



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
FACULTAD DE FILOSOFÍA Y LETRAS
COLEGIO DE GEOGRAFÍA**

**CLASIFICACIÓN DE ZONAS DE CAFÉ DE SOMBRA A PARTIR
DEL ANÁLISIS MULTITEMPORAL DE IMÁGENES LANDSAT 8
EN OCOSINGO, CHIAPAS.**

TESIS

QUE PARA OBTENER EL TÍTULO DE:

LICENCIADA EN GEOGRAFÍA

PRESENTA:

LAURA ALEJANDRA PÉREZ RÍOS

ASESOR:

MTRO. JOSÉ MAURICIO GALEANA PIZAÑA

CIUDAD UNIVERSITARIA, CIUDAD DE MÉXICO, 2019.





Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Agradecimientos

A la Universidad Nacional Autónoma de México, por permitirme la posibilidad de formarme como profesionista, por ayudarme a crecer como persona y por ser mi segunda casa.

A los profesores del Colegio de Geografía, por haber sido parte de mi formación y esforzarse en transmitirme su conocimiento.

Al Maestro, humilde profesor y ya casi Doctor José Mauricio Galeana Pizaña, por haberme permitido ser una más de sus rémoras tesis, por todo el apoyo brindado, por los comentarios y sugerencias sobre esta tesis, por sus palabras de aliento y, sobre todo, por la confianza.

A los miembros de sínodo:

Al Doctor Juan Manuel Núñez Hernández, por sus comentarios sobre esta tesis, sus recomendaciones y su apoyo.

A la Maestra Cecilia Gutiérrez Nieto, gracias por la rápida revisión y sus comentarios para mejorar mi trabajo.

Al Doctor Jesús Abraham Navarro Moreno, por la revisión detallada de este trabajo, por enseñarme lo bonito de la cartografía matemática y por brindarme su apoyo.

A la Doctora Rocío Marisol Alanís Anaya, por su tiempo para revisar mi trabajo y por hacer comentarios y recomendaciones sobre lo que no había tomado en cuenta.

Al CentroGeo y su comunidad, por recibirme y enseñarme tanto en tan poco tiempo, por las amistades y experiencias que se han generado, y por permitirme seguir creciendo como persona y profesionista.

Al CONACYT, por el apoyo brindado para realizar esta tesis a través del Proyecto 292474. “Estrategias multidisciplinares para incrementar el valor agregado de las cadenas productivas del café, frijol, mango, agave mezcalero y productos acuícolas (tilapia) en la Región Pacífico Sur a través de la ciencia, la tecnología y la innovación”

A mis padres, por el apoyo inmensurable que me han dado toda la vida, por todo su esfuerzo, por tanta paciencia y por haberme ayudado a llegar hasta aquí, éste también es su logro.

ÍNDICE

ÍNDICE	I	III CARACTERIZACIÓN GEOGRÁFICA DEL ÁREA DE ESTUDIO	64
ÍNDICE DE FIGURAS	II	3.1 Localización	66
ÍNDICE DE CUADROS	IV	3.2 Geología	67
INTRODUCCIÓN	1	3.3 Edafología	68
I EL CAFÉ COMO SISTEMA AGROECOLÓGICO	8	3.4 Climatología	69
1.1 El café	10	3.5 Hidrología	70
1.1.1 La producción de café	12	3.6 Vegetación	70
1.2 Sistema de producción nacional de café	14	IV METODOLOGÍA	73
1.2.1 El café mexicano en el mundo	19	4.1 Adquisición de imágenes satelitales y bases de datos	75
1.3 Chiapas como principal productor de café de México	21	4.2 Preprocesamiento de imágenes satelitales	77
1.3.1 La denominación de origen del Café Chiapas	22	4.3 Desmezclado espectral	77
1.3.2 Particularidades del Café Chiapas	23	4.4 Cálculo de índice de vegetación	80
1.4 La producción de café como sustento económico familiar	25	4.5 Análisis de separabilidad espectral	80
1.5 La producción de café en Ocosingo, Chiapas	29	4.6 Clasificación de imágenes satelitales con el método <i>Random Forest</i>	82
II PERCEPCIÓN REMOTA	31	4.7 Verificación	83
2.1 Percepción remota	33	4.8 Estimación de superficie de café en las clasificaciones	83
2.1.1 El espectro electromagnético	38	V RESULTADOS	84
2.1.2 La reflectancia espectral	42	5.1 Parcelas de café de Ocosingo	86
2.1.3 La separabilidad espectral	44	5.2 Vegetación natural de Ocosingo	87
2.2 Desmezclado espectral	45	5.3 Calibración radiométrica y corrección atmosférica	88
2.2.1 El método de Segregación Espectral Automatizada Monte Carlo (AutoMCU)	47	5.4 Imágenes de reflectancia	88
2.3 Las imágenes digitales	48	5.5 Imágenes de cobertura fraccional	88
2.3.1 Transformaciones de la imagen	50	5.6 Imágenes de vegetación fotosintética	88
2.4 Clasificación digital de imágenes	52	5.7 Índice de Vegetación de Diferencia Normalizada	95
2.4.1 Métodos de clasificación supervisada	53	5.8 Puntos de control	96
2.4.2 Métodos de clasificación no supervisada	54	5.9 Gráficas de separabilidad espectral	96
2.4.3 El método de clasificación <i>Random Forest</i>	55	5.10 Clasificación <i>Random Forest</i> con NDVI	96
2.5 La verificación de la clasificación	58	5.11 Clasificación <i>Random Forest</i> con Serie VI	96
2.5.1 Validación cruzada	58	5.12 Matrices de confusión	100
2.5.2 La matriz de confusión	58	5.13 Superficie de café en las clasificaciones	103
2.5.3 El índice Kappa	59	VI DISCUSIÓN	104
2.6 El Programa Landsat	60	CONCLUSIONES	114
2.6.1 Las imágenes Landsat 8	61	REFERENCIAS	117

ÍNDICE DE FIGURAS

Figura 1.1. <i>Coffea arabica</i> y <i>Coffea canephora</i> .	10	Figura 2.9. Interacción de la energía electromagnética.	43
Figura 1.2. Café cereza.	10	Figura 2.10. Tipos de reflexión.	43
Figura 1.3. La roya del café <i>Hemileia vastatrix</i> .	12	Figura 2.11. Firmas espectrales.	43
Figura 1.4. La broca del café <i>Hypothenemus hampei</i> .	12	Figura 2.12. Separabilidad espectral de dos tipos de vegetación.	45
Figura 1.5. Estacionalidad de la producción nacional de café cereza 2001-2015.	14	Figura 2.13. Concepto de píxel mezcla.	46
Figura 1.6. Plantación de café de sombra.	16	Figura 2.14. Desmezclado espectral lineal.	47
Figura 1.7. Plantación de café de sol.	16	Figura 2.15. Diagrama de dispersión de las bibliotecas espectrales de PV, NPV y S.	47
Figura 1.8. Principales estados productores de café cereza 2013-2016.	17	Figura 2.16. Comportamiento espectral de las fracciones de cobertura.	48
Figura 1.9. Superficie cosechada de café cereza 2002-2016.	17	Figura 2.17. Píxel	49
Figura 1.10. Producción de café cereza 2004-2016.	18	Figura 2.18. Organización de datos en una imagen digital.	49
Figura 1.11. Comercio exterior de café 2007-2017.	20	Figura 2.19. IFOV (Campo de visión instantáneo).	50
Figura 1.12. Precio del café cereza al productor 2001-2015.	20	Figura 2.20. Ejemplo de Clasificación en árbol.	56
Figura 1.13. Índices de precios al productor y al consumidor del café 2010-2015.	20	Figura 2.21. Estructura de una clasificación <i>Random Forest</i> .	57
Figura 1.14. Precios indicativos del café de la ICO 2009-2016.	21	Figura 2.22. Organización de una matriz de confusión.	59
Figura 1.15. Estacionalidad de la producción en Chiapas de café cereza 2001-2015.	21	Figura 2.23. Línea del tiempo del Programa Landsat.	62
Figura 1.16. Práctica de la caficultura en el estado de Chiapas.	26	Figura 3.1 Localización de Ocosingo.	66
Figura 1.17. Niños participando en la cosecha del café en el estado de Chiapas.	26	Figura 3.2 Geología de Ocosingo por tipo de roca.	67
Figura 1.18. Familia caficultora de Chiapas.	27	Figura 3.3 Tipos de suelo de Ocosingo.	68
Figura 1.19. Caficultores del municipio de Ocosingo que recibieron apoyos del gobierno.	30	Figura 3.4 Tipos de clima de Ocosingo.	69
Figura 2.1. Percepción remota	31	Figura 3.5 <i>Lacandonia schismatica</i> .	71
Figura 2.2. Formas de teledetección: reflexión; emisión; emisión-reflexión.	35	Figura 3.6 <i>Licania platypus</i> .	71
Figura 2.3. Componentes del proceso de percepción remota.	37	Figura 3.7 Tipos de vegetación de Ocosingo.	72
Figura 2.4. Onda electromagnética.	37	Figura 4.1. Esquema metodológico.	74
Figura 2.5. Longitudes y frecuencias de las ondas electromagnéticas.	37	Figura 4.2. Escenas Landsat 8 que cubren la zona de estudio.	75
Figura 2.6. El espectro electromagnético.	38	Figura 4.3. Perfil del valor de reflectancia de un píxel.	77
Figura 2.7. Los tipos de dispersión atmosférica.	40	Figura 4.4. Imágenes de Vegetación fotosintética que cubren la zona de estudio.	78
Figura 2.8. Ventanas atmosféricas.	41	Figura 4.5. Recorte de imágenes con base en la delimitación de Ocosingo.	79
		Figura 4.6. Recorte de Padrón Nacional Cafetalero con base en la delimitación de Ocosingo.	79

Figura 4.7. Recorte de Serie VI con base en la delimitación de Ocosingo.	79
Figura 4.8. Apilamiento de capas de Vegetación fotosintética.	80
Figura 4.9. Creación de puntos de control de vegetación natural sobre NDVI.	80
Figura 4.10. Apilamiento de las bandas con mayor separabilidad.	82
Figura 4.11. Proporción de puntos de control para entrenamiento y verificación de las clasificaciones.	82
Figura 5.1. Parcelas de café de Ocosingo.	86
Figura 5.2. Vegetación natural de Ocosingo.	87
Figura 5.3. Calibración radiométrica y corrección atmosférica de las imágenes del 2015.	89
Figura 5.4. Imágenes de reflectancia de la escena 2049.	90
Figura 5.5. Imágenes de reflectancia de la escena 2148.	91
Figura 5.6. Imágenes de cobertura fraccional de la escena 2049.	92
Figura 5.7. Imágenes de cobertura fraccional de la escena 2148.	93
Figura 5.8. Imágenes de vegetación fotosintética (PV) de Ocosingo.	94
Figura 5.9. NDVI de Ocosingo del año 2015.	95
Figura 5.10. Puntos de control de cada clase.	97
Figura 5.11. Gráfica de separabilidad espectral entre café y vegetación (del NDVI).	98
Figura 5.12. Gráfica de separabilidad espectral entre café y vegetación (de la Serie VI).	98
Figura 5.13. Clasificación de café y vegetación natural (del NDVI).	99
Figura 5.14. Clasificación de café y vegetación natural (de la Serie VI).	100
Figura 5.15. Superficie de café en las clasificaciones.	103
Figura 6.1. Comparación de superficies de café (hectáreas) entre las clasificaciones y los datos del SIAP.	108

ÍNDICE DE CUADROS

Cuadro 1.1. Variedades comerciales de la especie <i>Coffea arabica</i> .	11
Cuadro 1.2. Clasificación del Café Chiapas.	25
Cuadro 2.1. Aplicaciones de la percepción remota.	34
Cuadro 2.2. Elementos del proceso de percepción remota.	36
Cuadro 2.3. Valoración del índice Kappa.	60
Cuadro 2.4. Propiedades técnicas del Programa Landsat.	62
Cuadro 2.5. Características de los sensores OLI y TIRS.	63
Cuadro 4.1. Criterios de búsqueda de imágenes satelitales.	75
Cuadro 4.2. Fechas de las imágenes satelitales adquiridas.	76
Cuadro 4.3. Uso de Suelo y Vegetación de la Serie VI 2014.	76
Cuadro 4.4. Clases de cobertura fraccional.	78
Cuadro 4.5. Bandas que componen las imágenes de cobertura fraccional.	78
Cuadro 4.6. Comparación entre resolución de la imagen y puntos de control.	81
Cuadro 5.1. Matriz de confusión de la clasificación de café y vegetación natural (del NDVI).	101
Cuadro 5.2. Matriz de confusión de la clasificación de café y vegetación natural (de la Serie VI).	102
Cuadro 6.1. Comparación de los resultados de distintos métodos de clasificación de café.	109



INTRODUCCIÓN

El uso de la percepción remota ha funcionado como marco de trabajo para realizar investigaciones de distintas disciplinas que requieren el análisis e interpretación de la información espectral de la cobertura terrestre. Una de las técnicas de percepción remota más importante es la clasificación de imágenes de sensores remotos a partir de distintos métodos, pueden ser duros (asignan una clase a cada píxel) o suaves (asignan varias clases a un píxel y obtienen su grado de pertenencia a cada una), además, pueden ser clasificadores supervisados o no supervisados, es decir, con puntos de entrenamiento (conocimiento del terreno) o sin ellos (Chuvienco, 2010; Phiri & Morgenroth, 2017).

Las técnicas de percepción remota para clasificar los diferentes tipos de cobertura terrestre son muchas, algunas de ellas se han utilizado para clasificar el café de sombra y diferenciarlo de otro tipo de coberturas; este tipo de clasificaciones ha sido una línea de investigación relevante dada la importancia del sistema de producción de este grano a nivel mundial (Bolanos, 2007).

En algunos acercamientos que se han hecho a la clasificación de café, como en Kawakubo & Pérez (2016) y Gaertner, Genovese, Potter, Sewake & Manoukis (2017), los alcances de los métodos empleados y sus resultados expresados en porcentaje de confiabilidad global y el valor del

índice Kappa, son prueba de que la utilización de técnicas de percepción remota para este tipo de finalidades logra obtener resultados generalmente satisfactorios, pues los resultados de una clasificación son considerados aceptables a partir de 50 % de confiabilidad global (Reis, 2015) y de 0.21 de índice Kappa (Landis & Koch, 1977).

En Bolanos (2007), se analizaron dos métodos para clasificar café y otras siete clases (bosque, pastura, nubes, sombra, urbano, agua y otros) en Huila, Colombia, a partir de imágenes del sensor Landsat 7, los datos de elevación de la zona de estudio y 1429 puntos de control para evaluar la confiabilidad de los resultados de ambos métodos. La primera clasificación se realizó a través de un análisis basado en píxeles con el clasificador de árboles de decisión llamado *Rule-based*, con el que se obtuvo una confiabilidad global de 71.89% y un valor de índice Kappa de 0.66. El segundo método utilizado fue la clasificación basada en objetos, ésta tuvo como resultados un 74.88% de confiabilidad global y 0.70 de índice Kappa. Sin embargo, dentro de los resultados para cada una de las ocho clases, la clase de café es una de las que tienen menor confiabilidad (40%) e índice Kappa (0.38).

En Cordero-Sancho & Sader (2007) se utilizó una imagen Landsat 7, datos de elevación y

precipitación, datos topográficos de la *Shuttle Radar Topography Mission* (SRTM) y 293 puntos de control para clasificar la imagen en cinco clases, dos de las cuales pertenecían al cultivo de café. El modelo utilizado para clasificar la imagen fue el *Coffee Environmental Stratification Model* (CESM), que está basado en árboles de decisión estableciendo los datos de elevación y precipitación óptimos para el cultivo del café, con diferentes combinaciones de bandas espectrales con un índice de vegetación (NDVI), además utilizaron el clasificador de máxima verosimilitud y utilizaron los resultados para hacer la comparación con los de su modelo. La clasificación con el CESM tuvo como resultado 65.19% de confiabilidad global y 0.56 de índice Kappa, mientras que la clasificación de máxima verosimilitud obtuvo 63.14% de confiabilidad global y 0.53 de índice Kappa.

En Martínez-Verduzco, Galeana-Pizaña & Cruz-Bello (2011), se utilizaron dos métodos de clasificación para separar el café de sombra de la vegetación natural, y ambos resultados se integraron para generar un mapa de cobertura más certero. Se utilizaron imágenes SPOT 5, cuatro clases, 87 puntos de entrenamiento y 132 de verificación obtenidos en trabajo de campo. El primer método consistió en una clasificación supervisada con el clasificador de máxima verosimilitud, el segundo fue una clasificación

realizada a partir del mapeo participativo sobre un mosaico de ortofotos de la zona de estudio con habitantes expertos. Los resultados de la primera clasificación fueron de 73% de confiabilidad global y 0.7 de índice Kappa, se integraron ambas clasificaciones y el resultado obtuvo una confiabilidad global de 74.57% y un índice Kappa de 0.73.

Ortega, Komar, Price & Ventura (2012), plantearon cómo identificar el cultivo de café de sol y el café de sombra para distinguirlos de la vegetación natural a través de la biomasa asociada al cultivo. Para realizar la clasificación se utilizaron dos métodos, uno llamado *Cluster-Busting*, y uno en el que se utilizó una tabla de contingencia, un análisis de componentes principales y un análisis de clústeres; los resultados de ambas clasificaciones se utilizaron para hacer una clasificación híbrida. Usaron 3 imágenes multitemporales del sensor Landsat 5, 285 puntos de control y cuatro clases, dos de ellas de café; el resultado fue una confiabilidad global de 76.74% y un índice Kappa de 0.68, en donde las clases con resultados más bajos fueron las de café.

En Kawakubo & Pérez (2016), se utilizó una imagen del sensor LISS-III, 1600 puntos de control y ocho clases, tres de ellas de café (café seco o podado, café en producción y café

mezclado con otro tipo de vegetación). Se compararon 3 métodos para llegar a la clasificación final, primero aplicaron un análisis de componentes principales, posteriormente se realizó un análisis de desmezclado espectral (SMA) y una clasificación de minería de datos (*Data Mining*), con lo que obtuvieron una confiabilidad global de 73.7% y un índice Kappa de 0.67, con valores relativamente altos para las tres clases de café.

En Gaertner, Genovese, Potter, Sewake & Manoukis (2017), se analizaron dos métodos de clasificación utilizando dos imágenes WorldView-2. El primer método fue una clasificación de máxima verosimilitud con 2366 puntos de control y ocho clases, el segundo método fue una clasificación basada en objetos con 6446 puntos de control y diez clases. Los resultados del primer método tuvieron 68% de confiabilidad global y 0.60 de índice Kappa, el segundo, 81% de confiabilidad global y 0.75 de índice Kappa.

Con lo anterior en mente, surgió la problemática que abordó este trabajo de tesis. Los métodos de clasificación de imágenes satelitales han funcionado para identificar diversos tipos de uso de suelo y vegetación; sin embargo, cuando se requiere una clasificación más específica, los clasificadores tradicionales no logran identificar una diferencia suficiente entre los valores de

algunas clases y los confunde con otra. Por ello, los resultados finales no logran representar una clasificación precisa que indique con claridad dónde se encuentra una clase o tipo de vegetación específico, como lo es el café.

En este trabajo, se prueba y evalúa un método diferente y no tradicional de clasificación con la finalidad de discriminar el café de sombra y lograr diferenciarlo de la vegetación natural utilizando imágenes multitemporales. Es importante resaltar la multitemporalidad, pues la observación de un mismo sitio en tiempos distintos hace posible el reconocimiento de patrones del comportamiento espectral de coberturas valiosas como el café.

Además, se comparan los resultados con los de los métodos que ya han sido probados para este cultivo. También se analiza el alcance del clasificador utilizado, que es una combinación de una máquina de aprendizaje y un algoritmo probabilístico avanzado. Cabe destacar que este método híbrido no ha sido probado con anterioridad para clasificar café.

Se eligió este cultivo porque el café es uno de los productos de mayor consumo a nivel mundial y el cultivo más importante del estado de Chiapas (Instituto Nacional de Estadística y Geografía, 2017). Además de la importancia económica

estatal y nacional inherente, el café es muy relevante para los productores, pues la producción del grano funciona como su sustento familiar y su única vía de ingresos (Vargas, 2007). La caficultura en Chiapas es una de las principales actividades económicas de muchas familias, ha hecho que crezca el arraigo cultural por el café y por su cultivo, y ha constituido al café como uno más de sus rasgos culturales (Renard, 2002).

El grano mexicano ha logrado posicionarse en distintos países del mundo porque es cultivado bajo estrictas condiciones y óptimos procesos que aseguran una excelente calidad, aroma, sabor, cuerpo y acidez en el producto. Por esta razón, se otorgó la denominación de origen Café Chiapas a doce regiones que cultivan en el estado, entre ellas se encuentra Ocosingo, el municipio elegido como espacio de estudio de esta tesis (Diario Oficial de la Federación, 2003). Se eligió por sus características físicas, pues tiene las condiciones óptimas para el cultivo de café de sombra, además de eso es uno de los municipios del estado con mayor número de parcelas de café registradas y el de mayor superficie territorial, en la cual la distribución del cultivo es heterogénea.

La elección de hacer uso de la percepción remota en este trabajo se debe a que sus técnicas reducen considerablemente los costos y tiempos dedicados a la obtención de información, con ella

es posible obtener información espacial de una misma zona de estudio en distintas fechas. Aunque las imágenes satelitales, al ser el conjunto de diferentes bandas del espectro electromagnético, contienen información que es imperceptible para el ojo humano, al ser manejadas con técnicas de percepción remota es posible obtener mucha información sobre la superficie y aprovecharla para generar análisis y nueva información (Chuvienco, 2010).

Para esta tesis, se eligió la clasificación de imágenes multitemporales a partir de un desmezclado espectral y el clasificador *Random Forest*, éste último ha sido utilizado para clasificar distintas clases de cobertura, los resultados de este método obtenidos en algunos trabajos de investigación son bastante favorables en comparación con otros métodos de clasificación comúnmente utilizados. Sin embargo, no se ha profundizado el uso de esta técnica para la clasificación de cultivos específicos como el café.

Por ejemplo, en Rodríguez-Galiano, Ghimire, Rogan, Chica-Olmo & Rigol-Sánchez (2011), se utilizó la clasificación *Random Forest* para clasificar 14 clases (algunas fueron las zonas urbanas, coníferas, invernaderos, pastizales, suelos expuestos y agua), se usaron dos imágenes Landsat 5, variables topográficas como la elevación y 150 puntos de control para cada una

de las categorías, con esto se obtuvo una confiabilidad global de 92% y un índice Kappa de 0.92.

En Tian, Zhang, Tian, J & Sun (2016), se utilizaron imágenes del satélite Pléiades-1B y Landsat 8, se utilizaron 10 clases, entre ellas las Ciénegas, estanques, trigo, maíz, bosques, lagos, pastos y arbustos endémicos de China, se hizo un análisis de componentes principales y un NDVI que se unieron en una sola imagen. Se obtuvieron 182 puntos de control con los que se clasificó la imagen combinada usando *Random Forest*, el resultado fue 92.5% de confiabilidad global y 0.92 de índice Kappa.

En Ming, Zhou, Wang & Tan (2016), su utilizaron imágenes del satélite HJ-1B-CCD2 para clasificar en seis clases (zona urbana, suelo expuesto, cuerpos de agua, pastos, zona arbórea y cultivos), con 20 puntos de entrenamiento para cada clase y un total de 693 puntos de verificación. En la clasificación *Random Forest* se utilizaron variables que incluyen índices de vegetación, de agua y de nieve, así como un modelo digital de elevación y la pendiente. Los resultados fueron de 74.03% de confiabilidad global, con los valores más altos asociados a los cuerpos de agua y zonas arbóreas, mientras que los cultivos resultaron con valores de 35% de confiabilidad global.

En Kulkarni & Lowe (2016), se utilizó este método para clasificar imágenes Landsat 8 en cuatro clases (agua, vegetación, suelo y bosque), con 200 puntos de control para cada clase, se hizo la clasificación con 500 árboles y se comparó el resultado con otros métodos. La confiabilidad global que se obtuvo fue de 96% y un índice Kappa de 0.94, superando los resultados del clasificador de máxima verosimilitud y el *ID3 Tree*, con 92.5% y 0.89 para ambos casos.

La clasificación *Random Forest* se utilizó en Kelley, Pitcher & Bacon (2018) para clasificar diez clases de cobertura: café de sombra, milpa, pastizales, pino, pino-encino, bosques semi-húmedos, bosques caducifolios, construcciones, agua y campos de arroz. Se utilizaron 200 puntos de control para cada clase, datos de brillo, humedad, variables fisiográficas, topográficas y de temperatura, con esto se hicieron cinco combinaciones de variables con distintas bandas espectrales. Se obtuvieron cinco resultados, el mejor obtuvo una confiabilidad global de 90.5% y 0.89 de índice Kappa, mientras que el más bajo tuvo 65.6% y 0.62. Se observó que la variable de elevación fue la que presentó mayor importancia en la clasificación resultante.

Por lo anterior, en este trabajo se espera obtener una clasificación en la que se separe una imagen en dos clases, una de café de sombra y una de

vegetación natural, a partir del procesamiento de información que no ha sido utilizada con el clasificador *Random Forest* específicamente para el café.

En esta tesis se utiliza el análisis de la separabilidad espectral de cada clase y sus valores de vegetación fotosintética obtenidos por medio del desmezclado espectral. Se espera que la metodología aplicada sea capaz de adaptarse y reproducirse para otras zonas de cultivo de café y para clasificar otros tipos de cobertura. Está fundamentada en la **hipótesis** siguiente: la clasificación *Random Forest* aplicada a las imágenes con valores de vegetación fotosintética logra identificar de manera confiable el café de sombra de Ocosingo.

Para demostrar dicha hipótesis se planteó como **objetivo general** evaluar el alcance del método de clasificación *Random Forest* a partir de imágenes Landsat 8 para la identificación de zonas de café de sombra en Ocosingo, Chiapas. Dentro de éste, se plantearon cuatro **objetivos particulares** con los que se pretende: obtener valores de vegetación fotosintética de las imágenes Landsat 8 a partir del desmezclado espectral; clasificar imágenes de satélite a través

de información del Padrón Nacional Cafetalero, de un índice de vegetación y de una serie de Uso de Suelo y Vegetación; analizar la confiabilidad y el índice Kappa del método de clasificación *Random Forest* a partir de una matriz de confusión; y finalmente comparar los resultados con las superficies reportadas por la Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación a través del Servicio de Información Agroalimentaria y Pesquera.

La presente tesis se compone de siete capítulos. En el primero se expone la importancia del café a nivel mundial, nacional y familiar, así como su producción a través de la caficultura, en el segundo se explica la información básica para entender la percepción remota y algunas de sus técnicas; además, se explica cómo funciona el método de clasificación elegido para este trabajo y lo que hay detrás de los procesos utilizados. En el tercero se describen los aspectos físicos que integran el área de estudio; el cuarto capítulo se compone de la metodología seguida para la obtención de resultados; en el quinto se muestran los resultados alcanzados; el sexto desarrolla la discusión generada a partir de los resultados obtenidos y, finalmente, en el séptimo se muestran las conclusiones.



I EL CAFÉ COMO SISTEMA AGROECOLÓGICO

Contenido

Este primer capítulo se desarrolla en cinco subcapítulos, algunos de ellos contienen otros segmentos. En el primer subcapítulo se habla sobre el café, qué es, cuáles son sus variedades, las características más importantes, las especificaciones para su desarrollo y la manera en la que se clasifican sus granos; este segmento tiene una sección en la que se menciona la importancia de este producto, la caficultura y su proceso de producción. En el segundo subcapítulo se describe cómo se lleva a cabo la producción del café en México, las características de la producción, las regiones y los estados en donde se produce y los sistemas de cultivo que se practican; contiene un segmento en el que se explica la relación del café mexicano con otros países productores, la importación y exportación, así como el reconocimiento del café orgánico de México a nivel internacional. En el tercer subcapítulo se habla del estado de Chiapas y su importancia como principal productor de café del país, se muestran otras dos secciones; en la primera se explica la denominación de origen que adquirió el café producido en ese estado, en la segunda se muestran las características y particularidades del café chiapaneco. En el cuarto subcapítulo se menciona el papel que juega el café como sustento económico familiar de los caficultores y cómo han hecho del cultivo una característica cultural que los identifica. Finalmente, en el quinto subcapítulo se aterriza la producción de café del municipio de Ocosingo, Chiapas.

1.1 El café

Es un cultivo perenne producido principalmente en zonas altas de montaña; es un producto no perecedero, fácilmente almacenable y transportable, además en algunas zonas figura como un cultivo artesanal y orgánico (Rodríguez, 1998). Se desarrolla mejor en climas cálidos y semicálidos, con temperaturas que fluctúan entre 18° y 21° C; requiere de precipitaciones que varían entre 1 200 y 1 800 mm anuales, es decir, las condiciones ideales para su crecimiento se encuentran en la zona intertropical mundial (Aguirre, 1999).

Pertenece a la familia de las rubiáceas y existen más de treinta especies dentro del género *Coffea*, de las cuales sólo dos son relevantes para la comercialización: la *Coffea arabica* o café cereza, y la *Coffea canephora* o robusta (Figura 1.1) (Servicio de Información Agroalimentaria y Pesquera, 2016; Ruiz, 2018).



Figura 1.1. *Coffea arabica* y *Coffea canephora* (aquaspresso.co.za)

El café cereza (Figura 1.2) es la especie más reconocida y de mayor importancia, representa alrededor del 75% del café de todo el mundo, esto se debe a que la calidad de su cultivo es mayor, además produce un café fino y aromático más delicado. Por su parte, la especie *Coffea canephora* o robusta, es un producto con mayor contenido de cafeína, es más fuerte y ácida que la arábica, es también el grano comúnmente comercializado como café soluble o instantáneo y es adaptable a terrenos llanos (Ruiz, 2018).



Figura 1.2. Café cereza (inforural.com.mx)

El cafeto de arábica es un arbusto grande de hojas ovaladas color verde oscuro, produce un fruto ovalado que alcanza la madurez entre los siete y nueve meses, es una especie susceptible a las plagas y enfermedades. El cafeto de robusta es un arbusto o árbol pequeño de raíz poco profunda, su fruto es redondeado y puede tardar once meses en alcanzar la madurez; su semilla es más

pequeña y alargada que la semilla arábica, las variedades de café robusta son más resistentes a las plagas como la roya (Fideicomisos Instituidos en Relación con la Agricultura, 2016).

Cada especie contempla un gran número de variedades, las más comerciales son de la especie arábica, entre ellas se encuentran: Typica, Bourbon, Caturra, Mundo Novo, Pluma Hidalgo, Maragogipe, Garnica, Catucaí y Catimor, descritas en el cuadro 1.1 (Escamilla, 1993).

El café es un cultivo susceptible a infectarse de algunas enfermedades y plagas, una es la llamada roya (Figura 1.3), causada por el hongo *Hemileia vastatrix* que infecta las hojas de los cafetos

provocando su caída prematura. Se caracteriza por generar áreas claras en la superficie de las hojas y producir un polvillo color naranja en su parte posterior (Fideicomisos Instituidos en Relación con la Agricultura, 2016).

Otra enfermedad es la provocada por el insecto llamado broca del café (*Hypothenemus hampei*), éste es una plaga que afecta directamente al grano, se caracteriza por alimentarse de los granos (Figura 1.4), causar su pudrición y provocar la caída de los frutos (Servicio Nacional de Sanidad, Inocuidad y Calidad Agroalimentaria, 2017).

Cuadro 1.1. Variedades comerciales de la especie *Coffea arabica* (INIFAP, 2013; Escamilla & Zamarripa, 2016).

Variedad	Descripción
Catucaí (amarillo y rojo)	Cruza entre Icatú (Robusta, Bourbon y Mundo Novo) y Catucaí (Mundo Novo y Caturra), moderadamente resistente a la roya y tolerante a la sequía.
Garnica	Cruza entre Mundo Novo y Caturra, susceptible a la roya, requiere fertilización, ciclo productivo corto.
Geisha	Originaria de Etiopía, pertenece al subtipo Typica, su grano está catalogado como uno de los más exóticos. Es resistente a la roya pero susceptible a nemátodos.
Oro Azteca	Cruza entre Caturra Rojo y el Híbrido de Timor, resistente a la roya, semilla grande y susceptible a nemátodos.
Sarchimor	Proviene de la cruce entre las variedades Villa Sarchi 970/10 y CIFIC 8332, resistente a algunos nemátodos, su semilla es grande con resistencia a la roya.
Typica	Es la base para desarrollar otras variedades, bien adaptada a condiciones frías, de fruto grande, susceptible a nemátodos y muy susceptible a la roya. Tolerante a baja fertilidad y sequía.



Figura 1.3. La roya del café *Hemileia vastatrix* (scanews.coffee)



Figura 1.4. La broca del café *Hypothenemus hampei* (croplifela.org)

Existen distintas formas de clasificar este producto, puede clasificarse de acuerdo a su calidad en tres categorías comerciales: de calidad ejemplar (*exemplary*), de disponibilidad limitada con sabor superior y único; de primera calidad o calidad superior (*premium*), de disponibilidad moderada con sabor que varía entre bueno y muy

bueno; y de calidad normal (*mainstream*), de amplia disponibilidad y un sabor aceptable (Ruiz, 2018). También, puede clasificarse en cuatro categorías según precio y calidad, en orden descendente: suaves colombianos, otros suaves, brasileños naturales y robustas. Adicionalmente, el café se clasifica según lo establecido en la normatividad que cada país tiene vigente para este producto (Fideicomisos Instituidos en Relación con la Agricultura, 2016).

El café es distribuido y comercializado principalmente en dos presentaciones: como café tostado en grano entero o molido, que es envasado en bolsas o latas, y como café soluble normalmente envasado en frascos, adicionalmente, en años recientes se han vuelto populares otras presentaciones como las cápsulas, los sobres y las pastillas (Ruiz, 2018).

1.1.1 La producción de café

El café se consolidó como un producto importante mundialmente debido a los intereses económicos nacionales y del extranjero, tanto el mercado internacional como el local son dependientes de la producción campesina que conforma la base del circuito económico del café (Rodríguez, 1998).

Este grano es el segundo producto más comercializado del mundo (el primero es el petróleo), es también uno de los productos que poseen mayor valor dentro del mercado internacional, por lo que es crucialmente importante para la economía de muchos países en desarrollo. Para el año 2013, la Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO) dio un estimado de nueve millones de toneladas de café producidas en el mundo, donde los tres países productores más importantes fueron Brasil, Vietnam y Colombia, que en conjunto aportaron cerca del 60% de las exportaciones totales (Monroy, 2004; FAO, 2013; Ruiz, 2018).

La caficultura o cafeticultura, se refiere a la actividad de producción y comercialización del café, es considerada como una actividad estratégica fundamental en el sector agrícola, permite la integración de cadenas productivas, generación de divisas y empleos, y es una práctica con relevancia ambiental (Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación, 2015; Universidad Autónoma Chapingo, 2015).

La caficultura se volvió importante económica y socialmente en México a partir de finales del siglo XVIII, pues fue cuando comenzaron las primeras exportaciones de café, y desde el año 1877 México

es considerado como uno de los principales países productores de café del mundo (Renard, 1993; González, 2004).

Cuando la producción de café estaba regida por el sistema del Acuerdo Internacional del Café (AIC) desde 1962, éste no autorizaba la diferenciación de los granos de café de acuerdo con su calidad, por lo que se mezclaban granos de diferentes alturas y grados de humedad generando un café de calidad mediana y de menor precio en relación con los cafés de Centroamérica, fue después de la desaparición del AIC (en 1989) cuando en México el gobierno federal comenzó a apoyar al sector productor a través de órganos descentralizados de la SAGARPA y de la Asociación Mexicana de la Cadena Productiva del Café, con distintos programas en los que empezaron a separarse los granos de café de acuerdo a sus propiedades y se empezó a especializar la caficultura (Renard, 2002; Asociación Mexicana de la Cadena Productiva del Café, 2011).

Actualmente, las personas que dedican sus horas de trabajo a la producción de café requieren del total conocimiento del grano, saber cuándo hacer manualmente la cosecha, saber seleccionar los frutos, dominar las labores agronómicas, tener una concepción ecológica del proceso productivo y saber hacer la vinculación de su producto con los mercados nacionales y extranjeros (Centro de

Investigación, Capacitación, Asesoría y Promoción, 2007; Olivar & Bustamante, 2016). Son varios los eslabones que constituyen la cadena del café, a lo largo de ella existen etapas en las que se encuentran intermediarios: el primer eslabón lo conforman los pequeños productores, luego los comerciantes locales, el procesador, el exportador, el corredor de bolsa, las compañías multinacionales y los minoristas, para finalmente llegar a los consumidores (Ruiz, 2018).

1.2 Sistema de producción nacional de café

México cuenta con las condiciones ideales para desarrollar el cultivo del café, desde que el grano llegó proveniente del Caribe a inicios del siglo XIX, se extendió a las que hoy se conocen como zonas cafetaleras. En estas zonas se producen las dos especies principales de café, arábica y robusta; la primera ocupa el 97% de la superficie cafetalera (de la variedad *typica*) y la segunda solamente el 3% (Asociación Mexicana de la Cadena Productiva del Café, 2011; Ruiz, 2018).

En el país, la producción de café se desarrolla en ocho meses (entre octubre y mayo), y más del 80% de la producción nacional se cosecha entre diciembre y marzo (Figura 1.5) (Fideicomisos Instituidos en Relación con la Agricultura, 2016).

La producción se lleva a cabo principalmente en las zonas que se encuentran en las vertientes de las cadenas montañosas ubicadas en el centro y sur del país, en las zonas tropicales y subtropicales, principalmente bajo la sombra de la cubierta de un dosel de árboles. Los municipios cafetaleros se localizan, en mayor parte, en zonas tropicales húmedas (selvas altas y medianas) y subhúmedas (selvas bajas), aunque también es posible encontrarlos en zonas de bosque mesófilo (Toriello, 2014; Ruiz, 2018).

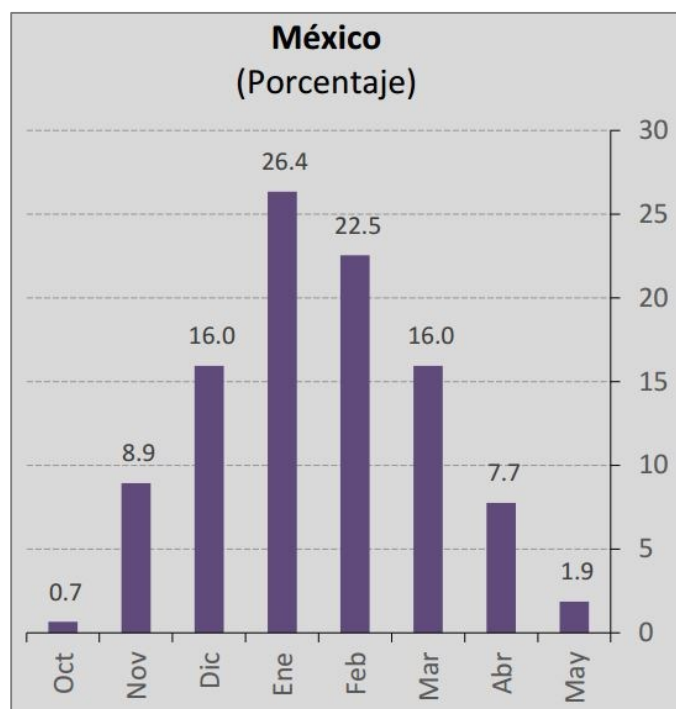


Figura 1.5. Estacionalidad de la producción nacional de café cereza 2001-2015 (FIRA, 2016).

Se estima que el 40% de la superficie cafetalera nacional se desarrolla en selvas altas y medianas, el 23% en bosques de pino y encino, el 21% en

selvas bajas caducifolias y el 15% en bosques mesófilos de montaña; esto muestra que las zonas en las que se practica la caficultura son ricas en flora y fauna (Bartra, 2006).

La producción de café en México se concentra principalmente en cuatro regiones: la Vertiente del Golfo de México (San Luis Potosí, Hidalgo, Querétaro, Puebla, Estado de México, Veracruz, Tabasco y una parte de Oaxaca); la Vertiente del Océano Pacífico (Colima, Guerrero, Jalisco, Nayarit, Michoacán, Morelos y parte de Oaxaca y del Estado de México); Región Soconusco (gran parte de Chiapas) y la Región Centro Norte de Chiapas (Servicio de Información Agroalimentaria y Pesquera, 2014).

En el país es posible encontrar dos tipos de sistemas de producción convencional de café: el café de sombra (Figura 1.6) y el café de sol (Figura 1.7). En la producción de café de sombra existe el policultivo tradicional y el policultivo comercial, en el primero se hizo la introducción del café bajo bosques y selvas que es donde se desarrolla con mayor presencia la cultura de las comunidades. En el policultivo comercial se introdujeron árboles que dieran sombra apropiada para el cultivo de café, generalmente en este sistema de producción las plantaciones son homogéneas. En el caso del café de sol, las plantas están expuestas directamente al sol sin ningún tipo de cobertura

de árboles que les proporcione sombra, además, los suelos presentan más erosión y las plantas de café necesitan agroquímicos para el control de plagas y obtener un buen rendimiento (Toriello, 2014).

Para su producción, trabajan más de quinientos mil productores residentes de quince entidades federativas. De las quince entidades productoras, más del 80% de la producción total nacional proviene de tres de ellas (Figura 1.8). En primer lugar, se encuentra Chiapas con el 41.0%, en segundo lugar, está Veracruz con 24.0% y en tercer lugar se ubica Puebla con el 15.3% de la producción, hasta el año 2016 el tercer lugar lo había ocupado el estado de Oaxaca (Figura 1.9) (Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación, 2018).

El café se ha convertido en uno de los productos más importantes para México y es considerado un cultivo estratégico para la economía del país. En términos nacionales, el café representa el 1.34% de la producción de bienes agroindustriales y el 0.66% del PIB agrícola. Del total de la superficie cultivada de café, el 3.24% representa al cultivo de café orgánico (Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación, 2018).



Figura 1.6. Plantación de café de sombra. (poabsestates.com).



Figura 1.7. Plantación de café de sol. (videoblocks.com).

Principales estados productores de café cereza, 2013/14-2015/16
(Miles de toneladas)

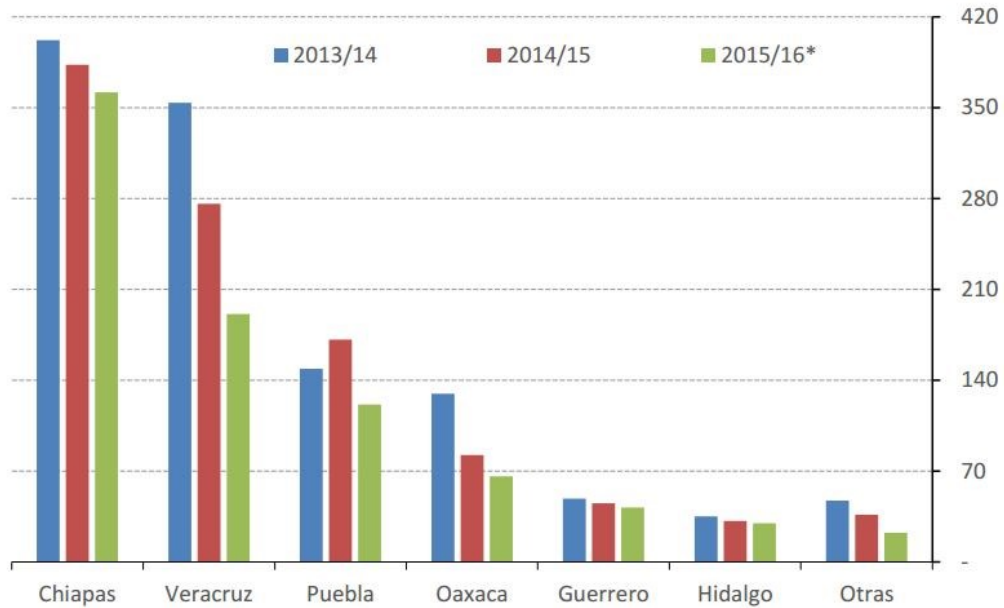


Figura 1.8. Principales estados productores de café cereza 2013-2016. (SIAP, 2016).

