



Universidad de San Carlos de Guatemala
Facultad de Ingeniería
Escuela de Estudios de Postgrado
Maestría en Ciencias en Geomática

**ANÁLISIS COMPARATIVO DE DOS MÉTODOS DE CLASIFICACIÓN DE
IMÁGENES PARA EL MAPEO FORESTAL EN GUATEMALA**

Ing. Edson Toniño Hernandez Mazariegos

Asesorado por la Mtra. Sara Michelle Catalán Armas

Guatemala, mayo de 2019

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA



FACULTAD DE INGENIERÍA

**ANÁLISIS COMPARATIVO DE DOS MÉTODOS DE CLASIFICACIÓN DE
IMÁGENES PARA EL MAPEO FORESTAL EN GUATEMALA**

TRABAJO DE GRADUACIÓN
PRESENTADO A LA JUNTA DIRECTIVA DE LA
FACULTAD DE INGENIERÍA
POR

ING. EDSON TONIÑO HERNANDEZ MAZARIEGOS
ASESORADO POR LA MTRA. SARA MICHELLE CATALÁN ARMAS

AL CONFERÍRSELE EL TÍTULO DE

MAESTRO EN CIENCIAS EN GEOMÁTICA

GUATEMALA, MAYO DE 2019

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA
FACULTAD DE INGENIERÍA



NÓMINA DE JUNTA DIRECTIVA

DECANO	Ing. Pedro Antonio Aguilar Polanco
VOCAL I	Ing. Jose Francisco Gómez Rivera
VOCAL II	Ing. Mario Renato Escobedo Martínez
VOCAL III	Ing. José Milton De León Bran
VOCAL IV	Br. Luis Diego Aguilar Ralón
VOCAL V	Br. Cristian Daniel Estrada Santizo
SECRETARIA	Inga. Lesbia Magali Herrera López

TRIBUNAL EVALUADOR QUE PRACTICÓ EL EXAMEN DE DEFENSA

DECANO	Mtro. Ing. Pedro Antonio Aguilar Polanco
DIRECTOR	Mtro. Ing. Edgar Darío Álvarez Cotí
COORDINADOR	Mtro. Ing. Marlon Antonio Pérez Turk
EXAMINADOR	Mtro. Ing. Hugo Humberto Rivera Pérez
SECRETARIA	Mtra. Inga. Lesbia Magali Herrera López

En mi calidad como Decano de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, luego de conocer la aprobación por parte del Director de la Escuela de Estudios de Postgrado, al Trabajo de Tesis de la Maestría en Ciencias en Geomática titulado: "ANÁLISIS COMPARATIVO DE DOS MÉTODOS DE CLASIFICACIÓN DE IMÁGENES PARA EL MAPEO FORESTAL EN GUATEMALA " presentado por el Ingeniero Agrónomo en Recursos Naturales Renovables Edson Toniño Hernández Mazariegos quien se identifica con Carné 200816770, procedo a la autorización para la impresión del mismo.

IMPRÍMASE.

"Id y Enseñad a Todos"

20/10/2019
Maestro. Ing. Pedro Antonio Aguilar Polanco
 Decano
 Facultad de Ingeniería
 Universidad de San Carlos de Guatemala

Guatemala, mayo de 2019.

Cc archivo/LZLA.

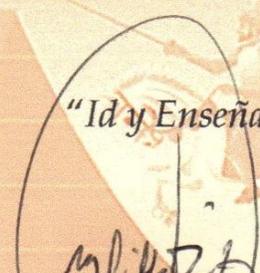


Ref.APT-2019-010

Como Coordinador de la Maestría en Ciencias en Geomática doy el aval correspondiente para la aprobación del Trabajo de Tesis titulado "ANÁLISIS COMPARATIVO DE DOS MÉTODOS DE CLASIFICACIÓN DE IMÁGENES PARA EL MAPEO FORESTAL EN GUATEMALA" presentado por el Ingeniero Agrónomo en Recursos Naturales Renovables Edson Toniño Hernández Mazariegos quien se identifica con Carné 200816770.

Atentamente,

"Id y Enseñad a Todos"


Maestro. Ing. Marlón Antonio Pérez Türk
Coordinador(a) de Maestría
Escuela de Estudios de Postgrado
Facultad de Ingeniería
Universidad de San Carlos de Guatemala



Guatemala, mayo de 2019.

Cc: archivo/L.Z.L.A.

Ref.APT-2019-010

En mi calidad como Asesora del Ingeniero Agrónomo en Recursos Naturales Renovables **Edson Toniño Hernández Mazariegos** quien se identifica con carné No. **200816770** procedo a dar el aval correspondiente para la aprobación del Trabajo de Tesis titulado **"ANÁLISIS COMPARATIVO DE DOS MÉTODOS DE CLASIFICACIÓN DE IMÁGENES PARA EL MAPEO FORESTAL EN GUATEMALA"** quien se encuentra en el programa de Maestría en Ciencias en Geomática en la Escuela de Estudios de Postgrado de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala.

Atentamente,

"Id y Enseñad a Todos"


Maestra. Inga. Sara Michelle Catalán Armas
Asesor(a)

Guatemala, mayo de 2019.

Cc: archivo/LZLA.

HONORABLE TRIBUNAL EXAMINADOR

En cumplimiento con los preceptos que establece la ley de la Universidad de San Carlos de Guatemala, presento a su consideración mi trabajo de graduación titulado:

ANÁLISIS COMPARATIVO DE DOS MÉTODOS DE CLASIFICACIÓN DE IMÁGENES PARA EL MAPEO FORESTAL EN GUATEMALA

Tema que me fuera aprobado por la Dirección de la Escuela de Estudios de Posgrado de la Facultad de Ingeniería, con fecha 24 de octubre de 2016.

Ing. Edson Toniño Hernandez Mazariegos

ACTO QUE DEDICO A:

- DIOS:** Mi fortaleza, de donde nace la voluntad de los triunfos, las lecciones y las bendiciones.
- MI MADRE:** Angélica Mazariegos, mi primer gran amor, presente en mis triunfos y también en mis derrotas, en quien ciegamente confío.
- MI PADRE:** Vicente Hernandez, mi padre y también el hombre a quien más admiro, un ejemplo de lucha, responsabilidad y superación.
- MI HIJO:** Jose Raul, quien ha llenado de alegría nuestras vidas, de quien hemos recibido las mejores lecciones de amor, esfuerzo y valentía. De donde nace mi orgullo de ser padre.
- MIS HERMANAS:** Nataly, Wendy y Beatriz, gracias por todo el apoyo incondicional, por ser parte importante de la familia y ser mis primeras amigas.

MI FAMILIA: Emma, Julio, abuelos y a toda mi familia, porque este triunfo es de todos, por ser mi inspiración.

MI AMADA: Karen Martínez, mi novia, mejor amiga, mi compañera, a quien admiro, respeto y amo infinitamente.

MIS AMIGOS: Ricardo, Mauro, Álvaro, Gabriela, Marielos y María Alejandra, amigos y profesionales de quienes frecuentemente aprendo.

ÍNDICE GENERAL

ÍNDICE ILUSTRACIONES.....	V
LISTA DE SÍMBOLOS.....	VII
GLOSARIO.....	IX
RESUMEN.....	XIII
PLANTEAMIENTO DEL PROBLEMA.....	XV
OBJETIVOS.....	XVII
RESUMEN MARCO METODOLÓGICO.....	XIX
INTRODUCCIÓN.....	XXI
1. MARCO TEÓRICO.....	1
1.1. Sensores remotos.....	1
1.1.1. Imagen de satélite.....	2
1.1.2. Pixel.....	3
1.2. Sistemas de Información Geográfica (SIG).....	4
1.3. Clasificación de imágenes.....	4
1.3.1. Clasificación de imágenes basada en píxeles.....	5
1.4. CLASlite.....	12
1.4.1. Vegetación fotosintética (VF).....	13
1.4.2. Vegetación no fotosintética (VNF).....	13
1.4.3. Superficie descubierta (S).....	14
1.5. Métodos de evaluación de mapas de cobertura forestal en Guatemala.....	14
1.5.1. Exactitud temática.....	15

1.6.	Antecedentes de la clasificación de imágenes en Guatemala	20
2.	APLICACIÓN DE LOS DOS MÉTODOS DE CLASIFICACIÓN DE IMÁGENES	25
2.1.	Área de estudio	25
2.2.	Descarga de imágenes	29
2.3.	Generación de composiciones de banda	31
2.4.	Áreas o muestras de entrenamiento	32
2.5.	Clasificación supervisada de la imagen “LC80210492016034LGN00”	33
2.6.	Clasificación CLASlite de la imagen “LC80210492016034LGN00”	34
3.	PRESENTACIÓN DE RESULTADOS	35
3.1.	Identificación de categorías de tierras forestales (bosque) y tierras agrícolas (no bosque)	35
3.1.1.	Clasificación supervisada	35
3.1.2.	Clasificación CLASlite	37
4.	DISCUSIÓN DE RESULTADOS	41
4.1.	Determinación de correspondencia espacial y cuantitativa de los resultados de la clasificación supervisada y clasificación con CLASlite	43
4.2.	Análisis de correspondencia espacial y cuantitativa de los resultados de la clasificación supervisada y clasificación con CLASlite	45

4.3.	Determinación de la exactitud de los resultados de la clasificación supervisada y clasificación con CLASlite	47
4.4.	Análisis de la determinación de la exactitud de los resultados de la clasificación supervisada y clasificación con CLASlite	49
CONCLUSIONES.....		53
RECOMENDACIONES.....		55
REFERENCIAS BIBLIOGRÁFICAS.....		57
ANEXOS.....		67

ÍNDICE ILUSTRACIONES

FIGURAS

1. Clasificación por mínima distancia	8
2. Clasificación por paralelepípedos	10
3. Clasificación por máxima probabilidad	12
4. Grilla Landsat para Guatemala.	26
5. Mapa del área de estudio Occidente.....	28
6. Buscador de imágenes EarthExplorer.....	29
7. Imagen LC80210492016034LGN00 combinación de bandas 5, 6 y 7... 31	
8. Distribución espacial de categorías de cobertura y uso de la tierra por el método de clasificación supervisada.....	36
9. Distribución espacial de categorías de cobertura y uso de la tierra por el método de clasificación CLASlite	38
10. Correspondencia espacial de dos métodos de clasificación	44

TABLAS

I.	Valoración del coeficiente kappa.	19
II.	Cálculo de tasa de deforestación dinámica forestal 2006-2010	27
III.	Metadatos escena Landsat "LC80210492014028LGN00"	30
IV.	Superficie de categoría resultado de la clasificación supervisada	37
V.	Superficie de categoría resultado de la clasificación CLASlite	39
VI.	Análisis de superficies de bosque y no bosque en otros proyectos....	42
VII.	Datos de superficie de la correspondencia espacial de los dos métodos de clasificación.....	45
VIII.	Evaluación de exactitud del método CLASlite	49
IX.	Evaluación de exactitud del método clasificación supervisada.....	49

LISTA DE SÍMBOLOS

Significado	Símbolos
Ha	Hectáreas
km ²	Kilómetros cuadrados
®	Marca registrada
%	Porcentajes

GLOSARIO

Algoritmo	Conjunto de instrucciones que ejecuta un ordenador para solventar un problema.
Bosque	Ecosistema en donde los árboles son las especies vegetales dominantes y permanentes.
Categorías de cobertura vegetal y uso de la tierra	Segmentación de los tipos de utilización agrícola, pecuaria y forestal que muestran la posibilidad de ser establecidos en el terreno.
Clasificación de imágenes	Proceso en el cual se asignan a las imágenes de satélite categorías de cobertura vegetal, uso del suelo, geología u otras características del terreno, a través de técnicas de teledetección, la cual se apoyó de algoritmos computacionales.
Cobertura vegetal	Manto vegetal de un territorio dado.
Deforestación	Tierras forestales convertidas a tierras agrícolas o no bosque.

Degradación	Diminución progresiva de la biomasa en los bosques sin llegar a convertirse en deforestación.
Imágenes de satélite	Relación que existe entre una posición específica, es decir, coordenadas geográficas o proyectadas en relación a un valor que capta el sensor en el momento de la toma de la imagen, este valor responde a la relación que existe entre la capacidad de un objeto de reflejar o absorber luz.
Landsat	Serie de satélites de observación de la tierra lanzados por el gobierno de los Estados Unidos de Norteamérica.
Sistemas de monitoreo forestal	Instrumento nacional que permite a los países evaluar una amplia gama de información sobre los bosques.
Tierras Forestales	Categoría de toda la tierra con vegetación boscosa coherente con los umbrales utilizados para definir las tierras forestales en el inventario nacional de gases de efecto invernadero. También incluye los sistemas con una estructura de vegetación que actualmente se encuentra por debajo, pero que potencialmente podría alcanzar

in situ los valores umbrales utilizados por un país, para definir la categoría de tierras forestales.

Tierras agrícolas

Tierra cultivada, incluidos los arrozales y los sistemas de agro-silvicultura donde la estructura de la vegetación se encuentra por debajo de los umbrales utilizados, para la categoría de tierras forestales.

Uso de la tierra

Categoría resultante de la acción del hombre sobre la cobertura terrestre producto de la interacción con el ecosistema como espacio de subsistencia.

RESUMEN

En los últimos años, los países de la región han optado por el establecimiento de sistemas de monitoreo forestal y hacen uso de los sensores remotos para la obtención de las imágenes de satélite, a través de algoritmos y programas computacionales logran determinar los cambios de cobertura forestal. Estos sistemas generan reportes de datos en períodos cortos de tiempo, lo cual representa un gran reto para los países que siguen utilizando metodologías convencionales para el monitoreo del recurso bosque. El reto para los países de la región es la selección y adaptación de uno de los diversos métodos que existen en la actualidad y que a la vez el método seleccionado cumpla con los estándares internacionales en términos de exactitud y este dentro de los límites de aplicabilidad del país.

Partiendo de la situación actual, se realizó el análisis comparativo de dos métodos de clasificación de imágenes, en el que se hizo uso del software Erdas Imagine para el método de clasificación supervisada (método convencional) utilizando el algoritmo de máxima verosimilitud; y el software CLASlite (método moderno) (Asner, 2014). Los resultados obtenidos se analizaron a partir de una matriz de confusión, donde se determinó un porcentaje de correspondencia espacial y los valores de exactitud. Para cada método, se contrastó con el índice Kappa, el cual permite identificar cuál de los dos métodos tiene mayor exactitud respecto a los datos de referencia obtenidos en campo.

