



Universidad Autónoma del Estado de México
Facultad de Geografía
Maestría en Análisis Espacial y Geoinformática

**Análisis Multitemporal de Cambio de Uso de Suelo en la
Zona Metropolitana de Tianguistenco, Estado de México**

Trabajo Terminal de Grado

Que para obtener el grado de
Maestra en Análisis Espacial y Geoinformática

Presenta:

ECATSIG. María de la Luz Pineda Monroy

Director:

Dr. Francisco Zepeda Mondragón

Codirectores:

Dra. Giovanna Santana Castañeda

Dra. Jocelyn Alejandra Cortez Núñez

Noviembre, 2024

ACRONIMOS

COESPO	Consejo Estatal de Población
APFyF	Área de Protección de Flora y Fauna
ESA	Agencia Espacial Europea
FAO	Organización de las Naciones Unidas para la Alimentación y la Agricultura
GFW	Global Forest Watch
GEE	Google Earth Engine
IPCC	Grupo Intergubernamental de Expertos sobre el Cambio Climático
ONU	Organización de las Naciones Unidas
PROFEPA	Procuraduría de Protección al Ambiente
CONANP	Comisión de Áreas Naturales Protegidas
INEGI	Instituto Nacional de Estadística y Geografía
RF	Random Forest
SVM	Support Vector Machine
CART	Árbol de Clasificación y Regresión
LULC	Land Use/ Land Cover – Uso de suelo/Cobertura del suelo
SIG	Sistema de Información Geográfica
ZMT	Zona Metropolitana de Tlanguistenco
NDVI	Índice de Vegetación de Diferencia Normalizada
NDBI	Índice de Diferencia Urbanizada Normalizada
PA	Precisión del Productor
UA	Precisión del Usuario
LCM	Land Change Modeler

Contenido

Resumen.....	1
Introducción	2
Planteamiento del problema	3
Justificación	5
Antecedentes	6
Contexto Internacional	6
Contexto Nacional.....	8
Objetivos	12
Objetivo General	12
Objetivos Específicos.....	12
1 CAPÍTULO 1. MARCO TEÓRICO.....	13
1.1 Contexto General de la Geografía	14
1.1.1 Conceptos del Análisis Espacial	15
1.2 Sistemas de Información Geográfica (SIG).....	16
1.2.1 Componentes de un SIG	16
1.2.2 Modelos de Representación de Información Geográfica	17
1.3 Teledetección.....	18
1.3.1 Elementos de la Teledetección	18
1.3.2 Espectro Electromagnético	19
1.3.3 Tipos de Resolución	20
1.3.4 Firmas espectrales.....	21
1.3.5 Sensores.....	22
1.3.6 Landsat 8 OLI/TIRS	22
1.3.7 Sentinel-2A	24
1.4 Inteligencia artificial.....	25
1.5 Aprendizaje automático.....	26
1.6 Google Earth Engine.....	26
1.6.1 Índice de Vegetación de Diferencia Normalizada (NDVI)	28
1.6.2 Índice de Diferencia Urbanizada Normalizada (NDBI)	28
1.6.3 Random Forest	29
1.6.4 Support Vector Machine	29
1.7 Cambio de Uso de Suelo	30
1.8 Modelo predictivo.....	31

2	CAPÍTULO 2. METODOLOGÍA.....	33
2.1	Ubicación Geográfica.....	34
2.1.1	Características Físico-Geográficas.....	36
2.1.2	Uso de Suelo	37
2.1.3	Ciénegas de Lerma.....	39
2.2	Método.....	39
2.3	Fase 1. Clasificación de las imágenes satelitales	41
2.4	Fase 2. Evaluación de la Confiabilidad Temática	44
2.4.1	Diseño de muestreo.....	44
2.4.2	Evaluación de los sitios de verificación.....	45
2.4.3	Análisis de datos.....	45
2.5	Fase 3. Cambio de uso de suelo	47
2.6	Fase 4. Modelo Predictivo – Cadena de Markov	49
3	CAPÍTULO 3. RESULTADOS	50
3.1	Clasificación de imágenes satelitales.....	51
3.1.1	Random Forest en imagen Landsat 8 de 2016.....	59
3.1.2	Support Vector Machine en imagen Landsat 8 de 2016	63
3.1.3	Random Forest en imagen Landsat 8 de 2022.....	67
3.1.4	Support Vector Machine en imagen Landsat 8 de 2022	71
3.1.5	Random Forest en imagen Sentinel-2 de 2016	75
3.1.6	Support Vector Machine en imagen Sentinel-2 de 2016.....	79
3.1.7	Random Forest en imagen Sentinel-2 de 2022	83
3.1.8	Support Vector Machine en imagen Sentinel-2 de 2022.....	87
3.2	Cambio de uso de suelo en la Zona Metropolitana de Tianguistenco.....	91
3.2.1	Ganancias y pérdidas	93
3.2.2	Intercambio entre categorías	95
3.2.3	Cambio total por categoría.....	96
3.2.4	Cambio de uso de suelo durante el periodo 2016 a 2022.....	100
3.3	Modelo predictivo de uso de suelo para 2028 mediante Cadenas de Markov 104	
	Conclusiones	110
	Recomendaciones	112
	Referencias.....	113
	Anexos.....	122

Índice de Figura

Figura 1.1 Modelos de representación geográfica	17
Figura 1.2.Elementos de la teledetección	19
Figura 1.3. Espectro Electromagnético	20
Figura 1.4. Firmas espectrales	22
Figura 2.1. Mapa de ubicación de la zona de estudio: Zona Metropolitana de Tianguistenco	35
Figura 2.2. Mapa de Uso de Suelo y Vegetación, Serie VII, INEGI.....	38
Figura 2.3. Esquema metodológico.....	40
Figura 2.4. Random Forest.....	43
Figura 2.5. Support Vector Machine.....	43
Figura 3.1. Core Editor de Google Earth Engine	52
Figura 3.2. Firmas espectrales en Landsat 8 de las clases de uso de suelo	54
Figura 3.3. Firmas espectrales en Sentinel-2 de las clases de uso de suelo	55
Figura 3.4. Interfaz de Semi-Automatic Classification Plugin en QGIS	58
Figura 3.5. Clasificación de uso de suelo para el año 2016 mediante Random Forest	62
Figura 3.6. Clasificación de uso de suelo para el año 2016 mediante SVM.....	66
Figura 3.7. Clasificación de uso de suelo para el año 2022 mediante RF.....	70
Figura 3.8. Clasificación de uso de suelo para el año 2022 mediante SVM.....	74
Figura 3.9. Clasificación de uso de suelo para el año 2016 mediante RF.....	78
Figura 3.10. Clasificación de uso de suelo para el año 2016 mediante SVM.....	82
Figura 3.11. Clasificación de uso de suelo para el año 2022 mediante RF.....	86
Figura 3.12. Clasificación de uso de suelo para el año 2022 mediante SVM.....	90
Figura 3.13. Interfaz de Terrset y la herramienta Land Change Modeler	92
Figura 3.14. Mapa de cambio de uso de suelo en la ZMT	103
Figura 3.15. Herramienta Markov en Terrset.....	104
Figura 3.16. Módulo CA_MARKOV para predicciones de cambio	106
Figura 3.17. Mapa de modelo predictivo al año 2028.....	109

Índice de Tabla

Tabla 1.1. Características de las bandas de Landsat 8 OLI/TIRS.....	23
Tabla 1.2. Características de las bandas del Sensor Sentinel 2A/ 2B.....	25
Tabla 2.1. Uso de suelo y vegetación, serie VII, INEGI.....	37
Tabla 2.2. Descripción de los procesos para la aplicación de la metodología	40
Tabla 2.3. Tamaño de la muestra por clase en función de p y B.....	45
Tabla 2.4. Matriz de confusión o matriz error	46
Tabla 2.5. Coeficiente Kappa	47
Tabla 2.6. Matriz de tabulación cruzada.....	48
Tabla 3.1. Nomenclatura de imágenes de satélite del proyecto	52
Tabla 3.2. RF en imagen Landsat 8 de 2016	60
Tabla 3.3. SVM en imagen Landsat 8 de 2016	64
Tabla 3.4. RF en imagen Landsat 8 de 2022	68
Tabla 3.5. SVM en imagen Landsat 8 de 2022	72
Tabla 3.6. RF en imagen Sentinel-2 de 2016.....	76
Tabla 3.7. SVM en imagen Sentinel-2 de 2016.....	80
Tabla 3.8. RF en imagen Sentinel-2 de 2022.....	84
Tabla 3.9. SVM en imagen Sentinel-2 de 2022.....	88
Tabla 3.10. Precisión general y Coeficiente Kappa en Landsat 8 y Sentinel-2	91
Tabla 3.11. Cambios de uso de suelo representativos en la ZMT.....	100
Tabla 3.12. Predicción de los usos de suelo al año 2028	106

Índice de Gráficas

Gráfica 3.1. Ganancias y pérdidas en hectáreas durante el periodo 2016 a 2022	93
Gráfica 3.2. Intercambio entre categoría por hectáreas durante el periodo 2016 a 2022	95
Gráfica 3.3. Cambio total para bosque	96
Gráfica 3.4. Cambio total para agricultura	97
Gráfica 3.5. Cambio total para zona urbana	98
Gráfica 3.6. Cambio total para cuerpos de agua	98
Gráfica 3.7. Cambio total para tular	99
Gráfica 3.8. Cambio total para pastizal	100

Resumen

La conversión y el aprovechamiento del suelo han provocado cambios en los ecosistemas terrestres. La expansión de las actividades agrícolas y urbanas han provocado una drástica disminución de la cobertura forestal y la degradación del suelo. Analizar los procesos detrás de estos cambios es crucial debido a sus efectos en el medio ambiente, la naturaleza, la economía y sociedad.