Superficie cosechada de café cereza en México, 2002/03-2015/16
(Miles de hectáreas)

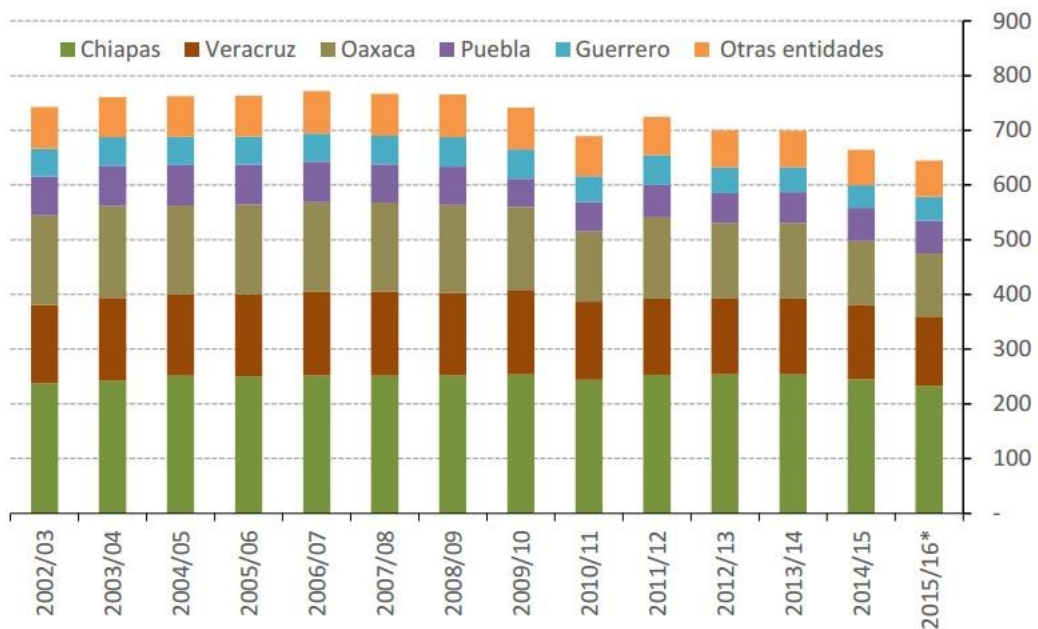


Figura 1.9. Superficie cosechada de café cereza 2002-2016. (SIAP, 2016).

Desgraciadamente, existe una tendencia decreciente en la producción nacional de este producto (Figura 1.10), debido principalmente a la reducción en la productividad de los cafetos durante los últimos años, que ha provocado que el rendimiento y la superficie cosechada hayan decrecido (Fideicomisos Instituidos en Relación con la Agricultura, 2016).

La presencia de la roya explica estas reducciones, a partir del año 2013 cuando se extendió en México y América Central, provocando millones de dólares en pérdidas. Por esta razón, la SAGARPA creó el Programa Integrado del Café (PIAC), con el que se buscaba aumentar la producción y competitividad a partir del desarrollo de viveros certificados, entrega de paquetes tecnológicos y rehabilitación de cafetales, además, en conjunto con la INIFAP se

buscaba obtener una semilla resistente a la roya (Oro Azteca) y distribuirla en las plantaciones (Fideicomisos Instituidos en Relación con la Agricultura, 2016).

En el país, existen alrededor de medio millón de productores de café, el 60% de ellos pertenecen a alguna comunidad indígena de 32 etnias diferentes, en conjunto, generan más de 121 millones de jornales anuales (Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación, 2010).

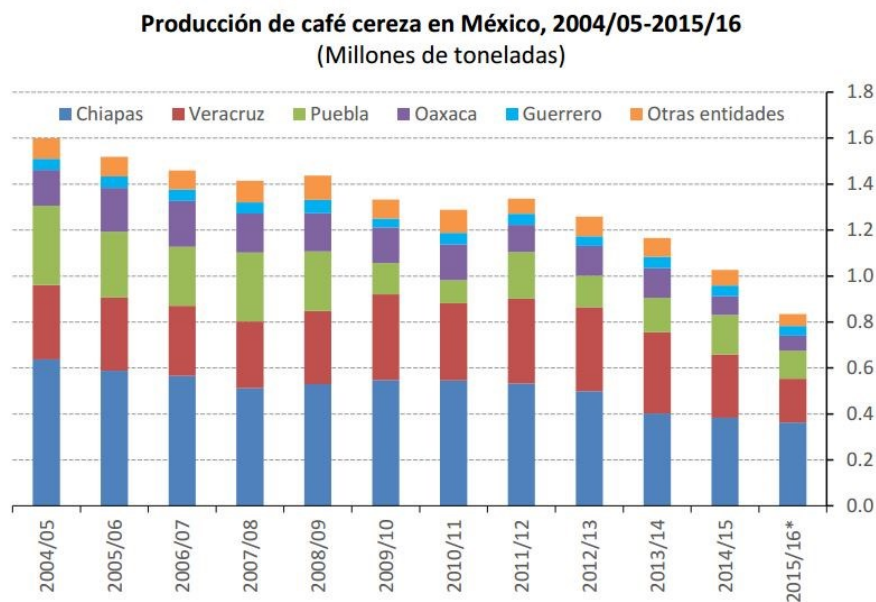


Figura 1.10. Producción de café cereza 2004-2016. (FIRA, 2016).

1.2.1 El café mexicano en el mundo

El café es un producto que proporciona las condiciones ideales para la cooperación internacional entre los países productores y los consumidores, se realizan intercambios de información entre países para identificar oportunidades y acciones que ayuden a mejorar las condiciones actuales y futuras del mercado mundial del café (Barrera, Parra, Herrera, Jarquín & Pohlan, 2004).

En México, del total de zonas conocidas con producción orgánica de café, alrededor del 80% están verificadas, la mayoría han sido certificadas por empresas extranjeras y en menor medida por empresas certificadoras nacionales (Gómez, Gómez Cruz & Schwentesius, 2001).

México es uno de los principales países productores de café orgánico en el mundo, el producto nacional se situó en el décimo lugar mundial de productores de café en el año 2017 (International Coffee Organization, 2018) y para julio del año 2018 descendió al décimo primer puesto (Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación, 2018).

Dentro de la producción total mundial de café, el producto mexicano representa alrededor del

2.4%, el país exporta alrededor de veintiocho mil toneladas a distintos países, principalmente a Estados Unidos con 53.85% del total de exportaciones, a países de la Unión Europea y otros más como Canadá, China, Cuba y Japón (Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación, 2018).

En el ciclo comercial 2015/16 las ventas del café de México en el extranjero llegaron a su nivel mínimo desde el ciclo comercial 2004/05, mientras que las importaciones incrementaron 8.2%. El 64.2% del café exportado fue grano verde sin tostar (Figura 1.11), el 6.2% fue tostado y molido y el 29.6% fue café soluble (Fideicomisos Instituidos en Relación con la Agricultura, 2016).

El precio del café mexicano (Figuras 1.12 y 1.13) alcanzó en Nueva York un precio de comercialización de 3,597 dólares por tonelada de café verde en el año 2016, mientras que, en México, hasta el año 2015 el precio para exportación se ubicó en 5,204 pesos por tonelada, el precio internacional del grano se muestra en la figura 1.14 (Fideicomisos Instituidos en Relación con la Agricultura, 2016).

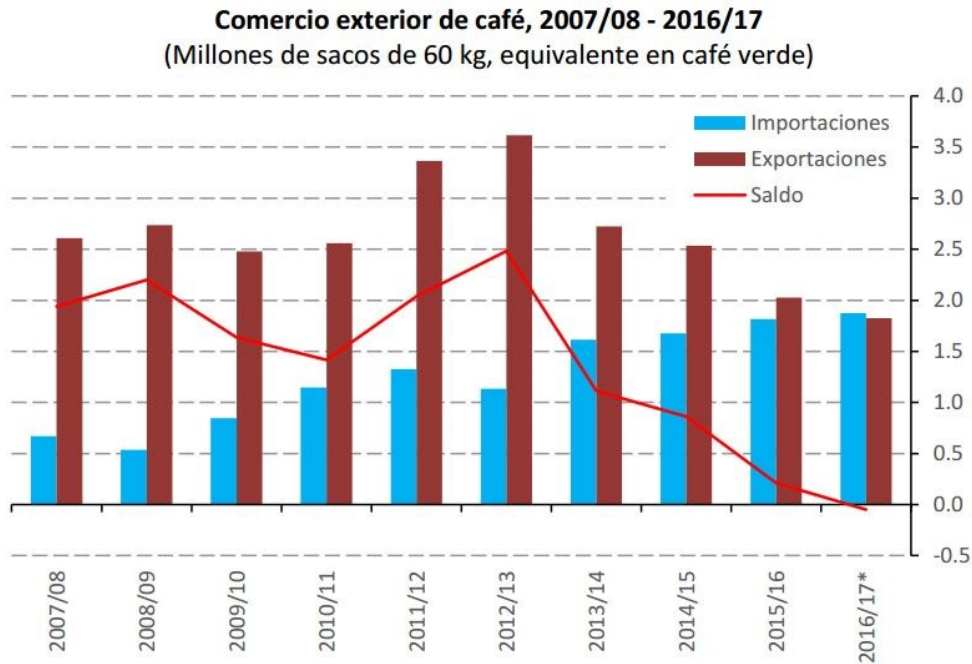


Figura 1.11. Comercio exterior de café 2007-2017. (SIAP, 2017).

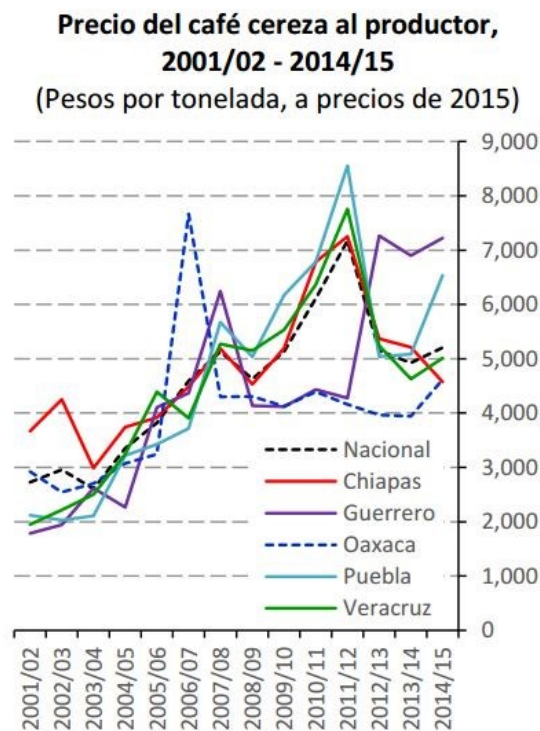


Figura 1.12. Precio del café cereza al productor 2001-2015. (SIAP-SAGARPA, 2016).

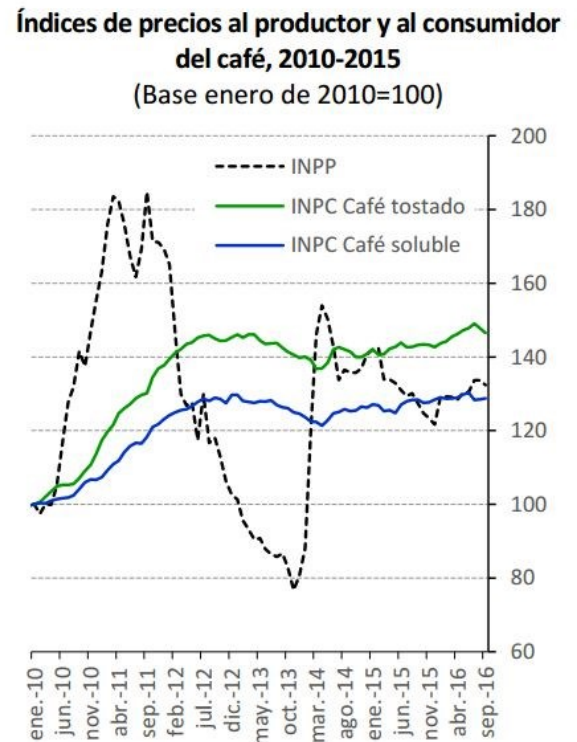


Figura 1.13. Índices de precios al productor y al consumidor del café 2010-2015. (INEGI, 2016).

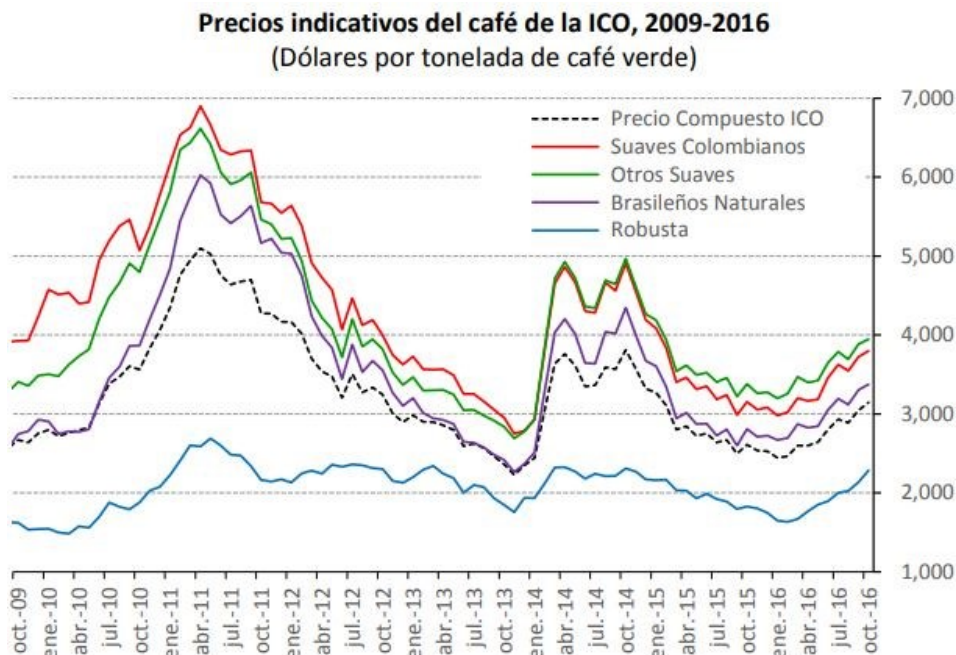


Figura 1.14. Precios indicativos del café de la ICO 2009-2016. (ICO, 2016).

1.3 Chiapas como principal productor de café de México

Chiapas es líder nacional en producción de café, con 357 mil 734 toneladas producidas en el año 2016, que representan el 41% del volumen nacional, predominan los cultivos de variedades de la especie *Coffea arabica* (Grupo Técnico Interdisciplinario de Café, 2017; Servicio de Información Agroalimentaria y Pesquera, 2017).

La producción de café de esta entidad presenta un rendimiento por hectárea mayor al promedio mundial, con condiciones semejantes a las que presentan Vietnam y China en el mismo rubro. Su mayor producción se obtiene en enero, febrero y marzo (Figura 1.15) (Flores, 2015).

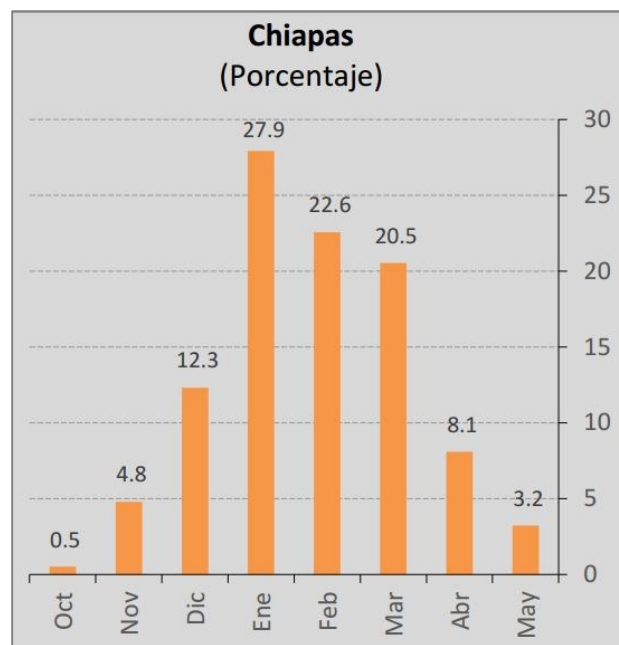


Figura 1.15. Estacionalidad de la producción en Chiapas de café cereza 2001-2015. (SIAP, 2016).

Los estados de Chiapas, Oaxaca y Veracruz concentran el mayor número de productores de

café, en conjunto suman alrededor del 74% de los caficultores del país, sin embargo, Chiapas es el estado que tiene la mayoría, con 36% del total; por esta razón, es el estado que concentra una tercera parte de la totalidad de las parcelas cafetaleras y también la mayor proporción de superficie destinada a la caficultura. Además, en Chiapas se encuentra el 41% de los hombres y el 24% de las mujeres que producen el grano, esto indica que también es este estado en el que se cuenta con la mayor fuerza de trabajo de este sector (Flores, 2015).

Chiapas se ha posicionado como el principal productor de café orgánico del mundo y del país, su producto se ha caracterizado por ser cultivado en su mayoría por hombres y mujeres indígenas bajo la sombra de árboles nativos y sin hacer uso de agroquímicos para evitar la contaminación del suelo (Mariscal, 2011). En el año 2013 este estado concentró el 35% de los beneficiarios totales del Programa Procampo Productivo que permitió otorgar apoyos directos a los ingresos de los productores para incrementar el volumen de producción por hectárea y fomentar la comercialización (Flores, 2015).

1.3.1 La denominación de origen del Café Chiapas

Una denominación de origen es el nombre atribuido a un producto procedente de una región geográfica del país específica (Secretaría de Economía, 2015), dicho producto debe cumplir con características, componentes, forma de extracción y procesos de producción o elaboración, atribuidos a factores naturales y humanos exclusivos de la región (Diario Oficial de la Federación, 2003).

La denominación de origen está vinculada a las tradiciones de los pueblos pertenecientes a cada región y al aprovechamiento que practica su población sobre las condiciones geográficas y los recursos naturales con la finalidad de generar un producto único; adicionalmente, se promueve y transmite el conocimiento sobre la conservación de los recursos locales y la producción sustentable a las nuevas generaciones. Estas denominaciones son protegidas con una declaración por parte del Instituto Mexicano de la Propiedad Industrial a petición de una persona física o moral, cuando un producto está respaldado por la tradición de los pueblos, el aprovechamiento de las condiciones geográficas y los recursos naturales particulares de una región que generan un producto único (Instituto Mexicano de la Propiedad Industrial, 2018).

1.3.2 Particularidades del Café Chiapas

En México se han realizado 16 declaraciones de denominación de origen, entre ellas, la denominación otorgada al café que se produce en Chiapas; específicamente en las regiones de Ángel Albino Corzo, Bochil, Comitán, Copainalá, Motozintla, Ocosingo, Ocozocoautla, Palenque, Pichucalco, San Cristóbal de las Casas, Tapachula y Yajalón. El 27 de agosto de 2003 se publicó en el Diario Oficial de la Federación (DOF) la declaración general de protección de la denominación de origen Café Chiapas, en la cual se enmarcan las características, componentes, forma de extracción y procesos de producción o elaboración, específicos y detallados que convierten al café de Chiapas en un producto genuino y particular (Diario Oficial de la Federación, 2003).

Los elementos establecidos en dicha declaración deben ser cumplidos en todo momento por los productores para que la denominación de origen sea respetada por el Instituto Mexicano de la Propiedad Industrial (IMPI). Lo anterior otorga al Café Chiapas un sistema de protección, asimismo, el reconocimiento como un producto con valor agregado y calidad certificada que posiciona a México como un país distinguido a nivel mundial por su café (Instituto Mexicano de la Propiedad Industrial, 2018).

En el DOF se han publicado todas las normas que contienen los requisitos y estándares que deben ser cumplidos por los productos, durante todo el proceso de producción del café para tener derecho a utilizar el nombre Café Chiapas. De acuerdo con lo establecido en la declaratoria publicada el 28 de agosto de 2003 en el DOF, el Café Chiapas es considerado como tal cuando el producto pertenece a la especie *Coffea arabica*, ya sea como café verde o como café tostado/molido y cumpla con las especificaciones sensoriales establecidas por los parámetros de medición del aroma, acidez, cuerpo y sabor, además de las especificaciones físicas de la planta, como son la altitud, el color, la humedad, el tamaño y la forma (Diario Oficial de la Federación, 2003).

El 26 de marzo de 2007 se publicó en el DOF la Norma Oficial Mexicana NOM-169-SCFI-2007, es aplicable al café verde, tostado, en grano o molido producido en Chiapas y está complementada por las siguientes normas publicadas en el DOF (Diario Oficial de la Federación, 2007):

NOM-030-SCFI-2006, Información Comercial - Declaración de cantidad en la etiqueta - Especificaciones. Publicada el 6 de noviembre de 2006.

NOM-037-FITO-1995, Por la que se establecen las especificaciones del proceso de producción y procesamiento de productos agrícolas orgánicos. Publicada el 23 de abril de 1997.

NOM-106-SCFI-2000, Características de diseño y condiciones de uso de la contraseña oficial. Publicada el 2 de febrero de 2001.

NMX-B-231-1990, Cribas para clasificación de materiales granulares. Declaratoria de vigencia publicada el 9 de enero de 1991.

NMX-EE-120-1981, Envase - Papel - Bolsas para envasar café - Dimensiones. Declaratoria de vigencia publicada el 9 de abril de 1981.

NMX-EE-178-1984, Envase - Textiles - Henequén - Sacos para envasar café - Especificaciones. Declaratoria de vigencia publicada el 16 de octubre de 1984.

NMX-F-013-SCFI-2000, Café puro tostado, en grano o molido sin descafeinar o descafeinado - Especificaciones y métodos de prueba. Declaratoria de vigencia publicada el 18 de agosto de 2000.

NMX-F-083-SCFI-1996, Alimentos - Determinación de humedad en productos alimenticios. Declaratoria de vigencia publicada el 14 de julio de 1996.

NMX-F-551-SCFI-2008, Café Verde - Especificaciones, preparaciones y evaluación sensorial. Publicada el 18 de febrero de 2009. (reemplazo de **NMX-F-551-1996-SCFI**, Café Verde - Especificaciones y Métodos de Prueba. Declaratoria de vigencia publicada el 11 de abril de 1997.

Las variedades de *Coffea arabica* consideradas como Café Chiapas son las de porte alto Bourbon, Maragogype, Mundo Novo y Typica, y las de porte bajo Catimor, Catuai, Caturra, Garnica y Pacamara, siempre y cuando sean producidas bajo las especificaciones de las normas anteriores (Diario Oficial de la Federación, 2003).

De acuerdo con la NOM-169-SCFI-2007, el Café Chiapas se clasifica en distintos tipos según la altura en la que se cosecha y las características de color, humedad, tamaño, forma y defectos de los granos, la clasificación se designa por su calidad, por su grado de transformación, por su grado de tueste y por su grado de molienda (Cuadro 1.2) (Diario Oficial de la Federación, 2007).

La producción de Café Chiapas se caracteriza por utilizar la sombra forestal facilitada por otros tipos de vegetación de la región y por cultivarse exclusivamente arriba de los ochocientos metros sobre el nivel del mar, de igual manera, se requiere de condiciones ambientales específicas como la precipitación de 2,500 a 5,000 mm anuales, la temperatura promedio anual de 18° a 25°C, suelos profundos con valores de pH de 4,5 a 7,0 y alto contenido de materia orgánica (Diario Oficial de la Federación, 2003).

Cuadro 1.2. Clasificación del Café Chiapas (DOF, 2007).

Tipo de clasificación	Tipo de café
Por calidad	Estrictamente altura Altura Extra prima lavado Prima lavado
Por grado de transformación	Café verde Café tostado: -en grano -molido
Por grado de tueste	Tostado claro Tostado medio Tostado oscuro
Por grado de molienda	Molido muy fino Molido fino Molido medio Molido grueso

Para que el producto chiapaneco siga conservando el título de Café Chiapas debe cumplir en todo momento con las normas establecidas por la ley correspondiente, de lo contrario el IMPI podrá hacer una declaración en

la que se solicite la modificación de las normas correspondientes, o de ser necesario, la cancelación de la denominación de origen (Diario Oficial de la Federación, 2003; Instituto Mexicano de la Propiedad Industrial, 2018).

1.4 La producción de café como sustento económico familiar

El café se ha vuelto un producto que logra integrar a miles de personas de todo el mundo, es un producto altamente comercializado que funciona como el sustento de miles de familias de productores, principalmente de familias campesinas e indígenas de distintas partes de América Latina, Asia y África (Moguel & Toledo, 1996; Olivar & Bustamante, 2016).

La Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación calcula un estimado de 125 millones de personas de 70 países que viven del cultivo de café, en los que se incluyen 26 millones de pequeños productores (Ruiz, 2018).

La caficultura (Figura 1.16) es una labor que ha transformado la economía de las regiones productoras en zonas en las que los ingresos totales de los productores están ligados especialmente a la unidad económica familiar, es decir, se basa en la utilización del trabajo de los

integrantes de la familia (Figura 1.17) sin ocupar el trabajo de personas externas asalariadas (Chayanov, 1985).

Cultivar café se ha vuelto la actividad que miles de familias (Figura 1.18) desarrollan para obtener el

sustento económico que necesitan para sobrevivir, lamentablemente estas personas están sujetas a varios factores como los precios del producto que se rigen por la bolsa de Nueva York (*Coffea arabica*) y de Londres (*Coffea canephora*), que los ubica en uno de los eslabones



Figura 1.16. Práctica de la caficultura en el estado de Chiapas. (conacytprensa.mx).



Figura 1.17. Niños participando en la cosecha del café en el estado de Chiapas. (swissinfo.ch).

más débiles de la cadena de producción en términos de ingresos y de bienestar, especialmente si hay bajos rendimientos y baja productividad de las parcelas o si las plantaciones se ven afectadas por plagas como la broca o la roya, además, la volatilidad de la producción es la causa de ciclos de escasez y desempleo cada año (Olivar & Bustamante, 2016; Ruiz, 2018).

La producción de café es una actividad que forma una gran parte de la vida personal y familiar de la población agricultora de este grano. En zonas caficultoras una de las formas de socialización más común dentro de la familia y en la

comunidad, es precisamente la enseñanza y el aprendizaje de las técnicas de producción de café, pues para ellos la organización es sinónimo de unidad y fuerza, además se considera como una vía para lograr progreso familiar. Los terrenos, parcelas o fincas de cultivo funcionan como un objeto de herencia familiar que resulta de muchos años de experiencia y práctica en la familia de origen, y que conlleva a la formación de un sentimiento de arraigo y orgullo hacia el café, así como la integración de su identidad a las actividades colectivas de la caficultura (Canabal, Contreras & León, 2006; Centro de Investigación, Capacitación, Asesoría y Promoción, 2007).



Figura 1.18. Familia caficultora de Chiapas. (migracionycafeblog.wordpress.com).

Las familias que laboran en la producción de café sostienen una organización social en la que a cada integrante de la familia se le asigna un rol de trabajo, en ocasiones sin importar el género y la edad, con la finalidad de mantener la supervivencia de la familia mediante la actividad cafetalera. Entre los productores es valorada esta organización como una responsabilidad colectiva hacia el entorno social y ambiental, al mismo tiempo, existe el compromiso de compartir los conocimientos y nuevas técnicas para aspirar a una mejor producción como comunidad (Centro de Investigación, Capacitación, Asesoría y Promoción, 2007).

La mayor parte de los pequeños productores de café son personas indígenas que viven en zonas con poca diferenciación socioeconómica y que preservan sus propias prácticas, así es como los campesinos aportan su sabiduría al proceso productivo del café, como mejorar la calidad de las plantas a través de injertos de variedades más resistentes (Escamilla & Zamarripa, 2002).

Las familias productoras de café conforman la unidad núcleo del sistema en el que toman parte las instituciones cafeteras, el Estado y el mercado mundial; por esto, los cambios económicos y técnicos afectan profundamente los ingresos de los productores del grano que se han vuelto

subordinados a las condiciones del mercado (Rodríguez, 1998).

Cuando en el cultivo existe una plaga como la broca o la roya, el cultivo está en riesgo de sufrir pérdidas que significan una disminución en los ingresos de las familias caficultoras, en vista de esto, es preciso tener en cuenta que el cultivo de café requiere de atención especial y de mejoras permanentes de las prácticas, como la renovación de plantas y el abonamiento de los suelos (Centro de Investigación, Capacitación, Asesoría y Promoción, 2007).

En México, el gobierno del estado de Chiapas, a través de la Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación y el Instituto del Café de Chiapas, publican convocatorias para los pequeños productores de café de ese estado para recibir apoyos monetarios con fines de impulsar su producción, a través de la adquisición de plantas de café mejoradas y paquetes tecnológicos (Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación, 2017).

Las familias caficultoras chiapanecas dependen en mayor medida de los ingresos que la producción de café les genera, los productores que reciben apoyos por parte del gobierno percibieron en el año 2016 más de 4 mil millones

de pesos para apoyar el aumento de su productividad (Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación, 2017; Instituto del Café de Chiapas, 2018).

1.5 La producción de café en Ocosingo, Chiapas

Para 2015, el municipio tenía una población de 218,893 habitantes, de los cuales alrededor del 60% eran indígenas con alto grado de marginación, la mayoría de ellos se dedicaban a la agricultura de autoconsumo con cultivos de granos básicos como el maíz y el frijol, hortalizas y recolección de frutos como el plátano, naranja, piña, mandarina, guanábana, zapote y mamey (Instituto Nacional de Estadística y Geografía, 2017).

Además de la agricultura, en Ocosingo existen las áreas naturales protegidas: Reserva de la Biosfera Lacan-tún, Área de Protección de Flora y Fauna Chan-kin, Monumento Natural Yaxchilán, monumento Natural Bonampak, Área de protección de Flora y Fauna Silvestre Nahá, Monumento Arqueológico Toniná, Área de Protección de Flora y Fauna Metzabok y Reserva Integral de la Biosfera Montes Azules, que constituyen un gran potencial natural aprovechado con el fomento al ecoturismo,

funcionan como una alternativa de sustento económico para sus habitantes (H. Ayuntamiento de Ocosingo, 2012; Comisión Nacional de Áreas Naturales Protegidas, 2016).

Sin embargo, en Ocosingo la caficultura es la actividad económicamente más importante para los pobladores, es practicada especialmente para la exportación de café orgánico y funciona como su fuente principal de ingresos (Vargas, 2007). En este lugar, el cultivo de café es el que genera mayores remuneraciones que otros productos como el frijol y el maíz, en un principio su producción era convencional; no obstante, en la última década se ha priorizado la caficultura con procesos orgánicos como medida de preservación de los recursos naturales y sus servicios ambientales (Vargas, 2007).

La producción orgánica del café se practica principalmente por los beneficios económicos que tienen los caficultores cuando reciben apoyos de los programas de gobierno (Figura 1.19) para poder vender su café a precio internacional, este precio se adquiere cuando el producto cumple con las especificaciones establecidas en la normatividad mexicana y tiene la denominación de origen Café Chiapas. Las familias caficultoras del municipio poseen en promedio entre 1 y 5 hectáreas de tierra que trabajan haciendo uso intensivo de la fuerza de trabajo familiar, que en

muy pocas ocasiones combinan con trabajo asalariado de personas guatemaltecas (Vargas, 2007).

En Ocosingo se produjo una cosecha de más de 7,398 hectáreas de café en 2011; sin embargo, después de la crisis ocasionada por la roya, en el año 2017 sólo se lograron cosechar 5,550 hectáreas, aún con esta situación, el café sigue siendo el cultivo con mayor valor de producción y

rendimiento del municipio (H. Ayuntamiento de Ocosingo, 2012; Servicio de Información Agroalimentaria y Pesquera, 2017).



Figura 1.19. Caficultores del municipio de Ocosingo que recibieron apoyos del gobierno. (incafech.gob.mx)



II PERCEPCIÓN REMOTA

Contenido

Este capítulo está dividido en seis apartados, que a su vez tienen subdivisiones. El primer apartado contiene la información básica sobre percepción remota, se explica qué es, cómo funciona y para qué se utiliza, así mismo se describen algunas de las técnicas más empleadas, también se aborda el espectro electromagnético, la reflectancia espectral y su separabilidad. El segundo contiene el desmezclado espectral y uno de sus métodos. En el tercero se detalla información sobre la imagen digital proveniente de un sensor remoto, y las transformaciones básicas que se pueden realizar sobre una imagen satelital. En el cuarto apartado se desarrolla el tema de las clasificaciones digitales, tanto los métodos supervisados como los no supervisados, además el método de clasificación *Random Forest*. El quinto está dedicado a la verificación de la clasificación, qué es y para qué se utiliza, así como la manera de hacerlo. Finalmente, el sexto apartado contiene una descripción sobre el programa de observación espacial del que proviene el satélite elegido para este trabajo, además se describen las características de las imágenes Landsat 8.

2.1 Percepción remota

La **percepción remota, teledetección o *Remote Sensing*** se define como la ciencia, técnica o arte con la que es posible la colección y observación de información de objetos, áreas o fenómenos de la superficie del planeta sin estar en contacto físico con ella (Chuvienco, 2010). Esto se realiza con la información que se obtiene a través de imágenes capturadas con aparatos o sensores remotos especializados (Figura 2.1) que se encuentran instalados en plataformas espaciales, aviones, dirigibles y otros medios como los diversos satélites que orbitan el planeta (Instituto Nacional de Estadística y Geografía, 2016).

Las imágenes capturadas permiten analizar la interacción que existe entre la energía electromagnética y el objeto de estudio, por medio de distintos métodos de preprocesamiento y procesamiento de la información (Lillesand, Kiefer & Chipman, 2008; Araque & Jiménez, 2009; Asner, Knapp, Balaji & Páez-Acosta, 2009; Chuvienco, 2010; Paine & Kiser, 2012). La percepción remota es una tecnología científica que puede ser utilizada para medir y monitorear características biofísicas y diversas actividades humanas (Jensen, 2000). Algunas de sus aplicaciones que se han practicado en diversas

disciplinas y ciencias del medio ambiente se muestran en el cuadro 2.1.

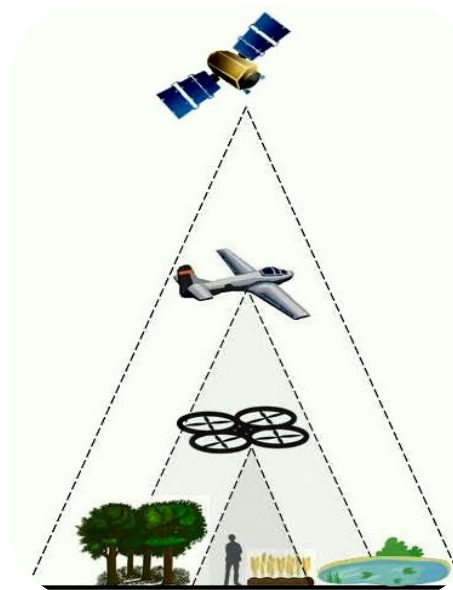


Figura 2.1. Percepción remota. (rezatec.com).

El uso de la percepción remota tiene algunas ventajas, pues facilita la reducción de tiempo y costos del proceso de obtención de la información, también es posible tener una cobertura global y periódica del objeto de estudio y de zonas con poca accesibilidad. Además, es posible analizar información de regiones no visibles por el ojo humano (Chuvienco, 2010).

De manera muy básica, un sistema de percepción remota está compuesto por tres principales elementos: el sensor, el objeto observado y el flujo energético; éste último es parte de las tres formas

Cuadro 2.1. Aplicaciones de la percepción remota (Modificado de Chuvieco, 1995).

Aplicaciones
Análisis en tiempo real de masas nubosas de escala media y pequeña
Cartografía temática, por ejemplo: <ul style="list-style-type: none"> – de áreas quemadas y seguimiento de los ritmos de repoblación natural – de la cobertura vegetal y uso del suelo – de cultivos por especies – geológica para la exploración mineral y petrolífera – de nuevos depósitos volcánicos – térmica de la superficie del mar
Control de la acumulación nival, de la fusión y de los cambios previsibles en la disponibilidad de energía hidroeléctrica
Control del movimiento de icebergs en zonas polares
Control de los movimientos del <i>Gulf-Stream</i> y otras corrientes marinas
Control de pastizales efímeros para estudiar efectos de la sequía y excesivo pastoreo
Estimación de modelos de escorrentía y erosión
Estudio de la erosión de playas y arenales
Inventario del agua superficial
Inventario regional del medio ambiente para preparar estudios de impactos ambientales
Medidas de aguas superficiales y humedales para evaluar la situación del hábitat para aves acuáticas
Predicción del rendimiento de cultivos
Rápida evaluación de condiciones de estrés en la vegetación, por efectos de la sequía o deforestación
Realización de inventarios forestales
Selección de rutas óptimas para nuevas vías de comunicación
Verificación de contenidos de salinidad en las principales corrientes de agua
Verificación y control de la calidad física del agua, turbidez y contenido de algas

de adquirir información y es un flujo que puede provenir de la reflexión de la luz solar sobre el objeto, de la energía que emite el objeto o de la energía que emite el sensor y después es reflejada (Figura 2.2) (Chuvienco, 2010).

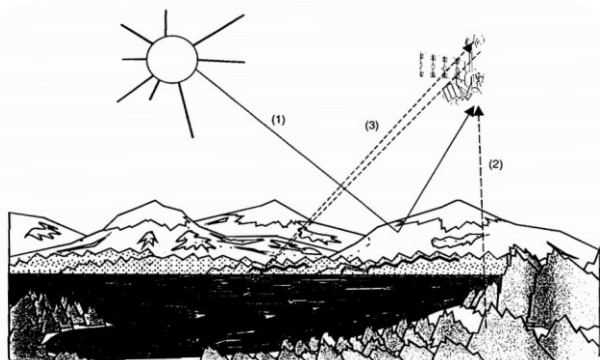


Figura 2.2. Formas de teledetección: (1) reflexión; (2) emisión; (3) emisión-reflexión. (Chuvienco, 2010).

La teledetección por reflexión es la forma de adquirir información de mayor importancia, pues la energía utilizada es la emitida por el Sol, éste ilumina la superficie y ésta a su vez refleja la energía según las características del objeto (Chuvienco, 2010). Acorde con Lira (2009) y el Instituto Nacional de Estadística y Geografía (2016), el proceso de percepción remota está compuesto por los diez elementos que se muestran en el cuadro 2.2 (Figura 2.3).

La radiación electromagnética de la energía constituye la base de los sistemas de percepción remota, esta radiación ha sido explicada como un

movimiento ondulatorio, en donde la energía es transmitida de un lugar a otro de acuerdo a una armonía continua que se mueve a la velocidad de la luz y que contiene un campo de fuerza eléctrico, así como uno magnético (Chuvienco, 2010).

Lo anterior se traduce como una onda electromagnética, que es considerada una vibración o perturbación emitida desde alguna fuente de energía, por medio de la cual la radiación es propagada a través del espacio y del tiempo. Los campos de fuerza de una onda electromagnética son fluctuantes y se propagan de manera oscilatoria, periódica y perpendicular a la dirección en la que viajan (Bolemon, 1985; Levin, 1999; Jensen, 2000; Máximo & Alvarenga, 2006).

Las ondas electromagnéticas son entonces flujos energéticos que constan, además de los campos de fuerza, de cuatro elementos primordiales: la longitud, la frecuencia, la amplitud y la velocidad (Figura 2.4). La longitud se describe como la distancia que existe entre dos crestas continuas de la onda y se expresa en micrómetros (μm) o nanómetros (nm), la frecuencia se refiere al número de oscilaciones o ciclos que pasan por un punto fijo en un lapso determinado y se expresa en Hertz (Hz) o ciclos por segundo, estos elementos son inversamente proporcionales (Figura 2.5), pues a mayor

longitud de onda, menor frecuencia y menor contenido energético (Jensen, 2000; Chuvieco, 2010).

Cuadro 2.2. Elementos del proceso de percepción remota. (INEGI, 2016; Lira, 2009).

Elemento	Descripción
La fuente de energía o iluminación	Es la radiación electromagnética emitida sobre un objeto, puede provenir de un sensor activo o pasivo. El sensor activo posee su propia fuente de iluminación, mientras que el sensor pasivo recibe la energía emitida por una fuente de iluminación natural externa.
La atmósfera	La energía interactúa con ella cuando viaja del sensor al objeto, así como a la inversa, pues dispersa y absorbe la energía.
El paisaje	Es el medio apreciable de la superficie terrestre.
La escena	Contiene el área en la que está enfocada la visión del sensor.
Los objetos	A partir de sus propiedades envían una respuesta al sensor en forma de energía reflejada.
El sensor remoto	Es un aparato o dispositivo con el que se captura y recoge la energía electromagnética reflejada por un objeto.
La plataforma o satélite	Es la base en la que se encuentra instalado el sensor remoto.
Transmisión, recepción y procesamiento	La energía recibida por el sensor se transmite a una estación receptora, en donde la información se procesa cualitativa y cuantitativamente.
Interpretación y análisis	Se basa en interpretar la imagen para extraer información de los objetos captados.
Usuario final	Hace uso de la información extraída a través de su aplicación con métodos de percepción remota.