El método de clasificación con CLASlite resulto en un rango “casi perfecto” (88 %) siendo este porcentaje de exactitud superior al método de clasificación supervisada con un rango “considerable” (80 %). Así mismo, el método CLASlite presenta mayor exactitud en la categorización de bosque (84 %) y un menor porcentaje de error (16 %) respecto al método de clasificación supervisada (79 % y 21 % respectivamente).

En este documento se presenta la investigación realizado en parte de la región de occidente de Guatemala, los resultados, conclusiones y recomendaciones que surgieron del proceso de investigación desarrollado.

PLANTEAMIENTO DEL PROBLEMA

En países de la región centroamericana, como es el caso de Guatemala, aún se hace uso del método de clasificación supervisada de imágenes para la obtención de categorías de cobertura y uso de la tierra; este método carece de actualización y poco análisis respecto a la evaluación de exactitud temática, esto conlleva a un elevado costo para su ejecución y la ocupación de un número considerable de profesionales. Esta problemática se debe a que no se han realizados estudios de las nuevas tecnologías que existen para la clasificación de imágenes de satélite, lo cual permitiría conocer el contexto de estos nuevos métodos y así determinar su exactitud y la viabilidad de su aplicación en el país.

Actualmente existe un creciente número de aplicaciones y usuarios de información geoespacial que hace necesario la exploración de nuevos métodos de clasificación de imágenes que optimicen el uso de los datos de entrada y mejoren la exactitud de sus resultados, algunos métodos como el de clasificación supervisada y no supervisada han sido ya ampliamente estudiados, pero el surgimientos de nuevos métodos, utilizando complejos y potentes algoritmos, requiere de análisis de sus resultados a partir de su aplicabilidad.

Considerando que los países de la región están optando por el establecimiento de sistemas de monitoreo forestal que permitan la estimación de la cobertura forestal y la dinámica de la misma (regeneración y deforestación), en períodos de tiempo cortos con resultados confiables y precisiones aceptables.

Estos sistemas de monitoreo deben de ser constantes y comparables para el reporte de diferentes métricas ambientales, que contribuirán a la conservación y mejor aprovechamiento de los bosques y que esto conlleve a reducir los efectos del cambio climático.

Como se mencionó en párrafos anteriores, en diversos países de la región aún se están utilizando métodos convencionales para la clasificación de imágenes lo que conlleva al poco aprovechamiento que se le puede dar a los nuevos recursos geospaciales, y el desconocimiento de nuevos métodos respecto a su aplicabilidad, eficiencia y exactitud en la clasificación de imágenes, por lo que con base a esta problemática, se plantean los siguientes cuestionamientos:

¿La aplicación del programa para la clasificación de imágenes CLASlite generará mejores resultados respecto a los métodos convencionales actualmente?

¿Existe correspondencia espacial entre los resultados obtenidos de la clasificación por el método CLASlite y el método de clasificación supervisada?

¿Cuál de los métodos evaluados (CLASlite y clasificación supervisada) tiene mayor exactitud en relación al método de validación?

OBJETIVOS

General

Evaluar el método de clasificación supervisada y el método de clasificación CLASlite en la identificación de categorías de tierras forestales (bosque) y tierras agrícolas (no bosque).

Específicos

1. Analizar si existe correspondencia de los resultados obtenidos entre los métodos de clasificación supervisada con el método de clasificación CLASlite.
2. Determinar cuál de las metodologías presenta mejores resultados de exactitud, mediante la evaluación de exactitud temática: matriz de error y estadísticos asociados.

RESUMEN MARCO METODOLÓGICO

En este apartado se hará una breve descripción metodológica a realizar, para la consecución de los objetivos de la tesis.

Se ha dividido en secciones de tal forma que en la primera se describe las operaciones previas para la generación de las áreas de entrenamiento que posteriormente se utilizaron para realizar la clasificación de las imágenes en dos métodos distintos. A continuación, se describe el procedimiento llevado a cabo en cada uno de los métodos de clasificación, máxima verosimilitud y CLASlite, y finalmente se describen las técnicas de control de error de cada uno de los métodos de clasificación que a su vez, sirve para comparar las aptitudes de los mismos.

La primera fase consistió en la recopilación y análisis de información, la cual incluye la búsqueda y descarga de la imagen a clasificar, combinación de bandas y selección de áreas o muestras de entrenamiento para la clasificación supervisada. La búsqueda y descarga de la imagen a clasificar, se realizó en los visualizadores Glovis y EarthExplorer del archivo de imágenes del Servicio Geológico de los Estados Unidos (USGS). La conversión de 16 bits a 8 bits y la combinación de bandas, se realizó con un software SIG. Las áreas de entrenamiento fueron obtenidas de recorridos en campo en la región de estudio, los cuales fueron revisados y validados respecto a la categoría asignada.

La segunda fase consistió en la aplicación de los métodos de clasificación de imágenes (clasificación supervisada y método CLASlite). Con la información previamente recolectada y analizada, se procedió a aplicar la metodología de clasificación de acuerdo a cada método de clasificación. Para el método de clasificación supervisada de imágenes, los datos fueron introducidas en el clasificador de imágenes Erdas Imagine, el cual construye reglas de decisión de pertenencia a cada clase (fase de entrenamiento), basándose en las muestras de entrenamiento y para el método de clasificación con CLASlite las imágenes crudas son el punto de partida para iniciar la clasificación, ya que "el software posee un banco de firmas espectrales para realizar la clasificación". (Asner, Tasar, Sousan, & Kanpp, 2013, pág. 11)

La tercera fase consistió en el contraste y análisis de resultados obtenidos de la clasificación de tierras forestales (bosque) y tierras agrícolas (no bosque), para ello se utilizaron muestras aleatorias obtenidas en campo para evaluar la clasificación de ambos métodos. La evaluación de resultados se realizó por medio del índice de evaluación Kappa y la matriz de confusión. Se compararon los resultados con ayuda del software ArcGIS, por medio de las herramientas de análisis espacial específicamente la herramienta Tabulate Área que permite realizar la estimación espacial o correspondencia espacial de dos métodos a evaluar (Stehman & Czaplewski, 1998).

INTRODUCCIÓN

El crecimiento constante de las tecnologías permite tener acceso a un gran contenido de imágenes de satélite, algunas de acceso libre como las imágenes obtenidas a partir del sensor Landsat. Este enorme archivo de imágenes en combinación con otras ciencias como la geomática permite obtener cartografía temática que contribuye con herramientas para la toma de decisiones; así como la aplicabilidad en diversos procesos tales como: ordenamiento territorial, gestión de riesgos, seguridad ciudadana, estudios sociales, arqueológicos, ambientales, dinámicas forestales, entre otros.

También existe un creciente número de aplicaciones y usuarios de información geoespacial lo que hace necesario la exploración de nuevos métodos de clasificación de imágenes que optimicen el uso de los datos de entrada y obtener mejor exactitud en los resultados. Algunos métodos como el de clasificación supervisada y no supervisada fueron ampliamente estudiados, pero el surgimiento de métodos utilizando nuevos algoritmos requiere el análisis de sus resultados, a partir de su aplicabilidad.

Este trabajo de investigación evalúa dos métodos de clasificación de imágenes de mediana resolución espacial, para la obtención de mapas de cobertura forestal y uso de la tierra, específicamente el método de clasificación supervisada utilizando el algoritmo de máxima verosimilitud y el método CLASlite. Se analizaron sus ventajas y desventajas y se evaluó su rendimiento con base a

los resultados obtenidos respecto a la categorización de cobertura forestal y no forestal.

La primera fase de la investigación se concentra en la adquisición de imágenes de mediana resolución del “sensor Landsat con un nivel de tratamiento 1T, el cual contiene correcciones de reflectancia atmosférica y topográfica”, la imagen adquirida tiene una proyección cartográfica UTM, referenciada a WGS84. Esta primera fase busca la consistencia de los datos y asegurar la calidad cartográfica de los archivos de entrada. Posteriormente, se establece una leyenda de trabajo con base a las capacidades de cada método de clasificación, mismas que sirven para la generación de las muestras de entrenamiento y extracción de firmas espectrales, las cuales son aplicables a los métodos de clasificación evaluados en esta investigación (Department of the Interior U.S. Geological Survey, 2017, pág. 3)

La siguiente fase consiste en la aplicabilidad de los dos métodos de clasificación de imágenes, para ello se hace uso del software Erdas Imagine para el método de clasificación supervisado utilizando el algoritmo de máxima verosimilitud; y el software CLASlite. Finalmente se desarrolla con base a protocolos de evaluación de exactitud temática cartográfica, un método para el análisis de los resultados para cada tipo de clasificación, esto se realiza con base a muestras de campo, las cuales serán contrastadas con los datos de clasificación. Los resultados de dicho contraste fueron analizados a partir de una matriz de confusión, donde se determina un porcentaje de correspondencia espacial, el cual es evaluado con el índice Kappa, que permite identificar cuál de los dos métodos presenta mejores resultados. (Stehman & Czaplewski, 1998)

1. MARCO TEÓRICO

Este capítulo hace referencia a la conceptualización y descripción de los aspectos importantes que intervienen en la clasificación de imágenes de satélite, para la obtención de categorías de cobertura y usos de la tierra.

1.1. Sensores remotos

También llamados instrumentos de teledetección, “son los dispositivos instalados en un satélite o plataformas móviles, capaces de generar imágenes, a partir de la energía electromagnética emitida o reflejada por los objetos de análisis” (Pinto, 2012, pág. 5).

Estos sensores se clasifican de acuerdo a su capacidad de detectar la energía reflejada o absorbida:

- Sensores pasivos: estos tipos de sensores no emiten energía y se limitan a detectar la energía del sol, absorbida o reflejada por los objetos (Pinto, 2012, pág. 5).
- Sensores activos: cumplen la misma función que los sensores pasivos, pero a diferencia de los anteriores tiene la capacidad de emitir su propia energía (Pinto, 2012, pág. 5).

Los sensores viajan a través de su órbita específica recolectando la radiación, a través de imágenes de los objetos situados sobre la superficie de la tierra. La información captada por estos sensores es almacenada y representada en formato digital, mediante imágenes de satélite (Plaza, 2015).

1.1.1. Imagen de satélite

Una imagen de satélite es la relación que existe entre una posición específica, es decir, coordenadas geográficas o proyectadas en relación a un valor que capta el sensor en el momento de la toma de la imagen, este valor responde a la relación que existe entre la capacidad de un objeto de reflejar o absorber luz (Sanz, 2015).

Estas imágenes son adquiridas a través de diferentes plataformas de almacenamiento como los son los que administra el Servicio Geológico de los Estados Unidos (USGS). La adquisición de estas imágenes es el punto de partida para la identificación de un área de estudio, el proceso continúa con la exploración de las imágenes, y finalmente se realiza el análisis con la ayuda de técnicas de teledetección y algoritmos computacionales.

Gonzales & Woods (2002) citado por Plaza (2015), las técnicas de teledetección y los algoritmos computaciones “analizan e interpretan la información contenida en pixeles, que agrupan características específicas del terreno, el término pixel proviene del vocablo *picture element*” (págs. 14-15).

1.1.2. Pixel

Chuvieco (1995) citado por Bautista & Palacio (2011), definen pixel como “la representación digital con posición geográfica específica que conforma una imagen de satélite o fotografía aérea, el mismo adquiere un valor denominado nivel digital, el que consiste en el arreglo espacial de números digitales que corresponde a la intensidad de luz o tonos de grises que percibe el sensor. (pág. 741)

Según Piedra Fernández (2008) en teledetección, percepción remota o fotointerpretación, “un pixel es la unidad mínima de interpretación y análisis del cual se hace uso para la representación coherente de una imagen de satélite, a través de categorías de uso y cobertura de la tierra, interpretación de unidades geológicas, cartográfica e identificación de recursos naturales, entre otros” (pág. 56)

En el pasado, los niveles digitales contenidos en un pixel eran analizados a través del sensor natural, es decir, el ojo humano; este proceso de interpretación demandaba recurso tiempo y recurso humano, lo cual dificultaba el avance de proyectos en períodos cortos de tiempo; en la actualidad los niveles digitales contenidos en los pixeles son analizados en periodos cortos de tiempo con la ayuda de algoritmos computacionales, utilizando sistemas de información geográfica (SIG).

1.2. Sistemas de Información Geográfica (SIG)

Los GIS por sus siglas en inglés (*Geographic Information System*), son en la actualidad una de las principales herramientas para el monitoreo de recursos naturales y con un creciente potencial de uso en el área industrial.

Estada (2014), define a los GIS como "un sistema computarizado que permite la entrada, almacenamiento, representación y salida eficiente de datos espaciales (mapas) y atributos (descriptivos) de acuerdo con especificaciones y requerimientos concretos" (págs. 157-158).

Para Olaya (2014), los SIG son: "una clase específica de los Sistemas de Información, donde básicamente el proceso es el mismo de lectura, edición, almacenamiento y, en términos generales, gestión de datos espaciales; con lo cual se tiene un mejor criterio para la toma de decisiones y gestión". (pág. 7)

El propósito de la Información geográfica está en conocer la ubicación sobre un elemento en la superficie de la Tierra, es decir, la información georreferenciada, como se indica en una de sus características. (Universidad de Murcia, 2006)

1.3. Clasificación de imágenes

Para este estudio se considera la clasificación de imágenes como el proceso en el cual se asignan a las imágenes satelitales categorías de cobertura

vegetal, uso del suelo, geología u otras características del terreno, a través de técnicas de teledetección, la cual se apoya de algoritmos computacionales.

1.3.1. Clasificación de imágenes basada en pixeles

En la actualidad existen diversos métodos para la clasificación de imágenes basada en pixeles, Reuter (2002), indica que estos métodos se basan en patrones que realizan la interpretación de la imagen con base a la “asignación de clases o categorías a cada uno de los pixeles de la misma”. (pág. 102)

Según Plaza (2015), la aplicación de estos métodos ofrece resultados importantes en distintas disciplinas entre las que se pueden mencionar el monitoreo forestal, la agricultura, hidrología, meteorología entre otras (págs. 52-54).

La clasificación basada en pixeles se puede dividir, según el algoritmo utilizado para la clasificación de imágenes:

1.3.1.1. Clasificación no supervisada

Este método clasificación se caracteriza por el desconocimiento de clases o categorías previamente, por lo que los criterios de clasificación se apoyan en técnicas estadísticas similitud.

Plaza (2015), define la clasificación no supervisada como la “identificación o separación de las clases o agrupaciones de forma automática utilizando para ello el criterio de similitud entre pixeles”. (pág. 54)

Según Pascual (2015), el método de clasificación no supervisada es el “proceso que permite la diferenciación de pixeles en clases o categorías de uso y cobertura de la tierra a través del algoritmo genérico de agrupamiento”. (págs. 37-41)

Dicho algoritmo en combinación con el algoritmo de densidad tiene la capacidad de simular mejores resultados, por lo que se recomienda la aplicación del híbrido algoritmo H-Density y algoritmo DHG (*Density-based Hierarchical Gaussian*) para obtener mejores resultados en relación a clasificación de imágenes de satélite (Pascual, 2015).