La metrópoli de Tianguistenco ha presentado importantes cambios de uso de suelo, principalmente en el bosque debido a la tala ilegal, la demanda requerida por el campo y el incremento de áreas urbanas. En este contexto, el objetivo principal de la investigación fue analizar el proceso de cambio de uso de suelo en la Zona Metropolitana de Tianguistenco durante el periodo 2016 y 2022 mediante el uso de métodos y técnicas de teledetección para identificar perturbaciones en las coberturas de suelo.

El desarrollo del proyecto se ha dividido en cuatro fases; la primera consiste en la clasificación de imágenes Landsat 8 y Sentinel-2 a partir del algoritmo Random Forest (RF) y Support Vector Machine (SVM) en la plataforma Google Earth Engine (GEE) para la obtención de mapas de uso de suelo; la segunda fase es la comparación de la información clasificada con otra de referencia muy fiable, en este contexto, imágenes de mayor resolución espacial y trabajo de campo; como tercera etapa se obtienen los cambios de uso de suelo, como pérdidas, ganancias, cambios netos e intercambios de categorías entre ambos periodos, y por último, la aplicación de un modelo predictivo para conocer el comportamiento a lo largo del tiempo.

SVM demostró un mejor desempeño con un coeficiente Kappa de 0.84 y 0.82 para los años 2016 y 2022 en imágenes Sentinel-2. En cuanto a los cambios de uso de suelo, la categoría de bosque fue la que tuvo mayor pérdida con 2,246 hectáreas, esto debido a la conversión en pastizales y agricultura, principalmente, es importante destacar que la zona urbana incrementó en 876 hectáreas durante el periodo de análisis. Por último, el modelo predictivo de Cadenas de Markov proyecta que, para el año 2028, se espera una disminución de 1,763 hectáreas de cobertura forestal, así como un crecimiento urbano de 844.14 hectáreas.

Los hallazgos en el estudio ponen de manifiesto la necesidad de modificar los patrones actuales de uso de suelo. Para garantizar lo anterior, es importante implementar políticas públicas y estrategias de gestión territorial para promover la conservación de los ecosistemas, la restauración de áreas afectadas y un uso sostenible de los recursos naturales.

Introducción

Como resultado de la conversión y el aprovechamiento del suelo, los ecosistemas terrestres han experimentado cambios importantes. La expansión agrícola y urbana ha provocado una drástica reducción o incluso la desaparición de los bosques. Las alteraciones en la cobertura y uso del suelo conducen a su deterioro y erosión. Estudiar los procesos que provocan estos cambios es fundamental por su impacto en los ámbitos ambiental, natural, económico y social (Galicia Sarmiento *et al.*, 2007).

La degradación del suelo se refiere al deterioro de su calidad, manifestado por una reducción en su capacidad y potencial para generar bienes y servicios de manera cualitativa y cuantitativa. Dichos procesos pueden afectar la productividad agrícola, el equilibrio de los ecosistemas y la disponibilidad de recursos naturales. La degradación puede estar provocada por diversos factores como la erosión, salinización, compactación y sus consecuencias pueden ser tanto ambientales como económicas y sociales. La degradación se clasifica según su naturaleza y las consecuencias que genera (García Orozco *et al.*, 2012).

Camacho Sanabria (2016), señala que el ser humano ejerce una influencia directa sobre el ambiente al explotar el suelo para la producción de bienes y servicios. Esta intervención se ha convertido en la principal causa de disminución de la diversidad biológica, el deterioro ambiental y la disminución de la capacidad de la tierra para sostener actividades productivas en grandes regiones del mundo.

Información precisa de cambios de uso de suelo pueden facilitar diversas actividades de investigación en temas como inundaciones, sequías, migración y cambio climático a diversas escalas. Los estudios con enfoque en el conocimiento de las dinámicas de cambio de uso del suelo tienen un amplio campo de aplicación, desde poder contar con

información actualizada y fiable tanto para ser parte de la elaboración de políticas públicas, realizar proyecciones a futuro y generar propuestas para el uso sostenible de los recursos, por lo anterior, es importante cartografiar y cuantificar los cambios ocasionados por el ser humano en los ecosistemas naturales (Basheer *et al.*, 2022).

Por lo tanto, el objeto de estudio es aplicar los algoritmos de clasificación y regresión Random Forest (RF) y Support Vector Machine (SVM) para la clasificación de la cartografía de uso de suelo mediante imágenes de satélite de los sensores Landsat 8 y Sentinel-2 en el programa Google Earth Engine (GEE). La finalidad es analizar los cambios en el uso de suelo de la Zona Metropolitana de Tianguistenco y generar un mapa predictivo que refleje estas transformaciones.

Planteamiento del problema

Según la Procuraduría de Protección al Ambiente - PROFEPA (2020), México posee una extensión de 138.8 millones de hectáreas con algún tipo de vegetación forestal. De esa superficie, 65.7 millones de ha corresponden a áreas arboladas que incluyen bosques, selvas, manglares y otras formaciones vegetales; 56.3 millones de ha (40.8%) están cubiertas por matorral xerófilo, mientras que el 11.5% restante, equivalente a 15.8 millones de ha, corresponde a otras áreas forestales.

La Organización de las Naciones Unidas para la Alimentación y la Agricultura - FAO (2023) advierte que las prácticas agrícolas y la deforestación alteran drásticamente los ecosistemas terrestres, causando la degradación de los suelos que conllevan la pérdida de su fertilidad, la erosión y una disminución significativa de la biodiversidad.

De acuerdo con García Orozco *et al.* (2012) los bosques de México y su pérdida van de un tercio a casi la mitad de la superficie total original, dicho fenómeno se ha visto incrementado a causa del crecimiento demográfico. De acuerdo con Global Forest Watch (GFW) citado en López Portillo (2019), México ocupa el puesto número decimoséptimo entre los países con mayor pérdida de cobertura forestal, con un total de 267,731 hectáreas.

Con base en la Protectora de Bosques - PROBOSQUE (2024), el Estado de México posee 1,080,374.90 hectáreas de superficie forestal, las cuales han sido degradadas por

diversas actividades humanas, como la deforestación, sobrepastoreo, prácticas agrícolas inadecuadas, incendios forestales y el evidente crecimiento urbano y necesidad de recursos naturales.

El Gobierno del Estado de México (2016) en su Decreto número 159 aprobó la declaratoria de Zona Metropolitana de Tianguistenco integrada por los municipios de Almoloya del Río, Atizapán, Capulhuac, Texcalyacac, Tianguistenco y Xalatlaco, para la coordinación de los planes, programas, y acciones de estos entre sí o del Estado y sus municipios, con planes federales o de la Ciudad de México; se publicó en la Gaceta del Gobierno.

La conformación de la Zona Metropolitana de Tianguistenco es resultado de procesos de transformación demográfica, económica y social, lo que también ofrece la oportunidad de impulsar el desarrollo de los municipios que la componen. No obstante, plantea un desafío en cuanto a la gestión territorial (COESPO, 2017).

En los últimos años, los municipios que conforman la metrópoli han experimentado cambios significativos, principalmente pérdidas de cobertura forestal ocasionada por la tala ilegal de los bosques, contaminación de los cuerpos de agua que forman parte del APFyF Ciéneegas del Lerma y Río Lerma, así como el decremento de los recursos hídricos, derivado de la sobreexplotación de los acuíferos, sumado a ello y como se mencionó en el párrafo anterior, la demanda requerida por el campo ha hecho que cubiertas, como pastizales y suelo urbano vayan en aumento (Pineda Monroy, 2019).

A nivel internacional y nacional, numerosos estudios han empleado imágenes satelitales y herramientas como QGIS, Terrset y GEE para analizar los cambios en el uso y cobertura del suelo, aplicando técnicas de teledetección avanzadas. Si bien existen un número importante de algoritmos para la clasificación de uso de suelo, derivado de la investigación previa, RF y SVM han demostrado tener una mayor precisión (Talukdar *et al.*, 2020; Ye *et al.*, 2021; Basheer *et al.*, 2022).

Para poder conocer los cambios en el uso de suelo en la Zona Metropolitana de Tianguistenco es fundamental contar con dos insumos principales, la cartografía de uso de suelo de los años 2016 y 2022, obtenida a través de una clasificación previa.

Posteriormente, utilizando el modelo predictivo Cadenas de Markov, es posible proyectar un escenario futuro para la zona de estudio.

Justificación

Los cambios en el uso de la tierra y cobertura del suelo son consecuencia de las actividades humanas en las que influyen factores ambientales, sociales, económicos, políticos e institucionales a nivel global, regional y local, sin embargo, los cambios también pueden ocurrir por procesos naturales, en conjunto, llegan a provocar deterioro ambiental y pérdida de la diversidad biológica. Se requiere la integración de las ciencias naturales y sociales, así como el reconocimiento del papel creciente de los factores globales para enfrentar el desafío (Lambin *et al.*, 2001; Montero Mora y Viales Hurtado, 2015; López Vazquez *et al.*, 2015).

El cambio de uso de suelo provocado por actividades humanas tiene influencia en procesos naturales como la recarga de acuíferos, escurrimiento superficial, deterioro del suelo, entre otros, en ello radica la importancia de conocer la dinámica de los cambios ocurridos durante diferentes periodos (Trucíos Cacicano *et al.*, 2011).

La cobertura de suelo tiene un papel importante para la planificación, supervisión y uso de los recursos naturales a partir del incremento de la demanda del hombre en la actualidad. Aunado a ello, los recursos naturales que brinda el área de estudio se han visto reducidos.

Como practica ancestral, en la región se lleva a cabo el trueque de leña, de esta forma, habitantes obtienen otros productos o sustento económico, sin embargo, la actividad no se encuentra regulada y tiene un impacto en el ambiente y suelo (Morales Sánchez, 2016; Caballero Martínez, 2017).

Los Programas de Ordenamiento Ecológico local de los municipios de Atizapán, Almoloya del Río, Capulhuac, Texcalyacac y Tianguistenco, expedidos entre 2015 y 2016 cuentan con Unidades de Gestión Ambiental (UGA), son unidades mínimas del territorio a las que se les asignan determinados lineamientos y estrategias ecológicas (SEMARNAT, 2022).