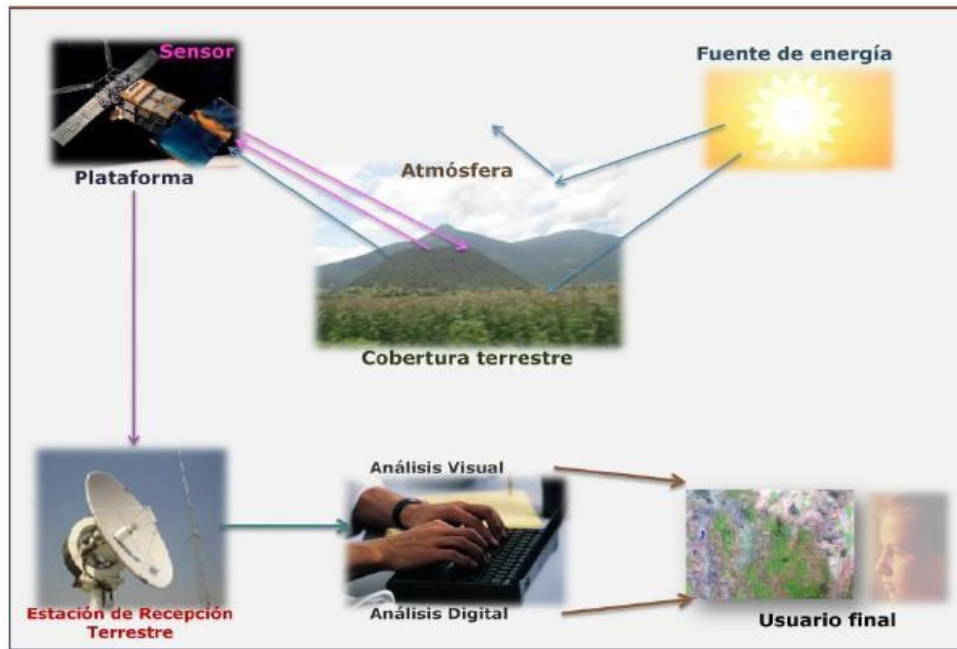


Figura 2.3. Componentes del proceso de percepción remota. (INEGI, 2016).

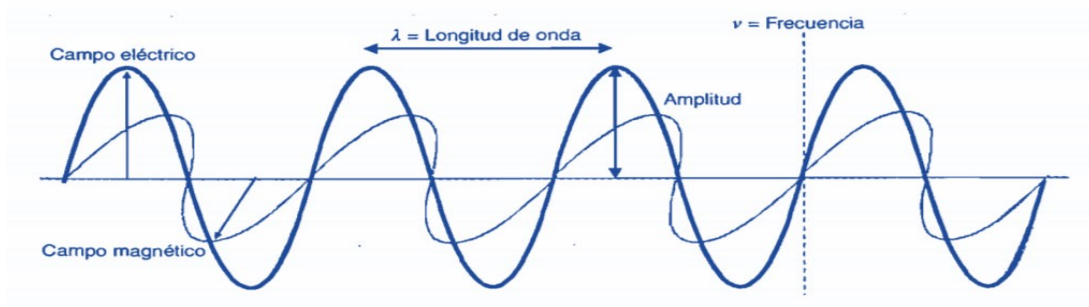


Figura 2.4. Onda electromagnética. (Chuvieco, 2010).

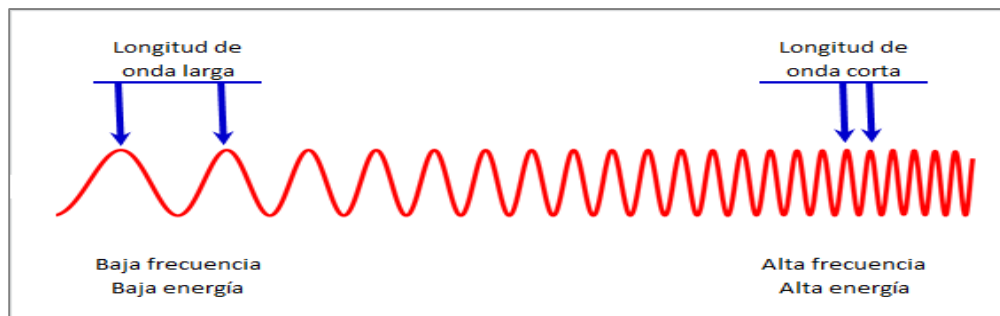


Figura 2.5. Longitudes y frecuencias de las ondas electromagnéticas. (wordpress.com).

La amplitud es el pico máximo de la cresta y es expresado en Watts por metro cuadrado (W/m^2), el elemento de velocidad de la onda está expresado como $3 \times 10^8 m/s$, también llamada velocidad de la luz (Jensen, 2000; Lillesand et al., 2008; Khorram, Koch, van der Wiele & Nelson, 2012).

2.1.1 El espectro electromagnético

Cualquier tipo de energía emitida en forma de radiación se puede describir a partir de su longitud de onda o frecuencia, como los rayos gamma, los rayos X, los rayos ultravioleta, la luz visible, la luz infrarroja, las microondas y las ondas de radio, que en conjunto, son ondas que conforman el espectro electromagnético (Figura 2.6) (DiBiase, 2018).

Este espectro se representa como la sucesión continua de valores de longitud de onda y está organizado en bandas, canales o regiones que van

desde las longitudes de onda corta que tienen como unidad de medida la micra (μm), hasta las longitudes más largas que se miden en centímetros, metros o en valores de frecuencia como los gigahercios (GHz). El espectro electromagnético clasifica, según las longitudes de onda, toda la energía que se mueve armónicamente a la velocidad constante de la luz (Jensen, 2000; Elachi & van Zyl, 2006; Chuvieco, 2010; United States Geological Survey, 2013).

La porción visible del espectro electromagnético es pequeña, pues la sensibilidad espectral del ojo humano sólo permite percibir la radiación que va desde $0.4 \mu m$ hasta $0.7 \mu m$ aproximadamente (Lillesand et al., 2008). El espectro visible está compuesto por tres bandas, la primera se denomina azul ($0.4 - 0.5 \mu m$), la segunda es verde ($0.5 - 0.6 \mu m$) y la tercera es la banda del rojo ($0.6 - 0.7 \mu m$) (Sagan, 1994; Lira, 2009).

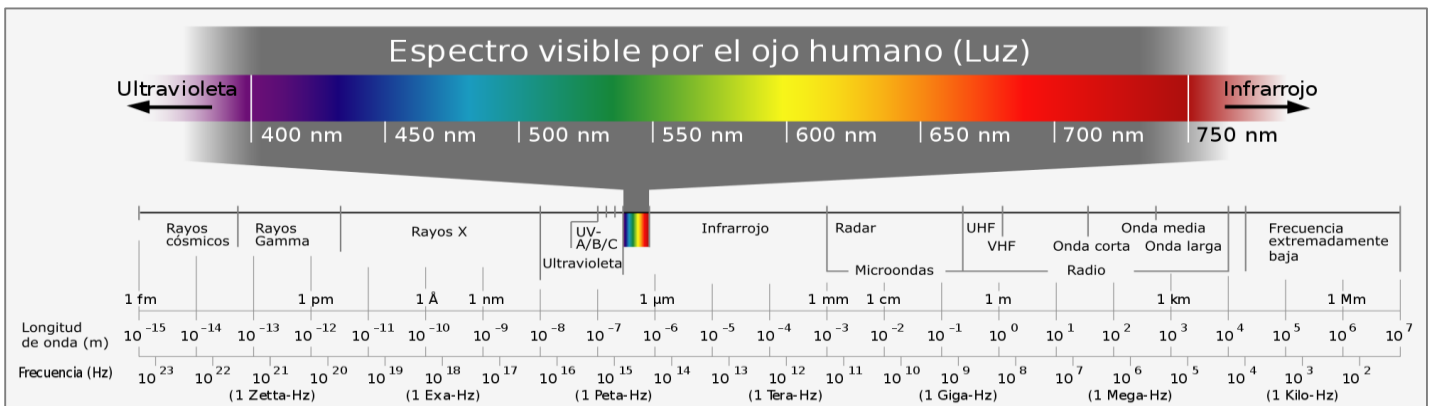


Figura 2.6. El espectro electromagnético. (Horst Frank, 2006).

Además del espectro visible, las otras bandas que son utilizadas con mayor frecuencia en percepción remota son principalmente cuatro, la banda del infrarrojo próximo o cercano que tiene una longitud de onda de 0.7 a 1.3 μm y que es especialmente utilizada para discriminar masas vegetales y humedad. El infrarrojo medio de 1.3 a 8 μm es una banda en la que se mezcla la reflexión de la luz solar y la emisión terrestre. El infrarrojo térmico o lejano que se encuentra en la longitud de onda de 8 a 14 μm contiene la región emisiva del espectro terrestre y permite la estimación de la temperatura del cuerpo emisor, y finalmente, la banda de las microondas con longitud de onda a partir de 1 mm hasta 1 m que es un tipo de energía que atraviesa mejor las nubes (Jensen, 2000; Chuvieco, 2010).

Toda la radiación electromagnética que reciben los sensores remotos atraviesa la atmósfera en su trayectoria de ida y vuelta, este paso por la atmósfera unido a las partículas suspendidas en ella afecta de diversas maneras al flujo radiante, como su velocidad, su longitud de onda, su intensidad y su distribución espectral (Jensen, 2000; Lillesand et al., 2008). Dentro de la composición de la atmósfera se encuentran distintos gases, vapor de agua y aerosoles, de estos componentes el dióxido de carbono, el ozono y el vapor de agua son los que tienen mayor interacción con la energía electromagnética y

producen modificaciones en la radiación que captura el sensor (Chuvieco, 2010).

Estas modificaciones son consecuencia de los efectos de refracción, dispersión y absorción de la atmósfera. **La refracción** se refiere al cambio de dirección y velocidad de cualquier onda cuando pasa de un medio a otro debido a sus diferentes densidades. Siempre que la energía electromagnética se propaga a través de la atmósfera a cualquier ángulo que no sea vertical, existirá la refracción, la cual puede ser predecible (Jensen, 2000).

La dispersión atmosférica (*scattering*) es la difusión impredecible de la radiación electromagnética a causa de las partículas suspendidas en la atmósfera, sus efectos se encuentran en cualquier imagen adquirida por medio de sensores remotos. La dispersión significa que la radiancia directa se reduce y la difusa aumenta debido a los aerosoles, al vapor de agua y otras partículas como el polvo y la humedad (Chuvieco, 2010).

El tipo de dispersión se define de acuerdo a la dimensión de las partículas, su abundancia en la atmósfera y la distancia que atraviesa la onda (Kidder & Haar, 1995; Jensen, 2000; Wallace & Hobbs, 2006), hay tres tipos de dispersión

esenciales, Rayleigh, Mie y la dispersión no selectiva (Figura 2.7).

La dispersión Rayleigh o dispersión molecular ocurre cuando las moléculas de los gases como el oxígeno y el nitrógeno tienen un diámetro menor al de la longitud de onda de la radiación interactuante. Este tipo de dispersión es responsable del color del cielo, pues a medio día la luz solar atraviesa una menor distancia en la atmósfera y es reflejada en las bandas del espectro visible y el ultravioleta, mientras que al atardecer y amanecer la luz atraviesa una mayor inclinación en la atmósfera en las bandas del rojo e infrarrojo cercano, esto provoca que el cielo se observe en tonalidades que van del rojo al naranja (Sagan, 1994; Kidder & Haar, 1995; Moran & Morgan, 1997; Wallace & Hobbs, 2006; Lillesand et al., 2008).

La dispersión Mie o dispersión no molecular existe cuando las partículas suspendidas en la atmósfera, como el polvo y el vapor de agua, tienen un diámetro igual o similar al de las longitudes de onda de la energía que interactúa con ellas y se reflejan en la banda verde y en la roja (Lillesand et al., 2008; Chuvieco & Huete, 2010).

La dispersión no selectiva tiene lugar en las partes más bajas de la atmósfera y sucede cuando

el diámetro de las partículas es más grande que la onda que las atraviesa, se llama no selectiva porque afecta a todas las longitudes de onda por igual y es por eso que las nubes o nieblas se observan de color blanco (Jensen, 2000).

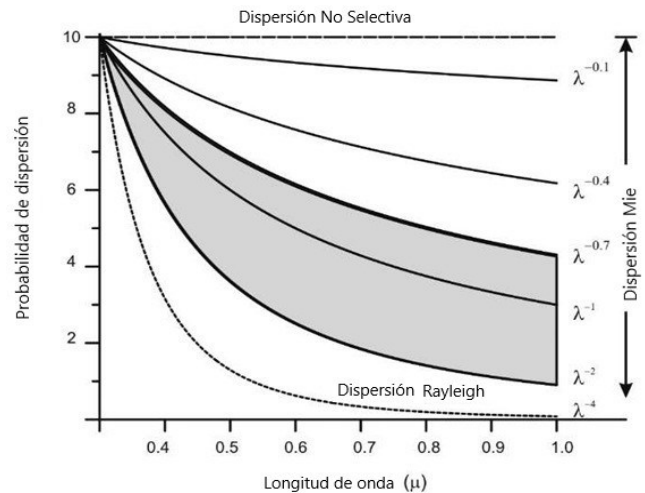


Figura 2.7. Los tipos de dispersión atmosférica.

La absorción atmosférica es el proceso por el cual la energía radiante se absorbe y se convierte en otro tipo de energía, ya sea en su paso por la atmósfera o al tener contacto con la superficie (Jensen, 2000), en otras palabras, es la energía electromagnética que se pierde por el contacto con los elementos de la atmósfera, principalmente oxígeno, ozono, dióxido de carbono, aerosoles y vapor de agua, y se convierte en energía calorífica (Lillesand et al., 2008; Richards, 2013).

La absorción provoca que no en todas las regiones del espectro electromagnético la atmósfera se comporte de manera transparente, es decir, no permite el paso de las ondas electromagnéticas. Las porciones del espectro o bandas en donde la energía radiante sí logra ser transmitida se llaman **ventanas atmosféricas** (Figura 2.8) y son fenómenos en los que existen procesos de dispersión y absorción que se deben tener en consideración para la elección del sensor que se utilice y su sensibilidad espectral, así como la presencia o ausencia de ventanas atmosféricas en

la región espectral de los objetos que se quieren observar (Lillesand et al., 2008; Aguirre, 2009). La energía que reciben los sensores remotos contiene errores generados en el recorrido que hace la radiación a través de la atmósfera, uno de estos errores provoca una modificación en el brillo observable en toda la imagen, pues en algunas de sus bandas se ve distribuido de manera distinta a la realidad del terreno, otro de los errores es a nivel píxel, en donde el brillo en una banda se distorsiona debido a la reflectancia del suelo (Richards, 2013).

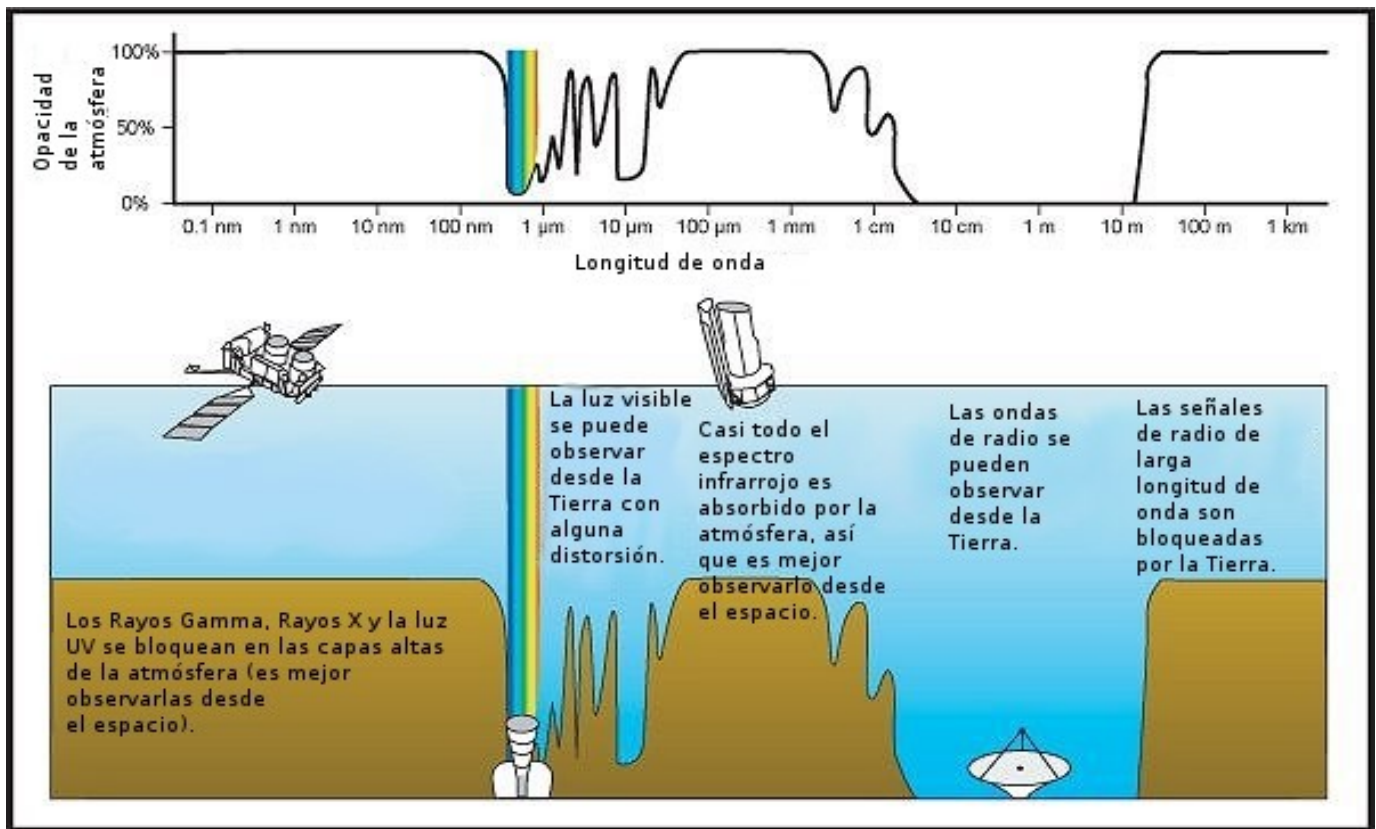


Figura 2.8. Ventanas atmosféricas. (NASA, 2010).

2.1.2 La reflectancia espectral

La reflexión, la absorción y la transmisión (Figura 2.9) son las tres interacciones que pueden llevarse a cabo a partir de la longitud de onda y la superficie en la que la energía electromagnética incide (Richards, 2013). La energía reflejada puede ser de dos tipos (Figura 2.10), especular (la energía incide en una superficie plana y actúa como espejo) o difusa (la energía incide en una superficie rugosa y se refleja en todas las direcciones). Las reflexiones difusas contienen las características espectrales de los objetos que reflejan la energía, es decir, el color de la superficie reflectante, es por esto que en percepción remota interesa más medir las propiedades de reflectancia difusa del terreno (Lillesand et al., 2008).

La reflectancia es el proceso en el cual la radiación rebota al toparse con cualquier objeto como un cuerpo de agua o nubes (Jensen, 2000).

La reflectancia superficial es la fracción de la energía radiante reflejada en la superficie del planeta que se compone de la reflexión difusa y especular de la energía incidente (Asner et al., 2009).

La reflectancia espectral es una propiedad de las características del terreno que puede ser medida de manera precisa y cuantitativa a través

de un sensor, es la reflectancia que define el color del objeto que se observa y es variante de acuerdo a la longitud de onda y a la capacidad de absorción y reflexión de dicho objeto. Un gráfico de la reflectancia espectral de un objeto se conoce como curva de reflectancia espectral, patrón de respuesta espectral o firma espectral (Figura 2.11), en estos gráficos es posible observar las variaciones de reflectancia e identificar objetos específicos (Lillesand et al., 2008; González, 2012; United States Geological Survey, 2013; DiBiase, 2018).

Las variaciones que existen entre los valores de las firmas espectrales hacen posible identificar y discriminar distintas cubiertas de la superficie terrestre como bosques, zonas agrícolas, suelos y formaciones geológicas, así como evaluar la condición en la que se encuentran, incluso es posible la identificación de distintas especies de vegetación como las coníferas y las caducifolias (González, 2012; DiBiase, 2018).

Una de las ventajas de ver más allá del espectro visible con percepción remota, es que se facilita la representación cartográfica de los objetos que se encuentran en la superficie terrestre en determinado tiempo, a partir de sus firmas espectrales, las cuales pueden ser utilizadas para la clasificación de los distintos elementos, objetos o materiales que se encuentran en la superficie

terrestre (Bertrand, 2007; Campbell & Wynne, 2011; United States Geological Survey, 2013).

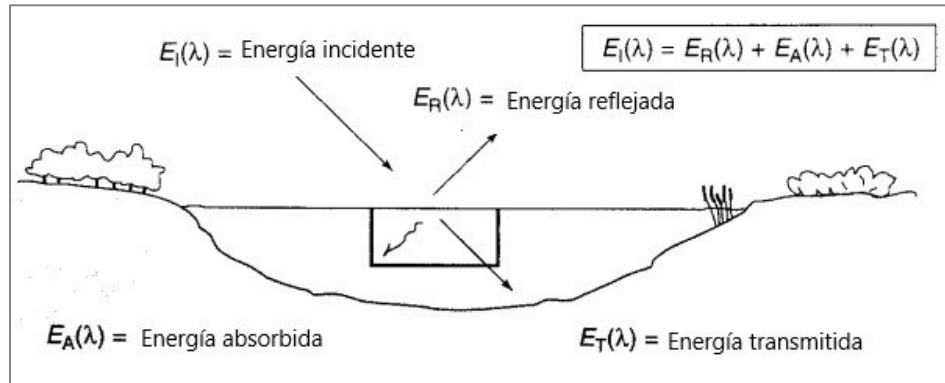


Figura 2.9. Interacción de la energía electromagnética. (Lillesand et al., 2008).

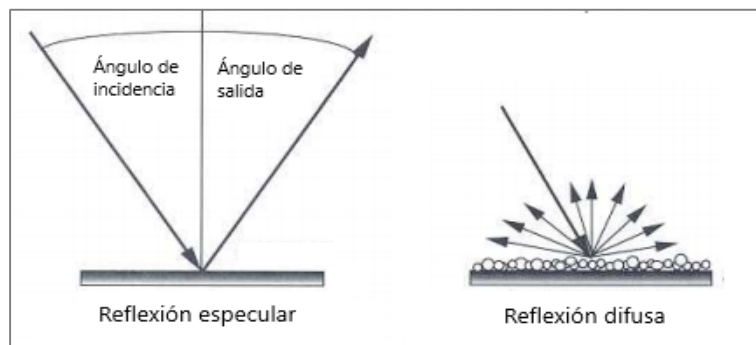


Figura 2.10. Tipos de reflexión. (Jensen, 2000).

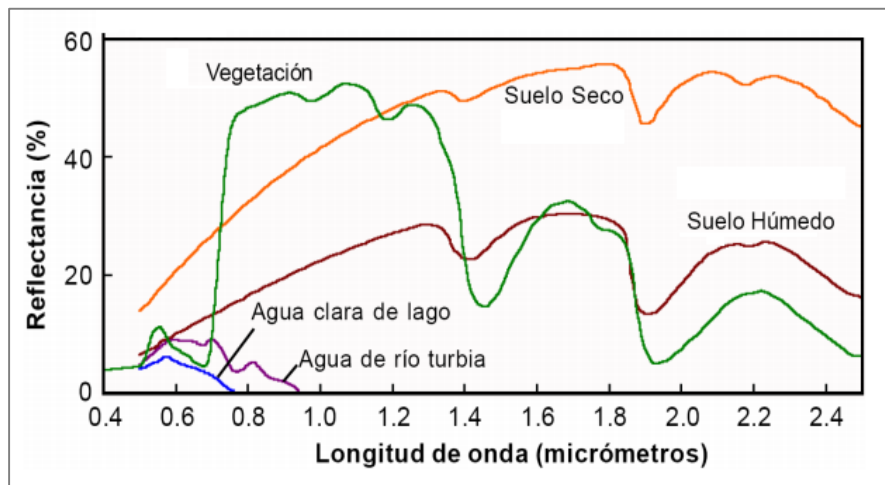


Figura 2.11. Firmas espectrales. (González, 2012).

Los tipos de superficie terrestre que son estudiados con mayor frecuencia incluyen la vegetación, el agua y los suelos expuestos. Para diferenciar entre distintas especies de vegetación por medio de las firmas espectrales, existen varios factores que influyen en la reflectancia, como la cantidad de humedad, la fenología y las características morfológicas de cada especie, así como la pendiente, la asociación con otras especies y las condiciones atmosféricas (Chuvienco, 2010).

La clorofila de la vegetación absorbe la energía en las longitudes de onda entre 0.45 y $0.67 \mu\text{m}$, por esta razón los ojos humanos perciben la vegetación sana de color verde, mientras que en la región del infrarrojo cercano del espectro la reflectancia de la vegetación sana aumenta dramáticamente en el rango de 0.7 a $1.3 \mu\text{m}$, en longitudes de onda más grandes la reflectancia decrece debido al agua contenida en las hojas (Lillesand et al., 2008).

El agua tiene una reflectividad limitada en comparación con la vegetación, pues absorbe la energía a partir de la longitud de onda del infrarrojo cercano, esto, además de los cuerpos de agua como los lagos, aplica para el agua contenida en el suelo y en la vegetación. La reflectancia del agua puede variar de acuerdo a su superficie, es decir, si tiene o no en alguna medida material

suspendido (orgánico o inorgánico) que la vuelva turbulenta, lo cual aumenta su capacidad reflectante. El monitoreo de la reflectancia del agua puede ayudar a determinar la presencia de contaminantes tales como el petróleo y los desechos industriales (Lillesand et al., 2008; Aguirre, 2009).

La firma espectral de los suelos se ve afectada por el contenido de humedad, la textura (la proporción de arena, limo y arcilla), la presencia de óxido de hierro y el contenido de materia orgánica. La presencia de humedad en el suelo disminuye su reflectancia, la presencia de óxido de hierro aumenta su reflectividad en la banda roja y mientras más materia orgánica contenga, menor será su reflectividad. En términos generales, los suelos más finos (predominantemente arenosos) tienden a tener mayor reflectividad espectral (Lusch, 1989; Curran, Foody, Kondratyev, Kozoderov & Fedchenko, 1990; Chuvienco, 2010).

2.1.3 La separabilidad espectral

Las propiedades espectrales de cubiertas como la vegetación, el suelo y el agua son normalmente separables espectralmente, esto depende de la longitud de onda en la que se observe su firma espectral. Los materiales u objetos que se pretenden diferenciar en una

imagen satelital deben ser espectralmente separables, casi en todas las ocasiones es en el rango del infrarrojo cercano donde las firmas espectrales son más fácilmente separables y presentan una mayor distancia espectral (Lillesand et al., 2008).

Un método para calcular la separabilidad espectral de los distintos tipos de cobertura de la superficie consiste en calcular el promedio y la desviación estándar de los valores de cada categoría de objetos y posteriormente graficar los resultados (González, Calera & Quintanilla, 1999), a partir del gráfico (Figura 2.12) es posible observar en qué banda o longitud de onda existe mayor distancia entre los valores de reflectancia espectral y analizar si es suficiente para que los modelos de clasificación no confundan los valores espectrales de categorías similares.

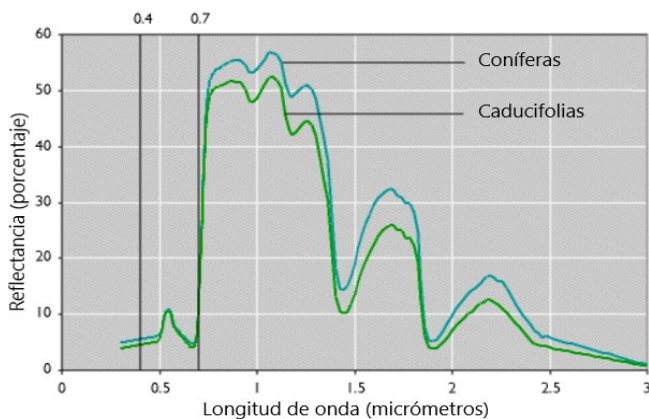


Figura 2.12. Separabilidad espectral de dos tipos de vegetación. (California Institute of Technology, 1999).

2.2 Desmezclado espectral

Las mediciones espectrales captadas por sensores remotos son inevitablemente una mezcla de los elementos encontrados en una determinada extensión espacial (González, 2012).

La resolución espacial de los sensores remotos no es capaz de separar espectralmente las distintas coberturas u objetos, el reconocimiento de que los píxeles que cubren un área de interés están muchas veces combinados por muchos componentes ha creado la necesidad de descomponer o desmezclar dichas combinaciones de objetos, estos objetos pueden cohabitar en un sólo píxel denominado píxel mezcla (Figura 2.13) o píxel mezclado (Keshava & Mustard, 2002; Sánchez & Plaza, 2010).

El desmezclado espectral es un procedimiento en el que el espectro medido de un píxel mezclado se convierte a firmas espectrales puras (*endmembers*) a partir de la estimación de la abundancia de cada firma a nivel subpíxel, entonces, un píxel mezclado es la composición de firmas espectrales puras expresadas en fracciones de abundancia que indican la proporción de cada *endmember* en el píxel, donde la suma de todos los *endmembers* debe ser el 100% en cada banda espectral. Los *endmembers* corresponden normalmente a objetos como la

vegetación, el agua, el suelo y el metal (Keshava & Mustard, 2002; Keshava, 2003).

Los píxeles mezclados representan un problema para la clasificación de imágenes, pues sus características espectrales no son representativas de todos los tipos de cobertura presentes en el píxel, al realizar una clasificación la característica espectral dominante en el píxel será la que determine la clase a asignar. El desmezclado espectral está diseñado para lograr una clasificación subpíxel que sea capaz de representar más certeramente la realidad de la cobertura terrestre (Lillesand et al., 2008; (Silván & López, 2012).

Una de las técnicas de desmezclado espectral es la lineal, que supone que los espectros captados por el sensor pueden ser expresados como una combinación lineal de *endmembers* ponderados a partir de su respectiva fracción de abundancia. El modelo de desmezclado espectral lineal puede ser representado de forma gráfica como un espacio bidimensional a través de un diagrama de dispersión entre dos bandas (Figura 2.14), en donde todos los puntos se encuentran dentro de un triángulo formado por *endmembers*, cualquier punto es una combinación lineal de sus extremos según el lugar que ocupen a partir del centroide del conjunto de puntos (Sánchez & Plaza, 2010; González, 2012).

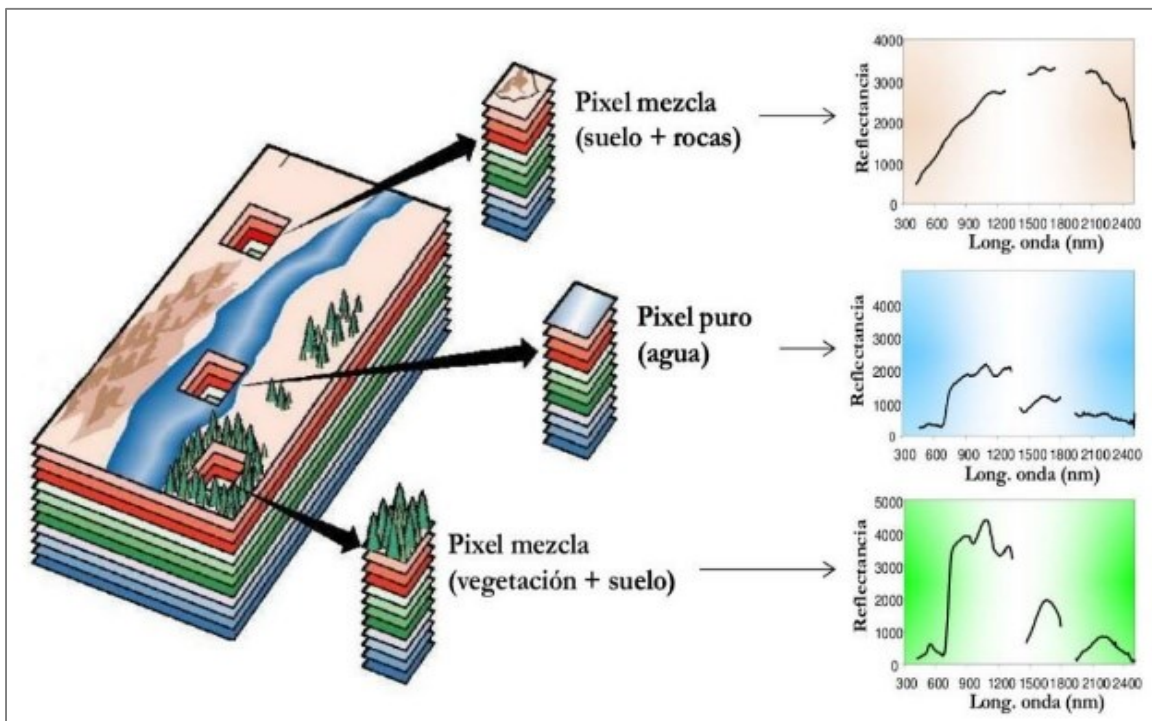


Figura 2.13. Concepto de píxel mezcla. (Sánchez & Plaza, 2010).

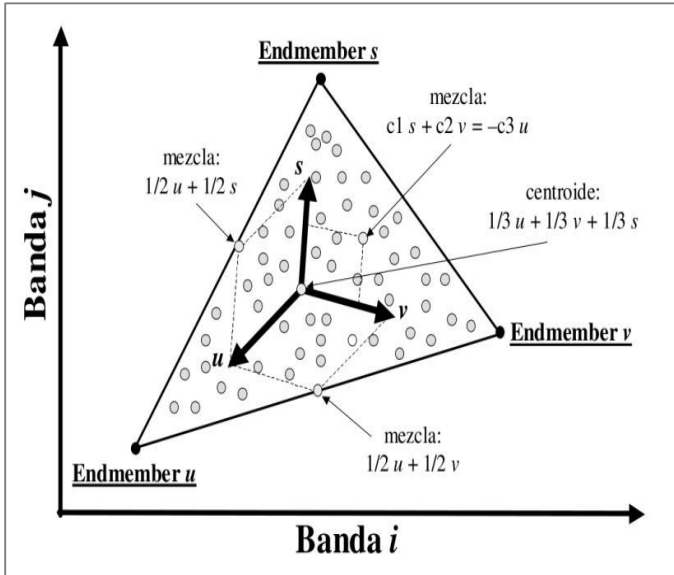


Figura 2.14. Desmezclado espectral lineal. (Sánchez & Plaza, 2010).

2.2.1 El método de Segregación Espectral Automatizada Monte Carlo <AutoMCU>

AutoMCU es un modelo de desmezclado espectral lineal que tiene como finalidad descomponer los píxeles de una imagen satelital en fracciones de cobertura de tres clases, nombradas por sus siglas en inglés como vegetación fotosintética (PV), vegetación no fotosintética (NPV) y superficie descubierta o sustrato expuesto (S). Las propiedades espectrales de la PV se basan en los pigmentos fotosintéticos de las hojas y su contenido de agua, en el caso de la NPV en los compuestos de carbón de la planta senescente o muerta, y para S las propiedades espectrales tienen que ver con suelo

mineral expuesto y varios tipos de roca (Asner et al., 2009).

Para lograr asignar una clase a cada píxel, el método utiliza una biblioteca espectral de cada clase que en total contienen más de 400,000 observaciones espectrales (Figura 2.15) (Asner et al., 2009; Chicas, Omine & Saqui, 2016). Cada biblioteca es una colección de espectros, normalmente tomados en un laboratorio bajo condiciones controladas o en trabajo de campo, para recoger la reflectancia espectral característica (*endmember*) de cada tipo de cobertura y elementos terrestres (Chuvieco, 2010; González, 2012).

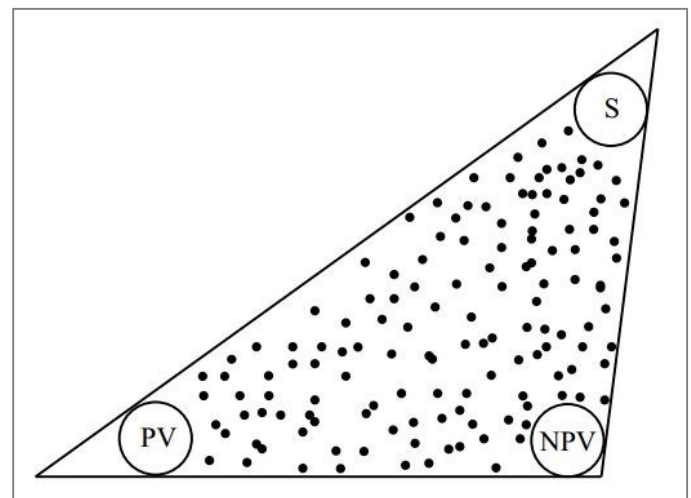


Figura 2.15. Diagrama de dispersión de las bibliotecas espectrales de PV, NPV y S. (CLASlite Team, 2013).

El AutoMCU utiliza los valores de reflectancia superficial de una imagen satelital y los compara con sus tres bibliotecas espectrales a través de un

proceso iterativo que normalmente se repite treinta veces, este proceso es calculado con el algoritmo probabilístico siguiente:

$$\begin{aligned} & \rho(\lambda)\text{píxel} \\ &= \Sigma [C_e \cdot \rho(\lambda)_e] + \varepsilon \\ &= [C_{pv} \cdot \rho(\lambda)_{pv} + C_{npv} \cdot \rho(\lambda)_{npv} + C_s \cdot \rho(\lambda)_s] + \varepsilon \end{aligned}$$

donde $\rho(\lambda)\text{píxel}$ es la reflectancia de la imagen, C_e es la solución de cada fracción de cobertura, $\rho(\lambda)_e$ es la reflectancia de los *endmembers* (e) de cada clase de cobertura (bibliotecas espectrales) en una longitud de onda λ y ε es un término de error. Este algoritmo se repite (itera) hasta que la solución converja en un valor medio para cada una de las tres fracciones de cobertura (Asner et al., 2009).

El resultado es una imagen que contiene siete bandas, donde las primeras bandas (Figura 2.16) son las fracciones de cobertura (banda 1 es S, banda 2 es PV y banda 3 es NPV) y están expresadas en porcentaje. Las siguientes tres bandas representan la incertidumbre de las clases (banda 4 es S, banda 5 es PV y banda 6 es NPV) y están expresadas como la desviación estándar de las iteraciones del modelo, y la séptima banda es el error total expresado como RMSE (*Root-Mean-Square Error*) de los *endmembers* modelados versus los observados (Asner et al., 2009).

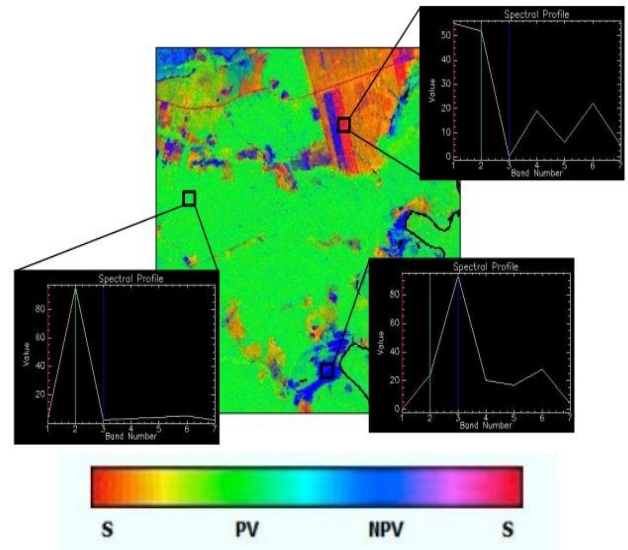


Figura 2.16. Comportamiento espectral de las fracciones de cobertura. (CLASlite Team, 2013).

2.3 Las imágenes digitales

Las imágenes que se obtienen a partir de sensores remotos son almacenadas en formato ráster, todas las imágenes derivadas del procesamiento en algún sistema de información geográfica, como las transformaciones o clasificaciones, en primera instancia tendrán este formato (Lillesand et al., 2008).

En percepción remota, las imágenes digitales, ráster o mapas de bits son el conjunto de parcelas del terreno que un sensor captura, el tamaño o área de la parcela es determinada por la resolución espacial de cada sensor y constituye la unidad mínima de información identificable almacenada en la imagen digital, dicha unidad tiene el mismo ancho y altura en toda la extensión

del ráster (Figura 2.17), se denomina píxel (*picture element*) y puede ser identificado como cada una de las celdas que forman una imagen (Chuvienco, 2010; Muñoz Bollas, 2011).

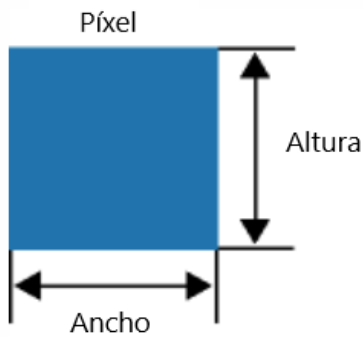


Figura 2.17. Píxel (celda).

El conjunto de píxeles o celdas de un ráster está organizado en una matriz con forma de rejilla o cuadrícula para cada banda (Figura 2.18), cada celda contiene un valor numérico que corresponde a la codificación de la reflectancia del terreno, este valor se denomina *Nivel Digital* o *Número Digital* (ND), y puede representar al fenómeno a través de valores positivos o negativos, enteros o flotantes, incluso es posible representar la ausencia de valores (*NoData*) contenidos en la imagen. Los píxeles, además de tener un ND tienen asignado un número de fila y columna que los ubica dentro de la matriz, mientras que el ráster está geográficamente referenciado. (Lira, 2009; Chuvienco, 2010; Muñoz Bollas, 2011; Environmental Systems Research Institute, 2016).

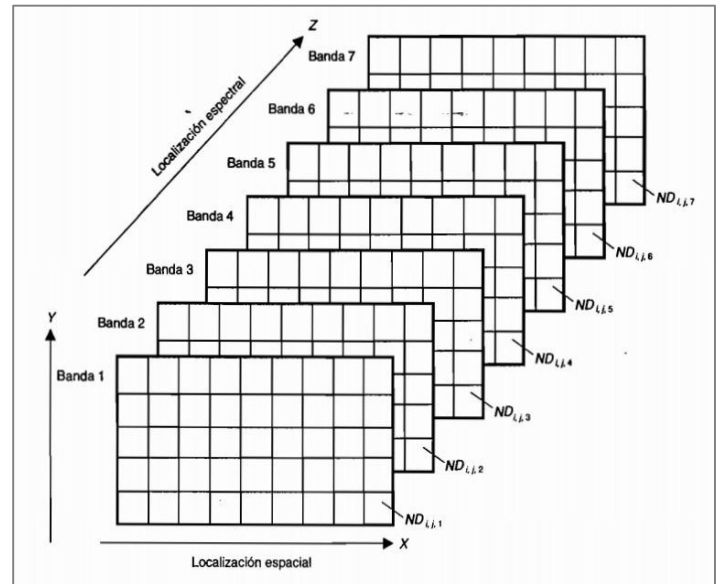


Figura 2.18. Organización de datos en una imagen digital. (Chuvienco, 2010).

La resolución de las imágenes de un sistema sensor se define como su habilidad para discriminar la información con detalle (Estes & Simonett, 1975), **la discriminación** se refiere a la capacidad de distinguir y diferenciar los objetos de la superficie, por su parte, **la información con detalle** se refiere al detalle espacial, al número y anchura de las bandas espectrales, a la temporalidad y a la capacidad para distinguir cambios en la energía que cada sensor detecta. En la literatura se conocen cuatro tipos de resolución: la espacial, la espectral, la radiométrica y la temporal (Tristán, Wainschenker & Doom, 2008; Chuvienco, 2010).

La resolución espacial se define de acuerdo a la distancia entre el sensor y la superficie terrestre, y el campo de visión instantáneo (*Instantaneous Field of View, IFOV*) del sensor, éste determina el tamaño de los píxeles de la imagen (Elachi & van Zyl, 2006). El IFOV (Figura 2.19) es la sección angular (en radianes) que el sensor observa en un momento determinado, es dependiente de la apertura del dispositivo óptico de cada sensor y de la altitud a la que se encuentra (Sobrino, 2001).