1.3.1.2. Clasificación supervisada

El método de clasificación supervisada utiliza algoritmos que operan con un conjunto de muestras que contienen información de clases existente sobre el terreno, las cuales son denominadas áreas o muestras de entrenamiento, estas son relevantes para la categorización de los diferentes tipos de cobertura vegetal.

Plaza (2015), define la clasificación supervisada como el “conocimiento previo de las clases existentes, a partir del cual se derivan los criterios de clasificación”. (pág. 54)

Chuvieco (2008), citado por Mejía & Moncayo (2012), indican que usualmente los métodos de clasificación supervisada se puede dividir en tres fases: Fase de entrenamiento que consiste en la asignación de muestras representativas para la clasificación. Fase de clasificación que consiste en la categorización de las muestras en clases determinadas por el usuario, y la fase de presentación de datos, que corresponde a la agrupación y representación de las categorías diferenciadas. (págs. 57-58)

Independientemente del método a utilizar es necesario definir un criterio estadístico para la fase de asignación, en dicha fase, el usuario hace la elección con base al análisis estadístico que mejor se adapte a sus intereses, según Chuvieco (1995), los criterios más comunes son:

1.3.1.2.1. Clasificador de mínima distancia

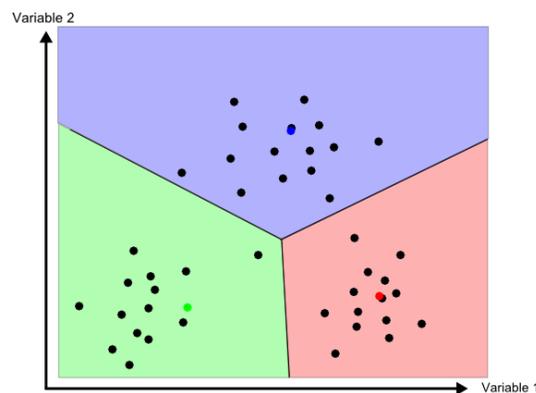
La mínima distancia es el criterio estadístico más sencillo, según Chuvieco (1995), este criterio se clasificación que se basa en la “lógica de incluir los pixeles más cercanos al pixel que contiene la categoría a evaluar”, esta cercanía es con base a los valores del pixel centroide y evalúa los pixeles que entran en este rango de mínima distancia en cada una de las bandas. (págs. 349-351)

Según Olaya (2014), el método de mínima distancia “se basa en conceptos geométricos dentro del espacio de atributos empleando únicamente la media de cada clase, prescindiéndose de la desviación típica”. (pág. 540)

Algunos autores sugieren modificaciones en el cálculo de la distancia, utilizando como el método alternativo el método distancia Manhattan, o longitud Manhattan (Olaya, 2014), así como considerar la varianza de las clases, (Chuvienco, 1995).

Con este método a todos los píxeles se asigna una clase, lo cual puede provocar errores de comisión, siendo esto una desventaja, ya que se puede asignar una clase incorrecta. En la figura 1 se ilustra la relación del método.

Figura 1. **Clasificación por mínima distancia**



Fuente: Olaya, 2014

En la figura 1, se observa agrupaciones de píxeles respecto a la distancia mínima. La desventaja de usar este criterio de clasificación es que todos los píxeles serán clasificados, lo cual no genera garantías de una buena clasificación, ya que habrá categorías que no fueron diferenciadas en la fase de entrenamiento, por lo que se considerará un problema y que es preferible dejar áreas sin clasificar que clasificarlas sin garantías (Universidad de Murcia, 2006).

1.3.1.2.2. Clasificador de paralelepípedos

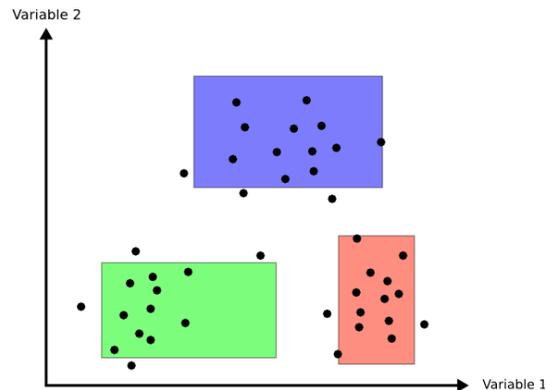
Este método establece segmentos con base a la dispersión permitida para cada categoría, es decir cada pixel es evaluado según su distancia espectral respecto a los pixeles con clase o categoría asignada, esta distancia está definida por la media de cada variable en la clase y su desviación típica. Según Olaya (2014), la longitud de los lados es igual a “dos veces la desviación típica, y se encuentran centrados en la media” . (págs. 539-540)

Este método utiliza las muestras o áreas de entrenamiento para clasificar la imagen multiespectral, analiza las muestras para cada clase generando una media para cada clase, así mismo calcula la desviación estándar para definir los límites. Definidos estos valores (media, límite inferior y superior) el algoritmo analiza cada pixel, si está contenido en alguno de los límites n-dimensionales, el pixel es asignado a la clase correspondiente, si no el pixel se agrupa en la categoría “sin clasificar” (Sanz, 2015).

Según Chuvieco (1995), este método es rápida ejecución, aplicado en la mayoría de casos como primer acercamiento para la diferenciación de categorías, así mismo resalta que este método presenta dos problemas principales. El primero es el diseño de la distancia o dispersión permitida, lo que conlleva que algunos pixeles puedan pertenecer a más de una categoría, es decir, el solape entre paralelepípedos, y la no asignación de categoría de algunos pixeles que no sean alcanzados por esta distancia o dispersión. (págs. 351-353)

La figura 2 ilustra la relación de este método (Universidad de Murcia, 2006).

Figura 2. **Clasificación por paralelepípedos**



Fuente: Olaya, 2014.

1.3.1.2.3. Clasificación por máxima verosimilitud

Este método compara cada pixel con la distribución normal de cada categoría, es decir, cada pixel tendrá una función de probabilidad respecto al vector de medias y a su matriz de varianza-covarianza, el cual es evaluado con cada categoría determinada por las áreas de enteramiento. El resultado es la probabilidad de pertenencia del pixel evaluado respecto a las categorías (Universidad de Murcia, 2006).

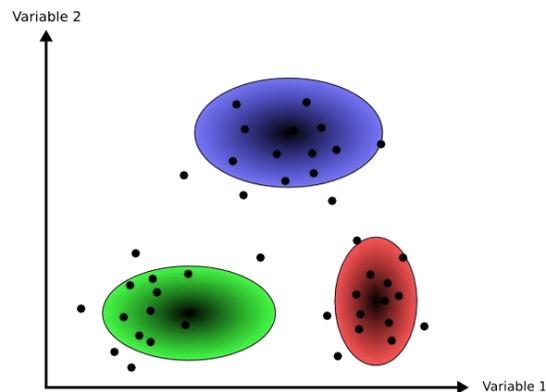
Según Chuvieco (1995), este método de clasificación es “el más complejo y el que demanda mayor volumen de cálculos”. Esto es debido a que para cada pixel se calcula la probabilidad de pertenencia respecto a cada clase. La probabilidad mayor indicara la clase a la que corresponde el pixel (págs. 353-361).

Para Olaya (2014), el método de máxima verosimilitud, según se apliquen o no umbrales de mínima probabilidad, encuentra cierta semejanza con los dos métodos estadísticos anteriores, es decir, si se establecen umbrales mínimos de probabilidad, los píxeles que no superen estos umbrales se agruparan en la categoría “Sin Clasificar”, similar característica al método de clasificación por paralelepípedos. Contrario a lo anterior, si no se establecen umbrales mínimos de probabilidad todos los píxeles serán categorizados según las muestras de entrenamiento, similar característica al método de clasificación de distancia mínima (págs. 541-542).

Actualmente es el método convencional más utilizado en los sistemas de monitoreo forestal, igualmente se considera el más complejo estadísticamente y el que demanda mayores recursos para su procesamiento computacionalmente, respecto a los métodos antes mencionados. Su alta demanda se debe a la robustez y resultados exactitud superiores a otros métodos convencionales (mínima distancia y paralelepípedos) y bastante cercanos a los métodos modernos de clasificación (SMAP, Random Forest, MARS, entre otros), según los análisis de evaluación de exactitud temática realizados por Del Toro Espin, Gomariz, Canovas, & Alonso (2013) y Quirós Rosado (2009) .

La figura 3 ilustra la relación estadística de este criterio de clasificación.

Figura 3. **Clasificación por máxima verosimilitud**



Fuente: Olaya, 2014.

1.4. CLASlite

CLASlite es un software diseñado por un equipo de investigadores del Instituto Carnegie para la Ciencia en la Universidad de Stanford. El objetivo principal de este software es la automatización de detección de zonas deforestadas y degradadas, a través de herramientas computacionales, firmas espectrales e imágenes de satélite (Asner, Tasar, Sousan, & Kanpp, 2013).

Según Asner y equipo de investigadores creadores de CLASlite, resaltan que CLASlite “es un software resultado de más de 10 años de investigación, obtención de firmas espectrales con sensores remotos y trabajo de campo para proveer una plataforma de mapeo satelital automatizado” (Asner, Tasar, Sousan, & Kanpp, 2013).

Una de las características principales de CLASlite es la biblioteca de firmas espectrales de bosques tropicales contenidas en el software. Dichas firmas fueron colectas por el equipo de investigadores en campo y a partir de técnicas de teledetección durante un periodo de 7 años (1996-2003) (Asner, Paez, Clark, Knapp, & Balaji, 2010).

Estas firmas espectrales fueron agrupadas en tres clases:

1.4.1. Vegetación fotosintética (VF)

Es la vegetación que se mantiene viva y contiene características espectrales asociadas con los pigmentos fotosintéticos de las hojas y del contenido de agua del dosel. Ajustada a las categorías del país representa la cobertura forestal (Asner, 2014).

1.4.2. Vegetación no fotosintética (VNF)

Representa la vegetación muerta o senescente, hojarasca, pastos cosechados, residuos de deforestación, residuos de manejo forestal intenso y otros materiales superficiales brillantes con características especiales asociadas a compuestos secos de carbono. Las firmas espectrales de la vegetación no fotosintética, fueron colectadas con la ayuda de un espectroradiómetro (Asner, Tasar, Sousan, & Kanpp, 2013).

En Guatemala esta categoría incluye áreas de pastoso cosechados en regiones de oriente y costa sur, áreas deforestadas e incendios forestales y zonas de inundación que provocan deforestación.

1.4.3. Superficie descubierta (S)

Representa los suelos con potencial agrícola, la infraestructura y el suelo mineral expuesto (Asner, 2014); estos datos fueron colectados con la ayuda de un espectroradiómetro que mide la radiación espectral en la longitud de onda visible e infrarrojo producida por cualquier cuerpo que reflecte la luz. La biblioteca de la superficie descubierta incluye las categorías de suelos minerales, materia orgánica superficial y condiciones de humedad (Asner, Tasar, Sousan, & Kanpp, 2013).

En Guatemala, esta categoría está compuesta por zonas de actividad volcánica, playas, zonas urbanas, zonas de rocas descubiertas, áreas kársticas, entre otras.

1.5. Métodos de evaluación de mapas de cobertura forestal en Guatemala

Castellanos y el equipo de investigadores (2011), del grupo de monitoreo de bosques en Guatemala, quienes han elaborado los mapas de cobertura forestal, indican que el método utilizado para la evaluación de exactitud temática, se realiza mediante el uso de puntos de control recolectados en campo y analizados sobre fotografías de alta resolución. El set de puntos es comparado

con imágenes de alta resolución proporcionadas por el visualizador de imágenes Google Earth; del cual se generan estadísticas que estimen el error del mapa, incluyendo matrices de contingencia, porcentaje de aciertos, precisiones de usuario y productor de mapa, entre otras. (págs. 29-32)

La exactitud de los mapas de cobertura forestal 2006 y 2010 ha obtenido resultado arriba del 92% y 91%, porcentaje considerado aceptable (Castellanos *et al.*, 2011).

1.5.1. Exactitud temática

La Unidad Administrativa Especial de Catastro Distal de Bogotá -UAECD- (2013), describe la exactitud temática como “el grado de fidelidad de los valores de los atributos asignados a los elementos en la base de datos con respecto a sus verdaderas características representadas en el mundo real”. (pág. 9)

Según la ISO 19113 (2002), exactitud temática “es la exactitud de los atributos cuantitativos o no cuantitativos de las correspondientes clasificaciones de los elementos y de sus relaciones”, es decir, son los elementos a valorar para determinar la calidad de los productos espaciales. (pág. 2)

Según Llactayo, Salcedo, y Ayala (2013), la exactitud temática “es la descripción cuantitativa y/o cualitativa de la realidad respecto a la clasificación computacional”. (pág. 11)

En términos generales, Stehman & Czaplewski (1998), recomiendan considerar tres aspectos para la evaluación de exactitud temática de productos cartográficos : a) selección de área, b) diseño de muestreo y c) análisis de datos. Dichos aspectos se deben realizar en el siguiente orden de evaluación (págs. 331-340):

1.5.1.2. Selección del área

Está conformado por el conjunto de muestras o áreas a ser evaluada y que está determinada por la intensidad de los cambios de bosque a no bosque, accesibilidad, costos, entre otros (Llactayo, Salcedo, y Ayala, 2013).

Para que los resultados de los métodos aplicados puedan ser considerados aplicables en el país, es necesario la definición de un área de estudio que sea representativa, e incluya la mayor variación categorías de cobertura y uso de la tierra. Otro variable a considerar es la superficie de cobertura de una imagen de satélite, es decir, el aprovechamiento máximo de una imagen para su clasificación (Del Toro Espin, Gomariz, Canovas, y Alonso, 2013)

1.5.1.2. Diseño del muestreo

Es importante definir el tamaño de la muestra con base a la población total (número de píxeles) esto se logra con la herramienta de Excel para el cálculo semiautomático de la muestra y el grado de confianza con el que se quiere evaluar.

Según Llactayo, Salcedo, y Ayala (2013), en los diseños de muestreo, las principales limitantes están fuertemente relacionadas a la disponibilidad económica y factor tiempo, lo cual circunscribe el uso de elementos estadísticos para determinar la muestra, siendo necesario la obtención del mayor número posible de elementos en campo. (pág. 13)

El tamaño de la muestra corresponderá a los puntos de control que serán contrastados con las imágenes o como se mencionó en el párrafo anterior, obtenidos de datos de campo, los cuales serán llamados datos de referencia (Llactayo, Salcedo, y Ayala, 2013).

