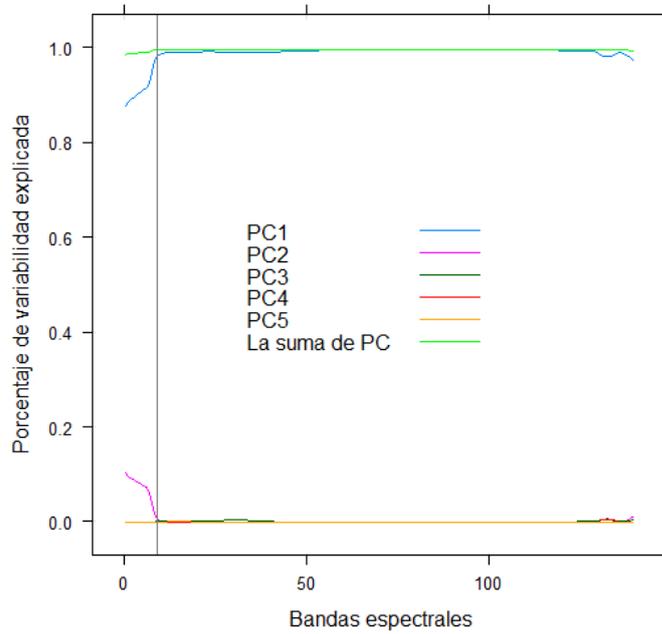


Figura 5.35. Porcentaje de variabilidad explicados por los primeros cinco componentes principales dentro de las bandas de la imagen de Cusco.

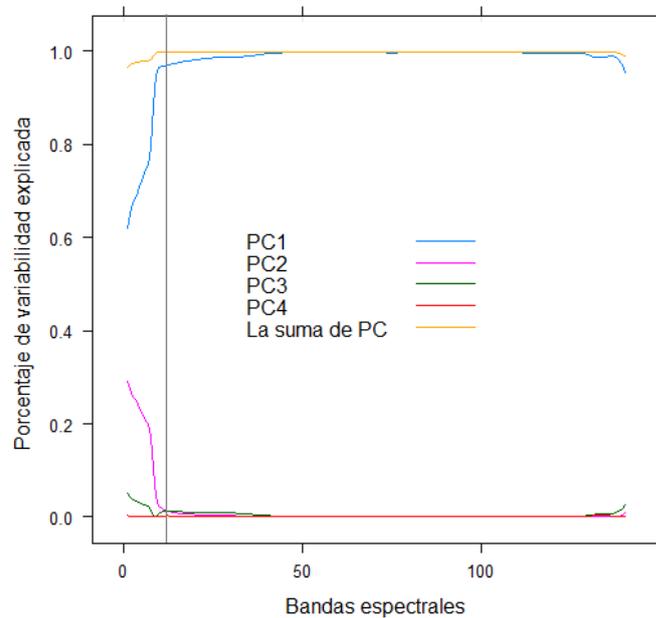


Figura 5.36. Porcentaje de variabilidad explicados por los primeros cuatro componentes principales dentro de las bandas de la imagen de Madre de Dios.

Los graficos anteriores muestra que el porcentaje de variabilidad explicada por el espectro de bandas, es mayor solo en las 6 primeras PCs en ambos casos de estudio.

Estos resultados lo podemos contrastar visualizando las imágenes de los 12 primeros PC's, tal como muestra la figura 5.36 y figura 5.37

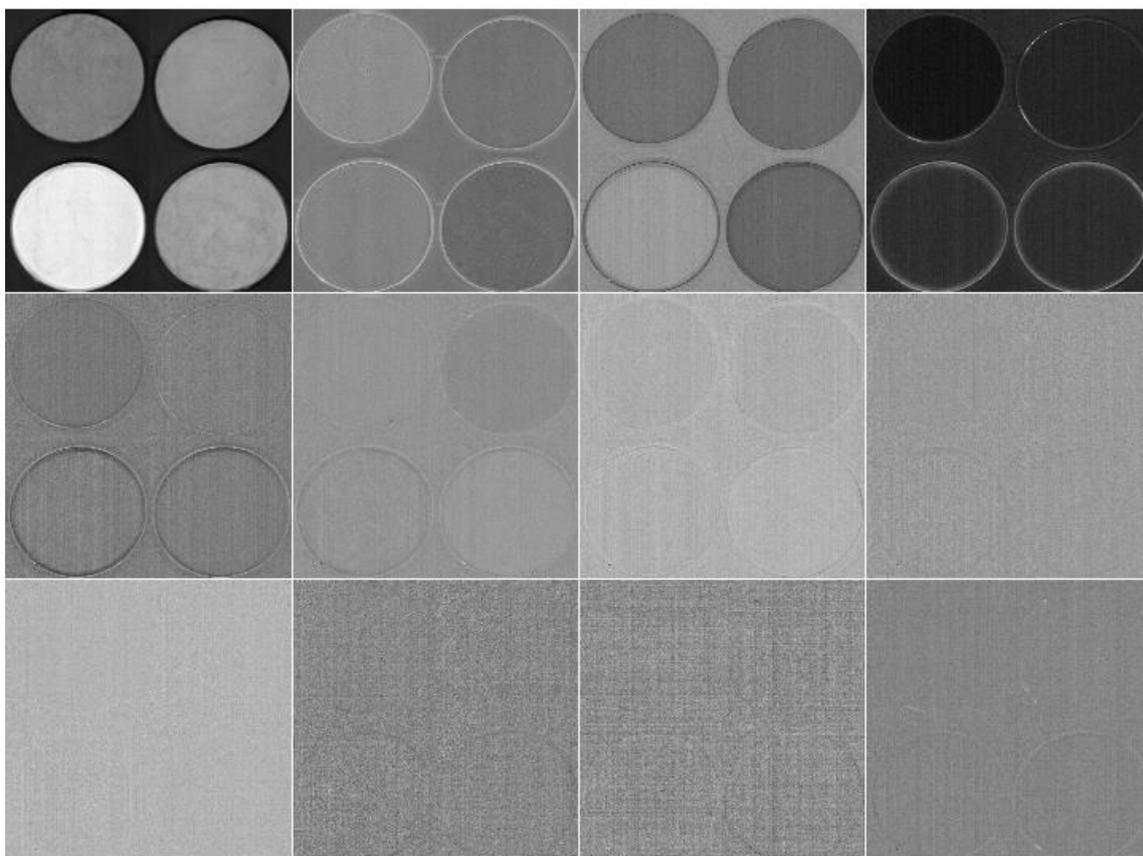


Figura 5.37. Visualización de izquierda a derecha de las 12 PCs de los suelos de Cusco.