La información más reciente referente a uso de suelo y vegetación es generada por el Instituto Nacional de Estadística y Geografía - INEGI, sin embargo, tiene errores de topología y un sesgo en cuanto al nivel de clasificación. En el presente estudio se aplicaron dos algoritmos de clasificación mismos que han demostrado ser eficaces en la clasificación de imágenes de satélite, aunado a ello, el uso de imágenes con diferente resolución espacial para resultados más detallados, a su vez, la comparación de dos periodos de tiempo para conocer la afectación en la ZMT a partir de su denominación como metrópoli.

La información generada a partir de la clasificación de imágenes satelitales, el análisis de cambios de uso del suelo y mapa resultado del modelo probabilístico (Cadenas de Markov) será fundamental para la toma de decisiones en temas ambientales, económicos y sociales.

Antecedentes

Un componente clave del proceso de investigación es entender el estado del arte en cuanto al cambio de uso de suelo, empleo de imágenes satelitales y modelos probabilísticos o tendenciales, para la generación de cartografía temática. Además, a partir de esta información, es posible generar nuevos análisis, como la identificación de pérdidas y ganancias en el uso y cobertura del suelo en periodos específicos, así como realizar proyecciones futuras, entre otras aplicaciones. A continuación, se ofrece un resumen de estudios a nivel nacional e internacional, enfocados en el análisis del cambio de uso de suelo mediante geotecnologías innovadoras y de vanguardia.

Contexto Internacional

A nivel internacional, se han realizado varios estudios sobre el cambio en el uso de suelo y la cobertura terrestre. Uno de estos se enfocó en la región de Xilingol, China, abarcando el periodo de verano de 2000 a 2020. Utilizando imágenes satelitales de Landsat 5 TM, 7 ETM+ y 8 OLI/TIRS, el estudio analizó cambios en tierras de cultivo, bosques, praderas, cuerpos de agua, áreas edificadas y tierras desérticas. El objetivo fue aplicar el algoritmo RF para generar mapas anuales de LULC y examinar los cambios dinámicos y espaciales de la región. Como métodos de validación, se emplearon análisis de

componentes principales y regresión lineal, utilizando la plataforma GEE. Los resultados mostraron que, en los últimos 20 años, la condición ecológica en Xilingol ha mejorado significativamente, especialmente en las áreas de bosques y pastizales (Ye *et al.*, 2021).

China ha liderado investigaciones sobre los cambios espaciotemporales en el uso de la tierra utilizando teledetección. Para esta investigación, se tomó como área de estudio Linyi, provincia de Shandong, empleando imágenes satelitales Landsat 5 TM, 7 ETM+ y 8 OLI, desde 1990 hasta 2020 en intervalos de 10 años. Con variables independientes como el DEM, pendiente y distancia a carreteras, se empleó la metodología CA-ANN del plugin MOLUSCE en QGIS para simular cambios futuros en el LULC. Los resultados muestran que Linyi experimentó un aumento considerable en superficies impermeables, que pasaron del 10.48% al 26.91%, mientras que la cobertura forestal disminuyó de 12.60% a 8.74%. Las predicciones para 2030-2050 confirman esta tendencia de crecimiento en áreas urbanizadas, a costa de la pérdida de bosques y espacios verdes (Muhammad *et al.*, 2022).

Rahnama (2021) llevó a cabo una investigación utilizando imágenes del sensor Sentinel-2A para analizar diferentes clases de suelo. El estudio se apoyó en el software ArcGIS y Terrset, implementando autómatas celulares y el modelo Cadenas de Markov. El objetivo principal fue simular los cambios LULC durante el periodo 2016-2020 y proyectar estos cambios hacia el año 2030. Para validar los resultados, se emplearon la tabulación cruzada y el índice Kappa, lo que permitió confirmar la precisión de los cambios detectados en las distintas clases de suelo estudiadas.

Se realizó un análisis exhaustivo sobre los cambios en el uso y cobertura del suelo (LULC) en la región del lago Urmia, utilizando imágenes satelitales de los sensores Landsat TM, ETM y TIRS para los años 2000, 2005, 2010, 2015 y 2019. El estudio se centró en diversas clases de suelo, como cuerpos de agua, áreas cultivadas, tierras secas y en barbecho, pastizales, asentamientos, tierras desoladas, orquíideas, salidas de roca y tierras saladas. Para llevar a cabo el análisis, se utilizó la plataforma GEE y se emplearon clasificadores avanzados como SVM, RF y CART. Además, se aplicó la teoría de Dempster-Shafer para realizar un análisis de incertidumbre. Entre los clasificadores utilizados, SVM mostró la mayor precisión. Los resultados revelaron una expansión

significativa de las tierras de cultivo y huertos de alta demanda de agua, lo que ha agravado la sequía en el lago Urmia, representando un problema crítico para el ecosistema de la región (Feizizadeh *et al.*, 2021).

En la ciudad de Timbú, capital de Bután, en Asia, se llevó a cabo un estudio para detectar y predecir cambios en el uso y cobertura de suelo utilizando Autómatas Celulares y Cadenas de Markov. Se emplearon imágenes de teledetección de los sensores Landsat 7 ETM+ y 8 OLI correspondientes a 2002 y 2018, las cuales fueron previamente preprocesadas para realizar una clasificación supervisada y generar mapas LULC. Posteriormente, se aplicó un análisis de cambios y una predicción hacia 2050 en el software Terrset. Los resultados revelaron un incremento del 12.77% en áreas construidas entre 2002 y 2018, mientras que la cobertura forestal disminuyó drásticamente en un 15.25%. La proyección a 2050 indica que la superficie construida consumirá gran parte del área de la ciudad, con una marcada reducción de bosques. Estos cambios se asocian con la migración rural-urbana y el rápido desarrollo económico de la región (Wang *et al.*, 2021).

En América del Sur, un estudio realizado en el Oeste de la Bahía, Brasil, utilizó imágenes satelitales de los sensores Landsat 5 TM, 7 ETM+ y 8 OLI/TIRS para analizar los cambios históricos en el uso de suelo entre 1990 y 2020. Este estudio se enfocó en diversas clases de suelo como la formación de bosques, sabanas, pastizales, mosaicos de cultivos y pasto, cultivos de secano y regadío, cuerpos de agua y áreas urbanas o edificios agrícolas. Para clasificar las imágenes se empleó el algoritmo RF e índice Kappa para validar los resultados, junto con un análisis de decisión multicriterio espacial que evaluó la idoneidad de la tierra. Los resultados mostraron un aumento significativo en el área agrícola, que creció en 3.14 millones de hectáreas durante el periodo estudiado, mientras que el área irrigada incrementó en 193,480 ha, lo que destaca la transformación del paisaje agrícola en la región (Pimenta *et al.*, 2021).

Contexto Nacional

En el contexto nacional, se realizó un estudio en la zona central montañosa de Guerrero utilizando imágenes satelitales de los sensores Landsat 5 TM y 8 OLI para monitorear los cambios en la cobertura del suelo entre 2011 y 2016. El análisis abarcó diversas

coberturas de suelo, tales como bosque de encino, bosque de pino, bosque mixto, bosque caducifolio, áreas agrícolas, suelo desnudo, cuerpos de agua y zonas urbanas o industriales. Para la clasificación, se empleó una metodología supervisada de máxima verosimilitud, complementada con el índice Tasseled Cap (TC), el índice de brillo-verdor-humedad y el análisis de Componentes Principales (PC). Como método de validación, se utilizó la matriz de confusión e índice Kappa, obteniendo una concordancia de 91.18 para 2011 y 87.53 para 2016, lo que se considera un nivel aceptable de precisión en la clasificación de los cambios detectados (Vázquez Jiménez *et al.*, 2021).

En un estudio realizado en el municipio de San Fernando, Tamaulipas, se evaluaron los cambios en el uso de la tierra utilizando imágenes satelitales de los sensores Landsat 5 TM y 8 OLI, correspondientes a los años 1987, 1997, 2007 y 2017. El análisis incluyó clases de suelo como áreas agrícolas, cuerpos de agua, zonas sin vegetación aparente y vegetación, además de nubes y sombras como categorías auxiliares. El método empleado fue el análisis espacial de combinación, junto con la segmentación de imágenes y el algoritmo SVM. Los resultados mostraron una clara tendencia hacia la reducción de la cobertura vegetal y de las áreas sin vegetación aparente, junto con un incremento significativo en el uso agrícola, evidenciando un proceso de transformación en el paisaje rural de la región en las últimas décadas (Salinas-Castillo *et al.*, 2020).

En la microcuenca de la Mixteca Oaxaqueña, se llevó a cabo una evaluación detallada de los cambios en el uso del suelo y la vegetación, utilizando imágenes satelitales de Landsat 8 OLI y Sentinel-2A correspondientes a los años 1995, 2010, 2013 y 2016. El estudio abarcó diversas clases de suelo, como áreas forestales, pastizales, tierras agrícolas, zonas sin vegetación, asentamientos, caminos y cuerpos de agua. Para el análisis, se aplicaron métodos de clasificación supervisada y no supervisada, respaldados por una validación en campo que aseguró la precisión de los resultados. Los hallazgos indicaron un incremento del 6.6% en la cobertura forestal entre 1995 y 2016, mientras que la superficie agrícola se redujo en un 8.4%. Estos resultados reflejan importantes transformaciones en el uso de la tierra en la región, proporcionando evidencia valiosa sobre la dinámica de los cambios en la cobertura vegetal a nivel nacional (Sandoval-García *et al.*, 2021).