1.5.1.3. Análisis de datos

Comprende el desarrollo de la matriz de confusión y el cálculo de las diversas medidas de exactitud, exactitud y error. El análisis de datos se evaluará a través de una matriz de confusión y el coeficiente de validación Kappa

1.5.1.3.1. Matriz de confusión

La matriz de confusión permite la evaluación del método de clasificación aplicado. Según Rodríguez (2011), esto se logra comparando el método aplicado con datos colectados en campo u obtenidos de fuentes con mayor exactitud como lo son interpretación o mapas obtenidos a partir de imágenes de satélite de alta resolución. (pág. 35)

Según Santos (2007) citado por Rodríguez (2011) una matriz de confusión es resultado de la comparación de las clases a evaluar, ordenadas vertical y horizontalmente. (pág. 35)

El arreglo de los datos no representa mayor problema, cuando se tiene presente este arreglo durante el cálculo de métricas. Algunos autores como Olofsson (2014), recomiendan ordenar los datos de manera que verticalmente queden los datos de la clasificación y en el eje horizontal se ordenen los datos de referencia. (págs. 50-51)

La matriz de confusión genera tres tipos de exactitud: exactitud global, exactitud del usuario, exactitud del productor. La tabla I ilustra el ordenamiento de los datos en la matriz de confusión (Rodríguez, 2011).

Tabla I. **Matriz de confusión**

Clasificación de mapa	Clasificación de referencia		
	Categorías	Tierras Forestales	Tierras Agrícolas
Bosque	n11	n12	n1j
No Bosque	n21	n22	n2j
Total	ni1	ni2	1

Fuente: elaboración propia, a partir de ejemplo de Rodríguez, 2011.

1.5.1.3.2. Coeficiente de validación Kappa

El coeficiente de validación Kappa es un estadístico que permite medir la concordancia de dos metodologías, la primera, a partir de los resultados

obtenidos del método de clasificación 1, 2, contrastándolos con imágenes de alta resolución (Rodríguez, 2011).

Moreno y Villon (2012), recomiendan para evaluaciones de métodos de clasificación de imágenes, “comparar y obtener un valor cuantitativo que indique el grado de acierto del modelo a evaluar”. (pág. 11)

El uso del coeficiente de Kappa es recomendable para este estudio, por su simplicidad y fácil interpretación, así como la aplicación en estudios similares como el realizado por Jena, Reyes, y Azucena (2003), los estudios citados en párrafos anteriores y en donde se ha logrado determinar el grado de acierto de los mapas, a través del análisis comparativo de los resultados obtenidos de la matriz de confusión. (págs. 61-62)

Para la evaluación de la exactitud temática de los mapas de cobertura forestal (bosque y no bosque), es necesario disponer de una escala para la interpretación del coeficiente de Kappa, como la propuesta por Llactayo, Salcedo, y Ayala (2013). En la Tabla II, se ilustra el ordenamiento del coeficiente de Kappa y su respectiva fuerza de concordancia. (pág. 20)

Tabla II. **Valoración del coeficiente Kappa**

Coeficiente de Kappa	Fuerza de Concordancia
0.00	Pobre
0.01 - 0.20	Leve
0.21 - 0.40	Aceptable
0.41 - 0.60	Moderada
0.61 - 0.80	Considerable
0.81 - 1.00	Casi Perfecta

Fuente: Llactayo, Salcedo, y Ayala, 2013.

1.6. Antecedentes de la clasificación de imágenes en Guatemala

Históricamente, los métodos de clasificación de imágenes se han actualizado constantemente a nivel mundial, esto promovido por el uso de las imágenes gratuitas del satélite Landsat puesto en órbita espacial desde 1972. Este satélite proporciona imágenes de mediana resolución a intervalos de tiempo frecuente y de casi cualquier parte del mundo (Pérez, 2006).

En Guatemala, se han realizado diversos intentos por estandarizar una metodología de clasificación de imágenes para la obtención de categorías de uso de suelo. Los primeros esfuerzos fueron realizados por la comunidad internacional en el año de 1988, a través de la FAO (Organización de las Naciones Unidas para la Agricultura y la Alimentación), el cual realizó los primeros intentos por mapear los bosques pero fue hasta 1999 que el Instituto Nacional de Bosques (INAB) continuó con estos esfuerzos, ese mismo año, el Ministerio de Agricultura, Ganadería y Alimentación (MAGA) generó un mapa de cobertura del país en el que se mapearon los bosques pero el enfoque principal de este mapa fue la caracterización de los diferentes sistemas de producción agrícola. Un año después dan inicio diversos estudios que culminarían en la generación de mapas de dinámica de la cobertura forestal 1991, 1996 y 2001, mismos que fueron liderados principalmente por la Universidad del Valle y el Instituto Nacional de Bosques (Castellanos *et al.*, 2011).

Fue a partir de estos esfuerzos que surge la necesidad de estandarizar una metodología aplicable al país, que pueda implementarse en futuras iniciativas de mapeo y a la vez ser reconocida como una metodología válida para iniciativas

globales como es el caso de los inventarios de gases de efecto invernadero (Castellanos *et al.*, 2011).

Actualmente se sigue utilizando la metodología de las dinámicas forestales que tiene como base el algoritmo de máxima verosimilitud. De los métodos de clasificación supervisada de imágenes, el de máxima verosimilitud es el más utilizado, debido a su robustez, porque los resultados que se obtiene a partir de su aplicación son más acertados (Chuvienco, 1990).

Este método fue adoptado en diversos estudios nacionales para el mapeo de bosque en Guatemala y en otros países de la región. En Guatemala se han aplicado métodos para la determinación de errores espaciales y la exactitud de los mapas. En la evaluación de los mapas de cobertura forestal y uso de la tierra de los años 2001 y 2006, se obtuvieron porcentajes de validación de 82 % y 92 % respectivamente. Para este estudio se propone la aplicación de buenas prácticas para la evaluación de productos cartográficos como el que ya se ha aplicado en proyectos nacionales y en otros países en el que se incluyen el índice Kappa para contrastar dos productos cartográficos y determinar su correspondencia espacial (Plata, 2010).

En otros países se han realizado análisis comparativos más profundos para evaluar la eficiencia del método de clasificación supervisada, por ejemplo, en México se ha reportado que esta clasificación tiene mejores resultados respecto de otros estudios utilizando diferentes metodologías, obteniendo porcentajes de evaluación de exactitud temática de 82.32 % para la clasificación supervisada, 78.72 % con el método interpretación visual y 73.18 % aplicando el método de clasificación no supervisada. Aun cuando el análisis realizado por Palacios en

1994, no se aplicó a metodológicas más actualizadas, genera importantes precedentes para realizar análisis comparativos con otros métodos en futuros estudios (Palacios, 1994).

En años posteriores se realizaron estudios que comprendían el análisis de los resultados, a partir de la clasificación supervisada en comparación con métodos más actuales. A diferencia de los estudios realizados por Palacios (1994), se obtuvieron mejores resultados en los métodos que utilizaron el algoritmo de segmentación de imágenes y el índice de vegetación diferencial normalizado NDVI (Pérez, 2014).

Este comportamiento se considera normal partiendo del conocimiento que las técnicas de percepción remota se encuentran en constante actualización; lo que promueve el surgimiento de nuevas metodologías que hacen uso de algoritmos cada vez más complejos y potentes en sus funciones; el método CLASlite es un ejemplo de la potencialidad de los nuevos algoritmos. Este método está integrado en un programa computacional con el mismo nombre CLASlite® diseñado principalmente para la identificación automática de la deforestación a través del uso de imágenes de satélite. El uso de este programa ha tenido un gran crecimiento de aplicabilidad en países como Colombia, Venezuela, Bolivia, Ecuador, entre otros de la región. Es utilizado por su simplicidad y su estrecha relación con los sistemas de monitoreo forestal para el reporte de emisiones de carbono, necesario para ingresar al sistema de beneficios financieros de los mercados de carbono (Gúzman, 2015).

Adicional a lo anterior, CLASlite tiene integrado módulos de corrección atmosférica de la imágenes descargadas y aunque esta característica no es

esencial para otros métodos si es parte fundamental y mejora los resultados utilizando el algoritmo de Monte Carlo, lo que hace suponer que este método facilita el procesamiento de las imágenes Landsat, para obtener mejores resultados (Pacheco, 2011).

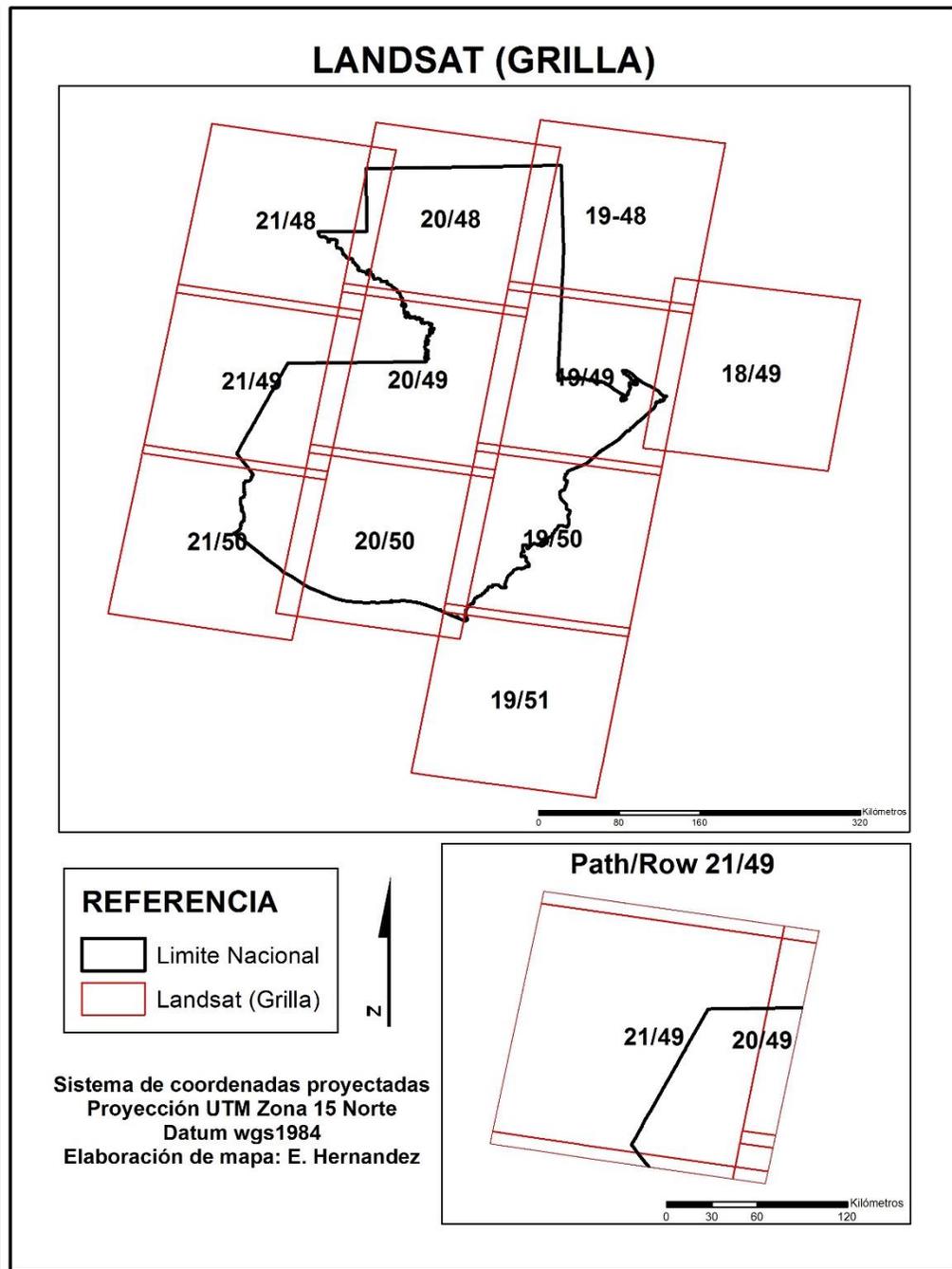
Se propone esta investigación para conocer nuevas alternativas de clasificación de imágenes, ya que estas permiten tener un diagnóstico y la dinámica respecto a distintos usos de la tierra, pero estos métodos deben pasar por procesos de análisis y evaluación respecto a su exactitud y aplicabilidad.

2. APLICACIÓN DE LOS DOS MÉTODOS DE CLASIFICACIÓN DE IMÁGENES

2.1. Área de estudio

Para la definición de área de estudio, se eligió el área con mayor representación de occidente contenida en una imagen Landsat. En tal motivo se utilizó el área contenida en el Path/Row 21-49 localizada en los límites de Guatemala. La figura 4 representa la grilla Landsat para el país y la selección del área Path/Row 21/49. Esta área fue definida de acuerdo su dinámica y crecimiento respecto al área de bosque. Según los datos de la dinámica forestal 2006-2010, la región de occidente es la que presenta mayor regeneración forestal en relación a otras áreas del país. En tal motivo, esta área es óptima para evaluar si los modelos pueden categorizar o clasificar las área de bosque para monitoreo (Castellanos *et al.*, 2011)

Figura 4. Grilla Landsat para Guatemala



Fuente: elaboración propia, a partir de datos USGS.

Se continúa con la determinación del área de estudio, para ello fue necesario calcular las tasas de deforestación de las regiones del país y elegir que se adecuara a los objetivos de la investigación. La tabla III presenta los resultados de estos cálculos.