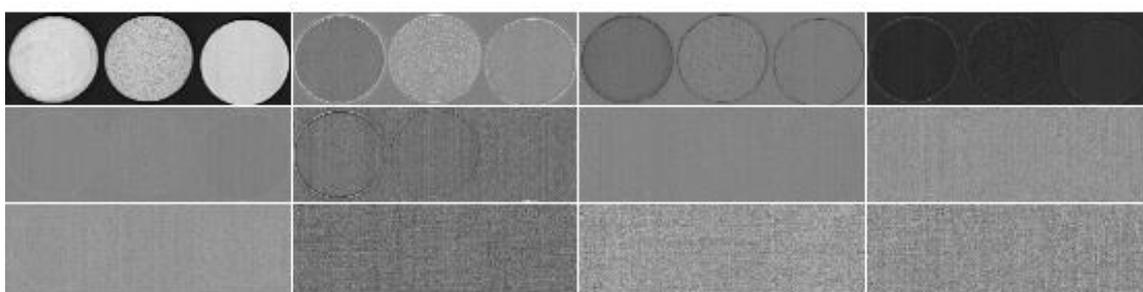


Figura 5.38. Visualización de izquierda a derecha de las 12 PCs de los suelos de Madre de Dios.

La figura 5.36, muestra claramente 5 bandas de componentes principales, lo que nos lleva a validar la regla de variabilidad empleada para predecir la cantidad de PC con mayor variabilidad. Del mismo modo la figura 5.37, muestra 4 bandas de componentes principales, esto demuestra que la regla de variabilidad es correcta.

Este proceso nos lleva a obtener dicho porcentaje de variabilidad explicada de los 7 primeros PCs en la siguiente tabla.

Tabla 5.11. Valores de Porcentaje de variabilidad PC para suelos de Cusco y su valor acumulado.

PC1	PC2	PC3	PC4	PC5	PC6	PC7
99.216	0.402	0.198	0.069	0.040	0.005	0.004
99.22	99.62	99.82	99.87	99.91	99.91	99.91

Tabla 5.12. Valores de Porcentaje de variabilidad PC para los suelos de Madre de Dios y su valor acumulado.

PC1	PC2	PC3	PC4	PC5
98.82	0.623	0.335	0.061	0.039
98.82	99.44	99.77	99.83	99.87

La variabilidad explicada de la primera componente, es el cuadrado de la primera desviación estándar. El segundo componente principal es el cuadrado de la segunda desviación estándar, y así sucesivamente.

A partir de los datos obtenidos en la tabla 5.11, se observa que las 7 primeras PC's contienen el 99.91% de toda la información de la imagen hiperespectral del suelo de Cusco. Del mismo modo se observa que en la tabla 5.12, se tiene las 5 primeras PC's conteniendo el 99.87% de información del suelo de Madre de Dios. Analizando los otros componentes PC2, PC3, PC4 etc... notamos que la información es alterada y en algunos casos hay presencia considerable de ruido.

Por lo tanto, estos componentes aportan información mínima, en algunos casos perjudican la visualización óptima de algún elemento importante de estudio, por ello solo se realizó el análisis y procesamiento del primer componente principal PC1 dejando de lado el resto de PC's.

El procesamiento del PC1 nos permite extraer el fondo, gracias a los valores de las unidades (scores) que se muestra en la figura 5.32, este proceso se realiza analizando histogramas de frecuencia para las dos imágenes, eliminando los scores diferentes a los elementos analizados.

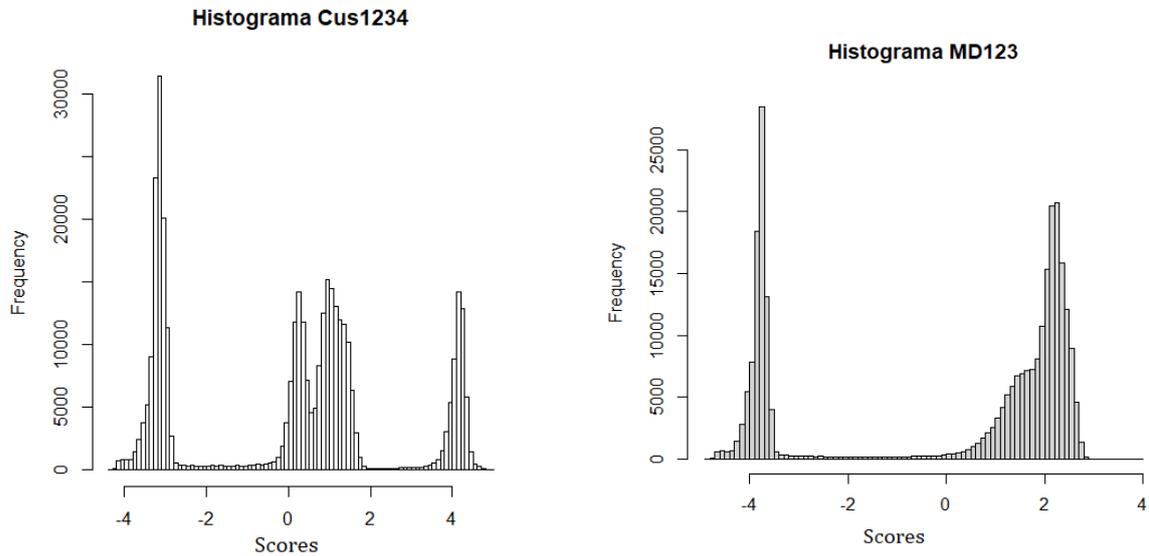


Figura 5.39. Histogramas de frecuencia de los scores de los suelos de Cusco y Madre de Dios.

Esta extracción del fondo (scores en el rango -4 a -3, en ambas figuras) del primer componente nos permitirá conocer los elementos distintos en cada imagen hiperespectral.

A continuación, se muestra el resultado del PC1 sin fondo, listo para la identificación de elementos distintos en cada imagen.