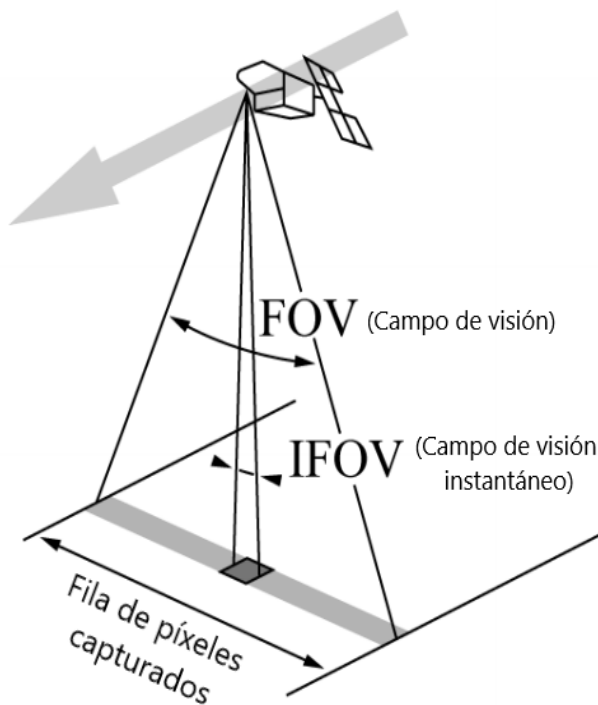


Figura 2.19. IFOV (Campo de visión instantáneo). (Dermanis & Biagi, 2002).

La resolución espectral de un sensor se define por el número de bandas espectrales que puede captar y por la longitud de las mismas, es decir, su capacidad de registrar simultáneamente el

comportamiento de los objetos en distintas longitudes de onda. Mientras más número de bandas proporcione, más fácil será la caracterización espectral de los distintos objetos que componen la superficie (Chuvienco, 2010; Richards, 2013).

La resolución radiométrica se refiere a la capacidad del sensor para detectar variaciones en la radiancia espectral que recibe, esto es, discriminar niveles de brillo que se representan con el número de bits. La cantidad de bits representa el número de grises diferenciados en una imagen digital y estos permiten diferenciar la superficie a través del nivel de brillo (Elachi & van Zyl, 2006; Chuvienco, 2010; Khorram et al., 2012).

La resolución temporal está definida por la frecuencia o periodicidad con la que un sensor cubre una misma escena o porción de la superficie terrestre. La frecuencia está determinada por las características orbitales del satélite o plataforma en la que se encuentra instalado el sensor (Sobrino, 2001; Chuvienco, 2010).

2.3.1 Transformaciones de la imagen

Algunas de las técnicas empleadas en percepción remota son las transformaciones de las imágenes que funcionan para observar con mayor facilidad algunos aspectos temáticos como

las cualidades biofísicas que tiene la superficie terrestre. Estas transformaciones son operaciones lineales (sumas ponderadas, multiplicaciones o divisiones) con las que se crean otras bandas a partir de combinaciones de las bandas originales de una imagen y permiten mejorar la disposición de los datos que ayudan a evidenciar ciertos fenómenos (Chuvienco, 2010).

Una de estas transformaciones es la del cálculo de algún índice de vegetación que son un parámetro obtenido a través de valores de reflectancia de cada píxel realizando operaciones algebraicas en dos o más bandas de una imagen de satélite, su principal función es mejorar la discriminación entre el suelo y la cobertura vegetal (Muñoz, 2013).

También ayudan a resaltar información sobre la biomasa, la densidad, las propiedades fotosintéticas y el contenido de agua en las hojas de la vegetación (Zha, Gao & Ni, 2003; Yengoh, Deni, Olsson, Tengberg & Tucker, 2015), en la mayor parte de los índices de vegetación se combina la banda roja (R) y la banda del infrarrojo cercano (NIR) del espectro después de haberlas corregido atmosféricamente, es decir, con valores de reflectancia (Chuvienco, 2010).

Dentro de los índices utilizados con mayor frecuencia se encuentra el DVI (Índice de

Vegetación Diferenciada) que es probablemente el índice más sencillo, pues sus valores se obtienen del cálculo de la resta de la banda del infrarrojo cercano (NIR) menos la banda roja (R). Es un índice que logra diferenciar el suelo desnudo, el agua y la vegetación. Otro índice de uso común es el PVI (Índice de Vegetación Perpendicular), que usa la distancia perpendicular (ángulo) que tiene cada píxel con respecto a la línea del suelo para calcular los valores de la vegetación a través de la multiplicación del seno de la banda del infrarrojo cercano (NIR) por el coseno de la banda roja (R) (Richardson & Wiegand, 1977).

El SAVI (Índice de Vegetación Ajustado al Suelo) es un índice que permite aislar la información del suelo sobre el que se encuentra la vegetación, es decir, minimiza el efecto del suelo en áreas con baja densidad de vegetación como las zonas áridas (Huete, 1988). Su cálculo se realiza restando la banda del infrarrojo cercano (NIR) menos la banda roja (R), esto dividido entre la suma del NIR más R más L (parámetro de la línea de suelo), y el resultado se multiplica por 1 más L. El valor de la línea de suelo varía de acuerdo a la densidad de la vegetación y en imágenes de satélite se toma como un valor de 0.5 (Cosero, Riaño & Chuvienco, 1998).

Otro de los índices más utilizados para gran variedad de aplicaciones es el NDVI (Índice de Vegetación de Diferencia Normalizada), que al igual que los anteriores, produce una escala lineal de medición que va del -1 al +1. Los píxeles con valores cercanos a 0.1 corresponden a un umbral crítico de la cubierta vegetal, los valores cercanos a -1 pertenecen a suelos desnudos, y valores entre 0.5 y 0.7 son considerados valores típicos de vegetación densa, lo que facilita su interpretación (Rouse, Haas, Schell, & Deering, 1974; Holben, 1986; Sánchez, Torres, Palacios, Aguilar, Pino & Granado, 2000; Chuvieco, 2010).

Sus valores se obtienen a través del cálculo de la relación diferencial entre las bandas del infrarrojo cercano y la banda roja de cualquier sensor (Sobrino, 2001; Zha et al., 2003; Evans & Lasaponara, 2012; Yengoh et al., 2015):

$$\text{NDVI} = \text{NIR} - \text{R} / \text{NIR} + \text{R}$$

El NDVI se ha utilizado ampliamente como indicador del vigor de la biomasa vegetal, en estudios relacionados con el comportamiento de cultivos y de vegetaciones boscosas (Gonzaga, 2014), sin embargo, es conocido que este índice comience a saturar cuando la cubierta vegetal es superior al 50%, por lo que puede llegar a sobre o subestimar algunas zonas (Gill, Phinn, Armston &

Pailthorpe, 2009; Wen, Lees, Jiao, Wanning & Shi, 2010).

Algunos parámetros de la cubierta vegetal que se han estimado de manera satisfactoria a partir del cálculo del NDVI incluyen el contenido de clorofila y agua en la hoja, el flujo neto de dióxido de carbono, la radiación fotosintéticamente activa, la productividad de la vegetación, el índice de área foliar y la dinámica fenológica (Chuvieco, 2008).

2.4 Clasificación digital de imágenes

La clasificación de una imagen es la categorización o agrupación automática de los píxeles de una imagen en conjuntos o grupos que corresponden a distintas clases de cobertura o temas (Lillesand et al., 2008; Lira, 2009; DiBiase, 2018), mediante el acomodo de los píxeles con valores similares de reflectancia espectral en grupos homogéneos (Campbell & Wynne, 2011) y dando como resultado una nueva imagen en la que cada píxel tiene el identificador de la clase en la que fue asignado.

La clasificación digital de imágenes puede ser interpretada como la fase culminante del procesamiento digital de imágenes, pues los datos adquiridos por el sensor no tienen mucho sentido por sí mismos, sino hasta ser convertidos a

categorías que sean fácilmente interpretables. Se trata de establecer un modelo entre medidas digitales de un momento determinado y los tipos de cubierta que representan (Chuvienco, 2010).

Existen dos tipos de clasificación, la primera es la clasificación supervisada y la segunda es la clasificación no supervisada, en algunos casos es posible utilizar ambas técnicas al hacer una combinación de modo híbrido. Se han formulado distintos algoritmos con los que funciona cada clasificador para realizar la asignación de clases en una imagen, la elección del tipo de clasificador se elige de acuerdo a la naturaleza de los datos, los medios disponibles y a la aplicación a la que están destinados los datos clasificados (Townshend & Justice, 1981; Lillesand et al., 2008).

2.4.1 Métodos de clasificación supervisada

En los **métodos de clasificación supervisada** los clasificadores “supervisan” el proceso de categorización de los píxeles al especificar por medio de un algoritmo los descriptores numéricos de varios tipos de cobertura llamados áreas de control o entrenamiento. En otras palabras, la clasificación supervisada parte del conocimiento previo de la zona de estudio y las áreas de control deben ser suficientemente representativas para cada una de

las categorías que se le indican al clasificador, ya que estas áreas son utilizadas para compilar una interpretación numérica que describe los atributos espectrales de cada clase (Lillesand et al., 2008; Chuvienco, 2010).

Este tipo de clasificación consta de tres fases, la primera es la fase de entrenamiento, en la que se identifican las áreas de control para generar una descripción numérica de los atributos espectrales de cada tipo de cobertura, la segunda fase es la de asignación, en la que cada píxel de la imagen es categorizado en la clase a la que más se asemeje, y la tercera fase es la de la verificación de los resultados (Chuvienco, 2010).

Dentro de la clasificación supervisada se encuentran tres clasificadores fundamentales, el primero y el más sencillo es el clasificador de mínima distancia, el cual asigna a cada píxel de la imagen la clase más cercana espectralmente a los valores del centroide de cada categoría, en donde la distancia espectral se mide normalmente al calcular la distancia euclidiana, pero sin ser sensible a la varianza (Chuvienco, 2010).

El segundo clasificador es el de paralelepípedos, cuya tarea es asignar una clase a los píxeles dependiendo si está dentro o no de una determinada distancia del centro de un paralelepípedo (área de entrenamiento

multidimensional). este clasificador no es sensible a la covarianza, lo cual puede resultar en confusión entre clases (Lillesand et al., 2008; Chuvieco, 2010).

El tercer clasificador es el más utilizado en percepción remota y se conoce como clasificador de máxima probabilidad, éste evalúa la media, la varianza y la covarianza de los valores espectrales de cada categoría asumiendo que la distribución de los puntos o áreas de entrenamiento es gaussiana (distribución normal), luego calcula la probabilidad de que el valor de cada píxel pertenezca a cada clase y le asigna la categoría que tenga la probabilidad más alta (Lillesand et al., 2008).

2.4.2 Métodos de clasificación no supervisada

Los **clasificadores no supervisados** no utilizan datos de entrenamiento como base para la clasificación y no requieren conocimiento previo del área de estudio, en lugar de esto, los métodos no supervisados realizan una búsqueda automática de grupos con valores espectrales homogéneos; es decir, utilizan algoritmos que examinan los datos de la imagen para clasificarlos en grupos de coberturas presentes en la escena. Si los valores espectrales de un píxel están dentro del rango de valores de un grupo de cobertura

encontrado en la escena, entonces a dicho píxel se le asigna ese grupo por tener valores espectrales similares (Chuvieco, 2010).

En este tipo de clasificación el resultado es una imagen con distintas clases que fueron espectralmente separadas, pero sin especificar sus identidades, por lo que esta información debe ser agregada a través de la comparación con información de referencia. Los resultados de los métodos de clasificación no supervisada están totalmente dispuestos a la interpretación del usuario (Lillesand et al., 2008; Chuvieco, 2010).

Entre los métodos de clasificación no supervisada existen varios algoritmos capaces de determinar grupos espectrales, a pesar de ello, son dos los principales clasificadores, el primero es el método de *K-means*, al cual se le especifica el número de grupos o conjuntos espectrales que se quieren obtener, luego el algoritmo arbitrariamente ubica ese número de centroides para cada grupo y calcula sus vectores medios, después asigna cada píxel al grupo cuyo vector medio sea más cercano y realiza una revisión para volver a clasificar hasta que no haya cambios significativos en la ubicación de los vectores medios de cada clase (Lillesand et al., 2008).

El segundo clasificador es una variante del anterior y es denominado *ISODATA (Iterative Self-*

Organizing Data Analysis), este algoritmo sigue casi los mismos pasos que el *K-means*, sin embargo, el clasificador *ISODATA* permite que el número de grupos cambie de una iteración a otra, al unir, separar o eliminar grupos (Lillesand et al., 2008).

En cada iteración se evalúan las estadísticas del grupo, si la distancia entre los puntos medios de dos grupos es mínima, se unen en un sólo grupo, en cambio si un grupo tiene una desviación estándar mayor al valor máximo predefinido, éste grupo es separado en dos, los grupos que tienen un menor número de píxeles al número determinado son eliminados. Finalmente, al igual que con *K-means* este proceso se repite hasta que no haya cambios significativos entre los valores de los grupos y los píxeles asignados (Lillesand et al., 2008).

2.4.3 El método de clasificación *Random Forest*

Existen otras técnicas de clasificación de imágenes que se idearon con la finalidad de no ser tan exigentes en cuanto al cálculo y sus requisitos para asignar clases. Una de esas técnicas es la de los clasificadores en árbol (árboles de decisión), pues son sencillos y bastante eficientes, en esencia, consisten en discriminar continuamente cada una de las clases o categorías con base en

valores espectrales, variables de textura, temporales, modelos de elevación o información auxiliar no obtenida por sensores como cartografía temática o límites administrativos (Chuvieco, 2010).

Los clasificadores en árbol tienen una estructura formada por nodos y vectores (ramas), establecen reglas en las que sólo hay dos opciones (dicotómicas) posibles de interpretación (Figura 2.20), es decir, cada clase es discriminada de acuerdo a las variables en las que tiene mejor discriminación y no toma en cuenta cada una de las variables indistintamente (Hansen, Dubayak & Defries, 1996; Friedl & Brodley, 1997; Chuvieco, 2010; Franco, Carrasco, Sánchez & Martínez, 2013).

Los clasificadores en árbol son considerados como un sistema experto o de aprendizaje automático (*Machine Learning*), en otras palabras, de acuerdo con los valores que observa, sus reglas permiten al algoritmo simular la toma de decisiones como si fuera un intérprete humano a partir de operadores lógicos condicionales de forma SI... ENTONCES... (IF... THEN...), es decir, el algoritmo se vuelve capaz de predecir un resultado a través de lo aprendido con los datos de entrenamiento (Richards, 2013).

El método de clasificación *Random Forest* (Bosques Aleatorios) consiste en un conjunto de

árboles de decisión que son independientes entre sí, donde el número de árboles a crear es definido por el usuario y cada árbol elige un subconjunto aleatorio del total de las variables con la misma

distribución de datos y nodos (profundidad) (Del Toro, Gomariz, Cánovas & Alonso, 2015).

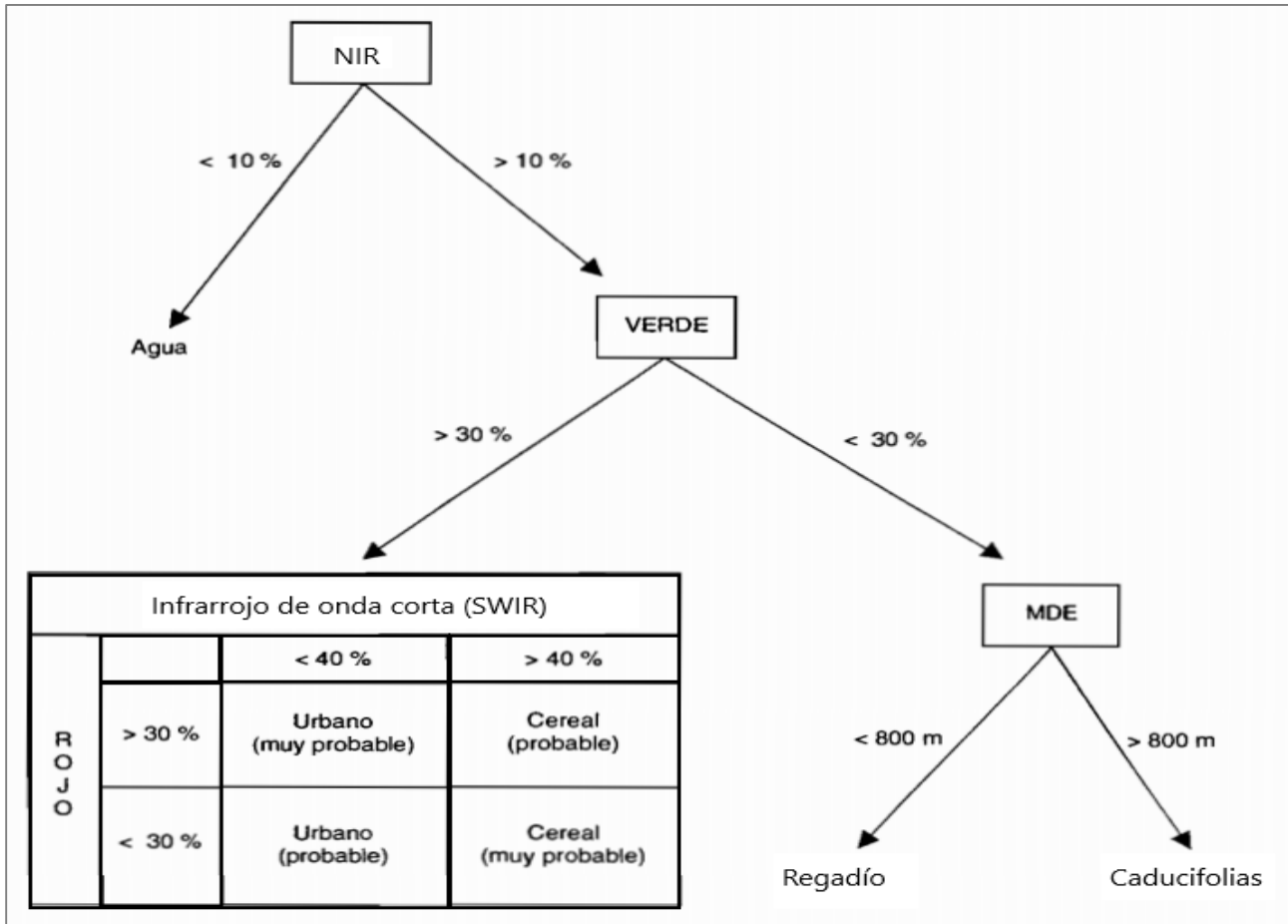


Figura 2.20. Ejemplo de Clasificación en árbol. (Chuvienco, 2010).

Los datos son tomados para generar el árbol y luego son devueltos al conjunto total, por lo que ningún dato es dejado sin tomarse en cuenta y un mismo dato puede ser utilizado por varios árboles. Cada árbol ofrece su voto para la clase más popular encontrada en el subconjunto

correspondiente, el total de los votos es promediado y con base en el resultado se asigna una clase a cada píxel de la imagen (Figura 2.21) (Breiman, 2001; Del Toro, Gomariz, Cánovas & Alonso, 2015).

El algoritmo que se utiliza para construir cada árbol de una clasificación con *Random Forest* es:

1. N es el número de nodos a construir.
2. M es el número total de datos en el clasificador.
3. m es el número determinado de datos aleatorios tomados de M para tomar una decisión en cada nodo.
4. Para cada nodo del árbol se toma una decisión con m hasta llegar a N .

Este proceso se lleva a cabo hasta cumplir con el número de árboles establecido. Al obtener el voto final de cada árbol, el clasificador promedia los resultados de cada uno para obtener el resultado final (Del Toro, Gomariz, Cánovas & Alonso, 2015).

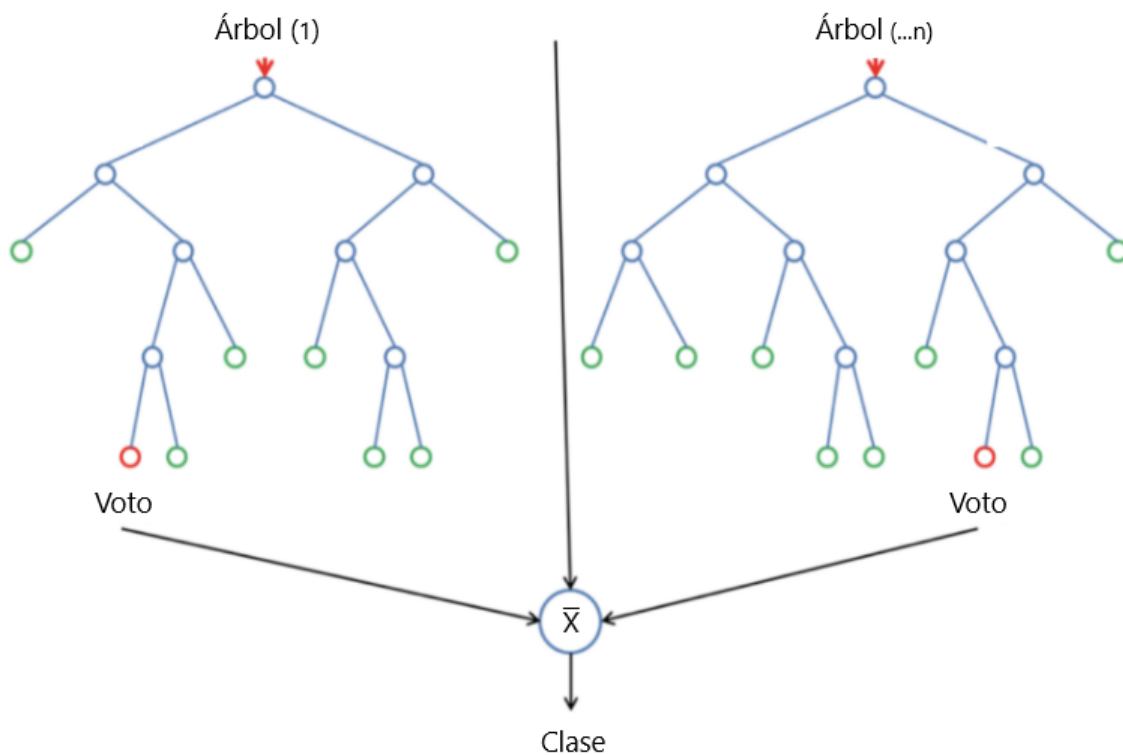


Figura 2.21. Estructura de una clasificación *Random Forest*. (Donges, 2018).

Una desventaja que puede tener este método es que caiga en el sobreajuste cuando se establece una profundidad demasiado grande en los árboles cuando la varianza es muy alta.

Es necesario contar con un número suficiente de datos, pues mientras más grande sea el árbol mejor será su resultado, aunque esto implica que el modelo tome más tiempo para terminar el proceso (Donges, 2018).

Por otro lado, las ventajas que tiene el método *Random Forest* incluyen su sencillez y flexibilidad, pues el usuario tiene el control de la profundidad de cada árbol y del número de datos que va a utilizar. También es considerado uno de los algoritmos de aprendizaje más certeros, al manejar cientos de variables sin excluir ninguna, además de tener la capacidad de estimar cuáles variables son las más importantes dentro de una clasificación (Donges, 2018).

2.5 La verificación de la clasificación

La última fase de la clasificación es la verificación, ésta es necesaria para comprobar la validez de un método de clasificación y la exactitud o precisión alcanzada por el resultado, además, contiene información sobre la magnitud de los errores de clasificación y sus patrones, esta fase ayuda a valorar qué tan concordante es la imagen clasificada con la realidad. Se requieren datos representativos de la realidad con los que sea posible hacer una comparación y estimar el error de los resultados a partir de un análisis estadístico. Los resultados de la verificación pueden ayudar a conocer si el método utilizado produce mejores resultados que otros métodos (Chuvienco, 2010; Mas, Pérez-Vega, Ghilardi, Martínez, Loya-Carrillo & Vega, 2014).

2.5.1 Validación cruzada

La validación cruzada (*Cross-Validation*) es una técnica utilizada principalmente en inteligencia artificial y predicción para la evaluación del resultado de un análisis estadístico, así como para calcular qué tan preciso es un modelo. En esta técnica se toma el total de los datos de validación (áreas de verificación) y se dividen en k número de subconjuntos y cada uno con la misma cantidad de datos. Consiste en repetir el cálculo de la media aritmética k veces y luego calcular la media de los resultados de cada iteración para tener sólo un resultado, donde cada iteración se realiza con un subconjunto diferente (Refaeilzadeh, Tang & Liu, 2009).

2.5.2 La matriz de confusión

Es una tabla o matriz cuadrada $n \times n$ también llamada matriz de error o tabla de contingencia, en la que se agrupan los errores y aciertos (acuerdos y desacuerdos) que se encontraron entre las categorías. Sus columnas indican la realidad a través de clases de referencia (áreas de verificación) y sus filas lo obtenido con el modelo de clasificación (Figura 2.22). La diagonal de la matriz señala las clases asignadas correctamente y los elementos que aparecen fuera de la diagonal (los marginales) señalan los errores de asignación (Chuvienco, 2010; Sánchez, 2016).

		Realidad	
		Positivos	Negativos
Modelo	Positivos	Verdaderos Positivos (VP)	Falsos Negativos (FN)
	Negativos	Falsos Positivos (FP)	Verdaderos Negativos (VN)

Figura 2.22. Organización de una matriz de confusión. (Zelada, 2017).

La matriz de confusión integra dos tipos de errores, el error de omisión (riesgos del usuario) y el error de comisión (riesgos del productor). Los errores de omisión (exclusión) corresponden a los elementos que pertenecen a una clase, pero no aparecen en ella y se incluyeron en otra, este error corresponde a los elementos no diagonales de las columnas. Los errores de comisión (inclusión) son los elementos que no pertenecen a una clase, pero aparecen en ella, corresponde a los elementos no diagonales de las filas (Lillesand et al., 2008; Sánchez, 2016).

Con esta matriz también es posible calcular la confiabilidad global (*overall accuracy*). Es la relación entre el total de los elementos y los que fueron asignados correctamente, permite conocer el grado global de acierto que tuvo el modelo (Lillesand et al., 2008; Chuvieco, 2010). La confiabilidad global es considerada como el porcentaje que representa la probabilidad de que un elemento haya sido clasificado correctamente,

es decir, que el clasificador le haya asignado al elemento la clase que le corresponde. Estos valores funcionan como indicadores de qué tan representativo de la realidad puede ser el resultado de un determinado método de clasificación (Alberg, Park, Hager, Brock & Diener-West, 2004).

Adicionalmente, es posible obtener la exactitud del usuario al dividir el número de elementos clasificados correctamente en cada categoría entre el número total de elementos que se clasificaron en cada categoría (el total de la fila), representa la probabilidad de que un píxel aleatorio haya sido correctamente asignado. Asimismo, se puede conocer la exactitud del productor al dividir el número de elementos clasificados correctamente en cada categoría entre el número total de la columna; esto representa la probabilidad de que un píxel aleatorio de las áreas de verificación de una clase haya sido asignado correctamente (Lillesand et al., 2008; Sánchez, 2016).

2.5.3 El índice Kappa

El índice o coeficiente Kappa es una medida estadística de la diferencia entre la concordancia del azar y lo observado, puede utilizarse para comparar la validez efectiva de diferentes métodos de clasificación (Lillesand et

al., 2008; Chuvieco, 2010). Es el valor que representa la concordancia que existe entre las muestras de dos observadores, puede tomar valores entre -1 y 1, pero normalmente se utiliza con valores desde 0, donde el grado de concordancia es mayor cuando es más cercano a 1 y menor cuando sea más cercano a 0, un valor de $k = 0$ representa una concordancia esperada a causa exclusiva del azar y un $k = 1$ representa la máxima concordancia entre la realidad y la imagen clasificada (Cerde & Villarroel, 2008).

Incorpora en su fórmula una corrección que excluye la concordancia debida al azar (López de Ullibarri & Pita, 2001). En su fórmula, este índice incluye la diferencia entre los aciertos totales observados y los totales de los elementos asignados incorrectamente en el numerador, mientras que en su denominador incluye la diferencia entre el total de observaciones y la sumatoria de los aciertos al azar (Cerde & Villarroel, 2008):

$$K = \frac{(\sum \text{aciertos totales}) - (\sum \text{aciertos al azar})}{(\text{total de observaciones}) - (\sum \text{aciertos al azar})}$$

Los resultados del cálculo del índice Kappa son valorados a partir de la escala del cuadro 2.3:

Cuadro 2.3. Valoración del índice Kappa. (Landis & Koch, 1977).

Índice Kappa	Fuerza de la concordancia
0,00	Pobre (<i>Poor</i>)
0.01 - 0,20	Leve (<i>Slight</i>)
0,21 - 0,40	Aceptable (<i>Fair</i>)
0,41 - 0,60	Moderada (<i>Moderate</i>)
0,61 - 0,80	Considerable (<i>Substantial</i>)
0,81 - 1,00	Casi perfecta (<i>Almost perfect</i>)

2.6 El Programa Landsat

En la década de los setenta, la agencia espacial norteamericana (National Aeronautics and Space Administration, NASA) puso en órbita el primer satélite del proyecto para la observación de los recursos terrestres (Earth Resources Technology Satellite) que más tarde tomaría el nombre de Programa Landsat (United States Geological Survey, 2018).

Este programa representa la colección continuamente adquirida más larga del mundo de datos terrestres de resolución moderada y es considerada una fuente de información básica en percepción remota. Sus imágenes proporcionan un recurso único para quienes trabajan en campos como la agricultura, la geología, la

silvicultura, la planificación regional, la educación, el mapeo e investigación sobre el cambio global (United States Geological Survey, 2018).

El primer satélite se denominó ERTS-1 que posteriormente actualizó su nombre a Landsat 1 y fue lanzado el 23 de julio de 1972. El siguiente satélite fue lanzado el 22 de enero de 1975, denominado Landsat 2, el 5 de marzo de 1978 se lanzó Landsat 3 como reemplazo de Landsat 1. Los tres primeros satélites llevaban un sensor llamado *Multispectral Scanner* (MSS) y tres cámaras *Return Beam Vidicon* (RBV). El satélite Landsat 4 fue lanzado el 16 de julio de 1982 y Landsat 5 el 1 de marzo de 1984, estos dos satélites contaban con un sensor multiespectral denominado *Thematic Mapper* (TM) para capturar bandas multiespectrales incluyendo una del infrarrojo térmico (United States Geological Survey, 2018).

El satélite Landsat 6 se lanzó el 5 de octubre de 1993, desgraciadamente hubo un fallo en el lanzamiento. La plataforma Landsat 7 se lanzó el

15 de abril de 1999, cuenta con un sensor llamado *Enhanced Thematic Mapper Plus* (ETM+) y Landsat 8 se lanzó el 11 de febrero de 2013, este satélite cuenta con un sensor que lleva por nombre *Operational Land Imager* (OLI) y otro llamado *Thermal Infrared Sensor* (TIRS). Para el año 2020 está previsto el lanzamiento del satélite Landsat 9. Las propiedades técnicas de cada satélite se observan en el cuadro 2.4 (Figura 2.23) (United States Geological Survey, 2018).

2.6.1 Las imágenes Landsat 8

Las imágenes del satélite Landsat 8 cubren una escena de 170 x 185 kilómetros y se conforman por once bandas espectrales recolectadas a través de sus dos sensores, nueve de ellas pertenecen al sensor OLI (*Operational Land Imager*) y dos al sensor TIRS (*Thermal Infrared Sensor*). Las características de ambos sensores se especifican en el cuadro 2.5. Las imágenes Landsat 8 están disponibles para descarga sin costo en distintas plataformas (United States Geological Survey, 2018).

Cuadro 2.4. Propiedades técnicas del Programa Landsat. (USGS, 2018).

Satélite	Lanzamiento	Fuera de servicio	Sensor	Resolución espacial (m ²)	Resolución temporal (días)	Altura (Km)
Landsat 1	23 de julio 1972	Enero 1978	MSS RBV	80 80	18	917
Landsat 2	22 de enero 1975	Febrero 1982	MSS RBV	80 80	18	917
Landsat 3	5 de marzo 1978	Septiembre 1983	MSS RBV	240 30	18	917
Landsat 4	16 de julio 1982	Diciembre 1993	MSS TM	80 30	16	705
Landsat 5	1 de marzo 1984	Enero 2013	MSS TM	80 30	16	705
Landsat 6	5 de octubre 1993	Lanzamiento fallido	ETM	30 Multiespectral 15 Pancromática	16	705
Landsat 7	15 de abril 1999	_____	ETM+	30 Multiespectral 15 Pancromática	16	705
Landsat 8	11 de febrero 2013	_____	OLI TIRS	30 Multiespectral 15 Pancromática 100 Infrarrojo térmico	16	705

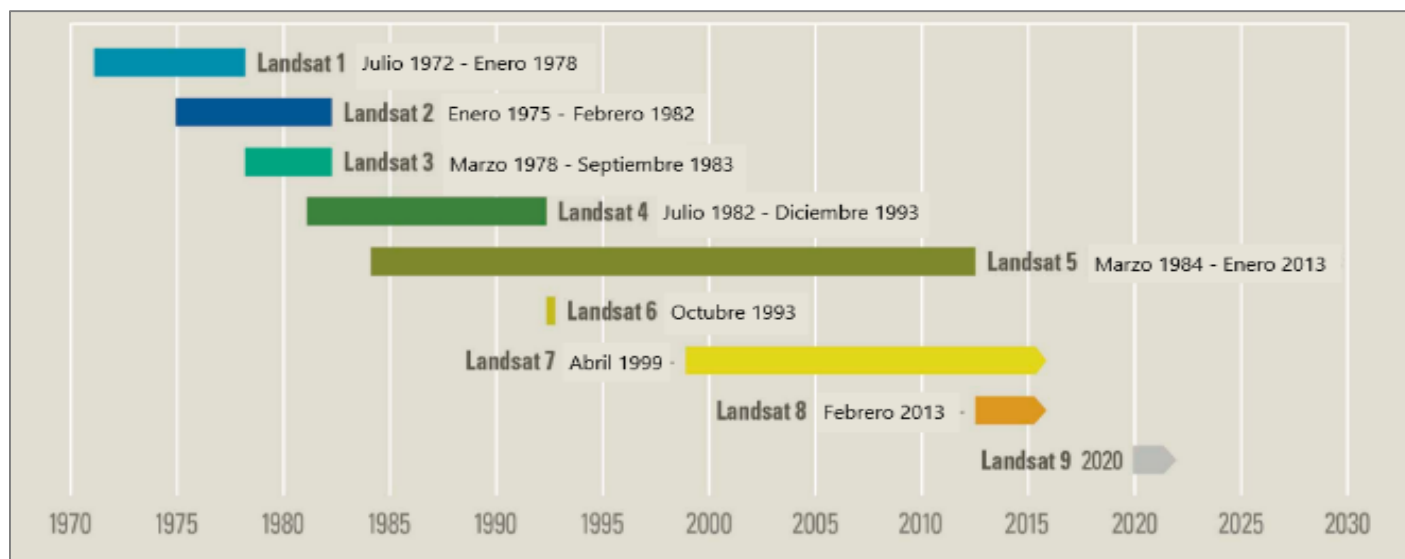


Figura 2.23. Línea del tiempo del Programa Landsat. (USGS, 2018).

Cuadro 2.5. Características de los sensores OLI y TIRS. (USGS, 2018).

Sensor	Resolución espacial (m ²)	Banda	Resolución espectral (μm)	Resolución radiométrica (Bits)
OLI	30	1 Coastal/Aerosol	0.43 - 0.45	12 (16 cuando se procesa a Nivel-1)
		2 Azul	0.45 - 0.51	
		3 Green	0.53 - 0.59	
		4 Rojo	0.64 - 0.67	
		5 Infrarrojo cercano (NIR)	0.85 - 0.88	
		6 Infrarrojo de onda corta 1 (SWIR 1)	1.57 - 1.65	
		7 Infrarrojo de onda corta 2 (SWIR 2)	2.11 - 2.29	
	15	8 Pancromática (PAN)	0.50 - 0.68	
	30	9 Cirrus	1.36 - 1.38	
TIRS	100	10 Infrarrojo térmico 1	10.6 - 11.19	
		11 Infrarrojo térmico 2	11.5 - 12.51	



III CARACTERIZACIÓN GEOGRÁFICA DEL ÁREA DE ESTUDIO

Contenido

Este capítulo contiene seis secciones, en la primera se describe la ubicación geográfica del municipio de estudio, las colindancias con otros municipios y su frontera con Guatemala. En la segunda sección se describe la geología de la zona y se muestran los principales tipos de roca que existen en el municipio, en la tercera sección, se señalan los tipos de suelo y los clasificadores primarios que ocupan un lugar en la zona de estudio, en la cuarta sección se describen los tipos de clima que persisten en Ocosingo, así como la precipitación que se presenta en el lugar para cada uno de los climas. La quinta sección contiene los ríos, lagos y arroyos que forman la red hidrológica del municipio; finalmente, en la última sección de este capítulo se muestra la flora más importante que predomina en el lugar y se ilustra el área que ocupa cada uno de los tipos de vegetación.

3.1 Localización

Ocosingo (Figura 3.1) es el municipio de mayor superficie territorial de los 124 que conforman el estado de Chiapas; se encuentra al oriente del estado y representa el 11.39% de la superficie total estatal con un área de 9520.117 km²; además en su territorio se encuentra la mayor parte de la superficie de la Selva Lacandona y el 26.01% de la superficie de la Región VI Selva (Instituto Nacional de Estadística y Geografía, 2017; Instituto Nacional para el Federalismo y el Desarrollo Municipal, 2018).

Este municipio se localiza en la región fisiográfica Montañas del Oriente, entre las coordenadas extremas 16° 54' 17" Norte y 92° 05' 40" Oriente, en su extremo norte limita con el municipio de Palenque, al suroeste con el municipio de Las Margaritas, al noroeste con los municipios de Chilón, Oxchuc, Altamirano y San Juan Cancuc, y en sus límites sur y este con la República de Guatemala (Instituto Nacional para el Federalismo y el Desarrollo Municipal, 2018). La altitud de Ocosingo varía entre los 25 y los 1966 m.s.n.m., con una media de 608 metros (Instituto Nacional de Estadística y Geografía, 2000).

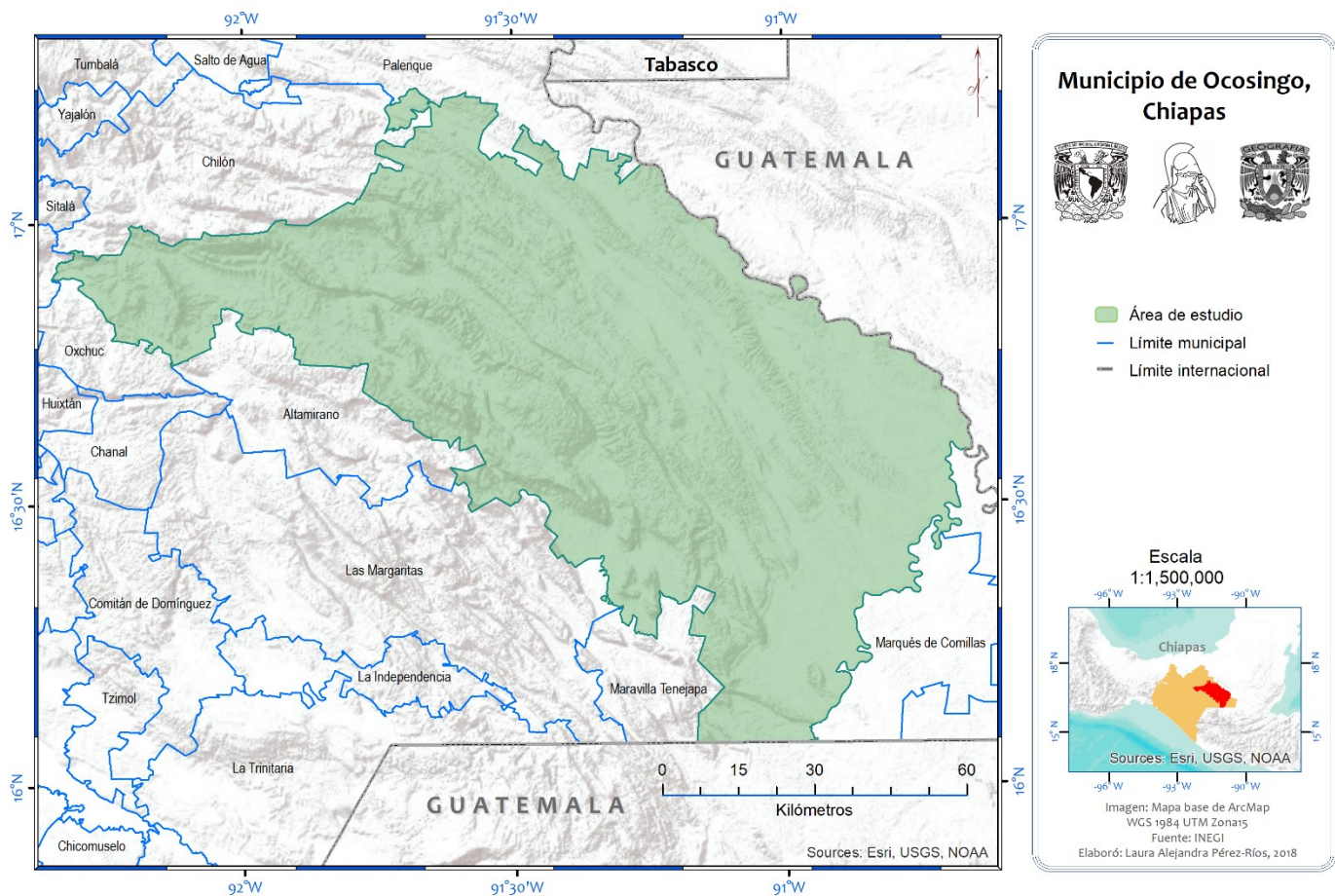


Figura 3.1 Localización de Ocosingo.

3.2 Geología

En su mayoría, la superficie de Ocosingo está formada por rocas sedimentarias de la era Cenozoica y Mesozoica, pertenecientes al periodo Cretácico, Paleógeno, y en menor medida, al Cuaternario. Las formaciones geológicas que predominan se componen por: la caliza (cz), la lutita-arenisca (lu-ar), la limolita-arenisca (lm-ar)

y las aluviales (al), entre otras como la caliza-lutita (cz-lu) y la lacustre (la) (Figura 3.2). En las tierras bajas y planas se encuentran lomeríos del noroeste y del este (Instituto Nacional de Estadística y Geografía, 1984; Universidad Tecnológica de la Selva, 2007).