Tabla III. **Cálculo de tasa de deforestación dinámica forestal 2006-2010**

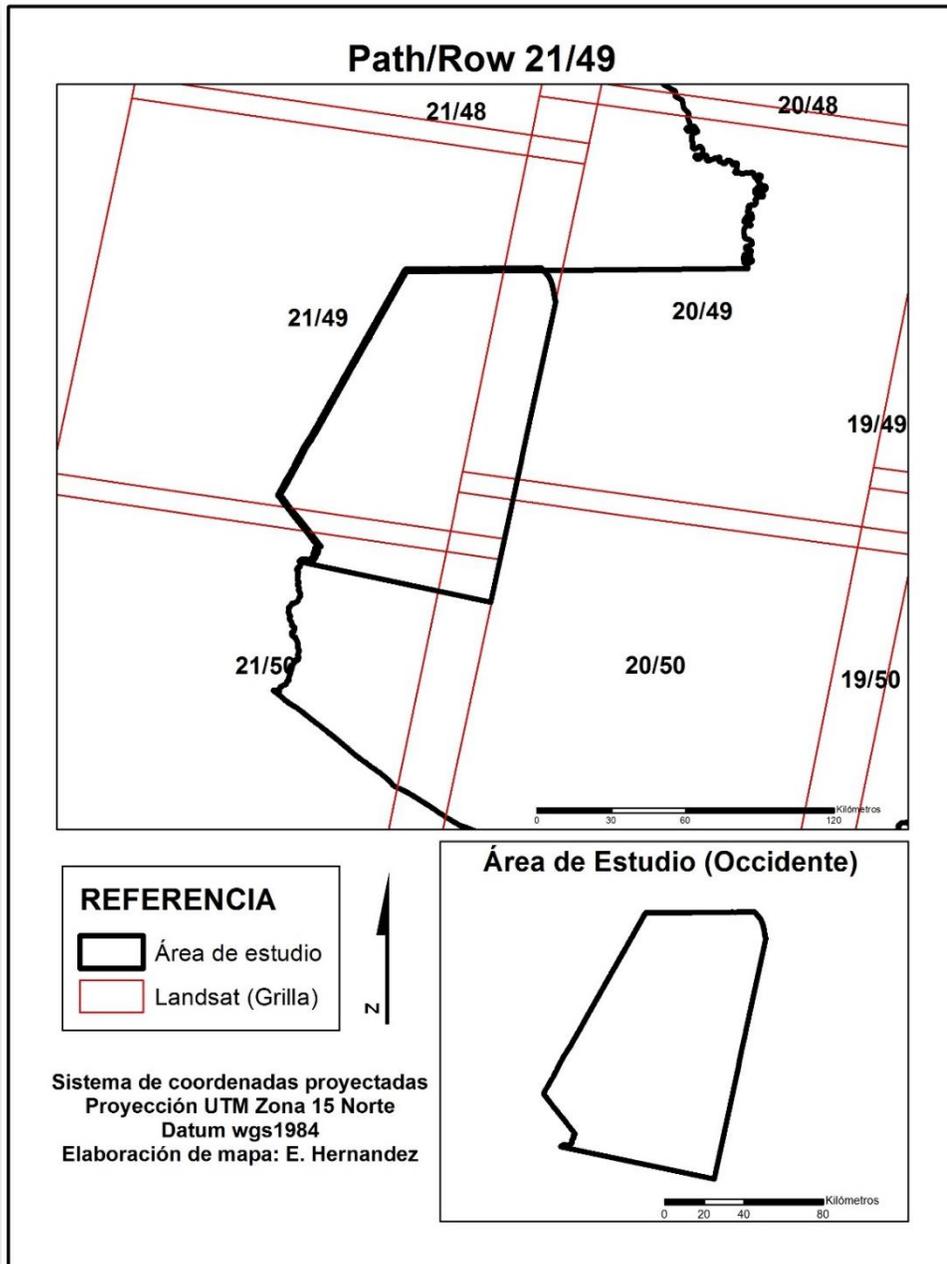
2006 - 2010						
CATEGORIAS	Centro - Oriente	Costa Sur	Occidente	Sarstún Motagua	Tierras Bajas del Norte	Total (has)
Total área deforestada	69,172.22	34,705.55	106,579.13	101,603.32	194,001.18	506,061.40
Cobertura Forestal	245,217.22	128,292.80	720,141.45	607,137.42	2,134,784.35	3,835,573.24
Total de tierras recuperadas	36,749.75	40,406.75	126,586.61	72,467.61	84,415.47	360,626.18
Cambio Neto	-32,422.48	5,701.21	20,007.48	-29,135.72	-109,585.71	-145,435.22
Cambio Anual (ha/año)	-8,554.74	1,504.28	5,279.02	-7,687.52	-28,914.44	-38,373.41
Tasa de Deforestación (%)	-3.46	1.17	0.73	-1.27	-1.35	-1.00

Fuente: elaboración propia. A partir de dinámica forestal 2001-2010 GIMBUT.

El área de estudio final tiene una extensión de 10,142Km², lo cual representa 9 % del total de superficie de la República de Guatemala (108,889Km²). A esta área fueron aplicados ambos métodos de clasificación.

La figura 5 hace referencia al mapa en el que se ubica el área de estudio, la cual se llamará seguidamente “Área de estudio Occidente”.

Figura 5. Mapa del área de estudio Occidente



Fuente: elaboración propia, a partir de datos USGS.

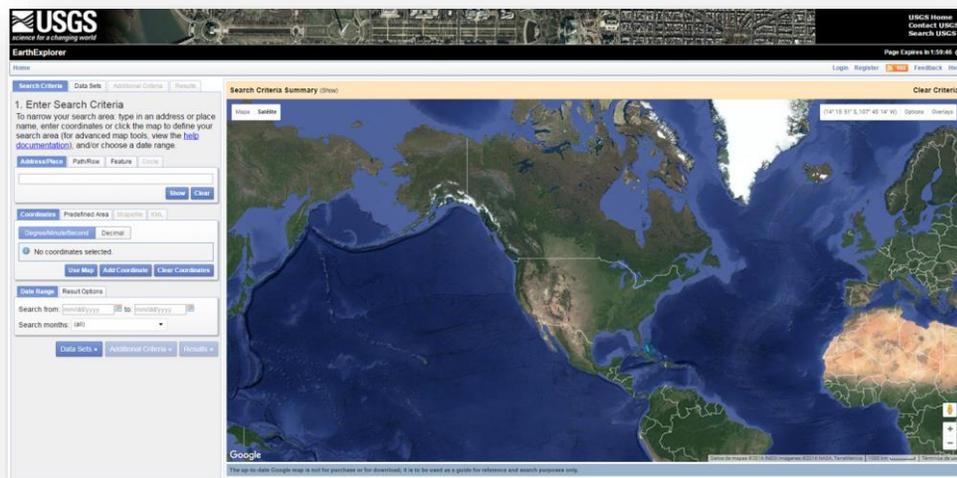
2.2. Descarga de imágenes

La búsqueda y la selección de la imagen, se realizó en los buscadores de imágenes del Servicio Geológico de los Estados Unidos. Específicamente el servidor de imágenes EarthExplorer. La selección de las imágenes, se realizó de acuerdo a las siguientes características:

- Datos colectados durante la estación seca de los años 2014-2016 (diciembre-abril) para garantizar consistencia fenológica.
- Ausencia de nubes y otros efectos atmosféricos como humo, bruma, etc.

La figura 6 muestra el buscador de imágenes Landsat del servicio Geológico de los Estados Unidos, específicamente la plataforma EarthExplorer.

Figura 6. Buscador de imágenes EarthExplorer



Fuente: EarthExplorer, USGS.

La imagen seleccionada es la que lleva por nombre “**LC80210492016034LGN00**”, la cual corresponde al año 2016, fue seleccionada esta fecha por la disponibilidad de datos de campo para este año. La fecha de obtención de la imagen corresponde al 03 de febrero de 2016, por tal motivo, se comprueba que la imagen cumple con la condición de estar entre las fechas de la época seca para el país (diciembre – abril). Porcentaje de nubes de 0.85 % cumpliendo con otro de los estándares el cual indica porcentaje de nubes menor al 10 %. Corresponde al satélite Landsat 8, sensor OLI_TIRS, tratamiento previo de la imagen L1T corrección geométrica), imagen multiespectral (11 bandas).

La tabla IV presenta los metadatos de la imagen seleccionada para esta investigación.

Tabla IV. **Metadatos escena Landsat “LC80210492016034LGN00”**

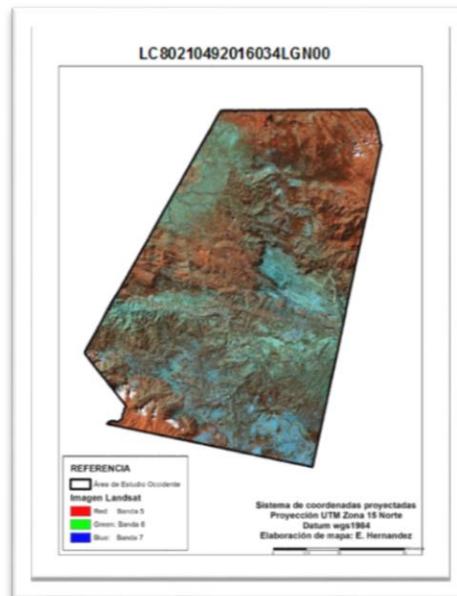
ATRIBUTO	VALOR
Escena Landsat	LC80210492016034LGN00
WRS Path	21
WRS Row	49
Día/Noche	Día
Nivel de tipo de datos	L1T
Sensor	OLI_TIRS
Fecha de adquisición	03/0122016
Cobertura de nubes de la escena	0.85
Proyección	UTM
Datum	WGS84
Elipsoide	WGS84
Zona UTM	15
Orientación	Norte
Multiespectral	11 Bandas
Resolución	30 metros
Pancromática	15 metros

Fuente: EarthExplorer, USGS.

2.3. Generación de composiciones de banda

La imagen seleccionada y descargada viene en un formato en el que cada banda viene separada; por ellos fue necesario realizar la combinación de las 7 bandas y su conversión de 16 bits a 8 bits. Este proceso se realizó con la ayuda del software ArcGis® y Erdas Imagine®. Las bandas utilizadas para la clasificación supervisada fueron las bandas 5, 6 y 7, (infrarrojo cercano, mediano y lejano respectivamente), esta combinación se eligió principalmente por referencia de los principales proyectos nacionales de clasificación de imágenes y por revisión literaria en la cual indica que esta combinación de infrarrojos permite mejor detección de bosque y no bosque.

Figura 7. Imagen LC80210492016034LGN00 combinación de bandas 5, 6 y 7



Fuente: elaboración propia.

2.4. Áreas o muestras de entrenamiento

Las áreas o muestras de entrenamiento obtenidas en campo fueron utilizadas para realizar la clasificación supervisada, a partir de estas áreas se construyen una regla de decisión basada en casos conocidos o conjunto de aprendizaje. Este conjunto de datos se llamarán áreas de entrenamiento (píxeles de la imagen) al cual fue asignado las clases a la que representan, según los datos recopilados en campo.

Las áreas de entrenamiento se obtuvieron a partir de recorridos de campo con la ayuda de un GPS en el área de estudio, dichos recorridos fueron realizados durante el mes de agosto 2016 y enero 2017. Utilizando la metodología de muestras de campo para las áreas de entrenamiento se dificulta la obtención de puntos en áreas de difícil acceso y un mayor a los obtenidos 628 muestras. Estas muestras fueron revisadas y validadas respecto a la categoría asignada. En el anexo 1, se presenta una breve descripción de las categorías identificadas en campo.

En total fueron evaluados 648 puntos, los cuales tenían la siguiente distribución: a) 325 áreas de entrenamiento correspondiente a bosque, b) 303 áreas de entrenamiento correspondiente a no bosque y se agregaron c) 20 áreas de entrenamiento correspondiente a las áreas con nubes. La revisión y validación de estos puntos consistió en contrastar con la imagen Landsat su posición, es decir, verificar que las áreas de entrenamiento recolectadas en campo se encontraran en regiones de la imagen que tuvieran potencial a ser clasificadas como bosque, las áreas que generaban cierta duda fueron eliminadas para reducir el ruido en la clasificación de la imagen.

De las 648 (628 recolectadas en campo y 20 correspondiente a nubes) áreas de entrenamiento, 334 fueron utilizadas para realizar la clasificación supervisada, ya que la clasificación con el software CLASlite tiene sus propias áreas de entrenamiento y 314 fueron utilizadas para validar ambos métodos.

2.5. Clasificación supervisada de la imagen “LC80210492016034LGN00”

Una vez definidas las áreas de entrenamiento con su asignación exacta a cada clase (bosque, no bosque y nubes), fueron introducidos en el clasificador de imágenes Erdas Imagine, estas fueron convertidas en un formato AOI nativo de este software. A partir de estas áreas, el software construye reglas de decisión de pertenencia a cada clase (fase de entrenamiento), basándose en las muestras de entrenamiento. Estas áreas extraen la información contenida en la imagen de satélite por cada pixel donde se posiciona esta información (firma espectral).

La firma espectral le indica al algoritmo que las distintas categorías tienen un rango de firma para que haga la separación semiautomáticamente en toda la imagen. La combinación de bandas utilizadas para esta clasificación fueron los infrarrojos (banda 5, 6 y 7), la cual permite diferenciar las áreas con cobertura vegetal y las áreas sin cobertura. Este software es utilizado actualmente en Guatemala, para la generación de mapas de cobertura forestal y uso de la tierra, por lo que se consideró idóneo su uso para esta investigación (Castellanos *et.al.*, 2011).

2.6. Clasificación CLASlite de la imagen “LC80210492016034LGN00”

La clasificación con el software CLASlite utiliza la imagen Landsat sin ningún tipo de tratamiento previo únicamente la corrección geométrica en tal motivo hace uso de las 11 bandas del sensor OLI_TIRS y los metadatos, los cuales sirvieron para realizar el proceso de corrección radiométrica que es uno de los módulos que tiene incorporado. Posterior a la corrección radiométrica, el programa realiza una fragmentación de la imagen basándose en la biblioteca de firmas espectrales contenidas en el programa siendo así los tres estratos clasificados; a) vegetación fotosintética (VF), b) vegetación no fotosintética (VNF) y c) superficie descubierta (S). La potencialidad del software se basa en la incorporación de probabilidad de acuerdo a las características de cada pixel utilizando el método Monte Carlo.

Adicional a la información contenida en cada una de las bandas de la imagen el software logra enmascarar las nubes, los cuerpos de agua y sombras provocadas por el relieve, lo cual facilita al algoritmo en la toma de decisiones a la hora de fragmentar la imagen. Finalmente, el programa utilizó el archivo de la imagen fragmentada y con el proceso de árbol de decisiones obtenido del algoritmo de Monte Carlo; clasifica la imagen tomando el fragmento VF y lo convierte a cobertura de bosque y no bosque para lo cual se asignaron los valores por defecto recomendado en el manual de usuario, donde $VF \geq 80$ es bosque y menor a este valor no bosque.

3. PRESENTACIÓN DE RESULTADOS

Este capítulo presenta los resultados obtenidos de los métodos de clasificación en el cual se presentarán datos cuantitativos de ambos métodos, y se compararán para evaluar y determinar la exactitud.