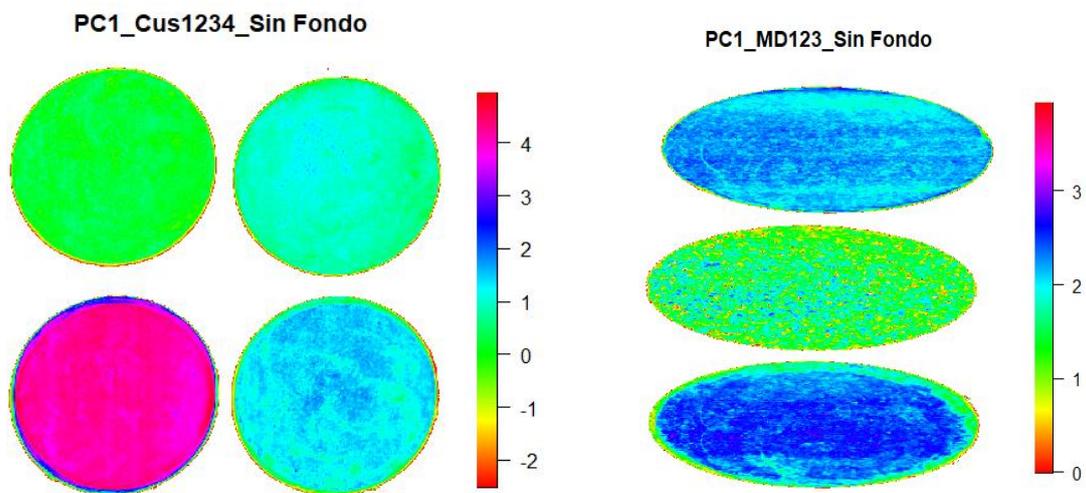


Figura 5.40. Imágenes hiperespectrales de los suelos de Cusco y Madre de Dios sin fondo.

Mediante la clasificación supervisada y utilizando herramientas en el software R studio, seleccionamos las regiones de interés (ROI's) en las imágenes hiperespectrales sin fondo de ambos suelos, esto con la finalidad de observar los espectros característicos de cada tipo de suelo.

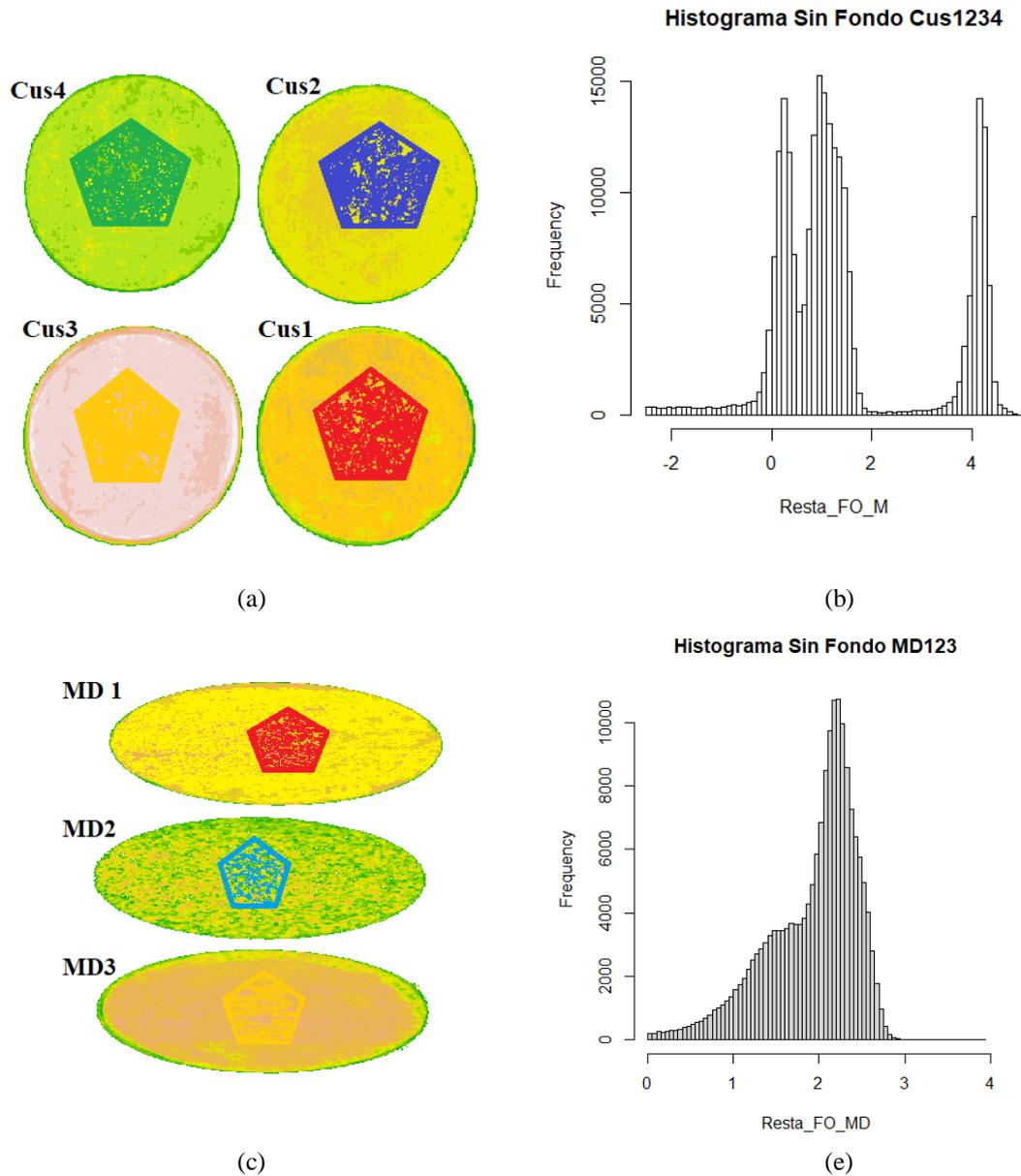
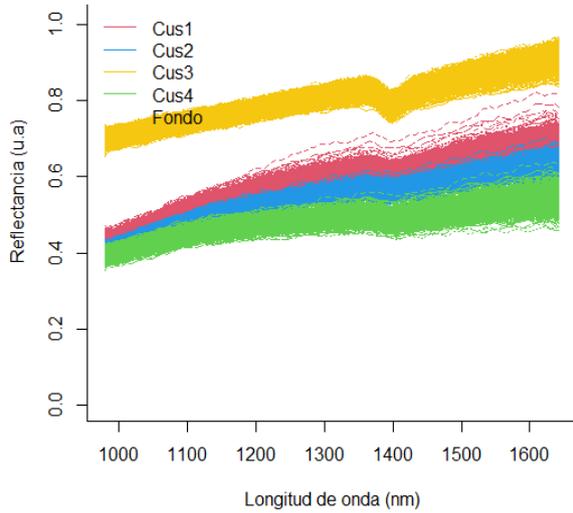


