



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

LICENCIATURA EN TECNOLOGÍAS PARA LA INFORMACIÓN EN CIENCIAS

Escuela Nacional de Estudios Superiores,
Unidad Morelia

MONITOREO DE LA DEGRADACIÓN FO-
RESTAL POR MEDIO DE ANÁLISIS DE SERIE
DE TIEMPO (2010-2020) DE ÍNDICES DE VE-
GETACIÓN, HUMEDAD Y FRACCIÓN DERI-
VADOS DEL SENSOR LANDSAT

TESIS

QUE PARA OBTENER EL TÍTULO DE

LICENCIADO EN TECNOLOGÍAS PARA LA
INFORMACIÓN EN CIENCIAS

P R E S E N T A

DANIEL DELGADO MORENO

DIRECTORA DE TESIS: DRA. YAN GAO

MORELIA, MICHOACÁN

MARZO, 2022



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

LICENCIATURA EN TECNOLOGÍAS PARA LA INFORMACIÓN EN CIENCIAS

Escuela Nacional de Estudios Superiores,
Unidad Morelia

MONITOREO DE LA DEGRADACIÓN FO-
RESTAL POR MEDIO DE ANÁLISIS DE SERIE
DE TIEMPO (2010-2020) DE ÍNDICES DE VE-
GETACIÓN, HUMEDAD Y FRACCIÓN DERI-
VADOS DEL SENSOR LANDSAT

TESIS

QUE PARA OBTENER EL TÍTULO DE

LICENCIADO EN TECNOLOGÍAS PARA LA
INFORMACIÓN EN CIENCIAS

P R E S E N T A

DANIEL DELGADO MORENO

DIRECTORA DE TESIS: DRA. YAN GAO

MORELIA, MICHOACÁN

MARZO, 2022



ESCUELA
NACIONAL
DE ESTUDIOS
SUPERIORES
mm
UNIDAD MORELIA

10
años
(2011-2021)

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
ESCUELA NACIONAL DE ESTUDIOS SUPERIORES UNIDAD MORELIA
SECRETARÍA GENERAL
SERVICIOS ESCOLARES

MTRA. IVONNE RAMÍREZ WENCE

DIRECTORA

DIRECCIÓN GENERAL DE ADMINISTRACIÓN ESCOLAR

PRESENTE

Por medio de la presente me permito informar a usted que en la **sesión ordinaria 11** del **H. Consejo Técnico** de la Escuela Nacional de Estudios Superiores Unidad Morelia, celebrada el día **01 de diciembre del 2021**, se acordó poner a su consideración el siguiente jurado para la presentación del Trabajo Profesional del alumno **Daniel Delgado Moreno** adscrito a la Licenciatura en Tecnologías para la Información en Ciencias, con número de cuenta **416109502**, quien presenta la tesis titulada: "Monitoreo de la degradación forestal por medio de análisis de serie de tiempo (2010-2020) de índices de vegetación, humedad y fracción derivados del sensor Landsat", bajo la dirección como **tutora** de la Dra. Yan Gao.

El jurado queda integrado de la siguiente manera:

Presidente:	Dr. Jean Francois Mas Causel
Vocal:	Dr. Víctor Hugo De la Luz Rodríguez
Secretario:	Dra. Yan Gao
Suplente 1:	Dr. Iván Franch Pardo
Suplente 2:	Dra. Michelle Farfán Gutiérrez

Sin otro particular, quedo de usted.

Atentamente
"POR MI RAZA HABLARA EL ESPIRITU"
Morelia, Michoacán a 14 de marzo del 2022.



DRA. YESENIA ARREDONDO LEÓN
SECRETARIA GENERAL

CAMPUS MORELIA

Antigua Carretera a Pátzcuaro N° 8701, Col. Ex Hacienda de San José de la Huerta
58190, Morelia, Michoacán, México. Tel: (443)689.3500 y (55)56.23.73.00, Extensión Red UNAM: 80614

www.enesmorelia.unam.mx

AGRADECIMIENTOS INSTITUCIONALES

Agradezco a la Universidad Nacional Autónoma de México por permitirme formar parte de esta gran casa de estudios, a la Escuela Nacional de Estudios Superiores, Unidad Morelia por todos los recursos y apoyos otorgados, y a la Licenciatura en Tecnologías para la Información en Ciencias por haberme brindado las herramientas y enseñanzas que me servirán para mi desarrollo profesional.

Agradezco a la Dra. Yan Gao, por confiar en mí y por permitirme trabajar con ella en la elaboración de este proyecto, además por siempre estar pendiente de mis avances y por apoyarme en mis dudas, brindándome sugerencias y en algunos momentos dándome ánimos para seguir trabajando. Al Dr. Víctor de la Luz por todo el apoyo brindado desde que inicie el proceso de la tesis hasta prácticamente la culminación de esta.

A todos mis profesores de la licenciatura por sus enseñanzas, apoyo y el tiempo brindado, a los profesores que estuvieron conmigo en la especialidad de ciencias de la Tierra, porque gracias a ellos logre confirmar mi gusto por el área.

Del mismo modo agradezco a los miembros del jurado, el Dr. Jean François, el Dr. Víctor de la Luz, la Dra. Yan Gao, el Dr. Iván Franch y la Dra. Michelle Farfán por aceptar revisar mi trabajo, así como por las aportaciones y sugerencias para mejorarlo.

AGRADECIMIENTOS PERSONALES

Agradezco a mi familia, en especial a mi mamá que siempre de alguna manera estuvo para apoyarme en mis estudios y crecimiento personal, agradezco a mi hermana por apoyarme cuando la necesito, a mis dos sobrinos, Diego y Vanessa por todos los momentos de alegría que me han regalado.

A mis amigos y a los compañeros que se volvieron amigos en el transcurso de la carrera, quienes siempre me dieron ánimos para no rendirme a pesar de las complicaciones académicas que presenté en el transcurso de esta y con quienes compartí grandes momentos dentro y fuera de las aulas.

A los profesores que en algún momento se tomaron el tiempo de ayudarme con mis dudas, problemas, que me motivaron a culminar con la carrera, que me dieron ánimos de no rendirme y por los que siempre creyeron en mí.

RESUMEN

El presente trabajo tiene el objetivo de estimar la degradación forestal en la cuenca del río Ayuquila a través de un análisis de tendencia de vegetación utilizando índices de vegetación de series de tiempo anuales (2010-2020) específicamente, el NDVI (índice de vegetación de diferencia normalizada), el NDMI (índice de humedad de diferencia normalizada) y el NDFI (índice de fracción de diferencia normalizada) derivados de imágenes de Landsat-5 TM y Landsat-8 OLI. Las imágenes con las cuales se trabajó en este proyecto fueron obtenidas de la plataforma llamada EarthExplorer, en total se obtuvieron 10 imágenes anuales desde 2010 hasta 2020 entre los meses de marzo y abril ya que se observó que en este periodo había menor nubosidad. Todas las imágenes fueron capturadas por satélites Landsat colección 1, donde 8 fueron capturadas por el satélite Landsat-8 OLI y 2 por Landsat-5 TM, las cuales fueron corregidas radiométrica y atmosféricamente con el fin de obtener el valor real de reflectancia. Este proceso de corrección se realizó mediante la aplicación del método Dark Object Substracion (DOS) mediante el uso de librerías como RStolbox del lenguaje R. Cabe mencionar que, para la obtención del NDFI, a diferencia del NDVI y NDMI, se requirió como insumo las imágenes fracción de suelo, sombra, vegetación y vegetación no fotosintética por lo que fue necesario aplicar antes el algoritmo de Spectral Mixture Analysis (SMA) el cual tiene el objetivo de descomponer los píxeles en las diferentes clases de las que se integran.

El análisis de la tendencia de la vegetación se llevó a cabo utilizando un modelo de regresión lineal para cada índice, mismos que fueron probados con la prueba de Mann-Kendall para determinar la significancia de los resultados. Se utilizó el coeficiente de pendiente para indicar la tendencia de la vegetación: una pendiente positiva indica aumento de vegetación mientras que una pendiente negativa indica la degradación de esta. Para la degradación forestal, solo se analizaron las tendencias significativas con pendiente negativa ($p < 0.05$). Además, para descartar tendencias negativas por deforestación, se aplicó una máscara de bosque para asegurarse de que las áreas degradadas detectadas estuvieran en la categoría de bosque al principio y al final del análisis. La máscara fue obtenida gracias a la aplicación de los métodos de

clasificación de Máxima Verosimilitud y Random Forest, donde para este estudio se obtuvieron mejores resultados con el primero.

Para validar los resultados obtenidos se aplicó un muestreo aleatorio estratificado donde el número de muestras se determinó mediante cálculos estadísticos. La evaluación de la precisión con un 95% de confianza, mostró que la estimación de la degradación forestal por series de tiempo del NDVI obtuvo la mayor precisión general con un $81.33\% \pm 6.23\%$, seguida por NDMI con $73.33\% \pm 7.07\%$ y finalmente NDFI con $72\% \pm 7.18\%$. El uso del NDVI como índice para medir la degradación forestal permitió identificar que un total de 191 hectáreas en la cuenca de Ayuquila presentaron degradación forestal.

ABSTRACT

The present work has the objective of estimating forest degradation in the Ayuquila river basin through a vegetation trend analysis using annual time series vegetation indexes (2010-2020), specifically, the NDVI (Normalized Difference Vegetation Index), NDMI (Normalized Difference Moisture Index), and NDFI (Normalized Difference Fraction Index) derived from Landsat-5 TM and Landsat-8 OLI images. The images with which I worked on this project were obtained from the platform called EarthExplorer, in total 10 annual images were obtained from 2010 to 2020 between the months of March and April since it was observed that there was less cloudiness in this period. All images were captured by Landsat collection 1 satellites, where 8 were captured by Landsat-8 OLI satellite and 2 by Landsat-5 TM, which were radiometrically and atmospherically corrected in order to obtain the real value of reflectance. This correction process was carried out through the application of the Dark Object Substration (DOS) method through the use of libraries such as RStolbox of the R language. It is worth mentioning that, in order to obtain the NDFI, unlike NDVI and NDMI, required as input the fraction images of soil, shade, vegetation and non-photosynthetic vegetation, so it was necessary to apply the Spectral Mixture Analysis (SMA) algorithm beforehand, which has the objective of decomposing the pixels into the different classes of which are integrated.

The vegetation trend analysis was carried out using a linear regression model for each index, which were tested with the Mann-Kendall test to determine the significance of the results. The slope coefficient was obtained to indicate the vegetation trend: a positive slope indicates an increase in vegetation while a negative slope indicates its degradation. For forest degradation, only significant trends with a negative slope ($p < 0.05$) are analyzed. Additionally, to rule out negative trends due to deforestation, a forest mask was applied to ensure that detected degraded areas were in the forest category at the beginning and end of the analysis. The mask was obtained thanks to the application of the Maximum Likelihood and Random Forest classification methods, where for this study better results were obtained with the last one.

To validate the results obtained, a stratified random sampling was applied where the number of samples was extended by means of statistical calculations. The evaluation of precision with 95% confidence, showed that the estimation of forest degradation by time series of the NDVI obtained the highest general precision with $81.33\% \pm 6.23\%$, followed by NDMI with $73.33\% \pm 7.07\%$ and finally NDFI with $72\% \pm 7.18\%$. The use of the NDVI as an index to measure forest degradation made it possible to identify that a total of 191 hectares in the Ayuquila basin presented forest degradation

ÍNDICE

CAPÍTULO 1. INTRODUCCIÓN **13**

1.1 ANTECEDENTES	13
1.1.1 CONTRIBUCIÓN DE LA PERCEPCIÓN REMOTA EN LA MEDICIÓN DE LA DEGRADACIÓN FORESTAL	14
1.1.2 CAUSAS DE LA DEGRADACIÓN FORESTAL	15
1.1.2.1 Degradación forestal por tala selectiva y quema forestal	16
1.1.2.2 Degradación forestal por sistema de cultivo roza, tumba y quema	17
1.1.2.3 Degradación forestal por recolección de leña y ganadería	18
1.2 MARCO TEÓRICO	19
1.2.1 USO Y CUBIERTA DE SUELO	19
1.2.2 LA DEGRADACIÓN FORESTAL	19
1.2.3 CAMBIO CLIMÁTICO	20
1.2.4 BOSQUES Y CAMBIO CLIMÁTICO	21
1.2.5 UNFCCC Y REDD+	21
1.2.6 PERCEPCIÓN REMOTA	22
1.2.7 ANÁLISIS DE MEZCLAS ESPECTRALES (SMA)	24
1.2.8 DEGRADACIÓN FORESTAL Y EL MÉTODO SMA	24
1.2.9 ANÁLISIS DE SERIE DE TIEMPO	29
1.2.10 ÍNDICES ESPECTRALES PARA MEDIR LA VEGETACIÓN	32
1.2.10.1 Índice de vegetación de Diferencia Normalizada (NDVI)	33
1.2.10.2 Índice de Diferencia Normalizada de humedad (NDMI)	33
1.2.10.3 Índice de Fracción de Diferencias Normalizadas (NDFI)	34
1.3 JUSTIFICACIÓN	36
1.4 OBJETIVO	37
1.4.1 OBJETIVO GENERAL	37
1.4.2 OBJETIVOS PARTICULARES	38

CAPÍTULO 2. MATERIALES Y MÉTODOS **39**

2.1 ÁREA DE ESTUDIO	39
2.2 DATOS	40
2.3 ANÁLISIS DE IMÁGENES SATELITALES	45
2.3.1 ADQUISICIÓN DE LA IMAGEN	46
2.3.2 PREPROCESAMIENTO	46
2.3.2.1 Corrección geométrica	46
2.3.2.2 Corrección atmosférica	47
2.3.3 CLASIFICACIÓN DE IMÁGENES	51
2.3.3.1 Clasificación no supervisada	52
2.3.3.2 Clasificación supervisada	53
2.3.4 INDICES DE VEGETACIÓN	55
2.3.4.1 NDVI	55

2.3.4.2	NDMI	56
2.3.4.3	NDFI	57
2.3.5	ANÁLISIS DE MEZCLAS ESPECTRALES (SMA)	58
2.3.6	ANÁLISIS DE LA TENDENCIA DE VEGETACIÓN CON EL MODELO DE REGRESIÓN LINEAL	60
2.3.6.1	Modelo de Regresión Lineal Simple	60
2.3.6.2	Pendiente de la recta y coeficiente de determinación	61
2.3.6.3	Prueba de Mann-Kendall	62
2.4	VERIFICACIÓN DE RESULTADOS	63
2.4.1	DISEÑO DE MUESTRAS	64
2.4.2	NÚMERO DE MUESTRAS	64
2.4.3	MATRIZ DE CONFUSIÓN	65
2.4.4	CONFIABILIDAD GLOBAL	66
2.4.5	CONFIABILIDAD DE USUARIO Y DE PRODUCTOR	67
2.4.6	INTERVALO DE CONFIANZA	68
2.4.7	EJEMPLO DEL CÁLCULO DE CONFIABILIDAD	68

CAPÍTULO 3. ANÁLISIS DE DATOS **70**

3.1	ADQUISICIÓN DE DATOS	70
3.2	PRE-PROCESAMIENTO DE DATOS	73
3.2.1	CORRECCIÓN GEOMÉTRICA	73
3.3	RECORTE Y SELECCIÓN DE ZONA DE ESTUDIO	74
3.3.1	CONVERSIÓN DE NÚMERO DIGITAL A RADIANCIA	75
3.3.2	CONVERSIÓN DE NÚMERO DIGITAL A REFLECTANCIA	76
3.3.3	CORRECCIÓN ATMOSFÉRICA POR EL MÉTODO DE DARK OBJECT SUBTRACTION (DOS)	76
3.4	SMA Y NDFI	81
3.4.1	IDENTIFICACIÓN DE MIEMBROS PUROS	81
3.4.2	LA APLICACIÓN DE SMA EN R	84
3.4.3	RESULTADOS DEL SMA	85
3.4.4	INDICES NDFI	86
3.4.5	CÁLCULO DE ÍNDICES NDVI Y NDMI	87
3.5	ANÁLISIS DE LA TENDENCIA DE VEGETACIÓN CON EL SERIE DE TIEMPO Y EL MODELO DE REGRESIÓN LINEAL	89
3.6	APLICACIÓN DE MÁSCARA	97
3.6.1	MAPEO DE LA COBERTURA DE SUELO CON CLASIFICACIÓN SUPERVISADA DE IMAGEN LANDSAT	97
3.6.2	EVALUACIÓN DE CONFIABILIDAD DE LA IMAGEN CLASIFICADA	106

CAPÍTULO 4. VERIFICACIÓN DE RESULTADOS **114**

4.1	DISEÑO DE MUESTRAS	114
4.2	MATRIZ DE CONFUSIÓN	114
4.3	CÁLCULO DE CONFIABILIDAD	120

CAPÍTULO 5. DISCUSIÓN **122**

5.1	COMPARACIÓN CON OTROS PROYECTOS	122
5.2	RESULTADOS OBTENIDOS NDVI Y NDMI	125
5.3	RESULTADOS OBTENIDOS POR EL NDFI	126
5.4	APLICACIÓN DE MÁSCARA	127
5.5	ANÁLISIS DE RESULTADOS OBTENIDOS POR EL NDVI	129
5.6	DISTRIBUCIÓN DE ÁREAS IDENTIFICADAS COMO DEGRADACIÓN	131
5.7	COMPARACIÓN DE RESULTADOS CON GLOBAL FOREST WATCH	133
CONCLUSIÓN		139
BIBLIOGRAFÍA		141

CAPÍTULO 1. INTRODUCCIÓN

1.1 ANTECEDENTES

Los bosques ocupan casi un tercio de la superficie terrestre, cubriendo casi cuatro mil millones de hectáreas. Un tercio de las áreas forestales actuales están formadas por bosque que se considera intacto mientras que los dos tercios restantes están sujetos a actividad antropogénica, es decir que el hombre ha intervenido en su desarrollo y como resultado su estabilidad tiene un futuro incierto. Cada año, en todo el mundo varios millones de hectáreas de bosque desaparecen debido a la deforestación y cambio de uso de suelo y otras tantas están sujetas a degradación (Blanc, et al., 2016).

La deforestación y la degradación forestal son procesos que generan una afectación en los bosques, sin embargo, ambos tienen características diferentes. Por un lado, la deforestación se refiere a la remoción completa de la cubierta forestal, es decir, áreas donde antes era bosque pasan a ser un área con otro tipo de uso de suelo, como área urbana, áreas de cultivo, suelo desnudo, etc. mientras que la degradación forestal se refiere a la remoción parcial o desgaste de la cubierta forestal, en otras palabras, áreas donde ocurre degradación forestal sigue siendo bosque, pero ya no conserva la misma calidad de antes. Esto generalmente sucede por la tala inmoderada de árboles, por incendios forestales, entre otros (Blanc, et al., 2016).

La conservación y cuidado de los bosques es de vital importancia para todos los seres que habitamos el planeta. La importancia de su cuidado no solo radica en la diversidad biológica, sino que los bosques también funcionan como regulador de temperatura en todo el planeta ya que estos pueden absorber grandes cantidades de dióxido de carbono (CO₂) el cual es un gas de efecto invernadero generado mayormente por actividades antropogénicas, esto tiene como una de sus principales consecuencias el cambio climático (Comisión Nacional Forestal, 2017).

Cuando los bosques son destruidos o sufren procesos de degradación, en particular cuando son quemados, se libera el carbono que estos habían absorbido y almacenado, estos procesos son uno de los principales causantes de las emisiones de gases de efecto invernadero que provocan a su vez el cambio climático (CONAFOR, 2013).

La degradación forestal contribuye en gran medida a las emisiones de gases de efecto invernadero sobre todo en áreas tropicales. Actualmente en México está en marcha el Programa REDD+ (Reducción de emisiones de gases efecto invernadero por deforestación, degradación y buen manejo de bosques) que busca la conservación de los acervos forestales contribuyendo de esta forma con el control de los gases de efecto invernadero. La degradación forestal se debe a cambios que ocurren dentro de los ecosistemas y que afectan negativamente la calidad de estos.

Es por ello por lo que tener información de las afectaciones forestales es de suma importancia ya que nos ayuda a tener un mejor control y cuantificación de los daños, de forma tal que se puedan aplicar medidas para el cuidado de los bosques. Tradicionalmente el monitoreo forestal se hacía mediante trabajo de campo, el cual usualmente requiere una inversión importante de recursos, además de la imposibilidad de abarcar grandes espacios y muchas veces la inaccesibilidad a algunas áreas. Sin embargo, en la actualidad existen novedosas técnicas para medir los cambios en el uso de suelo, como por ejemplo la teledetección o también llamada percepción remota la cual nos permite visualizar a través de imágenes satelitales estos cambios (Bowen, 2011).

1.1.1 Contribución de la percepción remota en la medición de la degradación forestal

La percepción remota es una herramienta muy importante y útil para medir la deforestación y la degradación forestal debido a su amplia cubierta y por la rápida periodicidad de observación de los sensores remotos, sin embargo, es más fácil detectar la deforestación que la degradación forestal mediante técnicas de percepción remota, esto debido a que en la deforestación se tiene un cambio de cubierta o uso de suelo que causa que la firma espectral de las zonas que presentan deforestación cambien y por ende sean detectadas relativamente fácil. Desafortunadamente, no sucede lo mismo en la detección de la degradación forestal

porque en sí, no se tiene un cambio de cubierta ni de uso de suelo, los cambios siguen siendo dentro del mismo bosque (por ejemplo, disminución de biomasa, cubierta forestal, en general se presenta una disminución de la calidad del bosque) por lo que la firma espectral de estas áreas puede variar muy poco y perderse entre las áreas de bosque no afectadas provocando que la identificación de las áreas degradadas sea más compleja.

A pesar de que ya se conocen las consecuencias en que afecta la degradación forestal, son pocos los estudios que se realizan sobre la identificación de esta problemática, aún en la actualidad es muy poco frecuente encontrar trabajos enfocados al monitoreo de la degradación forestal, la mayoría de los estudios se centran en identificar la deforestación. Esto se debe principalmente a que en el monitoreo la degradación forestal a diferencia del monitoreo de la deforestación, se tienen mayores dificultades a la hora de identificar cambios mediante técnicas de percepción remota. Esto como se mencionó anteriormente es porque, el área degradada aún permanece como bosque, pero con alguna alteración en su composición y función, lo cual dificulta en gran medida su observación. También hay que tener en cuenta que para un correcto monitoreo de la degradación forestal por medio de herramientas de percepción remota hay que considerar además de las áreas afectadas, la cantidad de biomasa perdida en dichas áreas, esto aumenta en cierto grado la dificultad en el análisis de la degradación, en comparación con la deforestación, ya que, en esta última, únicamente se toma en cuenta las áreas que han sido despejadas, es decir que ya no son más bosque. Otro factor importante a tomar en cuenta al monitorear la degradación forestal es el sistema de percepción remota a utilizar, el más común es Landsat, sin embargo presenta algunas limitaciones, y es que su resolución espacial es de 30 metros, lo cual dificulta la identificación de la degradación cuando esta ocurre en áreas pequeñas, y aunque existen otras imágenes con mejor resolución, estas resultan ser muy costosas y de difícil acceso, además de que suelen representar áreas más pequeñas, con lo que se requerirían mayor cantidad de imágenes a utilizar (Gao, et al., 2020).

1.1.2 Causas de la degradación forestal

Las causas de la degradación forestal son muy variadas, pueden ser ocasionadas por procesos naturales o por procesos antropogénicos, siendo estas últimas las más comunes. Algunos ejemplos pueden ser la tala inmoderada de los bosques, tala selectiva, incendios forestales, agricultura, sistemas de cultivo roza-tumba-quema, recolección de leña, por mencionar las más comunes. También si hablamos de causas naturales podemos mencionar las plagas, enfermedades, huracanes, y deslizamientos. Todas estas causas ocasionan daños en el ecosistema y provocan a su vez, aumentar la degradación forestal (CONAFOR, 2013). Por ejemplo, Souza et al. (2013) menciona que la tala selectiva y la quema se consideran las principales causas de la degradación forestal en los bosques de tierras altas del Amazonas y que incluso pueden provocar un daño mucho mayor que la deforestación. A continuación, se presentan algunos de los estudios que se han realizado con base en diferentes causas de degradación forestal.

1.1.2.1 Degradación forestal por tala selectiva y quema forestal

Souza et al. (2002) utilizan imágenes SPOT 4 para estudiar la degradación en la Amazonia la cual es mayormente causada por tala selectiva e incendios forestales. La tala selectiva degrada los bosques durante las actividades de cosecha, causando daños extensos a los árboles y suelos cercanos, aumentando el riesgo de extinción de especies, así como aumentando las emisiones de carbono. Además, la extracción de madera aumenta la susceptibilidad al fuego y los caminos abiertos, lo que fomenta el desarrollo no planificado. Para analizar la degradación forestal se definieron cuatro clases, bosque intacto, bosque talado, bosque degradado (por cuestiones de quemados) y bosque en regeneración. Para la clasificación de la imagen SPOT 4 los autores diseñaron un árbol de decisión definiendo las reglas que mejor separaban las clases. El resultado de este estudio mostró una confiabilidad del 86%, identificando el 35% del área forestal como intacta, 56% como área talada, y el 9% como áreas forestales degradadas. Estos resultados fueron comparados visualmente con otra imagen IKONOS llegando a tener una concordancia muy aceptable. La imagen IKONOS tiene una resolución de un metro, la cual es suficiente para realizar interpretaciones visuales.

En un estudio más reciente de Souza et al. (2013) donde utilizan datos de Landsat para cuantificar la deforestación y degradación forestal en la Amazonia desde el año 2000 hasta el 2010, en el cual aplican diferentes métodos como el análisis de mezclas espectrales, índice de fracción normalizada y árboles de decisión, logrando una precisión global del 92%, superior al 86% del estudio de Souza et al. (2002). Demostraron que 169,074 km² de extensión pasaron de ser bosque a zonas de agricultura, además 50,815 km² fueron sujetos a extracción de madera y/o incendios forestales, esto en tan solo 10 años. Un dato importante es que durante este mismo periodo de tiempo la deforestación anual disminuyó en un 46% mientras que la deforestación aumentó en un 20%.

1.1.2.2 Degradación forestal por sistema de cultivo roza, tumba y quema

Otra de las causas de degradación forestal es la antigua práctica de agricultura, roza, tumba y quema (RTQ), que en el contexto de cambio climático ha sido vista como una de las responsables de la deforestación y degradación forestal. Esta práctica genera grandes emisiones de gases de efecto invernadero. Dentro de un esquema de REDD+ la agricultura de roza tumba y quema debe ser considerada como degradación en lugar de deforestación. Los almacenes de carbono total de sitios que están en el sistema agrícola de RTQ en las selvas bajas caducifolias, pueden tener en promedio mucho más carbono que parcelas con agricultura permanente, y además sólo un 10% menos que los bosques con mejores condiciones de conservación (Paz Pellat, et al., 2015).

Borrego et al. (2015) mencionan que una de las principales causas de la degradación forestal son las actividades económicas como lo es el sistema de RTQ, ante esto se sugiere cambiar este tipo de agricultura a uno de agricultura permanente. Este estudio fue realizado en dos comunidades, Tonaya y Temazcal pertenecientes a la cuenca del río de Ayuquila en Jalisco, México, y toma un punto de vista económico, identificando el impacto económico que tiene la degradación forestal en las personas que dependen de los recursos naturales que ofrece el bosque. Su objetivo es estimar el costo que tienen las emisiones de carbono provocadas por el sistema de roza, tumba y quema. Este trabajo se realizó meramente con estudios de

campo y los resultados que se obtuvieron indican que las emisiones de carbono producidas por el sistema de RTQ son ligeramente superiores que el de la agricultura permanente.

Los sistemas de RTQ han sido usados por mucho tiempo, esto debido a que es un método de cultivo efectivo que permite la recuperación natural de la fertilidad del suelo durante la etapa de descanso, sin embargo, en los últimos años el constante uso de este sistema y la reducción del tiempo de descanso han acelerado la pérdida de nutrientes del suelo de los bosques y selvas donde se realizan estas prácticas. Si bien este método puede tener beneficios es importante no abusar de él y hacerlo con el debido cuidado para asegurar que la calidad de los bosques o áreas donde se practique este método sigan teniendo una calidad aceptable (García & Maass, 1998).

1.1.2.3 Degradación forestal por recolección de leña y ganadería

Las causas de la degradación forestal pueden ser muy variadas, algunas veces ocurre de forma muy abrupta como los casos de incendios o explotación forestal, sin embargo, en muchas ocasiones estas afectaciones ocurren en largos periodos de tiempo, de forma lenta y gradual, tal es el caso de la práctica de recolección de leña y ganadería. Es importante saber que, aunque este tipo de afectaciones sea lento, también es constante y siguiendo con estas prácticas a la larga provoca afectaciones muy severas a la salud forestal. Además, al ocurrir en lapsos de tiempo muy largos es más difícil detectar este tipo de daños a través de herramientas tecnológicas como la percepción remota. Esto a diferencia de otros tipos de degradación que ocurren muy repentinamente los cuales son vistos más fácilmente. Generalmente las actividades de recolección de leña y ganadería son realizadas por los propios habitantes de las áreas forestales o cercanas a estas, dichas actividades son realizadas con fines de su propio sustento y tal vez sin el afán de dañar la cobertura, sin embargo, el no poner atención en este tipo de actividades es perjudicial para los bosques y por consiguiente para las mismas personas que habitan en las áreas cercanas. Las consecuencias de este tipo de prácticas provocan tener bosques fragmentados donde las áreas abiertas muchas veces son ocupadas para el ganado evitando la regeneración de la cubierta foresta (Rojas, et al., 2012).

1.2 MARCO TEÓRICO

Detectar cambios en el uso de suelo es una tarea que se ha venido haciendo desde hace algunos años. Esto debido a la importancia de medir las afectaciones que se ocasionan, no solo ambientales sino también sociales y económicas. Muchos de los estudios que se realizan se enfocan en diferentes aspectos, por ejemplo, detectar crecimiento de áreas urbanas o agrícolas, pérdida de área forestal, o de cuerpos de agua, etc. Este tipo de trabajos se suele realizar mediante el uso de imágenes satelitales y técnicas de percepción remota. Además de la utilización o implementación de diferentes algoritmos computacionales.

En el caso de este proyecto se analizará, usando las herramientas de percepción remota y análisis estadístico, la degradación forestal y se tratará de medir las áreas afectadas. Es importante mencionar que detectar y monitorear la degradación forestal a través de imágenes satelitales es más complejo que identificar cambios de uso de suelo porque la degradación forestal es un cambio que se refleja en la estructura y la biomasa del bosque, sin presentar en sí, un cambio de cubierta.

1.2.1 Uso y cubierta de suelo

La cubierta de suelo se refiere al tipo de material presente en un espacio determinado, por ejemplo, agua, arena, bosque, concreto, etc. Mientras que el uso de suelo se refiere a la actividad desarrollada por las personas sobre esta superficie, por ejemplo, cultivos, ganado, asentamientos, etc. La magnitud y aceleración de los cambios de uso y cubierta de suelo en la superficie de la tierra han estado presentes en toda la historia de la humanidad es por ello por lo que tener la posibilidad de detectarlos y poder establecer una línea de referencia para poder medir y prevenir afectaciones posteriores provocadas por estos cambios es un tema de interés general (Jensen, 2015).

1.2.2 La degradación forestal

La degradación forestal se puede definir como un estado de sucesión detenida causada por procesos antropogénicos, donde los procesos ecológicos de los bosques se ven disminuidos

o severamente restringidos. La Organización de las Naciones Unidas para la Agricultura y la Alimentación (FAO), define la degradación forestal como "la reducción de la capacidad de un bosque para proporcionar bienes y servicios". Solo en los trópicos, se ha estimado que el área total de bosque degradado es de alrededor de 500 millones de hectáreas, es por ello por lo que implementar medidas para reducir la degradación forestal es un tema de mucha importancia. La degradación de los bosques tiene generalmente como consecuencia una pérdida de atributos forestales particulares; como biomasa, captura de carbono, regulación del agua, protección del suelo y conservación de la biodiversidad. Los bosques pueden ser considerados como ecosistemas dinámicos y complejos que, después de una perturbación, tienen procesos ecológicos de recuperación, pero cuando los bosques están degradados se provoca un estado de sucesión detenida, provocando pérdidas en la calidad de estos (Ghazoul, et al., 2015).

1.2.3 Cambio climático

El cambio climático es uno de los problemas ambientales de mayor importancia a escala global, esto debido a las implicaciones y afectaciones que causa, como cambios abruptos de temperatura, sequías, lluvias más intensas, desaparición de glaciares, entre muchas otras. Cada día se presentan variaciones en las condiciones del planeta, como temperatura, humedad, precipitaciones, viento, etc. Sin embargo, debe saberse que esto no es llamado clima, sino "estado del tiempo". Para hablar de clima se necesita tener registros y observaciones de las condiciones climáticas en un periodo mínimo de 30 años. El clima de la tierra ha cambiado muchas veces a lo largo de su historia. Estos cambios fueron ocasionados por procesos naturales, como erupciones volcánicas, cambios en la órbita, variaciones en la composición de la atmósfera, etc. Pero en los últimos años y gracias a la intervención del hombre, estos cambios climáticos han aumentado drásticamente su rapidez. Algunos de los factores más importantes para el cambio climático fue la revolución industrial que desde entonces por causa de la quema de combustibles fósiles (petróleo, gas, y sus derivados) ha provocado cambios en la composición de la atmósfera. Otro proceso que también ha aumentado su

aceleración es la pérdida de bosques y vegetación lo que provoca una disminución en la capacidad de remover o disminuir gases de efecto invernadero. Estos procesos han acelerado el cambio climático en la Tierra, siendo el calentamiento global una de las manifestaciones más evidentes (CONAFOR, 2013).

1.2.4 Bosques y cambio climático

Los bosques juegan un papel muy importante en la regulación del clima, debido a que estos tienen la capacidad de absorber el CO₂ mediante la fotosíntesis y almacenarlo durante largos periodos de tiempo. Si consideramos todo el carbono almacenado en los bosques, sería mayor que todo el carbono presente en la atmósfera. Con esto podemos darnos cuenta cuanto afecta los daños forestales al cambio climático. En México las causas de la deforestación y degradación forestal son muy diversas y muchas veces depende de la zona del país en la que se encuentre. Sin embargo, se puede decir que una de las principales causas de la deforestación es el aumento de la ganadería y la expansión de áreas urbanas, industriales y agrícolas, además de la tala ilegal, incendios, plagas y enfermedades. Mientras que la degradación forestal las principales causas son el pastoreo excesivo, tala selectiva, aumento de la agricultura, sistemas de cultivo, como roza, tumba y quema. Muchas de estas causas son debido al intento de las personas de cumplir o satisfacer sus necesidades. Tanto la degradación como la deforestación tienen importantes implicaciones en la salud forestal, ya que esta se ha visto disminuida notablemente en los últimos años, lo cual provoca afectaciones en el cambio climático (CONAFOR, 2013).

1.2.5 UNFCCC y REDD+

El cambio climático es una de las problemáticas ambientales que en la actualidad tienen las mayores afectaciones, es un tema que ha causado cada vez mayor preocupación a nivel global por lo que se han generado diversas iniciativas interesadas en disminuir estos daños. Una de estas iniciativas es la Convención Marco de Naciones Unidas sobre el Cambio Climático (en inglés UNFCCC - United Nations Framework Convention on Climate Change). La UNFCCC fue adoptada en 1992 y entró en vigor en 1994. Esta convención ha sido adoptada por 195

países, incluido México. El cambio climático al tener un alcance mundial también se requiere la cooperación a nivel mundial para dar solución a dicho problema. El fin último de la UNFCCC es lograr la estabilización de las concentraciones de gases de efecto invernadero las cuales son causadas principalmente por actividades antropogénicas. Todos los países que forman parte de la convención tienen la responsabilidad de cumplir con los compromisos para responder ante el cambio climático. Tienen también la obligación de reportar el inventario de las emisiones de gases de efecto invernadero, además de estas hay una extensa serie de protocolos que cada país debe cumplir [<https://unfccc.int/resource/docs/convkp/convsp.pdf>] (UNFCCC, 2007).

El cuidado de los bosques juega un papel fundamental para la mitigación de las emisiones de gases de efecto invernadero a nivel mundial, esto debido a su función de capturar dióxido de carbono. En México, esto implica un reto muy grande ya que los niveles de deforestación y degradación forestal son muy altos.

Otra organización preocupada por el cambio climático es REDD+ (Reducción de Emisiones de Carbono por Deforestación y Degradación Forestal) que es un programa internacional que promueve la toma de acciones para mitigar el cambio climático, dichas acciones se centran en la conservación y cuidado de las áreas forestales ya que se ha visto que la degradación forestal y la deforestación son uno de los principales causantes de los gases de efecto invernadero, REDD+ propone una serie de políticas, medidas y acciones que deben ser cumplidas para ayudar a mitigar estos gases. REDD+ puede definirse entonces, como un conjunto de estrategias orientadas a la mitigación y adaptación del cambio climático, esto por medio de un buen manejo de los recursos forestales logrando así, un desarrollo sustentable bajo en carbono (CONAFOR, 2017).

1.2.6 Percepción remota

La Percepción Remota es la técnica o conjunto de técnicas que permite medir y registrar la energía electromagnética reflejada o emitida por la superficie de la Tierra. La percepción remota es una técnica que ha evolucionado aceleradamente durante las últimas décadas con el desarrollo de la tecnología satelital. Sus orígenes se ubican con el inicio de las actividades

aeronáuticas a principios de siglo XX y evolucionaron con el desarrollo de las actividades espaciales. Los sistemas de percepción remota actuales constan de tres elementos, la fuente de iluminación, un sensor y el objeto observado (SEMAR, 2009).

Para que la percepción remota sea posible se requiere que exista algún tipo de interacción entre los objetos observados de la superficie y los sensores remotos. Esta interacción se conoce como la radiancia electromagnética (Henry, 2008).

La radiación se puede caracterizar en función de su longitud de onda o frecuencia. Debido a que los valores de la longitud de onda de la energía son continuos se establecen bandas en donde la radiación electromagnética tiene un comportamiento similar. La organización de estas bandas se denomina espectro electromagnético (Chuvienco, 1995).

La organización de bandas comúnmente aplicadas en la percepción remota está conformada por la bandas de espectro visible (0.4 – 0.7 μm) la cual es la única que puede ser vista por el ojo humano por lo que es la más utilizada, esta a su vez comprende a las bandas rojo, azul y verde, la banda de Infrarrojo cercano (IRC: 0.7 – 1.3 μm) la cual es útil para estudios de biomasa, Infrarrojo medio (IRM: 1.3 – 8 μm) permite medir niveles de humedad, Infrarrojo lejano o térmico (IRT: 8 – 14 μm) que sirve para medir la temperatura de la superficie terrestre y Microondas (M, por encima de un mm) las cuales son capaces de penetrar la atmósfera ante cualquier condición climática.

Los datos de una imagen satelital son capturados por un sensor remoto, estos son enviados a una estación donde se procesan para convertirse en imagen. Una ventaja de esto es que los satélites obtienen datos de forma periódica lo cual permite dar seguimiento al estudio de algún evento, además, con el desarrollo de la tecnología satelital, cada vez se cuenta con mayor resolución en las imágenes satelitales y un mejor aprovechamiento del número de bandas del espectro electromagnético (INEGI, 2019).

Una imagen satelital puede ser definida como una matriz tridimensional que está compuesta por píxeles, cada píxel tiene asociado un número digital el cual representa la unidad mínima de información de la imagen. Dos dimensiones de esta matriz representan las coordenadas geográficas mientras que la tercera dimensión representa su dimensión espectral, esto toma

sentido ya que la imagen satelital está compuesta por diferentes bandas las cuales representan diferentes longitudes de ondas del espectro electromagnético (Chuvieco, 1995).

1.2.7 Análisis de Mezclas Espectrales (SMA)

Una imagen satelital está conformada por un conjunto de píxeles, los cuales pueden ser considerados de dos tipos, píxeles “mixtos” y píxeles “puros”. Los píxeles puros también conocidos como miembros puros son aquellos que representan una sola clase, es decir que pueden ser considerados como bosque, cuerpo de agua u otros tipos de cubierta o uso de suelo, mientras que los píxeles mixtos son aquellos que tienen asociados dos o más clases diferentes. En una imagen satelital es muy común que existan píxeles mixtos, sobre todo en las áreas que son el límite entre una clase y otra. Por ello es muy importante tomar esto en cuenta a la hora de hacer una clasificación temática y para esto existe un algoritmo llamado análisis de mezclas espectrales (SMA, por sus siglas en inglés) el cual toma en consideración que la reflectancia de un píxel es la media de las reflectancias de cada componente dentro del píxel, donde cada componente está ponderada con base en la proporción de cobertura correspondiente. Este algoritmo se basa en un sistema de ecuaciones lineales y técnicas de regresión para obtener la proporción correspondiente a cada clase dentro del píxel (Mohammed, et al., 2006).

1.2.8 Degradación forestal y el método SMA

Generalmente la degradación forestal ocurre en áreas pequeñas por lo que no se puede observar de forma clara en imágenes satelitales de resolución media o baja. En el caso de México, la degradación forestal es difícil de cuantificar debido a que el acceso a imágenes de alta resolución es muy limitado, costoso o difícil de conseguir, aunado a esto, el uso de imágenes de alta resolución tiene otro inconveniente, entre mayor sea el área de estudio se requerirá mayor cantidad de imágenes satelitales lo cual aumenta aún más el costo, tanto económico como de procesamiento, esfuerzo y tiempo. Esto conlleva a buscar métodos alternativos que permitan estimar la degradación forestal (Leyva, et al., 2017). Hoy en día se tiene una gran variedad de instrumentos que proporcionan cada vez imágenes satelitales de

mayor resolución. Muchas de estas imágenes como por ejemplo imágenes Landsat y Sentinel, son puestas a disposición de cualquier usuario y se pueden obtener de forma gratuita [<https://www.inegi.org.mx/temas/imagenes/imgLANDSAT/>], [http://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-2/Introducing_Sentinel-2]. Existen también algunas otras que son de acceso restringido o con costos elevados como por ejemplo las imágenes SPOT. [<https://earth.esa.int/web/eoportal/satellite-missions/s/spot-6-7>] las cuales suelen tener una mucho mejor resolución. Sin embargo, como se puede apreciar en algunos trabajos, por ejemplo, el estudio realizado por Defries et al. (2020) o Blanc et al. (2016), aplicando cierto tipo de métodos sobre imágenes de acceso libre se pueden lograr muy buenos resultados.

Es importante mencionar que no solo basta con obtener imágenes satelitales para poder detectar cambios en el uso de suelo o en este caso medir la degradación forestal, también se requieren aplicar diferentes metodologías y procesos que ayuden a obtener información confiable. Las nuevas metodologías requieren de equipo de cómputo potente capaz de procesar imágenes de alta resolución con diferentes algoritmos orientados a recopilar la información requerida. Hoy en día existen algunos métodos que nos permiten lidiar con la resolución espacial de las imágenes satelitales para detectar cambios en escalas pequeñas, por ejemplo, el algoritmo Análisis de Mezclas Espectrales que descompone los píxeles de la imagen satelital en los diferentes elementos que lo componen de forma que se puedan detectar cambios en escalas menores. Algunos estudios para medir la degradación forestal, sobre todo en la Amazonía, han usado imágenes satélites en conjunto con algoritmos como el SMA para detectar y medir la degradación forestal, los cuales han obtenido buenos resultados.

Hoy en día existen algunos métodos que nos permiten lidiar con la resolución espacial de las imágenes satelitales para detectar cambios en pequeñas áreas, por ejemplo, el algoritmo SMA que descompone los píxeles de la imagen satelital en los diferentes elementos que lo componen de forma que se puedan detectar cambios en pequeñas áreas. Algunos estudios para medir la degradación forestal, sobre todo en la Amazonía, han usado imágenes satélites en conjunto con algoritmos como el SMA para detectar y medir la degradación forestal, los cuales han obtenido buenos resultados.

Frecuentemente la degradación forestal ocurre en áreas pequeñas, y teniendo en cuenta que las imágenes de Landsat (con las cuales se realizó este trabajo) tienen una resolución espacial de 30 metros, solo se podrían detectar áreas afectadas no menores a la que conforman un grupo de píxeles donde cada uno mide 0.09 hectáreas, por esta razón sería muy difícil detectar las áreas sujetas a degradación forestal siguiendo los métodos tradicionales, es por ello la importancia de aplicar metodologías, como el método SMA, con el cual se pueden obtener los componentes de cubierta de suelo dentro de un píxel y de esta manera detectar cambios en áreas más pequeñas a la resolución espacial de los datos Landsat.