A nivel local, en el Estado de México, se realizó en 2017 un estudio para analizar los cambios en el uso del suelo y la cobertura vegetal entre los años 2000 y 2015 en Progreso Hidalgo, en el municipio de Villa Guerrero. Para ello, se emplearon ortofotos e imágenes satelitales de Google Earth con el fin de crear mapas temáticos que incluyeron clases de uso del suelo como agricultura de temporal o riego, producción agrícola en invernaderos, áreas boscosas, cuerpos de agua y asentamientos humanos. La investigación abarcó dos periodos clave, 2000-2007 y 2007-2015, y se basó en la interpretación visual de insumos cartográficos para validar los mapas de uso de suelo determinando así el grado de confiabilidad temática. Los resultados mostraron que, entre 2000 y 2007, la agricultura de temporal y riego, junto con las áreas boscosas, disminuyeron en 23.04 y 16.76 hectáreas, respectivamente, mientras que la producción en invernaderos aumentó en 5.19 ha. En el periodo 2007-2015, la superficie agrícola perdió 18.71 ha, de las cuales 14.39 ha fueron convertidas en invernaderos, evidenciando una transformación significativa en el uso del suelo en la región (Camacho-Sanabria *et al.*, 2017).

En el oriente de la Zona Metropolitana de Toluca, que abarca los municipios de Lerma, Metepec, Ocoyoacac y San Mateo Atenco, se realizó una evaluación de los cambios en el uso del suelo y las dinámicas demográficas entre 1984 y 2017, utilizando ortofotografías e imágenes del mapa base de ArcGIS. La investigación, realizada por Adame Martínez *et al.* (2020), aplicó la técnica de fotointerpretación visual complementada con trabajo de campo para validar los resultados. En 1984, los principales usos del suelo eran las áreas agrícolas y los bosques. Sin embargo, para 2017, aunque la agricultura seguía siendo predominante, experimentó una disminución del 35.74%, al igual que los cuerpos de agua, que se redujeron un 31.77%. En contraste, las superficies artificiales crecieron un 235.84%, lo que evidencia una intensificación del proceso de urbanización y el desarrollo de infraestructuras en la región.

En su estudio, Jiménez López (2019) simuló el crecimiento del Área Metropolitana de Toluca empleando Cadenas de Markov y análisis de imágenes satelitales de los años 2003 y 2017 a partir del sensor Landsat. Estas imágenes fueron gestionadas mediante el proyecto CRISTALLER® del Colegio Mexiquense A.C. El objetivo principal fue

desarrollar una metodología capaz de proyectar la expansión urbana hacia 2031, aplicando una variación espacial de las Cadenas de Markov. Para evaluar la precisión del modelo, se utilizaron cuatro métricas que comparan los resultados de la simulación con la realidad, garantizando su fiabilidad y exactitud en las proyecciones. El estudio generó un mapa detallado que muestra la probabilidad de cambios en el uso del suelo a nivel de píxel, permitiendo identificar áreas ocupadas y no ocupadas en la ciudad de Toluca y se puede decir que la velocidad de crecimiento será acelerada al considerarse una ciudad millonaria con relación a sus habitantes.

La revisión de antecedentes nacionales e internacionales evidencia la amplia aplicación de las geotecnologías para el análisis de cambio de uso de suelo y los modelos predictivos, mismos que destacan la capacidad de estas herramientas para el monitoreo de las transformaciones del espacio geográfico. Sin embargo, se abre un área de oportunidad para el desarrollo de estudios a nivel local para desarrollar estudios y analizar los cambios de uso de suelo en la Zona Metropolitana de Tianguistenco desde su designación como metrópoli. A partir de estas observaciones, el presente estudio establece los siguientes objetivos.

Objetivos

Objetivo General

- Analizar el proceso de cambios de uso de suelo en la Zona Metropolitana de Tianguistenco mediante el uso de métodos y técnicas de teledetección para identificar las perturbaciones en las coberturas del suelo para el periodo 2016-2022.

Objetivos Específicos

- Evaluar la eficiencia de los algoritmos Random Forest y Support Vector Machine a través del uso de herramientas geotecnológicas para la obtención de las capas de uso de suelo.
- Estimar el cambio de uso de suelo mediante la aplicación de la matriz de tabulación cruzada para conocer ganancias y pérdidas en cada una de las coberturas.
- Implementar un modelo predictivo con el apoyo de un software SIG para identificar el comportamiento del proceso de cambio de uso de suelo.

CAPÍTULO 1.

MARCO TEÓRICO

A continuación, se presenta la base teórica y conceptual de la investigación, abarcando el contexto general de la geografía, con énfasis en el estudio del espacio geográfico y el análisis espacial. Se incluyen los SIG, que facilitan la aplicación de métodos y técnicas para desarrollar proyectos integrales y la teledetección, que permite obtener cartografía de uso de suelo mediante imágenes satelitales. Además, se exploran conceptos de inteligencia artificial y aprendizaje automático, junto con el uso de herramientas en la nube como GEE. Finalmente, se aborda el cambio de uso del suelo y los modelos predictivos.

1.1 Contexto General de la Geografía

A lo largo de la historia de la geografía, ciencia que tiene alrededor de 2,600 años de antigüedad, el objeto de estudio ha sido la tierra. No obstante, los geógrafos han estudiado diferentes cuestiones referentes a la tierra. Sin embargo, los geógrafos han abordado diferentes aspectos relacionados con ella. En 1884, el geógrafo y geólogo alemán Richthofen estableció que el objeto de estudio de la geografía no era el planeta en su totalidad, sino su superficie, conocida como *Endoberfläche*, es decir, la capa externa que recubre el planeta y donde interactúan los elementos bióticos, abióticos y humanos (Higueras Arnal, 2003).

Algunos autores definen a la Geografía de la siguiente manera:

Según Hartshorne, citado por Higueras Arnal (2003), el objetivo de la geografía es proporcionar descripciones e interpretaciones organizadas y precisas acerca de la variabilidad de la superficie terrestre.

La geografía se encarga de estudiar el espacio geográfico y la interacción entre el ser humano y la naturaleza, además de analizar los hechos, fenómenos y sus causas. Según López Trigal (2015), la geografía es la ciencia que se ocupa del territorio, el cual es concebido como una construcción social producto de las interacciones e interdependencias entre la sociedad y la naturaleza; también estudia, describe y representa los diferentes paisajes y espacios terrestres.

La geografía puede abordarse desde tres enfoques: el ecológico, que se centra en la relación hombre-medio; el corológico, que estudia la diferenciación espacial; y el

sistémico, que examina las leyes que rigen los patrones de distribución en el espacio (Buzai y Baxendale, 2010; Buzai, 2015).

1.1.1 Conceptos del Análisis Espacial

El geógrafo y climatólogo francés Emmanuel de Martonne [1873-1988] identificó los principios geográficos como la localización, conexión, extensión, complejidad, dinamismo, y globalidad territorial. Actualmente, estos conceptos han evolucionado hacia lo que hoy conocemos como análisis espacial, incorporando términos como Localización, Distribución Espacial, Interacción Espacial y Evolución Espacial (Buzai y Baxendale, 2013; Buzai, 2015).

La **localización** implica que todas las entidades y sus características tienen una posición definida dentro del espacio geográfico. Esta se puede entender desde dos enfoques complementarios: el espacio absoluto, que corresponde a un lugar fijo y específico determinado por la topografía local, y el espacio relativo, que se refiere a una posición dinámica y cambiante en relación con otros lugares.

La **distribución espacial** se refiere a cómo se organiza un conjunto de entidades del mismo tipo a lo largo del espacio geográfico. Esta distribución puede ser visualizada a través de representaciones vectoriales, como puntos, líneas o polígonos, o a través de representaciones ráster, que muestran localizaciones dispersas o zonas contiguas con clases similares.

La **asociación espacial** se centra en analizar las similitudes que surgen al comparar diferentes distribuciones espaciales. Esto se puede lograr con un análisis visual, que se realiza tras la superposición de ambas distribuciones en una representación cartográfica.

La **interacción espacial** analiza la estructura de un espacio en el que la localización, distancia y vínculos juegan un papel clave en la definición de los espacios funcionales. Los estudios de interacción espacial se enfocan en cuantificar los distintos tipos de relaciones horizontales que existen entre las entidades geográficas ubicadas en ese espacio.

La **evolución espacial** incluye la dimensión temporal, reconociendo los constantes cambios de un estado a otro. Los estudios geográficos se enfocan en el análisis del presente, pero toman en cuenta la relevancia del tiempo para lograr una comprensión integral del espacio geográfico.

1.2 Sistemas de Información Geográfica (SIG)

La Geografía se apoya de otras ciencias, técnicas y métodos para hacer investigación, como son los Sistemas de Información Geográfica.

Los Sistemas de Información Geográfica (SIG), según SIGSA (2022), proporcionan un marco integral para organizar, visualizar, comprender, recopilar, gestionar, analizar e interpretar datos de diversos tipos. Tienen la capacidad de analizar la ubicación espacial y de organizar la información en capas utilizando mapas y representaciones en 3D. Los SIG permiten descubrir información más detallada sobre los datos, como patrones, relaciones y situaciones, lo que facilita la toma de decisiones por parte del usuario.

Olaya (2014) señala que un SIG es una herramienta que facilita el análisis, presentación e interpretación de hechos relacionados con la superficie terrestre. Según Olaya, un SIG es un sistema de información compuesto por software y hardware, creado para la adquisición, mantenimiento y utilización de datos cartográficos, los cuales están relacionados por coordenadas espaciales o geográficas. INEGI (2014) define a los SIG son un conjunto de herramientas creadas para obtener, almacenar, recuperar y mostrar datos espaciales del entorno real.

1.2.1 Componentes de un SIG

De acuerdo con Olaya (2014), los SIG se componen de cinco elementos principales. El primero es el **dato**, que actúa como el insumo esencial y otorga sentido a los demás componentes del sistema, con características específicas como forma y propiedades. El segundo es el conjunto de **métodos**, que abarca las formulaciones y técnicas aplicadas a los datos. El **software** es la aplicación informática responsable de procesar los datos y ejecutar los métodos, permitiendo la manipulación de la información. El **hardware** hace referencia al equipo de cómputo necesario para operar el software SIG, cuya capacidad depende del tamaño de la comunidad de usuarios y del tipo de datos que se procesan.