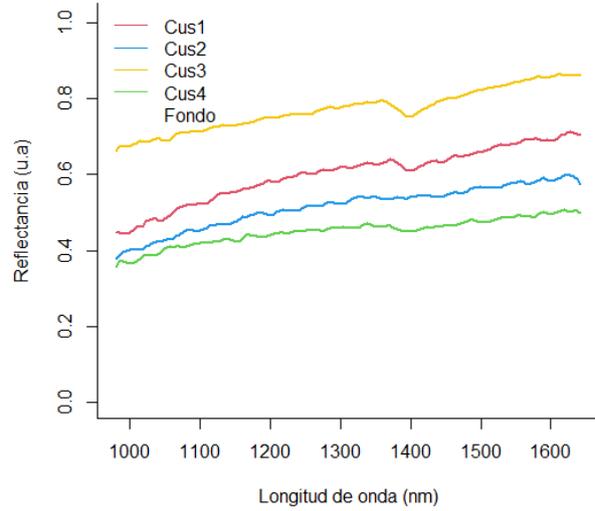
Figura 5.41. a) y c) Selección de las regiones de interés (ROI's) mediante clasificación supervisada de cada tipo de suelo de Cusco y Madre de Dios. b) y e) Histogramas de cada imagen hiperespectral sin fondo.

En ambas figuras a) y c) se observa la selección arbitraria de regiones que agrupan pixeles característicos de cada tipo de suelo, a su vez nos permitirá contabilizar o mostrar la abundancia basados en la cantidad promedio de espectros. Las figuras b) y e) muestran los

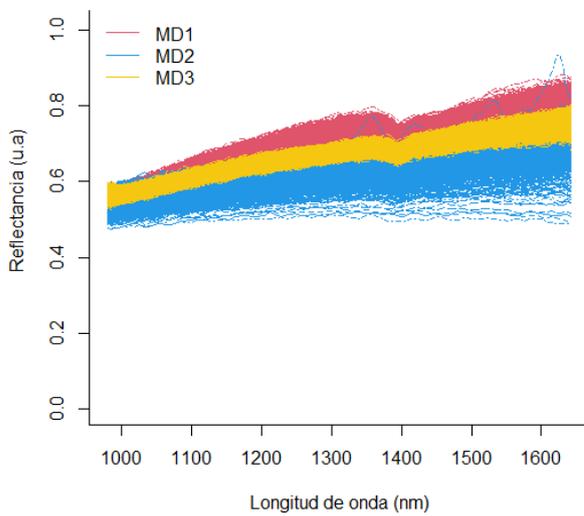
histogramas de los scores conteniendo solo información de los suelos, se observa la eliminación completa de los scores del fondo. En la figura 5.41, veremos la cantidad de firmas espectrales de cada región de interés seleccionado en la figura 5.40, que a su vez nos permitirá determinar la abundancia de cada tipo de suelo.



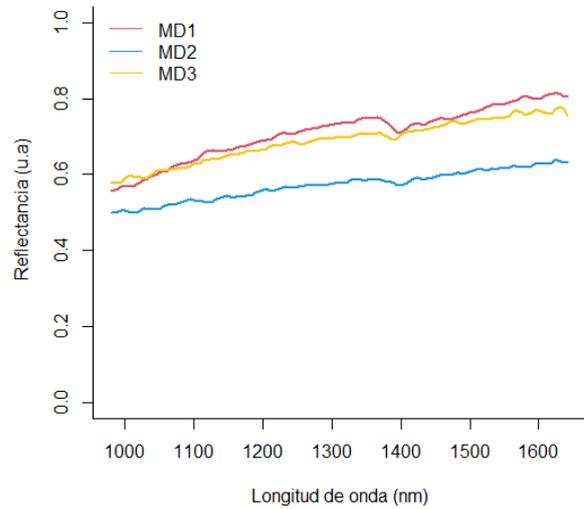
(a)



(b)



(c)



(d)

Figura 5.42. a) y c) Grupo de firmas espectrales de cada región de interés de cada suelo.

b) y d) Firmas espectrales representativas de cada tipo de suelo.

La figura a) y c) representa el grupo de firmas espectrales solo de las regiones de interés seleccionadas, dicho procedimiento se replica hasta cubrir la superficie de cada suelo con el

objetivo de contabilizar y extraer las firmas espectrales representativas.

Los espectros de la figura 5.41. (b) y (d) son los resultados finales de la clasificación de los suelos aplicando el algoritmo del PCA, cuyas bandas de absorción de cada espectro se encuentran en el mismo rango que los espectros estandarizados en la figura 5.22 y 5.23, cabe resaltar que en el PCA estas firmas son más planas y con menos profundidad de absorción de banda.

Los suelos Cus3 y MD1 conservan la banda característica, la amplitud es la misma que las firmas estandarizadas y las encontradas con el algoritmo PPI.

Los suelos Cus2, Cus4 y MD2 muestran espectros con una leve o casi nula banda de absorción, esto puede darse a diferentes factores, uno de ellos puede ser el haber promediado espectros distintos o de ruido.

Por otra parte, las regiones de interés (ROI's) permite cuantificar la abundancia (ver tabla 5.13 y 5.14) de cada tipo de suelo presente en cada imagen hiperespectral del PC1.

Tabla 5.13. Cuadro de estimación de abundancias obtenidos con el PC1 para los suelos de Cusco.

<b>Muestras</b>	<b>Color</b>	<b>Pixeles</b>	<b>% Área</b>
<b>PC1 – con fondo</b>	-	360 000	100 %
<b>PC1 – sin fondo</b>	-	236 880	65.8 %
<b>Cus1</b>	Rojo	49 265	13.69
<b>Cus2</b>	Azul	46 889	13.03
<b>Cus3</b>	Amarillo	56 987	15.83
<b>Cus4</b>	Verde	51 245	14.24
<b>Fondo</b>	-	-	-
<b>Mezcla de Pixeles</b>	Varios	32 494	9.026

Tabla 5.14. Cuadro de estimación de abundancias obtenidos con el PC1 para suelos de Madre de Dios.

<b>Muestras</b>	<b>Color</b>	<b>Pixeles</b>	<b>% Área</b>
<b>PC1 – con fondo</b>	-	270 000	100 %
<b>PC1 – sin fondo</b>	-	169 328	62.71%
<b>MD1</b>	Magenta	52 450	19.42
<b>MD2</b>	Verde	51 859	19.21
<b>MD3</b>	Amarillo	53 785	19.92
<b>Fondo</b>	-	-	-
<b>Mezcla Pixeles</b>	Varios	11 234	4.16

Los resultados de ambas tablas muestran la abundancia de cada tipo de suelo presente en las imágenes, para su mejor análisis haremos una comparativa de resultados entre los métodos aplicados.