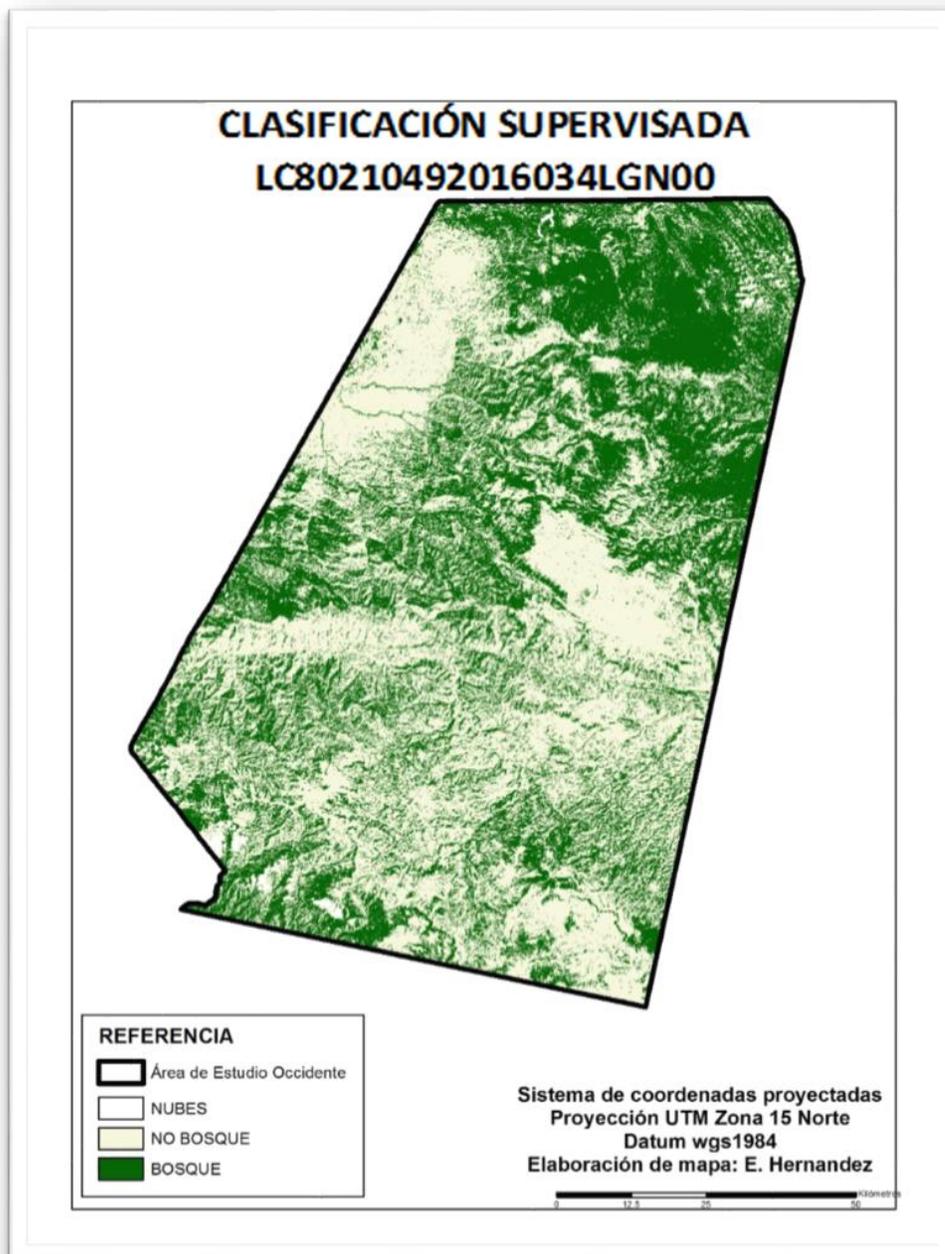
3.1. Identificación de categorías de tierras forestales (bosque) y tierras agrícolas (no bosque)

Para la evaluación de los dos métodos de clasificación fue necesario la segmentación de las imágenes en dos grandes categorías, mismas que son importantes para los sistemas de monitoreo forestal, siendo: a) tierras forestales o bosque y b) tierras agrícolas o no bosque.

3.1.1. Clasificación supervisada

Se utilizaron como insumos la imagen combinada (bandas 5, 6 y 7), áreas de entrenamiento y el software Erdas. Se aplicó el algoritmo de máxima verosimilitud para la obtención de dos categorías de uso y cobertura de la tierra siendo estas bosque (43.64 %) y no bosque (55.88 %); y dos categorías adicionales para la separación de las nubes (0.35 %) y sin clasificar (0.13 %). La figura 8 representa la distribución espacial de la clasificación supervisada en el área de estudio de occidente.

Figura 8. Distribución espacial de categorías de cobertura y uso de la tierra por el método de clasificación supervisada



Fuente: elaboración propia, a partir de clasificación supervisada.

La tabla V presenta los datos de superficie en kilómetros cuadrados (km²) resultante de la clasificación supervisada con el algoritmo de máxima verosimilitud. Esta tabla da como resultado mayor porcentaje de la categoría no bosque (55.88 %) en relación a las demás categorías, se logró el 99 % de clasificación de la imagen.

Tabla V. **Superficie de categoría resultado de la clasificación supervisada**

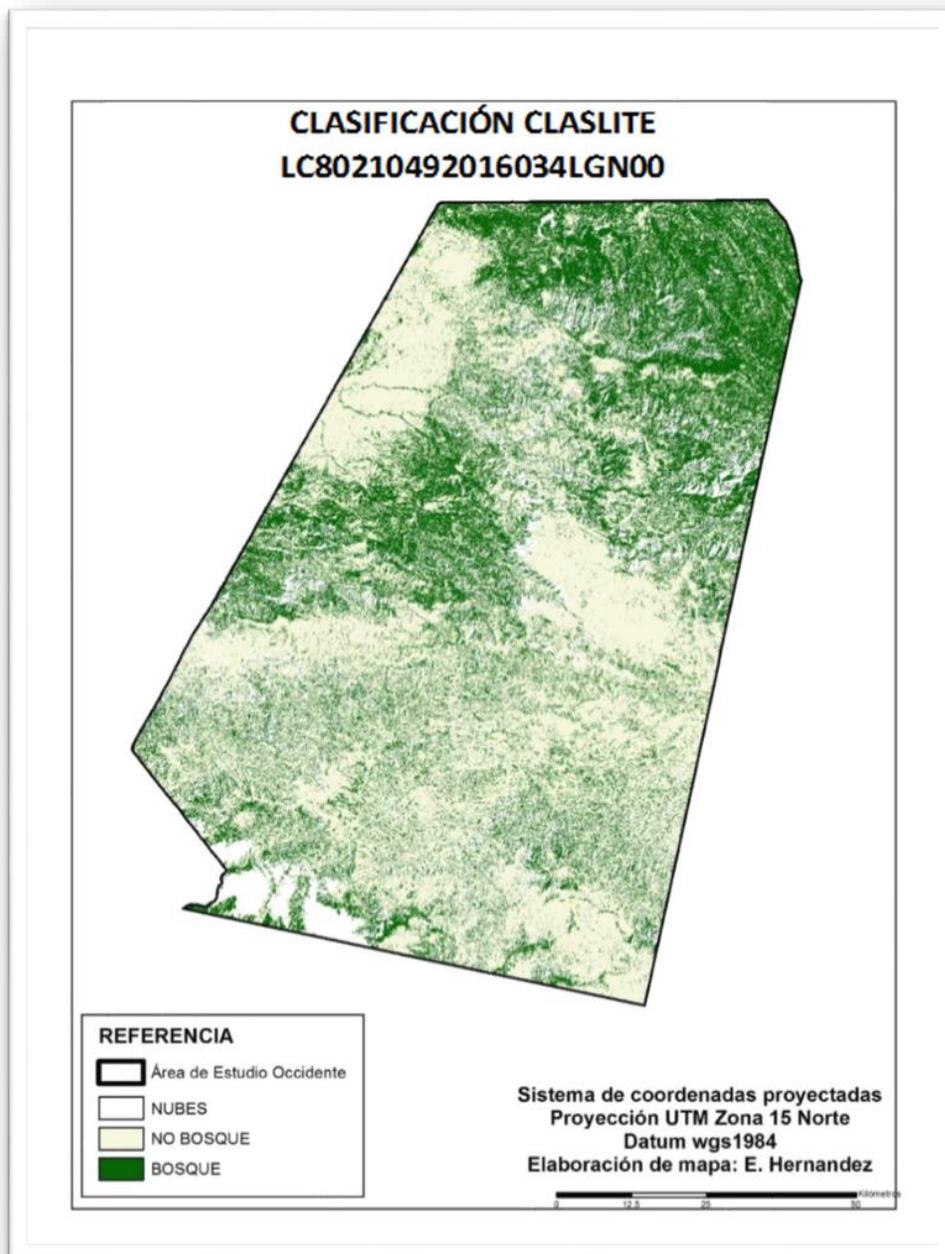
CLASIFICACIÓN SUPERVISADA	
Categoría	Superficie Km²
Sin Clasificar	13.00
Nubes	36.00
No Bosque	5,667.00
Bosque	4,426.00
Total	10,142.00

Fuente: elaboración propia, a partir de clasificación supervisada.

3.1.2. Clasificación CLASlite

Para esta clasificación se utilizaron las imágenes crudas del archivo “LC80210492016034LGN00”, el software realizó con base a sus diferentes módulos la categorización de dos categorías de uso y cobertura de la tierra: a) bosque (49.79 %) y b) no bosque (32.80 %); y dos categorías adicionales de clasificación: c) nubes (0.35 %) y sin clasificar (17.06 %). La figura 9 representa la distribución espacial de las anteriores categorías.

Figura 9. **Distribución espacial de categorías de cobertura y uso de la tierra por el método de clasificación CLASlite**



Fuente: elaboración propia, a partir de clasificación con CLASLite.

La tabla VI presenta los datos de superficie en kilómetros cuadrados (km²) resultante de la clasificación con CLASlite. Esta tabla da como resultado mayor porcentaje de la categoría bosque (49.79 %) en relación a las demás categorías, se logró el 99 % de clasificación de la imagen.

Tabla VI. **Superficie de categoría resultado de la clasificación CLASlite**

CLASIFICACIÓN CLASLITE	
Categoría	Superficie Km²
Sin Clasificar	1,730.00
Nubes	36.00
No Bosque	3,326.00
Bosque	5,050.00
Total	10,142.00

Fuente: elaboración propia, a partir de clasificación con CLASLite.

4. DISCUSIÓN DE RESULTADOS

Como se puede observar en las distintas tablas de resultados de ambos métodos ambos tuvieron la capacidad de clasificar el bosque y no bosque. Para la clasificación supervisada reportó mayor superficie (1,241km²) en relación a la categoría bosque. Así mismo, se puede analizar que la clasificación de CLASlite reporta mayor superficie (624km²) de bosque en relación a la clasificación supervisada. Otro valor a resaltar es que la clasificación con CLASlite reporta mayor superficie (1,723km²) de bosque en relación a la categoría no bosque.

Aún cuando es un análisis muy superficial, se puede concluir que ambos métodos, con diferencia en términos de superficie tiene la capacidad de categorizar bosque y no bosque, es importante resaltar que esta fase permitió identificar que ambos métodos presentan las siguientes limitantes:

El método de clasificación supervisada no excluye los factores como sombras provocado por el relieve, en tal motivo tiende a categorizar estas áreas oscuras como bosque.

El método de calificación con CLASlite realizó una exagerada discriminación de áreas con sombras provocadas por el relieve, en tal motivo, presenta un alto valor de superficie sin clasificar (1,730km²).

Otro análisis necesario fue obtener el área de bosque reportada en otros proyectos nacionales, con el fin de comparar con las categorías que se obtuvieron en ambos métodos evaluados. La tabla VII presenta los resultados de superficie de otros proyectos de categorización de cobertura y uso de la tierra específicamente los siguientes proyectos:

- Mapa de cobertura vegetal y usos de la tierra MAGA, 2010.
- Mapa de cobertura forestal y usos de la tierra INAB, CONAP, UVG, 2010.
- Mapa de cobertura forestal y usos de la tierra REDD+, 2014, (Aun no publicado).
- Mapa de bosques y usos de la tierra INAB, 2012.

Tabla VII. **Análisis de superficies de bosque y no bosque en otros proyectos**

COBERTURA Y USO DE LA TIERRA KM ²				
CATEGORIAS	MAGA 2010	LEDS 2010	REDD+ 2014 No publicado	INAB 2012
Sin clasificar	199	44	56	197
No Bosque	5,797	6,526	6,123	6,341
Bosque	4,145	3,572	3963	3,604
Total	10,142	10,142	10,142	10,142

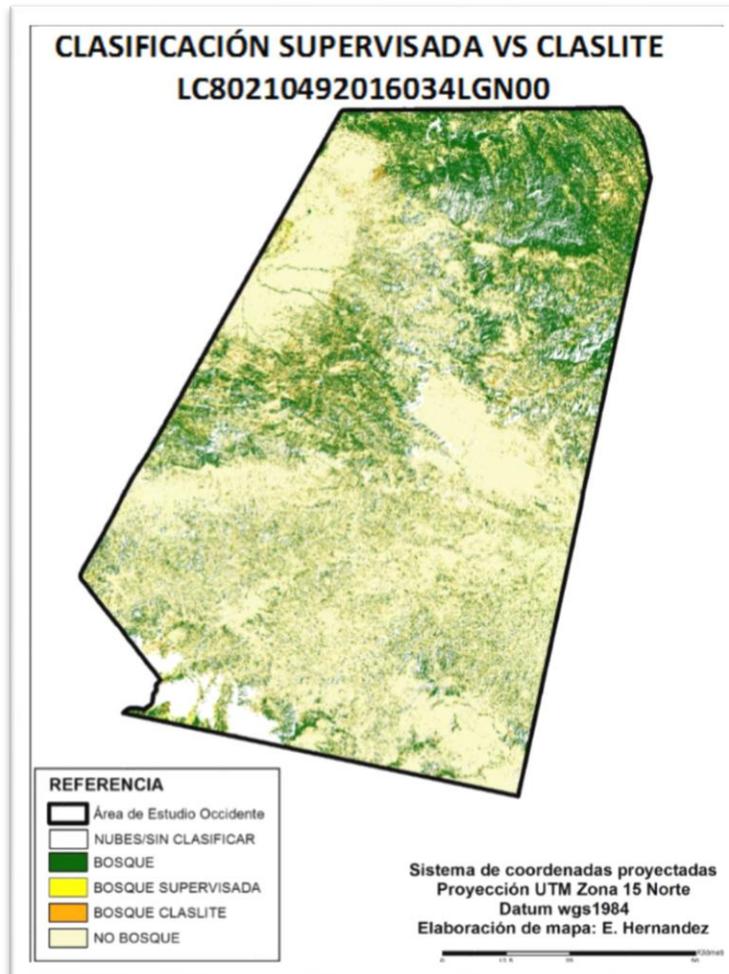
Fuente: INAB, CONAP, UVG MAGA, URL, MARN.