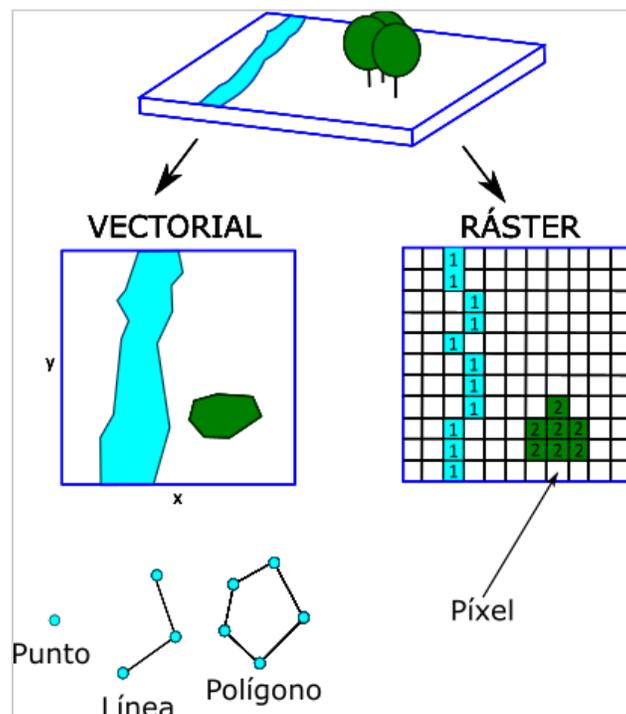
Finalmente, las **personas** son quienes diseñan, administran y utilizan el software, con cuatro tipos principales de usuarios: los técnicos informáticos que son especialistas en SIG pero no en su aplicación; los profesionales ocasionales, usuarios finales con conocimientos limitados en SIG pero expertos en su campo de aplicación; el público, que son los clientes del servicio SIG; y los especialistas en la aplicación, que dominan tanto en SIG como su campo de aplicación, con un alto nivel de conocimiento en la temática.

1.2.2 Modelos de Representación de Información Geográfica

Existen dos modelos de representación de información geográfica (Figura 1.1):

- **Modelo Ráster:** Se caracteriza por dividirse de forma sistemática en una serie de unidades mínimas, el modelo ráster se forma a partir de una matriz de celdas o píxeles dispuestos en un esquema de filas y columnas, en su interior recogen información que las describe (Olaya, 2014 y ESRI, 2016).
- **Modelo Vector:** Su principal característica es que se representa por entidades geométricas (puntos, líneas y polígonos).

Figura 1.1 Modelos de representación geográfica



Fuente: My GIS Notebook (2019)

1.3 Teledetección

Los Sistemas de Información Geográfica y la teledetección están estrechamente vinculados, ya que los SIG incluyen herramientas para gestionar, procesar y analizar los datos obtenidos por medio de la teledetección.

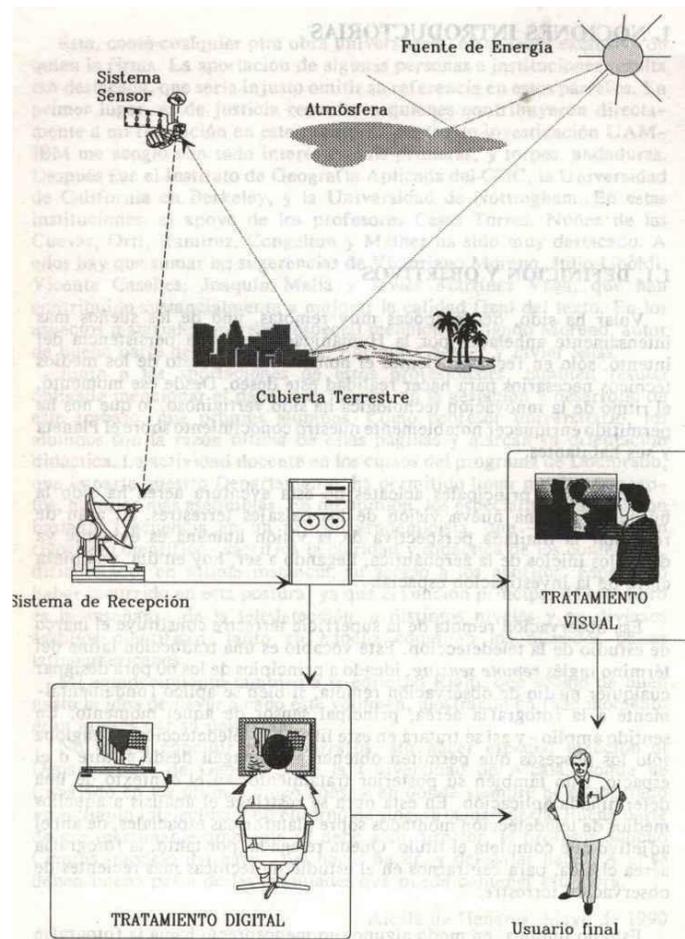
Se define a la teledetección “como aquella técnica que permite adquirir imágenes de la superficie terrestre desde sensores instalados en plataformas espaciales” (Chuvieco, 1995, p. 27). De acuerdo con Sacristán Romero (2007), la teledetección estudia las variaciones en las ondas electromagnéticas en términos espectrales, espaciales y temporales, destacando las correlaciones que existen entre estas ondas y las propiedades de los distintos materiales en la superficie terrestre.

1.3.1 Elementos de la Teledetección

Un sistema de teledetección espacial debe incluir los siguientes elementos (Figura 1.2):

1. **Fuente de energía:** Puede ser de origen natural o artificial. Este elemento se refiere al origen del flujo energético detectado por el sensor, que puede provenir de una fuente externa al sensor en el caso de la teledetección pasiva, o de un haz de energía emitido directamente por el sensor en la teledetección activa.
2. **Cubierta terrestre:** Está formada por diferentes tipos de superficies como vegetación, suelo, agua o áreas urbanas, que interactúan con la energía recibida de la fuente, ya sea reflejándola o emitiéndola según sus propiedades físicas.
3. **Sistema sensor:** Incluye tanto el sensor como la plataforma que lo soporta. Su función principal es captar la energía proveniente de la superficie terrestre, codificarla y almacenarla o transmitirla al sistema de recepción.
4. **Sistema de recepción-comercialización:** Es el lugar donde se recibe la información transmitida por la plataforma para su posterior procesamiento.
5. **Interprete:** Es responsable de analizar los datos captados, transformándolos en una clave temática que facilite la evaluación de un problema específico de estudio.
6. **Usuario final:** Es quien revisa el producto proporcionado por el intérprete y toma decisiones basadas en las conclusiones que derivan de la información analizada.

Figura 1.2. Elementos de la teledetección



Fuente: Chuvieco (1995)

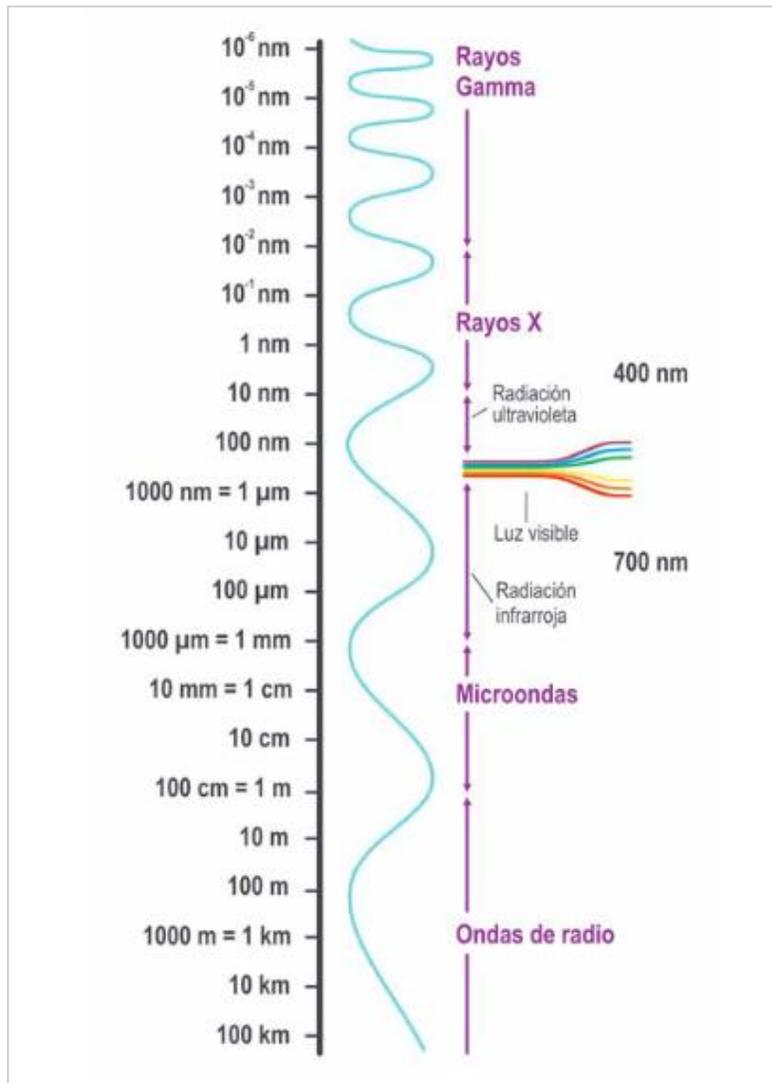
1.3.2 Espectro Electromagnético

Labrador García *et al.* (2012) sugiere que los ojos humanos pueden considerarse como sensores remotos, ya que son capaces de captar la luz reflejada por los objetos que nos rodean. Así mismo, Tempfli *et al.* (2009) explican que el rango total de longitudes de onda de la radiación electromagnética se conoce como espectro electromagnético, el cual es continuo y carece de límites definidos entre las distintas clases. Las diferentes partes del espectro incluyen los rayos gamma, los rayos X, la radiación ultravioleta, visible e infrarroja, así como las microondas y las ondas de radio.