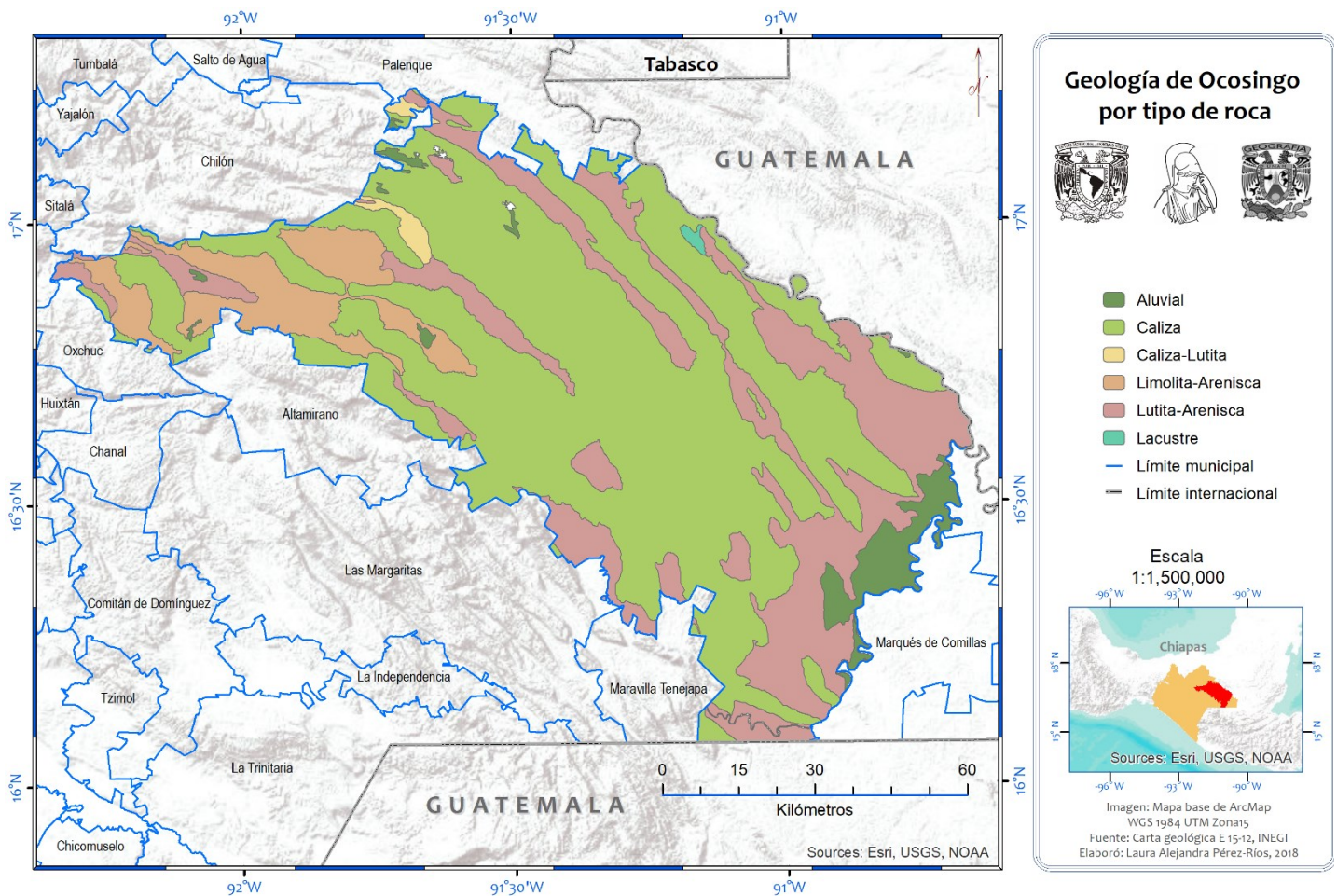


Figura 3.2 Geología de Ocosingo por tipo de roca.

3.3 Edafología

Las condiciones edafológicas de Ocosingo hacen que en el municipio se presenten diversos tipos de suelos (Figura 3.3), los que cubren la mayor cantidad de superficie son los suelos dominantes de tipo leptosol (LP), luvisol (LV) y phaeozem (PH), también hay otros tipos de suelo presentes en menor superficie como el cambisol (CM), gleysol (GL), regosol (RG), umbrisol (UM), vertisol (VR), acrisol (AC) y fluvisol (FL) (Instituto Nacional de Estadística y Geografía, 2007).

En Ocosingo, el leptosol tiene seis clasificadores primarios, de los cuales destacan el húmico, lítico, molihúmico y réndzico, el luvisol tiene once clasificadores primarios, de los cuales el vértico, profúndico, húmico y endoléptico son los que ocupan mayor superficie, el phaeozem tiene ocho clasificadores primarios, de los cuales predominan el endoléptico, epiléptico, húmico, lúvico y vértico (Instituto Nacional de Estadística y Geografía, 2007; 2017).

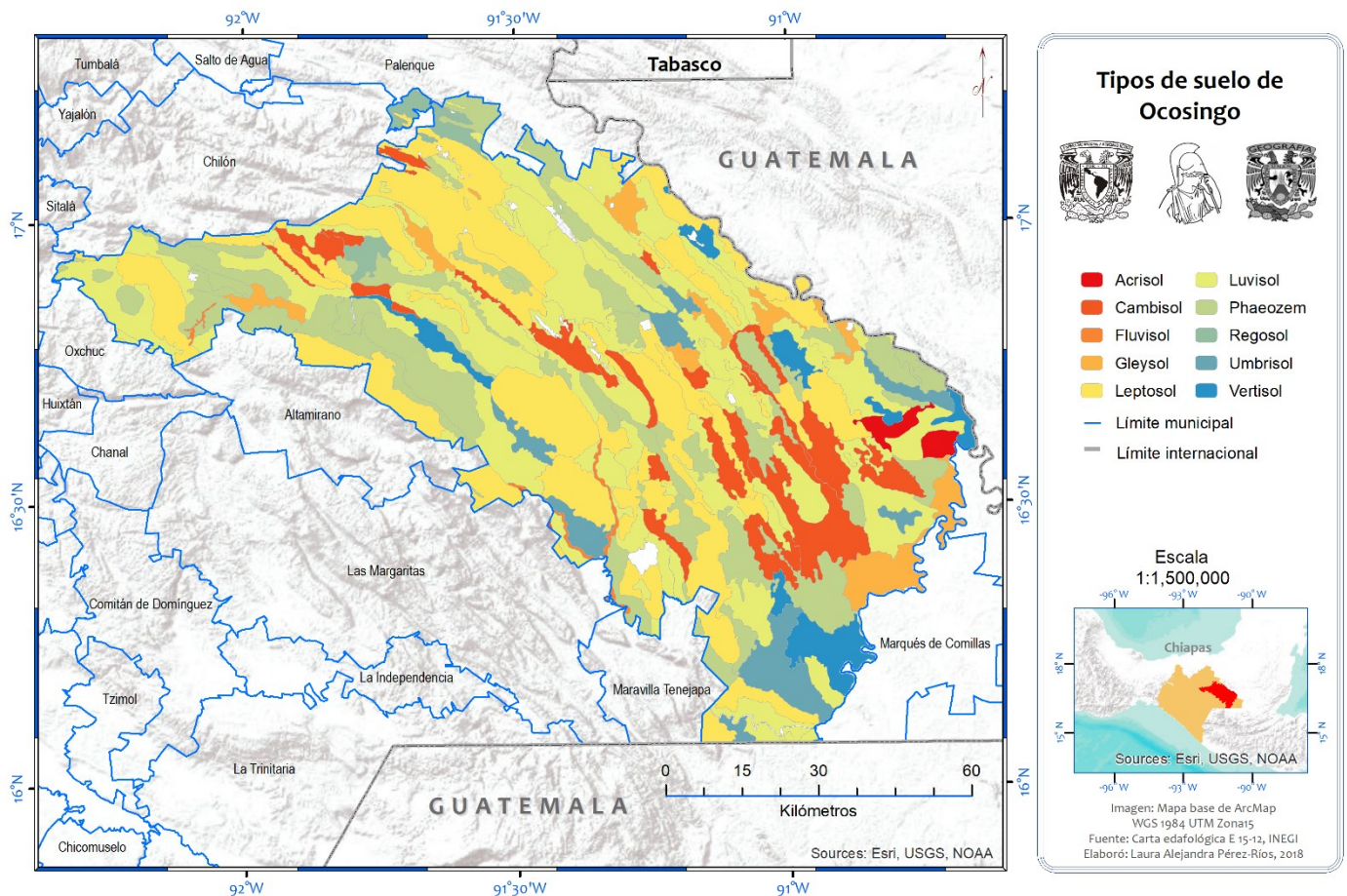


Figura 3.3 Tipos de suelo de Ocosingo.

3.4 Climatología

En Ocosingo existen cinco climas (Figura 3.4) de acuerdo con la clasificación de Köppen modificado por García (2004).

-El clima predominante es el cálido húmedo (Am), con temperatura media anual mayor de 22 °C y temperatura del mes más frío mayor de 18 °C hasta los 800 metros de altitud. La precipitación del mes más seco es menor de 60 mm, presenta lluvias en verano y tiene un porcentaje de

precipitación invernal del 5% al 10.2% del total anual (García, 2004).

-El semicálido húmedo del grupo C (A)C(m)(f) con temperatura media anual mayor de 18 °C, temperatura del mes más frío menor de 18 °C y temperatura del mes más caliente mayor de 22 °C entre los 800 y 1300 metros de altitud. Su precipitación anual es mayor de 500 mm y la precipitación del mes más seco es mayor de 40 mm, con lluvias de verano y un porcentaje de lluvia invernal mayor al 10.2% del total anual (García, 2004).

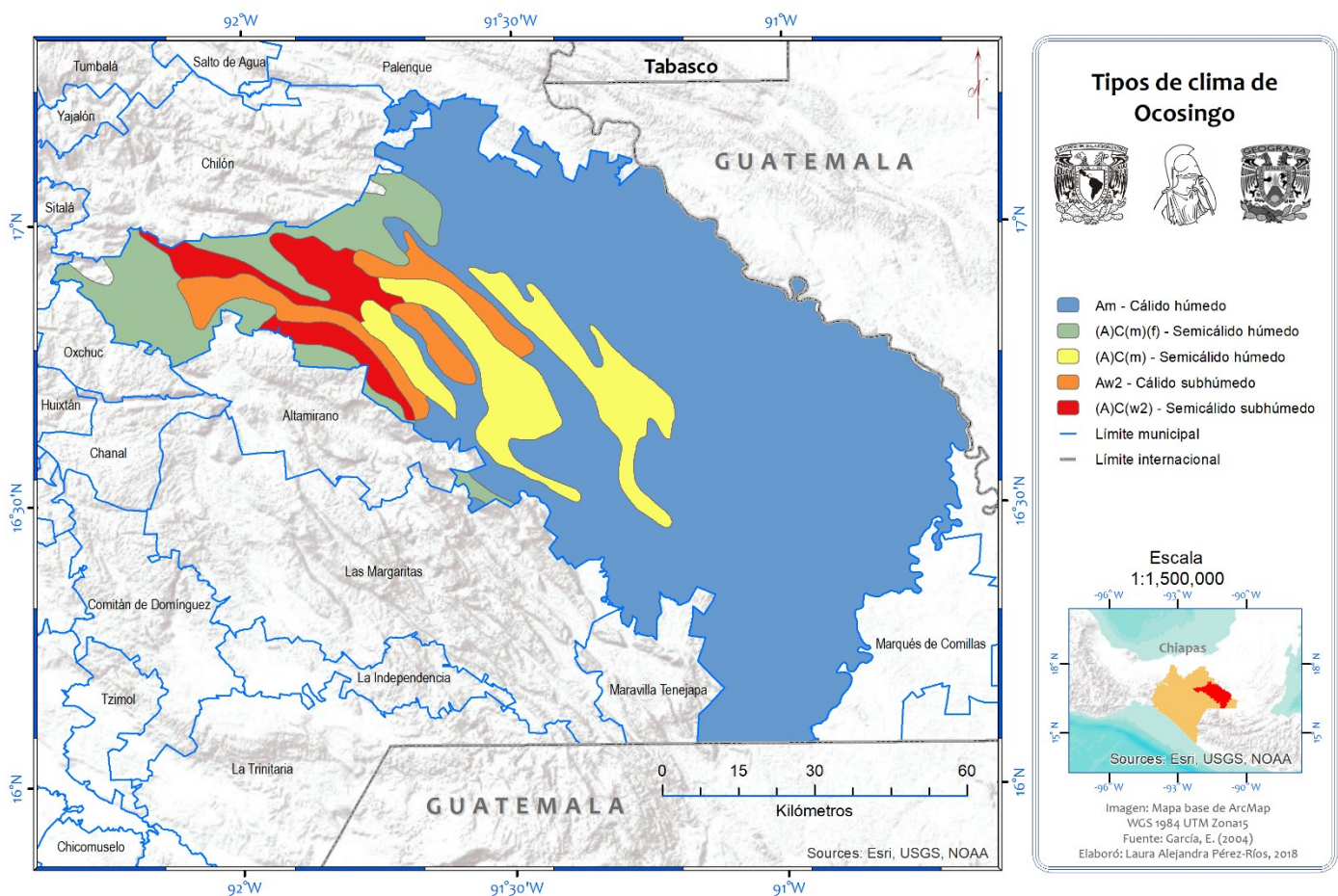


Figura 3.4 Tipos de clima de Ocosingo.

- El semicálido húmedo del grupo C (A)C(m) con temperatura media anual mayor de 18 °C, temperatura del mes más frío menor de 18 °C y temperatura del mes más caliente mayor de 22 °C entre los 800 y 1300 metros de altitud. Presenta lluvias de verano, con una precipitación del mes más seco mayor de 40 mm y un porcentaje de lluvia invernal del 5% al 10.2% del total anual (García, 2004).

-El cálido subhúmedo (Aw2), con temperatura media anual mayor de 22 °C y temperatura del mes más frío mayor de 18 °C. La precipitación del mes más seco varía entre 0 y 60 mm, presenta lluvias de verano con un índice de humedad mayor de 55.3 y un porcentaje de lluvia invernal del 5% al 10.2% del total anual (García, 2004).

-El semicálido subhúmedo del grupo C (A)C(w2), con temperatura media anual mayor de 18 °C, temperatura del mes más frío menor de 18 °C y temperatura del mes más caliente mayor de 22 °C. Su precipitación en el mes más seco es menor a 40 mm, tiene lluvias de verano con un índice de humedad mayor de 55 y un porcentaje de lluvia invernal del 5% al 10.2% del total anual (García, 2004).

Las temperaturas más altas se alcanzan desde mediados de marzo a finales de junio, los meses más lluviosos son junio, julio, agosto y septiembre

con una temporada menos húmeda (seca) de finales de diciembre a abril. La humedad relativa media anual es de 80% (Universidad Tecnológica de la Selva, 2007).

3.5 Hidrología

La red hidrológica presente en Ocosingo es grande, contiene una gran cantidad de arroyos, lagos y ríos. Los ríos principales son el Usumacinta y el Lacantún, este último representa el 53% de la cuenca del Usumacinta que al unirse con el sistema Grijalva forman la región hidrológica de mayor extensión de México. También se encuentran los ríos Jataté, Tzendales, Perlas y Lacanjá. Sus lagos son cársticos, de los cuales destacan Ocotal, El Suspiro, Miramar, Lacanjá, Orizaba, Maroma, Ojos Azules y Anaite (Universidad Tecnológica de la Selva, 2007; Comité Estatal de Información Estadística y Geográfica de Chiapas, 2018).

3.6 Vegetación

En Ocosingo se encuentran los mayores recursos forestales del estado de Chiapas, en la zona de la Selva Lacandona se cuenta con el 43.1% (por número de especies) de la flora total del estado y el 18.9% de la flora total nacional. La vegetación del municipio varía de acuerdo con la

fisiografía, el tipo de suelo y el tipo de clima (Universidad Tecnológica de la Selva, 2007).

Se extraen distintas especies de importancia cultural y económica, principalmente la *Swietenia macrophylla* (caoba), *Cedrela Odorata* (cedro rojo), la *Ceiba pentandra* (ceiba) y el *Quercus spp.* (roble), entre otros como la *Lacandonia schismatica* (Figura 3.5), *Licania platypus* (Figura 3.6), *Terminalia amazonia* y *Castilla elastica* (Universidad Tecnológica de la Selva, 2007).

De acuerdo con la clasificación de la vegetación de México de Rzedowski (2006), la vegetación predominante es la selva alta perennifolia y la selva mediana perennifolia, seguidas de otros tipos de vegetación como el bosque de pino-encino, el bosque mesófilo de montaña, el bosque ripario, los jimbales y las sabanas (Rzedowski, 2006).



Figura 3.5 *Lacandonia schismatica*.



Figura 3.6 *Licania platypus*.

Sin embargo, en la carta de Uso de Suelo y Vegetación Serie VI (Figura 3.7) del INEGI, la vegetación para el municipio de Ocosingo es distinta, aunque el tipo de vegetación predominante sigue siendo la selva alta perennifolia, de acuerdo al área que ocupan, le siguen los pastizales (naturales e inducidos), el bosque mesófilo de montaña, el bosque de pino-encino, el bosque de pino y la selva mediana subperennifolia, entre otros tipos de vegetación que se encuentran en una superficie no tan significativa como la sabana y el tular, aparte de esto, una parte considerable de la superficie está cubierta por agricultura de temporal anual (Instituto Nacional de Estadística y Geografía, 2014).

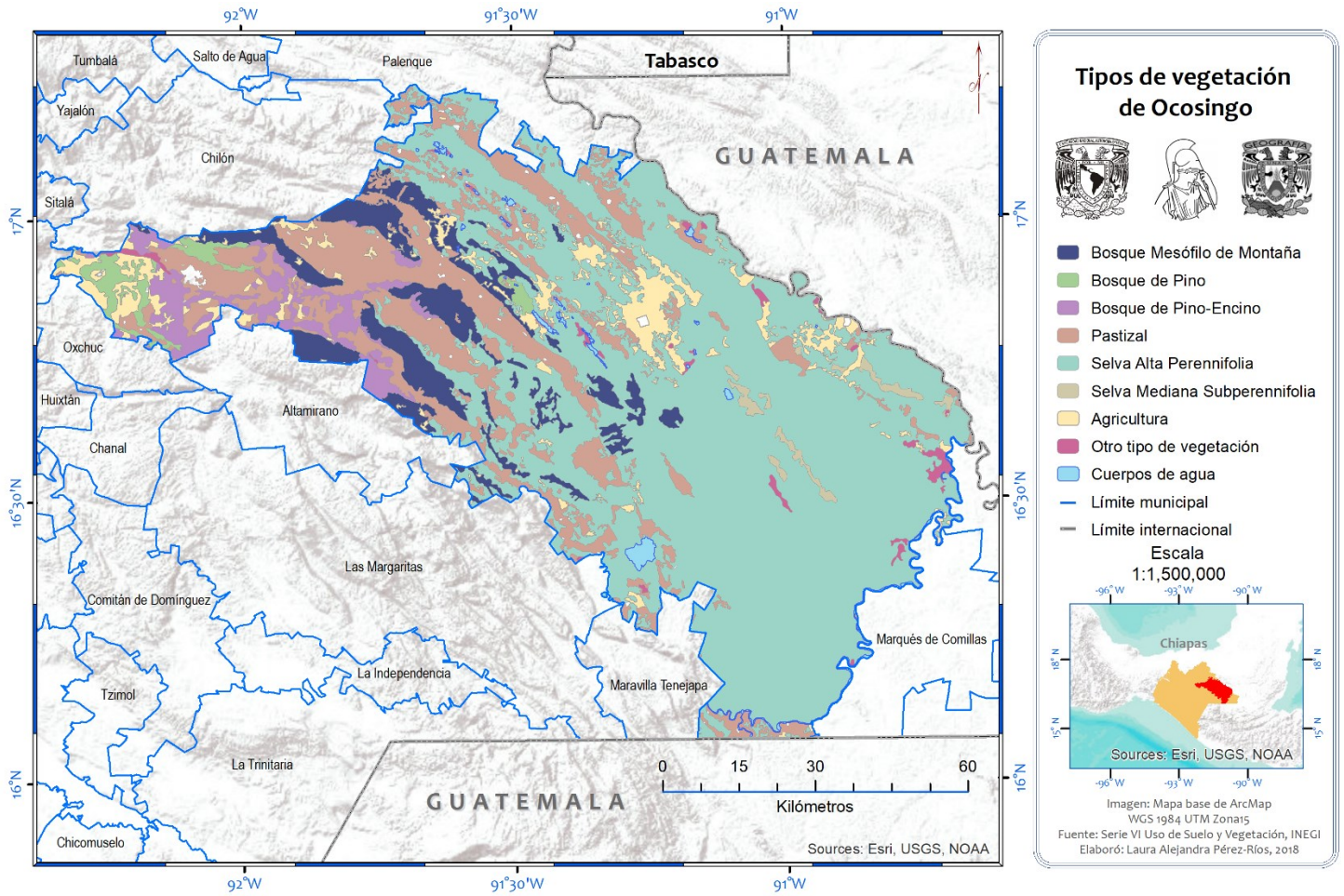


Figura 3.7 Tipos de vegetación de Ocosingo.



IV METODOLOGÍA

Contenido

En este capítulo se desarrollan los pasos seguidos para la obtención de los resultados de este trabajo de investigación. Se encuentra dividido en ocho apartados; en el primero se explica el proceso de selección y adquisición de todas las imágenes y los datos utilizados, después el preprocesamiento que se aplicó a cada imagen para la obtención de valores de reflectancia; en tercer lugar, se muestra cómo se realizó el desmezclado espectral para la obtención de la cobertura fraccional que incluye la vegetación fotosintética. Luego se describe cómo se hizo el cálculo de un índice de vegetación para la creación de una base de datos de vegetación natural, a continuación, cómo se realizó el análisis de la separabilidad espectral de cada imagen con el que se eligieron las imágenes a clasificar. En el sexto apartado se presenta la manera en la que se llevó a cabo la clasificación de las imágenes de vegetación fotosintética a través de las bases de datos de vegetación y café con el clasificador *Random Forest*, después se aborda la forma en la que se realizó la verificación de las dos clasificaciones a través del cálculo de matrices de confusión; finalmente, se calcula la superficie de café resultante en cada una de las clasificaciones (Figura 4.1).

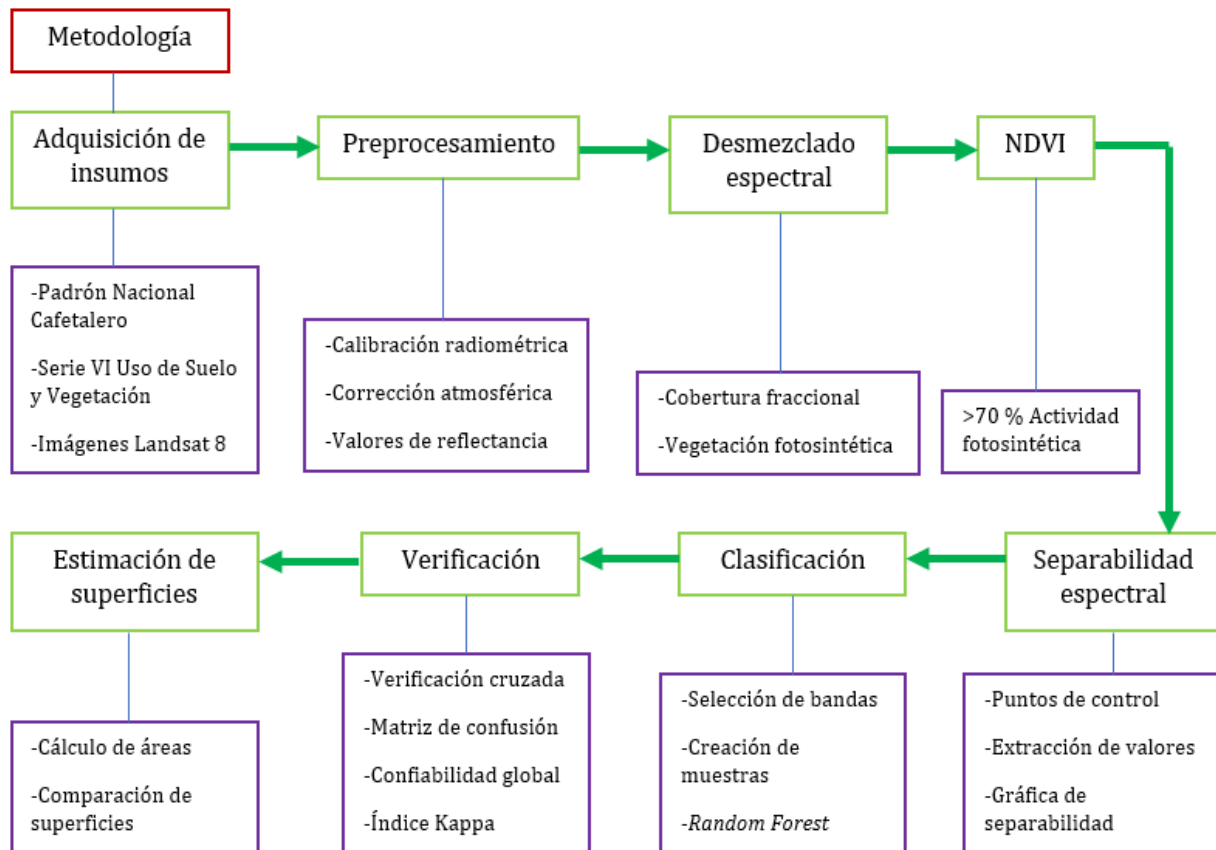


Figura 4.1. Esquema metodológico.

4.1 Adquisición de imágenes satelitales y bases de datos

Las imágenes satelitales utilizadas en este trabajo se obtuvieron a través de la plataforma de descarga EarthExplorer del Servicio Geológico de los Estados Unidos (USGS), en donde es posible la visualización y la descarga de imágenes de distintas misiones satelitales. En dicha plataforma se ingresaron los criterios de búsqueda necesarios (Cuadro 4.1) para optimizar los resultados de este trabajo, como el porcentaje de contenido de nubes y las fechas de adquisición de cada imagen.

Cuadro 4.1. Criterios de búsqueda de imágenes satelitales. (elaborado con base en EarthExplorer, 2018.).

Localización		
	Escena 1	Escena 2
Columna	21	20
Fila	48	49
Rango de fechas		
Temporada	Invierno - Primavera	
Años	2015 - 2017	
Set de datos	Landsat 8 OLI/TIRS	
Cubierta de nubes	Menos de 20%	

Fue necesario obtener la mayor cantidad de imágenes satelitales posibles de Ocosingo que cumplieran con los criterios anteriores, sin embargo, una sola escena del satélite Landsat 8

no cubre la zona de estudio en su totalidad, por ello se requirió la descarga de dos escenas de fechas similares para cubrir la zona (Figura 4.2).

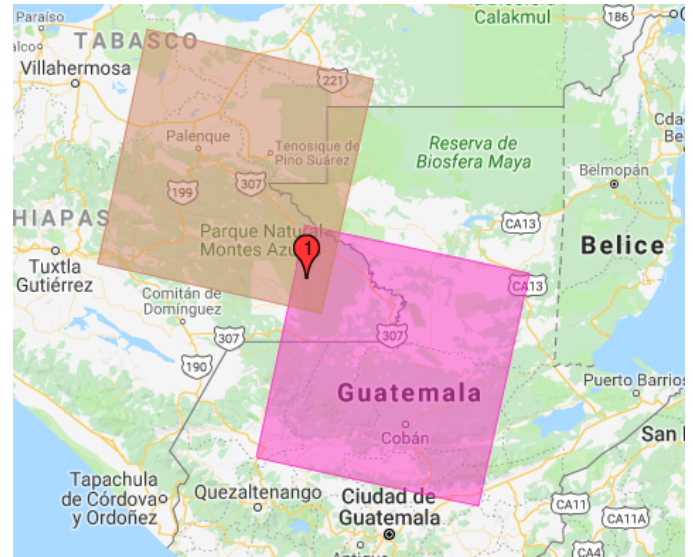


Figura 4.2. Escenas Landsat 8 que cubren la zona de estudio. (EarthExplorer, 2018).

Se utilizaron imágenes satelitales con fechas de invierno y primavera a causa del comportamiento espectral del café en la temporada seca, pues en dicha estación la planta de café no se ve igualmente mezclada con otro tipo de vegetación como en la temporada de lluvias (Moreira, Adami & Rudorff, 2004).

Se tomó en cuenta la referencia anterior como una posibilidad de que el café de sombra fuera identificado de manera más eficiente por el clasificador con imágenes de la temporada seca, además, la multitemporalidad utilizada en este trabajo hace posible el reconocimiento de los

valores fotosintéticos que tiene el café durante fechas similares en años distintos, para facilitar la identificación de su cultivo en otras zonas del área de estudio a través del clasificador.

A partir de lo anterior, se consiguieron dieciséis imágenes del satélite Landsat 8 con fechas de adquisición cercanas entre sí, pertenecientes a la temporada seca de los años 2015, 2016 y 2017 que cubren el área de interés (Cuadro 4.2).

Cuadro 4.2. Fechas de las imágenes satelitales adquiridas. (elaborado con base en EarthExplorer, 2018).

	Escena 1	Escena 2
Fechas	16 - 02 - 2015	25 - 02 - 2015
	03 - 02 - 2016	27 - 01 - 2016
	09 - 05 - 2016	18 - 05 - 2016
	03 - 12 - 2016	12 - 12 - 2016
	20 - 01 - 2017	29 - 01 - 2017
	25 - 03 - 2017	18 - 03 - 2017
	10 - 04 - 2017	03 - 04 - 2017
	26 - 04 - 2017	19 - 04 - 2017

Se usó la base de datos del Padrón Nacional Cafetalero (PNC) 2014 proporcionado por la Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación (SAGARPA) y por el Servicio de Información Agroalimentaria y Pesquera (SIAP). Esta base de datos contiene todas las parcelas en formato vectorial (polígonos) de los productores de café registrados ante SAGARPA.

Asimismo, se utilizó la información contenida en la Serie VI 2014 de Uso de Suelo y Vegetación (escala 1:250 000) proporcionada por el Instituto Nacional de Estadística y Geografía (INEGI) a través de su plataforma para la descarga de información. En esta base de datos se encuentra organizada la superficie nacional en las clases que se muestran en el cuadro 4.3, de éstas solamente se aprovecharon las categorías de vegetación natural.

Cuadro 4.3. Uso de Suelo y Vegetación de la Serie VI 2014. (INEGI, 2014).

Nombre de la clase	Usada
Acuícola	
Agricultura de riego	
Agricultura de temporal	
Asentamiento humano	
Bosque de coníferas primario	✓
Bosque de coníferas secundario	✓
Bosque de coníferas-latifoliadas primario	✓
Bosque de coníferas-latifoliadas secundario	✓
Bosque de latifoliadas primario	
Bosque mesófilo de montaña primario	✓
Bosque mesófilo de montaña secundario	✓
Cuerpo de agua	
Matorral xerófilo primario	
Matorral xerófilo secundario	
Mezquital primario	
Mezquital secundario	
Otro tipo de vegetación	
Pastizal inducido o cultivado	

Pastizal natural primario	
Pastizal natural secundario	
Plantación forestal	
Selva caducifolia y subcaducifolia primaria	
Selva caducifolia y subcaducifolia secundaria	
Selva perennifolia y subperennifolia primaria	✓
Selva perennifolia y subperennifolia secundaria	✓
Vegetación hidrófila primaria	
Vegetación hidrófila secundaria	
Área sin vegetación aparente	

4.2 Preprocesamiento de imágenes satelitales

Las imágenes Landsat 8 ya cuentan con georreferenciación y calibración geométrica, por lo que se procedió a ordenar las bandas de cada imagen de menor a mayor. Al contar con lo anterior, se realizó la calibración radiométrica y la corrección atmosférica de todas las imágenes; cabe destacar que durante este proceso las nubes, los cuerpos de agua y las sombras fueron enmascarados, así como cualquier otro tipo de artefacto en la atmósfera que pudiera provocar alguna perturbación en el cálculo de la reflectancia de cada imagen (Figura 4.3), a través del método de Carlotto (1999), el cual toma el valor promedio de las bandas visibles correspondientes a cada combinación de bandas de infrarrojo para repartir más homogéneamente los artefactos y después eliminarlos.

Al tener todas las imágenes ya calibradas y corregidas se llevó a cabo el cálculo de los valores de reflectancia superficial aparente de cada una de las dieciséis escenas para obtener la fracción de la energía radiante reflejada en la superficie.

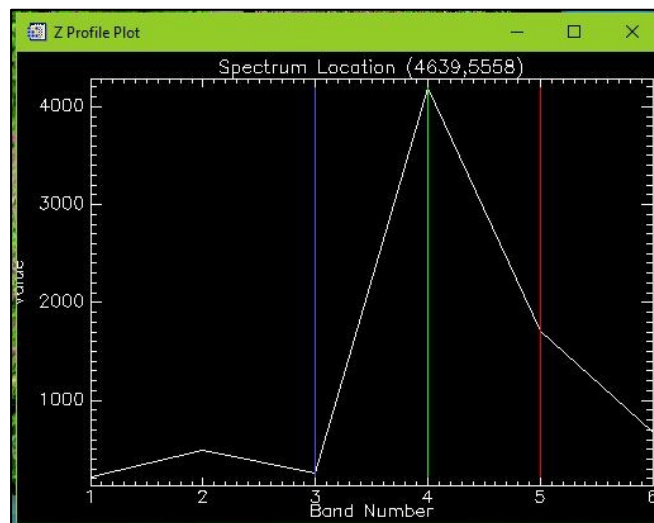
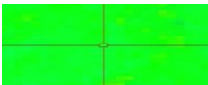
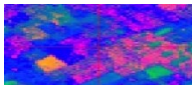
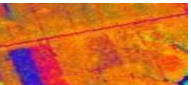


Figura 4.3. Perfil del valor de reflectancia de un píxel. (obtenido a partir de procesamiento en sistemas de información geográfica).

4.3 Desmezclado espectral

Se utilizaron las imágenes con valores de reflectancia que se obtuvieron anteriormente y, a partir de éstas, se calcularon los valores de cobertura fraccional de cada imagen. Fue en este paso cuando se aplicó el algoritmo del AutoMCU, que píxel por píxel clasificó cada imagen en alguna de las tres posibles clases de cobertura que se muestran en el cuadro 4.4.

Cuadro 4.4. Clases de cobertura fraccional. (obtenido a partir de procesamiento en sistemas de información geográfica).

Cobertura fraccional		
Vegetación fotosintética (PV)	Vegetación no fotosintética (NPV)	Sustrato expuesto (S)
		

Además de las tres bandas anteriores, cada una de las dieciséis imágenes que resultaron de este cálculo están compuestas por otras cuatro bandas (Cuadro 4.5). Dada la finalidad de este trabajo, solamente se utilizó la banda que contiene la cobertura fraccional de vegetación fotosintética (PV), misma que se extrajo de cada una de las imágenes.

Cuadro 4.5. Bandas que componen las imágenes de cobertura fraccional. (obtenido a partir de procesamiento en sistemas de información geográfica).

Banda	Nombre
1	Superficie descubierta
2	Vegetación fotosintética
3	Vegetación no fotosintética
4	Error de la superficie descubierta
5	Error de la vegetación fotosintética
6	Error de la vegetación no fotosintética
7	Error total RMSE

Como se mencionó antes, se necesitaron dos escenas de imágenes para cubrir la zona de estudio, que se muestra en color rojo en la Figura 4.4. Por esto, fue necesario unir los pares de imágenes para tener completa el área de interés. Este proceso se realizó a través de la creación de ocho mosaicos, uno por cada par de imágenes. Después, se delimitaron las imágenes a la zona de estudio, por esta razón cada mosaico se recortó con base en el polígono del municipio de Ocosingo (Figura 4.5). De igual modo, se recortaron las capas de información del Padrón Nacional Cafetalero 2014 (Figura 4.6) y de la Serie VI 2014 de Uso de Suelo y Vegetación del INEGI (Figura 4.7).

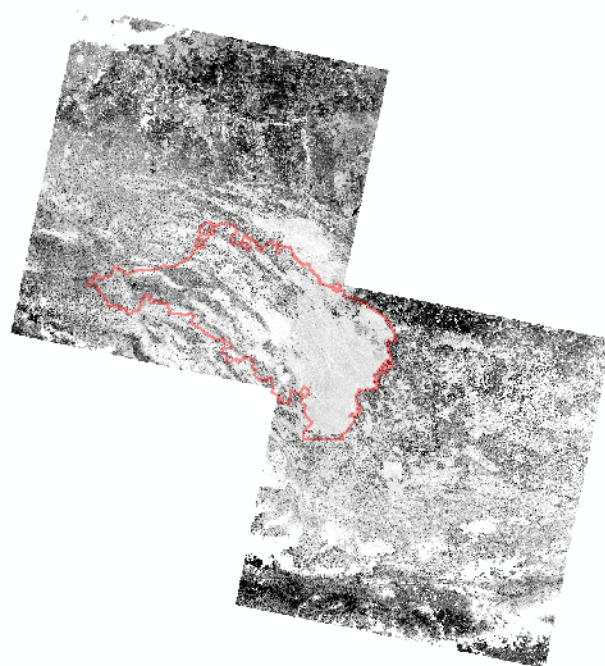


Figura 4.4. Imágenes de Vegetación fotosintética que cubren la zona de estudio. (obtenido a partir de procesamiento en sistemas de información geográfica).



Figura 4.5. Recorte de imágenes con base en la delimitación de Ocosingo. (Obtenido a partir de procesamiento en sistemas de información geográfica).

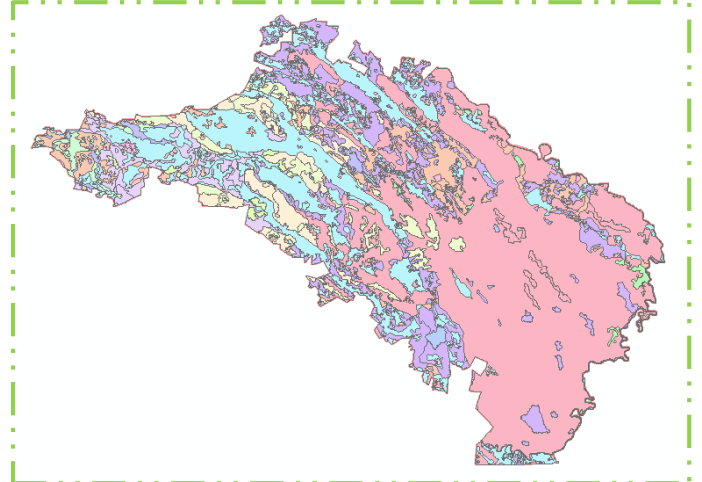


Figura 4.7. Recorte de Serie VI con base en la delimitación de Ocosingo. (Obtenido a partir de procesamiento en sistemas de información geográfica).

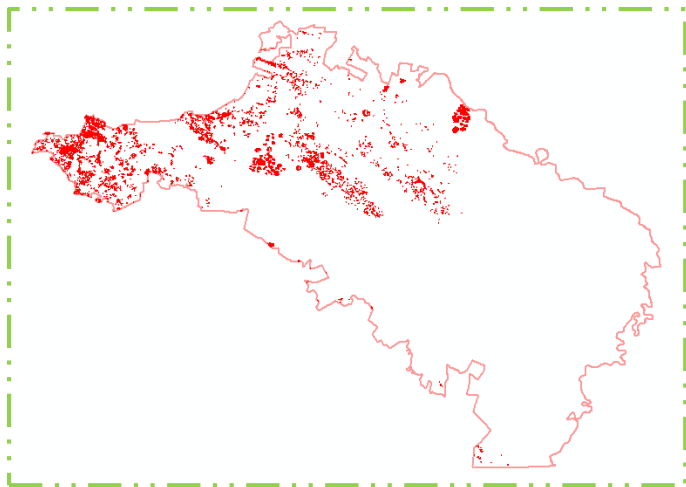


Figura 4.6. Recorte de Padrón Nacional Cafetalero con base en la delimitación de Ocosingo. (Obtenido a partir de procesamiento en sistemas de información geográfica).

Lo anterior se realizó para minimizar los tiempos de procesos posteriores, pues las imágenes son grandes y no es necesario contar con la totalidad del área que cubren, adicionalmente, con la información acotada a la zona de estudio se redujo también la cifra de parcelas de café a 19,000 polígonos.

Con las ocho imágenes de cobertura fraccional de Ocosingo en valores de vegetación fotosintética (PV), se unieron todas en una sola mediante un apilamiento de capas, en donde cada imagen se convirtió en una banda apilada en orden cronológico (Figura 4.8).



Figura 4.8. Apilamiento de capas de Vegetación fotosintética. (obtenido a partir de procesamiento en sistemas de información geográfica).

4.4 Cálculo del índice de vegetación

Se ejecutó la corrección radiométrica y la corrección atmosférica a ambas imágenes Landsat 8 del año 2015, se eligieron las imágenes de este año por su cercanía temporal con el Padrón Nacional Cafetalero 2014.

Posteriormente, se calculó el Índice de Vegetación de Diferencia Normalizada o NDVI (Rouse, Haas, Scheel & Deering, 1974) de ambas escenas a través del cálculo de sus bandas. Los resultados se unieron en una sola imagen por medio de un mosaico, y éste a su vez se recortó a la zona de estudio con el polígono de Ocosingo.

Se reclasificó el NDVI para utilizar únicamente la vegetación con más de 70% de actividad fotosintética.

4.5 Análisis de separabilidad espectral

Sobre la imagen del índice de vegetación (NDVI) previamente creada y ajustada, se crearon 19,000 puntos de control (polígonos) de treinta metros de diámetro sobre las áreas seleccionadas (Figura 4.9) con dos propósitos: tener una base de datos únicamente de vegetación natural con base en el NDVI, y que el tamaño de los polígonos coincidiera con el tamaño de los píxeles de las imágenes Landsat 8 (Cuadro 4.6) para mejorar la asignación de las clases.

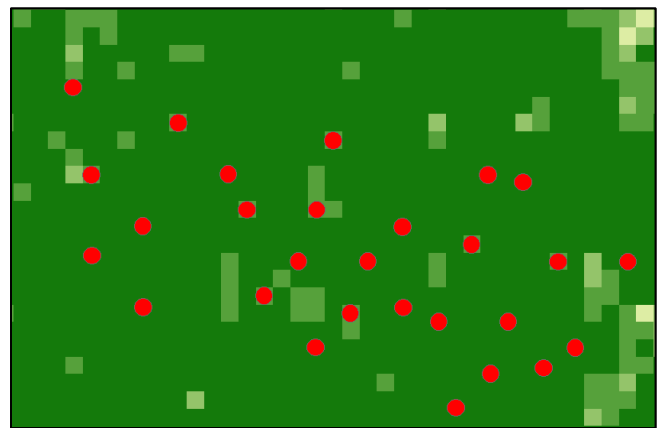
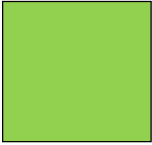

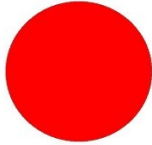


Figura 4.9. Creación de puntos de control de vegetación natural sobre NDVI. (obtenido a partir de procesamiento en sistemas de información geográfica).

Cuadro 4.6. Comparación entre resolución de la imagen y puntos de control. (obtenido a partir de procesamiento en sistemas de información geográfica).

Tamaños		
Resolución de un píxel Landsat 8	Punto de control de Vegetación natural	Punto de control de Café de sombra
30 metros	30 metros	30 metros
		

Cabe destacar que estos puntos de control pertenecen a áreas en las que no existe una sobreposición con los polígonos de café que contiene el Padrón Nacional Cafetalero y se procuró tener la misma cantidad de puntos de control (19,000) en las bases de datos.