### 5.5.3. Comparación de resultados de cada modelo

En la búsqueda de verificar el método adecuado en esta clasificación de suelos, se obtuvo resultados contundentes respecto a la comparativa con los valores estandarizados de cada suelo, la tabla 5.15 muestra dichos resultados.

Tabla 5.15. Cuadro comparativo de estimación de abundancias entre los diferentes métodos de clasificación.

Muestras/ algoritmo	Estandarizado %	PPI %	PCA %
<b>Cus1</b>	15.98	13.89	13.69
<b>Cus2</b>	16.06	13.26	13.03
<b>Cus3</b>	16.01	15.99	15.83
<b>Cus4</b>	15.90	14.36	14.24
<b>MD1</b>	21.15	20.94	19.42
<b>MD2</b>	21.13	20.54	19.21
<b>MD3</b>	21.49	20.87	19.92

Los valores que se visualizan en la tabla 5.15, están en porcentajes de área, es la representación de la abundancia en espectros que tiene cada suelo en cada imagen hiperespectral.

Se logra verificar que para el suelo Cus3 la abundancia obtenida por los dos métodos PPI y PCA son muy próximos a su valor estandarizado, del mismo modo el suelo MD1 presenta la misma tendencia.

Mientras que los suelos Cus2 y MD2, tienen una baja semejanza al valor estandarizado respectivamente, esto debido a que estos suelos contienen elementos o componentes muy heterogéneos.

En el caso del suelo MD2 se logró hasta identificar residuos de galena y calcita, en el suelo Cus2 se logró identificar residuos de vegetación, adicional a ello este suelo se asemeja más al suelo Cus4, debido a que ambos son de color oscuro.

### 5.5.4. Correlación

En esta instancia es importante analizar el signo y la tendencia de las variables obtenidas a partir de los algoritmos, por lo que aplicamos el cálculo del coeficiente de correlación lineal de Pearson que sirve para cuantificar tendencias lineales [94]. En las tablas 5.16 y 5.17, hacemos el cálculo de dichos coeficientes (valores estandarizados vs PPI) y (valores estandarizados vs PCA).

Tabla 5.16. Cuadro comparativo de estimación de abundancias entre valores estandarizados versus el algoritmo de PPI.

Muestras/ algoritmo	Estandarizado	PPI	%Error relativo	Correlación
Cus1	15.98	13.89	13%	<b>97.05 %</b>
Cus2	16.06	13.26	17%	
Cus3	16.01	15.99	0.1%	
Cus4	15.90	14.36	9.7%	
MD1	21.15	20.94	1.0%	
MD2	21.13	20.54	2.7%	
MD3	21.49	20.87	2.9%	

Tabla 5.17. Cuadro comparativo de estimación de abundancias entre valores estandarizados versus el PCA.

Muestras/ algoritmo	Estandarizado	PCA	%Error relativo	Correlación
Cus1	15.98	13.69	14.3%	<b>95.71 %</b>
Cus2	16.06	13.03	18.9%	
Cus3	16.01	15.83	1.1%	
Cus4	15.90	14.24	10.4%	
MD1	21.15	19.42	8.2%	
MD2	21.13	19.21	9.1%	
MD3	21.49	19.92	7.3%	

Los porcentajes de variabilidad obtenidos al aplicar el PPI son menores a los hallados con el PCA, dicha diferencia muestra un valor de correlación cercano al 100% o a 1, por lo que podemos decir que la correlación es muy fuerte para el algoritmo PPI con 97.05%.

Los suelos con mayor correlación son Cus3, MD1, MD3 y Cus1 respectivamente, mientras las de menor correlación son Cus2, Cus4 y MD2.

## 5.6. Discusión de resultados

Al analizar a los suelos en diferentes estados (natural, molido y empastillado) notamos que hay una banda (1415 nm) de absorción con una mayor amplitud para los suelos en estado natural, dicha banda se redujo al secarlo y moler los suelos.

La reducción de dimensionalidad con el MNF (figura 5.26 y 5.27) frente a la del PCA (figura 5.36 y 5.37) muestra claramente un mejor blanqueamiento de ruido, esto ayuda a no perder información que podría confundirse con ruido, el PCA solo promedia valores semejantes a la tendencia.

El considerar el fondo como una clase o componente adicional, el PPI y PCA muestran la misma tendencia en cuanto a variabilidad, pero era conveniente considerar solo los componentes de interés, el algoritmo de PPI utilizando el software ENVI procesa toda la información de la imagen considerando el fondo, en cambio utilizando el PCA con el software R Studio, permite eliminar los valores espectrales del fondo, tal como se muestra en las figuras 5.25 y figura 5.38.

Los métodos aplicados a la identificación y clasificación de suelos cumplen la función de identificar, diferenciar los tipos de suelos y algunos otros elementos presentes en la muestra. En la tabla 5.15, se registró los porcentajes de variabilidad de ambos algoritmos, siendo el algoritmo PPI el que mostró mejores resultados. Esto se corroboró en los cuadros comparativos de estimación de abundancias hallando los coeficientes de correlación, donde el algoritmo de PPI con un 97.05% muestra una fuerte correlación frente al PCA que tiene 95.71%.

Estos valores de correlación también validan los resultados de clasificación mostrados en la figura 5.30, comparándolos con la figura 5.38, que de manera cualitativa se ve una clasificación de suelos respecto a la tonalidad de colores que cada algoritmo asigna a cada suelo, donde una vez más el algoritmo de PPI muestra una mejor clasificación de los suelos.

Del mismo modo sucede para los espectros, el algoritmo PPI resalta por su similitud de firmas espectrales con cada banda de absorción (figura 5.31 y figura 5.19), en todos estos casos hacemos tal comparativa con las imágenes hiperespectrales de los suelos molidos.

## CONCLUSIONES

- Los resultados obtenidos permiten demostrar que a partir de la adquisición, procesamiento y análisis de imágenes hiperespectrales, es posible identificar y clasificar suelos procedentes de diferentes regiones.
- Este trabajo propone un esquema nuevo de analizar e identificar suelos usando imágenes hiperespectrales, ya sea suelos en su estado natural con abundante humedad, secos o molidos.
- Mediante la preparación de las muestras (secado y empastillado) es posible eliminar las bandas de absorción del agua, y reducir la cantidad de ruido o píxeles de sombra.
- Los espectros de reflectancia de las muestras estudiadas muestran que la absorción de las muestras a 1415 nm es una característica que permite diferenciar las diferentes muestras de suelos estudiados.
- Los espectros de reflectancia obtenidos corresponden a suelos esperados como se validó al compararlos con suelos de la librería Áster.
- El PCA permite un mejor manejo de información, pero en el proceso de tratamiento de reducción de dimensionalidad muestra una leve desventaja frente al MNF utilizado por el PPI. El algoritmo PPI mostrado presenta una mejor performance al momento de diferenciar las diferentes muestras en una imagen, con una correlación fuerte de 97.01%.
- Por lo tanto, los resultados obtenidos y la aplicabilidad de las HSI, el método del PPI es un método completamente viable para el análisis e identificación de suelos en diferentes condiciones, por lo que en primera instancia sirve para detectar elementos contaminantes en los suelos o ver el tipo de suelo que representa.
- Como continuación de este trabajo, la técnica de HSI se aplicará al campo biomédico en cooperación con la PUCP.
- En un trabajo posterior se puede incrementar la base de datos de imágenes hiperespectrales de otros suelos del Perú. Así mismo a partir de la identificación de los suelos puros se puede determinar la presencia de agentes contaminantes en los suelos.