Las regiones del espectro que más nos interesan son las visibles y las del infrarrojo. La región visible es una porción muy pequeña del espectro electromagnético, que abarca entre 0.4 μm y 0.7 μm . Dentro de esta, el azul se extiende desde 0.4 μm hasta 0.5 μm ,

el verde desde 0.5 μm hasta 0.6 μm , y el rojo desde 0.6 μm hasta 0.7 μm . Más allá del rojo se encuentra la región infrarroja (IR), que se divide en tres categorías: infrarrojo cercano (NIR) entre 0.7 y 1.3 μm , infrarrojo medio (SWIR) entre 1.3 y 3 μm , e infrarrojo térmico (TIR) entre 3 y 100 μm (Labrador García *et al.*, 2012). (Figura 1.3).

Figura 1.3. Espectro Electromagnético



Fuente: Labrador García *et al.* (2012)

1.3.3 Tipos de Resolución

De acuerdo con Chuvieco (1995) y Olaya (2014), entre las propiedades de un sistema de teledetección se encuentra la resolución, misma que depende de un efecto combinado de distintos componentes. Se distinguen cuatro resoluciones:

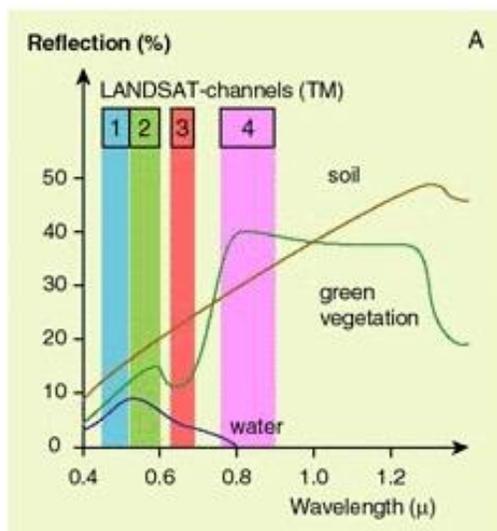
1. **Resolución Espacial:** Se refiere al tamaño del objeto más pequeño que puede identificarse en una imagen. Esta se mide en unidades de longitud, como milímetros en una foto o metros en el terreno, y está influenciada por la longitud focal de la cámara y su altura respecto a la superficie terrestre. La resolución espacial es fundamental en la interpretación de imágenes, ya que determina el nivel de detalle que estas proporcionan. Además, está estrechamente vinculada con la escala de trabajo y la fiabilidad en la interpretación de los datos.
2. **Resolución espectral:** se refiere a la cantidad y el ancho de las bandas espectrales que un sensor puede discriminar. Ambos factores determinan la resolución espectral del sensor. Cuantas más bandas proporcione un sensor, mayor será su capacidad para caracterizar espectralmente diferentes tipos de cubiertas. Por ejemplo, las imágenes en blanco y negro contienen una única banda, mientras que las imágenes a color incluyen tres bandas (rojo, verde y azul). Algunos sensores cuentan con bandas adicionales, como el infrarrojo, y a este tipo de imágenes se les denomina multiespectrales.
3. **Resolución radiométrica:** se refiere a la capacidad del sensor para detectar variaciones en la radiancia espectral que recibe, es decir, al número máximo de niveles digitales que puede distinguir en una imagen. La mayoría de los sistemas proporcionan 256 niveles por píxel (de 0 a 255). Cuanto mayor sea la precisión radiométrica, más detallada será la interpretación de la imagen.
4. **Resolución temporal:** se refiere a la frecuencia con la que el sensor captura imágenes de la misma área de la superficie terrestre. Esta frecuencia depende de las características orbitales de la plataforma, como su altura, velocidad e inclinación, así como del diseño del sensor.

1.3.4 Firmas espectrales

Es una representación gráfica, la cual muestra las características de la interacción entre la radiación electromagnética y la estructura de la planta. Cada tipo de cobertura terrestre

ya sea bosque, suelo cultivado, agua, nieve, áreas urbanas, entre otras, se distingue por la energía que reflejan o emiten en distintas longitudes de onda (Figura 1.4). Estos espectros característicos, que permiten identificar el tipo de cubierta observada, constituyen su firma espectral (Alonso *et al.*, 1999; Araque Torres y Jiménez, 2009).

Figura 1.4. Firmas espectrales



Fuente: European Space Agency [ESA] (2024)

1.3.5 Sensores

En la actualidad, el número de productos derivados de la teledetección ha aumentado considerablemente. Según Labrador García *et al.* (2012), para ese entonces ya operaban aproximadamente 120 satélites de teledetección. Hoy en día, esa cifra seguramente ha aumentado, junto con mejoras significativas en la calidad de los datos generados sobre la superficie terrestre.

Entre los sensores más destacados, tanto por su capacidad y accesibilidad, se encuentran Landsat y Sentinel, los cuales se describen a continuación:

1.3.6 Landsat 8 OLI/TIRS

Landsat es un programa de teledetección que ha puesto en órbita nueve satélites desde 1972 hasta 2022. Cuenta con un volumen de datos muy extenso y por ese motivo, Landsat es uno de los sensores más utilizado para elaborar estudios y monitorear los recursos de la tierra.

El satélite Landsat 8 fue puesto en órbita el 11 de febrero de 2013, está equipado con un radiómetro multicanal conocido como OLI (Operational Land Imager) y un radiómetro infrarrojo de dos canales TIRS (Thermal Infrared Sensor). El sensor OLI proporciona imágenes con una resolución de hasta 15 metros y permite realizar observaciones sobre la calidad del agua en lagos y zonas costeras, gracias a la incorporación de dos rangos de longitud de onda. Por su parte, TIRS permite realizar mediciones utilizando dos combinaciones de ocho bandas en el espectro infrarrojo. Al igual que su predecesor, Landsat 8 tiene un ciclo de repetición de 16 días (USGS, s.f.; EOS Data Analytics, 2023). Las características de las bandas del sensor Landsat 8 OLI/TIRS se presentan en la Tabla 1.1.

Tabla 1.1. Características de las bandas de Landsat 8 OLI/TIRS

SENSOR	Número de Banda	Nombre de la Banda	Longitud de Onda (μm)	Resolución (metros)	Aplicación
OLI	1	Aerosol costero	0.43 – 0.45	30	Estudios costeros y de aerosoles
	2	Azul	0.450 – 0.51	30	Distingue el suelo de la vegetación y la vegetación caducifolia de la vegetación de coníferas
	3	Verde	0.53 – 0.59	30	Destaca los picos de máxima vegetación, que son útiles para evaluar el vigor de las plantas
	4	Rojo	0.64 – 0.67	30	Distingue las laderas de vegetación
	5	Infrarrojo Cercano (NIR)	0.85 – 0.88	30	Destaca el contenido de biomasa y las costas
	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
	7	Infrarrojo de Onda Corta (SWIR 2)	2.11 – 2.29	30	Mejora de la lectura de la humedad del suelo y la vegetación y la

SENSOR	Número de Banda	Nombre de la Banda	Longitud de Onda (μm)	Resolución (metros)	Aplicación
					penetración a través de nubes finas
	8	Pancromática	0.50 – 0.68	15	
	9	Cirrus	.36 – 1.38	30	Detección mejorada de la contaminación por nubes
TIRS	10	TIRS 1	10.6 – 11.19	100	Mapeo térmico y humedad estimada del suelo
	11	TIRS 2	11.5 – 12.51	100	Mapeo térmico y humedad estimada del suelo

Elaboración propia con base en (EOS Data Analytics, 2023)

1.3.7 Sentinel-2A

El 23 de junio de 2015 fue lanzado el satélite Sentinel-2A, como parte del programa Copernicus de la Agencia Espacial Europea (ESA). Este satélite fue creado con el propósito de generar un volumen considerable de datos e imágenes. Está equipado con un sensor multispectral optoelectrónico que permite realizar capturas con resoluciones que varían entre 10 y 60 metros, utilizando 13 bandas espectrales que ofrecen una visión detallada de la superficie terrestre y la vegetación. Sentinel-2 forma parte de una misión compuesta por dos satélites idénticos (el Sentinel-2B fue lanzado en el primer trimestre de 2017), lo que permite una cobertura completa de la superficie terrestre cada cinco días. Las características de las bandas del sensor Sentinel-2A/2B se describen en la Tabla 1.2 (ESA, 2015).

Tabla 1.2. Características de las bandas del Sensor Sentinel 2A/ 2B

Nombre de la banda	Sensor	Número de la banda	Sentinel-2A		Sentinel-2B		Resolución (metros)
			Longitud de Onda Central (µm)	Ancho de Banda (µm)	Longitud de Onda Central (µm)	Ancho de Banda (µm)	
Coastal Aerosol	Multiespectral (MSI)	1	443.9	20	442	20	60
Blue		2	496.6	65	492.1	65	10
Green		3	560.0	35	559	35	10
Red		4	664.5	30	665	30	10
Vegetation Red Edge		5	703.9	15	703.8	15	20
Vegetation Red Edge		6	740.2	15	739.1	15	20
Vegetation Red Edge		7	782.5	20	779.7	20	20
NIR		8	835.1	115	833	115	10
Narrow NIR		8A	864.8	20	864	20	20
Water Vapor		9	945.0	20	943.2	20	60
SWIR – Cirrus		10	1373.5	30	1376.9	30	60
SWIR		11	1613.7	90	1610.4	90	20
SWIR	12	2202.4	180	2185.7	180	20	

Fuente: Elaboración propia con base en EOS Data Analytics (2023)

1.4 Inteligencia artificial

Estupiñan Ricardo *et al.* (2021) define la Inteligencia Artificial (IA) como una rama de la informática dedicada a crear máquinas y sistemas que puedan realizar tareas que habitualmente requieren la intervención de la inteligencia humana. Estas actividades abarcan, entre otras, la capacidad para resolver problemas, aprender, tomar decisiones

y comprender el lenguaje. La IA no se limita a crear sistemas que simplemente ejecuten comandos preestablecidos, sino que busca dotar a las máquinas de la capacidad de aprender y adaptarse a partir de la experiencia.