De la Serie VI del INEGI se utilizaron únicamente los polígonos de vegetación natural y debido a que estos son de dimensiones grandes, se obtuvo el centroide de cada uno de ellos y se creó una base de datos nueva con polígonos de treinta metros de diámetro por la razón mencionada antes. La cantidad de puntos de control es importante para la clasificación, pues se debe tener en cada clase una cantidad semejante para que el clasificador realice la asignación de clases de manera equilibrada.

Además, se crearon las bases de datos de vegetación natural con la finalidad de homogeneizar con el padrón de café, porque las bases de datos deben ser del mismo tipo (polígonos) como requisito para que el clasificador funcione correctamente. Luego se extrajeron los valores estadísticos (promedio) de los píxeles de cada una de las ocho bandas (imágenes) de vegetación fotosintética a través del cálculo de las estadísticas zonales de la imagen con las ocho bandas apiladas.

Se tomaron como base los polígonos de ambas bases de datos de vegetación natural, así como los polígonos de café. Posteriormente se extrajo la información con extensión .dbf y en formato de tabla. Se abrieron tres hojas de cálculo, una para cada base de datos, se ingresaron las ocho tablas correspondientes a cada clase y se calculó el promedio de los valores estadísticos obtenidos.

A continuación, se graficaron cronológicamente los resultados de las tres bases de datos y con ello se analizó la separabilidad espectral que existe entre la vegetación natural de la base de datos del NDVI y el café, y la separabilidad espectral entre la vegetación natural de la Serie VI y el café. Con estas gráficas se identificaron las tres imágenes en las que existe mayor separabilidad entre clases para ambos pares de muestras y se seleccionaron.

4.6 Clasificación de imágenes satelitales con el método *Random Forest*

Se abrieron las tres imágenes de vegetación fotosintética seleccionadas anteriormente y se apilaron en orden cronológico en una sola imagen de tres bandas (Figura 4.10) para que sobre ésta se realizara la asignación de clases por el clasificador.



Figura 4.10. Apilamiento de las bandas con mayor separabilidad. (obtenido a partir de procesamiento en sistemas de información geográfica).

Se dividió arbitrariamente la base de datos de cada clase en dos partes iguales (vegetación natural NDVI, vegetación natural Serie VI y café de sombra), una de las partes de cada clase se destinó al entrenamiento del clasificador y el resto se reservó para la posterior verificación de las clasificaciones (Figura 4.11).

■ Vegetación natural con NDVI
 ■ Vegetación natural con Serie VI
 ■ Café de sombra

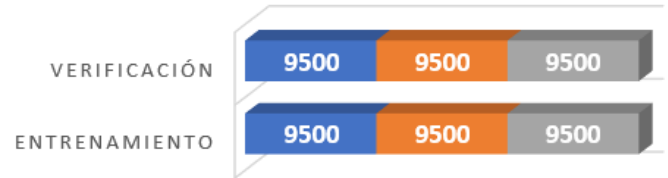


Figura 4.11. Proporción de puntos de control para entrenamiento y verificación de las clasificaciones.

En un administrador de muestras de entrenamiento se ingresaron todos los polígonos de entrenamiento de las dos combinaciones de clases y se obtuvieron las dos muestras con las que se realizaron las posteriores clasificaciones. Se creó una muestra con los polígonos de café y vegetación natural creada con la Serie VI, y otra con el café y la vegetación natural creada con el NDVI.

Para cada una de las clasificaciones, en las opciones de configuración del clasificador *Random Forest* se utilizó una de las muestras creadas, la imagen de tres bandas (imágenes) y se ajustó el número de árboles de predicción a cincuenta. Al término de ese proceso se generó como resultado un archivo en formato JavaScript Object Notation (JSON) con extensión de definición de clasificador de Esri (.ecd) para cada una de las clasificaciones. Los nuevos archivos .ecd contenían la información de atributos,

estadísticas y otra información que se necesitaba para que el clasificador de imágenes realizara la asignación de las clases sobre los píxeles de la imagen de tres bandas (imágenes) creada antes.

Para finalizar, se clasificó la imagen con las tres bandas (imágenes) de mayor separabilidad espectral y los archivos con extensión .ecd que calculó el clasificador *Random Forest*. En este último paso el clasificador asignó una de las dos clases posibles (vegetación natural o café de sombra) a cada píxel de la imagen de acuerdo a la información que cada archivo .ecd contenía, con lo que se obtuvieron las dos clasificaciones finales.

4.7 Verificación

Para cada una de las imágenes clasificadas se elaboró una validación cruzada. Se comenzó con la partición de los datos de las clases designados para verificación en veinte partes iguales y aleatorias. Con cada una de las partes de puntos, se verificaron las dos imágenes clasificadas correspondientes, a través del cálculo de matrices de confusión y se revisó si la clase asignada por el clasificador y la clase a la que en realidad pertenece cada píxel tuvieran coincidencia o fueran confundidas. Al obtener lo anterior, para cada clasificación se obtuvo el promedio de los resultados de las matrices,

únicamente de confiabilidad global e índice Kappa. Con estos datos se realizó el análisis de resultados para evaluar el método utilizado y para comparar los resultados obtenidos con los de otros métodos de clasificación utilizados para el café.

4.8 Estimación de superficie de café en las clasificaciones

Para conocer la superficie de café de sombra que fue asignada en cada una de las clasificaciones, se calculó el área de dicha clase por medio del conteo de píxeles pertenecientes al café, luego este número fue multiplicado dos veces por treinta debido a que la resolución de cada píxel es de treinta metros. Con esta multiplicación se obtuvo la superficie de la clase de café en metros cuadrados, y para realizar la comparación se convirtieron los valores de metros a hectáreas.

Finalmente, se hizo una gráfica en la que se muestra la comparación entre la superficie de café obtenida en ambas clasificaciones y la superficie sembrada de Ocosingo reportada por el Servicio de Información Agroalimentaria y Pesquera de los años 2015, 2016 y 2017.



V RESULTADOS

Contenido

En este capítulo se presentan los resultados que se obtuvieron siguiendo la metodología anteriormente descrita. Se divide en trece apartados, en el primero se muestran los polígonos que forman parte del Padrón Nacional Cafetalero 2014 y que pertenecen al área de estudio, el segundo muestra los polígonos de vegetación natural del municipio, obtenidos de la Serie VI de Uso de Suelo y Vegetación; el tercero muestra las imágenes del año 2015 antes y después de la calibración radiométrica y la corrección atmosférica. El cuarto apartado contiene las dieciséis imágenes de reflectancia superficial, el quinto apartado, las dieciséis imágenes de cobertura fraccional que resultaron del desmezclado espectral; en el sexto se muestran las imágenes de vegetación fotosintética delimitadas a la zona de estudio. En el séptimo apartado se presenta el índice de vegetación calculado para el año 2015 y su reclasificación para la zona de estudio, el octavo apartado muestra los puntos de control de cada clase utilizada para la clasificación; el noveno contiene las gráficas en las que se observa la separabilidad espectral entre el café y la vegetación. En el décimo se muestra la imagen clasificada con las parcelas de café y los polígonos de vegetación obtenidos a partir del índice de vegetación, en el decimoprimer se muestra la imagen clasificada con la vegetación seleccionada de la Serie VI, en el decimosegundo se presentan las matrices de confusión con los datos obtenidos de la verificación de las clasificaciones y sus valores de confiabilidad global e índice Kappa; finalmente, en el decimotercero se muestra el área de café que resultó en las clasificaciones.

5.1 Parcelas de café de Ocosingo

Se utilizaron los 19,000 polígonos (parcelas) de café que se encuentran dentro del territorio de Ocosingo. En la figura 5.1 se muestra la distribución espacial de los polígonos del Padrón Nacional Cafetalero (PNC) 2014, se observa que hay una concentración de parcelas al

centro y noroeste del municipio, mientras que en el sur no se observa una cantidad significativa de polígonos. Las parcelas son polígonos que varían en forma y tamaño, estos tienen un área mínima de 0.52 m², una máxima de 99,651.2 m² y un promedio de 3,687.3 m². En total, la superficie cubierta por café de acuerdo al PNC es de 7,360.3 hectáreas.

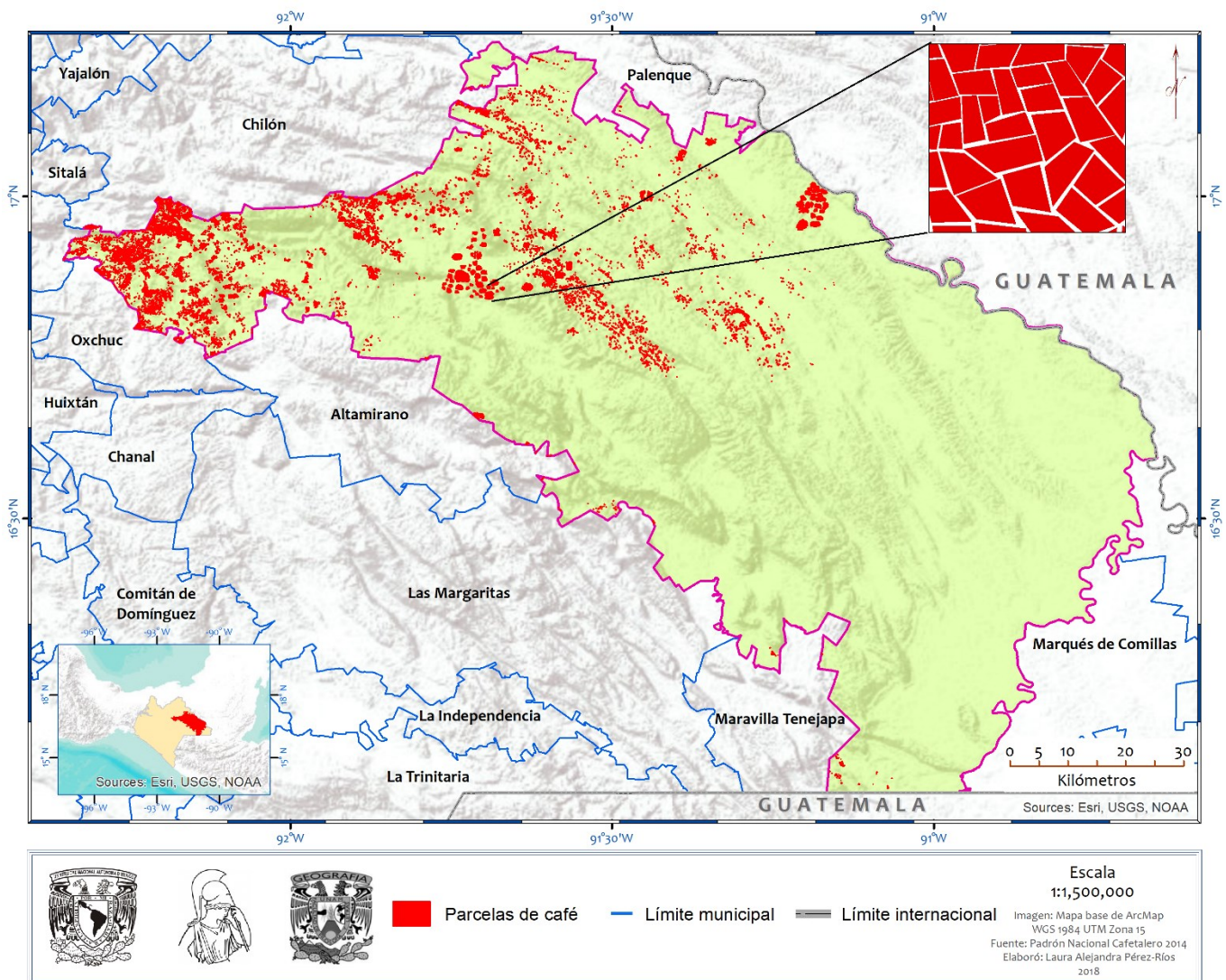


Figura 5.1. Parcelas de café de Ocosingo.

5.2 Vegetación natural de Ocosingo

La vegetación natural seleccionada de la Serie VI de Uso de Suelo y Vegetación se observa en la figura 5.2, la superficie cubierta por esta vegetación es de 694,740.2 hectáreas, donde la selva alta perennifolia ocupa la mayor área con 558,563.3 hectáreas, le sigue el bosque mesófilo

de montaña con 75,557.6 hectáreas, luego está el bosque de pino-encino con 37,976.6 hectáreas, después el bosque de pino con 14,735.1 hectáreas y finalmente, la selva mediana subperennifolia con 7,907.5 hectáreas. Las zonas que se observan vacías dentro del municipio pertenecen a pastizales, agricultura y a otro tipo de vegetación de menor superficie.

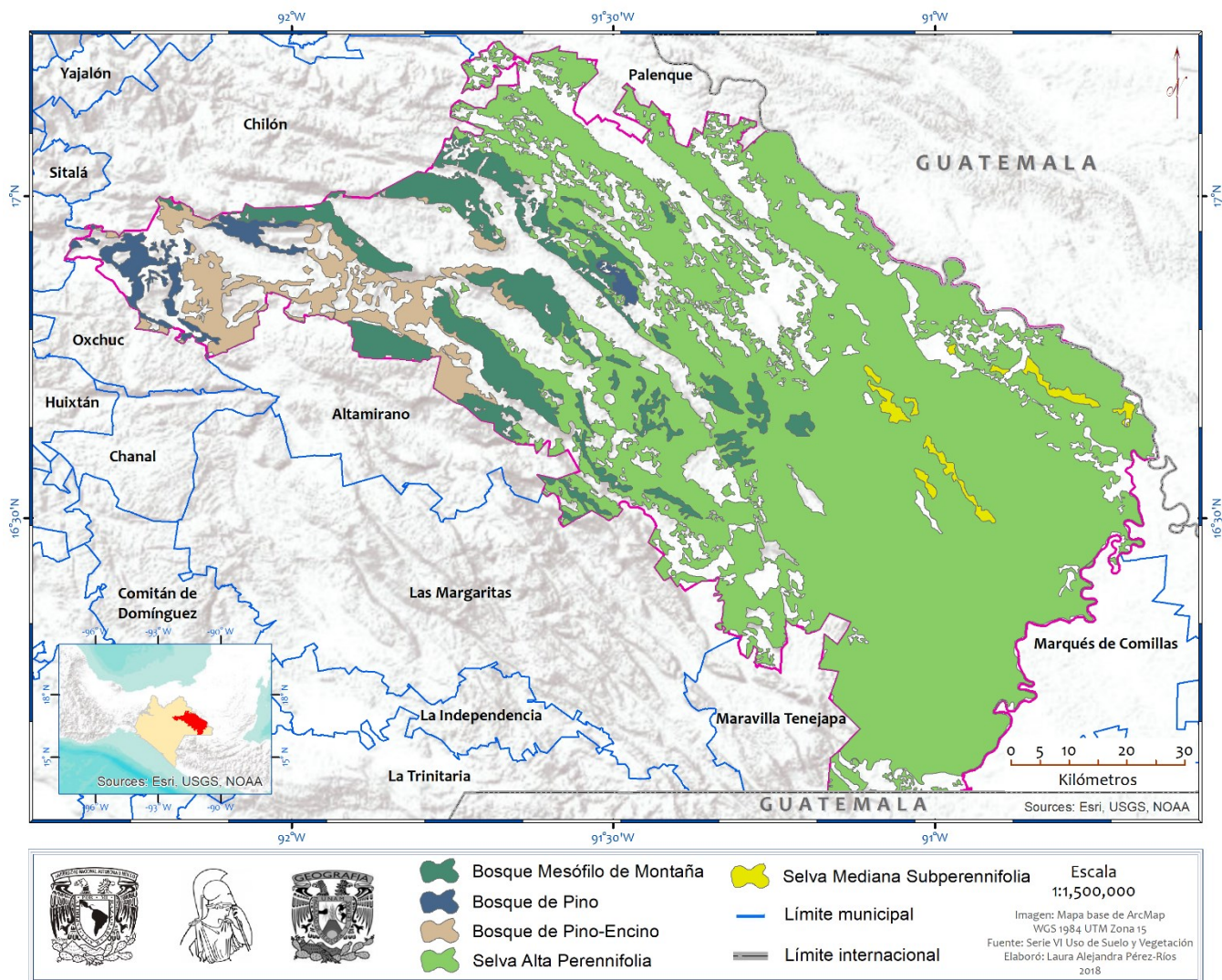


Figura 5.2. Vegetación natural de Ocosingo.

5.3 Calibración radiométrica y corrección atmosférica

Sobre las dos imágenes del año 2015 se hizo la calibración radiométrica y el cálculo de la corrección atmosférica (Figura 5.3), en donde se enmascararon las nubes y las sombras de cada escena (en blanco). La información de las imágenes crudas se convirtió en valores de radiancia y después en valores de reflectancia normalizada de 0 a 10,000 dejando fuera los valores negativos de los píxeles.

5.4 Imágenes de reflectancia

Se obtuvieron dieciséis imágenes con valores de reflectancia, ocho correspondientes a cada escena, con nubes, sombras y cuerpos de agua enmascarados (en blanco). Cada imagen se muestra en las figuras 5.4 (escena 2049) y 5.5 (escena 2148), con las fechas de captura.

5.5 Imágenes de cobertura fraccional

Del cálculo de desmezclado espectral que se hizo sobre las imágenes de reflectancia se obtuvieron dieciséis imágenes de cobertura fraccional que se muestran en las figuras 5.6 (escena 2049) y 5.7 (escena 2148). Las áreas en color verde pertenecen a la vegetación

fotosintética (PV), las azules a vegetación no fotosintética (NPV) y las rojas al sustrato expuesto o superficie descubierta (S). También se muestra la fecha de cada imagen, la parte de Ocosingo que cubre cada escena y en blanco los píxeles enmascarados.

5.6 Imágenes de vegetación fotosintética

De la banda de vegetación fotosintética (PV) extraída de las imágenes de cobertura fraccional, se recortó la zona de cada escena que cubre a Ocosingo y se unieron los pares en una sola imagen, en la figura 5.8 se muestran las ocho imágenes de vegetación fotosintética (PV) formadas por las dieciséis imágenes iniciales, a cada imagen se le asignó un número del 1 al 8 en orden cronológico. Los valores se encuentran en porcentaje; los píxeles en un color verde suave (cercano al blanco) tienen la menor cantidad de actividad fotosintética, los de color verde brillante tienen la mayor actividad fotosintética y los píxeles blancos pertenecen a las zonas enmascaradas.

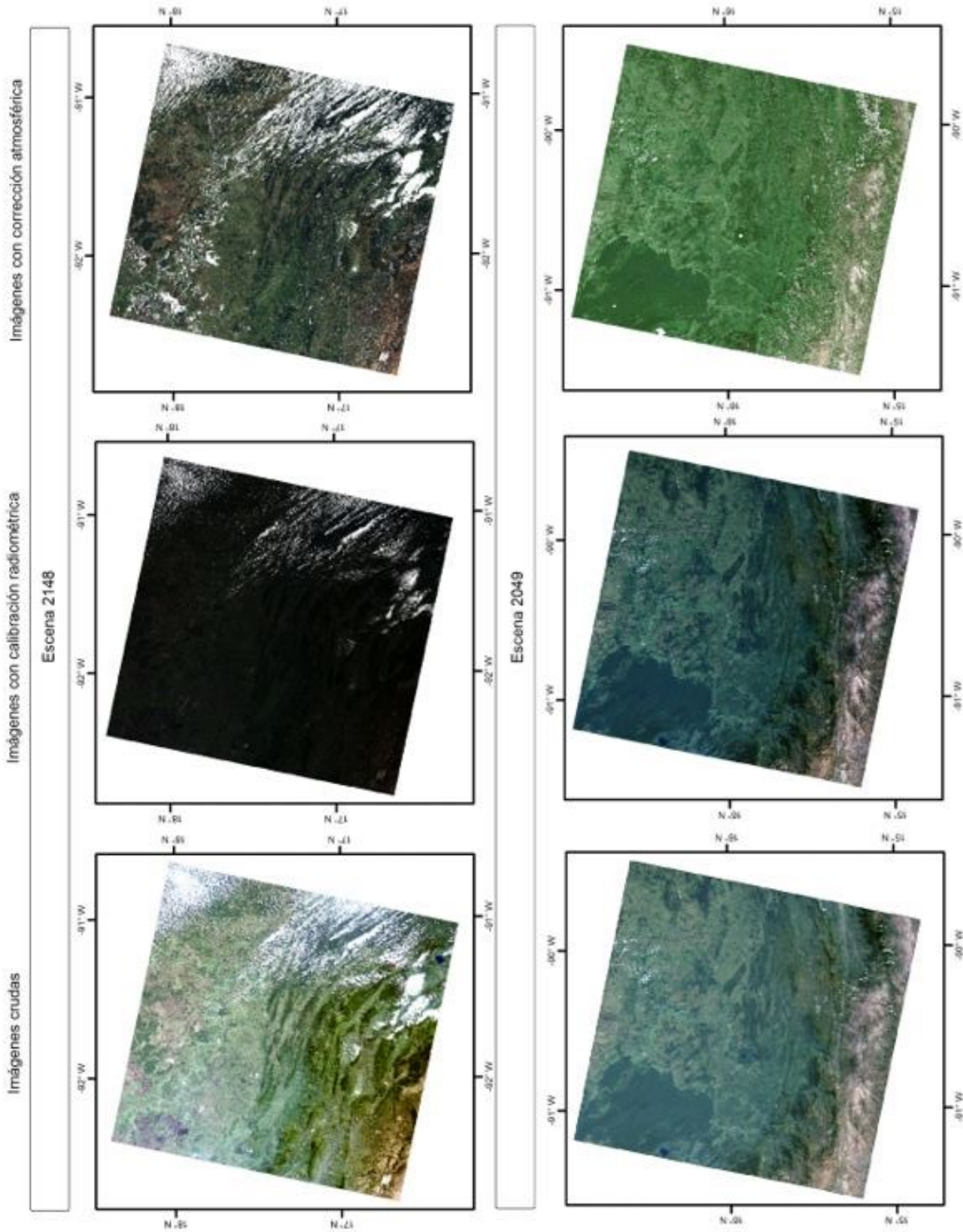


Figura 5.3. Calibración radiométrica y corrección atmosférica de las imágenes del 2015.

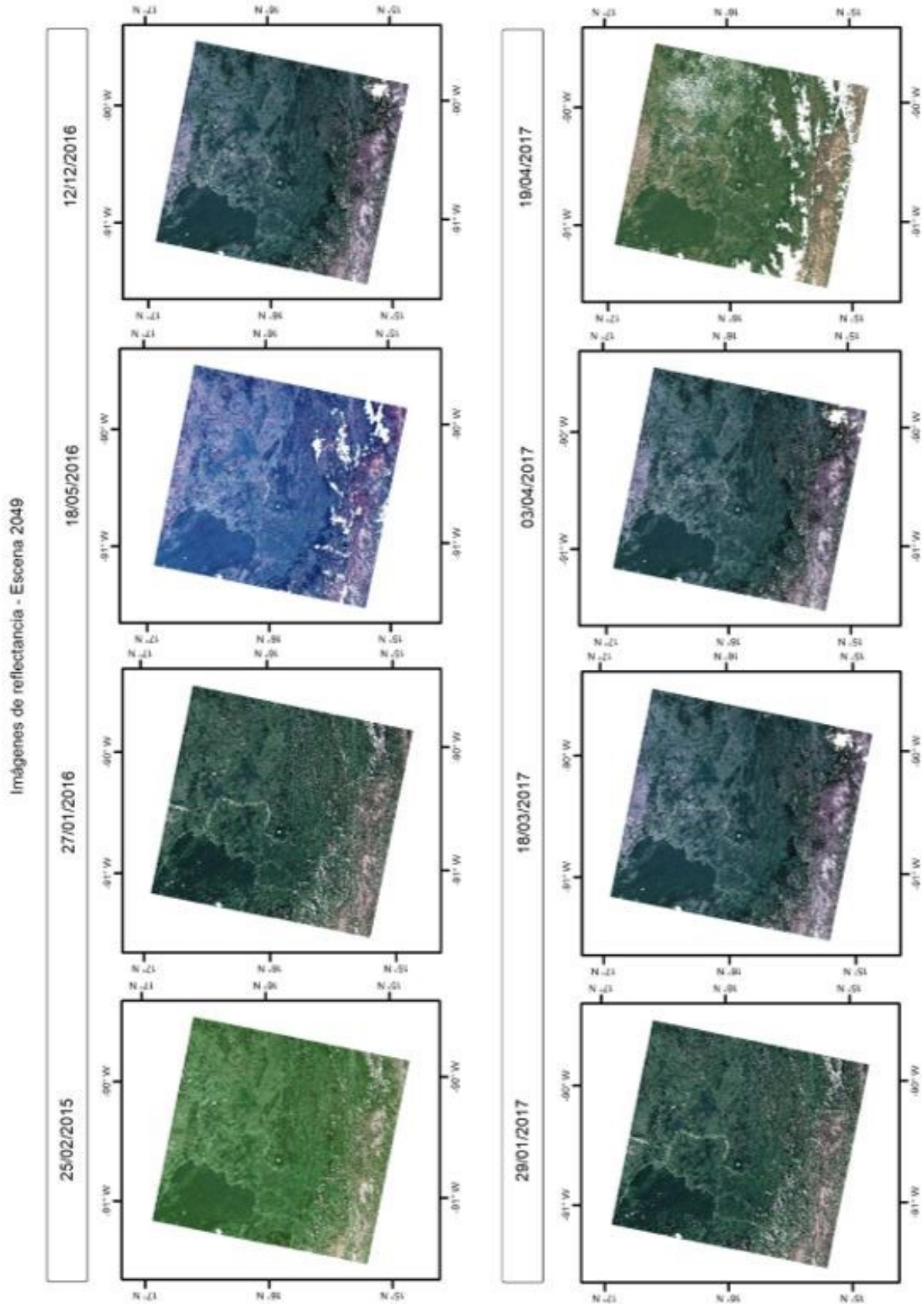


Figura 5.4. Imágenes de reflectancia de la escena 2049.

Imágenes de reflectancia - Escena 2148

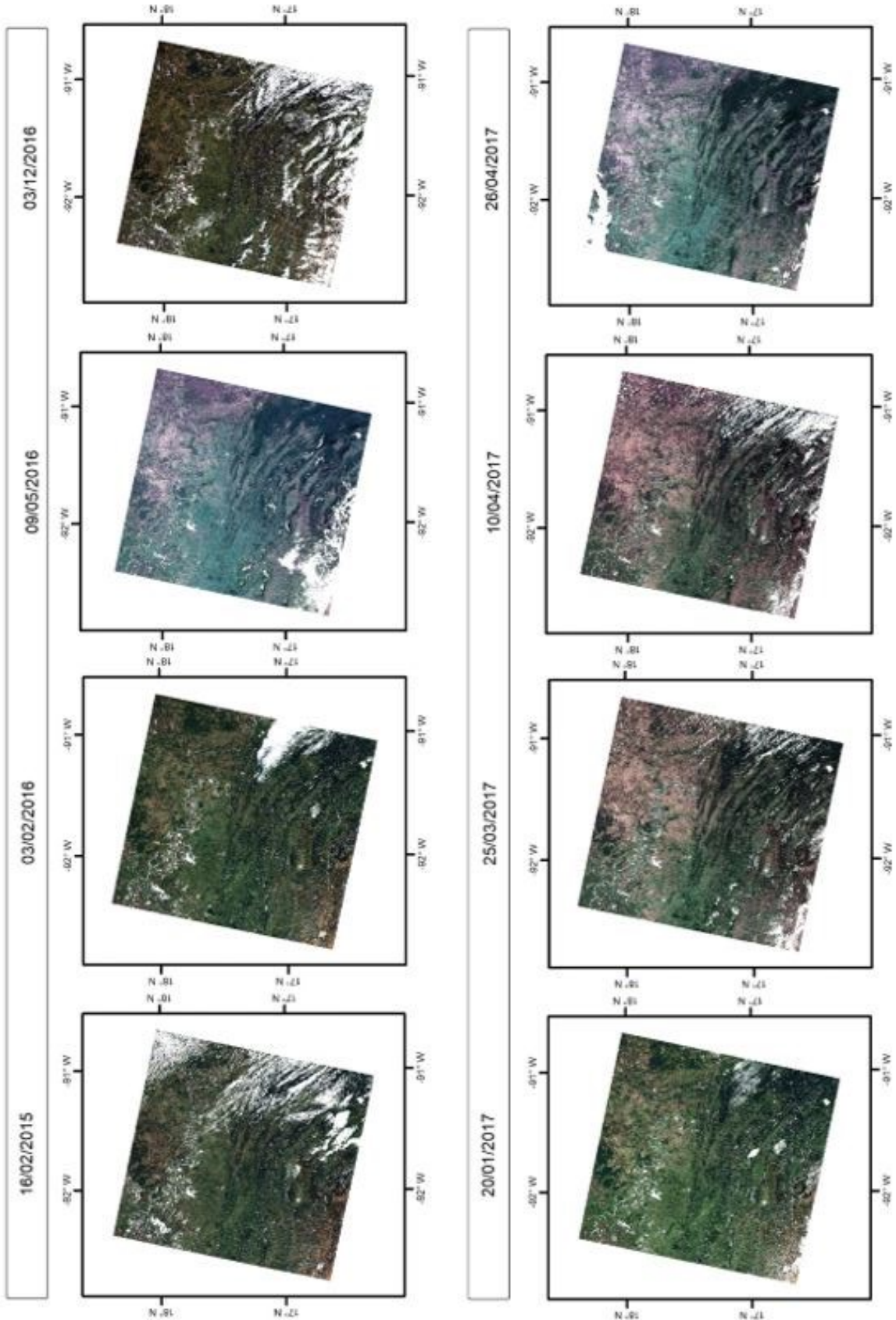


Figura 5.5. Imágenes de reflectancia de la escena 2148.

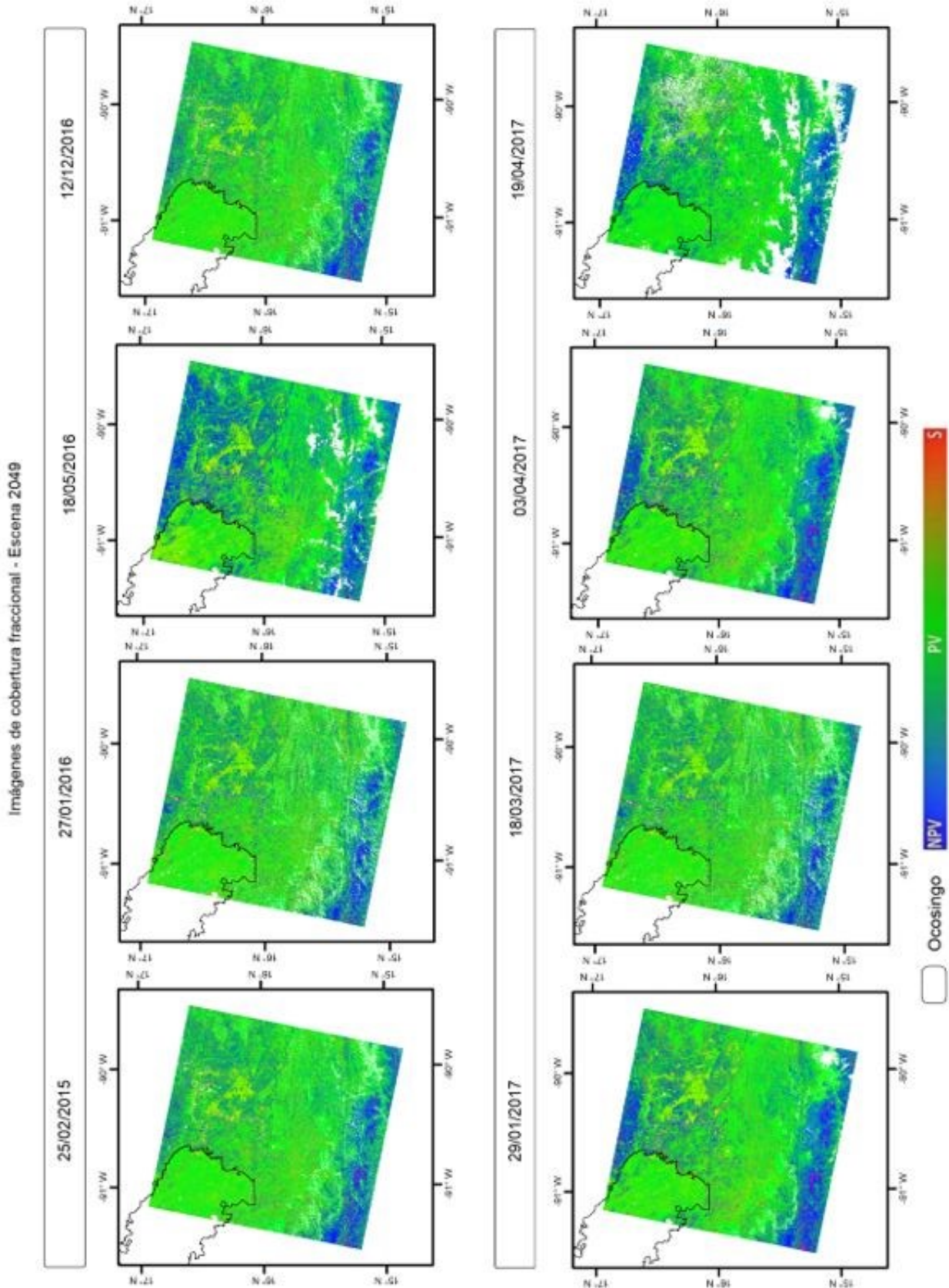


Figura 5.6. Imágenes de cobertura fraccional de la escena 2049.

Imágenes de cobertura fraccional - Escena 2148

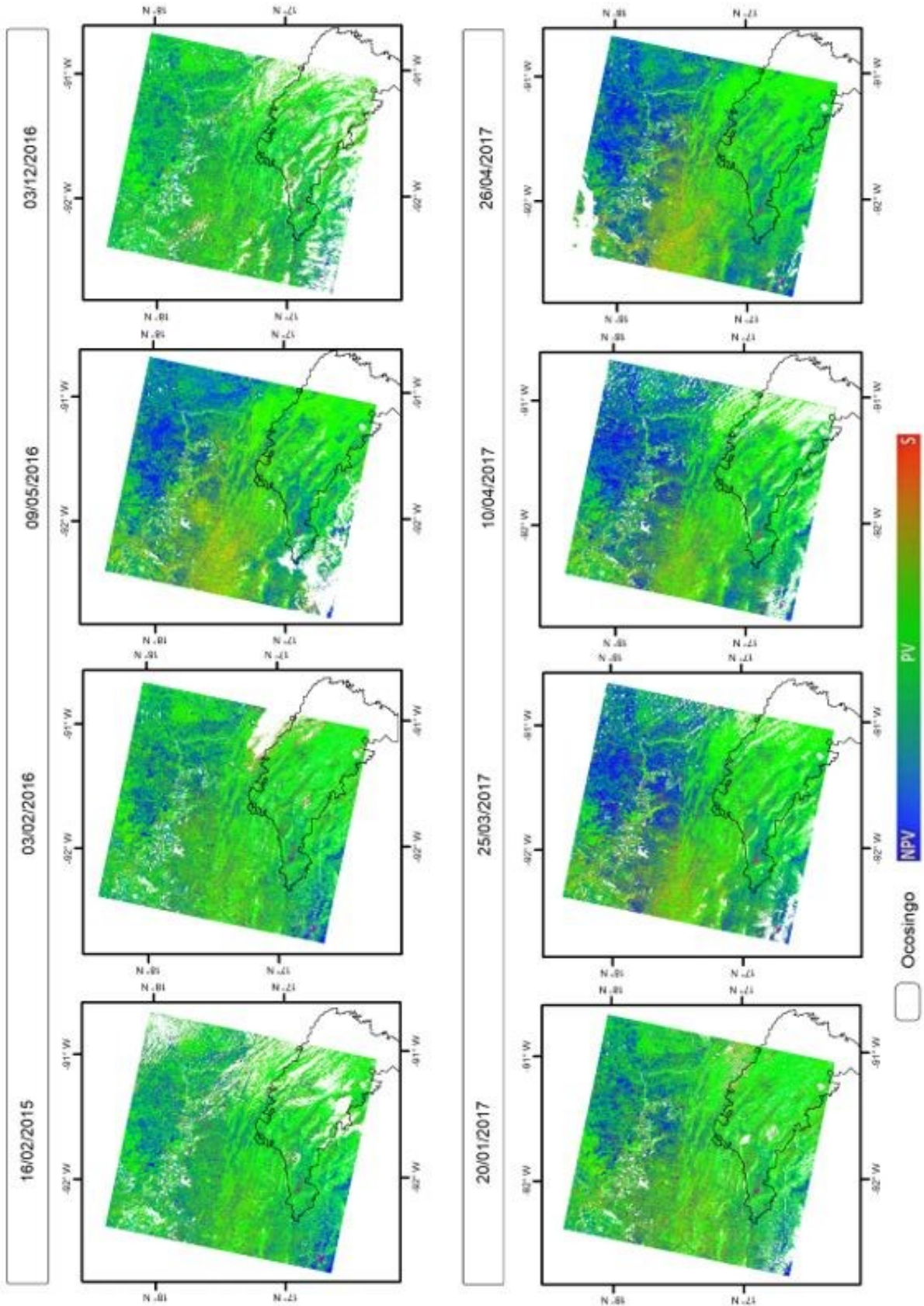


Figura 5.7. Imágenes de cobertura fraccional de la escena 2148.

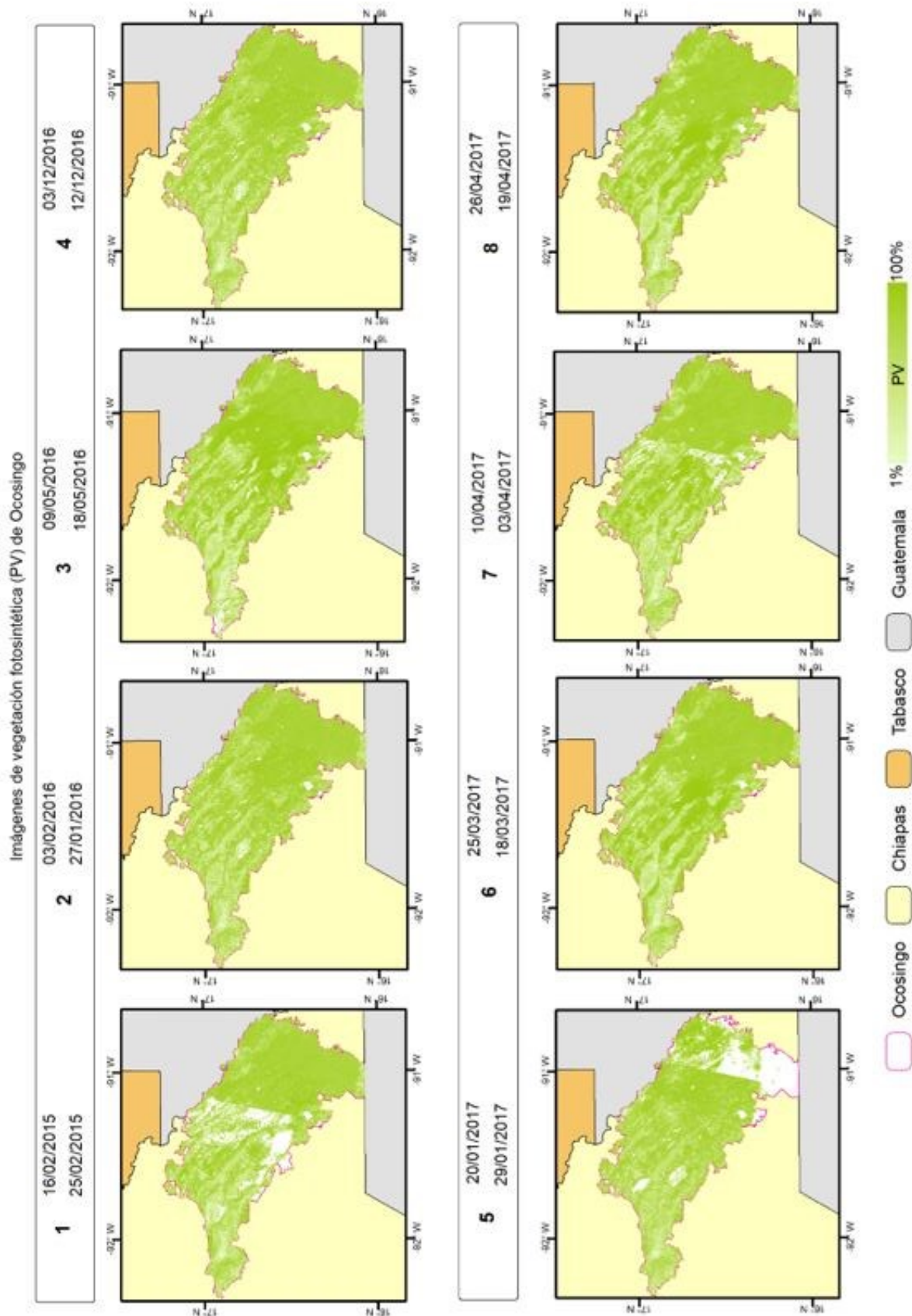


Figura 5.8. Imágenes de vegetación fotosintética (PV) de Ocosingo.

5.7 Índice de Vegetación de Diferencia Normalizada

más de 70% de actividad fotosintética y en blanco los píxeles enmascarados.