Este trabajo tiene la intención de hacer uso del algoritmo de análisis de mezclas espectrales el cual se basa en la hipótesis de que un píxel está compuesto por una o más componentes, por lo que en muchas ocasiones no se le puede asignar una sola clase (Defries, et al., 2000). Para cada píxel de la imagen se define el porcentaje de las clases que tiene asociadas, dicho porcentaje va de 0 a 100%, cuando el porcentaje es 100, significa que el píxel tiene asociada solo una clase, de otro modo tendrá asociadas dos o más, el SMA regresa imágenes fraccionadas para cada clase donde los píxeles contendrán el porcentaje que tienen de su respectiva clase. En el caso del análisis de degradación forestal, el SMA detectaría cuando un píxel contenga parte de áreas con bosque y otras con bosque degradado u otro, por ejemplo. Como se mencionó anteriormente este método ya ha sido implementado en algunos estudios, muchos de ellos en la Amazonia, dando resultados fiables.

Bullock et al. (2018) presentan un sistema que utiliza series temporales de datos de Landsat y el algoritmo de análisis de mezclas espectrales para detectar tanto la degradación como la deforestación en paisajes forestales en el estado de Rondônia en la Amazonia brasileña específicamente en el periodo de 1990 a 2013. El objetivo de este estudio fue desarrollar una nueva metodología que sea capaz de mapear y estimar la degradación y la deforestación en grandes áreas. Para su desarrollo, los datos obtenidos por Landsat fueron transformados en fracciones espectrales de miembros puros los cuales se usaron para obtener el porcentaje de miembros puros en cada píxel y posteriormente calcular el índice de degradación forestal. Los resultados del estudio lograron una precisión de usuario del 88% y 93.3% para la degradación y la deforestación respectivamente, mientras que la precisión de productor resultó

ser de 68.1% y 85.3% en cada caso. Con esto se puede ver que el error es mayor al intentar detectar degradación forestal.

Basuki et al. (2012) mencionan que una limitación principal para estimar la biomasa de los bosques es que no considera que los píxeles o valores de reflectancia tienen componentes espectrales mixtos. Ellos realizan un estudio en Labanan concession forest al este de Kalimantan, Indonesia. En su estudio ellos descomponen en componentes puros la reflectancia de cada píxel, esto antes de desarrollar su modelo de clasificación y lograr así una mayor precisión en la estimación de la biomasa aérea. La descomposición fue esencial para obtener la proporción de la vegetación dentro de los componentes mixtos, ya que esta es la variable finalmente necesaria para estimar la biomasa aérea. Al aplicar esto sobre una imagen Landsat 7 se obtuvieron imágenes fraccionadas por cada uno de los diferentes elementos puros que conformaban a cada píxel. Posteriormente fueron aplicados modelos de regresión logrando resultados mejores y se concluyó que descomponer las imágenes en imágenes de fracción de vegetación proporciona mejores estimadores de biomasa, por lo que sugieren que esta metodología debería usarse en estudios similares.

En el estudio desarrollado por Souza et al. (2005) se evalúan los cambios de reflectancia, vegetación e índices infrarrojos usando imágenes de fracciones derivadas del análisis de mezclas espectrales sobre un periodo de tiempo de 20 años y con tipos de bosques degradados bien caracterizados. Este estudio fue elaborado usando series temporales de imágenes Landsat, una por cada año de estudio. El área de estudio se encuentra en el estado de Mato Grosso, Brasil. Con el objeto de aplicar el modelo de análisis de mezclas espectrales se seleccionaron los miembros puros en función de la firma espectral y el contexto de la imagen (por ejemplo, los espectros de suelo se asocian principalmente con carreteras sin pavimentar y vegetación no fotosintética, VNP con pastos con vegetación senescente). Posteriormente se aplicó el modelo de mezclas espectrales para cada una de las imágenes exceptuando la imagen de la cual se extrajeron los miembros puros. Al final, se identificaron cuatro clases de bosques degradados: tala no mecanizada, tala gestionada, tala convencional, y tala y quema. Según pruebas estadísticas realizadas, fue posible diferenciar el bosque intacto de las clases

de bosque degradado, y entre sí las clases de bosque degradado. El uso de imágenes fraccionadas provenientes de la aplicación del modelo de análisis de mezclas espectrales fue fundamental para la obtención de tan buenos resultados.

En análisis de serie de tiempo a través de técnicas de regresión lineal ha sido ampliamente utilizado para identificar degradación forestal, tal es el caso del estudio desarrollado por Schneibe et al. (2017) cuyo objetivo era aplicar un enfoque de series de tiempo anual utilizando datos de Landsat en el centro-sur de Angola para evaluar el estado actual de degradación de los bosques de Miombo para ello, usaron el índice de Tasseled Cap Transformation. Con base en esta serie de tiempo, calcularon una regresión lineal para las áreas de bosque no afectadas por cambio de uso de suelo, y también para el período previo a la conversión de aquellas áreas que se utilizaron con fines de cultivo durante el tiempo de observación. Ellos generaron series de tiempo anuales de Landsat para el período de 1989 a 2013 donde para cumplir con su propósito, procesaron todos los datos de Nivel 1T disponibles en el periodo de mayo a junio (periodo donde se presenta menor nubosidad) aplicando correcciones geométricas, radiométricas y atmosféricas. Además del uso del algoritmo Fmask para identificar las nubes y la sombra de estas. Segmentaron el área de estudio para cada año, en áreas de bosque presuntamente estable y aquellas áreas que estaban bajo cultivo durante el tiempo de observación. Esto con la intención de evaluar las tendencias a largo plazo de los bosques supuestamente estables, que pueden incluir la degradación y regeneración de los bosques, que son demasiado sutiles para ser detectados utilizando métodos adaptados a la deforestación, así como para obtener información sobre el uso del bosque antes de la tala y sustitución del rodal. Utilizaron además el índice de perturbación (DI) el cual se desarrolló específicamente para identificar perturbaciones relacionadas con los bosques. El DI es una transformación lineal del índice Tasseled Cap reescalados y, por lo tanto, indica las disminuciones simultáneas de los valores de verdor y humedad. Posteriormente con el DI, aplicaron una serie de tiempo anual de 25 años sobre cada píxel de las imágenes Landsat, esto para derivar tendencias lineales con respecto al bosque presuntamente estable y también en la parte anterior a la tala y quema de la serie de tiempo para derivar información sobre los procesos de degradación que tuvieron lugar antes del uso agrícola. Los procesos de degradación se expresarían

entonces como un aumento de los valores de DI, mientras que las áreas de biomasa en regeneración muestran una disminución de los valores de este indicador. Al analizar los resultados, ellos encontraron que casi el 74% de los bosques de Miombo no mostraron ninguna tendencia significativa y, por lo tanto, se consideraron estables. El resto del bosque tuvo tendencias positivas significativas (13.3%), por lo que se consideró degradante, o tendencias negativas (12,8%), lo que sugiere acumulación de biomasa. Para evaluar los resultados obtenidos, compararon las áreas de degradación obtenidas, con las tendencias integrales latentes de MODIS, obteniendo una confiabilidad del 58%.

Teniendo en cuenta que la degradación forestal es un proceso que se va dando gradualmente a través del tiempo, es prudente analizar el cambio en la cubierta forestal en diferentes periodos. En el presente estudio se trabajará con 10 imágenes satelitales anuales, las cuales cubren la zona de estudio, desde 2010 hasta 2020. De esta forma se pretende realizar un análisis de serie de tiempo para monitorear durante este lapso los cambios en cuanto a degradación forestal se refiere.

1.2.9 Análisis de serie de tiempo

El análisis de serie de tiempo es un método estadístico con el cual se pueden determinar patrones de comportamiento sobre un conjunto de datos recolectados durante algún periodo de tiempo específico y con esto también es posible estimar comportamientos futuros (Levin & Rubin, 2004). De esta forma podemos aplicar este modelo para identificar patrones de comportamiento en el cambio de la cubierta forestal, identificando así, las áreas que presentan degradación forestal. Hoy en día, los métodos basados en series de tiempo han sido aplicados para detectar perturbaciones forestales (Bullock, et al., 2018).

En la actualidad existen algunos trabajos relacionados con el monitoreo de la degradación forestal mediante el uso de series temporales, tal es el caso del estudio realizado por Souza et al. (2013) donde tratan de monitorear la degradación forestal en la Amazonía durante 11 años, desde al año 2000 hasta 2010. Para esto ellos utilizan 1465 escenas de Landsat, las cuales fueron debidamente procesadas para la posterior aplicación del método SMA, el cual es un paso previo para el cálculo del índice de fracción de degradación normalizada (NDFI),

con este índice y con el análisis de series de tiempo lograron estimar el área de la Amazonia que presentaba deforestación, la cual ascendía a 169,074 km², mientras que la degradación alcanzaba casi 51,000 km². Otro estudio similar y más reciente es el de Bullock et al. (2018) donde se presenta también un sistema que utiliza series temporales de datos Landsat para detectar perturbaciones forestales, en dicho estudio las imágenes satelitales fueron transformadas por medio del algoritmo SMA a imágenes fracción, mismas que fueron usadas como insumo para calcular el NDFI, índice con el cual realizaron el análisis de series de tiempo durante el periodo de 1990 hasta 2013. La utilización de este índice fue debido a que se ha visto que es mejor que otros para detectar perturbaciones en los bosques tropicales.

En el trabajo desarrollado por Huang & Kong (2016) fue evaluada la dinámica de degradación de la tierra aplicando un análisis de serie de tiempo con el índice de vegetación de diferencia normalizada (NDVI). Este estudio se realizó en la región del Programa Forestal del Refugio Three-North en la región de China en el periodo de tiempo de 1982 a 2006. Los datos de NDVI utilizados en el estudio fueron obtenidos del grupo GIMSS basados en estimaciones de NOAA / AVHRR con resolución espacial de 8 km y temporal de 15 días. Uno de los objetivos principales de este estudio era evaluar la dinámica de degradación de la tierra utilizando la temporada de crecimiento de NDVI, donde la evaluación las tendencias temporales del NDVI se realizó mediante el cálculo de la pendiente de Theil-Sen (Sen), que es una estimación estadística no paramétrica robusta de la magnitud de la tendencia y es especialmente insensible a valores atípicos pequeños y valores de ruido. La significancia estadística de la tendencia de la pendiente se probó mediante la prueba de Mann-Kendall (MK), que se aplica ampliamente al análisis de tendencias de series de tiempo con datos no normales, esto se probó con un nivel de confianza del 95%. Para un mejor análisis, decidieron enmascarar los tipos de uso de tierra de asentamientos y tierras de cultivo, esto con el uso de tierra en el año de 2005, los resultados mostraron que el NDVI tuvo una tendencia creciente en aproximadamente el 13% del área de estudio mientras que el 6.20% mostró una disminución.

Otro estudio similar es el de Tarazona & Miyasiro (2020) cuyo objetivo era identificar la degradación forestal en una región al sur del Perú llamada “Madre de Dios”, usando el NDVI con el fin de identificar cambios graduales en la cubierta forestal. Las imágenes que utilizaron

fueron adquiridas por el sensor MODIS entre el periodo de 2000 y 2017, de las cuales para cada una se eliminaron las áreas que se presentaban nubosidad, nieve, o niveles altos de aerosoles, luego interpolaron los datos faltantes. Posteriormente y similar a lo que se realizó en el presente trabajo, realizaron un modelo de regresión lineal con el cuál fue determinada la cantidad de degradación forestal, esto mediante la suma de las tendencias negativas significativas (valor $p=0.05$) del análisis del NDVI, identificando que el 2.4% de la zona de estudio presentaba degradación forestal.

Wang & Myint (2016) realizaron también un análisis usando series de tiempo, ellos examinaron los patrones espaciotemporales de la deforestación y el flujo de carbono forestal en Myanmar de 2001 a 2010, utilizando imágenes de resolución moderada (MODIS), en específico la capa del porcentaje de cobertura forestal (The Percent Tree Cover, PTC) del producto MODIS, estas imágenes tienen una resolución espacial de 250 metros y el valor de píxel representa el porcentaje promedio anual de cobertura de árboles, con un rango válido entre 0 y 100. Para obtener el nivel de pérdida forestal implementaron un análisis de tendencias de serie de tiempo, utilizando como valores dependientes del modelo a los valores de los píxeles del PCT y analizaron contra la secuencia de años del periodo de estudio, esto mediante un análisis de regresión lineal simple. Cabe mencionar que sólo conservaron los píxeles que mostraron cambios estadísticamente significativos ($p < 0,05$). Con esto generan una imagen del coeficiente de la pendiente obtenida y una imagen del valor del coeficiente de determinación R^2 . Los resultados obtenidos indican que el área total de deforestación fue de 21,178.8 km², con una tasa de deforestación anual del 0.81%, y que la liberación total de carbono forestal fue de 20.06 millones de toneladas, con una tasa anual del 0.37%, encontrando fuertes relaciones positivas entre PTC y las pendientes de las existencias de carbono forestal. Según los valores obtenidos por R^2 , los modelos explicaron el 40,9%, de la variación total en el cambio de las existencias de carbono forestal con respecto al PCT.

También se han hecho estudios de este tipo para además de analizar la tendencia de la vegetación, predecir su comportamiento posterior, tal es el caso del estudio realizado por Anand et al. (2018), ellos trataron de estimar el cambio en la densidad del dosel forestal durante 2008 y 2018 analizando el NDVI en ambos años y donde posteriormente pretenden

predecir este mismo índice para el año 2028 por medio de la aplicación de un modelo de regresión lineal. Para este fin adquirieron dos imágenes Landsat, una para el año 2008 y otra para 2018, las cuales para su preprocesamiento se les aplicaron correcciones radiométricas y atmosféricas con el fin de obtener el valor real de reflectancia. Las imágenes después de ser corregidas fueron clasificadas en 5 categorías, bosque alto, bosque moderado, bosque bajo, pastizales y tierras desnudas, esto con el objetivo de analizar la variación cuantitativa en diferentes tipos de cubierta y ver el cambio exacto en la densidad de la cubierta forestal. Una vez clasificadas se hizo un análisis de serie de tiempo mediante una regresión lineal, utilizando el software MATLAB para establecer una relación entre el NDVI de 2008 y 2018. Donde a través esta regresión lineal se generó la predicción para el NDVI de 2028. Los resultados indican que la cubierta forestal alta disminuyó drásticamente, pasando de 181,079 km² en 2008 a 111,593 km² en 2018, mientras que la cobertura forestal media se ha reducido de 317.768 km² a 229.496 km² respectivamente. Con la disminución de la cobertura forestal alta y media, las tierras desnudas y los pastizales están aumentando significativamente. Y según la predicción realizada, la cual tuvo un valor R² de 0.617 de seguir esta tendencia, la pérdida de cubierta forestal seguirá aumentando.

El análisis de serie de tiempo se refiere al “conjunto de observaciones de una variable en diferentes periodos de tiempo. En el análisis de una serie de tiempo la variable se explica exclusivamente por su historia, es decir, cada dato está determinado por el simple paso del tiempo.” (Castillo & Guijarro , 2006). Con ayuda del análisis de serie de tiempo podemos detectar patrones de tendencia en los datos, y con esto estimar comportamientos futuros (Levin & Rubin, 2004).

1.2.10 índices espectrales para medir la vegetación

Uno de los principales fines que persiguen muchas de las investigaciones en evaluación y manejo de recursos naturales a través de percepción remota, es el monitoreo de la cubierta vegetal, lo cual es posible gracias a que los satélites como Landsat ofrecen imágenes multi-espectrales que ayudan a evaluar el estado de la vegetación, esto a través del cálculo de algunos índices (como el NDVI, NDMI y NDFI, explicados previamente), es decir operaciones

algebraicas aplicadas a los valores digitales de los píxeles en diferentes bandas de una misma imagen satelital (Muñoz , 2013).

1.2.10.1 Índice de vegetación de Diferencia Normalizada (NDVI)

Este índice de vegetación puede indicar la presencia y el vigor de la vegetación. En el análisis de serie de tiempo con el índice de vegetación se puede monitorear la dinámica de cambio en la vegetación a lo largo del tiempo.

El Índice de Vegetación de Diferencia Normalizada (NDVI, por sus siglas en inglés) es uno de los índices más utilizados para medir el nivel de vegetación presente en cada uno de los píxeles de una imagen satelital (Campos, et al., 2018). Este índice tiene un gran potencial para identificar variaciones en cuanto a estructura, composición y diversidad de los bosques. Mientras mayor sea el valor obtenido por el NDVI, mayor será el vigor de la vegetación presente y, por el contrario, mientras menor sea su valor, habrá menor cantidad o calidad de vegetación (Tun-Dzul, et al., 2008).

La percepción remota juega sin duda alguna un papel muy importante para monitorear cambios en la corteza forestal, sin embargo, gracias al uso de índices de vegetación, como el NDVI, podemos obtener información más certera del nivel o salud en la vegetación presente. El NDVI ha sido utilizado en varios estudios cuyo fin es identificar perturbaciones o cambios en los bosques. Por ejemplo, Maldonado et al. (2015) en su estudio realizaron una investigación sobre las afectaciones por plagas en un bosque piñonero ubicado en el municipio de Cuauhtémoc, Chihuahua, ellos realizaron un análisis temporal, de 2000 a 2014, usando el NDVI como indicador los cambios en la cubierta forestal. En su estudio, mencionan que las afectaciones sobre los bosques pueden ser monitoreadas de forma veraz utilizando series temporales con diferentes índices, en su caso el NDVI, además gracias a la buena resolución temporal de los satélites como Landsat que es de 16 días, se logra dar un seguimiento continuo al análisis temporal.

1.2.10.2 Índice de Diferencia Normalizada de humedad (NDMI)

El Índice de Diferencia Normalizada de Humedad (NDMI, por sus siglas en inglés) es un indicador numérico que se asocia con la humedad de la vegetación. El NDMI suele ser útil para identificar sequías y cambios sutiles en las condiciones de humedad de la vegetación (Santi, 2018). Aunque el NDMI es un índice menos conocido que otros, como el NDVI, algunos estudios han demostrado que es un indicador muy bueno para detectar cambios o perturbaciones forestales, incluso algunos estudios como el de Wilson & Sader (2002) han reportado la obtención de mejores resultados con el NDMI que con el NDVI, esto debido principalmente a que el NDMI tiene una mayor capacidad para detectar cambios más sutiles, como lo que ocurre en la degradación forestal (Jin & Sader, 2005).

A pesar de que son pocos los trabajos relacionados con el NDMI y el monitoreo de las perturbaciones forestales, es indudable que gracias a las imágenes satelitales obtenidas por diferentes satélites, y a la existencia de indicadores como el NDMI, es posible identificar cambios en la vegetación, tal como lo indican (Wilson & Sader, 2002) en su estudio, en el cual, tenían el objetivo de analizar diferentes técnicas para detectar niveles de perturbación en la cosecha en las masas forestales, esto por medio de diferentes indicadores, en específico el NDVI y el NDMI, donde a través de un análisis de serie de tiempo, encontraron que en NDMI tuvo mejores resultados para distinguir perturbaciones forestales.

Otro estudio similar es el elaborado por Jin & Sader (2005), quienes utilizaron imágenes Landsat e índices como el NDMI y el índice de “tasseled cap wetness” (TCW), para analizar las perturbaciones forestales. El objetivo de su investigación era comparar estos índices para identificar cual cometía mayor porcentaje de error. Para medir esto, realizaron un modelo de regresión lineal utilizando 5 imágenes Landsat de diferentes años (1988 – 2002), además de los valores obtenidos por el TCW y el NDMI, y aunque con el uso de ambos índices se obtuvieron resultados muy similares, se pudo notar que el NDMI resultó ser mejor para distinguir perturbaciones forestales.

1.2.10.3 Índice de Fracción de Diferencias Normalizadas (NDFI)

Tanto el NDVI como el NDMI son índices adecuados para identificar perturbaciones en la vegetación, sin embargo, estos dos índices toman como referencia el valor de reflectancia

del píxel, sin tomar en cuenta que un píxel puede ser la mezcla de uno o más tipos de cubierta, es por ello por lo que para medir la degradación forestal suele ser mejor utilizar otros índices espectrales como el Índice de Fracción de Diferencias Normalizadas (NDFI, por sus siglas en inglés), el cual se ha visto en algunas investigaciones ser mejor que otros a la hora de identificar perturbaciones en bosques tropicales (Bullock, et al., 2018).

Cabe destacar que, para la obtención de este índice, a diferencia del NDVI y NDMI no solo es necesario contar con diferentes bandas de una imagen satelital, sino que además es necesario obtener a partir de estas imágenes, las imágenes fracción de miembros puros o miembros puros, los cuales para el cálculo del NDFI, corresponden a las clases de vegetación verde, vegetación no fotosintética, suelo y sombra. Para la obtención de estas imágenes fracción es necesario realizar entonces un paso previo, la aplicación del algoritmo de análisis de mezclas espectrales (SMA, por sus siglas en inglés).

En los últimos años el NDFI ha sido cada vez más aplicado para realizar estudios donde se requiere analizar no solamente la deforestación, sino que también se quieren analizar perturbaciones dentro del mismo bosque, como ocurre en la degradación forestal, y es que gracias a las características que definen al NDFI, es posible tener mayor información sobre los componentes de cada píxel de la imagen satelital, logrando así identificar perturbaciones forestales en una resolución menor a la resolución espacial de la imagen satelital.

Muñoz et al. (2020) realizaron un estudio en una zona al noreste del Ecuador para medir la degradación forestal utilizando como indicador al NDFI, en su trabajo ellos evalúan la capacidad de detección de la degradación forestal mediante la aplicación de un análisis de serie de tiempo, desde el 2008 hasta el 2018, usando imágenes de Landsat 7 y 8, además del NDFI, donde como era de esperarse, tuvieron que aplicar el algoritmo SMA, identificando los cuatro componentes requeridos por el NDFI. De esta forma lograron observar que el NDFI tiene mayor sensibilidad para detectar cambios intermedios en la cubierta forestal, lo cual resulta ser muy oportuno ya que con este nivel de cambio se logró detectar mejor la degradación forestal.

Otro estudio relacionado con la degradación forestal y el uso del NDFI es el presentado por Bullock et al. (2018), donde a través de imágenes Landsat y la aplicación del SMA obtuvieron las imágenes fracción correspondientes para el cálculo del NDFI. Con esta información ellos realizaron un análisis de serie de tiempo desde 1990 hasta 2013 para detectar la degradación forestal en Rondonia, un estado ubicado en el oeste de Brasil. De esta forma ellos lograron obtener las zonas dentro del área de estudio que presentaban mayor degradación forestal. Además, también lograron identificar que la degradación empezó a aumentar durante la década de 1990 alcanzando su punto máximo, según su estudio, durante el periodo de 2005 a 2010.

1.3 JUSTIFICACIÓN

Como se ha mencionado anteriormente los bosques son de gran importancia no solo para la diversidad biológica, sino también funcionan como reguladores climáticos. Los bosques tienen la capacidad para absorber el dióxido de carbono (CO₂), el cual es un gas de efecto invernadero generado mayoritariamente por actividad antropogénica como procesos industriales, uso de combustibles fósiles, quema de bosques, entre otras. (CONAFOR, 2018)

Muchas veces la degradación forestal es difícil de percibir por lo que suele ser considerada como una etapa inicial de deforestación, y, por tanto, es una problemática que puede ser minimizada y considerarse una afectación con impacto poco importante a la salud de los bosques (Alonso, 2019).

Sin embargo, la degradación forestal es un problema que nos perjudica a todos los seres vivos ya que es uno de los principales causantes de las grandes emisiones de gases de efecto invernadero lo cual, a su vez, ocasiona el cambio climático. El cambio climático por emisiones de gases de efecto invernadero, constituye el principal desafío global en materia ambiental del siglo XXI, y representa una de las mayores amenazas para el proceso de desarrollo y el bienestar humano a mediano y largo plazo (CONAFOR, 2018).

Es por ello por lo que el presente proyecto tiene la intención de identificar y cuantificar las áreas que han sido afectadas por la degradación forestal, dándole así, la importancia que

tiene el llevar un adecuado control de la degradación y cuidado de las áreas forestales. Para este fin es importante hacer uso de las nuevas herramientas y técnicas enfocadas a extraer la mayor cantidad de información referente a la salud forestal, en específico, para este estudio la degradación forestal.

Algunas técnicas de percepción remota y análisis de imágenes satelitales permiten extraer información a partir de imágenes de satélite captadas por diferentes sensores remotos. Estas técnicas permiten identificar o clasificar los píxeles en respuesta a la firma espectral que refleja el área espacial a la cual pertenece el píxel. Sin embargo, dichas técnicas son consideradas “duras” ya que asignan al píxel una única clase y no toman en cuenta que un píxel puede estar compuesto por más de un tipo de cubierta (Giraldo & García, 2009).

Por lo anterior, en este proyecto se utilizarán otro tipo de técnicas que permitan identificar diferentes clases dentro de un mismo píxel como, por ejemplo, el SMA, el cual como ya se ha mencionado se basa en la hipótesis de que la radiancia de un píxel es una combinación lineal de las radiancias individuales de los diferentes componentes (píxeles puros) de los que está conformado el píxel.

En el análisis de imágenes satelitales, tener en cuenta que los píxeles pueden estar compuestos por diferentes clases es de mucha ayuda para detectar degradación forestal ya que esta suele ocurrir en áreas pequeñas comparadas con la resolución espacial de las imágenes satelitales por lo que las áreas afectadas pueden perderse en la imagen y dificultar su identificación.

1.4 OBJETIVO

1.4.1 Objetivo general

Detectar y cuantificar la degradación forestal con percepción remota, utilizando el análisis de tendencia de la vegetación con imágenes de serie de tiempo (2010-2020) de datos de índices de vegetación, humedad y degradación derivados de imágenes de Landsat.

1.4.2 Objetivos particulares

- Adquirir y procesar imágenes satelitales (Landsat-8 OLI, Landsat 7 ETM+) de múltiples fechas (2010-2020), para la zona de Ayuquila, Jalisco.
- Analizar los índices de vegetación, humedad y fracción para identificar afectaciones en la cubierta forestal, así como aplicar el modelo SMA para la obtención de imágenes fraccionadas de los componentes de vegetación, vegetación no-fotosintética, sombra y suelo desnudo.
- Analizar la tendencia de vegetación con imágenes de serie de tiempo (2010-2020) utilizando el modelo de regresión lineal para identificar áreas con cambios en el nivel de vegetación y de esta forma detectar mediante interpretación visual áreas con degradación forestal.

CAPÍTULO 2. MATERIALES Y MÉTODOS

2.1 ÁREA DE ESTUDIO

El área de estudio es la cuenca del río Ayuquila en el área del Pacífico occidental de México (Figura 1), la cual está integrada por los municipios de El Grullo, Autlán de Navarro, Unión de Tula, Tuxcacuesco, Tonaya, Zapotitlán de Vadillo, El Limón y Tolimán. El área tiene grandes variaciones topográficas, con una elevación que varía de 260 a 2500 metros por encima del nivel del mar. La precipitación anual en promedio varía de 800 a 1200 mm y la lluvia ocurre principalmente entre junio y octubre. La temperatura media mensual varía entre 18 y 22°C (Cuevas, et al., 1998). El área de estudio está compuesta tanto por bosque templado como bosque tropical seco, el bosque templado cubre alrededor del 12% de la cuenca mientras que el bosque tropical seco ocupa alrededor del 24% de la cuenca y se distribuye principalmente en las áreas más bajas (Ghilardi, et al., 2012). La composición principal del bosque templado incluye pinos (*Pinus spp.*), Abetos (*Abies spp.*) y robles (*Quercus spp.*) y se explota principalmente para madera, aunque recientemente se han establecido plantaciones de aguacate (*Persea americana*) en áreas anteriormente ocupadas por el bosque templado. El área cubierta por bosque seco tropical se utiliza principalmente para la agricultura migratoria, aunque también se realizan actividades de extracción de leña, pastoreo de ganado y extracción de postes para la construcción de cercas (Ghilardi, et al., 2012), (Jardel, et al., 2012).

La superficie total del área es de 417,554 hectáreas y representa el 5% de la extensión territorial del estado de Jalisco. Según datos de INEGI el 56% de la superficie corresponde a bosques y selvas, sin embargo, el 71% de esta área está en fase secundaria, esto principalmente debido por el aumento de actividades agrícolas y ganaderas, así como también por tala clandestina. El 31% del territorio corresponde a tierras agrícolas, 10% son pastizales y el restante son cuerpos de agua y área urbana. En cuanto a su clima podemos agruparlos en tres tipos, templado, cálido y semiseco, donde las áreas forestales conservan un clima mayormente templado subhúmedo (Santana, 2015).

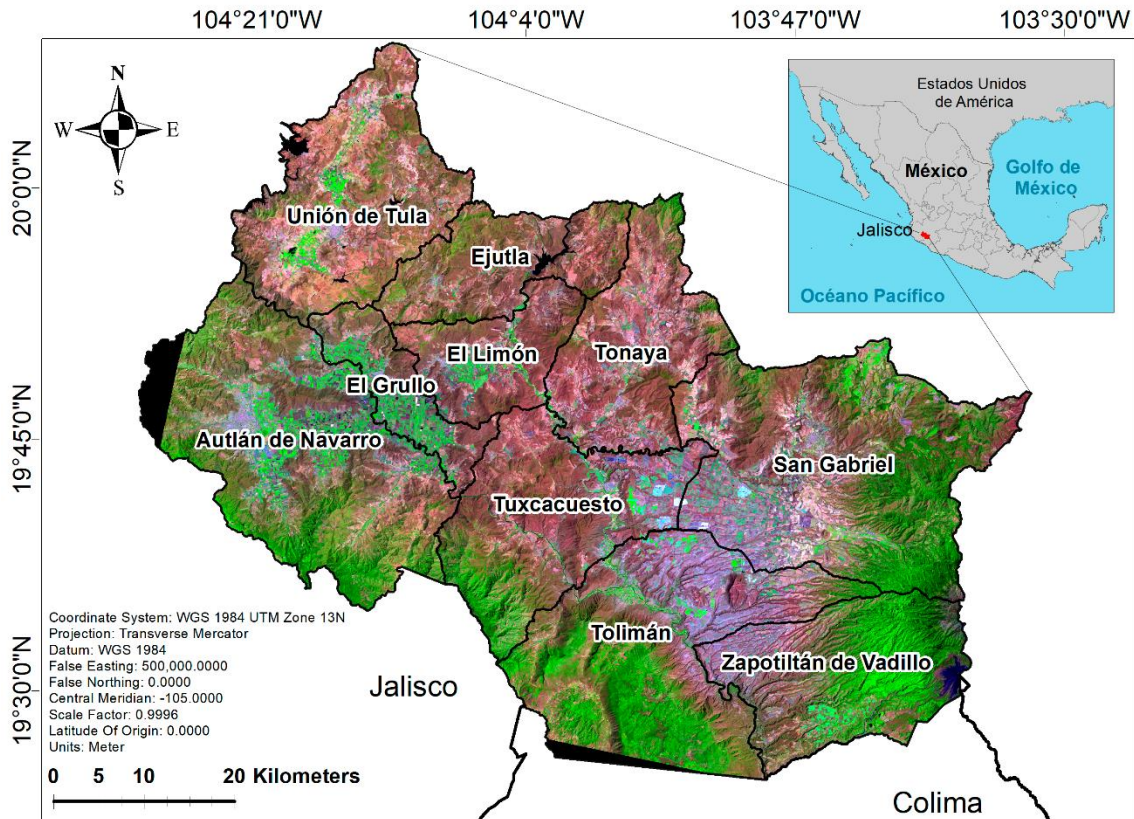


Figura 1. Imagen satelital de la cuenca de Ayuquila 2020
 (Combinación de bandas R = banda 6, G = banda 5, B = banda 4)

2.2 DATOS

Para la realización de este trabajo es necesario tener acceso a imágenes satelitales, en específico imágenes de Landsat 8, que son con las que se trabajó en este proyecto. En la actualidad existen diversas fuentes o plataformas de donde podemos obtener imágenes de diferentes satélites, incluido Landsat 8, sin embargo, una plataforma que permite su descarga de forma sencilla y gratuita es “EarthExplorer” [<https://earthexplorer.usgs.gov/>].

La interfaz de usuario EarthExplorer (EE) es una herramienta de búsqueda, descubrimiento y pedidos en línea desarrollada por el Servicio Geológico de los Estados Unidos (USGS). EE admite la búsqueda de inventarios de satélites, aviones y otros sensores remotos a través de capacidades de consulta interactivas y basadas en texto. Los usuarios registrados de EE tienen acceso a más funciones que los usuarios invitados (USGS, 2019).

En 2016 la USGS estructuró los archivos Landsat en forma de colecciones por niveles lo cual fue llamado como Landsat Collection 1. La implementación de Colecciones representa un cambio significativo en la gestión del Archivo Landsat al garantizar una calidad constante a través del tiempo y en todos los instrumentos, junto con cambios adicionales como metadatos y nombres de archivos. La colección 1 contiene todos los datos de nivel 1 adquiridos desde 1972 hasta el presente Landsat 8 los cuales, según la USGS son calibrados radiométricamente y ortorectificados utilizando puntos de control terrestre y datos del modelo de elevación digital. Además, también contiene otros dos niveles, el nivel 2 y el 3. Estos últimos dos niveles fueron desarrollados a partir de los datos de nivel 1, y tienen la finalidad de ser usados en diferentes tipos de aplicaciones científicas. Los productos de nivel 2 son datos de observación de series temporales de longitud, consistencia y continuidad suficientes para registrar los efectos del cambio climático, este nivel sigue el mismo estándar radiométrico que las escenas del nivel 1, sin embargo, difiere con las especificaciones geométricas ya que cuenta con información orbital menos precisa, nubosidad, entre otros factores. A su vez, los productos de nivel 2, sirven como insumo para productos de nivel 3 los cuales representan a las propiedades biofísicas de la superficie de la Tierra (USGS, 2019). Cabe mencionar que el nivel 1 consta de productos generados por el sensor de imágenes terrestres operacionales (OLI) / sensor infrarrojo térmico (TIRS) Landsat 8, el Mapper temático mejorado (ETM+) Landsat 7, el Mapper temático (TM) Landsat 4-5 y el Landsat 1-5 Instrumentos de escáner multiespectral (MSS) (USGS, 2019). En la siguiente tabla (Tabla 1) se muestra un resumen correspondiente a los diferentes niveles de la colección 1 de Landsat, así como una breve descripción de los productos de cada una.

Nivel	Producto	Descripción
1	Landsat 8 OLI/TIRS	El satélite Landsat transporta consigo dos instrumentos OLI y TIRS por sus siglas en inglés, Operational Land Imager y Thermal Infrared Sensor, respectivamente. El sensor OLI maneja nueve bandas espectrales las cuales van desde 0.43 μm hasta los 1.38 μm , mientras

		que el sensor TIRS va desde los 10.60 hasta los 12.51.
	Landsat 7 Enhanced Thematic Mapper Plus (ETM+)	Landsat 7 lleva el sensor ETM+, una versión mejorada de los instrumentos TM que estaban a bordo de Landsat 4 y Landsat 5. Los productos Landsat 7 se entregan como imágenes de 8 bits con 256 niveles de gris.
	Landsat 4-5 Landsat Thematic Mapper (TM)	El sensor Landsat TM se llevó a bordo desde julio de 1982 hasta mayo de 2012 con un ciclo de repetición de 16 días. El satélite comenzó las actividades de desmantelamiento en enero de 2013. Los archivos de datos de imagen Landsat 4-5 TM consisten en siete bandas espectrales. La resolución es de 30 metros para las bandas 1 a 7.
	Landsat 1-5 Multispectral Scanner (MSS)	Los datos de Landsat 1-5 han sido históricamente un desafío para los sistemas de procesamiento Landsat Nivel-1. A una cantidad significativa de los datos les faltan líneas de escaneo, y otros no son escenas nominales de datos completos. El esfuerzo por procesar los datos del Landsat MSS en la Colección 1 requirió cambios en el procesamiento de datos para influir en la precisión geométrica y radiométrica de los productos de datos de Nivel 1.
2	Reflectancia superficial	Mide la fracción de radiación solar entrante que se refleja desde la superficie de la Tierra hasta el sensor Landsat.

	Índices espectrales	Se encuentran disponibles mediciones de vegetación y humedad, junto con índices de relación de combustión
	Temperatura de la superficie	Representa la temperatura de la superficie de la Tierra en Kelvin (K).
3	Agua superficial	Describe la existencia y el estado del agua superficial.
	Cubierta de nieve	Indica el porcentaje de un píxel cubierto por nieve.
	Área quemada	Representa la clasificación de quemado por píxel y la probabilidad de quemado.

Tabla 1. Productos de Landsat Colección 1. Elaborada con datos de (USGS, 2019)

Las imágenes con las cuales se realizó este trabajo fueron imágenes captadas por sensor de Landsat 8. Este satélite fue lanzado el 11 de febrero de 2013 y órbita alrededor de la Tierra en una órbita sincrónica al Sol, casi polar, a una altitud de 705 km, inclinada a 98.2 grados, y rodea la Tierra cada 99 minutos. El satélite tiene un ciclo de repetición de 16 días. Los productos de Landsat antes de ser puestos a disposición de los usuarios son procesados usando parámetros estándar para obtener mejores resultados en cada escena. Uno de los procesos más importantes es la aplicación de técnicas para corregir geoméricamente a la imagen satelital, el cual está determinado por la existencia de puntos de control recopilados desde la superficie terrestre, la elevación obtenida a partir de modelos digitales de elevación, y otros datos obtenidos por el sensor, con esto se asegura que los píxeles de la imagen estén alineados y que su geolocalización sea la mejor posible, para así ser ubicados con una mejor precisión (USGS, 2017).

La USGS (2017) menciona los niveles de corrección y procesamiento de los productos de la colección de Landsat nivel 1.

Corrección de precisión del terreno: Calibrado y ortorectificado radiométricamente utilizando puntos de control del suelo (GCP, por sus siglas en inglés) y datos del modelo de elevación digital (DEM, por sus siglas en inglés) para corregir el desplazamiento en relieve.

Corrección sistemática del terreno: Calibrado radiométricamente con correcciones geométricas sistemáticas aplicadas utilizando los datos proporcionados por la nave espacial y los datos DEM para corregir el desplazamiento en relieve.

Corrección Sistemática Geométrica: Calibración radiométrica con solo correcciones geométricas sistemáticas aplicadas utilizando los datos de efemérides de la nave espacial.

Landsat 8 opera usando once bandas espectrales, incluida una banda panorámica y dos bandas térmicas (Tabla 2). El sensor OLI de Landsat 8 captura datos con una precisión radiométrica de 12 bits, con lo que se obtienen 4096 niveles diferentes de gris, lo cual es una diferencia bastante significativa con los 256 niveles de gris en de Landsat 1-7 de 8 bits. Esto permite una caracterización mejorada del estado y la condición de la cobertura del suelo (USGS, 2019).

Bandas		Longitud de onda (μm)	Resolución (m)	Aplicaciones
Banda 1	Aerosol costero	0.43 - 0.45	30	Estudios costeros y de aerosoles.
Banda 2	Azul	0.45 - 0.51	30	Cartografía batimétrica, que distingue el suelo de la vegetación y la vegetación caducifolia de la vegetación de coníferas.
Banda 3	Verde	0.53 - 0.59	30	Destaca los picos de máxima vegetación, que son útiles para evaluar el vigor de las plantas.
Banda 4	Rojo	0.64 - 0.67	30	Distingue las laderas de vegetación.
Banda 5	Infrarrojo cercano (NIR)	0.85 - 0.88	30	Destaca el contenido de biomasa y las costas.
Banda 6	Infrarrojo de Onda Corta 1 (SWIR 1)	1.57 - 1.65	30	Distingue la humedad del suelo y de la vegetación; penetra a través de nubes finas.
Banda 7	Infrarrojo de Onda Corta 2 (SWIR 2)	2.11 - 2.29	30	Mejora de la lectura de la humedad del suelo y la vegetación y la penetración a través de nubes finas.

Banda 8	Panromática (Pan)	0.50 - 0.68	15	Resolución de 15 metros, definición de imagen más nítida.
Banda 9	Cirros (Cirrus)	1.36 - 1.38	30	Mejor detección de la contaminación en cirros.
Banda 10	Sensor Térmico Infrarrojo 1 (TIRS 1)	10.60 - 11.19	100	Resolución de 100 metros, mapeo térmico y humedad estimada del suelo.
Banda 11	Sensor Térmico Infrarrojo 2 (TIRS 2)	11.50 - 12.51	100	

Tabla 2 Distribución de bandas de Landsat 8. Elaborada con datos de (Ariza, 2013)

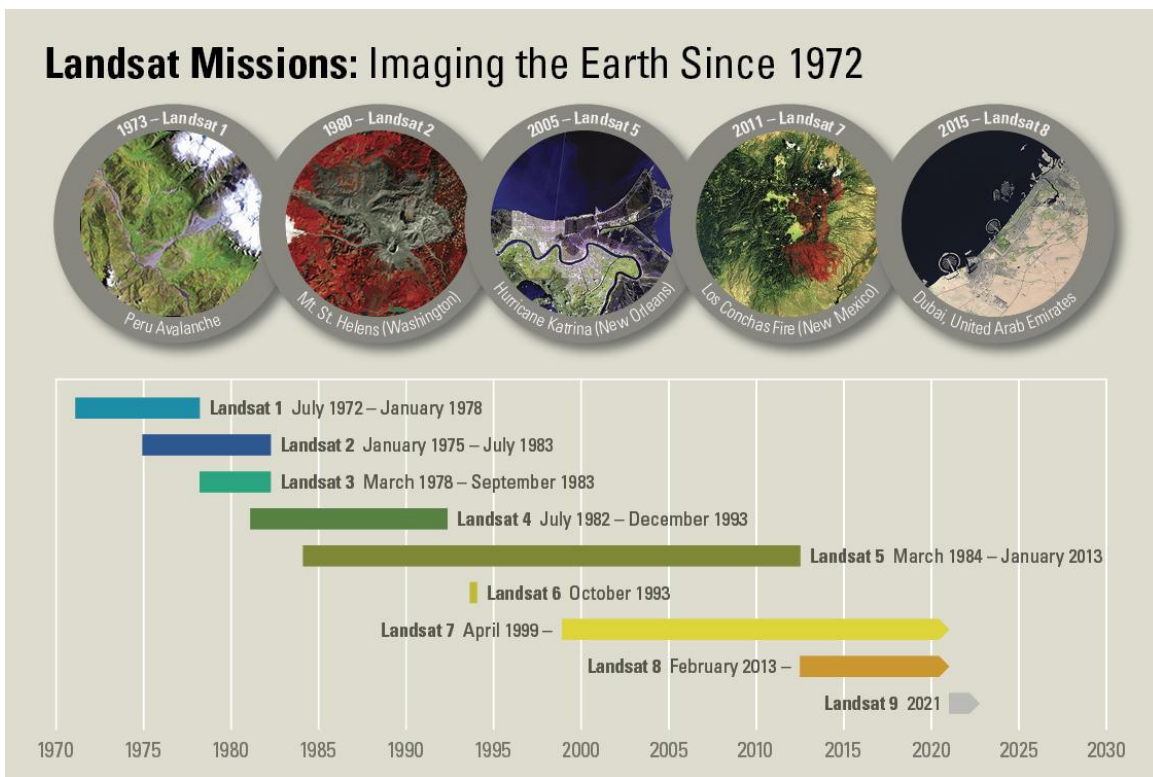


Figura 2. Misiones de Landsat. Obtenida de (USGS, 2019)

2.3 ANÁLISIS DE IMÁGENES SATELITALES

El analizar imágenes satelitales para extraer información de éstas, conlleva una serie de pasos que van desde la adquisición de las imágenes hasta la presentación de los resultados (Morales, 2013). A continuación, se mencionan cada una de estos.

2.3.1 Adquisición de la imagen

Se refiere al proceso de representar en una imagen digital los datos obtenidos por varios sensores remotos. Es importante primero definir la zona de estudio para así poder seleccionar la imagen que tenga información suficiente del área en la cual se va a trabajar. Una vez definida esta zona es importante seleccionar el sensor o sensores de los cuales se obtendrá la imagen satélite. En la actualidad se tiene la ventaja de que existen diferentes medios o fuente en los cuales se pueden obtener imágenes satelitales de diferentes sensores, algunos de ellos son; EarthExplorer [<https://earthexplorer.usgs.gov/>], GloVis [<https://glovis.usgs.gov/app>], Copernicus [<https://scihub.copernicus.eu/dhus/#/home>], entre otras.