La tabla VII permite analizar que los resultados se encuentran poco cercanos a los datos de superficie obtenido de los dos métodos de clasificación evaluados en esta investigación, siendo el reporte de MAGA, 2010 (4,145 km²) el más próximo al dato de superficie (4,426 km²) del método de clasificación supervisada.

4.1. Determinación de correspondencia espacial y cuantitativa de los resultados de la clasificación supervisada y clasificación con CLASlite

Como parte de unos de los objetivos de esta investigación es determinar la correspondencia espacial entre ambos métodos. En tal motivo fue necesario reclasificar las imágenes resultantes de los métodos en la primera etapa y dejar únicamente las categorías de bosque y no bosque. La continuación de este método fue la sobreposición de ambas imágenes clasificadas y extraer el área de coincidencia espacialmente. La figura10 representa la correspondencia espacial de ambos métodos de clasificación.

Figura 10. **Correspondencia espacial de dos métodos de clasificación**



Fuentes: elaboración propia, a partir de datos de clasificación evaluados.

En la figura 10 se puede observar que ambos métodos de clasificación tienen alta correspondencia espacial en las categorías de bosque y no bosque. Visualmente se puede determinar que prevalece el bosque obtenido con el software CLASlite sobre las áreas bosque que se obtiene con el método de clasificación supervisada.

Para fundamentar esta evaluación fue necesario cuantificar estas áreas de coincidencia espacial, por lo que se calcularon las 4 categorías resultantes y una categoría adicional, la cual corresponde a la sumatoria del área que no se calificaron en ambos métodos. La tabla VIII presenta los resultados cuantitativos de la correspondencia espacial.

Tabla VIII. **Datos de superficie de la correspondencia espacial de los dos métodos de clasificación**

CORRESPONDENCIA ESPACIAL KM ²	
CATEGORÍAS	SUPERFICIE
Nubes/Sin Clasificar	1,786.00
Bosque	2,424.00
Bosque Supervisada	446.00
Bosque CLASlite	888.00
No Bosque	4,597.00
TOTAL	10,142.00

Fuentes: elaboración propia, a partir de datos de clasificación evaluados.

4.2. **Análisis de correspondencia espacial y cuantitativa de los resultados de la clasificación supervisada y clasificación con CLASlite**

Con base a los resultados obtenidos en el apartado anterior, se puede concluir que espacialmente (visual) existe alta correspondencia espacial, se ve reflejado en el número de píxeles de coincidencia (7,801,500.00), lo cual corresponde a un 70 % de correspondencia espacial de la superposición de la imagen. Este porcentaje de correspondencia contrastado con los rangos del

coeficiente de Kappa se puede concluir que ambos métodos tienen un rango “considerable” de concordancia o correspondencia espacial.

Figura 11. **Análisis de correspondencia espacial entre ambos métodos evaluados**



Fuentes: elaboración propia, a partir de datos de clasificación evaluados.

Respecto a los resultados cuantitativos, del total de la imagen, la correspondencia espacial (coincidencia y sin coincidencia espacial) en kilómetros cuadrados es 8,356.00, lo cual representa el 82 % del total de la imagen clasificada. El 18 % (1,786.00 km²) restante corresponde a la sumatoria de área en ambos métodos de clasificación que no fueron clasificados o representa a la cobertura de nubes en la imagen.

Del total de bosque clasificado que coincide y no coincide en ambos métodos (3,756 km²) solo el 64 % (2,424 km²) de bosque coincide en ambas imágenes. El total del área de correspondencia espacial es de 8,356 km² de las

cuales el área de coincidencia para las dos categorías fue de 84 % repartidas de la siguiente forma: bosque 2,424.00 km² (29 %), no bosque 4,597.00 km² (55 %) y no correspondencia 1,334.00 km² (16 %). El porcentaje de correspondencia espacial cuantitativo es de 84 %, contrastado con los rangos del coeficiente de Kappa, se puede concluir que ambos métodos tienen un rango “casi perfecto” de concordancia o correspondencia cuantitativa.

4.3. Determinación de la exactitud de los resultados de la clasificación supervisada y clasificación con CLASlite

El método de validación o evaluación de exactitud en Guatemala es a través de la obtención de los datos de referencia. Los datos de referencia fueron obtenidos de recorridos de campo y su posterior verificación y análisis sobre imágenes de mayor resolución. Oloffson (2014) presenta diferentes alternativas para la evaluación de productos cartográficos, pero la decisión del insumo para la obtención de los datos de referencia dependerá de costos y tiempos. Una buena opción es el uso de imágenes de mejor resolución y el uso de la herramienta Google Earth, para la obtención de estas imágenes y la opción de análisis histórico, la cual permite la obtención imágenes de una fecha en particular. Para este estudio los datos de referencia fueron obtenidos de datos de campo durante período de agosto de 2016 a enero 2017.

Este proceso contempló la evaluación de cada uno de los métodos de clasificación, con el objetivo de conocer la exactitud de los productos obtenidos con base a experiencia de otros países en el marco normas ISO 19113, para evaluar exactitud temática de productos cartográficos, el cual contempló los siguientes pasos:

De las muestras obtenidas en campo, fueron separadas aleatoriamente el 50 % (314 muestras), las cuales fueron revisadas para reducir cualquier error del levantamiento. La aleatorización del universo total de los datos colectados fue realizada espacialmente con la ayuda del software ArcGis®. La distribución de la muestra fue del tipo aleatorio, el cual se seleccionó con la herramienta de “*Create Random Points*” de ArcGis®. Esta herramienta permite crear un archivo vectorial de puntos de manera aleatoria dentro de los límites establecidos por el área de estudio, a partir de los datos colectados en campo.

A este archivo vectorial fue agregado un campo de tipo texto, el cual contiene la información de bosque (código=1) y no bosque (código=2) obtenida en campo y verificada con imágenes de mayor resolución de Google Earth. La interpretación sobre las imágenes de Google Earth se realizó configurando este programa con el objetivo de obtener las imágenes de 2016 en los meses más cercanos a la fecha de la imagen. La siguiente fase de la evaluación de exactitud es la obtención de los datos de ambos métodos de clasificación, el cual fueron agregados a la capa de puntos para realizar el análisis.

Los resultados fueron evaluados a través de una matriz de confusión, que determina la exactitud de ambos métodos de clasificación en ambas categorías (bosque y no bosque); esto se realizó utilizando el programa Excel®, en el cual se genera una tabla dinámica para la separación de cada una de las categorías de acuerdo a los dos métodos de clasificación y los datos de referencia.

La tabla IX representa los resultados obtenidos para la evaluación de exactitud en ambos métodos de clasificación con CLASlite y los datos de referencia para dos categorías bosque y no bosque.

Tabla IX. **Evaluación de exactitud del método CLASlite**

Clasificación CLASlite	Datos de Referencia						
	Categorías	Bosque	No Bosque	Total general	Exactitud de Usuario	Error de Omisión	% de Coincidencia
	Bosque	94	17	112	84	16	30
	No Bosque	21	181	202	90	10	58
Exactitud Total						88	

Fuentes: elaboración propia, a partir de datos de clasificación evaluados.

La tabla X representa los resultados obtenidos para la evaluación de exactitud en ambos métodos de clasificación supervisada y los datos de referencia para dos categorías bosque y no bosque.

Tabla X. **Evaluación de exactitud del método clasificación supervisada**

Clasificación Supervisada	Datos de Referencia						
	Categorías	Bosque	No Bosque	Total general	Exactitud de Usuario	Error de Omisión	% de Coincidencia
	Bosque	73	19	93	79	21	23
	No Bosque	42	179	221	81	19	57
Exactitud Total						80	

Fuentes: elaboración propia, a partir de datos de clasificación evaluados.

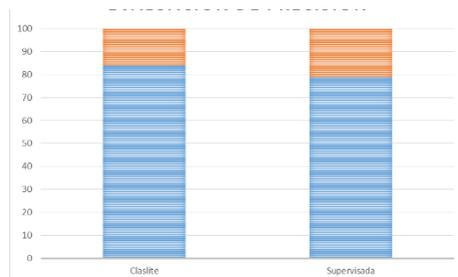
4.4. **Análisis de la determinación de la exactitud de los resultados de la clasificación supervisada y clasificación con CLASlite**

La matriz de confusión en las tablas IX y X nos presenta los resultados para cada clasificación. Ambos métodos superan el porcentaje de exactitud de 80 %. El método de clasificación con CLASlite presenta un porcentaje de exactitud de 88 % y el método de clasificación supervisada presenta una exactitud de 80 %.

Estos porcentajes de exactitud contrastada con los rangos del coeficiente de Kappa, se puede concluir que ambos métodos tienen un rango “considerable” (Clasificación supervisada) y “Casi Perfecta” (CLASlite) respecto a la exactitud de clasificación.

Los resultados también indican que el método CLASlite presenta mayor exactitud en la clasificación de Bosque (84 %) y un menor porcentaje de error (16 %) respecto al método de clasificación supervisada (79 % y 21 % respectivamente). Ambos métodos presentan mayor exactitud en los resultados de clasificación del área de no bosque siendo nuevamente CLASlite el de mayor exactitud respecto al método de clasificación supervisada 90 % y 81 % respectivamente).

Figura 12. **Análisis de exactitud entre ambos métodos evaluados**



Fuentes: elaboración propia, a partir de datos de clasificación evaluados.

Los mayores errores respecto a los datos de referencia fueron encontrados en las áreas oscuras o sombras provocadas por las nubes y el relieve (Clasificación supervisada) y en áreas consideradas para el país como vegetación secundarias o matorrales (CLASlite). Ambos métodos presentaron dificultad en el mapeo de bosque específicamente en zonas de bosque disperso y bosque de conífera siendo aún un reto en ambos métodos que debe ser

complementado con mayores datos de referencia, experiencia de campo y edición manual.

Los resultados de exactitud de ambos métodos se encuentran en rangos considerables respecto a la matriz de Kappa, por lo que pueden ser utilizados en combinación con postproceso (edición o corrección manual) para generar información temática cartográfica para la categorización de usos y cobertura de la tierra a nivel nacional.

CONCLUSIONES

1. Ambos métodos de clasificación (CLASlite y clasificación supervisada), tuvieron la capacidad de clasificar el bosque y no bosque en diferentes proporciones. La clasificación supervisada reportó mayor superficie (1,241.00 km²) en relación a la categoría no bosque. Así mismo, se puede analizar que la clasificación de CLASlite reporta mayor superficie de bosque (624.00 km²) en contraste con el otro método. Ambos métodos presentaron limitantes provocados por el relieve en el área de estudio lo que tiende a categorizar áreas oscuras como bosque y discriminar exageradamente las sombras lo que se ve reflejado en un alto valor superficial sin clasificar (1,730.00 km²).
2. Se realizó la sobreposición espacial de ambos métodos, con el objetivo de determinar la correspondencia espacial y cuantitativa. Los resultados tabulados en la tabla VII de los cuales se concluye que espacialmente (visual) existe “considerable” Concordancia o correspondencia espacial según la tabla de rangos del coeficiente de Kappa (70 % el total de pixeles coinciden espacialmente). El total de la imagen de correspondencia espacial (8,356km²), 84 % coincide, el cual contrastado con los rangos del coeficiente de Kappa, se puede concluir que ambos métodos tienen un rango “casi perfecto” de concordancia o correspondencia cuantitativa. Por ello se concluye que cuantitativamente y espacialmente el método de CLASlite puede sustituir el método tradicional de clasificación de imágenes

actualmente utilizado en Guatemala, por su fácil aplicabilidad y resultados en un período corto de tiempo.

3. Con el objetivo de conocer la exactitud de ambos métodos de clasificación, a partir de datos de referencia colectados en campo, los cuales fueron tabulados en una matriz de confusión, se concluye que el método de clasificación supervisada tiene un rango “considerable” (80 %), y la clasificación con CLASlite tiene un rango “Casi Perfecta” (88 %) respecto a los rangos del coeficiente de Kappa contrastados con los datos de exactitud. Adicionalmente, el método CLASlite presenta mayor exactitud en la categorización de bosque (84 %) y un menor porcentaje de error (16 %) respecto al método de clasificación supervisada (79 % y 21 % respectivamente).

RECOMENDACIONES

1. Ambos métodos evaluados dan resultados aceptables de clasificación sobresaliendo el método de clasificación con el software CLASlite, es importante aclarar que este estudio fue realizado en una sola región del país, para ello se recomienda realizar futuros ensayos en otras regiones del país con características biofísicas diferentes y categorías adicionales.
2. Antes de extender superficialmente su aplicación es importante resaltar que independiente, al método a utilizado es necesario realizar visitas de campo o garantizar la experiencia del intérprete que analizará el área además de validar la información con datos colectados en campo, según la disponibilidad de recurso económico y el tiempo.
3. Por su alto potencial como una alternativa para el monitoreo forestal, es recomendable generar un taller para su divulgación y que más usuarios puedan conocer del método y lo apliquen en otras regiones.

REFERENCIAS BIBLIOGRÁFICAS

1. Asner, G. (2014). CLASlite Project. *CLASlite Forest Monitoring*. California, USA.
2. Asner, G., Paez, G., Clark, J., Knapp, D., & Balaji, A. (2010). *CLASlite Sistema Amigable de Monitoreo de la Deforestación y Perturbación de Bosques Tropicales*. California, USA. Recuperado de: http://claslite.carnegiescience.edu/.static/themes/claslite/documents/es/Guia_del_Usuario_CLASlite.pdf
3. Asner, G., Tasar, E., Sousan, S., & Kanpp, D. (2013). *CLASlite Tecnología para el Monitoreo de Bosques. CLASlite Tecnología para el Monitoreo de Bosques*. Recuperado de: <ftp://dgc.stanford.edu/pub/clas/docs/Espa%F1ol/CLASlite%20Guia%20del%20Usario%20ESPA%D1OL.pdf>
4. Castellanos *et.al.* (2011). *Mapa de cobertura forestal de la República de Guatemala y dinámica forestal 200-2006*. Recuperado de: <http://glifos.concyt.gob.gt/digital/fodecyt/fodecyt%202009.28.pdf>
5. Castillejo, I. (2015). *Evaluación de métodos basados en píxeles y objetos para la clasificación de usos de suelo con imágenes de satélite*

Quickbird, para el seguimiento de medidas agroambientales y la optimización del uso de herbicidas con agricultura de precisión.