La IA se considera una ciencia interdisciplinaria que combina la informática, matemáticas, lógica, neurociencia, lingüística, psicología, entre otras disciplinas. De acuerdo con Sarmiento y Morales Medina (2021), el objeto de la IA es establecer las bases para el desarrollo de tecnologías que permitan a las máquinas operar de manera independiente, tomar decisiones y actuar sin intervención humana.

1.5 Aprendizaje automático

El aprendizaje es la capacidad de cambiar según estímulos externos y recordando aquellas experiencias anteriores. El aprendizaje automático (Machine Learning - ML) es un subapartado de la inteligencia artificial centrado en desarrollar sistemas que aprenden y tienen la capacidad de mejorar el rendimiento basado en los datos que procesan (Oracle, 2023).

De acuerdo con El Naqa y Murphy (2015), el aprendizaje automático es un campo en continuo desarrollo, centrado en algoritmos computacionales que buscan competir con la inteligencia humana al aprender del entorno. Este enfoque se aplica en áreas como la inteligencia artificial, la probabilidad y estadística, la informática y la teoría de la información, entre otras.

1.6 Google Earth Engine

La plataforma de computación en la nube (GEE) fue creada para almacenar y procesar grandes volúmenes de datos (a nivel de petabytes), facilitando el análisis y la toma de decisiones. Actualmente, GEE ofrece datos de diversos satélites, junto con conjuntos de datos vectoriales basados en SIG, información demográfica, meteorológica, modelos digitales de elevación y capas de datos climáticos. Además, es de fácil acceso y proporciona un entorno conveniente para el desarrollo interactivo de datos y algoritmos (Kumar y Mutanga, 2019).

GEE permite al usuario agregar y seleccionar sus propios datos y colecciones mientras utiliza los recursos de la nube de Google para realizar procesamientos personalizados. Esto resulta bastante benéfico para la comunidad investigadora ya que permite extraer una basta cantidad de datos para la detección de cambios de uso de suelo, mapear tendencias, cuantificar recursos, entre otras. Los softwares enfocados al procesamiento de información geográfica resultan de difícil acceso por los costos, además, al contener la información en la nube, no se requiere hardware con alto poder de procesamiento.

Los componentes principales de GEE se dividen en cuatro, el primero es la infraestructura de Google, la cual permite al usuario disponer de sus servidores, de esta manera se puede agilizar la velocidad de procesamiento de información, el segundo elemento es el conjunto de datos o datasets y hace referencia al contenido de imágenes de los sensores Landsat, Sentinel, MODIS y otros con las que cuenta GEE, los cuales se van actualizando diariamente (Perilla y Mas, 2020).

Uno de los componentes clave y tercer elemento de GEE es el Interfaz de Programación de Aplicaciones (API), permite la comunicación entre el usuario y los servidores de Google. El objetivo de la API de GEE es facilitar la creación de programas, la plataforma cuenta con un conjunto de funciones, métodos y algoritmos preprogramados que pueden ser invocados a través de líneas de código y crear los propios de acuerdo con las necesidades (Solórzano Villegas y Perilla Suárez, 2022).

GEE proporciona dos API, una en JavaScript y otra en Python. La API de JavaScript se puede utilizar a través de un navegador web y dispone de una extensa documentación y soporte para realizar diversos procedimientos. Por su parte, python debe trabajar desde su propia consola y realizar procedimientos más complejos, sin embargo, una desventaja es que requiere la instalación de varias bibliotecas en la computadora para trabajar sobre GEE y no se encuentra mucha documentación y ayuda disponible (Perilla y Mas, 2020; Solórzano Villegas y Perilla Suárez, 2022).

De acuerdo con Perilla y Mas (2020), el cuarto y último componente es el Code Editor, un entorno de desarrollo integrado en línea que reúne todos los elementos en un solo lugar. En este espacio, el usuario puede mediante scripts, acceder a los datos, procesar

y visualizar de forma virtual utilizando los servidores de Google, lo que permite obtener los resultados directamente en la nube.

1.6.1 Índice de Vegetación de Diferencia Normalizada (NDVI)

El NDVI se ha consolidado como uno de los índices más populares en el ámbito de la teledetección para el análisis de la vegetación. Esta técnica se basa en el uso de imágenes de percepción remota para evaluar la salud y vigor de las plantas. Su principio fundamental reside en la relación entre la luz reflejada por la superficie terrestre en dos regiones específicas del espectro electromagnético: el rojo visible y el infrarrojo cercano. Cuando se observa una baja reflectancia en la luz roja y una alta reflectancia en el infrarrojo cercano, esto indica una mayor actividad fotosintética, lo que sugiere la presencia de vegetación densa y saludable. El NDVI resulta una herramienta clave para obtener información sobre la vegetación, especialmente en la monitorización de ecosistemas, el seguimiento de cambios en la cobertura vegetal y la evaluación del estado de los cultivos (NASA Earth Observatory, s.f).

1.6.2 Índice de Diferencia Urbanizada Normalizada (NDBI)

El Índice de Diferencia Urbanizada Normalizada (NDBI) es una herramienta dentro de la teledetección que se emplea para resaltar y caracterizar áreas urbanizadas o construidas en el terreno. Este índice se calcula utilizando dos bandas espectrales: infrarrojo cercano (NIR) e infrarrojo de onda corta (SWIR). Esta combinación facilita la identificación y el resalte de zonas donde se encuentran estructuras urbanas y materiales construidos.

El NDBI utiliza proporciones específicas entre el NIR y el SWIR para minimizar los efectos adversos de las variaciones en la iluminación del terreno, así como para compensar los efectos atmosféricos que pueden afectar las mediciones desde la órbita satelital. Al hacerlo, se logra una mayor precisión en la identificación de las áreas urbanas, incluso en condiciones desafiantes como variaciones en la inclinación del terreno o presencia de contaminantes atmosféricos.

1.6.3 Random Forest

El método de aprendizaje automático RF es ampliamente aplicado en diversos campos, incluyendo la clasificación de imágenes, el mapeo de riesgos para desastres como deslizamientos de tierra, hundimientos e incendios forestales, así como en la exploración minera, entre otros (Feizizadeh *et al.*, 2021). Esta versatilidad se debe a su capacidad para manejar tanto tareas de clasificación como de regresión, lo que lo convierte en una herramienta poderosa en diferentes contextos.

Según IBM (s.f.), Random Forest está basado en el uso de múltiples árboles de decisión para realizar predicciones. Un árbol de decisión es un algoritmo de aprendizaje supervisado que, al ser no paramétrico, no requiere suposiciones previas sobre la distribución de los datos. Estos árboles tienen una estructura jerárquica compuesta por un nodo raíz (donde comienza la toma de decisiones), ramas (que representan las posibles opciones), nodos internos (que actúan como puntos de decisión intermedios) y nodos hoja (que representan los resultados o predicciones finales).

Feizizadeh, *et al.* (2021) refiere que los árboles de decisión se construyen utilizando una técnica de remuestreo con reemplazo, que incluye el muestreo aleatorio de los atributos y la selección de la mejor división entre esas variables en lugar de la mejor división entre todos los atributos. La asignación de la etiqueta de clase de una instancia desconocida se realiza mediante una estrategia de votación mayoritaria. Cada árbol de decisión en RF lleva a cabo la clasificación (voto), lo que significa que como resultado del modelo se elige la clase que obtiene más votos.

RF es un algoritmo robusto y flexible que, gracias a su capacidad para construir múltiples árboles de decisión independientes y combinar sus resultados, ofrece una alta precisión y generalización en una amplia gama de problemas del mundo real.

1.6.4 Support Vector Machine

El algoritmo SVM es una técnica de aprendizaje automático supervisado no paramétrico, lo que significa que no asume ninguna distribución específica de los datos. Este método es ampliamente utilizado para abordar problemas tanto de regresión como de clasificación. SVM incluye una fase inicial de entrenamiento en la que se utilizan datos

de entrada y salida relevantes para el problema que se está intentando resolver. Durante esta fase de entrenamiento, el objetivo es identificar los vectores de soporte, que son los puntos de datos más cercanos al límite de decisión entre las clases. Utilizando los datos de entrenamiento y aplicando algoritmos de optimización, el modelo de SVM determina el límite de decisión lineal óptimo que maximiza el margen entre las clases. Este margen se refiere a la distancia entre el límite de decisión y los vectores de soporte más cercanos. Al maximizar este margen, SVM garantiza que el modelo pueda generalizar mejor a nuevos datos, reduciendo la posibilidad de errores en la clasificación de ejemplos futuros (Feizizadeh *et al.*, 2021; Basheer, 2022).

1.7 Cambio de Uso de Suelo

La FAO define el suelo como un medio natural esencial para el crecimiento de las plantas, compuesto por horizontes formados a partir de la meteorización de materiales minerales, junto con materia orgánica, aire y agua. Este recurso natural es producto de la interacción a lo largo del tiempo de diversos factores como el clima, la topografía, la vegetación, la fauna, la actividad humana y las rocas, entre otros.

El Grupo Intergubernamental de Expertos sobre el Cambio Climático - IPCC (2012) describe el uso del suelo como el conjunto de actividades, insumos y disposiciones implementadas en un tipo específico de cobertura terrestre. En cuanto al cambio de uso del suelo, se refiere a las modificaciones en cómo los seres humanos gestionan o utilizan la tierra, lo que puede generar transformaciones en la cobertura de la superficie terrestre.