2.3.2 Preprocesamiento

Es indudable la importancia y crecimiento que ha tenido el empleo de imágenes satelitales. Actualmente muchas de estas imágenes son accesibles por casi cualquier usuario, sin embargo, el análisis de estas no es una tarea trivial, requiere de algunas técnicas de procesamiento y correcciones para poder extraer información verídica y oportuna. Es por esto por lo que no cualquiera tiene la oportunidad o habilidad de emplear este tipo de recursos. Antes de analizar una imagen satelital es necesario aplicar una serie de técnicas o procedimientos que permitan hacer cierto tipo de correcciones en estas, esto ayudará a que el proceso de reconocimiento e interpretación sea más preciso. Las diferentes técnicas para realizar este cometido se engloban en dos categorías, correcciones geométricas y correcciones atmosféricas (Hantson, et al., 2011).

2.3.2.1 Corrección geométrica

Como se ha mencionado anteriormente una imagen satelital está compuesta por píxeles, donde cada píxel tiene una ubicación, es decir, que el píxel se encuentra ubicado con base en su posición en las filas y columnas de la imagen satelital. Esta ubicación es conocida como el espacio imagen del píxel. Esta ubicación en sentido espacial no nos dice mucho. Por suerte, en la actualidad las imágenes satelitales como Landsat o Sentinel ya vienen definidas con las

coordenadas geográficas pertenecientes a estas, sin embargo, por el método de captura o sensor usado, puede haber ciertos errores geográficos, por ello es importante aplicar técnicas para corregir estos posibles errores. La corrección geográfica ajusta el espacio imagen a un nuevo sistema de referencia donde cada píxel de la imagen tenga asociado un valor XY correspondiente a las coordenadas geográficas que dependen de la proyección utilizada. Este nuevo espacio es denominado espacio de proyección. Para este fin existen dos técnicas principales, la rectificación y la ortorrectificación, ambas técnicas usan puntos de apoyo, los cuales son puntos de los cuales se conoce su ubicación exacta, con estos puntos se procede a corregir el resto de los píxeles de la imagen definiendo o modificando su posición asociada (Cuartero & Felicísimo, 2003).

2.3.2.2 Corrección atmosférica

Los sistemas de teledetección captan la radiación reflejada por los objetos en la corteza terrestre, desafortunadamente esta radiación puede verse obstruida por ciertos procesos atmosféricos como las nubes, gases atmosféricos, aerosoles, entre otros elementos. Por ello es importante tener en cuenta que no toda la energía que reciben los objetos estudiados es reflejada, una parte es absorbida por el propio objeto y otra tanta se pierde o es interrumpida por la atmósfera de la Tierra. Se puede decir entonces que la radiación medida por los sensores no es exactamente igual a la que es reflejada ya que la atmósfera produce cierto tipo de perturbaciones. Es importante entonces aplicar técnicas que permitan corregir el error producido por las interacciones en la atmósfera a la hora de recibir la energía reflejada (Castillo, et al., 2009).

La radiancia captada por los sensores remotos es transformada por medio de ecuaciones matemáticas y almacenada en una representación discreta de números digitales (ND) pero debido a que la escala de los ND depende completamente de la radiancia percibida por el sensor al momento de captura, no se puede comparar directamente firmas espectrales usando el ND de los píxeles de imágenes capturadas por sensores distintos o bien, por imágenes tomadas en diferente tiempo. Para poder realizar estas comparaciones es necesario conver-

tir el ND a una escala espectral normalizada, entonces la corrección atmosférica busca transformar el ND a un nivel de reflectancia con el cual ya es posible tener datos con características comparables. Para cumplir con este fin la corrección atmosférica convierte los ND en la radiancia intrínseca de cada píxel para posteriormente obtener la reflectancia correspondiente. Los ND son proporcionales a la radiancia capturada originalmente por el sensor, esto hace posible recuperar dicha radiancia a través de la aplicación de fórmulas de conversión junto con los datos de calibración mismos que generalmente vienen incluidos en la información o metadatos de la imagen satelital (Aguilar, et al., 2014). En la Figura 3 podemos observar el resultado de aplicar la corrección atmosférica a una imagen satelital y como puede observarse se obtiene una imagen con una mejor visibilidad.

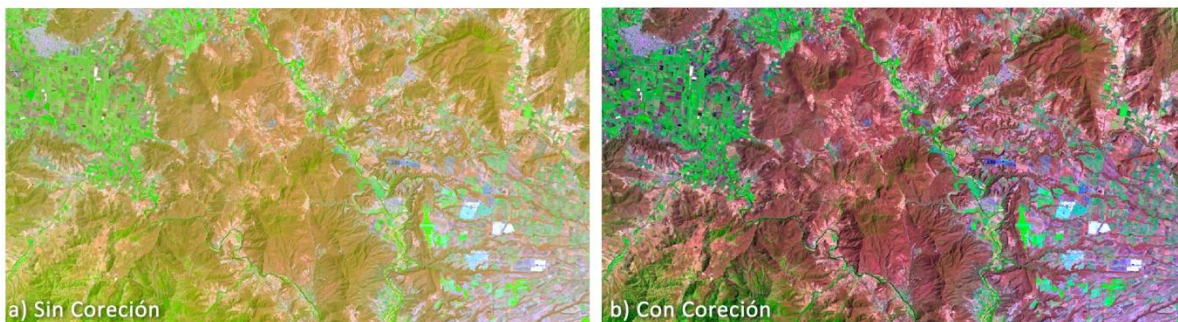


Figura 3. Ejemplo de corrección atmosférica (Combinación de bandas R = banda 6, G = banda 5, B = banda 4). La imagen de la izquierda representa una imagen antes de la corrección (píxeles representados con número digital) y la imagen de la derecha después de la corrección (píxeles representados por la reflectancia planetaria)

Para transformar a un nivel de radiancia los números digitales se pueden usar valores de reescalado que son proporcionados para cada banda en el archivo de metadatos de Landsat 8, con estos valores se puede aplicar la ecuación 1:

$$L_{\lambda} = M_L Q_{cal} + A_L \quad (1)$$

Donde:

$L_{\lambda} \rightarrow$ Radiancia espectral

$M_L \rightarrow$ Factor de reescalamiento multiplicativo

$A_L \rightarrow$ *Factor de reescalamiento aditivo*

$Q_{cal} \rightarrow$ *Valores de píxeles cuantificados y calibrados (ND)*

A su vez, también podemos obtener la reflectancia a partir de los ND utilizando factores de reescalado los cuales también son provistos en el archivo de metadatos de la imagen. La ecuación 2 nos permite obtener dicho valor a partir de los valores de reescalado.

$$\rho_\lambda = M_\rho Q_{cal} + A_\rho \quad (2)$$

Donde:

$\rho_\lambda \rightarrow$ *Reflectancia planetaria*

$M_\rho \rightarrow$ *Factor de reescalamiento multiplicativo*

$A_\rho \rightarrow$ *Factor de reescalamiento aditivo*

$Q_{cal} \rightarrow$ *Valores de píxeles cuantificados y calibrados (ND)*

Es importante hacer hincapié que los factores de reescalado multiplicativos y aditivos tanto para el cálculo de la radiancia como para el cálculo de la reflectancia son datos que vienen incluidos en el archivo de metadatos de la imagen. En dicho archivo viene además una descripción detallada y diferentes características propias de la imagen satelital. Este archivo siempre viene adjunto cuando se realiza la descarga de las imágenes. A continuación, se muestra un ejemplo de los factores de reescalado para la obtención de la radiancia y la reflectancia respectivamente. El factor multiplicativo para la radiancia es denominado en los metadatos como “RADIANCE_MULT_BAND_X” donde X es el número de banda, el factor aditivo es “RADIANCE_ADD_BAND_X”. En cuanto a los factores multiplicativos y aditivos para el cálculo de la reflectancia son “REFLECTANCE_MULT_BAND_X” y “REFLECTANCE_ADD_BAND_X” respectivamente.

RADIANCE_MULT_BAND_1 = 1.2832E-02	REFLECTANCE_MULT_BAND_1 = 2.0000E-05
RADIANCE_MULT_BAND_2 = 1.3140E-02	REFLECTANCE_MULT_BAND_2 = 2.0000E-05
RADIANCE_MULT_BAND_3 = 1.2109E-02	REFLECTANCE_MULT_BAND_3 = 2.0000E-05
RADIANCE_MULT_BAND_4 = 1.0211E-02	REFLECTANCE_MULT_BAND_4 = 2.0000E-05
RADIANCE_MULT_BAND_5 = 6.2485E-03	REFLECTANCE_MULT_BAND_5 = 2.0000E-05
RADIANCE_MULT_BAND_6 = 1.5539E-03	REFLECTANCE_MULT_BAND_6 = 2.0000E-05
RADIANCE_MULT_BAND_7 = 5.2376E-04	REFLECTANCE_MULT_BAND_7 = 2.0000E-05
RADIANCE_MULT_BAND_8 = 1.1556E-02	REFLECTANCE_MULT_BAND_8 = 2.0000E-05
RADIANCE_MULT_BAND_9 = 2.4421E-03	REFLECTANCE_MULT_BAND_9 = 2.0000E-05
RADIANCE_MULT_BAND_10 = 3.3420E-04	REFLECTANCE_ADD_BAND_1 = -0.100000
RADIANCE_MULT_BAND_11 = 3.3420E-04	REFLECTANCE_ADD_BAND_2 = -0.100000
RADIANCE_ADD_BAND_1 = -64.16146	REFLECTANCE_ADD_BAND_3 = -0.100000
RADIANCE_ADD_BAND_2 = -65.70212	REFLECTANCE_ADD_BAND_4 = -0.100000
RADIANCE_ADD_BAND_3 = -60.54394	REFLECTANCE_ADD_BAND_5 = -0.100000
RADIANCE_ADD_BAND_4 = -51.05406	REFLECTANCE_ADD_BAND_6 = -0.100000
RADIANCE_ADD_BAND_5 = -31.24254	REFLECTANCE_ADD_BAND_7 = -0.100000
RADIANCE_ADD_BAND_6 = -7.76974	REFLECTANCE_ADD_BAND_8 = -0.100000
RADIANCE_ADD_BAND_7 = -2.61882	REFLECTANCE_ADD_BAND_9 = -0.100000
RADIANCE_ADD_BAND_8 = -57.77915	
RADIANCE_ADD_BAND_9 = -12.21029	
RADIANCE_ADD_BAND_10 = 0.10000	
RADIANCE_ADD_BAND_11 = 0.10000	

Dark Object Subtraction

En la actualidad existen diferentes modelos que permiten cambiar el valor digital de la imagen satelital a su reflectancia, gracias a esto, se tienen varias técnicas que permiten corregir la dispersión atmosférica también conocida como turbulencia (o en inglés conocida como “haze”). La mayoría de estas técnicas se pueden agrupar en un método simple de sustracción de objetos oscuros (DOS, por sus siglas en inglés) o métodos más sofisticados que utilizan varios modelos de transmisión atmosférica o datos de campo in situ (Chavez, 1988). El método DOS es una técnica basada en imágenes para cancelar el componente de turbidez causado por la dispersión aditiva de la señal de detección remota, básicamente consiste en restar un valor DN constante para toda la imagen, lo cual supone un valor de turbidez constante en toda ella, cabe mencionar que esto no siempre es así, sin embargo, se logra una corrección de primer orden, que es mejor que ninguna. También es importante tener en cuenta que se usa una constante diferente para cada banda espectral, con un conjunto diferente de constantes usadas de imagen a imagen (Chavez, 1988). Recientemente se han desarrollado varios métodos para corregir o eliminar el componente de turbulencia, incluida la técnica de sustracción de objetos oscuros (DOS) ampliamente utilizada y existen otras varios modelos

basados en el método DOS (Dark Object Subtraction), el cual se basa en el supuesto de que algunos de los píxeles de la imagen son muy oscuros debido a que son completamente la sombra, generalmente causada por la topografía del terreno o por las nubes, además de algunos otros puntos en la tierra que por su naturaleza son completamente negros, es decir tiene 0% de reflectancia (Chavez, 1996). Una limitación de este modelo es que los valores de turbulencia se seleccionan de forma independiente para cada banda de la imagen. Sin embargo, existe otros modelos basados en la técnica DOS que permiten al usuario seleccionar un modelo de dispersión atmosférico relativo que sirve para predecir el valor de turbidez para todas las bandas. El método mejorado normaliza los valores de turbulencia pronosticados para los diferentes parámetros de ganancia y compensación utilizados por el sistema de imágenes. Una ventaja de la mayoría de los modelos basados en DOS es que no se requiere tener datos del lugar in situ, basta con los datos que vienen incluidos en la imagen satelital (Chavez, 1996).

2.3.3 Clasificación de Imágenes

Existen diferentes métodos de clasificación de imágenes satelitales, pero todos tienen como fin último asignar una clase temática para todos los píxeles de la imagen. Estas técnicas se pueden agrupar en dos métodos principales, la clasificación supervisada y clasificación no supervisada. La clasificación no supervisada se aplica cuando no se cuenta con conocimiento previo de la zona, es decir no se conocen las clases que se van a encontrar, este tipo de clasificaciones se basa en la información de cada píxel agrupándolos en clústeres o grupos con base en su nivel de similitud espectral (Jensen, 2015), algunos ejemplos de este tipo de algoritmos son K-means y el algoritmo ISODATA. Por otra parte, en la clasificación supervisada se tiene información previa de las clases que componen la zona de estudio (por ejemplo, agricultura, área urbana etc.), por lo que en la imagen satelital se deben tener ciertos píxeles ya identificados con su clase correspondiente, a estos píxeles se les conoce generalmente como datos de entrenamiento ya que permiten “entrenar” al modelo de clasificación para que con base en estos datos les asigne una clase a los píxeles restantes. Existen diferentes

algoritmos supervisados para la clasificación de imágenes satelitales, como por ejemplo el algoritmo de máxima verosimilitud.

2.3.3.1 Clasificación no supervisada

Algoritmo K-means

Uno de los algoritmos para agrupamiento o clustering más usados es el algoritmo K-means. Este algoritmo es un método de clasificación no supervisado que para dar la solución óptima requiere realizar numerosas iteraciones por lo que es considerado un método iterativo. Este algoritmo consiste en dos fases, la primera es seleccionar k centros aleatoriamente donde el valor de k es definido por el usuario o investigador. El siguiente paso es agrupar cada instancia o píxel en el caso de imágenes satelitales, al centro cuya distancia sea menor. Para este paso suele usarse la distancia euclidiana. Cuando cada instancia es asignada a su centro más cercano la primera iteración ha sido terminada, sin embargo, este método requiere de varias iteraciones para dar la mejor solución, por esta razón para las siguientes iteraciones los k centros son recalculados con base en las instancias que pertenecían a estos y se repite el ciclo, es decir se vuelven asignar las instancias al centro más cercano. Este proceso se repite hasta cumplir con un número de iteraciones previamente definido, o bien hasta alcanzar la solución óptima. K-means es considerado como un método muy sencillo y rápido por lo que se le ha dado muchas aplicaciones, sin embargo, debido a su naturaleza puede encontrar en algunas ocasiones soluciones incorrectas y es que al tomar los primeros k centros aleatorios puede dar respuestas muy diferentes dependiendo de la ubicación de estos (Shi, et al., 2010)

Algoritmo ISODATA

Un algoritmo usado en clasificaciones no supervisadas es el ISODATA (Iterative Self-Organizing Data Analysis Techniques) en el cual deben ser definidos una serie de parámetros como el número de clases, la cantidad de iteraciones, diferentes umbrales para agrupar, dividir o eliminar clases, entre otros, todos estos parámetros son necesarios para obtener el resultado de la clasificación, dicho resultado no es más que la agrupación de píxeles con base en su firma espectral, es decir se tiene una clasificación espectral sin embargo lo importante es saber a qué clase temática corresponde, esta tarea es propia del usuario o investigador quien

con base en su experiencia o mediante estudios de campo asigna la clase temática que corresponda. El algoritmo de agrupamiento ISODATA tiene como base el algoritmo k-Means, pero toma en cuenta más parámetros, el algoritmo ISODATA requiere ejecutar varias iteraciones en las cuales va eliminando grupos poco numerosos, fusionando grupos cercanos o bien dividiendo grupos dispersos. Este algoritmo es considerado un excelente método de agrupamiento siempre y cuando los parámetros que requiere estén bien definidos, de lo contrario el resultado no será muy confiable. Una de las principales ventajas del algoritmo es que introduce la división y fusión de grupos buscando grupos naturales presentes en los datos (Rendon, et al., 2015).

2.3.3.2 Clasificación supervisada

Algoritmo de máxima verosimilitud

El método de clasificación de máxima verosimilitud es uno de los métodos más usados de clasificación supervisada, esto debido a que es sencillo de aplicar y a que los resultados se pueden entender relativamente fácilmente, además suele devolver resultados muy fiables. Sin embargo, hay que tomar en cuenta que este método se basa en el supuesto de que todos los valores de reflectancia para cada clase siguen una distribución normal multivariada, siendo de esta manera que se podrán usar datos de entrenamiento para calcular el vector de medias y la matriz de covarianzas y a partir de estos estimar la probabilidad de que cierto píxel pertenezca a las diferentes clases, asignándole al píxel la clase que represente una mayor probabilidad (Del Toro , et al., 2015). Usualmente este método se enfoca en el teorema de Bayes el cual toma en cuenta el cálculo de ciertas probabilidades con las cuales se puede determinar a su vez la probabilidad de pertenencia del píxel a cada una de las clases. Las clases son denotadas como C_i donde $i = 1, 2, \dots, N$ y N representa el número de clases a las cuales un píxel con un vector de medias v puede pertenecer. Usando probabilidades condicionales podemos decir que un píxel pertenece a la clase i si y sólo si la probabilidad de que dicho píxel pertenezca a esa clase es mayor a la probabilidad de que este pertenezca a cualquier otra. Entonces tenemos que el píxel con un vector de medidas v es un miembro de la

clase C_i si $P(C_i|v)$ es la probabilidad mayor del conjunto. Tomando en cuenta que los valores de reflectancia siguen una distribución normal multivariada la tenemos que:

$$v \in C_i \text{ si } P(v|C_i)P(C_i) > P(v|C_j)P(C_j) \forall i \neq j$$

$P(v|C_i)$ es la probabilidad condicional que describe la posibilidad de encontrar un píxel en la posición v en cada una de las clases C_i . $P(C_i)$ representa la probabilidad de que el píxel pertenezca a la clase C_i . Estas probabilidades son estimadas a partir del conjunto de datos de entrenamiento para cada una de las clases. Para un espacio N dimensional la función de distribución normal multivariada es:

$$P(v|C_i) = (2\pi)^{-\frac{N}{2}} |Y_i|^{-\frac{1}{2}} e^{-\frac{1}{2}(v-Z_i)^T Y_i^{-1}(v-Z_i)}$$

Donde, Z_i y Y_i es el vector de medias y la matriz de covarianzas calculados a partir de los datos de entrenamiento para la clase C_i (Pushendra, et al., 2014).

Random Forest

Random Forest es un algoritmo de aprendizaje supervisado que se forma por la combinación de diferentes árboles de decisión, los cuales son generados a partir de subconjuntos aleatorios de un mismo conjunto de datos (Espinosa, 2019). Por otro lado, los árboles de decisión son un algoritmo que genera un árbol conformado por nodos, ramas y hojas. Los nodos representan las características, las ramas representan el camino a seguir con base en las características y por último las hojas representan el resultado de la clasificación (Hernández, et al., 2019).

Los árboles de decisión tienen la característica de que funcionan bastante bien con los datos de entrenamiento, pero a la hora de hacer predicciones con otros datos tienden a equivocarse con mucha frecuencia. En otras palabras, se puede decir que los árboles de decisión tienden a cometer overfitting (sobreajuste, en Español), por lo que una pequeña variación en los datos de entrenamiento puede generar un árbol totalmente diferente (Sotaquirá, 2021). Para solucionar esto, surge el algoritmo de random forest, el cual se compone de un conjunto de árboles de decisión combinados con bagging, esto último se refiere a que cada árbol se entrena con un subconjunto de características y de observaciones distintas, elegidas

de forma aleatoria, esto para disminuir el sobreajuste del modelo. Al usar bagging, se asegura que los distintos árboles sean entrenados con distintas porciones del mismo conjunto de datos para resolver un mismo problema. Las observaciones no estimadas en los árboles de decisión, que generalmente son una tercera parte de los datos, se utilizan para validar el modelo (Sotaquirá, 2021). Posteriormente, para hacer la predicción en un problema de clasificación, se asigna la categoría que haya sido predicha por la mayoría de los árboles del bosque aleatorio (Espinosa, 2019).

Entre las principales ventajas de este algoritmo es que es bastante fácil de comprender los resultados obtenidos, además de que se requiere poco trabajo de preprocesamiento de los datos, ya que no es necesaria la eliminación de datos faltantes, o la normalización de estos (Hernández, et al., 2019). Debido a sus ventajas y a la simplicidad que tiene comparada con otros algoritmos, Random Forest es una técnica ampliamente utilizada en muchos campos, como por ejemplo en la teledetección para clasificación de imágenes, en bancos para detección de fraudes, en medicina para identificar posibles enfermedades, en mercadotecnia para agrupar clientes, entre muchas otras (Espinosa, 2019).

2.3.4 Índices de vegetación

Los índices de vegetación son parámetros calculados a partir de los valores de reflectancia en diferentes longitudes de onda, con la característica de ser mayormente sensible a la cubierta vegetal, de esta forma este tipo de índices nos dirán que tanta vegetación está presente en cada píxel de una imagen satelital (Muñoz , 2013).

2.3.4.1 NDVI

Como se mencionó en el capítulo anterior el NDVI es uno de los índices más usados para medir la cantidad de vegetación presente en una imagen satelital. Los valores que puede tomar este índice están delimitados entre -1 y +1, donde los valores positivos van a referir a áreas de vegetación, y los negativos corresponden a otro tipo de cubierta, como nubes, nieve, agua, o suelo desnudo (Tun-Dzul, et al., 2008).

Para calcular el NDVI es necesario tener información espectral de las bandas roja e infrarroja, en Landsat 8, las bandas 4 y 5, respectivamente, una vez identificada esta información podemos calcular el NDVI usando la ecuación 3 (Maldonado, et al., 2015):

$$NDVI = \frac{\rho_{IR} - \rho_R}{\rho_{IR} + \rho_R} \quad (3)$$

Donde ρ_{IR} es la reflectividad en la región del infrarrojo cercano y ρ_R es la reflectividad en la región del rojo. La razón por la cual se utilizan estas bandas para calcular el NDVI, es debido a que la vegetación sana presenta altos niveles de clorofila, la cual absorbe la energía del sol en la banda roja y refleja la energía que cae en el espectro del infrarrojo cercano (Castro, 2013). Para calcular este índice usando datos Landsat 8 tendríamos que utilizar los valores de las bandas 4 y 5, mientras que para Landsat 7 se requieren los datos de las bandas 3 y 4, quedando los valores para Landsat 8 y Landsat 7 definidos como se muestran en las ecuaciones 4 y 5 respectivamente.

$$NDVI = \frac{Banda\ 5 - Banda\ 4}{Banda\ 5 + Banda\ 4} \quad (4)$$

$$NDVI = \frac{Banda\ 4 - Banda\ 3}{Banda\ 4 + Banda\ 3} \quad (5)$$

2.3.4.2 NDMI

El NDMI como se mencionó anteriormente suele ser útil para identificar sequías y cambios sutiles en las condiciones de humedad de la vegetación. Para el cálculo de este indicador es necesario contar con las bandas espectrales del infrarrojo cercano e infrarrojo de onda corta (Santi , 2018). El rango de valores que puede tomar el NDMI, al igual que en NDVI, es de -1 a 1, donde valores cercanos a 1, indican gran presencia de humedad en la vegetación, y entre más cercano a -1, menor nivel de humedad. Para realizar su cálculo se utiliza la ecuación 6 (Santi , 2018):

$$NDMI = \frac{NIR - SWIR}{NIR + SWIR} \quad (6)$$

Donde *NIR* es la reflectividad en la región del infrarrojo cercano y *SWIR* es la reflectividad en la región del infrarrojo de onda corta. La utilidad de la banda NIR es debido a que en esta región del espectro se identifican niveles altos de sequedad en la vegetación, mientras que en la región de la banda SWIR, se reflejan los cambios en el contenido de humedad (Santi , 2018). En Landsat 8, las bandas para calcular el NDMI corresponden a la banda 5 y banda 6, mientras que para Landsat 7 se requieren las bandas 4 y 5, quedando las fórmulas como se ve en la ecuación 7 y 8.

$$NDMI = \frac{Banda\ 5 - Banda\ 6}{Banda\ 5 + Banda\ 6} \quad (7)$$

$$NDMI = \frac{Banda\ 4 - Banda\ 5}{Banda\ 4 + Banda\ 5} \quad (8)$$

2.3.4.3 NDFI

El NDFI es un índice que cada vez más usado para medir los cambios en la cubierta forestal, y es que, como se mencionó anteriormente, este índice a diferencia del NDVI y NDMI, si toma en cuenta que un píxel puede ser las mezcla de más de una clase y por lo tanto logra identificar cambios en escalas más pequeñas. El primer paso para el cálculo del NDFI, es la aplicación del algoritmo SMA usando específicamente cuatro clases correspondientes a vegetación verde (GV), vegetación no fotosintética (NPV), sombra (Sh) y suelo (So), una vez obtenidas las imágenes fracción con estos cuatro miembros puros, es posible realizar el cálculo para obtener el NDFI, según la ecuación 9 (Souza, et al., 2013).

$$NDFI = \frac{GV_{sh} - (NPV + So)}{GV_{sh} + NPV + So} \quad (9)$$

Donde GV_{sh} es un coeficiente que relaciona el contenido de sombra con el de vegetación verde (ecuación 10), Muñoz et al. (2020) mencionan que experimentalmente se ha visto que este coeficiente permite diferenciar coberturas de bosque con otras coberturas vegetales.

$$GV_{sh} = \frac{GV}{100 - Sh} \quad (10)$$

Los valores del NDFI oscilan entre -1 y 1, donde valores cercanos a 1, representan áreas de bosque densas con poca o nula perturbación, valores más pequeños harán referencia a una disminución en la cubierta forestal, y a medida que el NDFI vaya disminuyendo pasará por niveles de bosques degradado hasta llegar a áreas sin vegetación, esto último para valores muy cercano a -1.

2.3.5 Análisis de Mezclas Espectrales (SMA)

Como se ha mencionado anteriormente una imagen satelital está formada por píxeles los cuales pueden ser de dos tipos, píxeles "puros" o píxeles "mezclados", el SMA considera que los píxeles formados por más de un tipo de cubierta de suelo son "mixtos". Esto pasa generalmente para los píxeles que se ubican en los bordes que unen diferentes tipos o usos de suelo. Por el contrario, los píxeles denominados "puros" están formados por un único tipo de cobertura del suelo (Blanc, et al., 2016).

La técnica de SMA aplica básicamente la técnica de regresión lineal múltiple. Esto proporciona la reflectancia del miembro final puro para cada banda, así como imágenes fraccionadas para cada clase de cobertura terrestre. La respuesta espectral de cada píxel de la imagen en cada banda espectral se puede considerar como una combinación lineal de la respuesta de cada fracción (o miembro final) contenida en la mezcla. Por lo tanto, cada píxel exhibe información sobre la proporción respectiva y la respuesta espectral de los diferentes componentes presentes. (Mohammed, et al., 2006).

El análisis de mezclas espectrales (SMA, del inglés, Spectral Mixture Analysis, ecuación 11) se basa en el supuesto de que la reflectancia en un píxel es el promedio de las reflectancias

de cada componente dentro del píxel ponderado por las respectivas proporciones de cada cubierta (Defries, et al., 2000).

$$R_i = \sum_{j=1}^Q r_{ij}x_j + e_i \quad (11)$$

Donde:

- $R_i \rightarrow$ *Reflectancia en la banda i*
- $r_{ij} \rightarrow$ *Reflectancia del componente j en la banda i*
- $x_j \rightarrow$ *Fracción de cobertura del componente j*
- $e_i \rightarrow$ *Término de error*
- $Q \rightarrow$ *Es el número de componentes*

Además, el modelo se basa en las restricciones de que la suma de las fracciones de cobertura de cada píxel debe de resultar en uno y cada una de estas fracciones debe ser mayor o igual a cero como se muestra en la ecuación 12:

$$\sum_{j=1}^Q x_j = 1 \quad (12)$$

$$0 \leq x_j \leq 1, j = 1, \dots, Q$$

Para la solución no restringida, la fracción x_j puede asumir valores negativos y no está limitada a sumar uno. Por lo tanto, los resultados de la solución sin restricciones no reflejan las verdaderas fracciones de abundancia de los miembros puros. Al resolver el sistema de ecuaciones antes planteado, se obtiene una imagen por cada componente básico definido que representa una estimación de la proporción de píxel ocupado por ese componente. Además, se obtiene una imagen del error cometido en la transformación (Quintano, et al., 2003).

El objetivo del algoritmo SMA es determinar la composición interna de los píxeles que se consideran mixtos, para hacer esto debe conocer las características espectrales de los píxeles

puros, conocidos también como miembros puros. Los miembros puros deben mostrar todas las clases de cobertura de la imagen con la que se esté trabajando (Blanc, et al., 2016).

El modelo SMA requiere de entrada los valores de los píxeles puros los cuales muestran la mejor representación para cada clase. Esta es considerada una de las etapas más importantes del modelo. El número total de miembros puros está limitado por las dimensiones espectrales del conjunto de datos de la imagen (Mohammed, et al., 2006).

2.3.6 Análisis de la tendencia de vegetación con el modelo de regresión lineal

El análisis de series de tiempo es una herramienta muy útil para identificar patrones en una serie de datos a través del tiempo, en otras palabras, por medio de una serie de observaciones de alguna variable, se trata de explicar el comportamiento o tendencia que esta va tomando durante algún periodo de tiempo. Para describir esta tendencia se pueden utilizar diferentes métodos, incluso se puede estimar de forma visual, sin embargo, esto no es muy recomendable ya que los resultados quedarían dados únicamente por la percepción del analista. Otra forma mucho más fiable de realizar este fin es aplicar el modelo de regresión lineal simple, de esta forma se obtiene una recta que mejor se ajusta a los datos históricos de la variable analizada, minimizando el error que se pueda producir (Levin & Rubin, 2004).

2.3.6.1 Modelo de Regresión Lineal Simple

Los modelos de regresión se encargan de medir la relación entre las variables estudiadas determinando así, un patrón que indique la tendencia de los datos, en la regresión lineal simple este patrón se determina por medio de una línea recta. Una vez obtenida esta recta, podemos también, corroborar los resultados, es decir medir el nivel de relación existente entre las variables, mediante un análisis de correlación. Para determinar la recta de regresión podemos usar la ecuación 13 (Levin & Rubin, 2004).

$$\hat{Y} = a + bX \quad (13)$$

Donde:

$$\hat{Y} = \text{Valor estimado de la variable dependiente}$$

$a = \text{ordenada } Y \text{ (el valor de cuando } X = 0)$

$b = \text{Pendiente de la recta de tendencia}$

$X = \text{Variable independiente (Tiempo en el análisis de tendencia)}$

Cabe resaltar que lo que queremos hacer es aplicar el modelo de regresión lineal a un análisis de serie de tiempo, por lo que nuestras variables dependientes, en el caso de nuestro estudio, vendría dada por los años en los que se está haciendo el análisis (2010 a 2020) y la variable independiente sería el valor estimado de vegetación presente.

2.3.6.2 Pendiente de la recta y coeficiente de determinación

Al aplicar el modelo de regresión lineal simple, lo que a nosotros nos importa más para identificar la posible degradación forestal, es la pendiente de la recta obtenida. Anteriormente habíamos dicho que el modelo de regresión lineal simple obtiene una recta que mejor se ajusta a los datos, donde uno de los elementos que la definen es la pendiente, donde una pendiente positiva, indica que la variable dependiente ha ido en aumento a lo largo del tiempo de estudio, en nuestro caso, sería que la cubierta forestal ha ido aumentando, por el contrario, una pendiente negativa hace referencia que el nivel de vegetación ha ido decayendo. Si la pendiente llega a ser cero, esto querría decir que el nivel de vegetación presente se ha mantenido estable a lo largo del tiempo. Para calcular la pendiente podemos hacer uso de la ecuación 14 (Levin & Rubin, 2004).

$$b = \frac{\sum XY - n \sum \bar{X}\bar{Y}}{\sum X^2 - n \sum \bar{X}^2} \quad (14)$$

Donde:

$b = \text{pendiente de la línea de estimación de mejor ajuste}$

$Y = \text{valores de la variable dependiente}$

$X = \text{valores de la variable independiente}$

$n = \text{número de datos en la serie de tiempo}$

Una vez que se ha aplicado el modelo de regresión lineal es importante también medir el nivel de relación que tienen las variables dependientes e independientes, y el coeficiente de determinación (R^2) es la principal forma en la que podemos medir el grado de asociación existente entre estas variables. El coeficiente R^2 toma valores entre 0 y 1, donde valores cercanos o iguales a 0, indican que no existe una correlación entre las variables y, por el contrario, cuando toma valores muy cercanos a uno, significa que la correlación es muy alta. Para calcular R^2 podemos usar la ecuación 15 (Levin & Rubin, 2004).

$$r^2 = \frac{a \sum Y + b \sum XY - n\bar{Y}^2}{\sum Y^2 - n\bar{Y}^2} \quad (15)$$

2.3.6.3 Prueba de Mann-Kendall

Existe también otra herramienta muy útil cuando se trabaja con series de tiempo, esta es la prueba de Mann-Kendall la cual es una prueba no paramétrica similar a un análisis de correlación de Pearson, esta prueba puede tomar valores de +1 a -1, donde valores positivos indican una tendencia creciente y valores negativos indican una tendencia decreciente. Además, cuanto mayor sea el valor absoluto obtenido por esta prueba mayor será la congruencia de la tendencia resultante (Priyanka, 2020). Una ventaja de usar esta prueba es que no requiere que los datos sigan una distribución normal o que sean lineales, únicamente requiere que estos no tengan correlación (Gocic & Trajkovic, 2012). La prueba de Mann-Kendall analiza cada valor comparándolo con los valores restantes, es decir, se calcula la diferencia de cada valor con cada uno de los valores de la serie de tiempo, para esto se ayuda de la ecuación 16 (Gocic & Trajkovic, 2012).

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^n \text{sign}(y_j - y_i) \quad (16)$$

Este valor S brinda información sobre la tendencia de la serie de tiempo, si S es positivo indica que la tendencia es creciente, cuando S es negativo indica que existe una tendencia decreciente. El valor S es necesario para obtener el valor del estadístico Mann-Kendall, que se calcula como indica la ecuación 17.

$$t = \frac{S}{n(n-1)/2} \quad (17)$$

Donde:

$n = \text{número de datos de la serie de tiempo}$

$$\text{sign}(y_j - y_i) \begin{cases} 1 \text{ si } (y_j - y_i) > 0 \\ 0 \text{ si } (y_j - y_i) = 0 \\ -1 \text{ si } (y_j - y_i) < 0 \end{cases}$$

2.4 VERIFICACIÓN DE RESULTADOS

Una clasificación de cualquier tipo no estaría bien hecha si no se presenta la validación de los resultados, es decir, indicadores que aseguren la confiabilidad de los datos. Entonces, si tomamos un resultado arrojado por la clasificación y nos vamos a comparar con los datos reales estos deben coincidir en la mayoría de los casos para asegurar un resultado confiable, de otra forma nuestra clasificación no estaría bien hecha y no serviría para tomar decisiones de ninguna índole. Para esto existen diferentes métodos aritméticos que, a partir de un conjunto de muestras de los datos, permiten calcular de forma numérica la confiabilidad del modelo, generalmente este resultado se expresa como un número entre cero y uno, o en forma porcentual, entre más cercano a uno más confiable será el resultado (Mas, et al., 2003).

La evaluación de la confiabilidad de las clasificaciones de imágenes satelitales consiste en comparar la información obtenida con puntos o áreas donde ya se conozca la clase a la cual pertenece, ya sea por medio de estudios de campo o a través de imágenes de más alta resolución o imágenes aéreas en donde sea más perceptible la clase a la cual pertenece un área (Mas, et al., 2003).

Según Stehman and Czaplewski (1998), el proceso de evaluación de la confiabilidad consiste en tres etapas principales;

1. El diseño de un muestreo para seleccionar las unidades o puntos de verificación.
2. La evaluación del sitio de verificación para obtener la clase a la que corresponde.
3. El análisis de los datos consiste en la elaboración de la tabla de confusión y el cálculo de los índices de confiabilidad.

2.4.1 Diseño de muestras

Esta es una etapa muy importante para la evaluación de la confiabilidad de un modelo de clasificación. Consiste básicamente en seleccionar un tipo de muestreo para elegir un número k de muestras que servirán como puntos de verificación a la hora de la validación. En términos de imágenes satelitales estas muestras pueden ser puntos, polígonos, píxeles o grupos de píxeles. Existen diferentes tipos de muestreo, entre los más conocidos están el aleatorio simple, aleatorio estratificado, sistemático y por conglomerados. Se debe ser muy cuidadoso al elegir el tipo de muestreo conveniente ya que los puntos de verificación deben contar con información suficiente de la clase a la que pertenecen para poder hacer la validación, además de estar bien distribuidos para abarcar a toda la zona de estudio (Mas, et al., 2003).

2.4.2 Número de muestras

Otra elección muy importante para la validación de la clasificación es la determinación del número de muestras que serán tomadas para la evaluación de la confiabilidad de los resultados obtenidos. Este número debe ser lo suficientemente grande para obtener una evaluación más precisa, sin embargo, entre más grande sea, más costoso será la validación (Mas, et al., 2003). Elegir el tamaño de la muestra debe tener un sustento estadístico por lo que generalmente se recomienda tomar al menos 50 muestras por cada clase (Sánchez Muñoz, 2016). Sin embargo, un método más sólido para obtener la cantidad de puntos de verificación es mediante la ecuación sugerida por Olofsson et al. (2014) y desarrollada por Cochran (1977) la cual es una ecuación comúnmente usada para calcular el tamaño de una muestra ya que se sustenta en diversos cálculos estadísticos que aseguran la buena representatividad de la población (Ren, et al., 2019).

$$n = \frac{(\sum W_i S_i)^2}{[S(\hat{\theta})]^2 + \left(\frac{1}{N}\right) \sum W_i S_i^2} \approx \left(\frac{\sum W_i S_i}{S(\hat{\theta})}\right)^2 \quad (18)$$

Donde N es el número total de unidades de muestreo, $S(\hat{\theta})$ es el error estándar estimado de la confiabilidad esperada, el cual generalmente se designa como 0.01, W_i es el área proporcional de la clase i , S_i es la desviación estándar del estrato i , $S_i = \sqrt{U_i(1 - U_i)}$, donde a su vez, U_i es la confiabilidad de usuario esperada para la clase i . Usualmente en los estudios espaciales, N es un valor muy grande, por lo tanto la ecuación puede simplificarse como se muestra a la izquierda de la ecuación 18 (Olofsson, et al., 2014).

2.4.3 Matriz de confusión

Esta matriz es útil para saber cuántas instancias de cada clase fueron asignadas a la clase correcta. La matriz de confusión es una matriz de tamaño $n \times n$ donde n es el número de clases. Sobre las columnas de la matriz se enlistan las clases reales y sobre las filas las clases obtenidas. Los elementos que aparecen en la diagonal principal de la matriz nos indican los elementos que fueron clasificados correctamente, es decir que se le asignó su clase real, todos los demás (los que están fuera de dicha diagonal) son errores cometidos por el modelo, los cuales pueden ser de dos tipos, error de omisión o error de comisión (Sánchez Muñoz, 2016).

CLASE	A	B	...	n	TOTAL FILA (t)
A	X_{11}	X_{12}	...	X_{1k}	$\sum_{i=1}^n X_{1i}$
B	X_{21}	X_{22}	...	X_{2k}	$\sum_{i=1}^n X_{2i}$
⋮	⋮	⋮	⋮	⋮	⋮
n	X_{n1}	X_{n2}	...	X_{nn}	$\sum_{i=1}^n X_{ni}$
TOTAL CO-LUMNA (T)	$\sum_{i=1}^n X_{i1}$	$\sum_{i=1}^n X_{i2}$...	$\sum_{i=1}^n X_{in}$	$\sum_{i=1}^n \sum_{j=1}^n X_{ij}$

Tabla 3. Matriz de confusión

2.4.4 Confiabilidad global

Se trata de uno de los indicadores de confiabilidad más usados, esto por su sencillez y por ser muy intuitivo, aunque este índice tiende a sobreestimar la bondad de la clasificación ya que no toma en cuenta los errores entre las diferentes clases. Se calcula como la suma de los elementos bien clasificados dividido entre el total de observaciones. La confiabilidad global toma valores entre cero y uno, siendo una mejor clasificación mientras más se acerque a uno. Esta medida suele ser acompañada por otros indicadores como la confiabilidad de usuario y de productor para evitar el problema de sobreestimar la bondad de la clasificación. (Sánchez Muñoz, 2016).

Cuando el tipo de muestreo realizado es aleatorio simple, la confiabilidad global se puede obtener dividiendo los elementos de la diagonal en la matriz de confusión entre el tamaño de la muestra, esto da como resultado la probabilidad para cualquier sitio en el mapa de ser correctamente clasificado (Mas, et al., 2003). De esta forma la confiabilidad global está dada por la ecuación 19.

$$\frac{1}{N} \sum_{i=1}^M X_{ii} \quad (19)$$

Donde:

$N \rightarrow$ *Tamaño de la muestra*

$M \rightarrow$ *Número de clases*

$X_{ii} \rightarrow$ *Elementos bien clasificados (elementos de la diagonal)*

2.4.5 Confiabilidad de usuario y de productor

La confiabilidad de usuario y de productor son dos indicadores que agregan más información a la confiabilidad global, ya que esta última en términos generales solamente nos dice que tan probable es que una instancia o elemento esté bien clasificado. Primero, la confiabilidad de productor se refiere a la probabilidad de que un elemento perteneciente a una clase esté asignado a la clase a la cual realmente pertenece. Se calcula como la razón entre el número de elementos bien clasificados y el número de elementos de la categoría de análisis. En otras palabras, se usa para calcular los errores omisión. Mientras que la confiabilidad de usuario se refiere a la probabilidad de que un elemento asignado a una clase cualquiera es en realidad perteneciente a esa clase. Este indicador se usa para calcular los errores de comisión (Sánchez Muñoz, 2016). La confiabilidad de usuario y de productor están dadas por las ecuaciones 20 y 21 respectivamente:

$$cu(j) = \frac{X_{jj}}{\sum_{i=1}^k X_{ji}} \quad (20)$$

$$cp(j) = \frac{X_{jj}}{\sum_{i=1}^k X_{ij}} \quad (21)$$

Donde:

$cp(j) \rightarrow$ *Confiabilidad de productor de la clase j*

$cu(j) \rightarrow$ *Confiabilidad de usuario de la clase j*

$X_{jj} \rightarrow$ *Píxeles bien clasificados de la clase j*

$X_{ji} \rightarrow$ *Elemento j de la fila de la clase i*

$X_{ij} \rightarrow$ Elemento j de la columna de la clase i

Como se mencionó anteriormente, los errores cometidos en una clasificación pueden ser de dos tipos, errores de comisión o errores de omisión, los primeros se refieren a clasificar erróneamente elementos en una clase cuando no pertenecen a ella y los de omisión se refieren a todos aquellos elementos que perteneciendo a una clase no están dentro de ella, ya que estos fueron clasificados en otra clase (Sánchez Muñoz, 2016).