## BIBLIOGRAFIA

- [1] H. R. Condit, "Spectral Reflectance of American Soils," *Photogramm Eng*, vol. 36, no. 9, pp. 955–966, 1970.
- [2] E. R. Stoner and M. F. Baumgardner, "Characteristic variations in reflectance of surface soils.," *Soil Science Society of America Journal*, vol. 45, no. 6, pp. 1161–1165, 1981.
- [3] E. Ben-Dor and A. Banin, "Visible and near-infrared (0.4-1.1  $\mu\text{m}$ ) analysis of arid and semiarid soils," *Remote Sensing of Environment*, vol. 48, no. 3, pp. 261–274, 1994.
- [4] J. Hill and B. Schütt, "Mapping complex patterns of erosion and stability in dry mediterranean ecosystems," *Remote Sensing of Environment*, vol. 74, no. 3, pp. 557–569, 2000.
- [5] E. Ben-Dor, K. Patkin, A. Banin, and A. Karnieli, "Mapping of several soil properties using DAIS-7915 hyperspectral scanner data - A case study over soils in Israel," *International Journal of Remote Sensing*, vol. 23, no. 6, pp. 1043–1062, 2002.
- [6] M. R. Nanni, J. A. M. Demattê, and P. R. Fiorio, "Análise discriminante dos solos por meio da resposta espectral no nível terrestre," *Pesquisa Agropecuaria Brasileira*, vol. 39, no. 10, pp. 995–1006, 2004.
- [7] FAO, *World 's Soil Resources*. 2015.
- [8] C. Canet and A. Camprubí, "Yacimientos minerales: Los Tesoros de la Tierra," *Boletín de la Sociedad Geológica Mexicana*, vol. 59, no. 2, pp. 221–222, 2007.
- [9] F. Van Der Meer, "Can we map swelling clays with remote sensing?," *ITC Journal*, vol. 1, no. 1, pp. 27–35, 1999.
- [10] M. F. Baumgardner, L. R. F. Silva, L. L. Biehl, and E. R. Stoner, "Reflectance properties of soils," *Advances in Agronomy*, vol. 38, no. C, pp. 1–44, 1986.
- [11] A. B. McBratney, M. L. Mendonça Santos, and B. Minasny, *On digital soil mapping*, vol. 117, no. 1–2. 2003.
- [12] P. Krishnan, J. D. Alexander, B. J. Butler, and J. W. Hummel, "Reflectance Technique for Predicting Soil Organic Matter<sup>1</sup>," *Soil Science Society of America Journal*, vol. 44, no. 6, p. 1282, 1980.
- [13] R. C. Dalal and R. J. Henry, "Simultaneous determination of moisture, organic carbon, and total nitrogen by near infrared reflectance spectrophotometry.," *Soil*

- Science Society of America Journal*, vol. 50, no. 1, pp. 120–123, 1986.
- [14] E. Ben-Dor and A. Banin, “Near-infrared analysis as a rapid method to simultaneously evaluate several soil properties,” *Soil Science Society of America Journal*, vol. 59, no. 2, pp. 364–372, 1995.
- [15] C.-W. Chang, A. L. David, J. M. Maurice, and R. H. Charles, “Analyses of Soil Properties,” *Soil Science Society of America Journal*, pp. 480–490, 2001.
- [16] K. D. Shepherd and M. G. Walsh, “Development of reflectance spectral libraries for characterization of soil properties,” *Soil Science Society of America Journal*, vol. 66, no. 3, pp. 988–998, 2002.
- [17] D. J. Brown, K. D. Shepherd, M. G. Walsh, M. Dewayne Mays, and T. G. Reinsch, “Global soil characterization with VNIR diffuse reflectance spectroscopy,” *Geoderma*, vol. 132, no. 3–4, pp. 273–290, 2006.
- [18] T. Terhoeven-Urselmans, H. Schmidt, R. Georg Joergensen, and B. Ludwig, “Usefulness of near-infrared spectroscopy to determine biological and chemical soil properties: Importance of sample pre-treatment,” *Soil Biology and Biochemistry*, vol. 40, no. 5, pp. 1178–1188, 2008.
- [19] R. A. V. Rossel and T. Behrens, “Using data mining to model and interpret soil diffuse reflectance spectra,” *Geoderma*, vol. 158, no. 1–2, pp. 46–54, 2010.
- [20] R. A. Viscarra Rossel *et al.*, “A global spectral library to characterize the world’s soil,” *Earth-Science Reviews*, vol. 155, pp. 198–230, 2016.
- [21] X. L. Xie and A. B. Li, “Identification of soil profile classes using depth-weighted visible–near-infrared spectral reflectance,” *Geoderma*, vol. 325, no. November 2017, pp. 90–101, 2018.
- [22] F. Chaudhry, C.-C. Wu, W. Liu, C.-I. Chang, and A. Plaza, *Pixel purity index-based algorithms for endmember extraction from hyperspectral imagery*, vol. 661, no. 2. 2006.
- [23] INEI (Instituto Nacional de Estadística e Informática), “Anuario de estadísticas ambientales,” *The British Journal of Psychiatry*, vol. 111, no. 479, p. 594, 2015.
- [24] W. Rodríguez and P. García, “Aplicaciones de Técnicas Espectroscópicas para el Análisis de Suelos,” vol. 12, no. 2, pp. 228–251, 2016.
- [25] Resonon, “SpectrononPro Manual,” p. 135, 2017.
- [26] W. M. Castro Silipu, “Aplicación de la tecnología de imágenes hiperespectrales al control de calidad de productos agroalimentarios de la región de Amazonas (Perú),”