Recuperado de:
<http://helvia.uco.es/xmlui/bitstream/handle/10396/5326/9788469447611.pdf?sequence=1>

6. Chuvieco Salinero, E. (2008). *Teledeteccion ambiental La Observación de la Tierra desde el Espacio*. (3). Barcelona, España. Recuperado de: <https://books.google.com.gt/books?id=aKsNXCVcTcQC>
7. Chuvieco, E. (1995). *Fundamentos de Teledetección Espacial*. *Fundamentos de Teledetección Espacial*. (2). Madrid, España. Recuperado de:
<http://pdfhumanidades.com/sites/default/files/apuntes/FUNDAMENTOS-DE-TELEDETECCION-EMILIO-CHUVIECO.pdf>
8. Del Toro Espin, N., Gomariz, F., Canovas, F., & Alonso, F. (2013). *Comparación de métodos de clasificación de imágenes de satélite en la Cuenca del Río Argos*. Alicante, España. Recuperado de: <https://dialnet.unirioja.es/descarga/articulo/5035579.pdf>
9. Department of the Interior U.S. Geological Survey. (2017). *LANDSAT COLLECTION 1 LEVEL 1 PRODUCT DEFINITION*. Estados Unidos. Recuperado de:

https://landsat.usgs.gov/sites/default/files/documents/LSDS-1656_Landsat_Level-1_Product_Collection_Definition.pdf

10. Estrada, F. (2014). *GIS (Geographic Information System)*. Quito, Ecuador. Recuperado de: <http://repositorio.utn.edu.ec/bitstream/123456789/3631/7/capitulo%206.pdf>
11. FAO. (2013). *Sistemas Nacionales de Monitoreo de los Bosques*. Roma, Italia.
12. FCPF. (2015). *Niveles de Referencia en Costa Rica*. San Jose, Costa Rica.
13. Garcia, J. (2008). *Comparación de metodologías para el mapeo de la cobertura y uso del suelo en el sureste de México*. Recuperado de: http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0188-46112008000300002
14. Gonzales, R., & Woods, R. (2002). *Digital Image Processing*. (2). (P. E. Asia, Ed.) Michigan, Estados Unidos. Recuperado de: http://web.ipac.caltech.edu/staff/fmasci/home/astro_refs/Digital_Image_Processing_2ndEd.pdf
15. Gonzalez Aguayo, R. (s.f.). *Diccionario de terminos*. Buenos Aires, Argentina.

16. Guatemala, C. d. (1996). *Ley Forestal de Guatemala. Guatemala, Guatemala, Guatemala.*
17. Gúzman, J. (2015). *Análisis Multitemporal de las Zonas Forestales en la Zona INTAG Ecuador.* Recuperado de: http://issuu.com/unigis_latina/docs/guzman
18. INEGI. (2018). *Uso potencial de suelo.* Recuperado de: <http://www.beta.inegi.org.mx/temas/mapas/usopsuelo/>
19. Instituto de Geografía Universidad Autónoma de México. (2011). *Centro de Investigaciones en Ingeniería Ambiental.* (F. Bautista Zuñiga, J. Palacios Prieto, & H. Delfin Gonzales, Edits.) Recuperado de: http://www.ciga.unam.mx/publicaciones/images/abook_file/tmuestreo.pdf
20. International Standard. (2002). *ISO 19113 Geographic Information Quality Principles.* Suecia. Recuperado de: <https://www.sis.se/api/document/preview/903487/>
21. IPCC. (2006). *Guidelines for National Greenhouse Gas Inventories, Prepared by the National Greenhouse Gas Inventories Programme.* Japon.

22. Jena, F., Reyes, J., & Azucena, P. (2003). *Evaluación de la confiabilidad temática de mapas o de imágenes clasificadas*. Recuperado de: <http://www.revistas.unam.mx/index.php/rig/article/download/30414/28236>

23. Llactayo, W., Salcedo, K., & Ayala, E. (2013). *Evaluación de la Exactitud Temática del Mapa de Deforestación*. Recuperado de: <http://www.minam.gob.pe/ordenamientoterritorial/wp-content/uploads/sites/18/2013/10/Protocolo-Validacion-Mapa-Deforestacion.pdf>

24. Mejía Ramos, X. A., & Moncayo Silva, P. J. (2012). *Evaluación de métodos de clasificación de imágenes satelitales de moderada y alta resolución para la obtención de mapas de uso y cobertura del suelo basados en píxeles y orientados a objetos: Estudio de caso en sub escena de la parroquia Shushufindi, Prov.* Recuperado de: <http://repositorio.espe.edu.ec/handle/21000/5528>

25. Moreno, R., & Villon, C. (2012). *Detección de Zonas de Deforestación y Estimación de Biomasa en la Selva Peruana*. Recuperado de: <http://en.calameo.com/read/00197320997b277ab5669>

26. Olaya, V. (2014). *Sistemas de Información Geográfica*. Recuperado de: <http://volaya.github.io/libro-sig/index.html>

27. Olofsson, P. (2014). Good practices for estimating area and assessing accuracy of land change. Recuperado de: http://reddcr.go.cr/sites/default/files/centro-de-documentacion/olofsson_et_al._2014_-_good_practices_for_estimating_area_and_assessing_accuracy_of_land_change.pdf

28. Pacheco, C. (2011). *Analisis de la deforestacion en Venezuel*. Recuperado de: http://dspace.uah.es/dspace/bitstream/handle/10017/16601/Tesis_Carlos%20Pacheco_2011.pdf?sequence=1

29. Palacios, J. (1994). *Clasificación espectral automática vs. Clasificación visual: Un ejemplo al sur de la Ciudad de México*. Recuperado de: <http://www.scielo.org.mx/pdf/igeo/n29/n29a2.pdf>

30. Pascual, D. (2015). *Algoritmos de Agrupamiento basados en densidad y Validación de clusters*. Recuperado de: <http://www.cerpamid.co.cu/sitio/files/DamarisTesis.pdf>

31. Perez, C. (2006). *Teledeteccion: Nociones y aplicaciones*. Recuperado de: https://books.google.com.gt/books?id=SfrGxbO1DT0C&printsec=frontcover&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false

32. Perez, J. (2014). *Segmentación y clasificación de imágenes SAR en zonas de inundación en Colombia*. Recuperado de: <http://csifesvr.uan.edu.co/index.php/ingeuuan/article/download/266/pdf>
33. Piedra Fernandez, J. A. (2008). *Aplicacion de los sistemas Neurodifusos a la Interpretación Automatica de Imágenes de Satelite*. Almeria, España. Recuperado de: <https://books.google.com.gt/books?id=1Yw-AQAAQBAJ>
34. Pinto, F. (2012). *Geomática Tecnologías De Punta*. (Palibrio, Ed.) Recuperado de: <https://books.google.com.gt/books?id=tiEG2MFpqXwC>
35. Plata, W. (2010). *Descripcion, Analisis y Simulacion del Crecimiento Urbano Mediante Tecnologias de la Informacion Geografica*. Recuperado de: <http://dspace.uah.es/dspace/bitstream/handle/10017/9762/Tesis%20Final.pdf?sequence=1>
36. Plaza, A. (2015). *Proposición, Validación y Prueba de una Metodología Morfológica para el Análisis de Datos Hiperespectrales que Integra Información Espacial y Espectral*. Recuperado de: http://www.umbc.edu/rssi/pl/people/aplaza/Papers/Spanish/Thesis_aplaza.pdf

37. Quiros Rosado, E. M. (2009). *Clasificación de Imágenes Multiespectrales ASTER mediante funciones adaptativas*. Cáceres, España. Recuperado de: http://dehesa.unex.es/bitstream/handle/10662/457/TDEUX_9788469263167.pdf?sequence=1&isAllowed=y
38. Reuter, A. (2002). *Teledetección Forestal*. Teledetección Forestal. Buenos Aires, Argentina. Recuperado de: <http://fcf.unse.edu.ar/archivos/lpr/pdf/p11.PDF>
39. Rodríguez, A. (2011). *Metodología para detectar cambios en el uso de la tierra utilizando los principios de la clasificación orientada a objetos, estudio de caso piedemonte de Villavicencio, Meta*. Recuperado de: <http://www.bdigital.unal.edu.co/5241/1/andresfeliperodriguezvasquez.2011.pdf>
40. Santos, C. (2007). *Extracción de atributos de forma y selección de atributos usando algoritmos genéticos para clasificación de regiones*. San Pablo, Brasil. Recuperado de: <http://mtc-m16b.sid.inpe.br/col/sid.inpe.br/mtc-m17@80/2007/04.04.16.56/doc/publicacao.pdf>
41. Sanz, C. (2015). *Un Método de Clasificación Aplicado al Análisis de Imágenes Hiperespectrales*. Recuperado de: http://sedici.unlp.edu.ar/bitstream/handle/10915/2213/Documento_completo.pdf?sequence=14

42. Stehman, S., & Czaplewski, R. (1998). *Design and Analysis for Thematic Map Accuracy Assessment: Fundamental Principles*. Recuperado de:
http://www.fs.fed.us/rm/pubs_other/rmrs_1998_stehman_s001.pdf
43. Tiscornia, G. (2011). *Análisis temporal del uso del suelo en el Departamento de Soriano*. Recuperado de:
<http://ambiente.fcien.edu.uy/tesis/Guadalupe%20Tiscornia.pdf>
44. Unidad Administrativa Especial de Catastro Distrital (UAECD). (2013). *Procedimiento para Evaluar y Reportar Datos Espaciales*. Recuperado de: https://www.ideca.gov.co/sites/default/files/PPIG-06-Procedimiento_Evaluar_Reportar_Calidad_Datos_Espaciales_V2%201_2013.pdf
45. Universidad de Murcia. (2006). *SIG y Teledetección en la Universidad de Murcia*. Recuperado de:
http://www.etisig.catamarca.gov.ar/index.php?option=com_content&view=article&id=86&Itemid=103
46. Velasquez, M. (2010). *Estimación del Volumen de Bosques de Pino en Honduras, utilizando imágenes del sensor MODIS, como herramienta para ordenación del territorio*. Recuperado de:
http://faces.unah.edu.hn/mogt/images/stories/PDF/Tesis/07_Tesis_Miguel_Velasquez_2010.pdf

ANEXOS

DESCRIPCIÓN DE LOS CÓDIGOS

CATEGO = Categoría general

SUB_CATEGO = Clasificación de acuerdo a las características generales.

CATEGORÍA DE NO BOSQUE

SUB_CATEGO = PASTOS_GRANOSBA

Áreas del altiplano que va desde San Marcos, Quetzaltenango, Totonicapán hasta los Cuchumatanes, que son las regiones más secas del área de estudio en la que predominan pastos y granos básicos. Desde la parte más húmeda que está del otro lado de los Cuchumatanes, buscando la frontera norte de Huehuetenango, Barillas.

SUB_CATEGO = CAFÉ_SINSOMBRA

Esta sub categoría es de la misma manera para la parte de la costa de San Marcos, y el centro de Huehuetenango, donde la estación seca es más marcada a diferencia de la parte más húmeda de Huehuetenango, donde hay más café bajo sombra.

SUB_CATEGORÍA = CAFÉ_CONSOMBRA

Estos puntos fueron tomados en la parte norte de Huehuetenango, principalmente en el área de Barillas, ya que allí es donde hay más presencia de humedad y el tipo de cultivo es más diverso y se encuentran más especies de árboles.

SUB_CATEGORÍA = SUELO ROCOSO

Estos puntos que no fueron muchos corresponden a las rocas, que se encuentran en la meseta de los Cuchumatanes y algunas faldas de la misma cordillera.

SUB_CATEGORÍA = PASTO_GRA_SUEHUM

Esta es la misma categoría igual primera de pastos y granos básicos, con la diferencia que está se encuentra en la parte norte de Huehuetenango, desde San Juan Ixcoy hasta Barillas, debido a que existe una mayor cantidad de humedad.

SUB_CATEGORÍA = ZONAS_URBA

Se tomaron puntos de las zonas urbanas para tener una representación de categoría, que no fueron muchos pero que si existen zonas de importancia como Huehuetenango, San Marcos, Parte de Xela y Totonicapán.

CATEGORÍA DE BOSQUE

SUB_CATEGO = BOS_LATI_DEN

Estos puntos fueron los primeros que se tomaron, se encuentran al norte de Huehuetenango en su mayoría, ya que es una de las zonas muy húmedas, donde hay una gran cantidad de estos bosques todavía con características muy particulares.

SUB_CATEGO = BOS_CON_DEN

Esta es otra categoría de bosque en su mayoría de la misma región, donde hay una densidad muy alta y también existen bosque de una edad adulta.

SUB_CATEGO = BOS_LAT_NODEN

Estos son puntos de la misma región anterior, que se clasificó así por su poca densidad, pero que sí tienen presencia en la misma área y que muchos de ellos son bosque degradados.

SUB_CATEGO = BOS_MIX_DEN

Estos son puntos que fueron tomados en su mayoría en la parte media en regiones húmedas que la región fisiográfica ha marcado la combinación de los dos tipos de bosque y que tienen características de ser muy denso.

SUB_CATEGORÍA = BOS_CON

Estos puntos fueron tomados en la parte sur del área de estudio, donde existe una mayor presencia de coníferas, su densidad es significativa, que no es tan notoria como la subcategoría anterior, pero que tampoco está tan degradada como una categoría más adelante, que se encuentra en la parte media y oriental del área de estudio.

SUB_CATEGORÍA = BOS_MIX

De igual manera que la subcategoría anterior es una clasificación de bosque mixto menos denso que el anterior, por lo general está más en la parte sur del área de estudio donde la humedad es menor que en la parte norte y los bosques han tenido un mayor grado de intervención.

SUB_CATEGORÍA = BOS_MIX_CON_RALO

Estos puntos se refieren más a una zona que está siendo muy degradada y muy seca, que tienen en su mayoría bosque de conífera con una combinación de Quercus en menor proporción se encuentra en los límites de los Cuchumatanes buscando Totonicapán.

SUB_CATEGORÍA = BOS_CON_RALO

Esta es otra categoría que se quiso subdividir, ya que está en la zona húmeda a seca de la parte de Nentón y los Huistas del departamento de Huehuetenango, donde son bosques de coníferas, pero de una densidad más rala que los que están en la parte más alta.

SUB_CATEGO =BOS_DE_RIBERA

Esta es una categoría pequeña que solo se tomaron 8 puntos, ya que está ubicada en la zona seca del departamento de Huehuetenango, especialmente en el municipio de Nentón, pero que si es notorio el bosque en las riberas de los ríos que bajan de la parte alta de los Cuchumatanes.