2.4.6 Intervalo de confianza

Una vez evaluada la confiabilidad es importante mencionar el intervalo de confianza, el cual en el caso de este estudio sería un intervalo de confianza a partir de una proporción ya que lo que importa es identificar el intervalo de confianza para la confiabilidad de la proporción de los pixeles bien clasificados. En la ecuación 22 se muestra la forma en la que se realiza este cálculo. Este método asume que la proporción fue tomada a partir de un muestreo aleatorio y para poder ser aplicada se debe cumplir que $n\bar{p} > 5$ y $n(1 - \bar{p}) > 5$ (Newcombe & Merino , 2006).

$$\bar{p} \pm Z \sqrt{\frac{\bar{p}(1 - \bar{p})}{n}} \quad (22)$$

Donde \bar{p} es la proporción de la muestra, la cual se calcula a su vez como el cociente del número de elementos que cumplen con una condición (en nuestro caso los elementos bien clasificados) entre el tamaño de la muestra, Z es el valor crítico para para la distribución normal estandarizada la cual generalmente se establece en 1.96 para determinar un nivel de confianza del 95% (Newcombe & Merino , 2006) y por último n es el tamaño de la muestra.

2.4.7 Ejemplo del cálculo de confiabilidad

A partir de los datos de la figura 4 se procederá a determinar la confiabilidad global, la de usuario y la de productor, esto a fin de ilustrar el cálculo de estos indicadores. En la matriz de ejemplo, T representa la cifra total de las columnas y t representa la cifra total de las filas.

		Terreno				t
		A	B	C	D	
M a p a	A	5	0	0	2	7
	B	0	5	1	0	6
	C	1	1	4	0	6
	D	1	0	0	5	6
T		7	6	5	7	25

Figura 4. Ejemplo de matriz de confusión.

Confiabilidad global:

$$\frac{5 + 5 + 4 + 5}{25} = \frac{19}{25} = 0.76$$

Confiabilidad de usuario para la clase C:

$$\frac{4}{6} = 0.66$$

Confiabilidad de productor para la clase C:

$$\frac{4}{5} = 0.80$$

Intervalo de confianza para la confiabilidad global de 95%:

$$0.76 \pm 1.96 \sqrt{\frac{0.76(1 - 0.76)}{25}} = 0.1674 \rightarrow 0.76 \pm 0.16$$

Se cuempe que $25(0.76) > 5$ y $25(1 - 0.76) > 5$

CAPÍTULO 3. ANÁLISIS DE DATOS

3.1 ADQUISICIÓN DE DATOS

Las imágenes con las cuales se trabajó en este proyecto fueron obtenidas de la plataforma llamada EarthExplorer, en total se obtuvieron 10 imágenes anuales desde 2010 hasta 2020. Todas las imágenes fueron capturadas por satélites Landsat colección 1, de las cuales 8 fueron capturadas por el satélite Landsat-8 OLI/TIRS y 2 por el satélite Landsat-5 TM. Se hubiera preferido trabajar con un mismo satélite, pero debido al periodo de estudio (2010 a 2020) no fue posible ya que Landsat-8 OLI/TIRS que es el satélite en funcionamiento más reciente, fue lanzado hasta febrero de 2013, por lo que no hay información de este satélite de 2010 a 2012. Para este periodo se pensó en usar Landsat-7 ETM, pero debido a una falla del SLC (Scan Line Corrector) que presentó este satélite desde 2003, las imágenes capturadas contienen datos faltantes ya que solo se cuentan con el 78% de sus píxeles (USGS, 2017). Por esto se llegó a la decisión de usar imágenes de Landsat-5 TM por lo que para 2010 y 2011 se trabajó con imágenes de este satélite, desafortunadamente, aunque este satélite cesó completamente su funcionamiento en 2013, para 2012 no se cuenta con datos en el área de estudio de nuestro proyecto por lo que no quedó otra opción más que omitir información para ese año.

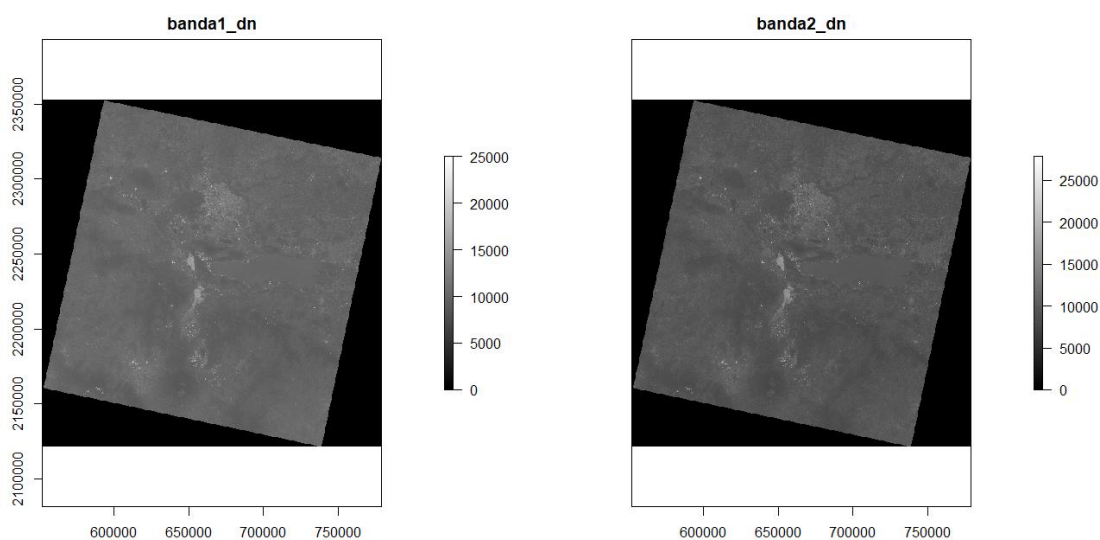
Se trató de que la fecha de adquisición de las imágenes fuera a mediados de marzo ya que en esta época las condiciones climáticas son más adecuadas permitiendo que las imágenes satelitales presenten la menor perturbación posible. En la Tabla 4 se menciona la fecha y el satélite de las imágenes obtenidas para cada año.

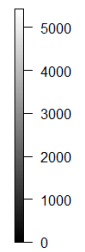
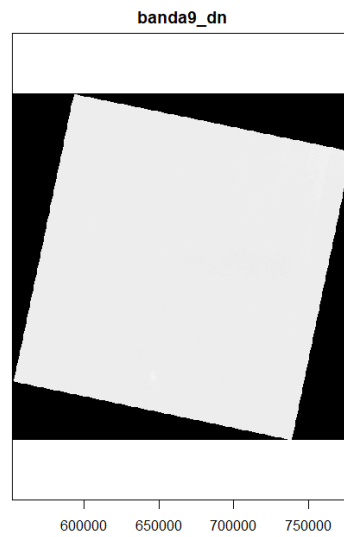
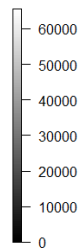
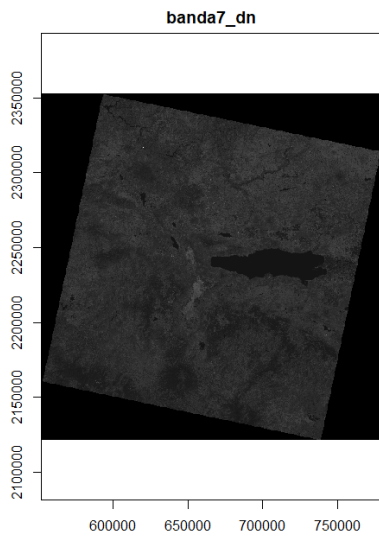
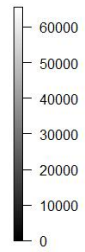
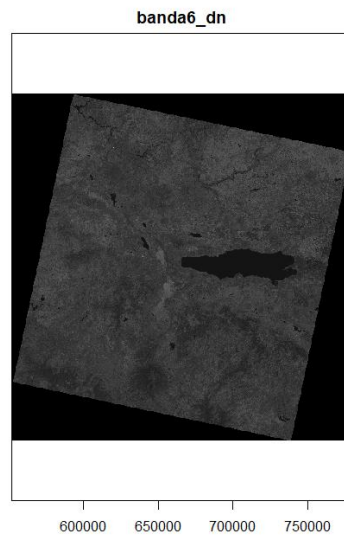
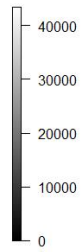
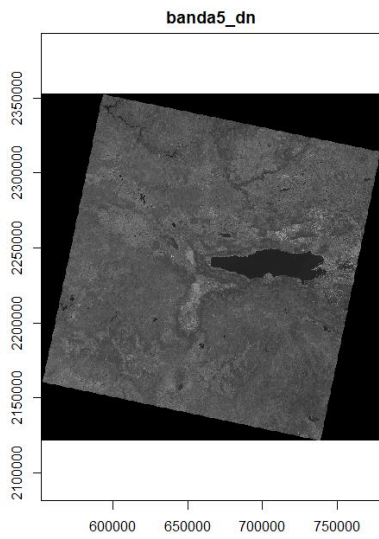
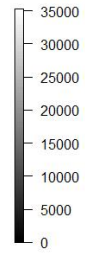
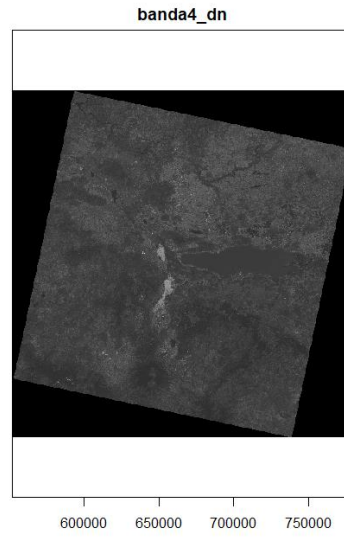
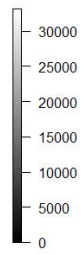
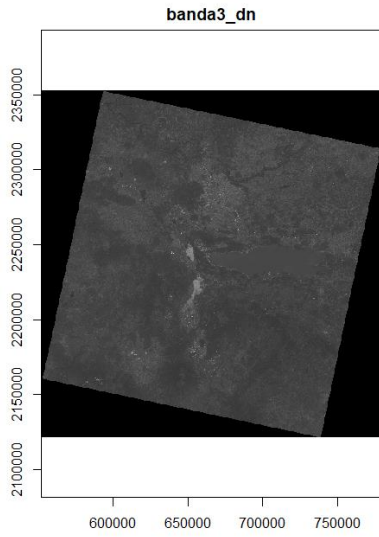
<i>Año</i>	<i>Fecha de adquisición</i>	<i>Satélite</i>
2010	25/01/2020	Landsat-5 TM
2011	01/03/2011	Landsat-5 TM
2013	10/06/2013	Landsat-8 OLI/TIRS
2014	25/03/2014	Landsat-8 OLI/TIRS

2015	28/03/2015	Landsat-8 OLI/TIRS
2016	30/03/2016	Landsat-8 OLI/TIRS
2017	01/03/2017	Landsat-8 OLI/TIRS
2018	20/03/2018	Landsat-8 OLI/TIRS
2019	23/03/2019	Landsat-8 OLI/TIRS
2020	25/03/2020	Landsat-8 OLI/TIRS

Tabla 4 Imágenes Landsat obtenidas

A manera de ejemplo se procede a graficar las bandas espectrales de una de las imágenes descargadas, obtenida por Landsat-8 OLI/TIRS la cual cuenta con 11 bandas espectrales, donde una de ellas es una banda pancromática y otras dos son bandas térmicas. En la Figura 5 se puede observar estas bandas, a excepción de la banda 8 que debido a que tiene otra resolución no fue posible graficar junto con las demás bandas. La resolución espacial de dicha imagen es de 30 metros y la zona a la cual representa es la cuenca de Ayuquila ubicada en el estado de Jalisco al occidente de México.





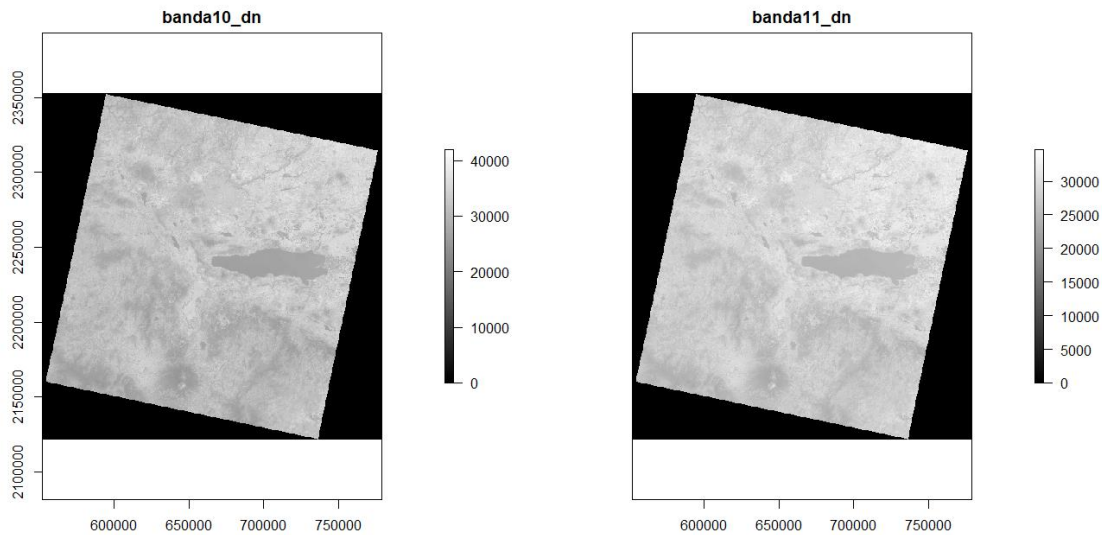


Figura 5. Bandas espectrales de la imagen Landsat 2020

3.2 PRE-PROCESAMIENTO DE DATOS

Para poder extraer información útil de la imagen satelital descargada es necesario hacer cierto tipo de correcciones antes de aplicar cualquier otro método. Esto con la finalidad de que, al realizar procesos posteriores, como el SMA, no se tengan tantos errores, evitando así, pérdidas de tiempo y eficiencia en la finalización del proyecto. Las principales correcciones que se deben aplicar a estas imágenes son las correcciones geométricas y las atmosféricas. El proceso que se realizó para aplicar este tipo de correcciones en las imágenes satelitales obtenidas se describe en los siguientes párrafos.

3.2.1 Corrección geométrica

Generalmente cuando se descarga una imagen satelital de Landsat esta ya viene corregida geométricamente, es decir que cada píxel ya tiene asignado su ubicación geográfica. Para el caso de este proyecto, las imágenes obtenidas no fueron la excepción por lo que no se tuvieron que hacer modificaciones en este sentido. La proyección cartográfica de las imágenes obtenidas está representada en coordenadas UTM Zona 13 usando como datum WGS84 y la ubicación correspondiente dicha imagen dada por su path/row es 29/46 respectivamente.

3.3 RECORTE Y SELECCIÓN DE ZONA DE ESTUDIO

Una vez descargadas las imágenes se pudo identificar que estas abarcaban mucha más área de la necesaria ya que ocupan todo el centro de Jalisco y parte de los estados vecinos y como se puede observar en la Figura 6, la cuenca de Ayuquila es un área relativamente pequeña, abarcando 4,175.54 km² en comparación con los 31,000 km² de área que abarca aproximadamente una imagen satelital de Landsat (Ariza, 2013), por lo que fue necesario hacer un recorte de la imagen para obtener solamente el área de interés para este trabajo. Para la extracción del área se utilizaron librerías del lenguaje de programación R que trabajan con rasters, como la librería raster y rgdal, en específico se usaron las herramientas de crop y mask para poder recortar las imágenes a partir de un archivo shape del área de la cuenca de Ayuquila.

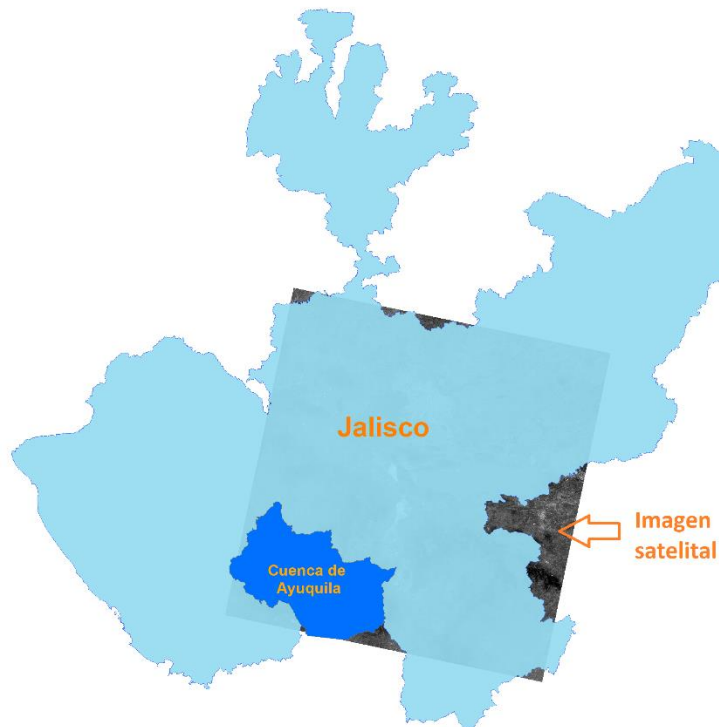


Figura 6. Área de la cuenca de Ayuquila en comparación con la imagen satelital

3.3.1 Conversión de número digital a radiancia

En el capítulo anterior se mencionó como los sistemas de teledetección captan la radiación reflejada por los objetos en la superficie de la tierra, además de que la imagen satelital no es representada por este valor de radiación, sino por un número digital discreto, el cual puede ser reconvertido a su nivel de radiancia a través del número digital y valores de reescalado proporcionados en el archivo de metadatos que viene adjunto al descargar la imagen satelital, dichos valores para la imagen del año 2020 se muestran en la Tabla 5. Para ejemplificar este proceso, se muestra el cálculo para la reconversión de la radiancia a partir del número digital y los valores de reescalado, esto se hace para la banda 2 pero el proceso es el mismo para todas las bandas y todas las imágenes. También es importante mencionar que esto se debe hacer en toda la imagen, es decir, a cada píxel se le debe aplicar la ecuación (23).

$$Rad_{b2} = b2_{dn} \times 1.2922 \times 10^{-2} + (-64.60856) \quad (23)$$

donde $b2_{dn}$ representa el número digital de dl pixel

FACTOR DE REESCALADO MULTIPLICATIVO PARA EL CÁLCULO DE LA RADIANCIA	FACTOR DE REESCALADO ADITIVO PARA EL CÁLCULO DE LA RADIANCIA
RADIANCE_MULT_BAND_1 = 1.2619E-02	RADIANCE_ADD_BAND_1 = -63.09354
RADIANCE_MULT_BAND_2 = 1.2922E-02	RADIANCE_ADD_BAND_2 = -64.60856
RADIANCE_MULT_BAND_3 = 1.1907E-02	RADIANCE_ADD_BAND_3 = -59.53623
RADIANCE_MULT_BAND_4 = 1.0041E-02	RADIANCE_ADD_BAND_4 = -50.20431
RADIANCE_MULT_BAND_5 = 6.1445E-03	RADIANCE_ADD_BAND_5 = -30.72253
RADIANCE_MULT_BAND_6 = 1.5281E-03	RADIANCE_ADD_BAND_6 = -7.64041
RADIANCE_MULT_BAND_7 = 5.1505E-04	RADIANCE_ADD_BAND_7 = -2.57523
RADIANCE_MULT_BAND_8 = 1.1363E-02	RADIANCE_ADD_BAND_8 = -56.81746
RADIANCE_MULT_BAND_9 = 2.4014E-03	RADIANCE_ADD_BAND_9 = -12.00706
RADIANCE_MULT_BAND_10 = 3.3420E-04	RADIANCE_ADD_BAND_10 = 0.10000
RADIANCE_MULT_BAND_11 = 3.3420E-04	RADIANCE_ADD_BAND_11 = 0.10000

Tabla 5. Factores de reescalado para el cálculo de la radiancia de la imagen de 2020

3.3.2 Conversión de número digital a reflectancia

Para poder comparar o identificar los diferentes elementos que conforman nuestra imagen satelital, es necesario tener información de la firma espectral de cada uno, lo cual no es posible usando sólo el número digital de los pixeles, por lo que es importante transformarlo a una escala espectral normalizada, como lo es el nivel de reflectancia (ecuación 24). Este proceso de reconversión es muy importante al momento de realizar una clasificación ya que cuando cada píxel es representado por su valor de reflectancia se tiene información más cercana a la realidad y además permite poder comparar las diferentes firmas espectrales de la imagen satelital. Para ejemplificar el cálculo de la reflectancia, al igual que en el cálculo de la radiancia, se hará para la banda 2, pero de igual forma tener en cuenta que el procedimiento es el mismo en todas las bandas.

$$Ref_{b2} = b2_{dn} \times 2 \times 10^{-5} + (-0.1) \quad (24)$$

donde $b2_{dn}$ representa el número digital de cada pixel

FACTOR DE REESCALADO MULTIPLICATIVO PARA EL CÁLCULO DE LA REFLECTANCIA	FACTOR DE REESCALADO ADITIVO PARA EL CÁLCULO DE LA REFLECTANCIA
REFLECTANCE_MULT_BAND_1 = 2.0000E-05	REFLECTANCE_ADD_BAND_1 = -0.100000
REFLECTANCE_MULT_BAND_2 = 2.0000E-05	REFLECTANCE_ADD_BAND_2 = -0.100000
REFLECTANCE_MULT_BAND_3 = 2.0000E-05	REFLECTANCE_ADD_BAND_3 = -0.100000
REFLECTANCE_MULT_BAND_4 = 2.0000E-05	REFLECTANCE_ADD_BAND_4 = -0.100000
REFLECTANCE_MULT_BAND_5 = 2.0000E-05	REFLECTANCE_ADD_BAND_5 = -0.100000
REFLECTANCE_MULT_BAND_6 = 2.0000E-05	REFLECTANCE_ADD_BAND_6 = -0.100000
REFLECTANCE_MULT_BAND_7 = 2.0000E-05	REFLECTANCE_ADD_BAND_7 = -0.100000
REFLECTANCE_MULT_BAND_8 = 2.0000E-05	REFLECTANCE_ADD_BAND_8 = -0.100000
REFLECTANCE_MULT_BAND_9 = 2.0000E-05	REFLECTANCE_ADD_BAND_9 = -0.100000

Tabla 6. Factores de reescalado para el cálculo de la reflectancia de la imagen de 2020

3.3.3 Corrección atmosférica por el método de Dark Object Subtraction (DOS)

En la Figura 1 se representa la imagen satelital de la cuenca de Ayuquila sin corrección atmosférica, por lo que los pixeles están representados por su número digital, es decir por valores discretos y no por su nivel de reflectancia, para poder confirmar esto, en la Figura 7

se pone la información y características propias de dicha imagen, por ejemplo, podemos ver que el tipo de imagen es un RasterBrick, es decir un ráster con varias capas, también se ve la dimensión o tamaño de la imagen, su resolución espacial, que es de 30 metros, la extensión que abarca y el sistema de coordenadas usado, en los últimos tres valores, que son los que nos interesan para ver por qué valores están representados los píxeles de la imagen satelital, los cuales en este caso están representados por números digitales (DN), se puede ver también el valor digital mínimo y máximo que toman los píxeles en cada banda, de esta forma confirmamos que se trata de una imagen satelital que aún no está corregida atmosféricamente.

```
> print(imagen)
class       : RasterBrick
dimensions  : 2700, 3270, 8829000, 10  (nrow, ncol, ncell, nlayers)
resolution  : 30, 30  (x, y)
extent      : 554985, 653085, 2146425, 2227425  (xmin, xmax, ymin, ymax)
crs         : +proj=utm +zone=13 +datum=WGS84 +units=m +no_defs +ellps=WGS84 +towgs84=0,0,0
source      : C:/Users/danii/AppData/Local/Temp/RtmpCCVgpb/raster/r_tmp_2020-04-12_122240_13056_42962.grd
names       : B1_dn, B2_dn, B3_dn, B4_dn, B5_dn, B6_dn, B7_dn, B9_dn, B10_dn, B11_dn
min values  : 0, -30877, -19568, -27242, -32592, -32554, -32675, 0, -32767, 0
max values  : 30745, 27629, 30046, 26363, 31399, 32764, 32132, 5341, 32767, 31339
```

Figura 7. Características de la imagen de 2020 antes de la corrección atmosférica

Para aplicar el proceso de corrección atmosférica a la imagen satelital, fue necesario utilizar algunas funciones de la librería RStoolbox Versión 0.2.6 en R [<https://cran.r-project.org/web/packages/RStoolbox/RStoolbox.pdf>] como estimateHaze y radCor, la primera de estas funciones permitió obtener el valor digital de los objetos oscuros, así como la frecuencia de cada uno de estos para las bandas en el espectro de onda visible e infrarrojo (Bandas de la dos a la siete, en Landsat 8), esto último es debido a que la turbidez atmosférica se presenta en mayor medida a estas bandas. La Figura 8 muestra los resultados obtenidos.

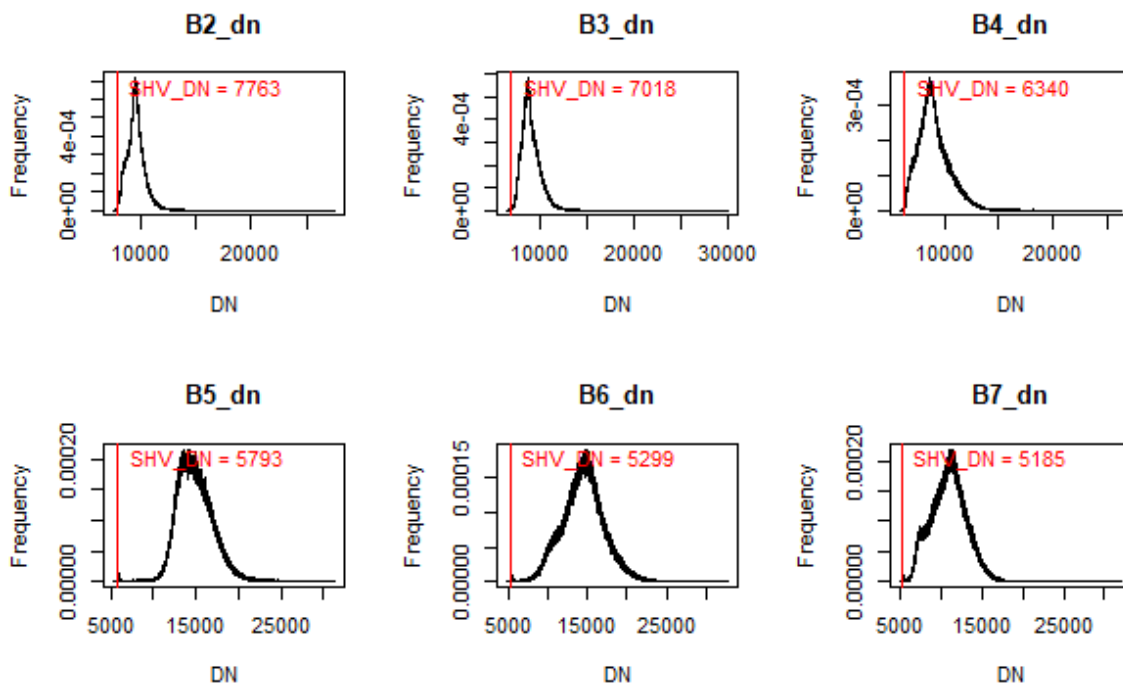


Figura 8. Frecuencia de objetos oscuros en la imagen de 2020

Después de haber calculado el número digital de los objetos oscuros para cada banda, se procedió a realizar la corrección atmosférica usando la función radCor, la cual como también ya mencionamos pertenece a la librería RStoolbox de R. Esta función tiene implementados varios métodos de calibración y corrección de datos Landsat, entre los cuales podemos encontrar “rad”, “apref”, “costz”, “dos” y “sdos”, los últimos tres métodos sirven para aplicar la corrección atmosférica por el método de sustracción de objetos oscuros (DOS, por sus siglas en inglés) (Leutner, et al., 2019), este método como se explicó en el capítulo anterior, se basa en el supuesto de que la reflectancia de los objetos oscuros es muy cercana a cero, y para objetos completamente negros, es cero, esto es un componente muy importante en la dispersión atmosférica por lo que el método trata de obtener el valor digital de los objetos más oscuros (donde el usuario da la proporción estimada) para posteriormente eliminar dicho valor o valores de toda la imagen satelital (Chavez, 1996). Sin embargo, el método que se utilizó en este proyecto fue el método SDOS el cual es igual al método DOS, pero con la característica de que este método requiere como parámetro los valores de turbidez para

cada banda por separado (Leutner, et al., 2019), dichos valores ya han sido calculados con la función estimateHaze, en la Figura 8, podemos ver el nivel de turbidez en cada banda, se denotan como SHV_DN.

Una vez aplicada la función radCor con el método sdos, los píxeles de la imagen ya se cambian por un valor digital (DN) discreto sino por el nivel de reflectancia (SRE, Surface Reflectance) de cada uno, esto mismo lo podemos comprobar en la Figura 9, donde si comparamos con la Figura 7, vemos el cambio en estos valores. De igual forma, se pueden visualizar estos cambios en la Figura 10, donde se grafican cada una de las bandas con su correspondiente nivel de reflectancia, esta imagen es similar a la Figura 5, pero recordemos que en esta última sus valores vienen dados por el número digital de los píxeles y no de la reflectancia, además de que está aún no está recortada al área de interés por lo que se ve la imagen completa. También se puede notar que solo se grafican las bandas de la dos a la siete, esto es porque la turbidez atmosférica afecta casi exclusivamente el rango de longitud de onda visible, y por lo general solo se estima la neblina en bandas azules, verdes y rojas y ocasionalmente también en la banda de infrarrojo cercano (Leutner, et al., 2019).

Recordemos que todos estos procesos de corrección, así como los posteriores donde aplicaremos el algoritmo SMA se realiza en cada una de las imágenes, desde 2010 hasta 2020, por lo que se tienen algunas pequeñas variaciones dependiendo del satélite del cual se obtuvo la imagen satelital, ya que cambian un poco las bandas usadas por cada uno, que, aunque esto es algo lógico, es importante tenerlo presente a la hora de trabajar con estas imágenes.

```
> print(imagen_sdos)
class      : RasterStack
dimensions : 2700, 3270, 8829000, 6  (nrow, ncol, ncell, nlayers)
resolution : 30, 30  (x, y)
extent     : 554985, 653085, 2146425, 2227425  (xmin, xmax, ymin, ymax)
crs        : +proj=utm +zone=13 +datum=WGS84 +units=m +no_defs +ellps=WGS84 +towgs84=0,0,0
names      : B2_sre, B3_sre, B4_sre, B5_sre, B6_sre, B7_sre
min values : 0, 0, 0, 0, 0, 0
max values : 0.4752918, 0.5571616, 0.4814372, 0.6106327, 0.6226911, 0.5941030
```

Figura 9. Características de la imagen de 2020 después de la corrección atmosférica

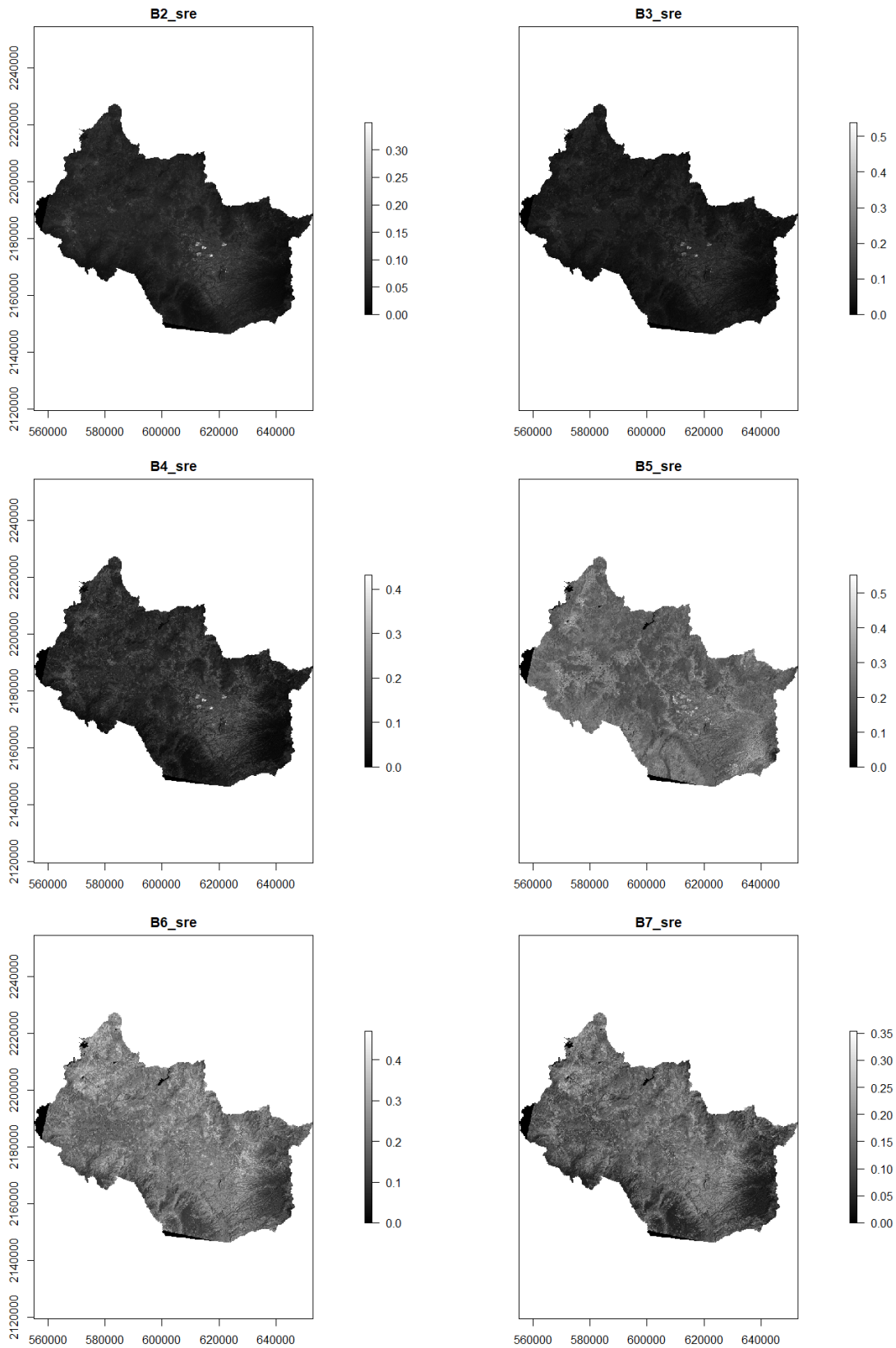


Figura 10. Bandas espectrales representadas por su nivel de reflectancia, del año 2020

3.4 SMA Y NDFI

Hasta este punto, se han realizado todas las correcciones necesarias de las imágenes satelitales, ahora es importante conocer qué áreas dentro de estas imágenes representan áreas como vegetación, vegetación no fotosintética, sombra y suelo, esto con el fin de identificar los píxeles puros, mismos que servirán como insumo para la aplicación del algoritmo de análisis de mezclas espectrales (SMA) con el cual se pretende detectar la degradación forestal a través de un análisis de serie de tiempo con el indicador NDFI y para obtener una mejor comprensión de los resultados también se realizará con los índices NDMI y NDWI.

3.4.1 Identificación de miembros puros

Haciendo referencia a lo expuesto en el capítulo uno, cabe recordar que la degradación no se refiere a lo mismo que la deforestación, ya que está no es un cambio de uso de suelo, sino que solo se presenta un desgaste o daño a la cubierta forestal, algunas de las causas que generan degradación también fueron expuestas en ese capítulo, sin embargo, es necesario mencionar que una de las principales es el sistema de cultivo roza-tumba-quema, conocido también como “coamil” en la zona de Ayuquila, este sistema se practica generalmente en laderas, donde es difícil trabajar con tractores o con arado de tierra, recordemos que el proceso que sigue este sistema es cortar y dejar secar la vegetación para posteriormente quemarla, después viene el cultivo de las semillas, las cuales generalmente en esta zona, son de Maíz, y una vez que se cosecha, se deja descansar el terreno para recuperar la fertilidad del suelo (Parera, et al., 2001). La duración de la etapa de descanso es variable, pero generalmente va de los 8 a los 10 años. En esta etapa de descanso crece de nuevo la vegetación, pero los ejidatarios no permiten que crezcan árboles muy grandes. Cabe destacar que este sistema de agricultura se da con una mayor frecuencia en bosques tropicales secos (selva baja) que, en bosques de encino y bosques templados, donde se potencializa aún más en áreas cercanas a las áreas urbanas, áreas con una menor inclinación y áreas que presentan una mayor elevación. Debido a las características de este sistema, no podemos hablar de deforestación, ya que el cambio de uso de suelo que se da no es permanente, esto es porque

con el regreso de la cubierta forestal, también regresa la fertilidad del suelo, por ello se dice que el sistema de agricultura roza-tumba-quema es un proceso que causa degradación pero no deforestación, ya que aunque la cubierta forestal regresa, cada vez su calidad o salud disminuye (Mwampamba, et al., 2015).

Como se comentó anteriormente, para aplicar el algoritmo SMA, es necesario identificar los miembros puros, los cuales están dados por píxeles que solo representan una clase. Para esto, se seleccionó una parte de la imagen satelital en donde se pudo identificar con una mejor visibilidad dichos píxeles, para la selección de esta área se tomó en cuenta que la zona tuviera partes de vegetación, vegetación no fotosintética, sombra y suelo, esto con el fin de identificar los píxeles puros que serán necesarios para detectar la degradación forestal. En la Figura 11 se puede observar la parte de la imagen extraída, además ahí mismo se identifican las clases mencionadas previamente.

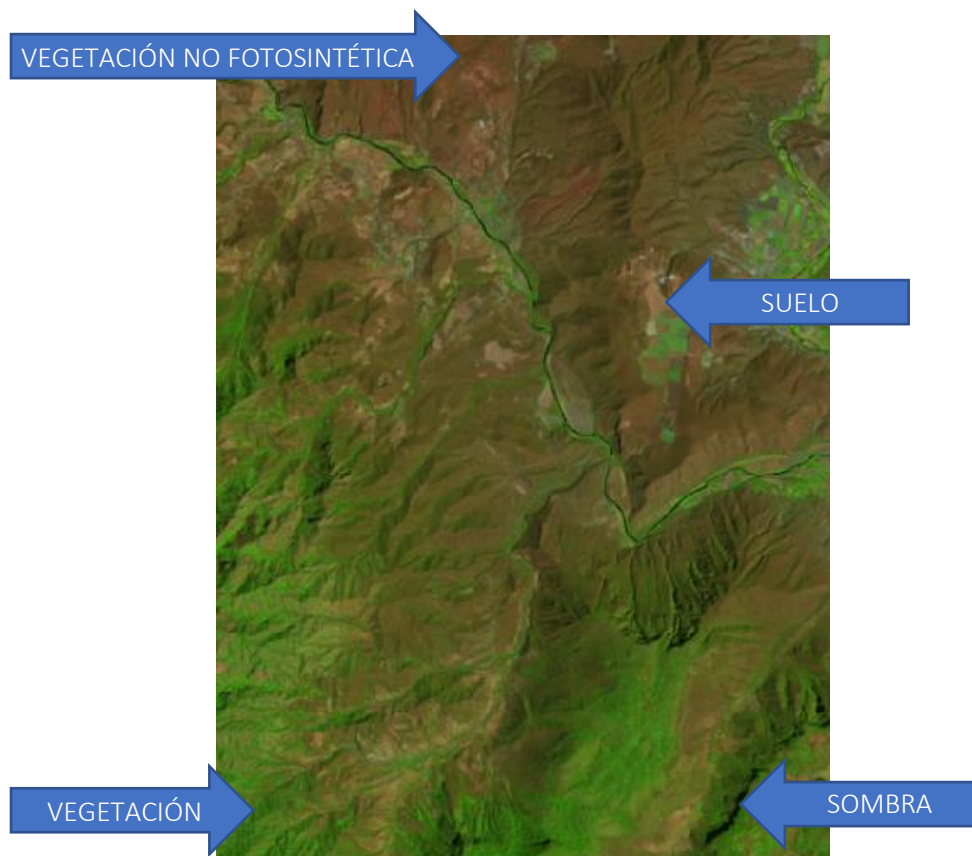


Figura 11. Identificación visual de clases

Una vez identificadas las diferentes clases existentes (vegetación, vegetación no fotosintética, sombra y suelo), se procedió a identificar los valores de los píxeles puros que mejor representaban a estas clases, para esto se identificó de forma visual aquellos píxeles que solo estaban compuestos por una sola clase. En Figura 12 se muestra la ubicación de estos píxeles y en la Tabla 7 podemos ver sus valores espectrales en las seis bandas de las que está compuesta la imagen. La selección de estos píxeles se tuvo que realizar en cada una de las imágenes, es decir desde 2010 hasta 2020, sin embargo, para fines demostrativos sólo se presentan aquí los píxeles puros del año 2020.



Figura 12. Selección de los miembros puros

<i>Banda</i>	Vegetación	Vegetación no fotosintética	Sombra	Suelo
Banda 2	0.0266059	0.0562809	0.015832	0.0798195
Banda 3	0.0359942	0.0593035	0.00605572	0.0985562
Banda 4	0.0225965	0.0745362	0.00585612	0.135588
Banda 5	0.429711	0.185667	0.0382653	0.257492
Banda 6	0.150784	0.264382	0.0181904	0.34676
Banda 7	0.0609395	0.169963	0.00808075	0.245529

Tabla 7. Valores espectrales de los miembros puros para el año 2020

3.4.2 La aplicación de SMA en R

Recordemos que el algoritmo SMA requiere como parámetros los valores espectrales de los miembros puros para cada una de las clases, esto para poder descomponer la imagen satelital en las imágenes fracción correspondientes a dichas clases. Como estos valores ya fueron obtenidos (Tabla 7) ya es posible aplicar el algoritmo el cual viene implementado en el lenguaje de programación R y se puede encontrar dentro de la librería RStoolbox. El algoritmo es llamado en R como MESMA (Multiple Endmember Spectral Mixture Analysis) y requiere como parámetros: la imagen satelital multibanda a analizar, la matriz con los valores de los miembros puros en donde las columnas representan la banda espectral y las filas representan la clase, el método para realizar la desmezcla espectral (actualmente esta función sólo reconoce el método NNLS, Non Negative Least Squares), la cantidad de iteraciones (por defecto son 400) y la tolerancia (número cercano a cero, por defecto es $1e-8$). Una vez que se establecen todos los parámetros necesarios y ejecutado el algoritmo, este regresará un RasterBrick el cual contendrá una banda por cada miembro final, donde el valor de cada píxel estará representado por la probabilidad de presencia estimada del miembro final, dicho valor estará dentro del intervalo de 0 a 1, también incluirá otra banda que representará el error cometido (RMSE). Cabe mencionar que el valor de los pixeles puede salir un poco del rango antes dicho, esto dependiendo del número de iteraciones establecidas, así como del error

permitido, por lo que es importante ir modificando estos parámetros hasta encontrar mejores resultados (Leutner, et al., 2019).

```
1  mesma(img, em, method = "NNLS", iterate = 400, tolerance = 1e-08,  
2  ..., verbose)
```

Figura 13. Ejemplo de uso de SMA obtenido de (Leutner, et al., 2019)

- img: Imagen satelital
- em: matriz con valores de miembros puros
- method: método de desmezcla
- iterate: número de iteraciones
- tolerance: tolerancia mínima
- verbose: valor lógico, si es verdadero imprime el proceso de ejecución

3.4.3 Resultados del SMA

Después de haber aplicado el algoritmo SMA en R, se obtuvo como se esperaba un RasterBrick compuesto de cinco bandas, donde las primeras cuatro representan las imágenes fracción de cada una de las clases analizadas (vegetación, vegetación no fotosintética, sombra y suelo) y la última banda representa el error cometido (RMSE). Donde, en la Figura 14 se observan las imágenes fracción obtenidas en diferentes tonalidades, los píxeles más oscuros representan una proporción mayor a la clase correspondiente y, por el contrario, los más claros sugieren la poca o nula presencia de esta. Cabe resaltar que este cálculo se hizo para los 10 años que abarca nuestro estudio, es decir, desde 2010 a 2020.