- 2016.
- [27] Y. Qi, J. L. Darilek, B. Huang, Y. Zhao, W. Sun, and Z. Gu, “Evaluating soil quality indices in an agricultural region of Jiangsu Province, China,” *Geoderma*, vol. 149, no. 3–4, pp. 325–334, 2009.
- [28] C. Castilla, A. Gutiérrez, and L. Ramírez, “Taller Latinoamericano Globalsoilmap.net Atlas de Suelos LAC. Red de Cultivos Permanentes. Rio de Janeiro,” p. 26, 2010.
- [29] MTC (Ministerio de Transportes y Comunicaciones), “Manual de ensayo de materiales,” vol. I, 2016.
- [30] J. Gómez Sanchis, “DETECCIÓN AUTOMÁTICA DE PODREDUMBRES EN CÍTRICOS MEDIANTE PROCESADO AVANZADO DE IMÁGENES HIPERESPECTRALES.,” *Universidad de Valencia*, vol. 29, no. 14, pp. 235–246, 2010.
- [31] V. Hanna Castro, “EVALUACIÓN DE TECNOLOGÍAS HIPERESPECTRALES EN LA CARACTERIZACIÓN MINERAL DE YACIMIENTOS PARA APLICACIONES GEOMETALÚRGICAS: CASO APLICADO A MINA FLORIDA, DISTRITO MINERO ALHUÉ, REGIÓN METROPOLITANA, CHILE,” 2017.
- [32] P. Bajorski, “STATISTICS FOR IMAGING, OPTICS, AND PHOTONICS,” 2011.
- [33] C. I. Chang, *Hyperspectral Data Processing: Algorithm Design and Analysis*. 2013.
- [34] A. M. Baldridge, S. J. Hook, C. I. Grove, and G. Rivera, “The ASTER spectral library version 2.0,” *Remote Sensing of Environment*, vol. 113, no. 4, pp. 711–715, 2009.
- [35] A. Linzt, J and D. S. Simonett, “Remote sensing of environment,” pp. 609–611, 1976.
- [36] J. B. Campbell and R. H. Wynne, *Introduction to REMOTE SENSING*. 2011.
- [37] F. F. Sabins, “Remote Sensing: Principles and Interpretation,” pp. 635–635, 1996.
- [38] M. Borengasser, W. S Hungate, and R. Watkins, “Hyperspectral remote sensing : principles and applications,” *Current Science*, vol. 94, no. 9, pp. 1115–1116, 2008.
- [39] B. Valdés Moreno, “Identificación de lineamientos y minerales asociados a alteraciones hidrotermales en la Bahía de la Paz, Baja California Sur, México, mediante imágenes satelitales Landsat ETM,” p. 81, 2014.
- [40] M. Hongn, S. Flores Larsen, M. Altamirano, M. Gea, and L. Saravia, “Estudio De

- Reflectancia De Espejos Para Un Concentrador Solar Fresnel Lineal,” *Avances en Energías Renovables y Medio Ambiente*, vol. 15, pp. 191–197, 2011.
- [41] J. M. A. Gallego, “Espectrometría de imagen en el infrarrojo: una contribución al estudio de las combustiones y a la teledetección de incendios forestales,” p. 4545, 2002.
- [42] J. Manuel Aranda Gallego and J. Meléndez Sánchez Fernando López Martínez, “Espectrometría de imagen en el infrarrojo: una contribución al estudio de las combustiones y a la teledetección de incendios forestales,” 2002.
- [43] B. Park and R. Lu, *Hyperspectral imaging technology in food and agriculture*. 2015.
- [44] Resonon Inc., “SpectrononPro Manual Release 5.0,” 2016.
- [45] M. Antonio Plaza, “Proposición, Validación y Prueba de una Metodología Morfológica para el Análisis de Datos Hiperespectrales que Integra Información Espacial y Espectral,” 2002.
- [46] E. S. Bernabé, “Procesado de imágenes hiperespectrales,” pp. 65–91, 2016.
- [47] G. ElMasry, M. Kamruzzaman, D. W. Sun, and P. Allen, “Principles and Applications of Hyperspectral Imaging in Quality Evaluation of Agro-Food Products: A Review,” *Critical Reviews in Food Science and Nutrition*, vol. 52, no. 11, pp. 999–1023, 2012.
- [48] D. Ayala Martini, “Automatización del análisis de imágenes hiperespectrales para identificación de aptitud de patatas,” Universidad Pública de Navarra, Pamplona, España, 2018.
- [49] J. Hernández, “Características de los sensores hiperespectrales para el sensoramiento remoto,” *HySpex*, 2015.
- [50] E. S. Bernabé, “Procesado de imágenes hiperespectrales,” *ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA Y SISTEMAS DE TELECOMUNICACIÓN*, pp. 65–91, 2016.
- [51] R. O. Green *et al.*, “Imaging spectroscopy and the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS),” *Remote Sensing of Environment*, vol. 65, no. 3, pp. 227–248, 1998.
- [52] C. Gonzalez Calvo, “Procesamiento a bordo de imágenes hiperespectrales de la superficie terrestre mediante hardware reconfigurable,” p. 219, 2011.
- [53] R. Smith, “Introduction to hyperspectral imaging,” *Microimages*. Retrieved on June, pp. 1–24, 2006.

- [54] E. Ivorra Martínez, “Desarrollo de técnicas de visión hiperespectral y tridimensional para el sector agroalimentario,” p. 202, 2015.
- [55] M. H. Jafari, E. Nasr-Esfahani, N. Karimi, S. M. R. Soroushmehr, S. Samavi, and K. Najarian, “Extraction of Skin Lesions from Non-Dermoscopic Images Using Deep Learning,” *arXiv*, pp. 10–11, 2016.
- [56] A. Esteva *et al.*, “Dermatologist-level classification of skin cancer with deep neural networks.,” *Nature*, vol. 542, no. 7639, pp. 115–118, 2017.
- [57] R. X. Xu *et al.*, “Developing digital tissue phantoms for hyperspectral imaging of ischemic wounds.,” *Biomedical optics express*, vol. 3, no. 6, pp. 1433–45, Jun. 2012.
- [58] O. Marrero, Gustavo - Sarmiento, “PROYECTO HELICOID: detección de tumores cerebrales mediante imágenes hiperespectrales,” 2018.
- [59] S. C. Jay, R. L. Lawrence, K. S. Repasky, and L. J. Rew, “Detection of leafy spurge using hyper-spectral-spatial-temporal imagery,” *International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 4374–4376, 2010.
- [60] M. Kanning, I. Kühling, D. Trautz, and T. Jarmer, “High-resolution UAV-based hyperspectral imagery for LAI and chlorophyll estimations from wheat for yield prediction,” *Remote Sensing*, vol. 10, no. 12, pp. 1–17, 2018.
- [61] “Centro Nacional de Investigación y Educación de Great Rivers :: Inicio.” [Online]. Available: <http://www.ngrrrec.org/>. [Accessed: 02-Mar-2020].
- [62] N. G. L. E. R. L. C. I. for G. L. R. US Department of Commerce, “Hyperspectral imaging.”
- [63] D. Slaughter, “Nondestructive Maturity Assessment Methods for Mango: A Review of Literature and Identification of Future Research Needs,” *University of California, Davis*, pp. 1–18, 2009.
- [64] “Aplicaciones de cámaras de imágenes hiperespectrales | Estudios de caso de Resonon.” [Online]. Available: <https://resonon.com/applications>. [Accessed: 02-Mar-2020].
- [65] FAO - UNESCO, *Mapa mundial de suelos. Volumen IV América del Sur*. 1971.
- [66] P. R. Steinhagen and D. C. Baumgart, “Grundlagen des Mikrobioms,” *Internist*, vol. 58, no. 5, pp. 429–434, 2017.
- [67] J. Castellanos, “Propiedades físicas del suelo y el crecimiento de las plantas,” *Intagri*, no. Cuadro 1, 2014.