En la figura 5.9 se muestra el resultado del cálculo y reclasificación del NDVI obtenido con las imágenes del año 2015 unidas y recortadas a Ocosingo. En verde se observa la vegetación con

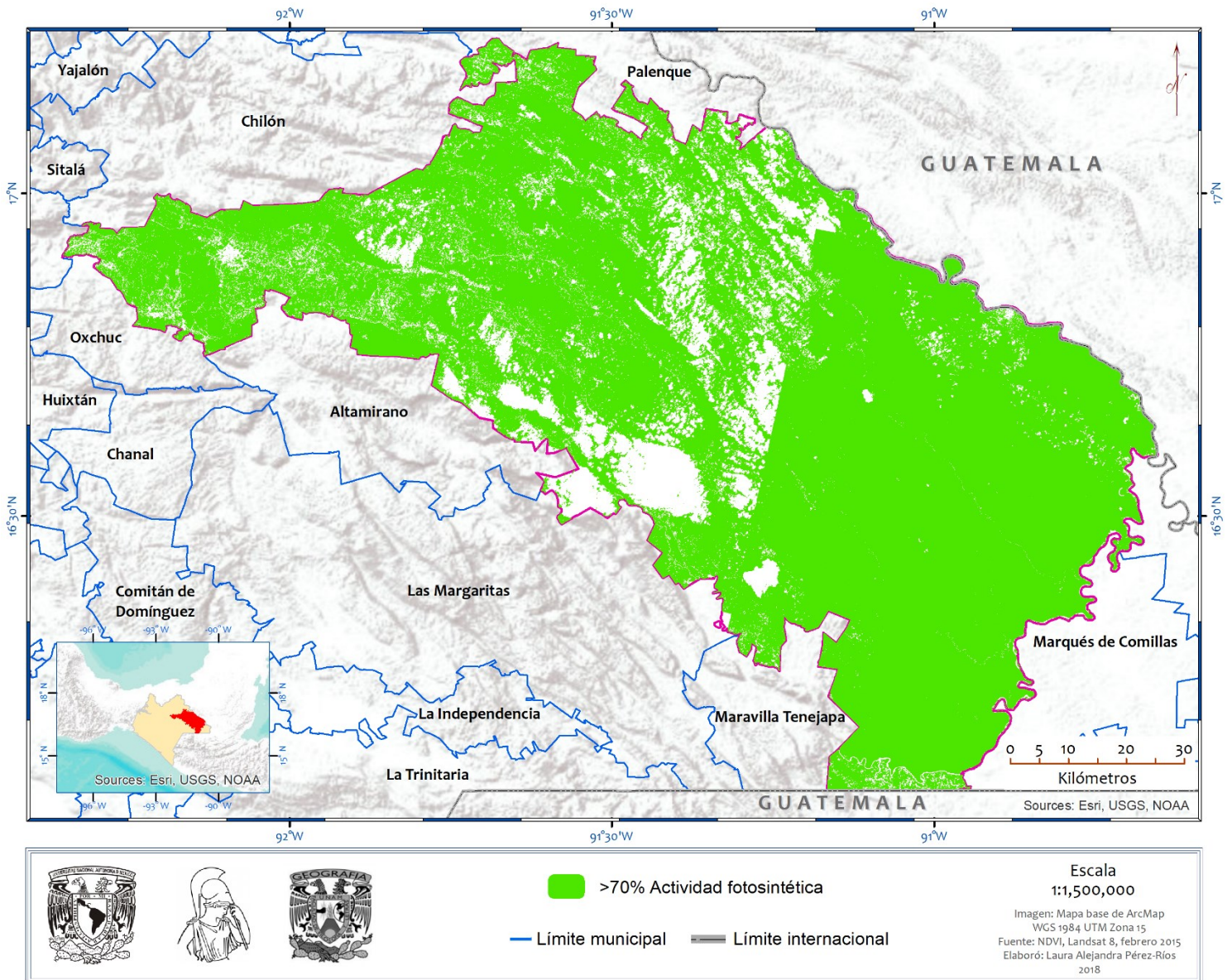


Figura 5.9. NDVI de Ocosingo del año 2015.

5.8 Puntos de control

En la figura 5.10 se muestra la distribución espacial de los 19,000 puntos de control que se crearon para cada clase y que se dividieron a la mitad aleatoriamente para crear puntos de entrenamiento y puntos de verificación. En color rojo se muestran los 9,500 puntos de entrenamiento de cada clase y en verde los 9,500 de verificación.

5.9 Gráficas de separabilidad espectral

Con los puntos de control se hizo el cálculo de los promedios de vegetación fotosintética y las gráficas generadas se muestran en las figuras 5.11 y 5.12. La primera muestra la separabilidad espectral entre el café y la vegetación obtenida del NDVI, y la segunda muestra la separabilidad entre el café y la vegetación obtenida de la Serie VI. En ambas gráficas las imágenes 3, 6 y 8 tienen la mayor diferencia de porcentaje de PV, por lo que esas tres fueron las elegidas para la clasificación *Random Forest*. En la primera gráfica la diferencia entre clases fue de 61.53% en la imagen 3, 64.81% en la imagen 6 y 61.33% en la imagen 8, mientras que en la segunda gráfica la diferencia fue de 18.37%, 16.57% y 17.59% respectivamente.

5.10 Clasificación *Random Forest* con NDVI

La imagen que resultó al clasificar las imágenes 3, 6 y 8 con los puntos de entrenamiento de café y de vegetación del NDVI se muestra en la figura 5.13, en rojo se observa lo clasificado como café y en verde lo clasificado como vegetación, al igual que en las demás imágenes, el blanco ocupa el lugar de los enmascaramientos, en este caso de las tres imágenes utilizadas.

5.11 Clasificación *Random Forest* con Serie VI

La imagen 5.14 muestra la imagen obtenida de la clasificación de las tres imágenes seleccionadas utilizando los puntos de entrenamiento de café y vegetación de la Serie VI, al igual que en la imagen anterior, los píxeles blancos están enmascarados, el rojo muestra los píxeles clasificados como café y el verde la vegetación natural.

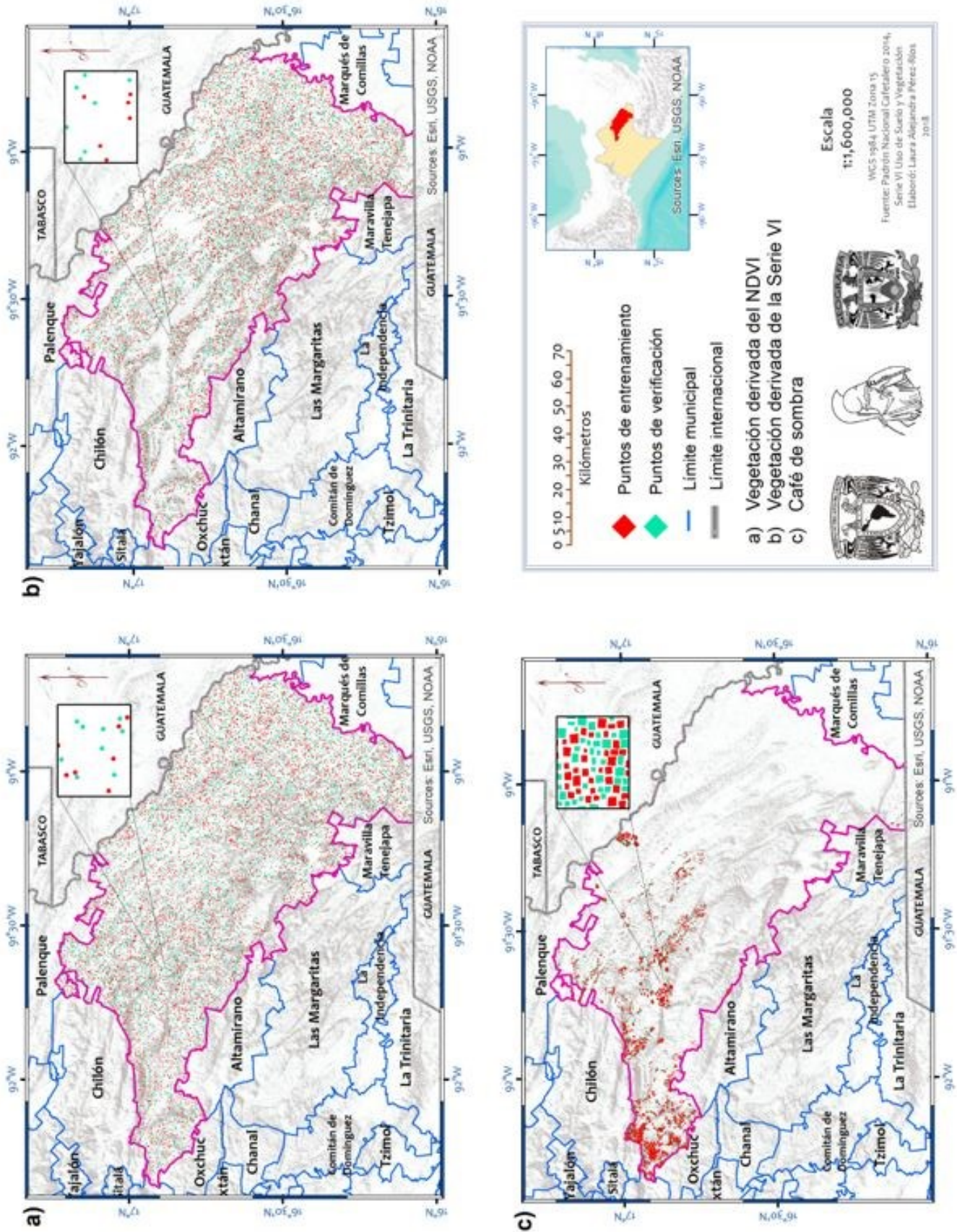


Figura 5.10. Puntos de control de cada clase.

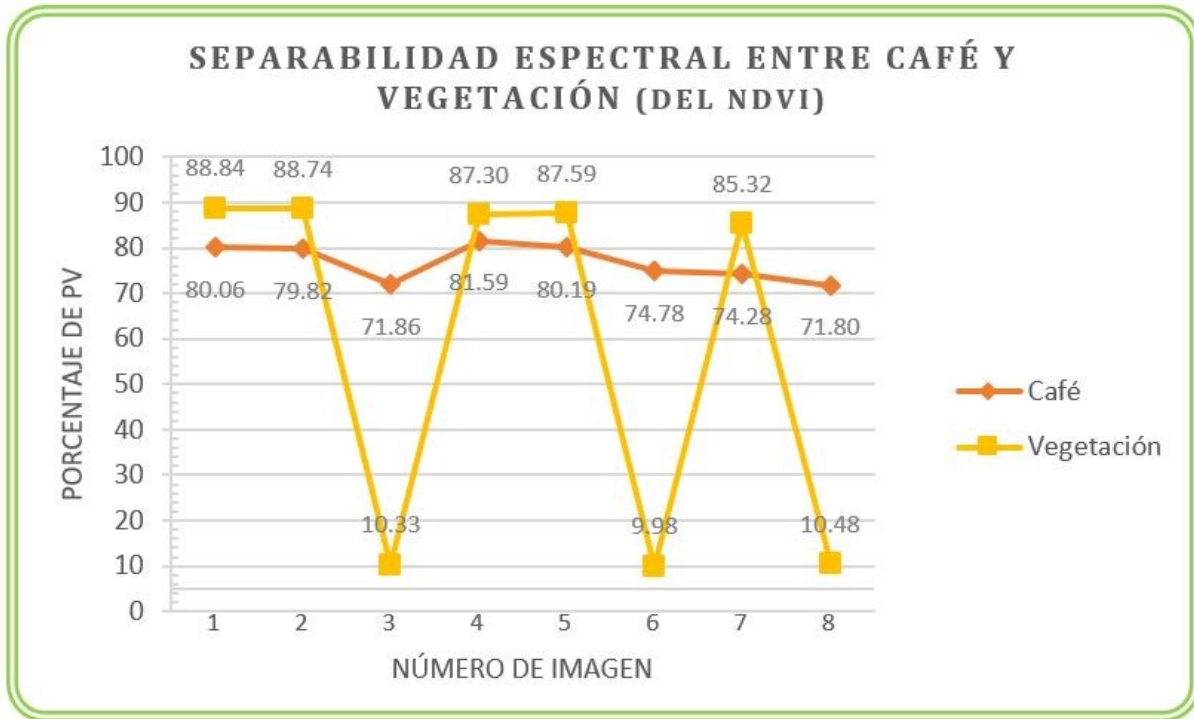


Figura 5.11. Gráfica de separabilidad espectral entre café y vegetación (del NDVI).

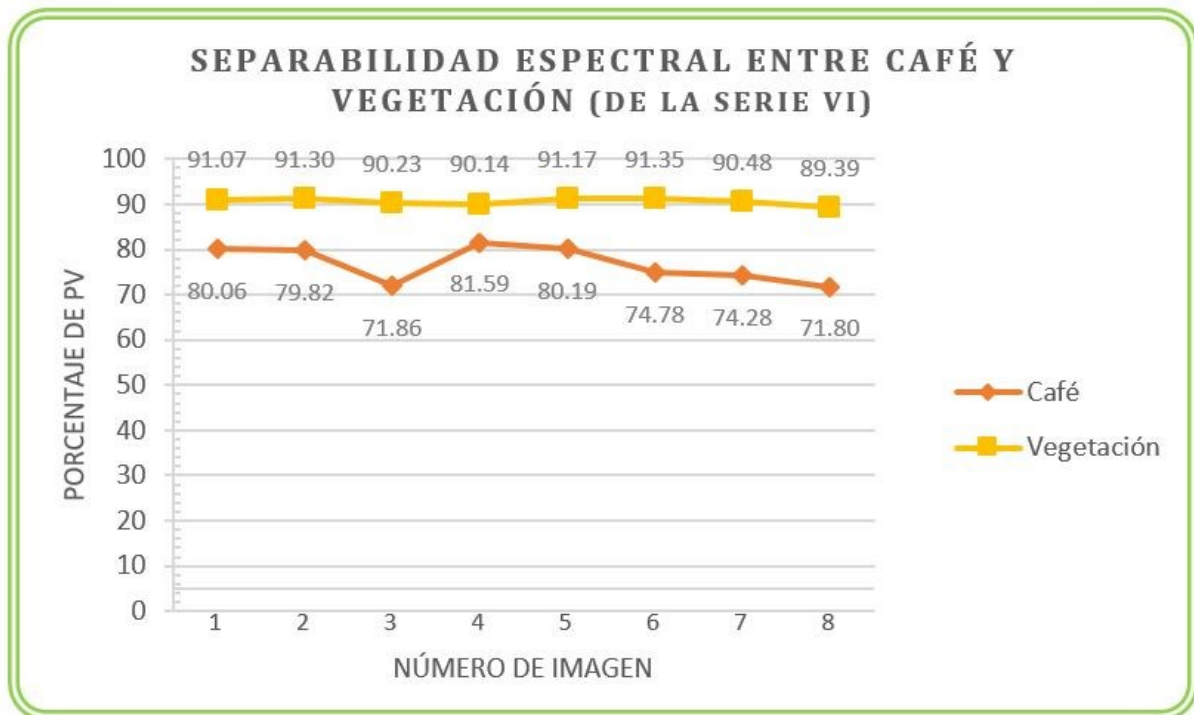


Figura 5.12. Gráfica de separabilidad espectral entre café y vegetación (de la Serie VI).

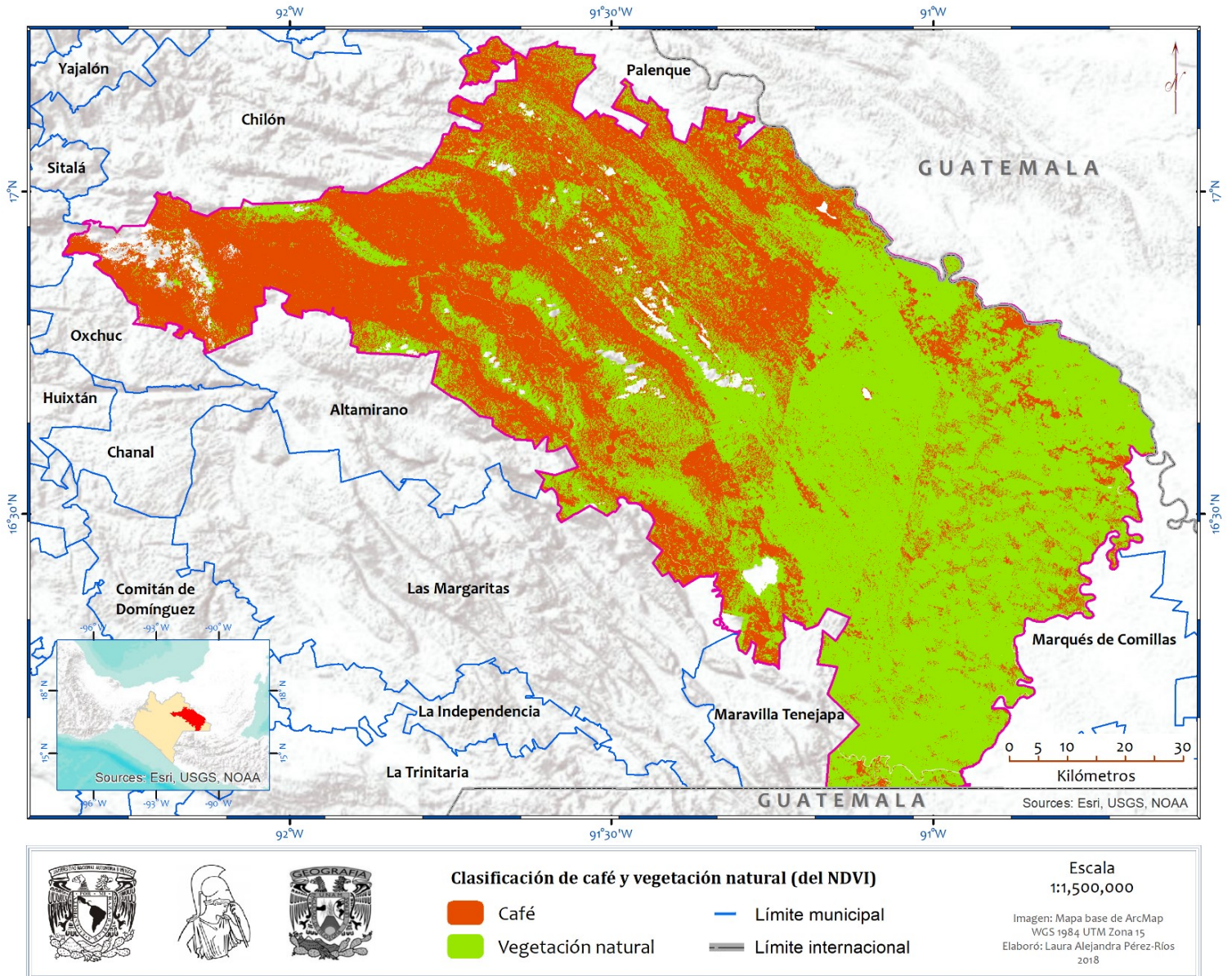


Figura 5.13. Clasificación de café y vegetación natural (del NDVI).

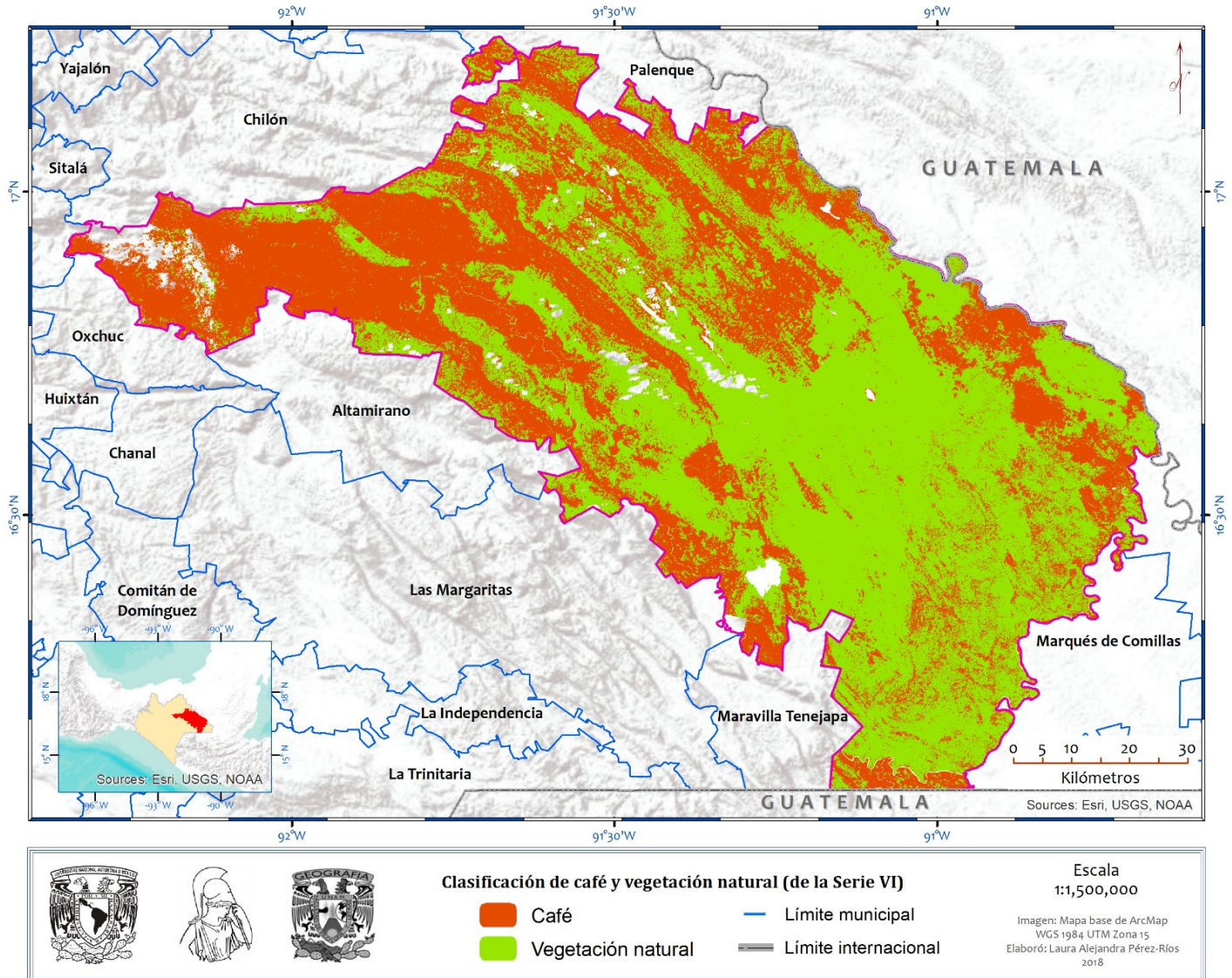


Figura 5.14. Clasificación de café y vegetación natural (de la Serie VI).

5.12 Matrices de confusión

En los cuadros 5.1 y 5.2 se muestran los resultados de los cálculos de las matrices de confusión de cada clasificación. En la primera clasificación el resultado de confiabilidad global alcanzó el 72.9% y el índice Kappa tuvo un valor de 0.42 que indica una concordancia moderada.

De los 9,500 puntos de control de cada clase utilizados para la verificación, en la primera clasificación sólo se obtuvieron valores de 5,418 puntos de verificación café, de los cuales 3,516 fueron correctamente clasificados y 9,359 de vegetación, de los cuales 7,257 fueron correctamente asignados. En total, de los 19,000

puntos de verificación se obtuvieron los valores de 14,777 puntos, el resto no se clasificó en ninguna de las dos clases porque se encuentran en las zonas enmascaradas. En la segunda clasificación los valores de confiabilidad global e índice Kappa descendieron considerablemente, con resultados de 43.6% y -0.005 respectivamente. De los puntos de verificación de café se clasificaron correctamente 3,911 de 5,418 totales, de la clase de vegetación se clasificaron

correctamente 2,538 de 9,344. En total, se obtuvieron 14,762 valores de los 19,000 puntos de verificación utilizados. Se muestran también los errores de omisión y comisión de ambas clasificaciones expresados en porcentajes, así como las exactitudes del productor y del usuario.

Cuadro 5.1. Matriz de confusión de la clasificación de café y vegetación natural (del NDVI).

Confiabilidad global	0.72903	72.903%		
Índice Kappa	0.421			
		Realidad		
		Café	Vegetación	Total
Clasificación	Café	3,516	2,102	5,618
	Vegetación	1,902	7,257	9,159
	Total	5,418	9,359	14,777
Error de comisión		%		
	Café	37.415		
	Vegetación	20.766		
Error de omisión		%		
	Café	35.105		
	Vegetación	22.459		
Exactitud del productor		%		
	Café	64.894		
	Vegetación	77.540		
Exactitud del usuario		%		
	Café	62.584		
	Vegetación	79.233		

Cuadro 5.2. Matriz de confusión de la clasificación de café y vegetación natural (de la Serie VI).

Confiabilidad global	0.43686	43.686%			
Índice Kappa	-0.005				
Clasificación		Realidad			
		Café	Vegetación	Total	
		Café	3,911	6,806	10,717
		Vegetación	1,507	2,538	4,045
	Total	5,418	9,344	14,762	
Error de comisión		%			
	Café	63.506			
	Vegetación	37.255			
Error de omisión		%			
	Café	27.814			
	Vegetación	72.838			
Exactitud del productor		%			
	Café	72.185			
	Vegetación	27.161			
Exactitud del usuario		%			
	Café	36.493			
	Vegetación	62.744			

5.13 Superficie de café en las clasificaciones

Se hizo el cálculo del área cubierta por café de acuerdo al resultado de las clasificaciones y se graficó el resultado (Figura 5.15). En la clasificación con la vegetación del NDVI se obtuvo una superficie de café de 13,169.9 hectáreas y

para la clasificación con la Serie VI la superficie fue de 13,542.9 hectáreas, es decir, en la segunda clasificación se asignaron 372.9 hectáreas más de café que en la primera, que equivale a 124,328 píxeles adicionales que fueron asignados a la clase de café.

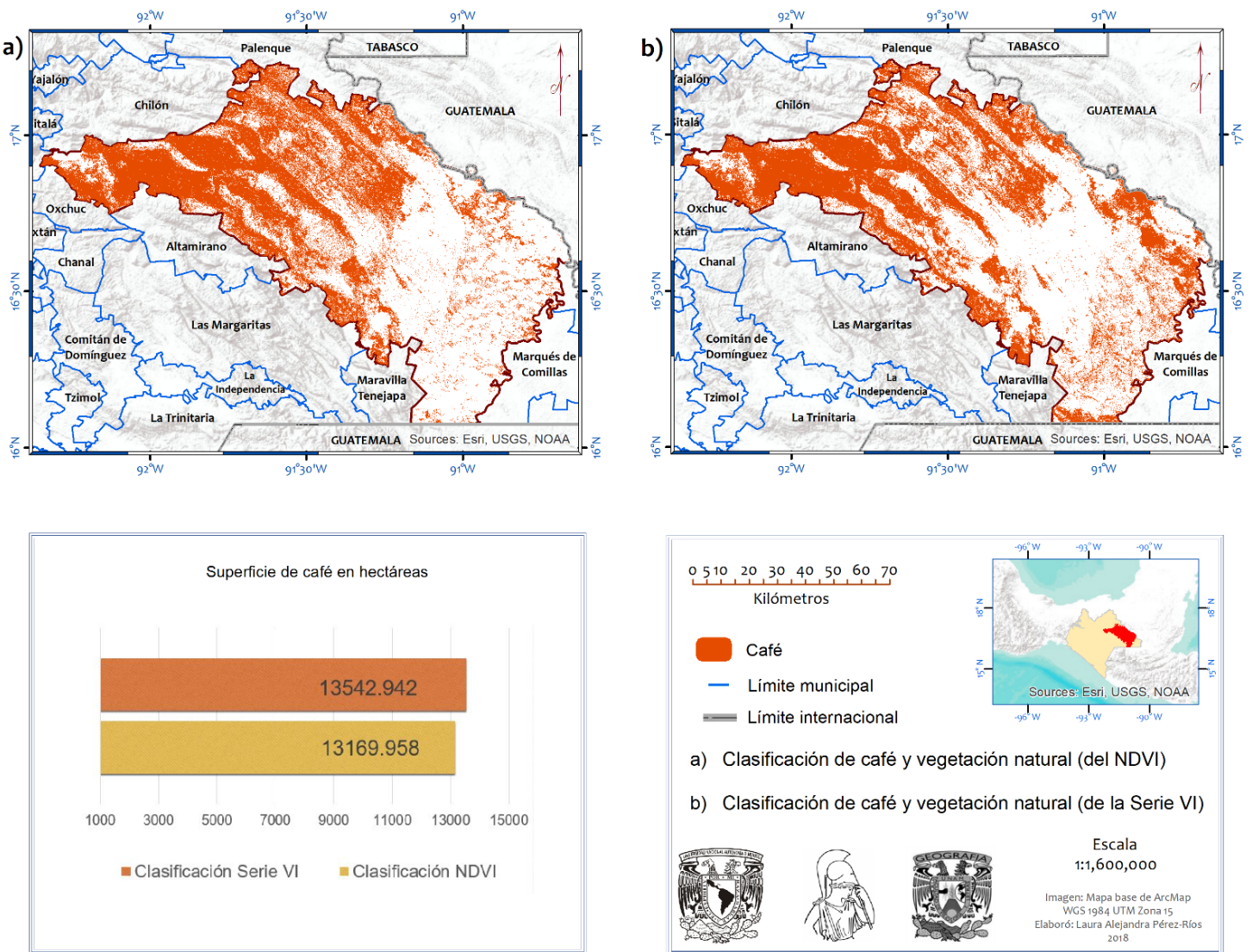


Figura 5.15. Superficie de café en las clasificaciones.



VI DISCUSIÓN

Se observa que las zonas donde hay parcelas de café se encuentran mayormente donde la vegetación dominante es el pastizal cultivado, esto quiere decir que supuestamente sería fácil distinguir el café de la demás vegetación natural de acuerdo al comportamiento espectral de cada clase, que como menciona Lillesand et al., (2008), es lo que permite que exista separabilidad espectral. Lo anterior se distingue en las imágenes de cobertura fraccional (Ver figuras 5.6 & 5.7), especialmente en la banda de PV, pues casi la totalidad de los píxeles de la zona de estudio se encuentra en valores de vegetación fotosintética, excepto los que pertenecen a pastizales que en la temporada de la captura de las imágenes se encontraban como vegetación no fotosintética o senescente.

De la banda utilizada (PV), los únicos píxeles que no tuvieron valores fueron los que estaban enmascarados (nubes y sombras) y los pocos píxeles encontrados como superficie descubierta (sin vegetación), pues los de vegetación no fotosintética se mostraron con bajo porcentaje de PV (Ver figura 5.8), esto implicó que el clasificador asignara una de las dos clases (café o vegetación natural) a los píxeles que no tuvieron valores de vegetación fotosintética.

En el resultado del NDVI (Ver figura 5.9), una gran parte de la superficie presentó píxeles que se

enmascararon por tener nubes, esto pudo haber influido en la colocación de los puntos de control de vegetación natural, pues en esas zonas cubiertas por nubes es muy probable que también hubiera vegetación con más de 70% de actividad fotosintética que pudo haber quedado fuera del análisis. Autores como Bolanos (2007), evitan esta pérdida de puntos de control de vegetación al clasificar las nubes y las sombras como dos clases más y utilizar puntos específicos para dichas clases.

Los puntos de control fueron creados pensando que a mayor cantidad de puntos mayor certeza estadística del clasificador (Oniga, Breaban & Statescu, 2018), además no se colocaron sobre cualquier lugar o zona aleatoria, sino que se distribuyeron en todo el municipio para cubrir la mayor superficie posible (Sedorovich, O'Hara & Schukman, 2010), tanto de vegetación derivada del NDVI como de la vegetación seleccionada de la Serie VI de Uso de Suelo y Vegetación.

En cuanto a la separabilidad espectral entre el café y la vegetación, cuando se utilizó el NDVI las diferencias entre los promedios del valor de la PV de cada imagen fueron muy grandes en las tres imágenes seleccionadas (3, 6 y 8), mientras que en las demás imágenes la diferencia no pasaba de 11% (Ver figura 5.11), esto quiere decir que al menos en esas tres imágenes se tuvo una buena

separabilidad espectral. En el caso de las diferencias espectrales entre el café y la vegetación de la Serie VI (Ver figura 5.12), la vegetación en las ocho imágenes tuvo porcentajes de PV más altos que los obtenidos con la vegetación del NDVI, sin embargo, la separabilidad entre clases fue mucho menor con la Serie VI que con el NDVI.

Esto puede explicarse por el comportamiento espectral del café y la vegetación en las fechas de captura de estas tres imágenes (Moreira et al., 2004), correspondientes a mayo, marzo y abril respectivamente, pues las demás imágenes no son de estos meses, excepto la imagen número 7 que también es de abril y es la siguiente con mayor diferencia espectral.

La primera clasificación se hizo con la vegetación del NDVI, en la imagen que resultó (Ver figura 5.13), se observó que los píxeles clasificados como café se encuentran principalmente sobre las zonas de pastizal, bosque de pino, bosque de pino-encino y bosque mesófilo de montaña de acuerdo a la clasificación de la vegetación de la Serie VI. En el caso de los bosques tendría sentido por los requerimientos altitudinales que requiere el cultivo de café (Diario Oficial de la Federación, 2003). El sistema de producción en las zonas de pastizal sería de café de sol, mientras que en los

bosques sería un sistema de sombra (Toriello, 2014).

En la segunda clasificación (Ver figura 5.14) pasó lo mismo que en la anterior, con la diferencia de que en las zonas donde la selva es la vegetación predominante el clasificador asignó la clase de café a muchos más píxeles que en la primera clasificación, esto se explica porque los valores de vegetación fotosintética de esas zonas tuvieron menor separabilidad espectral, de acuerdo con lo observado en la figura 5.12, los valores de PV de la segunda clasificación tuvieron mayor probabilidad de ser confundidos al momento de la asignación de clases.

Visualmente las dos imágenes que resultaron de las clasificaciones lucen parecidas, aunque con algunas diferencias en la distribución de la clase de café; por ello se esperaban resultados de verificación similares. Sin embargo, cuando se calcularon las matrices de confusión, los valores de confiabilidad global e índice Kappa difirieron considerablemente.

La primera clasificación tuvo 72.9% de confiabilidad global, porcentaje considerado aceptable, y un índice Kappa de 0.42 (Ver cuadro 5.1) que indica que la clasificación tuvo una fuerza de concordancia moderada y mejor que aceptable. En contraste, la segunda clasificación

no tuvo buenos resultados, en la matriz de confusión (Ver cuadro 5.2) obtuvo 43.6% de confiabilidad global y un valor de índice Kappa de -0.005, que indica una fuerza de concordancia extremadamente baja.

Lo anterior puede explicarse por dos razones, la primera es que los píxeles que en realidad pertenecen a vegetación hayan sido erróneamente clasificados como café o viceversa, porque no existió suficiente separabilidad espectral entre ellos, como le sucedió a Bolanos (2007). Para la segunda razón hay que recordar que las tres imágenes que se utilizaron para la clasificación fueron de los años 2016 y 2017, por lo que es posible que los píxeles clasificados como café fueron correctamente asignados y su cultivo se ha comenzado a practicar en zonas en las que hasta el año 2014 (año del padrón) no habían sido registradas como parcelas de cultivo de café, pues éste es un cultivo que se ha ido expandiendo de acuerdo a las necesidades del mercado (Meyfroid, Tan & Viet, 2013).

En la figura 6.1 se observan las hectáreas de café sembradas y reportadas en el Servicio de Información Agroalimentaria y Pesquera (SIAP) de los años 2014 al 2017, y las hectáreas de café obtenidas en las dos clasificaciones. En los años 2014, 2015 y 2016 la superficie sembrada de acuerdo al SIAP fue la misma con 8,507 hectáreas,

mientras que en el año 2017 la superficie disminuyó 2,781 hectáreas, lo que equivale a una reducción del 32,7%. La clasificación hecha con la Serie VI obtuvo un resultado de 59% más superficie que en los datos de los años 2014 al 2016 y 136.5% más que en el año 2017. La clasificación con el NDVI obtuvo 54.8% más superficie que lo reportado de 2014 al 2016 y 130% más que en el año 2017.

Lo anterior pudo deberse a dos principales razones, una es que en el municipio existen más zonas de cultivo de café además de las que se han reportado oficialmente porque es café producido únicamente para autoconsumo y el SIAP reporta sólo las superficies del café que se comercializa, la otra razón es que las zonas de café excedentes son de la vegetación que tiene valores de actividad fotosintética muy similares a los del café; por lo que el clasificador confundió la clase de vegetación natural con la clase de café y la clasificó como tal. Estas confusiones suceden siempre que las distintas clases o categorías presentan valores espectrales muy similares entre sí (Chuvieco, 2010).

Los resultados de las clasificaciones de café hechas por otros autores (Cuadro 6.1), en general tienen valores más altos de confiabilidad global (desde 63.14% hasta 90.5%) e índice Kappa, que los obtenidos en las dos clasificaciones realizadas;

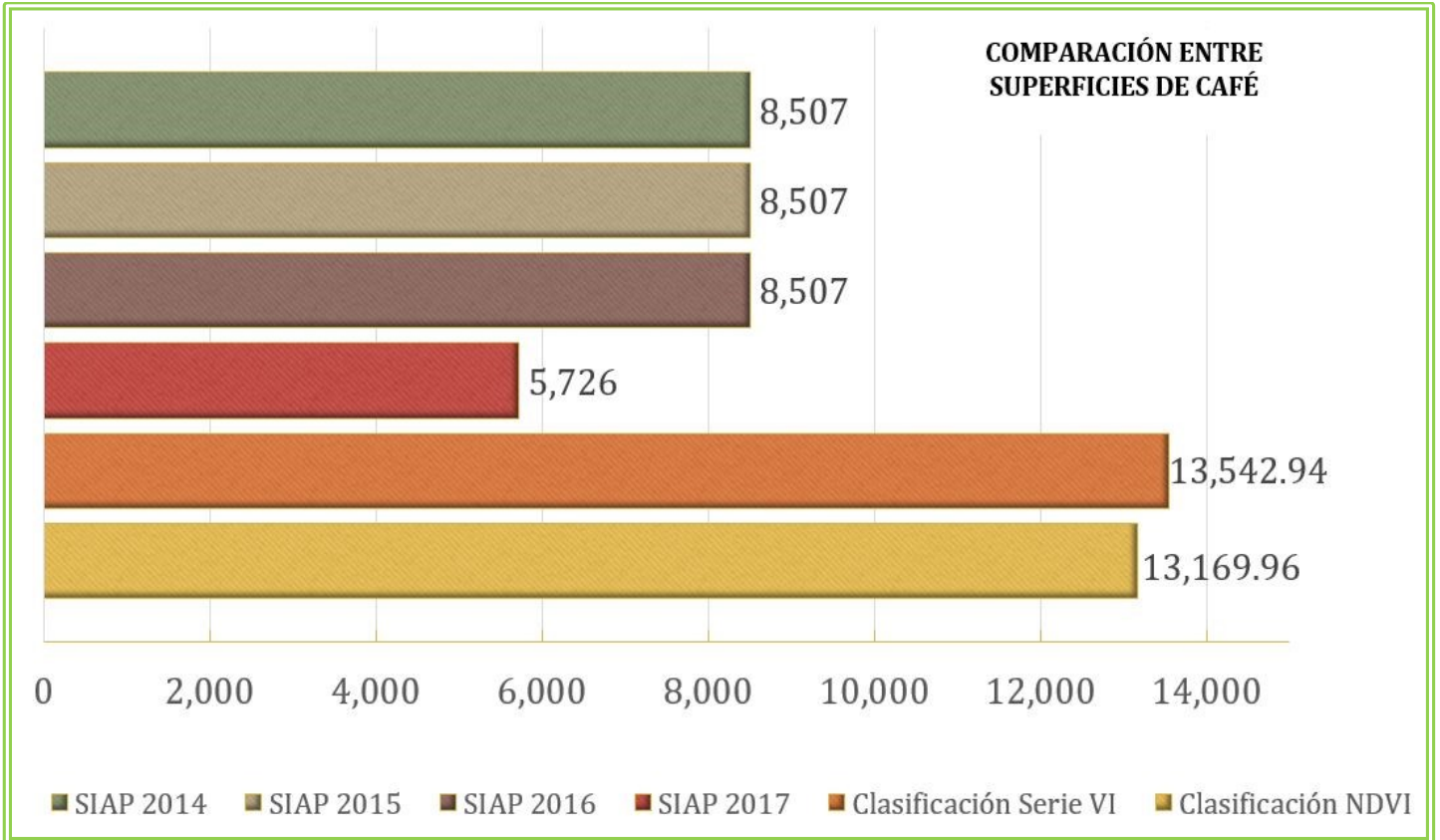


Figura 6.1. Comparación de superficies de café (hectáreas) entre las clasificaciones y los datos del SIAP.

sin embargo, sus resultados contemplan varias clases, y no sólo dos como en este trabajo.

Por lo anterior, en algunas categorías utilizadas por otros autores el resultado fue muy bueno para clases como agua y nubes, y en otras fue más bajo como en los pastizales, la agricultura y el café, a pesar de esto, en sus matrices de confusión los valores bajos de ciertas clases no afectaron mucho sus resultados finales porque los valores de las demás clases fueron muy altos.

A diferencia de lo que se hizo en este trabajo, otros autores (Bolanos, 2007; Cordero-Sancho & Sader, 2007; Ortega-Huerta et al., 2012; Kelley et al., 2018) utilizaron más datos para ayudar al clasificador a tener un mejor resultado, y comprobaron que incluyendo la elevación y otros datos fisiográficos hacen que la clasificación sea más exacta. Aquí solamente se utilizaron los porcentajes de vegetación fotosintética de la zona de estudio obtenidos del desmezclado espectral.

En el cuadro 6.1 se muestran los acercamientos que se han tenido para clasificar café en distintas clases.

Cuadro 6.1. Comparación de los resultados de distintos métodos de clasificación de café.

Artículo	Datos utilizados	Método(s) de clasificación	Resultados
<i>Using Google Earth Engine to Map Complex Shade-Grown Coffee Landscapes in Northern Nicaragua.</i> (Kelley, Pitcher & Bacon, 2018)	-143 Imágenes Landsat 8 <i>Top of Atmosphere</i> (TOA) -200 puntos de control -10 clases -Datos de brillo, humedad, topografía y temperatura.	-Transformación lineal <i>Kauth-Thomas</i> (KT) -Clasificación <i>Random Forest</i>	Confiabilidad global: 90.5% Kappa: 0.89
<i>Vegetation classification of Coffea on Hawaii using WorldView-2 satellite imagery.</i> (Gaertner, Genovese, Potter, Sewake & Manoukis, 2017)	-2 imágenes WorldView-2 -2366 puntos de control (ML) -6446 puntos de control (OBIA) -8 clases (ML) -10 clases (OBIA)	1-Máxima verosimilitud (ML)	Confiabilidad global: 68% Kappa: 0.60
		2-Clasificación basada en objetos (OBIA)	Confiabilidad global: 81% Kappa: 0.75
<i>Mapping coffee crops in southeastern Brazil using spectral mixture analysis and data mining classification.</i> (Kawakubo & Pérez, 2016)	-1 imagen LISS-III -1600 puntos de control -8 clases -4 <i>endmembers</i>	-Componentes principales - <i>Desmezclado espectral</i> (SMA) -Clasificación <i>Data Mining</i> (DM)	Confiabilidad global: 73.7% Kappa: 0.67
<i>Mapping coffee plantations with Landsat imagery: an example from El Salvador.</i> (Ortega-Huerta, Komar, Price & Ventura, 2012)	-3 imágenes Landsat 5 -285 Puntos de control -4 clases -Datos de elevación	-Clasificación ISODATA -Componentes principales -Análisis de clústeres	Confiabilidad global: 76.74% Kappa: 0.68
<i>Coupling Community Mapping and supervised classification to discriminate Shade coffee from Natural vegetation.</i> (Martínez-Verduzco, Galeana-Pizaña & Cruz-Bello, 2011)	-Imágenes SPOT 5 -219 puntos de control -4 clases -ortofotos	1-Máxima verosimilitud	Confiabilidad global: 73% Kappa:0.70
		2-Mapeo participativo	Confiabilidad global: 74.57% Kappa:0.73
<i>Using Image Analysis and GIS for Coffee Mapping.</i> (Bolanos, 2007)	-1 imagen Landsat 7 -1429 puntos de control -8 clases -Modelo de elevación digital	1-Clasificación basada en píxeles	Confiabilidad global: 71.89% Kappa: 0.66
		2-Clasificación basada en objetos (OBIA)	Confiabilidad global: 74.88% Kappa: 0.70
<i>Spectral analysis and classification accuracy of coffee crops using Landsat and a topographic environmental model.</i> (Cordero-Sancho & Sader, 2007)	-1 imagen Landsat 7 -293 puntos de control -5 clases - <i>Shuttle Radar Topography Mission</i> (SRTM) -Datos de elevación y precipitación	1- <i>Coffee environmental stratification model</i>	Confiabilidad global: 65.19% Kappa: 0.56
		2-Máxima verosimilitud	Confiabilidad global: 63.14% Kappa: 0.53
El presente trabajo de tesis: <i>Clasificación de zonas de café de sombra a partir del análisis multitemporal de imágenes Landsat 8 en Ocosingo, Chiapas.</i>	-16 imágenes Landsat 8 -19,000 puntos de control por clase -2 clases -Índice de vegetación (NDVI) -Serie de Uso de Suelo y Vegetación	1-Desmezclado espectral -Clasificación <i>Random Forest</i> con NDVI	Confiabilidad global:72.9 % Kappa: 0.42
		2-Desmezclado espectral -Clasificación <i>Random Forest</i> con Serie de Uso de Suelo y Vegetación	Confiabilidad global: 43.6% Kappa: -0.005

Se observan siete artículos, cada uno con los datos que se utilizaron, como el tipo y la cantidad de imágenes satelitales, los puntos de control tomados, el número de clases en las que se clasificó y otros datos ingresados a los clasificadores; como la elevación, precipitación y temperatura, así como el clasificador empleado para realizar su respectiva clasificación. Se muestran los resultados obtenidos a través de su porcentaje de confiabilidad global y su valor de índice Kappa, al final del cuadro se muestran también los resultados obtenidos para las dos clasificaciones realizadas en este trabajo de tesis.