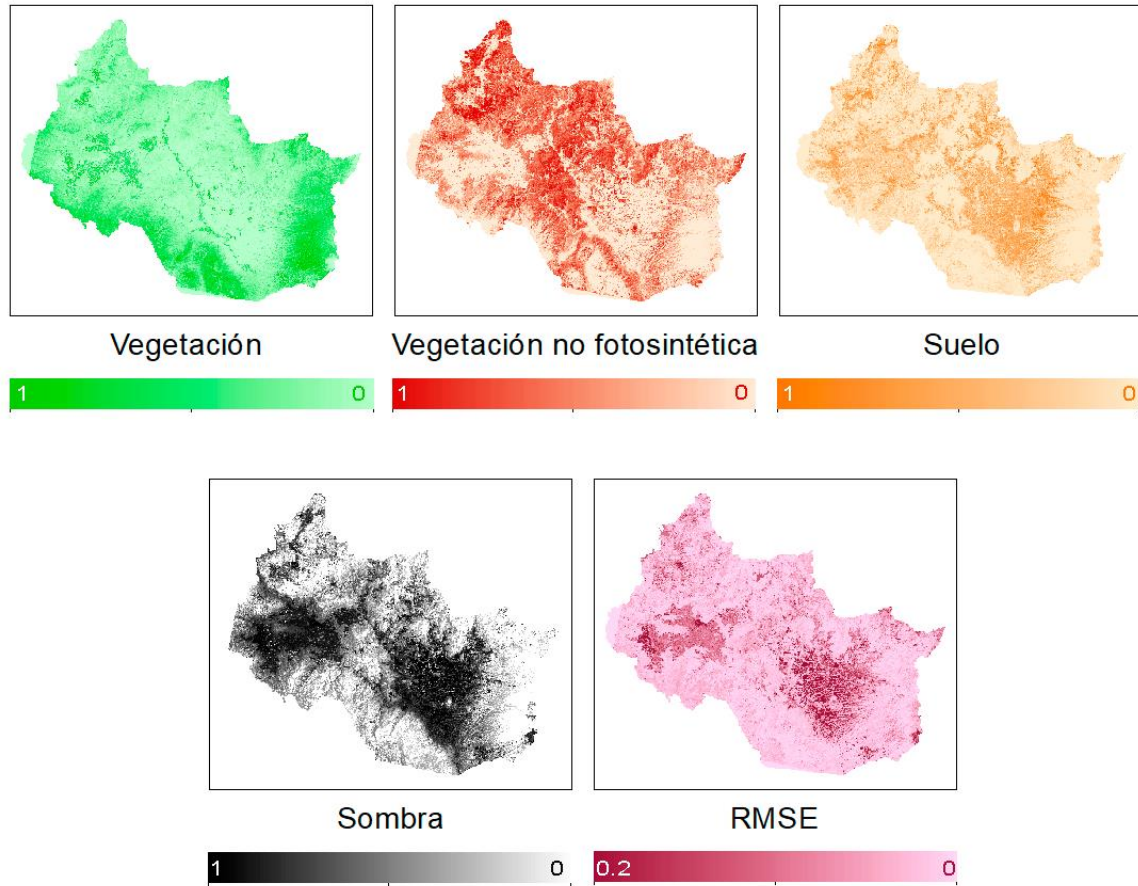


Figura 14. Resultados SMA 2020

3.4.4 Índices NDFI

Una vez obtenidas las imágenes fracción de las clases de vegetación, vegetación no fotosintética, sombra y suelo, se procedió a calcular el índice NDFI, el cual, como se explicó en el capítulo anterior es muy útil para medir los cambios en la cubierta forestal ya que toma en cuenta la consideración de que un píxel es la mezcla de diferentes clases. El resultado se muestra en escala de grises, donde los colores más claros representan valores más altos de este indicador, y los colores más oscuros representan valores más bajos.

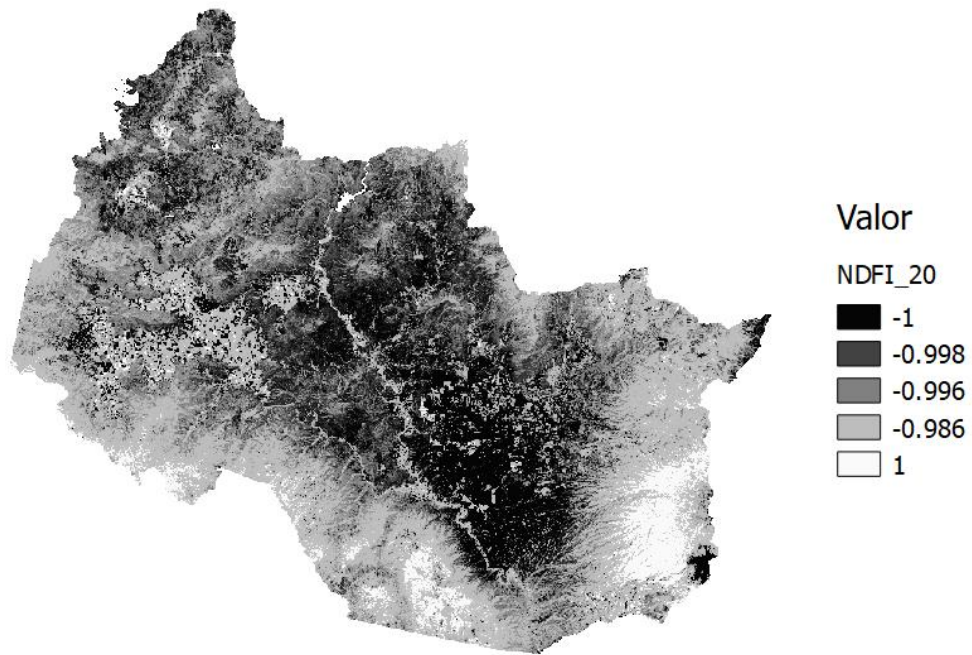


Figura 15. NDFI 2020

3.4.5 Cálculo de Índices NDVI y NDMI

Para obtener una explicación más clara de lo obtenido por el NDFI, se procedió también a calcular los índices NDMI y NDVI. Mismos que nos indican el nivel de humedad y vegetación presente en la imagen satelital, de esta forma podremos interpretar de mejor manera los resultados obtenidos.

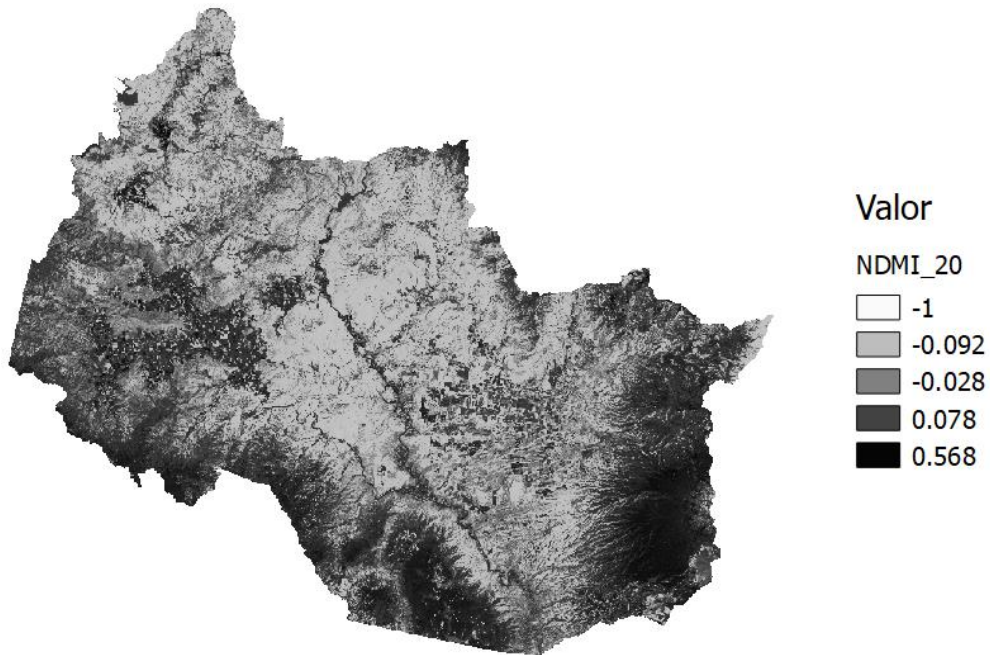


Figura 16. NDMI 2020

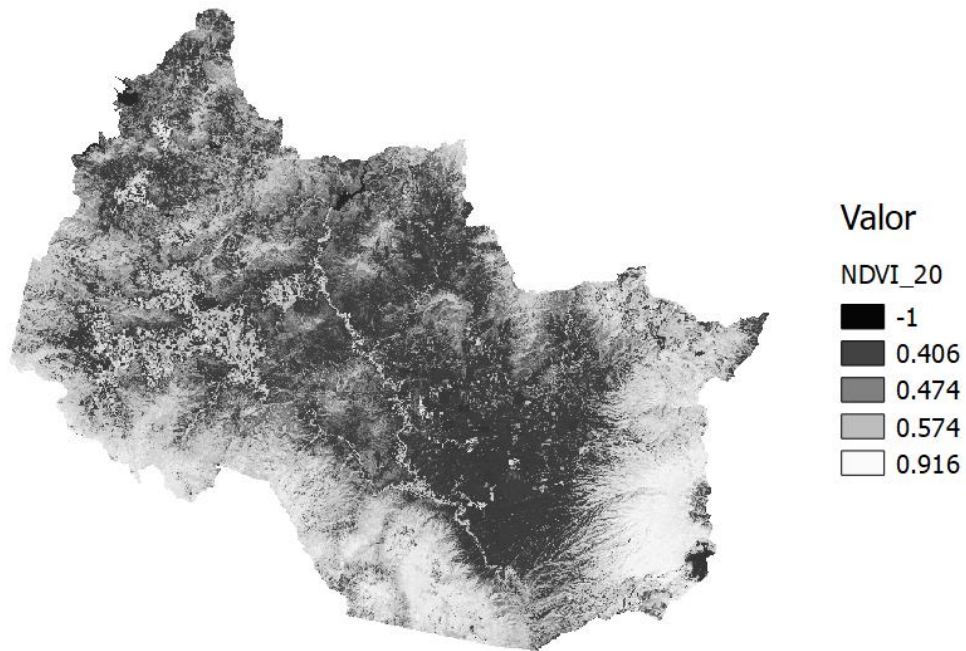


Figura 17. NDVI 2020

3.5 ANÁLISIS DE LA TENDENCIA DE VEGETACIÓN CON EL SERIE DE TIEMPO Y EL MODELO DE REGRESIÓN LINEAL

Todos los pasos anteriores se tuvieron que realizar 10 veces, una vez por cada año de nuestro estudio, desde 2010 hasta 2020. De esta forma se obtuvieron 10 imágenes del NDFI, 10 del NDMI y 10 del NDVI, además se tomaron también las 10 imágenes fracción de la clase de bosque. Con dichos datos se realizó un análisis de serie de tiempo por cada índice, es decir, 4 en total. Este análisis se realizó mediante la utilización de la función llamada “lm” propia del lenguaje R, la cual es una función que permite ajustar modelos lineales, que resultó bastante útil para realizar las regresiones necesarias para nuestro análisis de serie de tiempo. Los parámetros que nos pide la función lm es la variable dependiente y la variable independiente, las cuales en nuestro estudio son los valores que toma el índice, ya sea el NDVI, NDMI, NDFI o bien la imagen fracción de vegetación y el tiempo, respectivamente. Resaltando que los valores para cada índice vienen dados por una matriz que contiene los valores de cada uno de los píxeles que conforman la imagen satelital, los cuales son en total 4,302,860. Entre los valores que regresa la función lm de R, está la pendiente, el coeficiente de determinación r^2 y el nivel de significancia. Con los resultados obtenidos de la pendiente se pudieron determinar los píxeles sujetos a una posible degradación forestal, y además se logró identificar que tan bien se ajustaron los datos a nuestro modelo, esto mediante los valores obtenidos por el coeficiente de determinación y el de significancia. En la Figura 18 podemos observar un ejemplo de la aplicación del modelo lm en el lenguaje R, recalcando que este ejemplo es solo para un píxel, y en el desarrollo del estudio, se tuvo que aplicar para los más de 4 millones de ellos. Se puede observar que la función nos regresa los valores que resultan de interés para este estudio, por ejemplo, el valor de la pendiente, que es de -0.17005, lo que indica que hubo un decremento del nivel de vegetación en ese píxel. También nos indica el valor obtenido por el coeficiente de determinación R^2 que en este caso es de 0.7155, lo que nos indica un nivel aceptable de correlación entre el tiempo y el NDFI, por último, el valor p, que es de 0.002042 el cual es menor a 0.05, lo que indica una significancia bastante alta. Para ver esto de forma visual, en la Figura 19 se muestra la gráfica de los valores del NDFI durante el periodo de estudio, así como también la recta de la regresión para estos puntos. Se puede

apreciar como los puntos se ajustan bastante bien a la recta de regresión y podemos ver como esta, tiene una pendiente negativa, indicando en este caso que el píxel tuvo una disminución en el porcentaje de vegetación presente.

```

Call:
lm(formula = x ~ año)

Residuals:
    Min       1Q   Median       3Q      Max
-0.040962 -0.028403 -0.001649  0.020632  0.050758

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.789852   0.023525  33.575 6.76e-10 ***
año         -0.017005   0.003791  -4.485 0.00204 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.03444 on 8 degrees of freedom
Multiple R-squared:  0.7155,    Adjusted R-squared:  0.6799
F-statistic: 20.12 on 1 and 8 DF,  p-value: 0.002042

```

Figura 18. Ejemplo de aplicación de lm en R

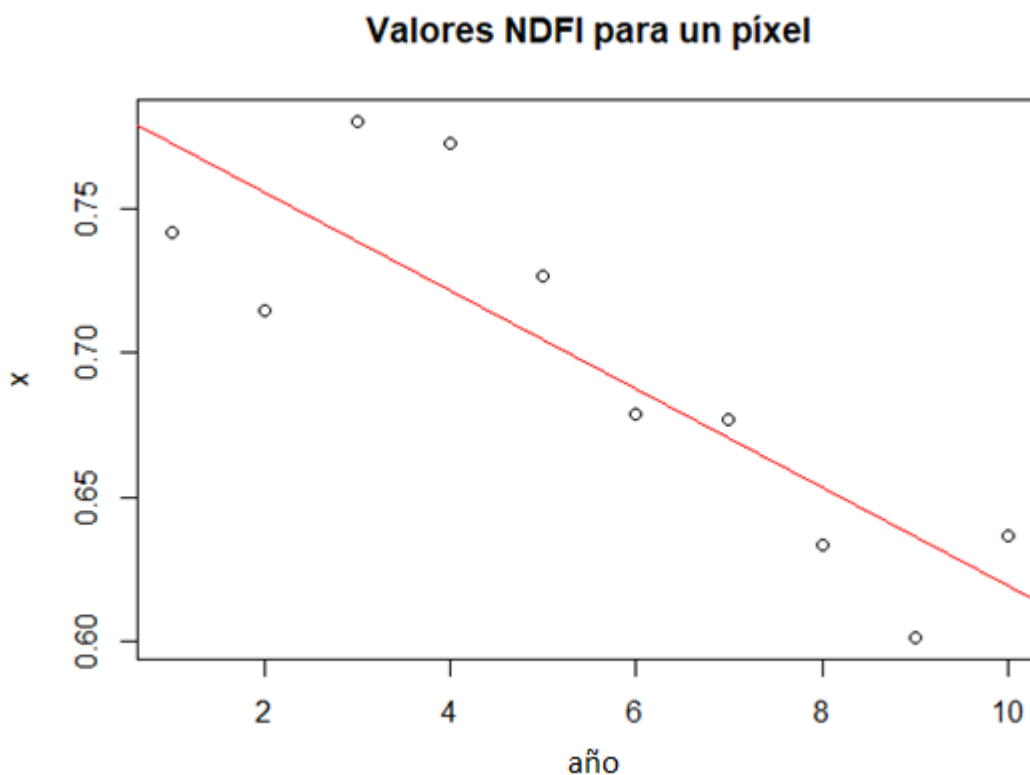


Figura 19. Cálculo de la pendiente del NDFI para un píxel de la imagen

Para la pendiente de la recta de tendencia se obtuvieron 4 imágenes, una por cada índice y otra de la imagen fracción de vegetación, en las cuales, como se muestra en la Figura 20, se identifican aquellas áreas que presentan degradación forestal (pendiente negativa), áreas que se han mantenido estables (pendiente cercana a cero) y áreas que han presentado una regeneración en cuanto a degradación se refiere (pendiente positiva). En dicha imagen se puede notar que los píxeles con valores negativos en el NDFI son mayores en cantidad con respecto al NDVI y NDMI, una posible explicación de esto sería debido a que el NDFI como se mencionó con anterioridad es mejor para identificar degradación forestal que otros índices. También se puede observar que para la imagen fracción de vegetación se detectan aún más píxeles con pendiente negativa. Por otro lado, se observa que el NDVI logra identificar mayor cantidad de píxeles con valores positivos para la pendiente. En la Figura 21 se muestran los resultados para el nivel de significancia, donde los píxeles con valores comprendidos entre 0 y 0.05 representan una buena relación entre la variable dependiente e independiente, en este caso entre valor del índice (NDVI, NDMI y NDFI) o imagen fracción y el tiempo. En la imagen, se alcanza a percibir que dichos valores se encuentran localizados espacialmente en las áreas que hay mayor cobertura forestal.

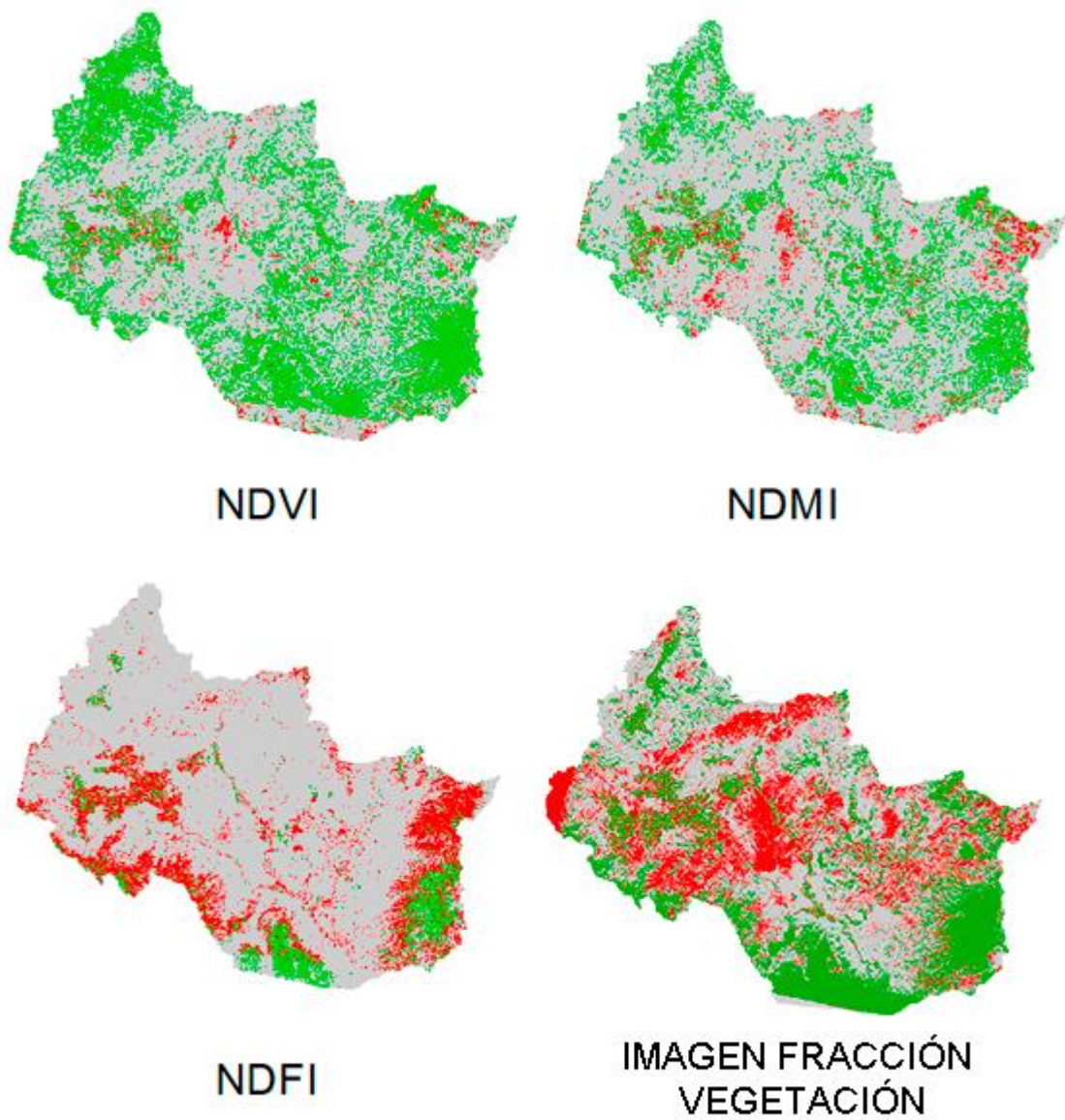


Figura 20. Valores de la pendiente

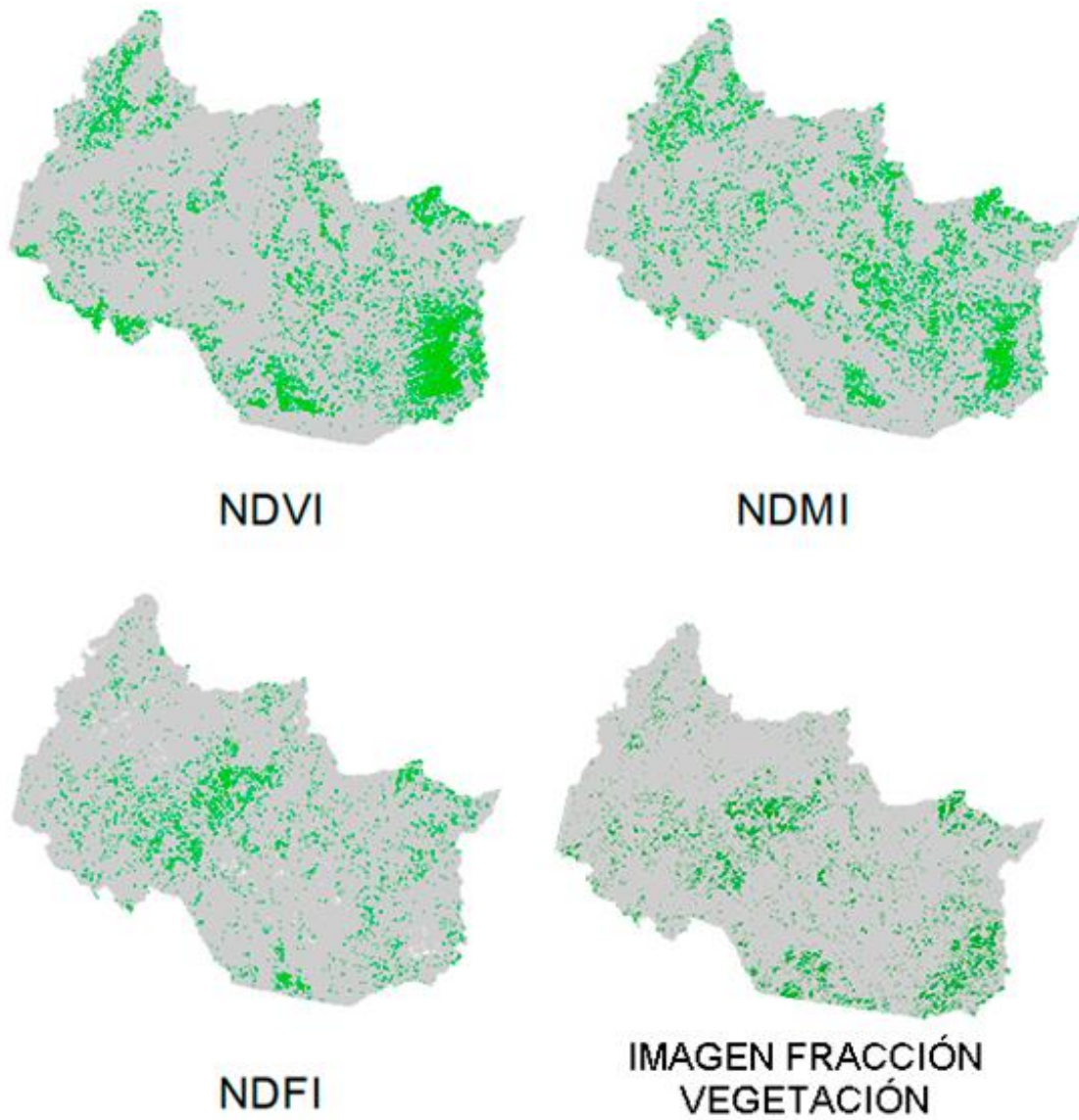


Figura 21. Nivel de significancia

En la Figura 22 se puede observar el coeficiente de determinación R^2 obtenido a partir del análisis de serie de tiempo para el NDVI, NDMI, NDFI e imagen fracción de vegetación. En dicha imagen se puede notar que al igual que para el nivel de significancia, los valores más altos tienden a estar dentro de áreas de bosque y vegetación. También se puede ver que la cantidad de píxeles con un valor mayor del coeficiente de determinación es superior en el NDVI y NDMI que para el NDFI. Por último, en la Figura 23 se puede observar el resultado de la prueba de Mann-Kendall para el análisis de serie de tiempo, donde los valores más rojos indican una fuerte significancia decreciente y por el contrario los valores más verdes indican una fuerte significancia creciente. También se puede notar que al igual que en la pendiente, las áreas de color rojo son mayores en el NDFI, que, en los otros índices, una explicación de esto sería muy similar a lo que se explicó anteriormente en el cálculo de la pendiente. De igual forma se observa que hay más píxeles de color verde oscuro en el NDVI y NDMI.

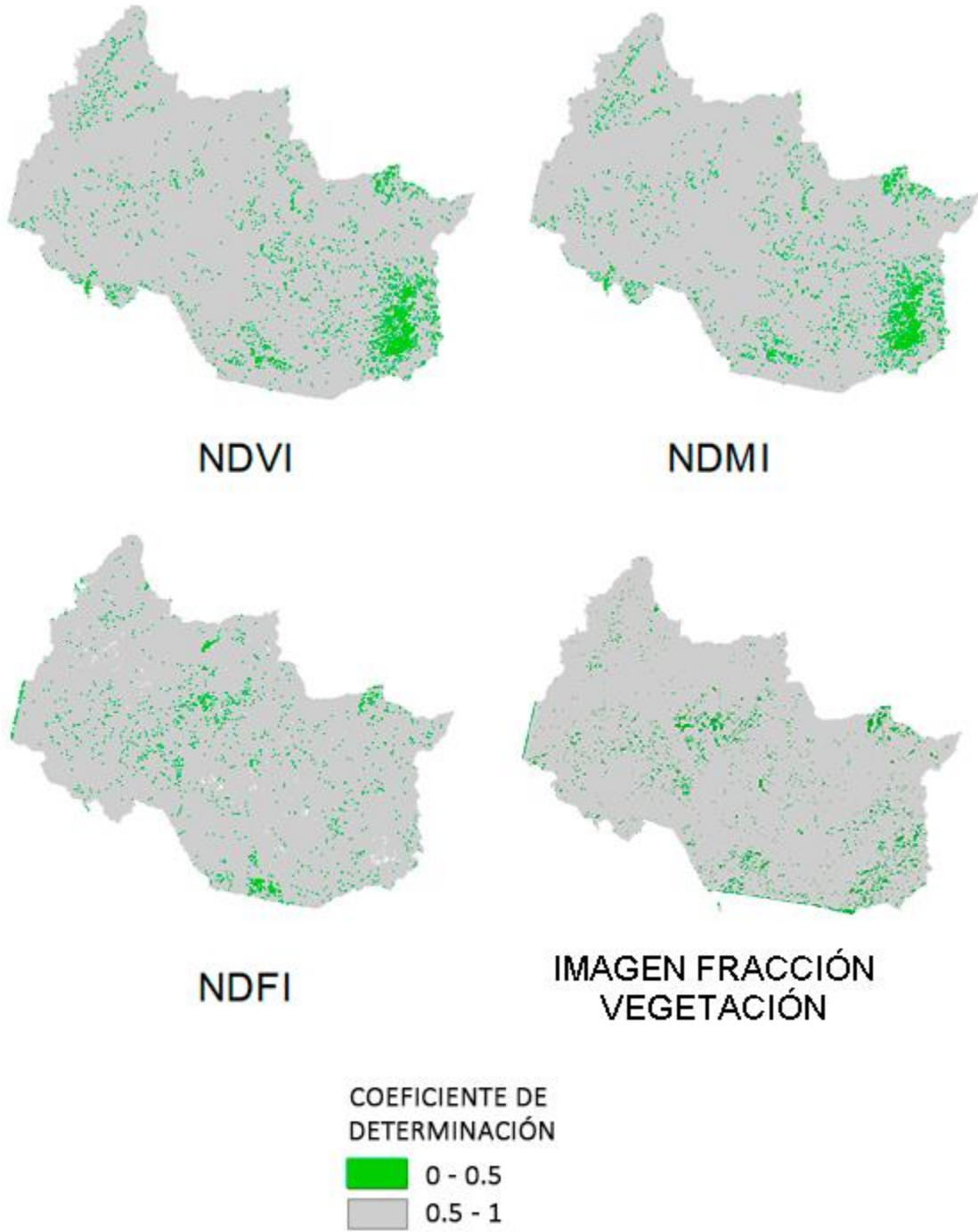


Figura 22. Coeficiente de determinación R^2

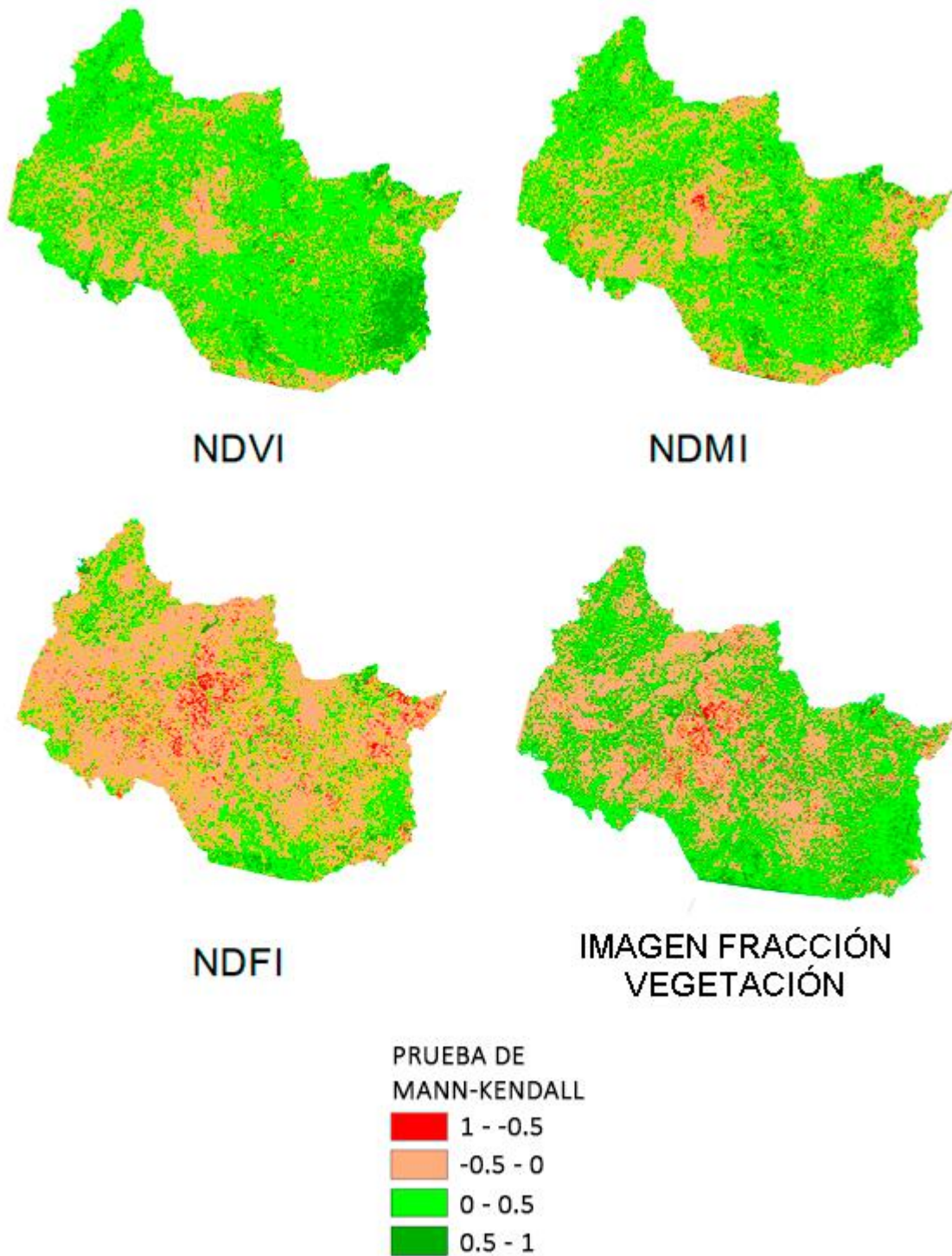


Figura 23. Prueba de Mann-Kendall

3.6 APLICACIÓN DE MÁSCARA

Teniendo en cuenta que para el análisis de degradación forestal se requieren únicamente aquellas áreas que desde 2010 hasta 2020 continúan representando la clase de bosque, se llegó a la conclusión de aplicar una máscara para eliminar de la zona de estudio todos aquellos píxeles que no fueran de esta clase. La máscara se aplicó tomando en cuenta tanto a la imagen de 2010 como a la de 2020. Esto debido a que, si solo usáramos una, por ejemplo, la de 2010, si bien, tendríamos la seguridad de que todos los píxeles fueron píxeles de bosque y podríamos medir como este ha perdido densidad y calidad, también, estaríamos admitiendo aquellos píxeles que para 2020 ya no representan a la clase de bosque, es decir aquellos píxeles en lo que hubo un cambio, en otras palabras, en las áreas en las que hubo deforestación, lo cual no es el interés de este estudio. En cambio, si se toman también los datos de la imagen de 2020 para la aplicación de la máscara, tendríamos la seguridad de que durante el periodo de estudio los píxeles seleccionados siguen teniendo la clase de Bosque, y de esta forma si podemos medir la degradación que estos han sufrido con el transcurso del tiempo.

3.6.1 Mapeo de la cobertura de suelo con clasificación supervisada de imagen Landsat

Para la aplicación de la máscara se realizó la clasificación de la imagen satelital usando dos algoritmos de clasificación supervisada, el de máxima verosimilitud y random forest, esto con el fin de comparar cuál de los dos métodos daba mejores resultados. Para poder realizar dicha clasificación se tuvieron que obtener datos de entrenamiento, los cuales fueron obtenidos mediante la selección de áreas dentro de la imagen satelital que formaran parte de las principales clases que conformaban la zona de estudio, estas clases fueron obtenidas mediante interpretación visual y con ayuda de datos de INEGI, 2009, los cuales fueron obtenidos del portal de CONABIO [http://www.conabio.gob.mx/informacion/metadata/gis/usv250s6gw.xml?_httpcache=yes&_xsl=/db/metadata/xsl/fgdc_html.xsl&_indent=no]. En total se seleccionaron 78 polígonos divididos entre las diferentes clases identificadas, los cuales se muestran en la Tabla 8.

Clase	Datos de entrenamiento 2010	Datos de entrenamiento 2020
Bosque	10	10
Cuerpo de Agua	1	2
Selva Baja	22	23
Áreas sin vegetación aparente	2	2
Vegetación con cultivo	10	11
Vegetación seca	10	11
Vegetación sin cultivo	15	15
Vegetación de temporal	6	6
Área urbana	2	2

Tabla 8. Clases identificadas

Después de la selección de los datos de entrenamiento, es importante medir qué tan diferente es la respuesta espectral de cada una de las clases, para esto se procedió evaluar la separabilidad de los datos de entrenamiento, creando 1366 puntos aleatorios dentro de los polígonos previamente obtenidos y se extrajeron sus valores espectrales correspondientes a las 6 bandas usadas en la imagen. Estos valores fueron promediados con respecto a su clase, con lo cual se realizó una gráfica en la que se puede observar que tan bien resulta su separabilidad, además también se calculó la desviación estándar para identificar si existe solapamiento entre los valores obtenidos. Este proceso se realizó tanto para los datos de 2010 como para los de 2020 (Figura 24 y Figura 25). Con dicha gráfica se puede observar que la respuesta espectral de los diferentes polígonos de entrenamiento para las diferentes clases y en las diferentes bandas que conforman la imagen satelital si presentan diferencias considerables por lo que se procedió a realizar la clasificación con mayor seguridad de obtener buenos resultados.

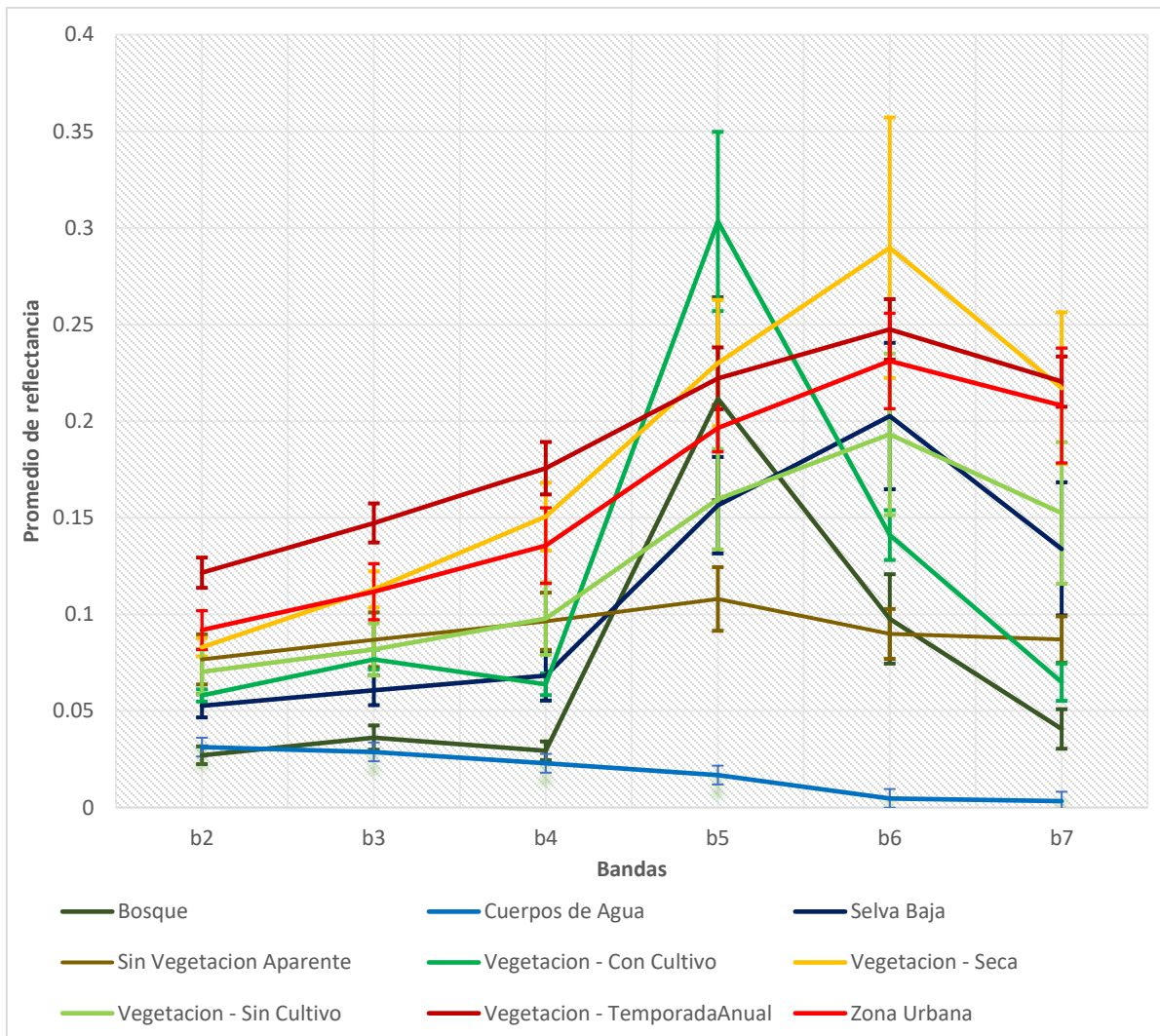


Figura 24. Información espectral de datos de entrenamiento 2010. (Basado en 1366 pixeles en total de las clases de entrenamiento, de los cuales corresponden a bosque: 200, selva baja: 415, vegetación seca: 157, vegetación con cultivo: 164, vegetación sin cultivo: 229, vegetación de temporada anual: 55, sin vegetación aparente: 19, cuerpo de agua: 100 y zona urbana: 27)

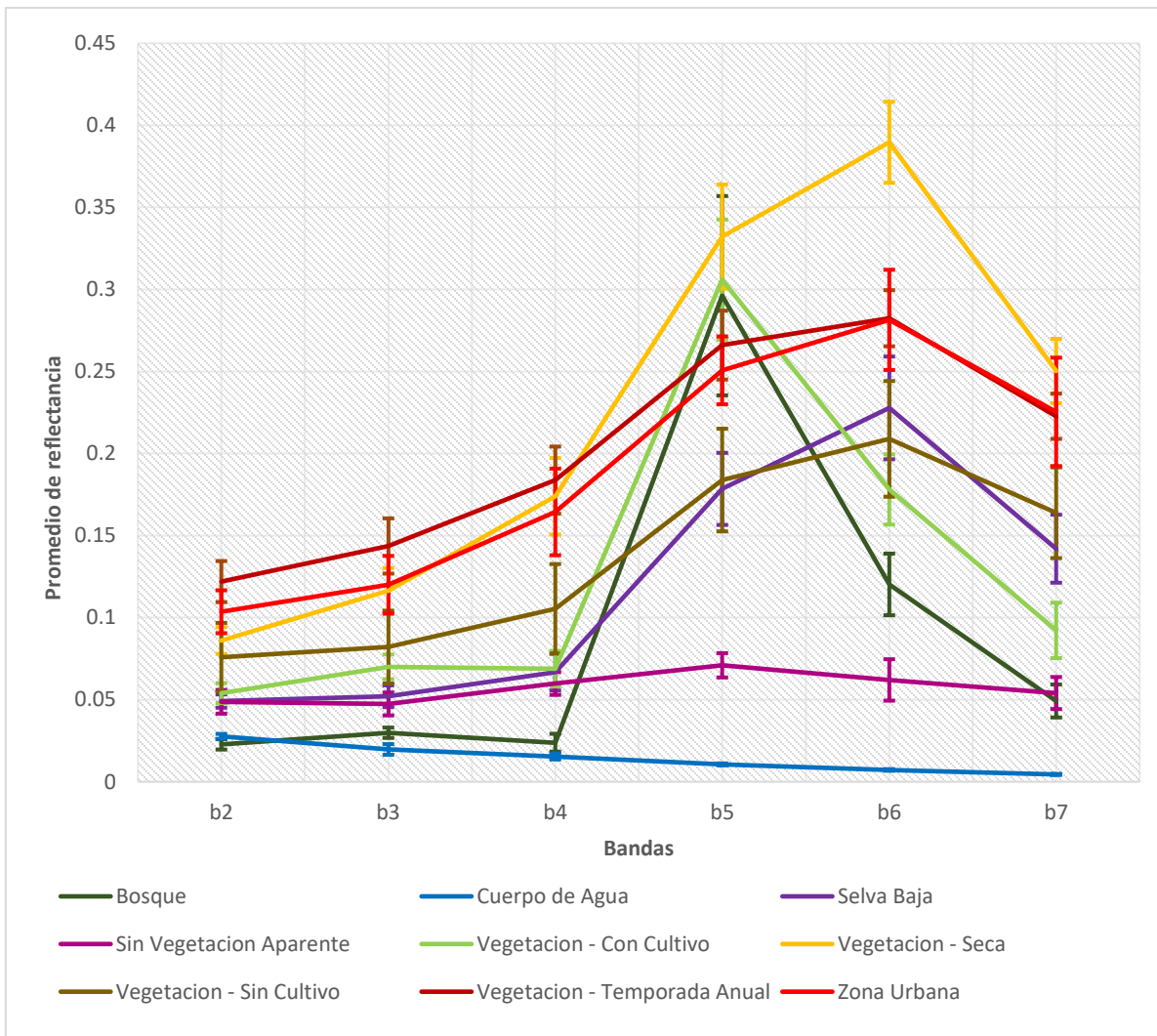


Figura 25. Información espectral de datos de entrenamiento 2020. (Basado en 1366 píxeles en total de las clases de entrenamiento, de los cuales corresponden a bosque: 200, selva baja: 415, vegetación seca: 157, vegetación con cultivo: 164, vegetación sin cultivo: 229, vegetación de temporada anual: 55, sin vegetación aparente: 19, cuerpo de agua: 100 y zona urbana: 27)

Clasificación en R usando algoritmos de MLC y Random Forest

Los resultados de la clasificación por el algoritmo de máxima verosimilitud se pueden observar en la Figura 26, así mismo los resultados del algoritmo de Random Forest se muestran en la Figura 27. A simple vista los resultados de ambos métodos de clasificación y para los dos años resultan muy similares, sin embargo, existen varios píxeles que difieren en ambos métodos de clasificación, por ejemplo, el algoritmo de random forest para 2010, se detectan píxeles de cuerpo de agua en gran parte de la zona boscosa, lo cual no hace el algoritmo de

máxima verosimilitud, además se nota que el algoritmo de máxima verosimilitud identifica menor cantidad de píxeles de área urbana. Cabe mencionar que para 2010, el porcentaje de píxeles que difieren en ambos métodos de clasificación es el 18.71%, en tanto para 2020 fue mayor con un 25.98%. Debido a diferencias en los resultados es importante medir cuál de los dos algoritmos obtuvo una mayor confiabilidad y con base en eso seleccionar la mejor clasificación para la aplicación de la máscara en los dos años, 2010 y 2020. Es importante resaltar que para aplicar la máscara únicamente se requiere seleccionar aquellos píxeles de las imágenes que sean de la clase de bosque o selva baja, ya que estos representan realmente las zonas en las que se basa nuestro estudio, por ello los píxeles que no sean de alguna de estas dos clases, serán todos tomados como “otro”. Al realizar esta reclasificación las imágenes para el algoritmo de Máxima Verosimilitud y Random Forest quedan como se muestra en la Figura 28 y Figura 29, respectivamente.