- [68] W. W. EMERSON, “the Structure of Soil Crumbs,” *Journal of Soil Science*, vol. 10, no. 2, pp. 235–244, 1959.
- [69] O. Piedrahita, “Introducción pH del Suelo,” *Magnesio Heliconia S.A.*, vol. 14, no. cuadro 1, 2009.
- [70] J. Castellanos, “La salinidad de los suelos, un problema que amenaza su fertilidad,” *México: Instituto para la Innovación Tecnológica en Agricultura, Intagri.*, 2016.
- [71] R. Casas-Flores, “El suelo de cultivo y las condiciones climáticas,” p. 236, 2012.
- [72] USDA NRCS, “United States Department of Agriculture Natural Resources Conservation Service. Soil Taxonomy: A Basic System of Soil Classification for Making and Interpreting Soil Surveys,” *USDA Natural Resources Conservation Service*, vol. 436, no. Edition 2, 1999.
- [73] O. Mehl, H. ; Peinado, “Introducción a la reflectancia espectral de la Vegetación , Suelo Y Agua,” no. 1987, 1987.
- [74] S. y M. F. B. Eric, “Physiochemical, site, and bidirectional reflectance factor characteristics of uniformly moist soils,” vol. 69, no. 1, pp. 1–16, 1980.
- [75] “Minagri, 20.” [Online]. Available: <http://minagri.gob.pe/portal/objetivos/43-sector-agrario/suelo/330-clasificacion>. [Accessed: 25-Apr-2019].
- [76] A. A. Villalobos Pinzón, “Procesamiento digital de imágenes hiperespectrales,” Universidad de Sevilla, 2015.
- [77] S. T. Monteiro, “Hyperspectral Image Classification of Grass Species in Northeast Japan,” no. August, 2008.
- [78] F. Cao, Z. Yang, J. Ren, M. Jiang, and W. Ling, “Does Normalization Methods Play a Role for Hyperspectral Image Classification ?,” no. 2, pp. 2–7.
- [79] A. A. Naeini, M. Babadi, and S. Homayouni, “ASSESSMENT OF NORMALIZATION TECHNIQUES ON THE ACCURACY OF HYPERSPECTRAL DATA CLUSTERING,” vol. XLII, no. October, pp. 7–10, 2017.
- [80] Y. Li, W. Xie, and H. Li, *Author ’ s Accepted Manuscript Hyperspectral image reconstruction by deep To appear in : Pattern Recognition network for classification*. Elsevier, 2016.
- [81] W. Dai, I. Selesnick, J.-R. Rizzo, J. Rucker, and T. Hudson, “A nonlinear generalization of the Savitzky-Golay filter and the quantitative analysis of saccades,” *Journal of Vision*, vol. 17, no. 9, p. 10, 2017.

- [82] J. Chen, P. Jönsson, M. Tamura, Z. Gu, B. Matsushita, and L. Eklundh, “A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter,” *Remote Sensing of Environment*, vol. 91, no. 3–4, pp. 332–344, 2004.
- [83] J. Steinier, Y. Termonia, and J. Deltour, “Comments on Smoothing and Differentiation of Data by Simplified Least Square Procedure,” *Analytical Chemistry*, vol. 44, no. 11, pp. 1906–1909, 1972.
- [84] A. Savitzky and M. J. Golay, “Smoothing and Differentiation of Data by Simplified Least Squares Procedures,” *Analytical Chemistry*, vol. 36, no. 8, pp. 1639–1643, 1964.
- [85] R. J. Barnes, M. S. Dhanoa, and S. J. Lister, “Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra,” *Applied Spectroscopy*, vol. 43, no. 5, pp. 772–777, 1989.
- [86] A. Plaza and C. I. Chang, “Impact of initialization on design of endmember extraction algorithms,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 11, pp. 3397–3407, 2006.
- [87] P. J. Burón González, “Aplicación del Análisis en Componentes Independientes a la Clasificación de Texturas,” pp. 9–14, 2015.
- [88] A. A. Green, M. Berman, P. Switzer, and M. D. Craig, “A Transformation for Ordering Multispectral Data in Terms of Image Quality with Implications for Noise Removal,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 26, no. 1, pp. 65–74, 1988.
- [89] L. Parra, C. Spence, P. Sajda, and G. M. D. First, “Unmixing Hyperspectral Data.”
- [90] A. Farooq and F. Qurat-ul-ain, “Algorithm and n-Dimensional Visualization for ETM+ Image Analysis: A Case of District Vehari,” *Global Journal of Human Social Science Arts & Humanities*, vol. 12, no. 15, pp. 23–32, 2012.
- [91] C. González, J. Resano, D. Mozos, A. Plaza, and D. Valencia, “FPGA Implementation of the pixel purity index algorithm for remotely sensed hyperspectral image analysis,” *Eurasip Journal on Advances in Signal Processing*, vol. 2010, 2010.
- [92] J. Theiler and G. Gisler, “A Contiguity-Enhanced K-Means Clustering Algorithm for Unsupervised Multispectral Image Segmentation,” vol. 836, 1997.
- [93] BRACK, Antonio y MENDIOLA, Cecilia, “Enciclopedia ECOLOGÍA DEL

PERÚ.,” 1997. [Online]. Available: <http://www.peruecologico.com.pe/libro.htm>.  
[Accessed: 29-Apr-2019].

- [94] H. Lalinde *et al.*, “Sobre el uso adecuado del coeficiente de correlación de Pearson: definición, propiedades y suposiciones,” 2018.