A partir de lo resumido en el cuadro 6.1 es posible observar que el método utilizado para clasificar también tuvo un papel determinante para la exactitud de los resultados, pues los autores que utilizaron el clasificador de máxima verosimilitud (Cordero-Sancho & Sader, 2007; Martínez-Verduzco et al., 2011; Gaertner et al., 2017) obtuvieron resultados más bajos que los obtenidos con otros métodos. Sin embargo, lo que parece determinar en mayor medida la calidad de los resultados es la cantidad de imágenes utilizadas, la resolución de las mismas, y en menor medida, la cantidad de puntos de control, pues el mejor resultado se obtuvo por Kelley et al. (2018) con pocos puntos de control, pero decenas de imágenes satelitales, mientras que el peor lo

obtuvieron Cordero-Sancho & Sader (2007) con una sola imagen.

De acuerdo con los insumos y métodos utilizados por los autores del cuadro mencionado en el párrafo anterior, se hizo una comparación de los resultados a partir de que tan mejor o peor fue el valor de la confiabilidad global y el índice Kappa que obtuvieron contra lo que se obtuvo en las dos clasificaciones realizadas en esta tesis, principalmente, lo que se obtuvo en la clasificación con el NDVI.

Como se dijo antes, en Kelley et al. (2018) se reporta el mejor resultado al clasificar café, su confiabilidad global alcanzó el 90.5% y su índice Kappa fue casi perfecto (Landis & Koch, 1977) con un valor de 0.89, el clasificador que utilizaron fue el mismo que se utilizó en esta tesis (*Random Forest*). La gran diferencia entre los resultados de la tesis (72.9% y 0.42) y los de Kelley et al. (2018) se debe a que ellos utilizaron 143 imágenes satelitales (Landsat 8), mientras que en la tesis sólo se usaron 16, además, clasificaron en 10 clases con 200 puntos de control, mientras que en la tesis se clasificó en 2 clases con 19,000 puntos, a pesar de ello, el factor que elevó la calidad de los resultados fue la integración de datos fisiográficos de la zona de estudio, los cuales no fueron contemplados en las clasificaciones hechas en esta tesis.

El siguiente trabajo con mejores resultados fue el reportado por Gaertner et al. (2017), en donde uno de los dos métodos utilizados obtuvo 81% de confiabilidad global y 0.75 de índice Kappa, en este caso se utilizó la clasificación basada en objetos (OBIA), con 6446 puntos de control para clasificar dos imágenes WorldView-2 en 10 clases. El resultado fue mejor que el obtenido en las clasificaciones de la tesis porque el tipo de imágenes utilizadas tienen una resolución espacial mucho más fina, con bandas que van de 0.46 m a 2.4 m (DigitalGlobe, 2018), mientras que las imágenes Landsat 8 tienen 30 m, esto permitió que los valores espectrales de cada clase se identificaran con mayor claridad, reduciendo la confusión del clasificador.

El siguiente trabajo con mejores resultados es el de Ortega-Huerta et al. (2012), donde se reporta 76.74% de confiabilidad global y 0.68 de índice Kappa, aquí se utilizó la clasificación ISODATA para clasificar 3 imágenes Landsat 5 en cuatro clases, a partir de 285 puntos de control y datos de elevación de la zona de estudio, estos resultados son mejores que el obtenido en la tesis, lo cual puede explicarse porque se tomó en cuenta el factor de elevación que en la tesis no se usó y porque el método ISODATA asignó mejor las clases que el *Random Forest*.

Bolanos (2007) obtuvo una clasificación con 74.88% de confiabilidad global y 0.70 de índice Kappa, con una clasificación basada en objetos (OBIA) y una sola imagen Landsat 7, a través de 1429 puntos de control y un modelo digital de elevación la autora clasificó la imagen en 8 clases. El resultado es mejor que el obtenido en la tesis principalmente por la incorporación del modelo de elevación y por el método utilizado, pues en Gaertner et al. (2017) se observó que este método también dio buenos resultados.

A través del método de máxima verosimilitud, Martínez-Verduzco et al. (2011) clasificaron en 4 clases con ortofotos, imágenes del satélite SPOT 5 y 219 puntos de control, obtuvieron resultados de 73% de confiabilidad global y 0.70 de índice Kappa, además, hicieron otra clasificación con los mismos insumos, pero con el método de mapeo participativo, sus resultados aumentaron a 74.57% de confiabilidad global y 0.73 de índice Kappa. Los valores anteriores fueron mejores que los obtenidos en la tesis, esto puede explicarse porque las imágenes utilizadas tienen mejor resolución espacial y porque se tuvo mayor certeza sobre los puntos de control de cada clase a partir del mapeo participativo.

El último trabajo en el que los resultados reportados fueron mejores que los encontrados en la tesis fue en Kawakubo & Pérez (2016),

donde la confiabilidad global alcanzó 73.7% y el índice Kappa tuvo un valor de 0.67. Para la clasificación utilizaron una sola imagen del sensor LISS-III, 1600 puntos de control y 4 *endmembers*, a través del desmezclado espectral y una clasificación *Data Mining* se clasificó la imagen en 8 clases. Estos resultados pueden explicarse porque la resolución de la imagen es ligeramente mejor (23.5 m) que la de las imágenes utilizadas en la tesis (30 m), y porque se tuvieron valores espectrales específicos para algunas de sus clases, lo cual no se tuvo en los insumos utilizados para las clasificaciones hechas en la tesis, además, el resultado de confiabilidad global es el más cercano al obtenido en la mejor clasificación de la tesis, esto puede explicarse porque en el método incorporaron un tipo de desmezclado espectral.

Bolanos (2007) también hizo una clasificación que obtuvo una confiabilidad global de 71.89% y un índice Kappa de 0.66, con los mismos insumos señalados antes, clasificó en 8 clases con el método de clasificación basada en píxeles. A pesar de que utilizó un modelo de elevación digital, los resultados fueron más bajos que los obtenidos en su otra clasificación, su confiabilidad global fue más baja que la que se obtuvo en la primera clasificación de la tesis, esto puede deberse a que el método que utilizó fue más propenso a confundir las clases. Sin embargo, el valor de su

índice Kappa fue más alto que el de 0.42 obtenido en la primera clasificación de la tesis.

La siguiente clasificación que tuvo valores más bajos de confiabilidad global que los obtenidos en esta tesis fue la de Gaertner et al. (2017), utilizaron las mismas imágenes WorldView-2 para clasificar en 8 clases con 2366 puntos de control. El método que utilizaron fue el de la clasificación de máxima verosimilitud, con el que tuvieron resultados mucho menores que los que tuvo su otra clasificación, con una confiabilidad global de 68% y un índice Kappa de 0.60, esto pudo deberse a que utilizaron menos puntos de control y menos clases que en su otra clasificación. La confiabilidad global de su clasificación con máxima verosimilitud fue peor y su índice Kappa fue mejor que los resultados de la primera clasificación de la tesis.

Los dos métodos de clasificación realizados por Cordero-Sancho & Sader (2007) fueron los que obtuvieron los resultados de clasificación más bajos de la literatura citada. Su primera clasificación se hizo con el modelo *Coffee environmental stratification*, con una imagen Landsat 7 y 293 puntos de control para clasificar en 5 clases, ingresaron datos auxiliares como la elevación y la precipitación, sin embargo, el resultado de confiabilidad global fue de 65.19% y el índice Kappa de 0.56, es decir, fue peor que lo

obtenido en la primera clasificación de esta tesis. La segunda clasificación la hicieron con los mismos insumos, pero con el clasificador de máxima verosimilitud. Los resultados fueron más bajos que los anteriores, con una confiabilidad global de 63.14% y un índice Kappa de 0.53, por lo que fue también una peor clasificación que la hecha con el NDVI en esta tesis. Esto pudo ocurrir porque utilizaron únicamente una imagen y pocos puntos de control a diferencia de otros autores (Bolanos, 2007; Kawakubo & Pérez, 2016) que utilizaron una imagen, pero muchos puntos de control.

A pesar de todo lo anterior, la menos confiable de todas las clasificaciones fue la segunda que se hizo en esta tesis, pues con 16 imágenes, 19,000 puntos de control y el apoyo de una Serie de Uso de Suelo y Vegetación, la clasificación en dos clases utilizando el método *Random Forest* sólo logró obtener 43.6% de confiabilidad global y un índice Kappa de -0.005, es decir, que esta clasificación no es confiable y mucho menos concordante.

Para cerrar, hay que destacar que, a pesar de la mala clasificación obtenida con uno de los métodos utilizados, se logró un avance para discernir entre la vegetación natural y los cultivos de café que se encuentran en el estado de Chiapas. Además, se pudo observar la posible expansión

que este cultivo ha tenido dentro de Ocosingo desde el último padrón cafetalero, lo cual implica que lo obtenido en este trabajo tiene utilidad contextual del café en México, principalmente del café orgánico (como el Café Chiapas), que cumple con estándares y normas (Diario Oficial de la Federación, 2003) que han llevado a México a ser uno de los principales exportadores de este tipo de café a nivel mundial (International Coffee Organization, 2018). Con lo anterior, queda claro que mapear café de sombra a través de la clasificación de imágenes satelitales es un esfuerzo fundamental para la agricultura nacional.



CONCLUSIONES

El cultivo de café está posicionado como uno de los cultivos más importantes y una de las formas de sustento económico de cientos de familias en todo el mundo, por lo que tener un mejor conocimiento sobre su distribución en zonas como la estudiada en este trabajo es fundamental para tener un mejor control de la producción del grano sin necesidad de acudir al lugar para corroborar los datos. Por esta razón, el uso de técnicas de percepción remota para clasificar e identificar de manera confiable los lugares en los que existe café o cualquier otro cultivo, es una manera eficaz y poco costosa de lograrlo (Chuvienco, 2010), especialmente si se obtienen resultados de confiabilidad suficientemente buenos o aceptables.

En esta tesis se hicieron dos intentos de clasificar café y se obtuvieron resultados muy diferentes para cada uno de ellos. Los puntos de control de vegetación que se derivaron del NDVI fueron los que ayudaron a que una de las clasificaciones fuera relativamente buena, es decir, los valores de vegetación con más de 70% de actividad fotosintética resultaron ser más eficientes para clasificar vegetación que los polígonos de vegetación natural propuestos en la Serie VI de Uso de Suelo y Vegetación del INEGI.

Los resultados alcanzados fueron buenos y aceptables en la primera clasificación, por lo que

se puede decir que la hipótesis planteada para esta tesis se demostró parcialmente. Hay que tener en cuenta que las clasificaciones de café que se han hecho por los autores citados no sólo contemplan dos clases como en este trabajo (café y vegetación natural).

El método de clasificación *Random Forest* tuvo buen resultado y fue similar al obtenido con otros métodos; sin embargo, se encontraron varias maneras en las que podría mejorarse el resultado, una de ellas sería utilizar más clases de cobertura que no sean necesariamente de vegetación, sino que se tomen en cuenta otras como cuerpos de agua y zonas urbanas. También debe mencionarse que la clasificación pudo haber mejorado con el uso de datos auxiliares como lo hicieron Kelley et al. (2018) quienes así afinaron la votación final de los árboles de decisión del clasificador.

Este trabajo deja claro que utilizar únicamente los porcentajes de Vegetación Fotosintética (PV) no logra superar los resultados alcanzados con otros métodos en la segunda clasificación, pues la separabilidad espectral entre sus clases no muestra suficiente diferencia, como se observó en la primera clasificación. Entonces, el uso exclusivo de PV sólo sería suficiente para una buena clasificación si cada clase tuviera valores muy diferentes entre sí, y de esta manera se

podría evitar la confusión del clasificador que sucedió en la segunda clasificación.

Como se necesitaron dos escenas Landsat 8 para cubrir la zona de estudio, se observaron las diferencias en las imágenes de vegetación fotosintética. Aunque sus fechas de captura fueron lo más cercanas posibles, algunas escenas se encontraron con algunas nubes y esto ocasionó diferencias en los promedios del porcentaje de PV utilizado para encontrar la separabilidad espectral de las clases.

Por lo mencionado, se recomienda que las zonas de estudio se encuentren dentro del área cubierta por una sola imagen y que ésta tenga poco o nada de cubierta nubosa, o bien, que las imágenes que cubran la zona sean de la misma fecha de captura. Aparte de esto, utilizar imágenes con una mejor resolución que la que tienen las imágenes Landsat 8 podría mejorar la asignación de clases que hace el clasificador *Random Forest*.

Se considera que la clasificación mejoraría si los puntos de control fueran más precisos, es decir, que fueran recogidos en campo para tener mayor certeza de que sí pertenecen a la clase en la que se agruparon. Aunque esto significaría realizar trabajo de campo que implica más costos, valdría la pena abordar este análisis con la certeza de tener puntos de café que existen en la realidad.

Aunado a esto, los puntos de control deberían corresponder a las fechas de captura de las imágenes satelitales.

Además, sería necesario obtener puntos de verificación en campo para conocer si la clasificación realmente confundió las clases o si el cultivo de café ha ido expandiéndose en Ocosingo. Otro factor que pudo haber afectado la clasificación es que dentro de las parcelas del PNC se encuentren cultivos de café de sol y café de sombra combinados, pues sus valores espectrales son distintos y esto pudo ocasionar ruido en el clasificador.

Se propone que el uso del clasificador *Random Forest* se practique y se pruebe con distintas combinaciones de datos auxiliares para encontrar con cuáles se obtiene el resultado más exacto y confiable para la clasificación de café, pues demostró tener un buen resultado en la primera prueba de clasificación que se hizo en este trabajo.

Finalmente, hay que resaltar la labor de la geografía dentro del estudio del café, pues es a partir de esta disciplina que es posible entender de manera integral la dinámica en la que se desenvuelven los aspectos físicos y sociales que caracterizan a cada territorio y que dan pie a la caficultura que se desarrolla en el país.



REFERENCIAS

- Alberg, A., Park, JW., Hager, B., Brock, M. & Diener-West, M. (2004) *The Use of "Overall Accuracy" to Evaluate the Validity of Screening or Diagnostic Tests*. Blackwell Science Inc. Journal of general internal medicine. V:19 pp. 460-465.
- Aguirre, F. (1999). *Introducción al estudio de la cafeticultura*. Revista Vinculando. Disponible en: vinculando.org
- Aguirre, R. (2009). *Conceptos de geomática y estudios de caso en México*. Geografía para el Siglo XXI. UNAM: Instituto de Geografía.
- Araque, L. & Jiménez, A. (2009). *Caracterización de firma espectral a partir de sensores remotos para el manejo de sanidad vegetal en el cultivo de palma de aceite*. Palmas Vol.30 No. 3.
- Asner, G., Knapp, D., Balaji, A. & Páez-Acosta, G. (2009). *Automated mapping of tropical deforestation and forest degradation*. Journal of Applied Remote Sensing 3:033543.
- Asociación Mexicana de la Cadena Productiva del Café (AMECAFE). (2011). *Plan de Innovación en la cafeticultura de México*.
- Barrera, J., Parra, M., Herrera, O., Jarquín, R. & Pohlen, J. (2004). *Plan Estatal de Manejo Agroecológico del Café en Chiapas: Guía hacia una cafeticultura sustentable*. México: Colegio de la Frontera Sur & Comisión para el Desarrollo y Fomento del Café de Chiapas.
- Bartra, A. (2006). *Virtudes económicas, sociales y ambientales del café certificado. El caso de la coordinadora estatal de productores de café de Oaxaca*. En: B. Canabal, G. Contreras, & A. León, *Estrategias Económicas y Procesos Culturales* (pág. 436). México: Plaza y Valdés.
- Bertrand, C. & Bertrand, G. (2007). *Geografía del medio ambiente. El sistema GTP: Geosistema, territorio y paisaje*. Manuales de humanidades: Universidad de Granada.
- Bolanos, S. (2007). *Using Image Analysis and GIS for coffee Mapping*. (Tesis de maestría). McGill University, Montreal, Canadá.
- Bolemon, J. (1985). *Physics: An Introduction*. Englewood Cliffs, New Jersey: Prentice Hall.
- Breiman, L. (2001). *Random Forests*. Machine Learning, 45(1), 5-32.
- Campbell, J. B. & Wynne, R. H. (2011). *Introduction to remote sensing*. Guilford Press.

- Canabal, B., Contreras, G. & León, A. (2006). *Diversidad rural: estrategias económicas y procesos culturales*. México: Universidad Autónoma Metropolitana.
- Carlotto, M. (1999). *Reducing the effects of space-varying, wavelength-dependent scattering in multispectral imagery*. International Journal of Remote Sensing.
- Centro de Investigación, Capacitación, Asesoría y Promoción (CICAP). (2007). *Café Femenino. Experiencias de mujeres emprendedoras*. Chiclayo, Perú.
- Cerdeña, J. & Villarroel, L. (2008). *Evaluación de la concordancia inter-observador en investigación pediátrica: Coeficiente de Kappa*. Revista chilena de Pediatría. 79 (1): 54-58
- Chayanov, A. (1985). *Sobre la teoría de los sistemas económicos no capitalistas*. Bureau d'Etudes Cooperatives et Comunnautaires. Francia.
- Chicas, S., Omine, K. & Saqui, P. (2016). *CLASlite algorithms and social surveys to assess and identify deforestation and forest degradation in Toledo's protected areas and forest ecosystems, Belize*. Applied Geography 75, pp. 144-155.
- Chuvieco, E. (1995). *Fundamentos de teledetección espacial*. Rialp, S. A. Madrid, España.
- Chuvieco, E. (2008). *Teledetección ambiental. La observación de la Tierra desde el espacio*. España: Ariel Ciencia. 3era edición.
- Chuvieco, E. (2010). *Teledetección ambiental. La observación de la Tierra desde el espacio*. Barcelona: Ariel S. A.
- Chuvieco, E. & Huete, A. (2010). *Fundamentals of Satellite Remote Sensing: An Environmental Approach*. CRC press.
- Comisión Nacional de Áreas Naturales Protegidas (CONANP). (2016). *Sistema Nacional de Áreas Protegidas (SINAP)*. Publicado el 22/07/2016. Disponible en: www.gob.mx
- Comité Estatal de Información Estadística y Geográfica de Chiapas (CEIEG). (2018). *Información estadística y sociodemográfica: Ocosingo*. Gobierno del estado de Chiapas.
- Cordero-Sancho, S. & Sader, S. (2007). *Spectral analysis and classification accuracy of coffee crops using Landsat and a topographic-environmental model*. International Journal of Remote Sensing, 28:7, 1577-1593

- Cosero, D., Riaño, D. & Chuvieco, E. (1998). *Estimación de la Humedad de la Vegetación Usando Imágenes de Satélite e Índices Meteorológicos de Peligro de Incendios*. En: Serie Geográfica: Departamento de Geografía, Universidad de Alcalá, España. Vol. 7 59-72.
- Curran, P., Foody, G., Kondratyev, K., Kozoderov, W., & Fedchenko, P. (1990). *Remote Sensing of Soils and Vegetation in the USSR*. Londres: Taylor and Francis.
- Del Toro, N., Gomariz, F., Cánovas, F. & Alonso, F. (2015). *Comparación de métodos de clasificación de imágenes de satélite en la cuenca del río Argos (Región de Murcia)*. Boletín de la Asociación de Geógrafos Españoles N. 67, pp. 327-347.
- Diario Oficial de la Federación (DOF). (1981). *Norma Mexicana NMX-EE-120-1981, Envase - Papel - Bolsas para envasar café - Dimensiones*. Publicada el 09/04/1981. Disponible en: dof.gob.mx
- Diario Oficial de la Federación (DOF). (1984). *Norma Mexicana NMX-EE-178-1984, Envase - Textiles - Henequén - Sacos para envasar café - Especificaciones*. Publicada el 16/10/1984. Disponible en: dof.gob.mx
- Diario Oficial de la Federación (DOF). (1991). *Norma Mexicana NMX-B-231-1990, Cribas para clasificación de materiales granulares*. Publicada el 09/01/1991. Disponible en: dof.gob.mx
- Diario Oficial de la Federación (DOF). (1996). *Norma Mexicana NMX-F-551-1996-SCFI, Café verde - Especificaciones y Métodos de Prueba*. Publicada el 11/04/1996. Disponible en: dof.gob.mx
- Diario Oficial de la Federación (DOF). (1996). *Norma Mexicana NMX-F-083-SCFI-1996, Alimentos - Determinación de humedad en productos alimenticios*. Publicada el 14/07/1996. Disponible en: dof.gob.mx
- Diario Oficial de la Federación (DOF). (1997). *Norma Mexicana NMX-F-551-1996-SCFI, Café Verde - Especificaciones y Métodos de Prueba*. Publicada el 11/04/1997. Disponible en: dof.gob.mx
- Diario Oficial de la Federación (DOF). (1997). *Norma Oficial Mexicana NOM-037-FITO-1995, Por la que se establecen las especificaciones del proceso de producción y procesamiento de productos agrícolas orgánicos*. Publicada el 23/04/1997. Disponible en: dof.gob.mx
- Diario Oficial de la Federación (DOF). (2000). *Norma Mexicana NMX-F-013-SCFI-2000, Café puro tostado, en grano o molido sin descafeinar o descafeinado - Especificaciones y métodos de prueba*. Publicada el 18/08/2000. Disponible en: dof.gob.mx

- Diario Oficial de la Federación (DOF). (2001). *Norma Oficial Mexicana NOM-106-SCFI-2000, Características de diseño y condiciones de uso de la contraseña oficial*. Publicada el 02/02/2001. Disponible en: dof.gob.mx
- Diario Oficial de la Federación (DOF). (2003). *Declaración general de protección de la denominación de origen Café Chiapas*. Publicado el 27/08/2003. Disponible en: dof.gob.mx
- Diario Oficial de la Federación (DOF). (2006). *Norma Oficial Mexicana NOM-030-SCFI-2006, Información Comercial - Declaración de cantidad en la etiqueta - Especificaciones*. Publicada el 06/11/2006. Disponible en: dof.gob.mx
- Diario Oficial de la Federación (DOF). (2007). *Norma Oficial Mexicana NOM-169-SCFI-2007, Café Chiapas - Especificaciones y métodos de prueba*. Publicada el 26/03/2007. Disponible en: dof.gob.mx
- Diario Oficial de la Federación (DOF). (2008). *Norma Mexicana NMX-F-551-SCFI-2008, Café Verde - Especificaciones, preparaciones y evaluación sensorial*. Publicada el 18/02/2009. Disponible en: dof.gob.mx
- DiBiase, D. (2018). *The Nature of Geographic Information*. The Pennsylvania State University. Recuperado de: <https://www.e-education.psu.edu/natureofgeoinfo/>
- DigitalGlobe (2018). *Resources: Satellite Information*. Disponible en: <https://www.digitalglobe.com/resources/satellite-information>
- Donges, N. (2018). *The Random Forest Algorithm*. SAP Machine Learning Foundation, CODE University, Berlin.
- Elachi, C. & van Zyl, J. (2006). *Introduction to the physics and techniques of remote sensing*. Vol. 28. John Wiley & Sons.
- Environmental Systems Research Institute. (2016). *¿Qué son los datos ráster?* Esri.
- Escamilla, P. (1993). *El café cereza en México, tecnología de la producción*. Universidad Autónoma Chapingo, México: Grupo Esmeralda, 116 pp.
- Escamilla, E. & Zamarripa, A. (2002). *Variedades de café en México. Origen, características y perspectivas*. Universidad Autónoma Chapingo.
- Estes, J. & Simonett, D. (1975). *Fundamentals of image interpretation*. En R. G. Reeves (ed.). *Manual of Remote Sensing*. Falls Church, American Society of Photogrammetry. pp. 869-1076.

- Evans, D. & Lasaponara, R. (2012). *Uncovering Angkor: Integrated Remote Sensing Applications in the Archaeology of Early Cambodia*. En: Lasaponara, R. & Masini, N. (eds.) *Satellite remote sensing: A new tool for archaeology*. Vol. 16. Springer & Business Media.
- Fideicomisos Instituidos en Relación con la Agricultura (FIRA). (2016). *Panorama Agroalimentario Café 2016*. Dirección de Investigación y Evaluación Económica y Sectorial.
- Flores, F. (2015). *La producción de café en México: ventana de oportunidad para el sector agrícola de Chiapas*. Universidad Autónoma de Nuevo León. Espacio I+D, Innovación más Desarrollo, Vol. IV, No. 7.
- Franco, A., Carrasco, J., Sánchez, G. & Martínez, J. (2013). *Decision Tree based Classifiers for Large Datasets*. Computación y Sistemas Vo. 17 No. 1, pp. 95-102.
- Friedl, M. & Brodley, C. (1997). *Decision tree classification of land cover from remotely sensed data*. Remote Sensing of Environment, 61, pp. 399-409.
- Gaertner, J., Genovese, V., Potter, C., Sewake, K. & Manoukis, N. (2017). *Vegetation classification of Coffea on Hawaii using WorldView-2 satellite imagery*. Journal of Applied Remote Sensing.11:4, 046005
- García, E. (2004). *Modificaciones al sistema de clasificación climática de Köppen*. Instituto de Geografía, UNAM.
- Gill, T., Phinn, S., Armston, J. & Pailthorpe, B. (2009). *Estimating tree-cover change in Australia: challenges of using the MODIS vegetation index product*. International Journal of Remote Sensing. 30 (6), 1547–1565.
- Gómez, L., Gómez Cruz, M. & Schwentesius, R. (2001). *Desafíos de la agricultura orgánica*. México: Universidad Autónoma Chapingo.
- Gonzaga, C. (2014). *Aplicación de Índices de Vegetación Derivados de Imágenes Satelitales Landsat 7 ETM+ y ASTER para la Caracterización de la Cobertura Vegetal en la Zona Centro de la Provincia De Loja, Ecuador*. Universidad Nacional de La Plata. Argentina.
- González, P. (2004). *El café en México*. México desconocido: México.
- González, C. (2012). *Procesamiento a bordo de imágenes hiperespectrales de la superficie terrestre mediante hardware reconfigurable*. Madrid.
- González, A., Calera, A. & Quintanilla, A. (1999). *Separabilidad y evolución multitemporal de la respuesta espectral e índices de vegetación en masas forestales mediterráneas*. España: Instituto de Desarrollo Regional.

- Grupo Técnico Interdisciplinario de Café. (2017). *Manual para la producción de semilla certificada de café en México. Plan Integral de Atención al Café*. 36 pp.
- H. Ayuntamiento de Ocosingo, Chiapas. (2012). *Plan de Desarrollo Municipal 2012-2015*.
- Hansen, M., Dubayak, R. & Defries, R. (1996). *Classification trees: an alternative to traditional land cover classifiers*. International Journal of Remote Sensing, 17, pp.1075-1081.
- Holben, B. (1986). *Characteristics of maximum value composite images from temporal AVHRR data*. International Journal of Remote Sensing, 7, pp. 1417-1434.
- Huete, A. (1988). *A soil-adjusted vegetation index (SAVI)*. Remote Sensing of Environment, vol. 25, issue 3, pp. 259-309.
- Instituto del Café de Chiapas (INCAFECH). (2018). *Programa de apoyos a pequeños productores componente PROCAFÉ e impulso productivo al café*. Gobierno del estado de Chiapas.
- Instituto Mexicano de la Propiedad Industrial (IMPI). (2018). *México ya tiene 16 Denominaciones de Origen*. Publicado el 19/03/2018. Disponible en: www.gob.mx
- Instituto Nacional de Estadística y Geografía (INEGI). (1984). *Continuo Nacional del Conjunto de Datos Geográficos de la Carta Geológica E 15-12 Chiapas Escala 1:250 000, Serie I*.
- Instituto Nacional de Estadística y Geografía (INEGI). (2000). *Modelos Digitales de Elevación Las Margaritas. E 15-12 Chiapas Escala 1:250 000*.
- Instituto Nacional de Estadística y Geografía (INEGI). (2007). *Continuo Nacional del Conjunto de Datos Vectoriales Edafológicos Las Margaritas E 15-12 Chiapas Escala 1:250 000, Serie II*.
- Instituto Nacional de Estadística y Geografía (INEGI). (2014). *Conjunto de Datos Vectoriales de Uso de Suelo y Vegetación Serie VI*.
- Instituto Nacional de Estadística y Geografía (INEGI). (2016). *Elementos de percepción remota*. Disponible en: inegi.org.mx
- Instituto Nacional de Estadística y Geografía (INEGI). (2017). *Anuario estadístico y geográfico de Chiapas 2017*. Gobierno del estado de Chiapas.

- Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias (INIFAP). (2013). *El sistema producto café en México: problemática y tecnología de producción*. Veracruz, México. 462 pp.
- Instituto Nacional para el Federalismo y el Desarrollo Municipal (INAFED). (2018). *Enciclopedia de los Municipios y Delegaciones de México. Estado de Chiapas: Ocosingo*. Disponible en: www.inafed.gob.mx
- International Coffee Organization (ICO). (2018). *Historical Data on the Global Coffee Trade*. Disponible en: www.ico.org
- Jensen, J. (2000). *Remote Sensing of the Environment. An Earth Resource Perspective*. Upper Saddle River, New Jersey: Prentice Hall.
- Kawakubo, F. & Pérez, R. (2016). *Mapping coffee crops in southeastern Brazil using spectral mixture analysis and data mining classification*. International Journal of Remote Sensing, 37:14, 3414-3436.
- Kelley, L., Pitcher, L. & Bacon, C. (2018). *Using Google Earth Engine to Map Complex Shade-Grown Coffee Landscapes in Northern Nicaragua*. Remote Sensing, 10, 952, doi:10.3390/rs10060952
- Keshava, N. (2003). *A Survey of Spectral Unmixing Algorithms*. Lincoln Laboratory Journal Vol. 14 No. 1.
- Keshava, N. & Mustard, J. (2002). *Spectral Unmixing*. IEEE Signal Processing Magazine 19, N. 1.
- Khorram, S., Koch, F., van der Wiele, C. & Nelson, S. (2012). *Remote sensing*. Springer. Science & Business Media.
- Kidder, S. & Haar, T. (1995). *Satellite meteorology: an introduction*. Gulf Professional Publishing.
- Kulkarni, A. & Lowe, B. (2016). *Random Forest Algorithm for Land Cover Classification*. Computer Science Faculty Publications and Presentations.
- Levin, N. (1999). *Fundamentals of remote sensing*. 1st Hydrographic Data Management Course, International Maritime Academy, Trieste, 76.
- Lillesand, T., Kiefer, R. & Chipman, J. (2008). *Remote sensing and image interpretation*. Sixth Edition, New York: John Wiley & Sons.
- Lira, J. (2009). *La percepción remota: nuestros ojos desde el espacio*. México: Fondo de Cultura Económica.
- López de Ullibarri, I. & Pita, S. (2001). *Medidas de concordancia: el índice de Kappa*. Cad Aten Primaria 6, pp. 169-171.

- Lusch, R. (1989). *Spectral characteristics of land surfaces*. En M. D. Nellis, R. Lougeay & K. Lula (eds.), *Current Trends in Remote Sensing Education*, Hong-Kong, Geocarto International Centre, pp. 81-86.
- Mariscal, A. (2011). *El café orgánico de Chiapas crece a contracorriente y sin incentivo*. Disponible en: <https://expansion.mx/nacional/2011/01/25/el-cafe-organico-de-chiapas-crece-a-contracorriente-y-sin-incentivos>
- Martínez-Verduzco, G. C., Galeana-Pizaña, J. M., & Cruz-Bello, G. M. (2011). *Coupling Community Mapping and supervised classification to discriminate Shade coffee from Natural vegetation*. *Applied Geography*, 34, 1-9. doi: 10.1016/j.apgeog.2011.10.001
- Mas, J., Pérez-Vega, A., Ghilardi, A., Martínez, S., Loya-Carrillo, J. & Vega, E. (2014). *A Suite of Tools for Assessing Thematic Map Accuracy*. *Geography Journal*. Volume 2014, Article ID 372349, 10 pages <http://dx.doi.org/10.1155/2014/372349>
- Máximo, A. & Alvarenga, B. (2006). *Física general*. México: Oxford.
- Meyfroid, P., Tan, V. & Viet, H. (2013). *Trajectories of deforestation, coffee expansion and displacement of shifting cultivation in the Central Highlands of Vietnam*. *Global Environmental Change*. <http://dx.doi.org/10.1016/j.gloenvcha.2013.04.005>
- Ming, D., Zhou, T., Wang, M. & Tan, T. (2016). *Land cover classification using random forest with genetic algorithm-based parameter optimization*. *Journal of Applied Remote Sensing* 10(3). DOI: 10.1117/1.JRS.10.035021
- Moguel, P. & Toledo, V. (1996). *El café en México, ecología, cultura indígena y sustentabilidad*. *Ciencias*, 43: 40-51.
- Monroy, M. (2004). *¿Qué hay detrás de una taza de café? Datos básicos sobre café en México*. *Revista sobre Desarrollo sustentable, Vinculando*.
- Moran, J. & Morgan, M. (1997). *Meteorology: The Atmosphere and the Science of Weather*. Prentice Hall.
- Moreira, M., Adami, M., Rudorff, B. (2004). *Análise espectral e temporal da cultura do café em imagens Landsat*. *Pesquisa Agropecuária Brasileira*, 39 (3), 223-231
- Muñoz Bollas, A. (2011). *Geodesia y Cartografía*. En: Pérez Navarro, A. (Coord) *Introducción a los sistemas de información geográfica y geotelemática*. Editorial UOC.

- Muñoz, P. (2013). *Apuntes de teledetección: índices de vegetación*. Centro de Información de Recursos Naturales. Chile.
- Olivar, A. & Bustamante F. (2016). *Comprendiendo la situación de los trabajadores del café en haciendas y pequeñas fincas familiares*. Specialty Coffee Association of America. Solidaridad Network.
- Oniga, V., Breaban, A. & Statescu, F. (2018). *Determining the Optimum Number of Ground Control Points for Obtaining High Precision Results Based on UAS Images*. Proceedings, 2, 352, doi:10.3390/ecrs-2-05165
- Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO). (2013). FAOSTAT. Commodity balances. Coffee and products.
- Ortega-Huerta, M., Komar, O., Price, K., & Ventura, H. (2012). *Mapping coffee plantations with Landsat imagery: an example from El Salvador*. *International Journal of Remote Sensing*. 33. 220-242. 10.1080/01431161.2011.591442.
- Paine, D. & Kiser, J. (2012). *Aerial photography and image interpretation*. Nueva Jersey, Estados Unidos de América: John Wiley & Sons, Inc.
- Phiri, D. & Morgenroth, J. (2017). *Developments in Landsat Land Cover Classification Methods: A Review*. *Remote Sens*. 2017, 9, 967; doi:10.3390/rs9090967
- Refaeilzadeh, P., Tang, L. & Liu, H. (2009). *Cross-Validation*. En: Liu, I. & Özsu, M. (eds.) *Encyclopedia of Database Systems*. Boston: Springer.
- Reis, P. (2015). *Error and uncertainty in the accuracy assessment of land cover maps*. NOVA Information Management School.
- Renard, M. (1993). *El Soconusco, una economía cafetalera*. México: Universidad Autónoma Chapingo.
- Renard, M. (2002). *El café en la frontera sur: vocación y destino*. En; Pohlen Jurgen (coord.), *México y la cafecultura chiapaneca*. Alemania: Shaker Verlag.
- Richards, J. (2013). *Remote Sensing Digital Image Analysis: An Introduction*. Fifth Edition. Berlin: Springer-Verlag.
- Richardson, A. & Wiegand, C. (1977). *Distinguishing vegetation from soil background information*. *Photogramm. Eng.* 43:1541-1552.

- Rodríguez, F. (1998). *La organización social de los productores de café*. INNOVAR, revista de ciencias administrativas y sociales. No. 12, julio - diciembre.
- Rodríguez-Galiano, V., Ghimire, B., Rogan, J., Chica-Olmo, M. & Rigol-Sánchez, J. (2011). *An assessment of the effectiveness of a random forest classifier for land-cover classification*. ISPRS Journal of Photogrammetry and Remote Sensing 67, 93-104. doi:10.1016/j.isprsjprs.2011.11.002
- Rouse, J., Haas, R., Scheel, J. & Deering, D. (1974). *Monitoring Vegetation Systems in the Great Plains with ERTS*. Proceedings, 3rd Earth Resource Technology Satellite (ERTS) Symposium, vol. 1, p. 48-62.
- Ruiz, L. (2018). *Impacto del comercio justo en la calidad de vida de los pequeños productores de café orgánico en México y en Perú*. Universidad Nacional Autónoma de México.
- Rzedowski, j. (2006). *Vegetación de México*. 1ra. Edición digital, Comisión Nacional para el Conocimiento y Uso de la Biodiversidad. México.
- Sagan, C. (1994). *A Pale Blue Dot*. Nueva York: Random House.
- Sánchez, J. (2016). *Análisis de Calidad Cartográfica mediante el estudio de la Matriz de Confusión*. Pensamiento Matemático, Vol VI, Número 2, pp. 9-26.
- Sánchez, S. & Plaza, A. (2010). *Implementación Paralela del Algoritmo Pixel Purity Index para Análisis Hiperespectral en GPUs*. Congreso español de informática.
- Sánchez, E., Torres, M., Palacios, A., Aguilar, M., Pino, S. & Granada, L.. (2000). *Comparación del NDVI con el PVI y el SAVI como Indicadores para la Asignación de Modelos de Combustible para la Estimación del Riesgo de Incendios en Andalucía*. Tecnologías Geográficas para el Desarrollo Sostenible. Universidad de Alcalá. 164-174.
- Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación (SAGARPA). (2010). *Respalda SAGARPA certificación de café orgánico en Chiapas*.
- Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación (SAGARPA). (2015). *¿Qué es la caficultura?* Publicado el 06/07/2015. Disponible en: www.gob.mx
- Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación (SAGARPA). (2018). *México, onceavo país productor mundial de café*. Publicado el 02/03/2018. Disponible en: www.gob.mx

- Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación (SAGARPA) & Servicio de Información Agroalimentaria y Pesquera (SIAP). (2014). *Padrón Nacional Cafetalero*. México.
- Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación (SAGARPA) & Servicio de Información Agroalimentaria y Pesquera (SIAP). (2017). *Atlas agroalimentario 2017*. México.
- Secretaría de Economía (SE). (2015). *Denominaciones de Origen #OrgulloDeMéxico*. Publicado el 14/09/2015. Disponible en: www.gob.mx
- Sedorovich, A., O'Hara, C. & Schukman, K. (2010). *Image accuracy checkpoint collection guide*. Penn State University.
- Servicio de Información Agroalimentaria y Pesquera (SIAP). (2014). *Sistemas Producto Café*.
- Servicio de Información Agroalimentaria y Pesquera (SIAP). (2016). *Café cereza*. Disponible en: www.gob.mx
- Servicio de Información Agroalimentaria y Pesquera (SIAP). (2017). *Avance de Siembras y Cosechas. Resumen nacional por estado*. Disponible en: Infosiap.siap.gob.mx
- Servicio Nacional de Sanidad, Inocuidad y Calidad Agroalimentaria (SENASICA). (2017). *Broca del café*. Publicado el 09/06/2017. Disponible en: www.gob.mx
- Silván, J. & López, P. (2012). *Clasificación subpixel desde una perspectiva ontológica*. GEOcibernética: Innovando en Geomática para la Sociedad. Año 1, Número 1, Artículo 5.
- Sobrino, J. (2001). *Teledetección*. Universidad de Valencia.
- Tian, S., Zhang, X., Tian, J. & Sun, Q. (2016). *Random Forest classification of wetland landcovers from Multi-sensor data in the arid region of Xinjiang, China*. Remote Sensing, 8(11), 954; doi:10.3390/rs8110954
- Toriello, J. (2014). *Las crisis en las Organizaciones Campesinas de productoras de café que participan en el Mercado Solidario (Comercio Justo). El caso de dos organizaciones campesinas en las regiones Sierra y Soconusco del estado de Chiapas*. Universidad Nacional Autónoma de México.
- Townshend, J. & Justice, C. (1981). *Information extraction from remotely sensed data. A user view*. International Journal of Remote Sensing, 2, pp. 313-329.
- Tristán, P., Wainschenker, R. & Doom, J. (2008). *Normalización de imágenes satelitales en el análisis multi-temporal*. En: *IX Workshop de Investigadores en Ciencias de la Computación*.

- United States Geological Survey (USGS). (2013). *Landsat 7 Science Data Users Handbook*. Disponible en: <https://landsat.usgs.gov/landsat-7-data-users-handbook>
- United States Geological Survey (USGS). (2018). *Landsat Missions*. Disponible en: <https://landsat.usgs.gov/landsat-missions>
- Universidad Autónoma Chapingo. (2015). *Líneas de estrategia para el impulso de la cafeticultura mexicana*. CENIDERCAFÉ, CENACAFÉ, Universidad Autónoma Chapingo. Huatusco, Veracruz.
- Universidad Tecnológica de la Selva. (2007). *Programa de reforestación en la región Selva Lacandona del estado de Chiapas*. Gobierno del estado de Chiapas.
- Vargas, P. (2007). *Mujeres cafetaleras y producción de café orgánico en Chiapas*. Universidad Autónoma Metropolitana Azcapotzalco. El cotidiano, vol. 22. Núm. 142, marzo-abril 2007, pp. 74-83.
- Wallace, J. & Hobbs, P. (2006). *Atmospheric science: an introductory survey*. Vol. 92. Academic press.
- Wen, Z., Lees, B. Jiao, F., Wanning, L. & Shi, H. (2010). *Stratified vegetation cover index: A new way to assess vegetation impact on soil erosion*. China: Catena. 83. 87-93.
- Yengoh, G., Deni, D., Olsson, L., Tengberg, A. & Tucker, C. (2015). *Use of the Normalized Difference Vegetation Index (NDVI) to Assess Land Degradation at Multiple Scales: Current Status, Future Trends and Practical Considerations*. Springer.
- Zha, Y., Gao, J. & Ni, S. (2003). *Use of normalized difference built-up index in automatically mapping urban areas from TM imagery*. International Journal of Remote Sensing. 24(3), 583-594.