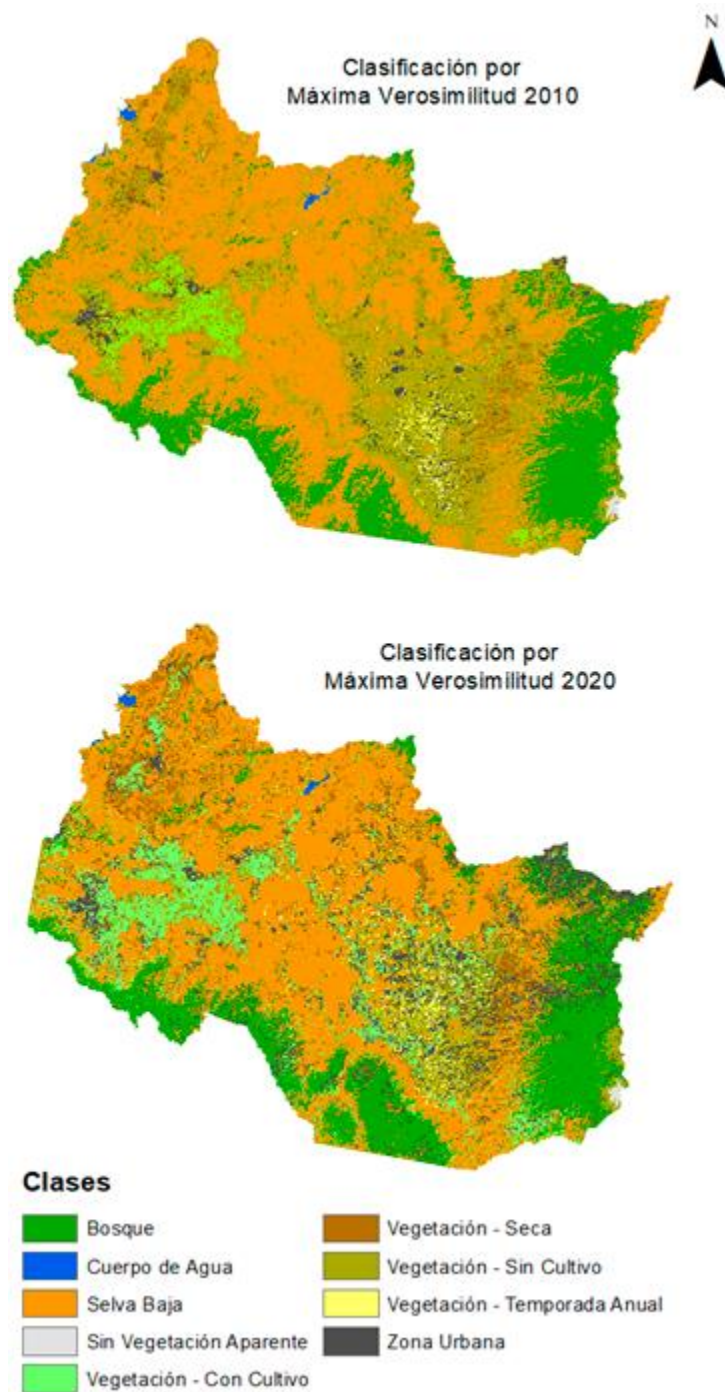


Figura 26. Resultado de la clasificación usando el algoritmo de Máxima Verosimilitud

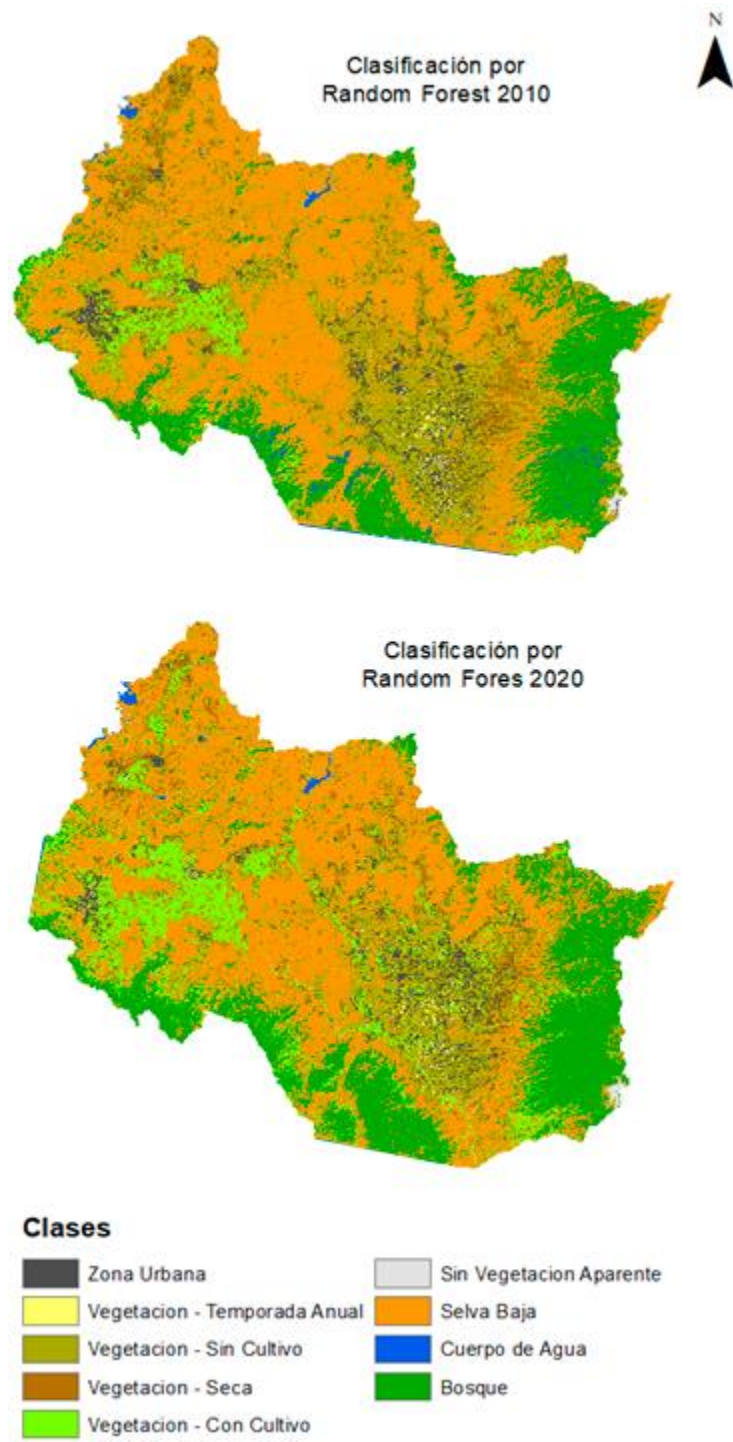


Figura 27. Resultado de la clasificación usando el algoritmo de Random Forest

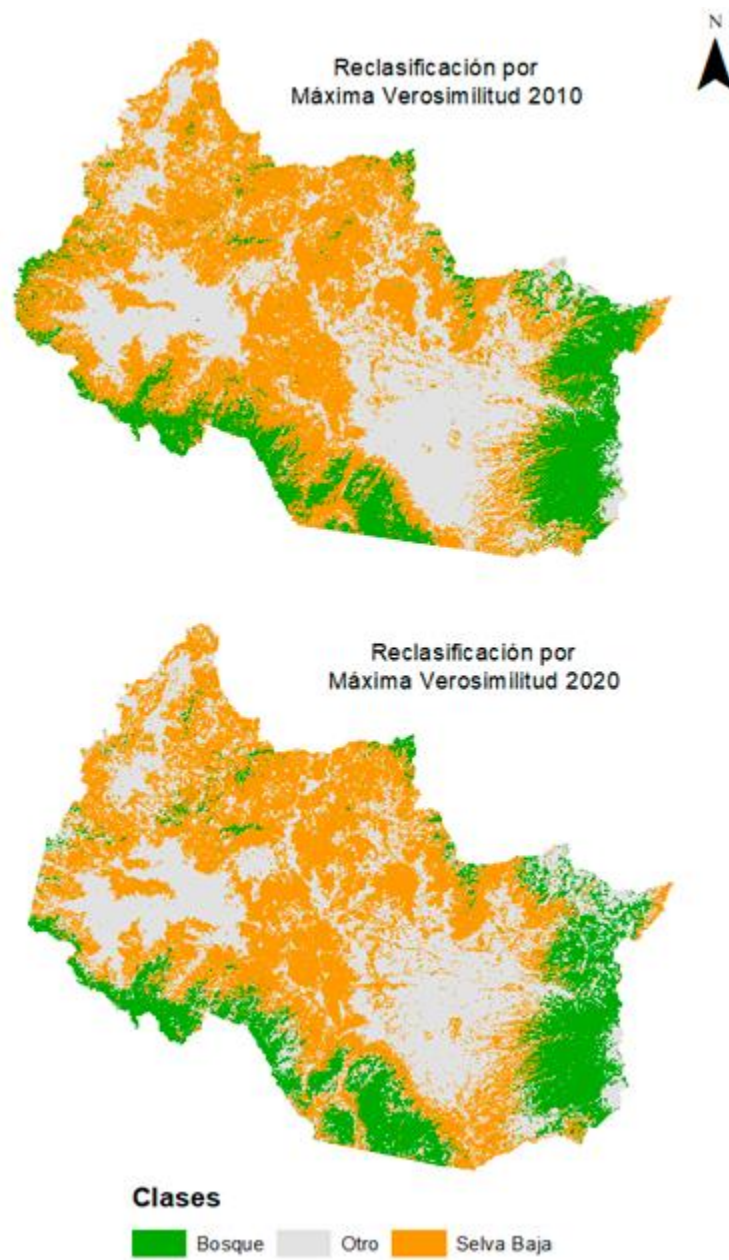


Figura 28. Reclasificación de la imagen del algoritmo de Máxima Verosimilitud

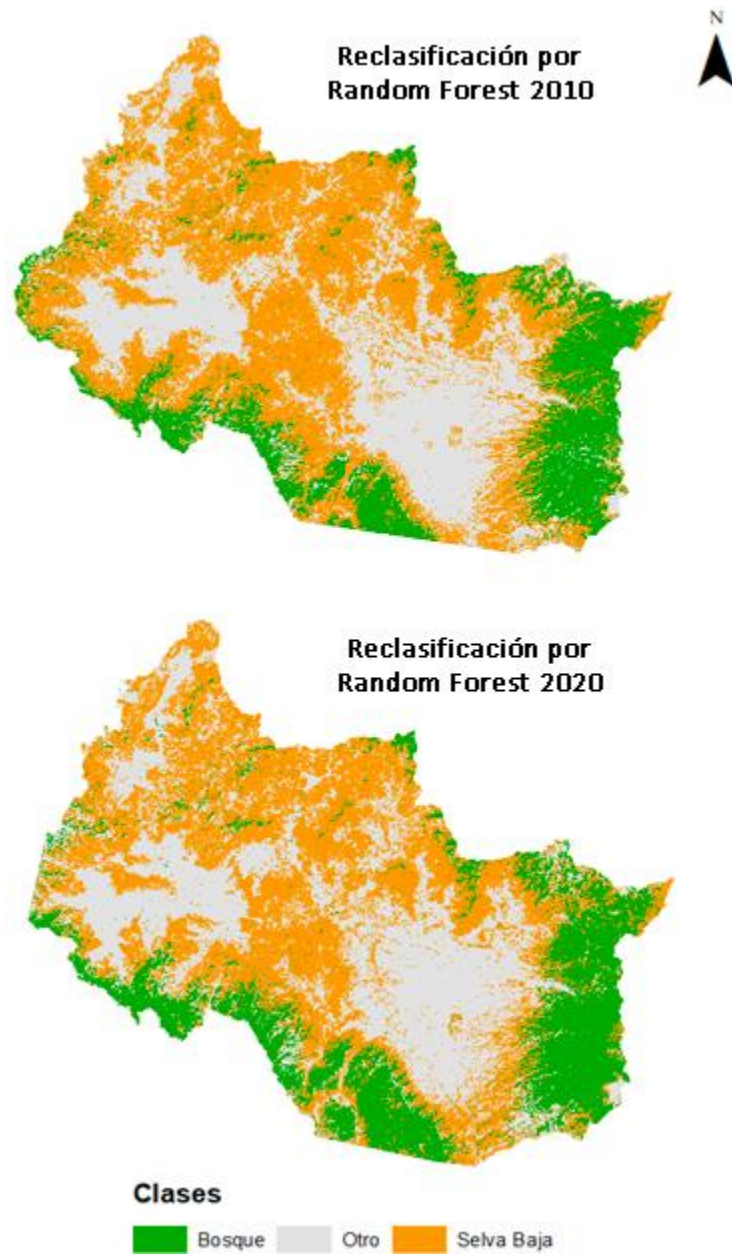


Figura 29. Reclasificación de la imagen del algoritmo de Random Forest

Aunque falta evaluar la clasificación de los métodos para seleccionar el que tenga mejor confiabilidad, en ambos se puede observar que para 2020 hay áreas en las cuales se ve a simple vista que la clase de bosque ha disminuido bastante, así como también otras áreas donde la selva baja lo ha hecho. En total, para la clasificación con el algoritmo de máxima verosimilitud se obtuvo que de 2010 a 2020 se ha perdido 4.21% de bosque y selva baja, mientras que el

resultado por el algoritmo de random forest indica una pérdida de 3.10%. Con esto se puede confirmar que desde 2010 hasta 2020 la cubierta forestal sí ha tenido un impacto negativo en la zona de estudio, o bien que alguna de las clasificaciones tenga error omisión o comisión en las categorías forestales, es por ello la importancia de validar estos resultados por lo que se procedió a medir la confiabilidad global, de usuario y de productor.

3.6.2 Evaluación de confiabilidad de la imagen clasificada

Una vez obtenida la reclasificación con las clases de interés (bosque, selva baja y otros), es el momento en el cual se tiene que evaluar la confiabilidad para identificar cuál de las dos clasificaciones resulta ser más conveniente aplicar para la selección de la máscara. Para esto fue necesario obtener una muestra con puntos de verificación en cada una de las clases obtenidas, lo cual se determinó mediante la aplicación del muestreo aleatorio estratificado sugerido por Olofsson et al. (2014).

$$n \approx \left(\frac{\sum W_i S_i}{s(\hat{\theta})} \right)^2 \approx \left(\frac{0.1542 \times 0.2179 + 0.4472 \times 0.3570 + 0.3985 \times 0.2179}{0.01} \right)^2 \approx 784.93$$

En nuestro caso, lo primero que se hizo para calcular n , fue determinar el área de cada una de las clases, resultando que la clase bosque ocupaba 640 Km², selva baja 1856 Km² y la clase de otro 1654 Km². Por lo que los valores proporcionales quedan como 0.154216, 0.447289 y 0.398554, respectivamente. Estos datos corresponden a la clasificación por el algoritmo de Máxima Verosimilitud, sin embargo, los datos por el algoritmo de Random Forest varían por muy poco, por lo que el cálculo de n se hizo con estos datos. Después se definió la confiabilidad de usuario para las clases de bosque, otro y selva baja, las cuales se establecieron como 0.95, 0.95 y 0.85 respectivamente. Posteriormente se calculó S_i quedando para la clase de bosque y la clase otro como $S_i = \sqrt{0.95(1 - 0.95)} = 0.2179449$ mientras que para la clase de selva baja quedo como $S_i = \sqrt{0.85(1 - 0.85)} = 0.3570714$. Por último, el error estándar se dejó como se sugiere en 0.01, dando como resultado $n = 785$.

Después de haber calculado n , se procedió a determinar la cantidad de puntos de verificación en cada una de las clases, y de forma que se tuvieran suficientes puntos en cada una de

ellas, se tomó la decisión de hacerlo de forma proporcional al área de cada una, dando como resultado 121 puntos para la clase de bosque, 351 para selva baja y 313 para la clase otro, una vez definido esto, se realizó el muestreo aleatorio estratificado donde los puntos resultantes fueron analizados mediante interpretación visual con ayuda de Google Earth y QGIS, de esta forma se logró identificar la clase real a la que pertenecían. De manera similar se aplicó el mismo procedimiento para los datos de 2020, obteniendo un total de 770 puntos para la validación. En la Figura 30 se puede observar la distribución de los puntos de verificación obtenidos, así como la clase a la que corresponden para ambos años.

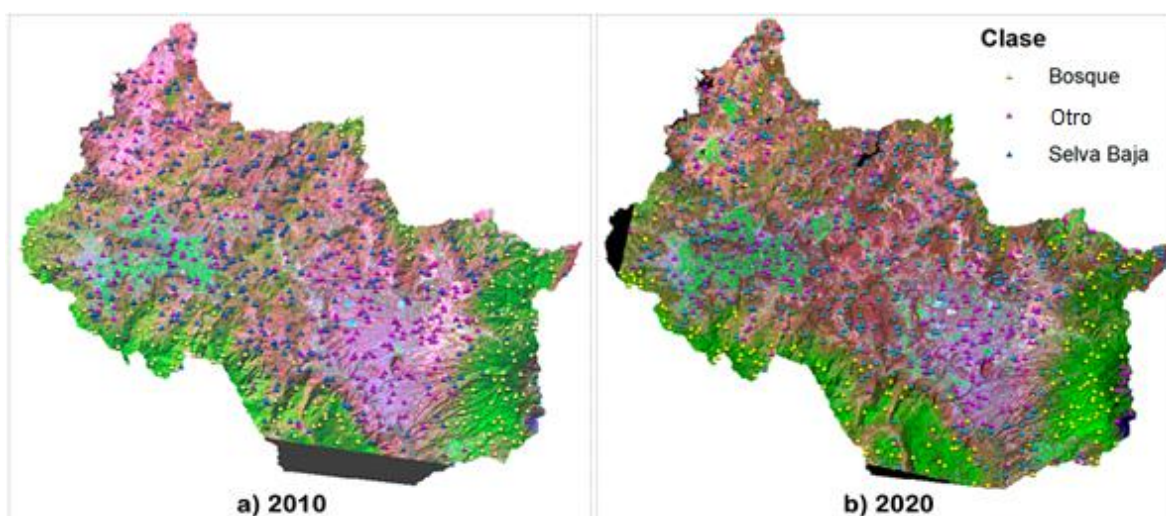


Figura 30. Distribución de los puntos de verificación 2010 y 2020

Con los puntos de verificación se pudo calcular la confiabilidad para las dos clasificaciones realizadas, tanto para 2010 como para 2020. En ambos años los dos métodos tuvieron resultados muy similares, siendo un poco mayor la confiabilidad global de la clasificación por el algoritmo de máxima verosimilitud con un 88% en 2010 y un 89% en 2020, en comparación con el 86% obtenido por la clasificación con el algoritmo de random forest en 2010 y un 86.88% en 2020. Sin embargo, además del calcular la confiabilidad global y para tener una mayor certeza de los resultados, es bueno realizar también el cálculo de la confiabilidad de usuario y la de productor, donde para esto es necesario primero realizar la matriz de confusión, en este caso para los dos métodos de clasificación utilizados. En la Tabla 9 y en la Tabla 10 se tienen las matrices de confusión para 2010 y en la Tabla 11 y Tabla 12 se muestran las matrices para 2020. En dichas matrices se pueden observar en la diagonal los valores bien

clasificados, los demás valores corresponden a los puntos de verificación que no fueron clasificados correctamente. Por ejemplo, en la matriz de confusión de 2010 para el algoritmo de máxima verosimilitud (Tabla 9) se puede notar que, en la clase de bosque, 27 puntos fueron confundidos con la clase de selva baja y 15 con la clase de otro. Algo similar ocurre con la matriz de confusión de la clasificación con Random Forest para 2010 (Tabla 10) ya que tiende también a confundir algunos de los valores de bosque con selva baja, esto podría parecer un tanto preocupante, sin embargo, debido a que lo que se quería hacer con la clasificación era obtener una máscara de bosque y selva baja, no representaría un mayor problema para nuestro objetivo.

		Referencia			SUMA REAL
		Bosque	Otro	Selva Baja	
Predicción	Bosque	119	2	0	121
	Otro	15	259	39	313
	Selva Baja	27	10	314	351
SUMA MAPA		161	271	353	785

Tabla 9. Matriz de confusión de la clasificación con el algoritmo de Máxima Verosimilitud 2010

		Referencia			SUMA REAL
		Bosque	Otro	Selva Baja	
Predicción	Bosque	123	2	3	128
	Otro	16	244	41	301
	Selva Baja	22	25	309	356
SUMA MAPA		161	271	353	785

Tabla 10. Matriz de confusión de la clasificación con el algoritmo de Random Forest 2010

		Referencia			SUMA REAL
		Bosque	Otro	Selva Baja	
Predicción	Bosque	123	3	1	127
	Otro	17	265	31	313
	Selva Baja	12	17	301	330
SUMA MAPA		152	285	333	770

Tabla 11. Matriz de confusión de la clasificación con el algoritmo de Máxima Verosimilitud 2020

		Referencia			SUMA REAL
		Bosque	Otro	Selva Baja	
Predicción	Bosque	137	9	5	151
	Otro	18	240	36	294
	Selva Baja	12	21	292	325
SUMA MAPA		167	270	333	770

Tabla 12. Matriz de confusión de la clasificación con el algoritmo de Random Forest 2020

Con los valores obtenidos en la matriz de confusión se procedió a calcular la confiabilidad de usuario y de productor, mismas que servirán para tener un mayor sustento de la precisión de las clasificaciones realizadas. Donde cómo se puede ver en la Tabla 13 para 2010 en el algoritmo de máxima verosimilitud, la confiabilidad de productor de la clase bosque, es un tanto baja con respecto a las demás y eso es debido a los varios puntos de verificación que realmente pertenecían a esta clase pero que fueron confundidos con otra, mayormente con la clase de selva baja. No obstante, se puede ver también que la confiabilidad de usuario de esta misma clase, para ambas clasificaciones de 2010 fue bastante buena (81.98% y 96.09%), esto hace referencia a que los píxeles clasificados como bosque, en su mayoría realmente si eran bosque. Para las otras dos clases, los valores de la confiabilidad de usuario y de productor resultan ser bastante aceptables, ya que toman valores mayores al 80% de confiabilidad e incluso algunos mayores del 90%. Por ejemplo, la confiabilidad de productor de la clase otro en 2010 resulta ser de 95.57%, un valor muy alto, por lo que se tiene la seguridad de que la mayoría de los píxeles de esa clase, fueron correctamente clasificados. Si comparamos los resultados de ambas clasificaciones de 2010, podemos observar que los valores obtenidos por el algoritmo de máxima verosimilitud tienden a ser un poco más altos, lo cual hace sentido, ya que la confiabilidad global de la clasificación por máxima verosimilitud fue la mayor con un 88%. Para la confiabilidad de usuario y de productor en la clasificación de 2020 se puede observar que toman valores muy cercanos a los obtenidos en la clasificación de 2010, pero en este caso todos los valores fueron mayores a 80%, lo cual nos da una mayor certeza de los resultados obtenidos.

CLASE	MÁXIMA VEROSIMILITUD				RANDOM FOREST			
	2010		2020		2010		2020	
	USUARIO	PRODUCTOR	USUARIO	PRODUCTOR	USUARIO	PRODUCTOR	USUARIO	PRODUCTOR
BOSQUE	81.98%	73.91%	96.85%	80.92%	96.09%	76.39%	90.72%	82.03%
OTRO	82.74%	95.57%	84.66%	92.98%	81.06%	90.03%	81.63%	88.88%
SELVA BAJA	89.45%	88.95%	91.21%	90.39%	86.79%	87.53%	89.84%	87.68%

Tabla 13. Confiabilidad de usuario y de productor

Con base en los resultados obtenidos se puede ver que, para ambos años, 2010 y 2020 el algoritmo de máxima verosimilitud tiende a dar resultados un tanto mejores en cuanto a la confiabilidad, por ello se decidió seleccionar las clasificaciones obtenidas por este algoritmo para la aplicación de la máscara. Para esto se utilizaron las reclasificaciones obtenidas en pasos previos (mostrados en la Figura 28) de las cuales se eliminó la clase de otro, posteriormente se unieron las clases de bosque y selva baja quedando ambas como una sola clase, el resultado se puede observar en la Figura 31, donde se puede apreciar el cambio que ha tenido la cubierta forestal de 2010 a 2020.

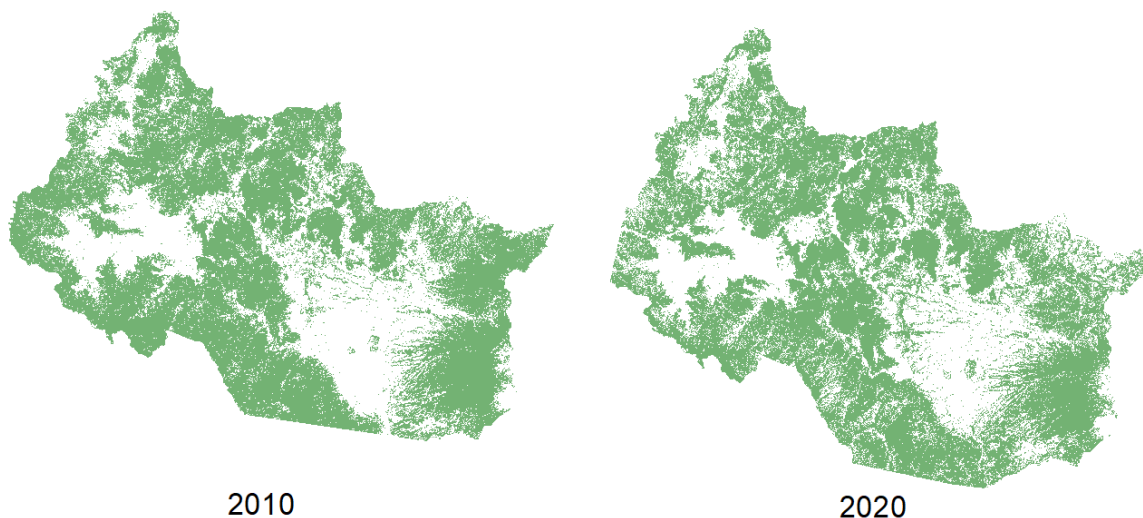


Figura 31. Bosque 2010 y 2020

Ahora bien, para la aplicación de la máscara se procedió para cada índice (NDVI, NDMI y NDFI) a seleccionar únicamente los píxeles que tuvieran una tendencia negativa y que para 2010 fueran bosque, esto para analizar únicamente las áreas que presentaron una disminución de cubierta forestal, sin embargo, las áreas podrían haber sido afectadas tanto por procesos de degradación como de deforestación, y en este estudio únicamente nos interesa las áreas que presentaran degradación, es por ello que se tuvieron que eliminar también las áreas que para 2020 ya no fueran de la clase bosque, de esta manera tenemos la seguridad que todos los píxeles resultantes son píxeles que en 2010 eran bosque, y que para 2020 lo siguen siendo, pero presentan una disminución de biomasa, es decir fueron sujetos de degradación forestal. En la Figura 32 se muestran los resultados obtenidos, donde se puede observar que el NDFI tuvo una mayor tendencia a obtener valores negativos, por su parte el NDVI y el NDMI se ven muy similares, sin embargo, se puede apreciar que el NDMI detecta mayor cantidad de píxeles con pendiente negativa.

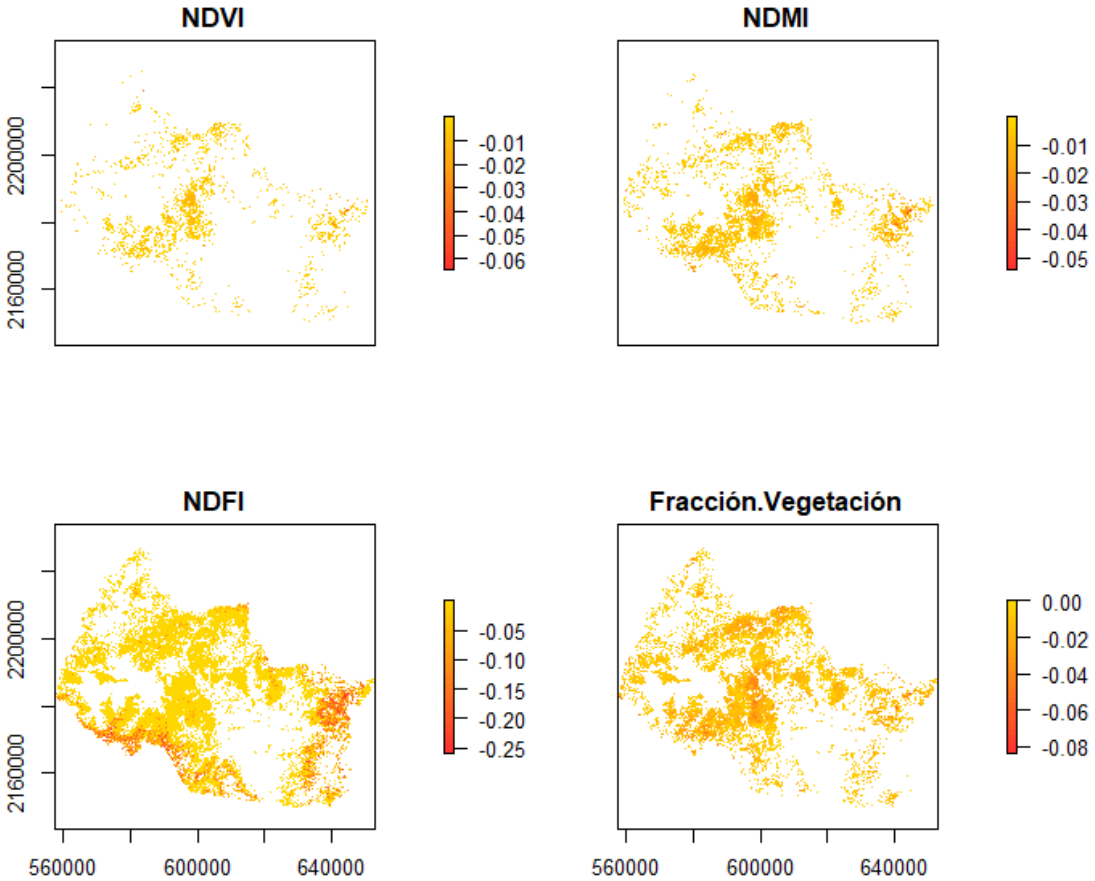
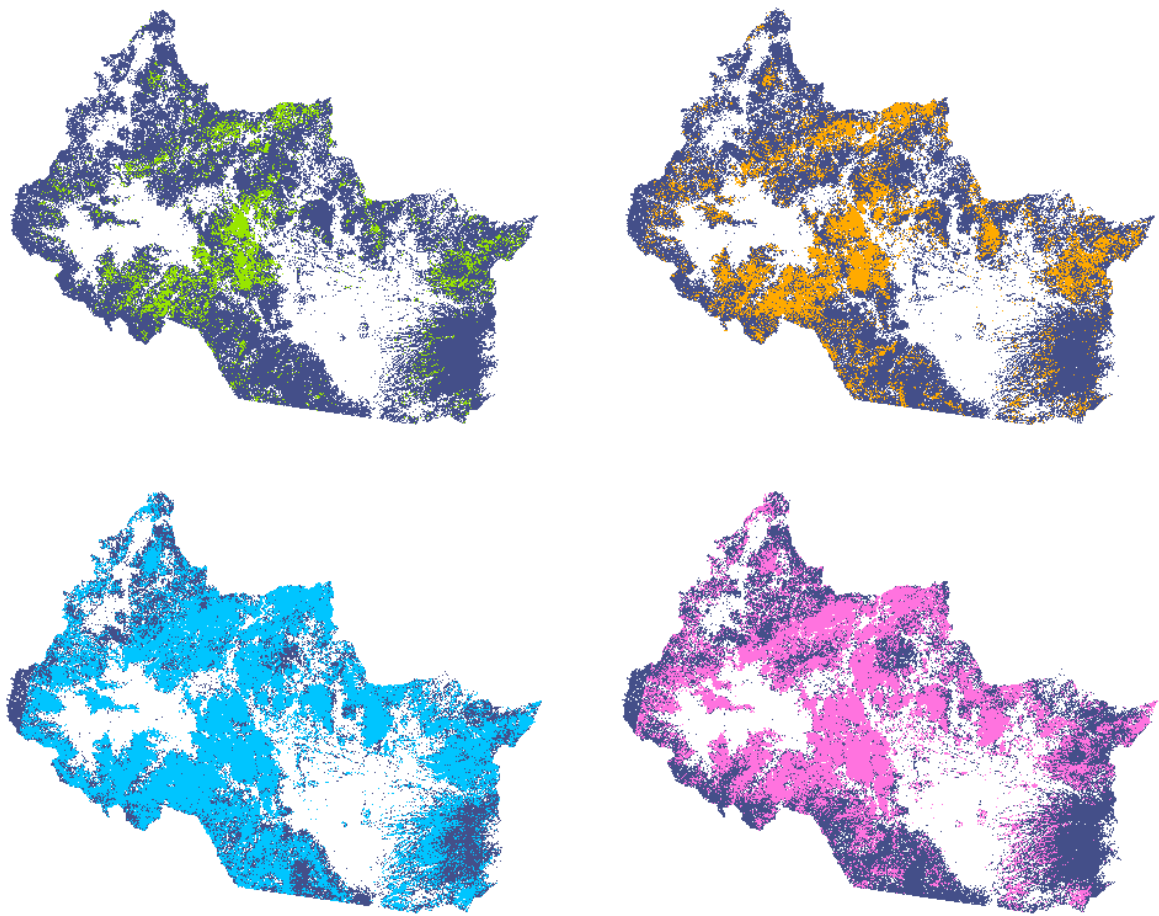


Figura 32. Máscaras por índice

Como se mencionó anteriormente, el NDFI parece identificar mayor cantidad de píxeles con una tendencia negativa, es decir píxeles que fueron sujetos a procesos de degradación forestal. Esto se pudo comprobar calculando el área de degradación identificada por cada índice donde para el NDFI fue de 1826.9 km², mientras que para el NDVI fue de tan solo 356.96 km², una diferencia bastante considerable con respecto al NDFI y también respecto al NDMI donde para este el área total fue 691.19 km² casi el doble de lo obtenido por el NDVI, por último la imagen fracción de vegetación también identificó una gran cantidad de píxeles con tendencia negativa, el área total de este fue de 1293.50 km².

Es importante señalar que el área total identificada como bosque para 2010 era de 2946.62 km², por lo que el bosque identificado por los diferentes índices como degradado resulta bastante significativo, por ejemplo, si tomamos el área identificada por el NDVI como degradación y tomamos el total de área de bosque en 2010, tenemos que de ese total y según el NDVI el 12.21% corresponde a degradación forestal, haciendo hincapié en que el porcentaje restante, es decir el 87.79% no corresponde únicamente a áreas deforestadas sino que también son áreas que presentaron un aumento en la densidad de bosque o bien que se conservan estables durante el periodo de este estudio. Estas áreas fueron omitidas de nuestro estudio ya que para este el objetivo era cuantificar las áreas de bosque degradado. Ahora bien, para el NDMI, el porcentaje de áreas degradadas con respecto al total de bosque en 2010 fue de 23.45%, casi el doble que el NDVI, como se mencionaba anteriormente, para el NDFI fue de 61.99%, casi dos terceras partes de la cubierta forestal en 2010, esto resulta un valor demasiado grande, sin embargo, se tendría que evaluar que tan fiables son estos resultados para tener una mayor seguridad de ellos, por último para la imagen fracción de vegetación resultó en un porcentaje de 43.89%, aunque menor que el NDFI, sigue siendo un valor muy considerable. En la Figura 33 se muestra de forma visual el área que abarcan las diferentes máscaras con respecto al bosque en 2010, donde se aprecia de mejor forma lo explicado previamente. Por ejemplo, vemos que la máscara de NDFI abarca la mayor parte del bosque, mientras que la máscara de NDVI abarca una parte mucho más pequeña.



Leyenda

-  Máscara NDVI
-  Máscara NDMI
-  Máscara NDFI
-  Fracción de Vegetación
-  Bosque 2010

Figura 33. Comparación entre bosque 2010 y máscaras

CAPÍTULO 4. VERIFICACIÓN DE RESULTADOS

4.1 DISEÑO DE MUESTRAS

Para la evaluación de los resultados del análisis de serie de tiempo, se procedió a identificar dentro de las máscaras de los índices obtenidos anteriormente aquellas áreas cuya tendencia negativa fuera significativa, es decir que su valor de significancia fuera menor que 0.05, las áreas que cumplieran con esa condición fueron clasificadas como áreas donde el análisis si encontró una disminución en la cubierta forestal, y por otro lado, las áreas cuya significancia no fuera relevante se tomaron como áreas en donde no hubo cambio significativo es decir áreas que permanecieron constantes durante el periodo de estudio. Para una correcta evaluación se crearon puntos aleatorios dentro de cada índice, donde la cantidad de puntos se determinó de la misma forma en que se hizo para la validación de la máscara, es decir siguiendo el proceso del muestreo aleatorio estratificado sugerido por Olofsson et al. (2014). La única diferencia con respecto a la evaluación de la máscara es que aquí, únicamente se tenían dos clases, la clase “Degradación” y la clase “Sin cambio”, dando como resultado la creación de 150 puntos por cada uno de los índices. En la siguiente ecuación se muestra el cálculo de los puntos para el NDMI, donde el área degradada corresponde a 13 km^2 y el área sin cambio fue de 678 km^2 además se estableció la confiabilidad de usuario como 0.950 y 0.985 respectivamente.

$$n \approx \left(\frac{\sum W_i S_i}{S(\hat{\theta})} \right)^2 \approx \left(\frac{0.0188133 \times 0.2179449 + 0.98118667 \times 0.1215525}{0.01} \right)^2 \approx 150$$

4.2 MATRIZ DE CONFUSIÓN

Después de obtener los 150 puntos aleatorios estos fueron analizados mediante el software de Google Earth Pro para identificar los puntos en los que realmente si se tenía degradación forestal y en cuales no había cambio alguno. En las figuras 34, 35y 36 se muestran algunos ejemplos con puntos de verificación en los cuales el modelo identificó correctamente áreas

con degradación forestal para los diferentes índices (NDVI, NDMI y NDFI). Además de ver los cambios en diferentes años, en las imágenes también se adjunta información referente al nivel de significancia, coeficiente de determinación, así como la recta de regresión obtenida por el análisis de serie de tiempo, en la cual se puede observar que todas las pendientes son negativas, esto debido a que los puntos aleatorios fueron generados en las máscaras obtenidas de cada índice. El punto de verificación en cada imagen está representado por un punto de color amarillo y el cuadro blanco a su alrededor representa el área que abarca un píxel.

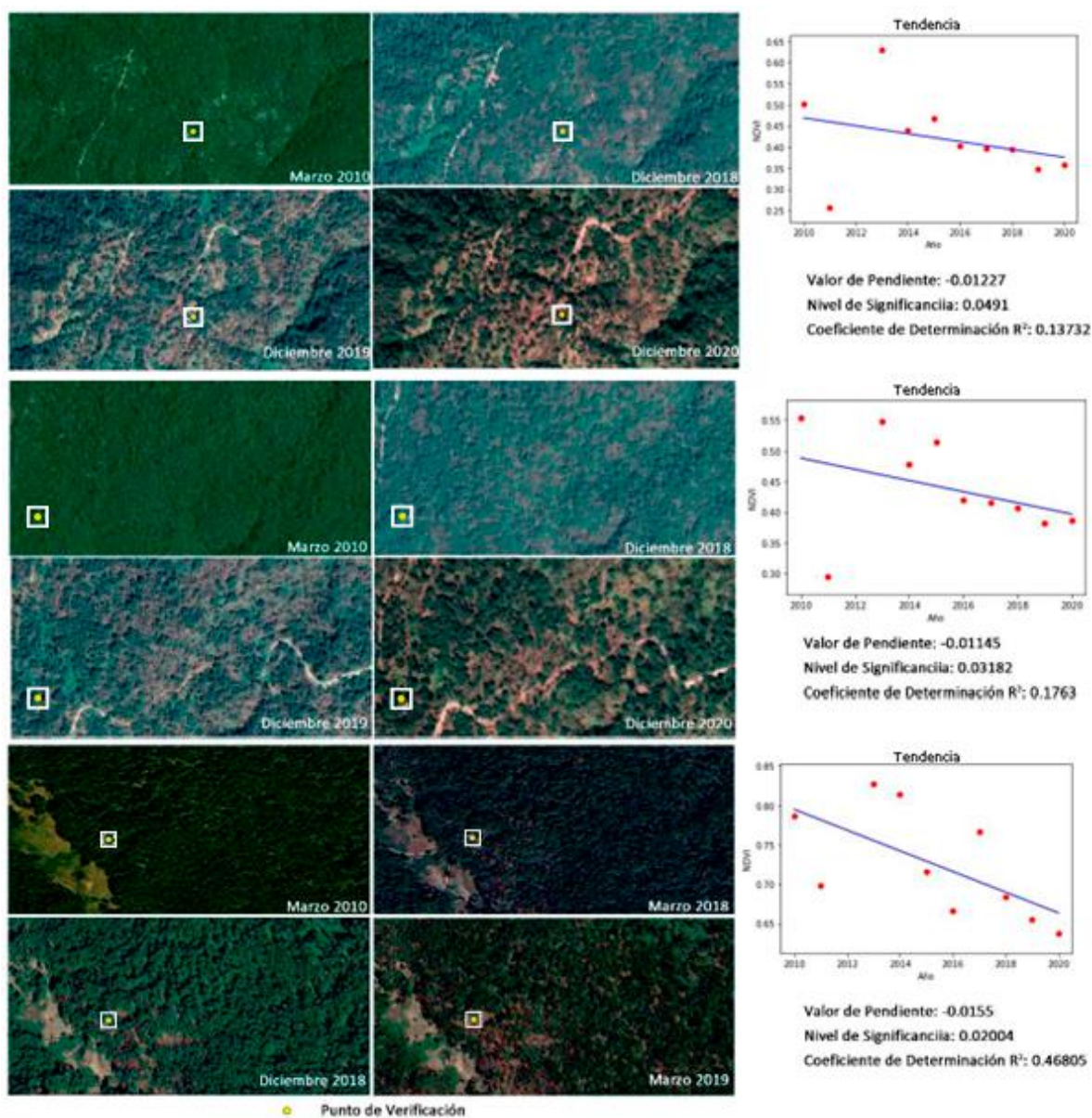


Figura 34. Ejemplos de Zonas con degradación Identificadas por el NDVI

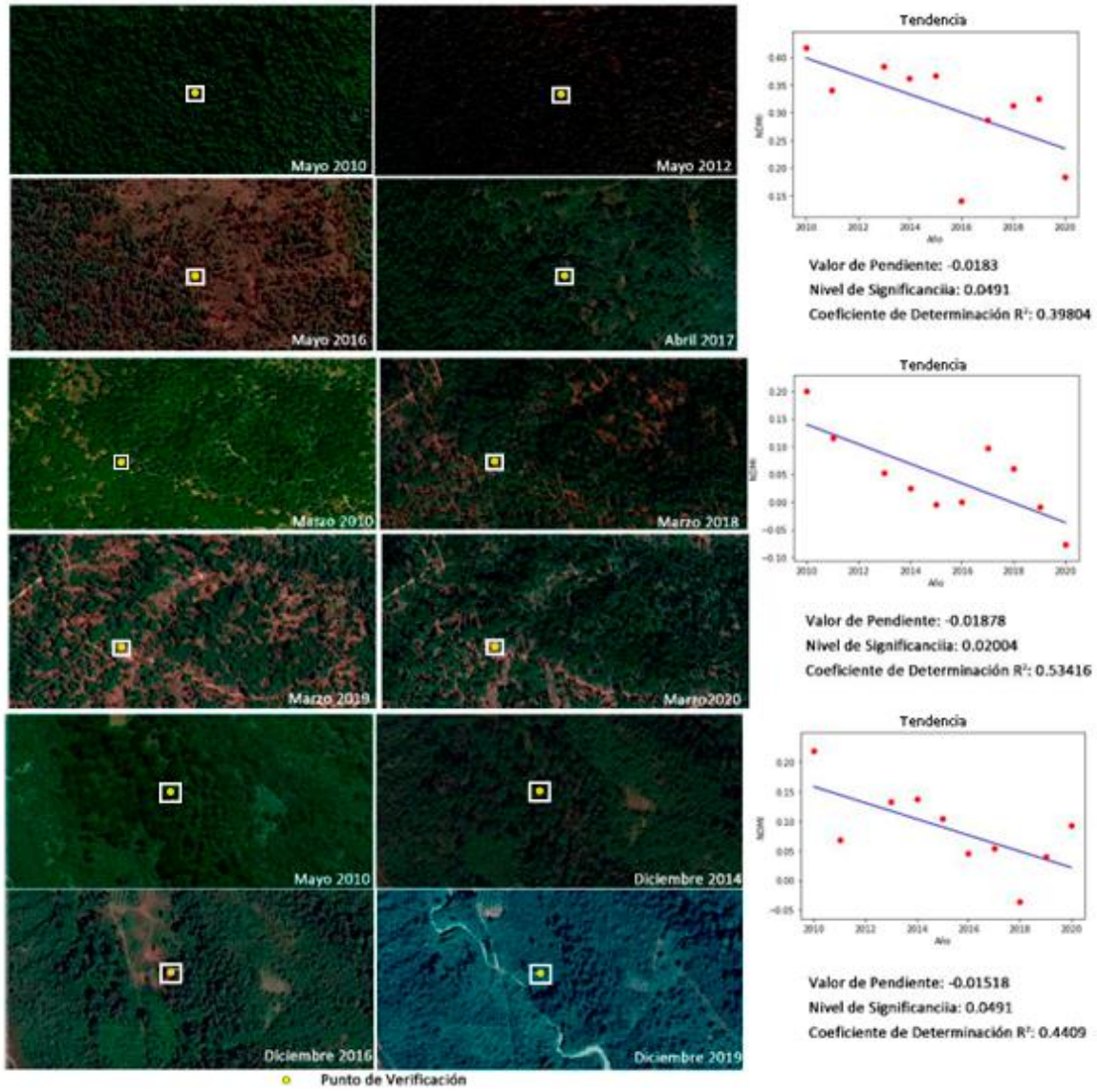


Figura 35. Ejemplos de Zonas con degradación Identificadas por el NDMI

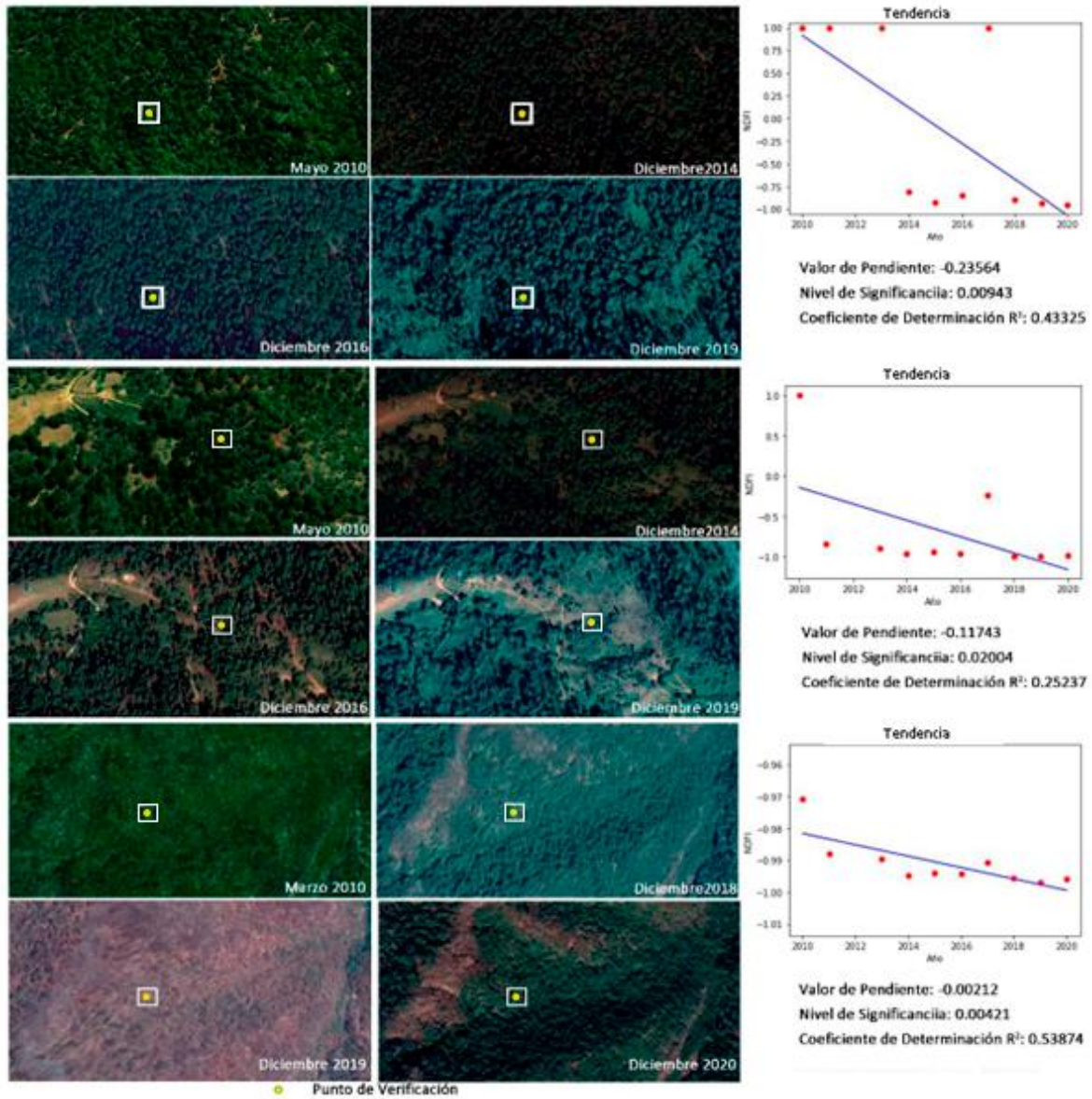


Figura 36. Ejemplos de Zonas con degradación Identificadas por el NDFI

Para tener una mejor apreciación de los resultados, también se integran imágenes de las áreas con puntos de verificación en las cuales no se presentó ningún cambio, mismas que se pueden apreciar en las figuras 37, 38 y 39 para los tres índices. En dichas imágenes se observa que el valor de la pendiente, aunque aún es negativo, tiende a ser muy cercano a cero, esto debido a que no se identifica un cambio o disminución de la cubierta forestal, es decir no se identifican áreas con degradación. También se puede apreciar que los valores obtenidos por en NDFI son un tanto dudosos, ya que tanto en la Figura 36 como en la Figura 39

los valores obtenidos por este índice la mayoría de las veces son o muy cercanos a 1 o muy cercanos a menos 1, lo cual no explica bien la variación de los cambios en el análisis de serie de tiempo. Por lo que desde este punto se podría inferir que el NDFI fue el índice que dio resultados menos certeros, sin embargo, para confirmar esto se procedió con la información recabada por el análisis y realizar el cálculo de la confiabilidad para cada uno de los índices.

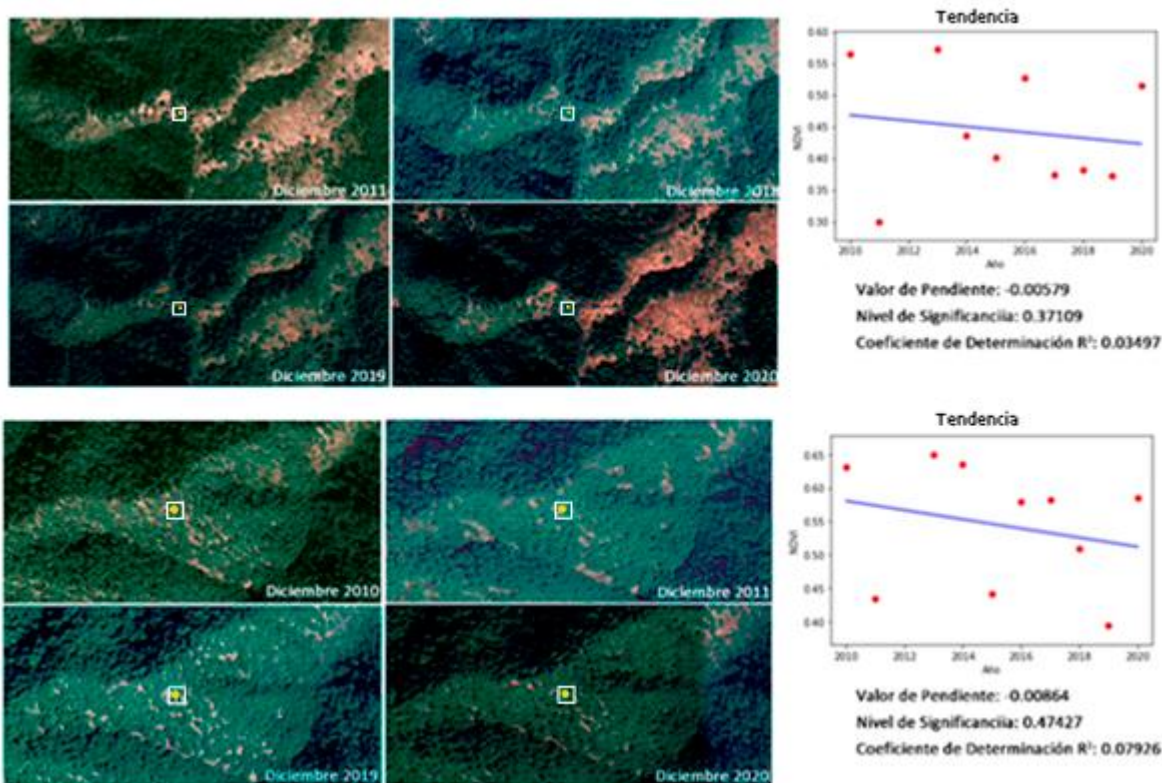


Figura 37. Ejemplos de Zonas sin cambio Identificadas por el NDVI

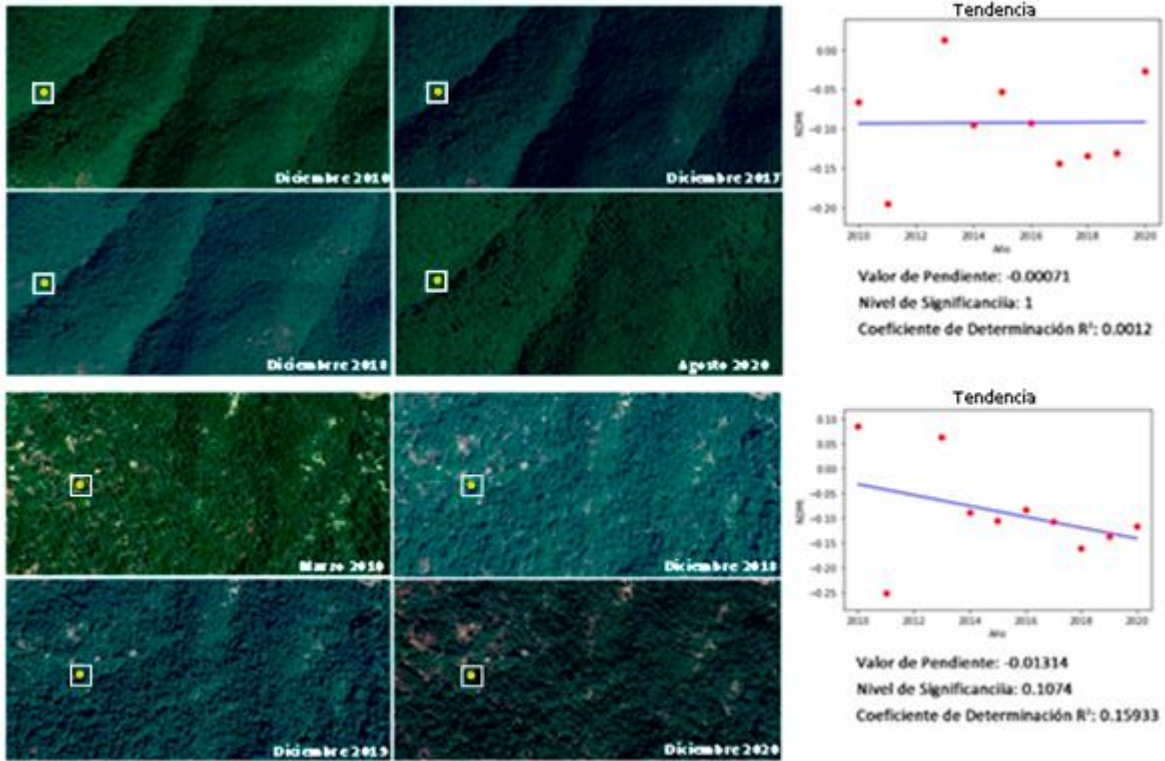


Figura 38. Ejemplos de Zonas sin cambio Identificadas por el NDMI

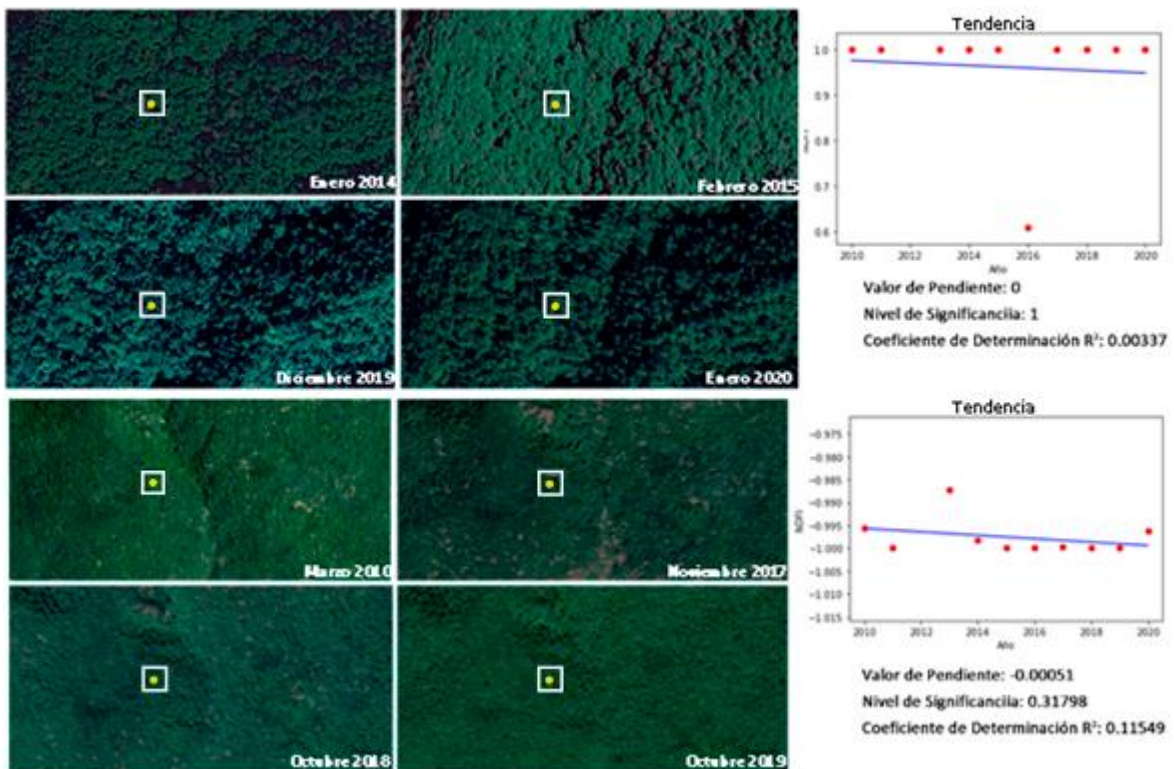


Figura 39. Ejemplos de Zonas sin cambio Identificadas por el NDFI

NDVI		Referencia		SUMA REAL
		Degradación	Sin Cambio	
Predicción	Degradación	64	11	75
	Sin Cambio	17	58	75
SUMA CLASIFICACIÓN		81	69	150

Tabla 14. Matriz de confusión NDVI

NDMI		Referencia		SUMA REAL
		Degradación	Sin Cambio	
Predicción	Degradación	47	28	75
	Sin Cambio	12	63	75
SUMA CLASIFICACIÓN		59	91	150

Tabla 15. Matriz de confusión NDMI

NDFI		Referencia		SUMA REAL
		Degradación	Sin Cambio	
Predicción	Degradación	39	36	75
	Sin Cambio	6	69	75
SUMA CLASIFICACIÓN		45	105	150

Tabla 16. Matriz de confusión NDFI

4.3 CÁLCULO DE CONFIABILIDAD

Para evaluar la confiabilidad del análisis se realizaron las matrices de confusión de los tres índices (Tabla 14, 15 y 16) donde se puede observar a simple vista que el índice con mayor confiabilidad es el NDVI, ya que tiende a cometer menos errores que el NDMI y el NDFI. Sin embargo, para corroborar esto, se procedió a calcular las diferentes medidas de confiabilidad, donde para el NDVI con un intervalo de confianza del 95% se obtuvo una confiabilidad global de $81.33\% \pm 6.23\%$, con lo que se sugiere que la mayoría de los píxeles fueron clasificados en su categoría correcta, mientras que para el NDMI la confiabilidad global fue de $73.33\% \pm 7.07\%$ un valor un tanto menor que lo obtenido por el NDVI, y por último, el NDFI, que aunque en un inicio se esperaba que tuviera mejores resultados, terminó siendo el índice que tuvo una menor confiabilidad en este análisis, quedando por debajo del NDMI con un $72\% \pm 7.18\%$. De igual modo que para la evaluación de la confiabilidad de la máscara, se dio también a la tarea de calcular la confiabilidad de usuario y de productor para tener mayor certeza de los resultados obtenidos en las clasificaciones, en la Tabla 17 se muestran los resultados obtenidos, donde como ya se asumió con anterioridad, los valores obtenidos por el

NDVI resultan ser más consistentes, ya que tanto la confiabilidad de usuario y de productor para las dos clases rondan el 80% de confiabilidad, siendo la más baja la confiabilidad de usuario para la clase de “sin cambio” con un 77.33%, lo cual indica que hay un ligero aumento en el error de comisión para esta clase, es decir que varios de los píxeles que eran de la clase “sin cambio” los clasificó como degradación. Sin embargo, el valor obtenido sigue siendo muy cercano al 80%, un valor relativamente alto. A su vez, para el NDMI y el NDFI se tienen algunos valores muy bajos, por ejemplo, la confiabilidad de usuario para la clase “Degradación” en el NDMI, resulta en un 64.66% un valor bastante bajo, lo cual indica que el error de comisión para esta clase fue muy alto, ya que como se puede ver en la matriz de confusión de este índice (Tabla 15) 28 de los 75 puntos de verificación que clasificó como “Degradación” realmente eran de la clase de “Sin cambios”. En el mismo índice vemos también que la confiabilidad de productor de la clase “Sin cambios” fue de igual forma muy baja, obteniendo un 69.23%, por lo que el error de omisión para esta clase también fue considerablemente alto, ya que de los 91 puntos de verificación que eran de esta clase, omitió 28, casi la tercera parte. Por último, para el NDFI, que aunque tuvo un confiabilidad de usuario para la clase “sin cambio” muy alta la de la clase “Degradación es extremadamente baja, con un 52% y es que casi la mitad de los datos que clasificó como “degradación” realmente no presentaban un cambio, por lo que también afectó la confiabilidad de productor de la clase “Sin cambio” ya que de los 105 puntos de verificación que eran de la clase “Sin cambio” únicamente clasificó bien 69, confundiendo los 36 restante con la clase de “Degradación”.

INDICE	CONFIABILIDAD DE USUARIO		CONFIABILIDAD DE PRODUCTOR	
	DEGRADACIÓN	SIN CAMBIO	DEGRADACIÓN	SIN CAMBIO
NDVI	85.33%	77.33%	79.01%	84.05%
NDMI	62.66%	84.00%	79.66%	69.23%
NDFI	52.00%	92.00%	86.66%	65.71%

Tabla 17. Confiabilidad de usuario y productor para el NDVI, NDMI y NDFI

CAPÍTULO 5. DISCUSIÓN

Los índices usados para este análisis de serie de tiempo, el NDVI, NDMI y NDFI, son indicadores comúnmente usados para analizar el nivel de vegetación, donde los más comunes son el NDVI y el NDMI, siendo el más reciente el NDFI, aunque como se menciona en capítulos anteriores, el NDFI es un índice que en los últimos años se ha utilizado cada vez con mayor frecuencia para analizar cambios en la cubierta forestal y que ha dado muy buenos resultados (Schultz, et al., 2016). Por esta razón, en un inicio se creyó conveniente para este estudio el uso del NDFI y se esperaba que tuviera mejores resultados que el NDVI y NDMI, sin embargo, al analizar los resultados, el índice que presentó mayor confiabilidad fue el NDVI, con un 81.33% de confiabilidad global, dejando atrás al NDMI por 8% y al NDFI por 9.33%. Por lo que el NDFI resultó ser en este estudio el índice con menor confiabilidad global. También como se puede observar en Tabla 17 la confiabilidad de usuario y de productor para el NDVI es para las dos clases analizadas (degradación y sin cambio) muy cercana al 80% por lo que se puede concluir que los resultados para este índice fueron bastante buenos. Esto no ocurrió para el NDFI ni NDMI ya que sus valores de confiabilidad tanto de usuario como de productor en algunos casos son menores al 70%. Siendo mucho más bajos los valores obtenidos por el NDFI, por ejemplo, la confiabilidad de usuario para la clase de “Degradación” resultó en un 52%, un valor extremadamente bajo, que nos indica que el error de comisión o la detección falso positivo de esta clase fue muy alto (42%) es decir, que muchos píxeles clasificados como “Degradación” realmente eran de la clase “Sin cambio”, lo cual se puede comprobar en la Tabla 16.

5.1 COMPARACIÓN CON OTROS PROYECTOS

Como este, otros estudios han utilizado técnicas de regresión lineal para identificar cambios y afectaciones en la cubierta forestal a través de análisis de serie de tiempo, tal es el caso de Zhu et al. (2016) cuyo objetivo era analizar las tendencias de verdor con base en imágenes satelitales Landsat en Guangzhou, China. Similar al presente trabajo, también se basaron en

un modelo de regresión lineal simple, pero usando únicamente dos índices de vegetación, el NDVI y el EVI. En su análisis encontraron que los valores del NDVI tenían un sesgo mayor que los valores de EVI. Argumentan que esto pudo deberse a que tuvieron un sesgo muy alto en las bandas visibles, y aunque también se presentó en las bandas de onda más larga fue mucho menor. Mencionan también que el sesgo pudo estar relacionado con los diferentes métodos de corrección atmosférica utilizados ya que esta tiene un mayor impacto en los valores de reflectancia superficial resultantes para las bandas visibles y que el sesgo en los valores de EVI es menor que los valores de NDVI ya que los sesgos de la banda azul y la banda roja se cancelan entre sí durante su cálculo. Debido a esto, utilizaron únicamente el EVI como indicador de verdor, contrario a nuestro estudio que aunque no se calculó el EVI, de los tres índices que usamos (el NDVI, NDMI y NDFI), el NDVI resultó ser el que mejor resultados obtuvo, sin embargo, no tan precisos como lo fue el EVI en su estudio ya que ellos alcanzaron una confiabilidad global del 90.31%, una confiabilidad de usuario 91.97% y por último la confiabilidad de productor de 87.05% y en este estudio el NDVI obtuvo una confiabilidad global del 81.33%.

Prajesh et al. (2019) en su estudio tenían la intención de monitorear el crecimiento y desarrollo de la vegetación en los distritos de Tamil Nadu, India, usando imágenes de MODIS y a través de un análisis de serie de tiempo sobre los índices NDVI y NDWI (índice de agua de diferencia normalizada). El análisis de serie de tiempo también se determinó utilizando un modelo de regresión lineal y para detectar cambios en la vegetación, al igual que en este trabajo, se extrajo del modelo el valor de la pendiente. Los resultados que obtuvieron con el NDVI mostraron un mayor porcentaje de áreas con tendencias positivas (86.52%) durante la temporada de lluvia en comparación con la temporada seca (59.72%). Por otro lado, el NDWI durante la temporada de lluvias mostró un porcentaje mayor de área con tendencias positivas (90.29%) mientras que, durante la temporada de seca, un mayor porcentaje del área cae bajo una tendencia negativa (59.75%). De esta forma se puede decir que el NDWI es más sensible a cuestiones como precipitación o temperatura, por lo que el NDVI en este caso, al igual que en el presente estudio, resultó ser un buen estimador para identificar cambios en la vegetación y aunque en este proyecto no se usó el NDWI, que de haberse usado muy

posiblemente hubiera detectado valores muy bajos debido a las fechas de obtención de las imágenes, el NDVI logro cumplir con el objetivo.

Otro estudio similar es el de Bullock et al. (2018) en donde ellos, además de implementar un análisis de serie de tiempo para analizar la degradación forestal en el estado de Rondônia, también tomaron en consideración que la resolución espacial de las imágenes puede ser una limitante para detectar perturbaciones en los bosques, sobre todo cuanto esta ocurre en áreas pequeñas. Por esta razón, al igual que en este estudio, tuvieron la intención de aplicar el algoritmo SMA para analizar la degradación forestal en áreas menores a las de la resolución espacial de las imágenes satelitales. De esta forma los autores calcularon el NDFI con el objetivo de identificar la diferencia entre áreas de bosque intactas, bosque perturbado y bosque despejado. Al aplicar el análisis de serie de tiempo de 1990 a 2013 usando el NDFI encontraron que aproximadamente el 6% del bosque había sido degradado y aproximadamente la cuarta parte deforestado. Los resultados obtenidos tuvieron una confiabilidad global de 90.8%, en cuanto para la degradación y deforestación se tuvo una confiabilidad de usuario de 88.0% y 93.3% respectivamente y una confiabilidad de productor de 68.1% y 85.3%. Con esto se puede decir que, como se pensaba en un inicio en este proyecto, el NDFI si es índice adecuado para medir las perturbaciones forestales, sin embargo, en este trabajo los resultados obtenidos no fueron los que se esperaban, esto puede ser debido a posibles errores en la selección de los miembros puros, por lo que los resultados se vieron afectados.

Es importante reconocer que, como estos, hay otros estudios que han utilizado un análisis de serie de tiempo por medio de la aplicación de regresiones lineales para detectar perturbaciones en la cubierta forestal, esto a través del análisis de diferentes índices de vegetación, humedad, agua, entre otros. La mayoría de los estudios revisados siguen un proceso similar, y aunque algunos usan otras herramientas y hacen uso de diferente tipo de imágenes, la mayoría coincide en la aplicación de técnicas de corrección y preprocesamiento, obtención de los índices de interés (NDVI, NDMI, MDFI, NMWI, etc.), clasificación para identificar las áreas de bosque y vegetación, aplicación de una máscara y la aplicación del análisis de serie de tiempo por medio de una regresión lineal. Todos estos pasos también fueron aplicados en el presente estudio ya que son los pasos principales para aplicar este tipo de análisis.

Además, los estudios revisados también coinciden en que la forma para detectar áreas con degradación es identificar los píxeles que presenten una tendencia negativa y significativa en el análisis de serie de tiempo.

Algo importante de mencionar es que en algunos estudios revisados como Tarazona & Miyasiro (2020) y Zhu et al. (2016), comentan que este tipo de análisis funciona bien para áreas que no están experimentando cambios sustanciales en la cobertura del suelo, pero para áreas caracterizadas por cambios importantes en la cobertura del suelo, este enfoque puede proporcionar resultados engañosos o incompletos ya que para los lugares que experimentan cambios en la cobertura del suelo, el cambio abrupto puede sesgar el análisis de las tendencias del verdor. Por lo tanto, para una cuantificación es importante hacer la distinción entre los cambios abruptos causados por el cambio de cobertura terrestre de los cambios graduales en lugares donde no se ha producido un cambio de cobertura terrestre.

5.2 RESULTADOS OBTENIDOS NDVI Y NDMI

El NDVI es un indicador muy útil para medir la presencia y el vigor de la vegetación, lo cual en nuestro estudio era muy importante para estimar la cantidad de biomasa en los bosques. Además, como ya se ha mencionado anteriormente el NDVI es uno de los índices más usados para medir el nivel de vegetación y tiene un gran potencial para identificar variaciones en cuanto a estructura, composición y diversidad de los bosques. Por otro lado, el NDMI es un índice que mide el nivel de humedad en la vegetación, lo cual es útil para identificar sequías y cambios sutiles en las condiciones de humedad de la vegetación. Además de esto, como se mencionó anteriormente, existen algunos estudios han demostrado que es un indicador muy bueno para detectar cambios o perturbaciones forestales (Wilson & Sader, 2002). A pesar de ello, para analizar la degradación forestal en nuestro estudio no resultó ser muy preciso, ya que como se muestra en la Tabla 17 aunque los resultados fueron muy similares a los del NDVI, debido a que ambos índices tienden a detectar degradación forestal en las mismas zonas, el error de comisión (falsos positivos) del NDMI fue mucho más alto que el del NDVI (14.67 vs 37.34), esto debido a que las áreas que el NDMI identificó como degradación eran

más extensas que las identificadas por el NDVI, y al analizar los resultados se vio que gran parte del área que el NDVI identificó como degradación, no presentaban realmente una disminución en el nivel de vegetación.

5.3 RESULTADOS OBTENIDOS POR EL NDFI

Aunque el NDVI y NDVI suelen ser buenos indicadores para medir la vegetación, tienen la “desventaja” de que toman como referencia para su cálculo el valor de reflectancia de cada píxel, y no toman en cuenta que un solo píxel de una imagen satelital, debido a la resolución espacial, puede estar conformado por más de una clase, por esta razón se tuvo la intención de que además de estos índices también se usó el NDFI para medir la degradación forestal, ya que este último, tiene la ventaja de que si considera la premisa de que un píxel puede estar compuesto por varias clases. Por esta razón se esperaba en un inicio que este índice daría mejores resultados en el presente análisis, sin embargo, como ya se ha comentado fue el índice que menor confiabilidad tuvo. Esto pudo ser debido a que para su cálculo se requiere un análisis muy minucioso de la zona de estudio, además es importante resaltar que, para la obtención de este índice, a diferencia del NDVI y NDVI, es necesario contar además de las diferentes bandas de la imagen satelital, con los miembros puros o píxeles puros, los cuales son el insumo para obtener las imágenes fracción, que para el cálculo del NDFI corresponden a las clases de vegetación verde, vegetación no fotosintética, suelo y sombra.

Para obtención de estas imágenes fracción, es necesario aplicar el algoritmo SMA, el cual se encarga de descomponer los píxeles en las diferentes clases que lo componen. Esto hace que el cálculo del NDFI sea un tanto más complicado, ya que se deben tener muy bien identificadas las clases a analizar y los píxeles puros, mismos que sirven como insumo para el algoritmo SMA. En este estudio, aunque estos se trataron de identificar de forma muy cuidadosa pudieron no haber sido los que mejor representaban a los píxeles puros (Vegetación verde, Vegetación no fotosintética, sombra y suelo). La elección de los píxeles puros se hizo por medio de interpretación visual sobre la misma imagen satelital, sin embargo, para este tipo de análisis suele ser más conveniente tener datos de campo que nos den mejor certidumbre

de la elección de estos píxeles. Ahora bien, además de que obtener los datos in situ requería un desplazamiento a la zona de estudio, es importante recordar que este, es un análisis de serie de tiempo, por lo cual no podríamos obtener información de años previos, únicamente se hubieran podido obtener los datos del año 2020 (año en el cual se inició este trabajo). Otra posible solución podría haber sido usar imágenes de mayor resolución que la de Landsat (recordemos que Landsat tiene una resolución espacial de 30 metros) y de esta forma la selección de los píxeles puros sería más precisa. Sin embargo, aunque actualmente existen imágenes gratuitas y con una mayor resolución como las de Sentinel-2, no fue posible hacer uso de ella debido a que su disponibilidad fue a partir de 2015 lo cual no concuerda nuestro periodo de análisis. Además de esto es importante considerar que debido a una mayor resolución espacial estas imágenes tienden a abarcar una menor área, y el poder necesario para su procesamiento se incrementa bastante. En resumen, se podría atribuir los bajos resultados del NDFI a los insumos que se usaron para su determinación.

5.4 APLICACIÓN DE MÁSCARA

Otro aspecto importante en el análisis de los resultados es considerar que para la identificación de la degradación forestal se procedió a aplicar una máscara con la cual se asegurara que todos los píxeles del estudio fueran de la clase de bosque, para esto se realizaron dos clasificaciones, una para 2010 y otra para 2020. Con los resultados obtenidos se prosiguió a realizar una reclasificación (Tanto para 2010 como para 2020) con las clases de interés, en este caso con las clases de bosque, selva baja y otros. Ahora bien, para asegurar que únicamente se tuvieran datos de degradación forestal, con la clasificación de 2010 y 2020 se seleccionaron únicamente los píxeles donde para ambos años corresponden a la clase bosque y selva baja (posteriormente se tomó con una sola clase, bosque), de lo obtenido se eliminaron todas aquellas áreas que en el análisis de serie de tiempo tuvieran una pendiente positiva, ya que estos representan una posible regeneración de la cubierta forestal, lo cual es fuera del tema de esta tesis. De forma tal que al final se quedaron únicamente los píxeles que tanto en 2010 y en 2020 eran de la clase de bosque (según las clasificaciones realizadas) y que además en el análisis de serie de tiempo mostraron tener una disminución en su nivel

de vegetación (Esto se identificó mediante los valores obtenidos por el NDVI, NDMI y NDFI). De esta forma se obtuvieron las máscaras para el análisis (una por cada índice).

Dicho esto, es importante señalar que las clasificaciones realizadas para la obtención de la máscara se realizaron con el algoritmo de máxima verosimilitud, el cual es un algoritmo de clasificación supervisada, lo que quiere decir que para la clasificación se tuvieron que usar datos de entrenamiento, en otras palabras, se tuvieron que seleccionar polígonos que representan la información espectral de las diferentes clases de las que se conforma la zona de estudio (Ver Tabla 8), estos polígonos sirvieron de insumo para el algoritmo de clasificación, los cuales también fueron obtenidos por medio de interpretación visual, con lo que también podrían estar sujetos a errores de percepción, sin embargo, en este caso para asegurarnos de que la selección de estos polígonos tuviera buenos resultados se elaboró un análisis para medir la información espectral de las clases y ver que tan bien se diferenciaba una de otra (Figura 24 y 25), con esto se tuvo una mayor seguridad de que el algoritmo de máxima verosimilitud lograría clasificar con mayor exactitud las diferentes categorías.

Los resultados obtenidos de ambas clasificaciones resultaron ser bastante aceptables ya que para 2010 la confiabilidad obtenida fue de 88% y para 2020 de 89%, sin embargo, no son perfectos, por lo que se tiene que considerar que la máscara obtenida puede tener un nivel de error. Por ejemplo, si analizamos la confiabilidad de usuario y de productor (Tabla 13 Y 15) podemos ver que para 2010 la confiabilidad de productor de la clase bosque es un tanto baja, con un 73.91%, lo cual nos indica que el error de omisión de esta clase es de 26.09%, es decir que varios de los pixeles de bosque fueron clasificados en otra categoría, algo similar pasa con los resultados de la clasificación de 2020. Sin embargo, si analizamos las matrices de confusión (Tabla 9 y 11) podemos ver que de los pixeles de bosque templado que fueron confundidos con otra clase, la mayoría fueron clasificados como selva baja, por lo que no resulta tan grave, ya que para la máscara los pixeles que se quieren seleccionar son tanto los de selva baja como los de bosque, No obstante sigue existiendo error y se debe tener en cuenta que existen pixeles de la clase de bosque que fueron omitidos y algunos otros (muy pocos en realidad) que no eran pixeles de bosque, pero que fueron clasificados como tal.

5.5 ANÁLISIS DE RESULTADOS OBTENIDOS POR EL NDVI

Debido a que el mejor resultado cuanto a la identificación de la degradación forestal fue obtenido por el NDVI se procedió a analizar únicamente los valores arrojados por este índice, donde para una mejor visualización de estos resultados, se muestra la Figura 40 , en la cual de color verde se muestra la zona de bosque presente en 2010, de color amarillo la máscara del índice, y por último de color rojo se muestran las áreas en las que el modelo identificó degradación forestal. El área total degradada identificada fue de 1.91 kilómetros cuadrados, que corresponde a 191 hectáreas, donde a simple vista se ilustran dos grupos principales, el primero de ellos al centro de la zona de estudio, el cual corresponde principalmente a áreas de selva baja, y el otro grupo está un poco más al Noreste donde corresponde principalmente a áreas de bosque pino encino, Además de estos, también se tienen algunos otros puntos distribuidos en diferentes áreas al norte y otros pocos al sur. En la Figura 41 se aprecian los principales grupos de degradación forestal, donde se integra una visualización de cómo era en 2010 y el cambio que tuvo para 2020, en ambos grupos se puede notar que si existe una disminución considerable de la cubierta forestal.

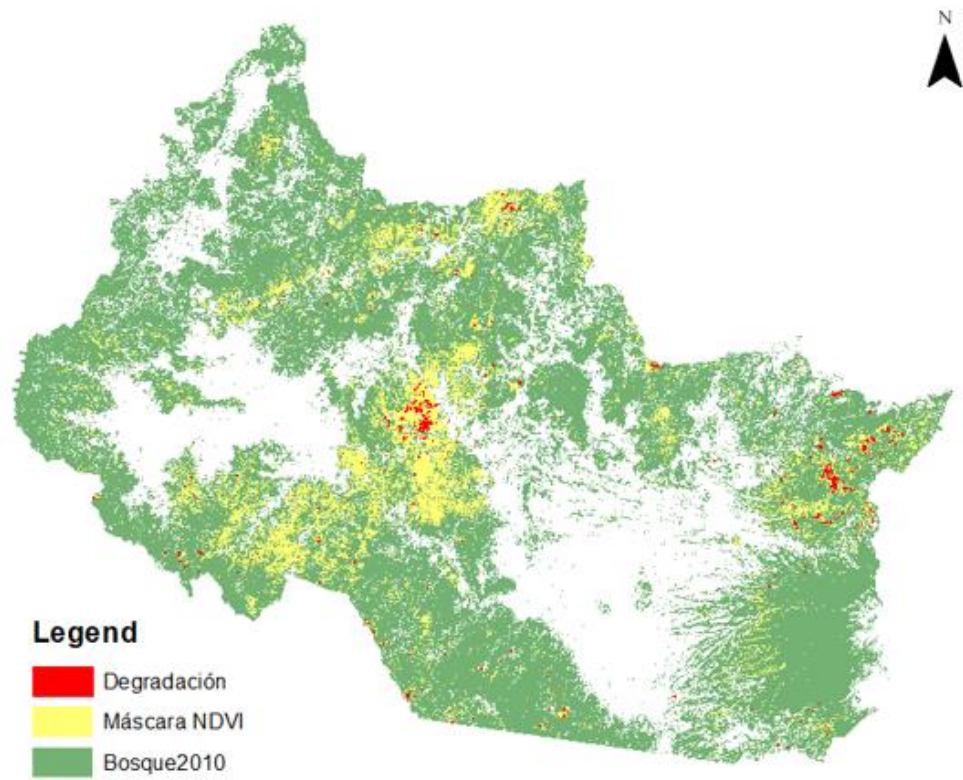


Figura 40. Degradación forestal detectada por el NDVI

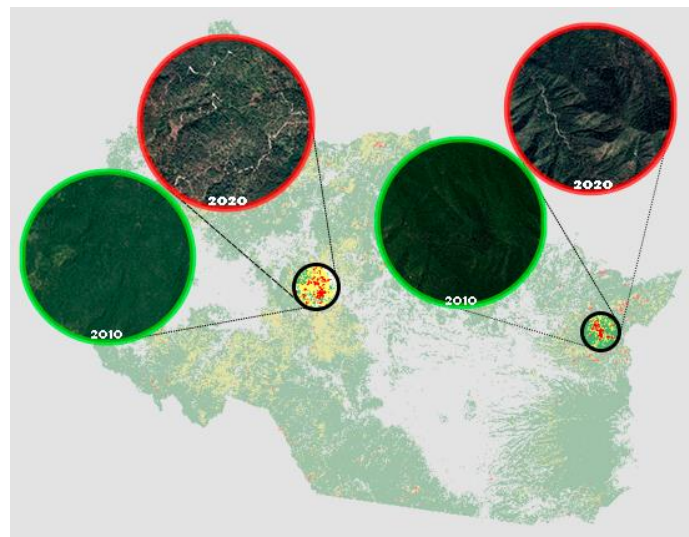


Figura 41. Visualización de las principales áreas con degradación forestal detectado por el NDVI

5.6 DISTRIBUCIÓN DE ÁREAS IDENTIFICADAS COMO DEGRADACIÓN

Dentro de los dos grupos principales mencionados anteriormente se puede observar (Figura 42) que en el primero de ellos (Grupo A) los píxeles tienden a estar un tanto dispersos, salvo en los que aparecen en la zona sur del mismo grupo donde por medio de interpretación visual, se pudo notar que la mayoría de ellos están unidos en subgrupos de entre 30 y 70 píxeles mientras que los píxeles restantes de este mismo grupo son muy variados y aunque la mayoría tienden a estar muy cercanos unos de los otros, gran parte son píxeles que se encuentran sin agrupar. Por otro lado, en el segundo grupo principal identificado (grupo B), los píxeles tienden a estar mejor agrupados, formando mayor cantidad de subgrupos los cuales contienen entre 5 y 55 píxeles, estos subgrupos se encuentran mucho más cercanos entre sí, que los subgrupos del Grupo A, y aunque también hay píxeles que se encuentran sin agrupar, estos son muy pocos.

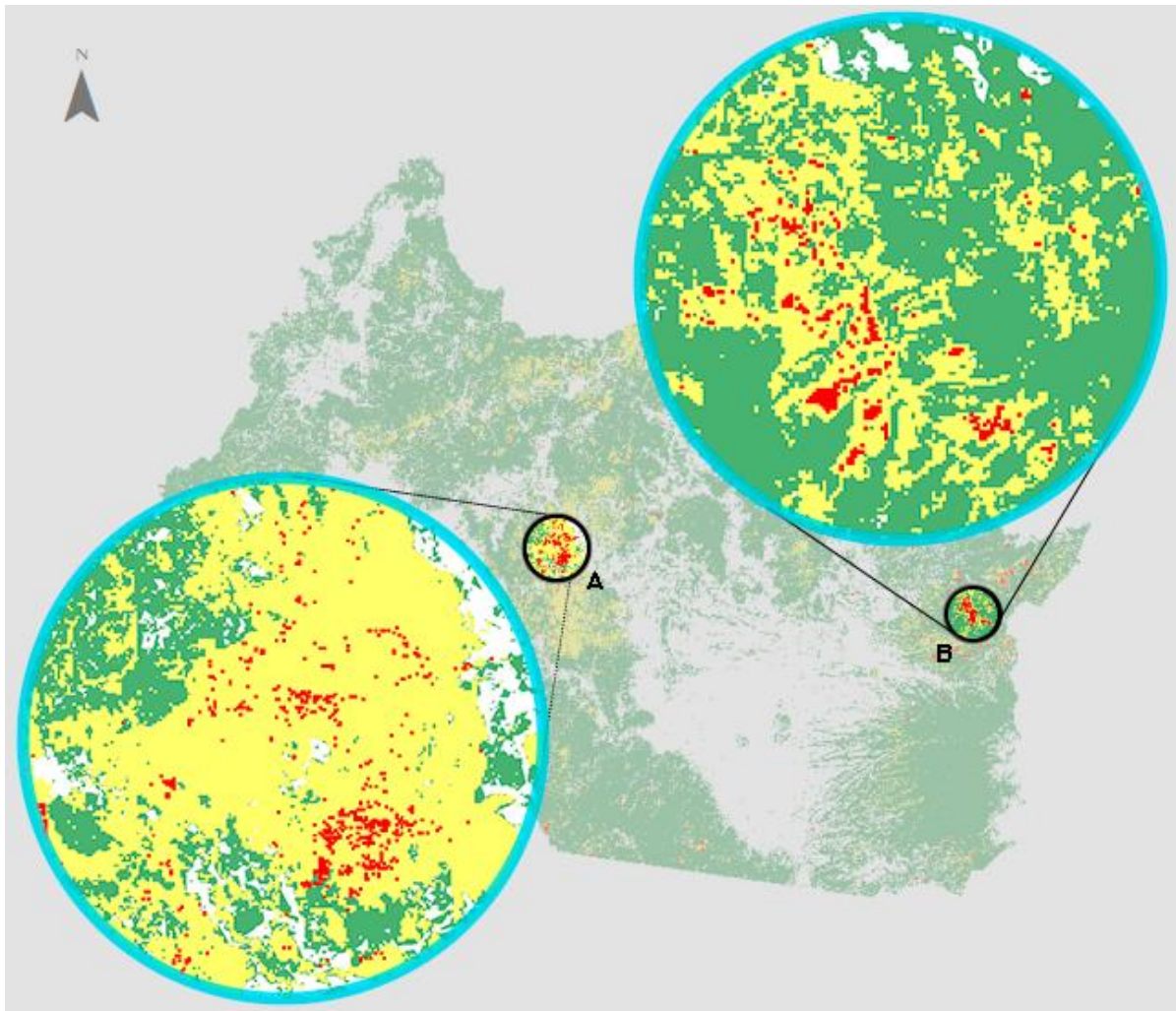


Figura 42. Distribución de áreas con degradación identificadas por el NDVI

Además de estos dos grupos, también existen otras áreas en donde el modelo identificó degradación forestal, sin embargo, estas áreas son muy dispersas y pequeñas, ya que muchas de ellas corresponden al tamaño de un solo píxel y de los grupos que se forman la mayoría son más pequeños que los subgrupos formados en los grupos principales, ya que estos van desde 2 hasta 13 píxeles. No obstante, es importante no dejar de lado a estos últimos ya que, aunque sean grupos pequeños o píxeles independientes, también representan áreas donde según los resultados obtenidos por el modelo existe degradación forestal.

5.7 COMPARACIÓN DE RESULTADOS CON GLOBAL FOREST WATCH

Para reafirmar los resultados obtenidos se realizó una comparación con la información que ofrece “Global Forest Watch” (GFW), la cual es una plataforma en línea que cuenta con datos actualizados de los cambios y afectaciones de la cubierta forestal en varias partes del mundo. Dicha plataforma se compone por más de 100 organizaciones que ponen a su disposición los datos e información necesaria para poder monitorear los bosques. GFW hace uso de datos satelitales, diversos algoritmos y la potencia del cómputo en la nube, para ofrecer un conjunto de herramientas de libre acceso las cuales están puestas a disposición para permitir que cualquier usuario pueda obtener información sobre el cambio forestal (Global Forest Watch, s.f.).

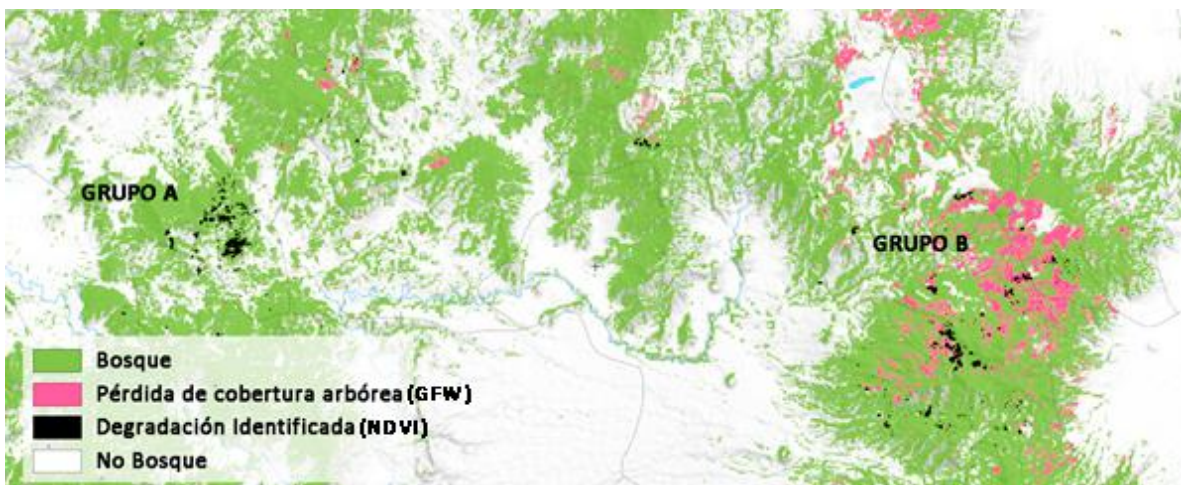


Figura 43. Datos de la plataforma GFW comparado con los resultados obtenidos por el NDVI

En la Figura 43 se muestran los grupos analizados anteriormente, el grupo A y el grupo B, los cuales corresponden a las áreas en las que el modelo detectó mayor degradación forestal (Representado de color negro), también se muestran los datos ofrecidos por la plataforma “Global Forest Watch”, donde se pueden ver aquellas áreas que se conforman de bosque (color verde) y las áreas que según los datos de esta plataforma tienen pérdida de cubierta forestal (color rosa). Como se puede ver en dicha imagen el grupo B parece corresponder con las áreas que la plataforma GFW identifica como pérdida forestal, sin embargo, el grupo

A parece no tener relación alguna con los datos de la plataforma a pesar de que como se había visto en la Figura 41 si se puede notar una disminución bastante considerable en la cubierta forestal de esa zona, no obstante es importante reconocer que la información arrojada por GFW corresponde a la pérdida total de la cubierta forestal, lo cual en esta zona no es el caso, ya que aunque no conserva la misma calidad, sigue siendo bosque. Ahora bien, los resultados obtenidos por el modelo de análisis de serie de tiempo del NDVI que forman parte del grupo B tienden a estar más cercanas a las áreas con pérdida de cobertura forestal identificadas por la plataforma GFW y aunque como se puede ver en la Figura 44 en muchas de las zonas si se intersecan ambas, existen también muchas otras que, aunque están muy cerca no son correspondidas. Esto nos podría indicar que en áreas donde se ha tenido deforestación o pérdidas de cobertura arbórea es también común encontrar áreas cercanas a ellas que, aunque siguen siendo bosque estas son sujetas a degradación forestal.

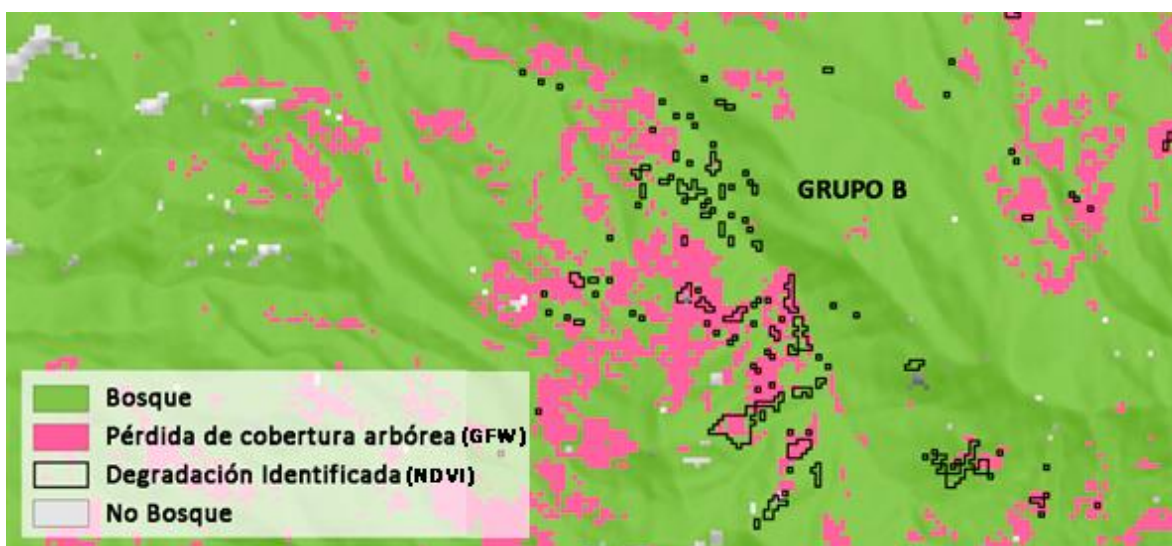


Figura 44. Comparación de Grupo B con GFW

Si analizamos el área total de la cuenca de Ayuquila en la plataforma GFW se obtiene que de 2010 a 2020 en toda la zona se ha perdido un total de 3.91kha, tomando en cuenta que este valor no solo incluye degradación, sino también procesos de deforestación, tala, incendios, entre otras perturbaciones que afectan a la cubierta forestal. En la Figura 45 se observa cómo se fue perdiendo por año la cubierta forestal en la zona de la cuenca de Ayuquila, dónde cómo se puede ver, de 2010 a 2016 la pérdida parece estar en un rango de entre 100 y 200

hectáreas perdidas por año, pero en 2017 estas se incrementan hasta 377 hectáreas, en 2018 y 2019 sigue incrementando, llegando hasta 1820 hectáreas. En la Figura 46 se muestra cómo se distribuyen las áreas que presentan pérdida de cobertura forestal, donde se puede observar que la mayoría se ubica en la zona noreste de la cuenca de Ayuquila.

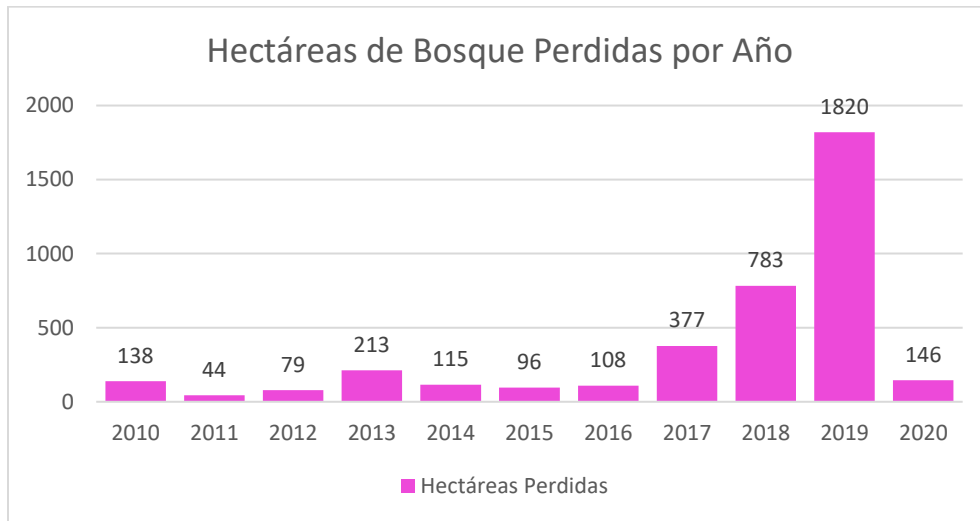


Figura 45. Área de cubierta forestal pérdida por año en la Cuenca de Ayuquila (en ha/año), según datos de GFW

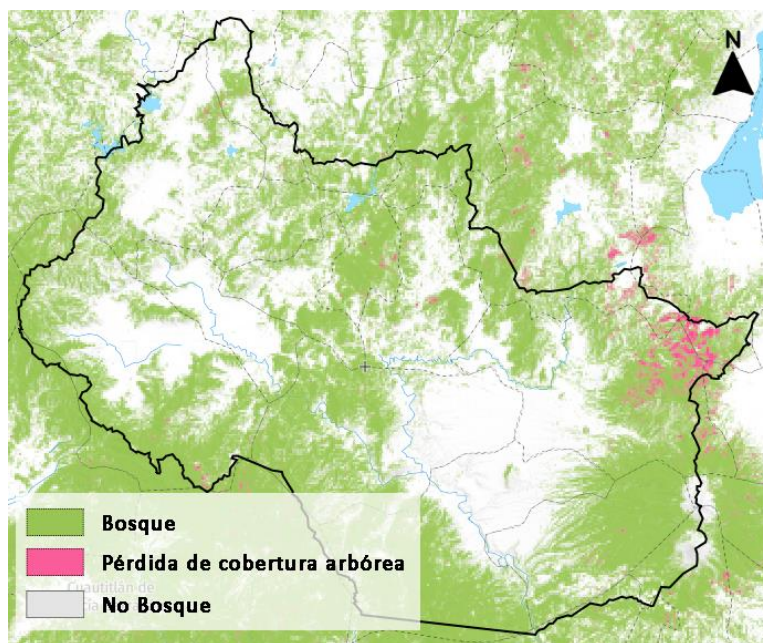


Figura 46. Pérdida de cubierta forestal en la Cuenca de Ayuquila, según datos de GFW

Además de realizar el análisis en toda la cuenca, también se procedió a revisar en la plataforma GFW, la pérdida de cubierta forestal que ésta identifica en los dos grupos principales antes mencionados. Como se muestra en la Figura 43 y como se mencionaba anteriormente, en el grupo A pareciera no tener mucha relación como los datos arrojados por la plataforma, sin embargo, si se detecta pérdida de cobertura forestal, ya que de 2010 a 2020 en esa zona, GFW detecta que se perdieron 4 hectáreas de bosque. En la Figura 47 se muestra cómo se distribuye esta pérdida en el transcurso del lapso de nuestro estudio, y en la Figura 48 se muestra cómo se distribuyen estas áreas dentro del grupo A.

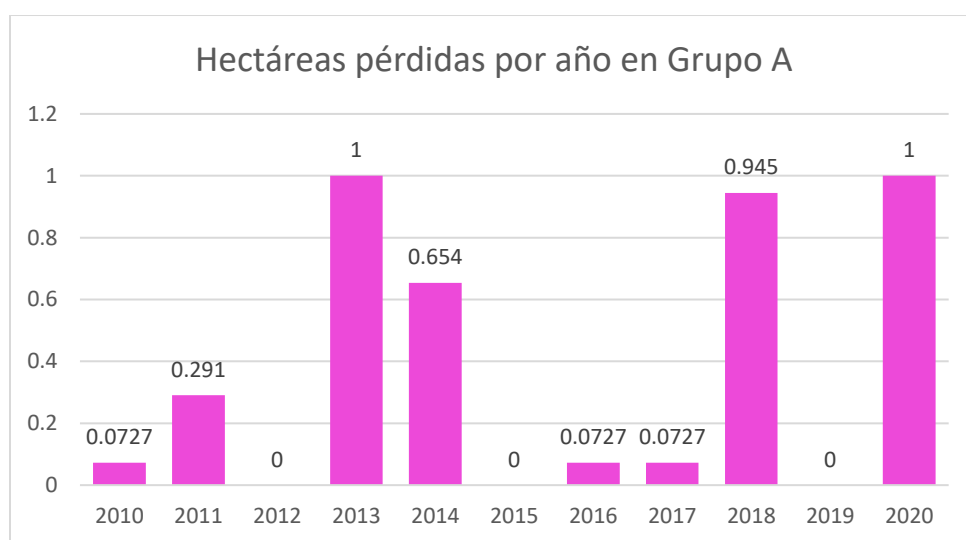


Figura 47. Área de cubierta forestal en el Grupo A pérdida por año, según datos de GFW

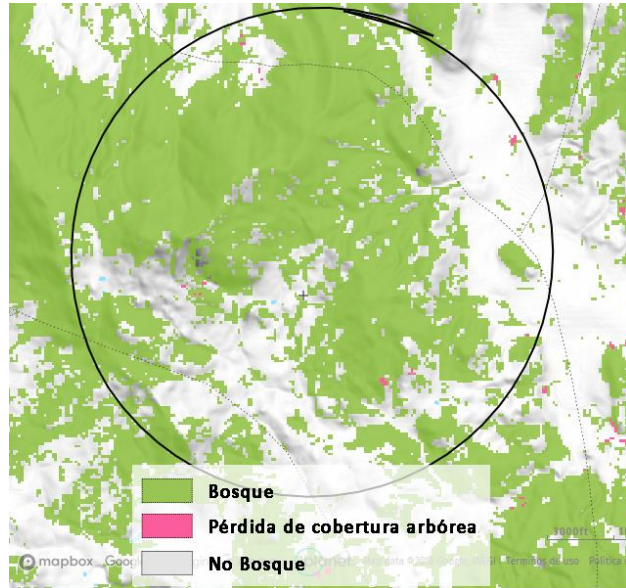


Figura 48. Pérdida de cubierta forestal en Grupo A, según datos de GFW

Para el grupo B en la plataforma si se muestra una pérdida muy considerable, y está según datos de esta asciende a 317 hectáreas, mismas que se distribuyen como indica la Figura 49, donde se puede ver que las pérdidas más considerables se tuvieron en 2017, un poco en 2018 y aumentando considerablemente en 2019. Además, en la Figura 50 se muestran las áreas de este grupo en donde se identifica la pérdida de cobertura forestal.

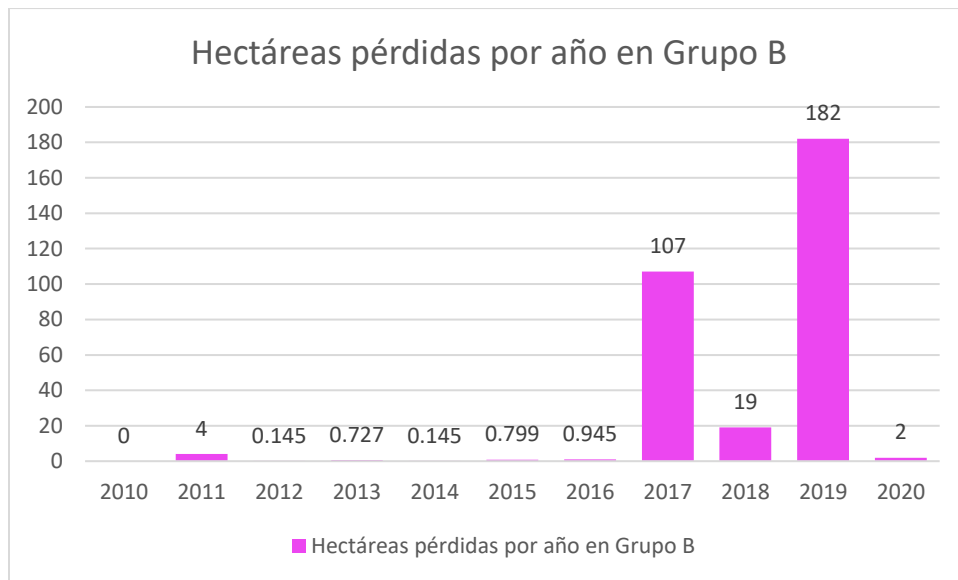


Figura 49. Área de cubierta forestal en Grupo B pérdida por año, según datos de GFW

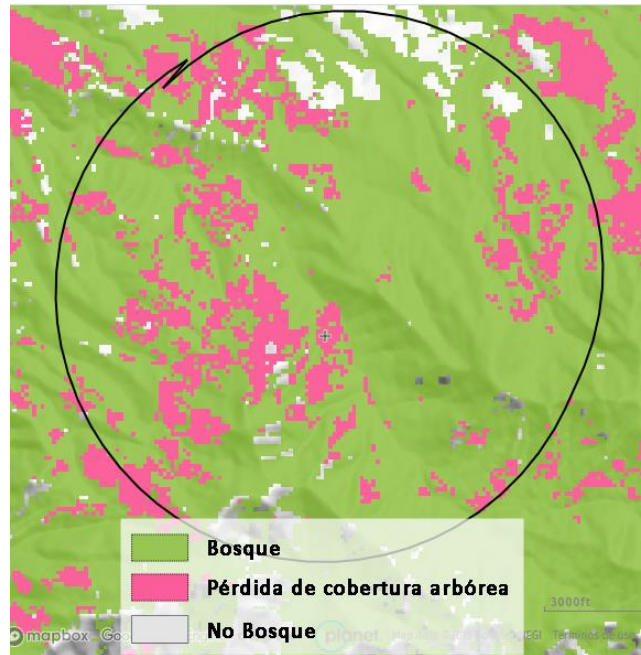


Figura 50. Pérdida de cubierta forestal en Grupo B, según datos de GFW

CONCLUSIÓN

La degradación forestal juega un papel muy importante en las emisiones de gases de efecto invernadero y en el cambio climático, es por ello, por lo que este trabajo tuvo como principal objetivo estimar la degradación forestal en la cuenca del Río Ayuquila a través de un análisis de serie de tiempo anual de los índices NDVI, NDMI y NDFI derivados de imágenes Landsat para el periodo 2010 - 2020. Para ello se obtuvieron 10 imágenes Landsat colección 1. La fecha de adquisición de todas las imágenes se concentra en los meses de febrero a abril que corresponde a temporada de seca para así evitar los falsos cambios causados por la estacionalidad de la vegetación. A las imágenes se les aplicaron correcciones radiométricas y atmosféricas, con el objetivo de obtener su valor de reflectancia en el nivel de superficie.

Los índices de NDVI y NDMI se correlacionan con el verdor y biomasa de la vegetación, mientras que el índice de fracción (NDFI) se usa para medir los cambios en la vegetación, cambios relacionados con la degradación. Para obtener el NDFI, se requirió calcular las imágenes fracción de suelo, sombra, vegetación y vegetación no fotosintética, las cuales se obtuvieron mediante la aplicación del algoritmo de mezcla espectral (SMA), para el cual se identificaron a su vez los miembros puros o “endmembers” de las imágenes de fracción requeridas (suelo, sombra, vegetación y vegetación no fotosintética) para la obtención del NDFI.

Para identificar la degradación forestal, se aplicó sobre la serie de imágenes anuales de los índices un modelo de regresión lineal simple, con el fin de identificar la tendencia en el nivel de vegetación, humedad o fracción según el índice. Además, para cumplir con el objetivo de este análisis y que los resultados obtenidos fueran confiables, se extrajeron únicamente los valores de las pendientes significativas ($p < 0.05$) del modelo de regresión, donde únicamente se tomaron en cuenta las que tuvieran un valor negativo ya que esto indicaría áreas que durante el periodo de estudio (2010 – 2020) tuvieron una disminución en el nivel de vegetación. Para descartar áreas cuya disminución se debiera a procesos de deforestación, se aplicó una máscara de bosque, de tal forma que solo nos enfocamos en aquellas áreas que tanto al inicio como al final, representaban a la clase de bosque. Esta máscara se realizó

mediante la aplicación de una clasificación supervisada, en específico con el método de máxima verosimilitud. Al realizar todo este proceso pudimos alcanzar el objetivo que se buscaba en un inicio, el cual era detectar todas aquellas áreas de la cuenca del río Ayuquila que presentaran degradación forestal, sin tomar en cuenta las áreas con deforestación. Con la verificación de los resultados se pudo corroborar que el análisis de serie de tiempo con el NDVI fue bastante certero.

La evaluación de la precisión mostró que el índice que logró la mayor confiabilidad fue el NDVI con un $81.33\% \pm 6.23\%$, de confiabilidad global, seguido por el NDMI con $73,33\% \pm 7.07\%$ y finalmente el NDFI con $72.00\% \pm 7.18\%$. El uso del NDVI como índice para medir la degradación forestal a través de un análisis de series de tiempo implementado a través de un análisis de regresión permitió identificar que un total de 191 hectáreas en la cuenca de Ayuquila presentaron degradación forestal durante el periodo de 2010 - 2020. Con base en los resultados obtenidos se pudo observar que el NDVI y NDMI son comparables ya que ambos representan la abundancia y la salud de la vegetación. Sin embargo, con el NDFI en este estudio, no se lograron los resultados esperados, las posibles razones incluyen la dificultad de localizar los miembros puros de las imágenes Landsat, mismos que son necesarios para calcular las imágenes fracción de suelo, sombra, vegetación y vegetación no fotosintética, que a su vez son el insumo para el cálculo del NDFI.

BIBLIOGRAFÍA

- Aguilar, H., Mora Zamora, R. & Vargas Bolaños, C., 2014. Metodología para la corrección atmosférica de imágenes Aster, RapidEye, Spot 2 y Landsat 8 con el módulo Flaash del software Envy. *Revista Geográfica de América Central*, Issue 53, pp. 39-59.
- Alonso, J., 2019. *Deforestación y degradación, las mayores amenazas de los bosques en América Latina*. [En línea]
Available at: <https://www.forbes.com.mx/deforestacion-y-degradacion-las-mayores-amenazas-de-los-bosques-en-america-latina/>
[Último acceso: 16 Diciembre 2019].
- Anand, A., Singh, S. K. & Kanga, S., 2018. Estimating the change in Forest Cover Density and Predicting NDVI for West Singhbhum using Linear Regression. *International Journal for Environmental Rehabilitation and Conservation*, IX(1), pp. 193-203.
- Ariza, A., 2013. *Instituto geográfico Agustín Codazzi*. [En línea]
Available at: <http://www.un-spider.org/sites/default/files/LDCM-L8.R1.pdf>
[Último acceso: 29 Abril 2020].
- Ariza, A., 2013. *INSTITUTO GEOGRÁFICO AGUSTÍN CODAZZI*. [En línea]
Available at: <http://www.un-spider.org/sites/default/files/LDCM-L8.R1.pdf>
[Último acceso: 03 Enero 2020].
- Barrera, R. A., 2015. *Procesamiento, interpretación y análisis de imágenes satelitales en la banda de los 400 a 700 nanómetros para aplicaciones de percepción remota*, México DF: s.n.
- Basuki, T. M. y otros, 2012. The potential of spectral mixture analysis to improve the estimation accuracy of tropical forest biomass. *Geocarto International*, XXVII(4), p. 329–345.
- Blanc, L., Gond, V. & Ho Tong Minh, D., 2016. Remote Sensing and Measuring Deforestation. En: *Land Surface Remote Sensing*. s.l.:s.n., pp. 27-53.
- Borrego, A., Salinas Melgoza, M. & Skutsch, M., 2015. El valor de las emisiones de carbono asociadas al sistema agrícola de roza, tumba y quema. *Programa Mexicano del Carbono*, pp. 512-520.
- Bowen, L. M. R., 2011. Segmentación y clasificación de imágenes satelitales para determinar la cobertura de suelo. *LA TÉCNICA: REVISTA DE LAS AGROCIENCIAS*, pp. 38-44.
- Bullock, E. L., Woodcock, C. E. & Olofsson, P., 2018. Monitoring tropical forest degradation using spectral unmixing and Landsat time series analysis. *Remote Sensing of Environment*.
- Campos, A., Guerrero, E. & Gines, E., 2018. Evolución de la cobertura vegetal en el Parque Nacional Cerros de Amotape de Tumbes utilizando el Índice de Vegetación de Diferencia Normalizada (NDVI): 2000 - 2014. *Revista de Investigación Científica*, pp. 47-56.

- Candra, D. S., Phinn, S. & Scarth, P., 2016. Cloud and cloud shadow masking using multi-temporal cloud masking algorithm in tropical environment. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volumen XLI-B2, pp. 95-100.
- Castillo, I. M. & Guijarro, M. G., 2006. *Estadística descriptiva y cálculo de probabilidades*. Madrid: Pearson.
- Castillo, E., R., P., De Luis Ruiz, J. M. & Sánchez, J. M., 2009. La corrección atmosférica en la teledetección aerotransportada. *Teledetección: Agua y desarrollo sostenible. XIII Congreso de la asociación Española de teledetección*, Volumen XIII, pp. 409-412.
- Castro, D. R. M., 2013. Monitores de bosques utilizando NDVI REDEDEde RapidEye. *Revista Geoespacial*, Issue 10, pp. 58-71.
- Chavez, P. S. J., 1988. An Improved Dark-Object Subtraction Technique for Atmospheric Scattering Correction of Multispectral Data. *Remote sensing of environment*, Issue 24, pp. 549-479.
- Chavez, P. S. J., 1996. Image-Based Atmospheric corrections - Revisited and improved. *Photogrametric Engineering & Remote Sensing*, 62(9), pp. 1025-1036.
- Chuvieco, E., 1995. *Fundamentos de teledetección espacial*. Madrid: RIALP, S.A..
- Cochran, W. G., 1977. *Sampling Techniques*. Tercera ed. Nueva York: s.n.
- Comisión Nacional Forestal, 2017. *Gobierno de México*. [En línea]
Available at: <https://www.gob.mx/conafor/documentos/bosques-y-cambio-climatico-23762>
- CONAFOR, 2013. *Bosques, cambio climático y REDD+ en México*. [En línea]
Available at:
<http://www.conafor.gob.mx:8080/documentos/docs/35/4034Gu%C3%ADa%20B%C3%A1sica%20de%20Bosques,%20Cambio%20Clim%C3%A1tico%20y%20REDD+%20.pdf>
- CONAFOR, 2017. *Estrategia Nacional para REDD+*. Issue 1.
- CONAFOR, 2018. *Reducción de Emisiones por Deforestación y Degradación Forestal y la conservación e incremento de acervos de carbono forestal (REDD+)*. [En línea]
Available at:
<http://www.conafor.gob.mx:8080/documentos/docs/1/7618REDD+%20en%20M%C3%A9xico.pdf>
[Último acceso: 16 Diciembre 16].
- Cuartero, A. & Felicísimo, A., 2003. Rectificación y ortorrectificación de imágenes de satélite: análisis comparativo y discusión. *Revista internacional de ciencia y tecnología de la información geográfica*, Issue 3, pp. 45-57.
- Cuevas, R. G., Núñez López, N. M., Guzmán Hernández, L. & Santana Michel, F. J., 1998. El bosque tropical caducifolio en la reserva de la biosfera sierra de Manantlan, Jalisco-Colima, México. *Boletín IBUG*, 5(1-3), pp. 445-491.
- Defries, R. S., Hansen, M. C. & Townshend, J. R. G., 2000. Global continuous fields of vegetation characteristics: A linear mixture model applied to multi-year 8 km AVHRR data. *International Journal of Remote Sensing*, p. 1389–1414.

- Del Toro , N., Gomariz, F., Cánovas, F. & Sarría, F. A., 2015. Comparación de métodos de clasificación de imágenes de satélite en la cuenca del río argos (Región de Murcia). *Boletín de la Asociación de Geógrafos Españoles*, Issue 67.
- Espinosa, J. J. Z., 2019. Aplicación de algoritmos Random Forest y XGBoost en una base de solicitudes de tarjetas de crédito. *Ingeniería Investigación y Tecnología*, XXI(3), pp. 1-16.
- García, F. & Maass, J. M., 1998. Efecto de la transformación de la selva a pradera sobre la dinámica de los nutrientes en un ecosistema tropical estacional en México. *Boletín de la Sociedad Botánica de México*, Issue 62, pp. 39-48.
- Ghazoul, J., Burivalova, Z., Garcia Ulloa, J. & King, L. . A., 2015. Conceptualizing Forest Degradation. *Trends in Ecology & Evolution*, XXX(10), pp. 622-632.
- Ghilardi, A. y otros, 2012. Analisis de cambio de cobertura y uso del suelo, escenario de referencia de car-bono y diseño preliminar del mecanismo de monitoreo, reporte y verification en los diez municipios de la Junta Intermunicipal del Rio Ayuquila. *Centro de Investigaciones en Geografía Ambiental (CIGA)*.
- Giraldo, J. & García, S., 2009. Modelo de análisis de subpíxeles aplicado a la cartografía de inundaciones en zonas de llanura. *Teledetección: Agua y desarrollo sostenible*, Volumen VIII, pp. 569-572.
- Global Forest Watch, s.f.. *Global Forest Watch*. [En línea]
Available at: <https://www.wri.org/initiatives/global-forest-watch>
[Último acceso: 20 08 2021].
- Gocic, M. & Trajkovic, S., 2012. Analysis of changes in meteorological variables using Mann-Kendall and Sen's slope estimator statistical tests in Serbia. *Global and Planetary Change*, Issue 100, pp. 172-182.
- Hantson, S. y otros, 2011. Cadena de pre-procesamiento estándar para las imágenes Landsat del Plan Nacional de Teledetección. *Revista de Teledetección*, Issue 36, pp. 51-61.
- Henry , V. V., 2008. *Instituto Colombiano de Geología y Minería*. [En línea]
Available at:
<http://recordcenter.sgc.gov.co/B12/23008002524448/documento/pdf/2105244481102000.pdf>
- Hernández, R. M., Martínez, C. A. P. & Navarro , S. G. J., 2019. Redes neuronales y árboles de decisión para la clasificación de objetos astronómicos. *Research in Computing Science*, 7(148), pp. 477-489.
- Huang, S. & Kong, J., 2016. Assessing Land Degradation Dynamics and Distinguishing Human-Induced Changes from Climate Factors in the Three-North Shelter Forest Region of China. *International Journal of Geo-Information*.
- INEGI, 2019. *INEGI*. [En línea]
Available at: <https://www.inegi.org.mx/app/buscador/default.html?q=percepcion+remota>

- Jardel, E. y otros, 2012. Estudio para determinar el estado de degradación y potencial de restauración en ecosistemas forestales bajo manejo de la Cuenca Baja Del Rio Ayuquila – Parte 1, Informe final.
- Jensen, J. R., 2015. *Introductory Digital Image Processing*. Tercera ed. s.l.:Pearson.
- Jin, S. & Sader, S. . A., 2005. Comparison of time series tasseled cap wetness and the normalized difference moisture index in detecting forest disturbances. *Remote Sensing of Environment*, Issue 94, pp. 364-372.
- Krammer, T. A. & Saeedi, P., 2018. Improving Landsat 8 Cloud Detection Algorithms via a new Snow Identification and Separation Algorithm. En: H. R. Arabnia, . M. R. Grimaila, D. D. Hodson & F. G. Tinetti, eds. *Proceedings of the 2018 International Conference on Scientific Computing*. s.l.:s.n., pp. 27-33.
- Leutner, B., Horning, N., Schwalb-Willmann, J. & Hijmans, R. J., 2019. *cran.r-project*. [En línea] Available at: <https://cran.r-project.org/web/packages/RStoolbox/RStoolbox.pdf> [Último acceso: 01 05 2020].
- Levin, I. R. & Rubin, R. D., 2004. *Estadística para administración y economía*. Séptima ed. México: Pearson.
- Leyva, Á. y otros, 2017. Monitoreo de la degradación forestal en México con base en el inventario nacional forestal y de suelos (Infys). *Madera y Bosques*, pp. 69-83.
- Maldonado, J. D., Alatorre, L. C. & Torres, M. E., 2015. Análisis de las tendencias del NDVI con imágenes satelitales en Cuauhtémoc, Chihuahua (2000-2014). *Memorias de resúmenes en extenso SELPER-XXI-México-UACJ-2015*.
- Mas, J. F., Díaz Gallegos, J. R. & Pérez Vega, A., 2003. Evaluación de la confiabilidad temática de mapas o de imágenes clasificadas: una revisión. *Boletín del Instituto de Geografía, UNAM*, Issue 51, pp. 53-72.
- Mohammed, M. A., Rokhmatuloh, Salem, Z. E. & Javzandulam, T., 2006. *Linear Mixture Model Applied to the Land-Cover Classification*. [En línea] Available at: <http://www.jeionline.org/index.php?journal=mys&page=article&op=view&path%5B%5D=200600071>
- Morales, J., 2013. *Obtención y clasificación de píxeles mezclados generados por regiones de transición en imágenes satelitales*, Puebla: s.n.
- Muñoz , P. A., 2013. *Centro de Información de Recursos Naturales*. [En línea] Available at: <http://bibliotecadigital.ciren.cl/bitstream/handle/123456789/26389/Tema%20Indicadores%20de%20vegetaci%C3%B3n%20Pedro%20Mu%C3%B1oz%20A.pdf?sequence=1&isAllowed=y> [Último acceso: 30 Octubre 2020].
- Muñoz, E. A., Zozaya, A. & Lindquist, E., 2020. Satellite Remote Sensing of Forest Degradation Using NDFI and the BFAST Algorithm. *IEEE Latin America Transactions*, XVIII(7), pp. 1288-1295.

- Mwampamba, T. H., Salinas, M., Larrazabal, A. & Skutsch, M., 2015. *Opciones en REDD+ para ejidos en la cuenca de Ayuquila, Jalisco*, Morelia, Michoacan: s.n.
- Newcombe, R. G. & Merino, C. S., 2006. Intervalos de confianza para las estimaciones de proporciones y las diferencias entre ellas. *SciELO*, 23(2), pp. 141-154.
- Olofsson, P. y otros, 2014. Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment*, pp. 42-57.
- Parera, M., Bustos, H. & Gerritsen, P., 2001. Hacia una conservación basada en la gente. *Ecología política*, Issue 21, pp. 43-50.
- Paz Pellat, F., C. J., González, W. & Torres Alamilla, R., 2015. Estado Actual del Conocimiento del Ciclo del Carbono y sus Interacciones en México: Síntesis a 2015. *Programa Mexicano del Carbono*, pp. 292-299.
- Prajesh, P. y otros, 2019. Monitoring and mapping of seasonal vegetation trend in Tamil Nadu using NDVI and NDWI imagery. *Journal of Applied and Natural Science*, XI(1), pp. 54-61.
- Priyanka, J. F. V., 2020. *Freygeospatial*. [En línea]
Available at: <https://freygeospatial.github.io/PM25-TimeSeries-R-Tutorial/>
[Último acceso: 18 Mayo 2021].
- Pushpendra, S., Vivekanand, T. & Anil, K., 2014. Analysis of supervised maximum likelihood classification for remote sensing image. *IEEE International Conference on Recent Advances and Innovations in Engineering*.
- Qiu, S., Zhu, Z. & He, B. B., 2019. Fmask 4.0: Improved cloud and cloud shadow detection in Landsats 4-8 and Sentinel-2 imagery. *Teledetección del medio ambiente*, Volumen 231.
- Quintano, y otros, 2003. Estimación del área quemada por incendios forestales Empleando el modelo de mezclas espectrales desde una Perspectiva multitemporal. *Teledetección y Desarrollo Regional. X Congreso de Teledetección.*, pp. 175-178.
- Rendon, E., Mendoza, M., Cisniego, R. & Carbajal, G., 2015. Una versión modificada del algoritmo de agrupamiento Isodata. *Revista Tecnología e Innovación*, II(5), pp. 942-957.
- Ren, H., Cai, G. & Du, M., 2019. Surface Heterogeneity-Involved Estimation of Sample Size for Accuracy Assessment of Land Cover Product from Satellite Imagery. *Sensors*, pp. 1-20.
- Rojas, Y., Loguercio, G., Nieto, V. & Bahamondez, C., 2012. *Análisis de la degradación forestal en el marco de REDD+*, Chile: s.n.
- Sánchez Muñoz, J. M., 2016. Análisis de la calidad cartográfica mediante el estudio de la Matriz de Confusión. *Pensamiento Matemático*, VI(2), pp. 9-26.
- Santana, J. A. M., 2015. *JIRA.org*. [En línea]
Available at: https://www.jira.org.mx/imagenes/trasparencia/Programa_de_inversion_2016.pdf
[Último acceso: 04 Noviembre 2020].

- Santi , N. A. R., 2018. *UNIVERSIDAD TÉCNICA ESTATAL DE QUEVEDO*. [En línea]
Available at: <https://repositorio.uteq.edu.ec/bitstream/43000/3961/1/T-UTEQ-0107.pdf>
[Último acceso: 27 10 2020].
- Schneibe, A. y otros, 2017. Using Annual Landsat Time Series for the Detection of Dry Forest Degradation Processes in South-Central Angola. *Remote Sensing*, pp. 1-14.
- Schultz, M. y otros, 2016. Performance of vegetation indices from Landsat time series in deforestation monitoring. *International Journal of Applied Earth Observation and Geoinformation*, Issue 52, pp. 318-327.
- SEMAR, 2009. *Percepción Remota*. [En línea]
Available at: <http://2006-2012.semar.gob.mx/informacion-sector/ciencia/ermexs/temas-de-interes/384-percepcion-remota.html>
- Shi, N., Guan , y. & Liu , X., 2010. Research on k-means Clustering Algorithm. *2010 Third International Symposium on Intelligent Information Technology and Security Informatics*, pp. 63-67.
- Sotaquirá, M., 2021. *Codificandobits*. [En línea]
Available at: <https://www.codificandobits.com/blog/bosques-aleatorios/>
[Último acceso: 27 Octubre 2021].
- Souza, C., Firestone, L., Moreira Silva, L. & Roberts, D., 2002. Mapping forest degradation in the Eastern Amazon from SPOT 4 through spectral mixture models. *Remote Sensing of Environment*, pp. 496-506.
- Souza, C. M., Roberts, D. A. & Monteiro, A. L., 2005. Multitemporal Analysis of Degraded Forests in the Southern Brazilian Amazon. *Earth Interactions*, IX(19).
- Souza, C. M. y otros, 2013. Ten-Year Landsat Classification of Deforestation and Forest Degradation in the Brazilian Amazon. *Remote Sensing*, Issue 5, pp. 5493-5513.
- Stehman, S. , V. & Czaplewski, R. L., 1998. Design and Analysis for Thematic Map Accuracy Assessment: Fundamental Principles. *ELSEVIER*, Volumen 64, pp. 331-344.
- Tarazona, Y. & Miyasiro, M., 2020. Monitoring tropical forest degradation using remote sensing. Challenges and opportunities in the madre de dios region, Peru. *Remote Sensing Applications: Society and Environment*, Issue 19, pp. 1-13.
- Tun-Dzul, F. J., Hans , V., Durán , R. G. & Schmook, B., 2008. Estructura arbórea y variabilidad temporal del NDVI en los “bajos inundables” de la península de Yucatán, México. *Polibotánica*, Issue 25, pp. 69-90.
- UNFCCC, 2007 . [En línea]
Available at: https://unfccc.int/resource/docs/publications/unitingonclimate_spa.pdf
[Último acceso: 26 Enero 2020].

USGS, 2017. *Misiones Landsat*. [En línea]

Available at: [usgs.gov/land-resources/nli/landsat/landsat-8?qt-science support page related con=0#qt-science support page related con](https://www.usgs.gov/land-resources/nli/landsat/landsat-8?qt-science_support_page_related_con=0#qt-science_support_page_related_con)

USGS, 2019. *USGS*. [En línea]

Available at: <https://www.usgs.gov/land-resources/nli/landsat>
[Último acceso: 19 Marzo 2020].

USGS, 2019. *USGS science for a changing world*. [En línea]

Available at: <https://www.usgs.gov/earthexplorer-0>
[Último acceso: 18 Diciembre 2019].

Wang, C. & Myint, S. W., 2016. Environmental Concerns of Deforestation in Myanmar 2001–2010. *Remote Sensing*.

Wilson, E. H. & Sader, S. A., 2002. Detection of forest harvest type using multiple dates of Landsat TM imagery. *Remote Sensing of Environment*, Issue 80, p. 385–396.

Zhu, Z. y otros, 2016. Including land cover change in analysis of greenness trends using all available Landsat 5, 7, and 8 images: A case study from Guangzhou, China (2000–2014). *Remote Sensing of Environment*, Issue 185, pp. 243-257.