El cambio de uso de suelo se refiere a la evolución de la vegetación original hacia otro uso o al deterioro de su calidad a través de procesos físicos, químicos y biológicos. De acuerdo con la FAO, la degradación del suelo es el deterioro en su estado de salud, lo que reduce la capacidad del ecosistema para generar bienes y servicios para los humanos. Por otro lado, la erosión del suelo implica la pérdida de la capa superficial junto con sus nutrientes, y es un proceso natural en las zonas montañosas que tiende a empeorar por malas prácticas de manejo.

De acuerdo con Jiménez-Moreno *et al.* (2011), existen tres criterios para clasificar los métodos de detección de cambios en el uso del suelo. El primero se relaciona con

técnicas para identificar dichos cambios, las cuales pueden estar vinculadas a modificaciones en la reflectividad o en las texturas de las imágenes, siendo estas separables y afectadas por limitaciones espaciales, espectrales, temáticas y temporales. Entre los métodos utilizados se encuentran el álgebra de mapas, transformaciones matemáticas, post-clasificación y modelación, como la evaluación multicriterio y autómatas celulares.

El segundo criterio se enfoca en la temporalidad de la información abarcando el análisis bitemporal, análisis multitemporal y predicción de cambios a futuro. Finalmente, el tercer criterio de clasificación se refiere al tipo de datos utilizados y se pueden agrupar en análisis de datos continuos (álgebra de mapas, transformaciones matemáticas y modelación) y categóricos (post-clasificación con matrices de cambio o Cadenas de Markov).

1.8 Modelo predictivo

Según la Real Academia Española, un modelo es un esquema teórico, generalmente expresado en forma matemática, que representa un sistema o una realidad compleja y facilita la comprensión y el análisis de su comportamiento. Un modelo de predicción en el estudio de cambios en los usos del suelo tiene como objetivo principal anticipar, prevenir, comprender las transiciones entre diferentes tipos de uso del suelo, así como mitigar las dinámicas que pueden volverse insostenibles para el crecimiento de las ciudades (Aguilera Benavente, 2006).

Conforme a Henríquez Ruiz y Azócar García (2007), existen dos tipos de modelos: estáticos y dinámicos. Los modelos estáticos utilizan datos de entrada y salida que corresponden al mismo instante en el tiempo, permitiendo combinar diversas entradas para obtener un resultado relevante en ese momento específico. En cambio, en los modelos dinámicos la salida se produce en un momento distinto al de los datos de entrada, este modelo ayuda a predecir el comportamiento futuro y la proyección de posibles consecuencias de las transformaciones proyectadas.

En el ámbito de los modelos predictivos para cambios de uso del suelo, existen dos enfoques principales. El primero incluye modelos basados en regresión, que establecen

relaciones entre un amplio rango de variables predictivas y las probabilidades de cambio de uso del suelo. El segundo enfoque corresponde a los modelos de transición espacial, que comprende técnicas estocásticas como las Cadenas de Markov y Autómatas Celulares, donde se asume explícitamente que las áreas vecinas ejercen una influencia significativa en la probabilidad de transición del área o celda central.

CAPÍTULO 2.

METODOLOGÍA

2.1 Ubicación Geográfica

En noviembre de 2016, la LIX Legislatura del Estado de México aprobó la declaratoria de una nueva zona metropolitana que comprende los municipios de Almoloya del Río, Atizapán, Capulhuac, Texcalyacac, Tianguistenco y Xalatlaco, nombrándola “Zona Metropolitana de Tianguistenco” (COESPO, 2017).

La Zona Metropolitana de Tianguistenco está ubicada en la región centro de México, cuenta con una superficie de 303.77 km². Las coordenadas geográficas que delimitan la región son:

- Latitud norte

Latitud máxima: 19°14'58.96”

Latitud mínima: 19°05'11.52”

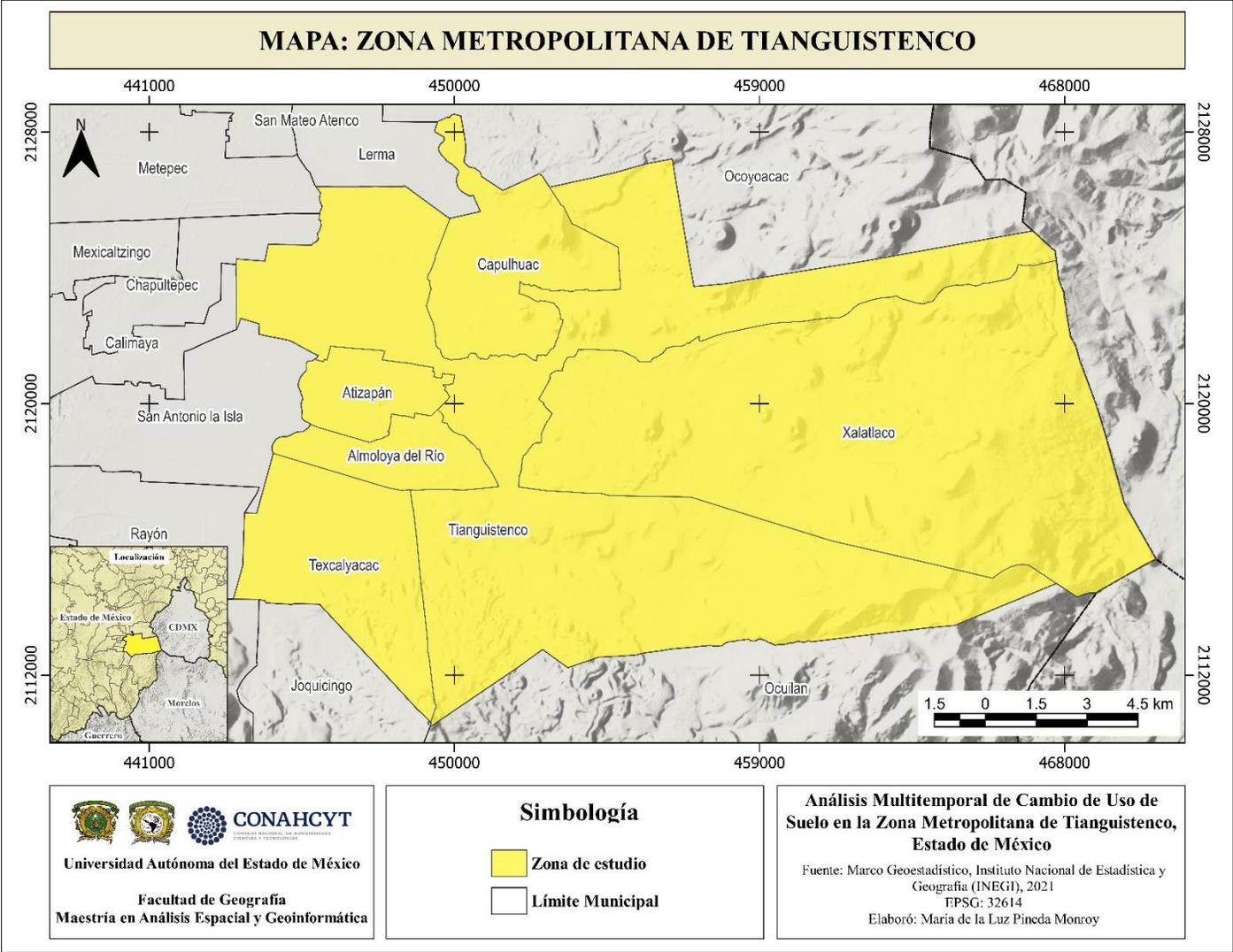
- Longitud oeste

Longitud máxima: 99°16'43.30”

Longitud mínima: 99°32'14.56”

El desarrollo de esta metrópoli está relacionado con su proximidad a la Ciudad de México y la zona industrial del Valle de Toluca, lo que ha impulsado el crecimiento económico de la región (COESPO, 2017) (Figura 2.1).

Figura 2.1. Mapa de ubicación de la zona de estudio: Zona Metropolitana de Tianguistenco



Fuente: Elaboración propia con base en INEGI

2.1.1 Características Físico-Geográficas

La información físico-geográfica de la Zona Metropolitana de Tianguistenco es de importancia para un análisis multitemporal de cambio de uso de suelo, ya que proporciona el contexto necesario para entender cómo las dinámicas naturales y antropogénicas han influido en los cambios de cobertura y uso de suelo a lo largo del tiempo.

El relieve de la ZMT presenta una combinación de zonas montañosas y llanuras, destacando los municipios de Xalatlaco, Tianguistenco y Texcalyacac, el primero alberga más de una docena de volcanes inactivos, entre los que destacan el Tezontle, Muñeco y Jumento, con altitudes que superan los 8,850 metros sobre el nivel del mar (msnm). En contraste, las llanuras están vinculadas a las “Ciénegas del Lerma”, con altura promedio de 2,500 msnm, reflejando la diversidad topográfica de la región. Geológicamente, la metrópoli se sitúa en la Faja Volcánica Transmexicana, cuyo origen se remonta a los periodos Plioceno y Mioceno. La región está compuesta por rocas clásticas, volcanoclásticas y volcánicas de la era terciaria, junto con rocas ígneas intrusivas como el basalto y la andesita. Además, la zona lacustre está formada por materiales aluviales ricos en sedimentos. En cuanto a la edafología, los suelos predominantes son el andosol húmico y andosol mólico. El andosol húmico es rico en materia orgánica pero pobre en nutrientes y tiene un nivel de acidez elevado. En contraste, el andosol mólico también es rico en materia orgánica, pero cuenta con mayor contenido de nutrientes y una acidez baja, lo que lo hace ideal para la agricultura. Otros tipos de suelos presentes en la región son andosol, feozem, vertisol, gleysol húmico, histosol, histosol eútrico, litosol y leptosol. El clima de la ZMT es templado subhúmedo con lluvias en verano y la presencia de microclimas debido al terreno accidentado, particularmente en las zonas forestales de la región. La temperatura media anual es de 14.7° C, con un máximo de 23° C y mínimo de 3.1°C. La precipitación media anual oscila entre los 700 y 1500 milímetros.

2.1.2 Uso de Suelo

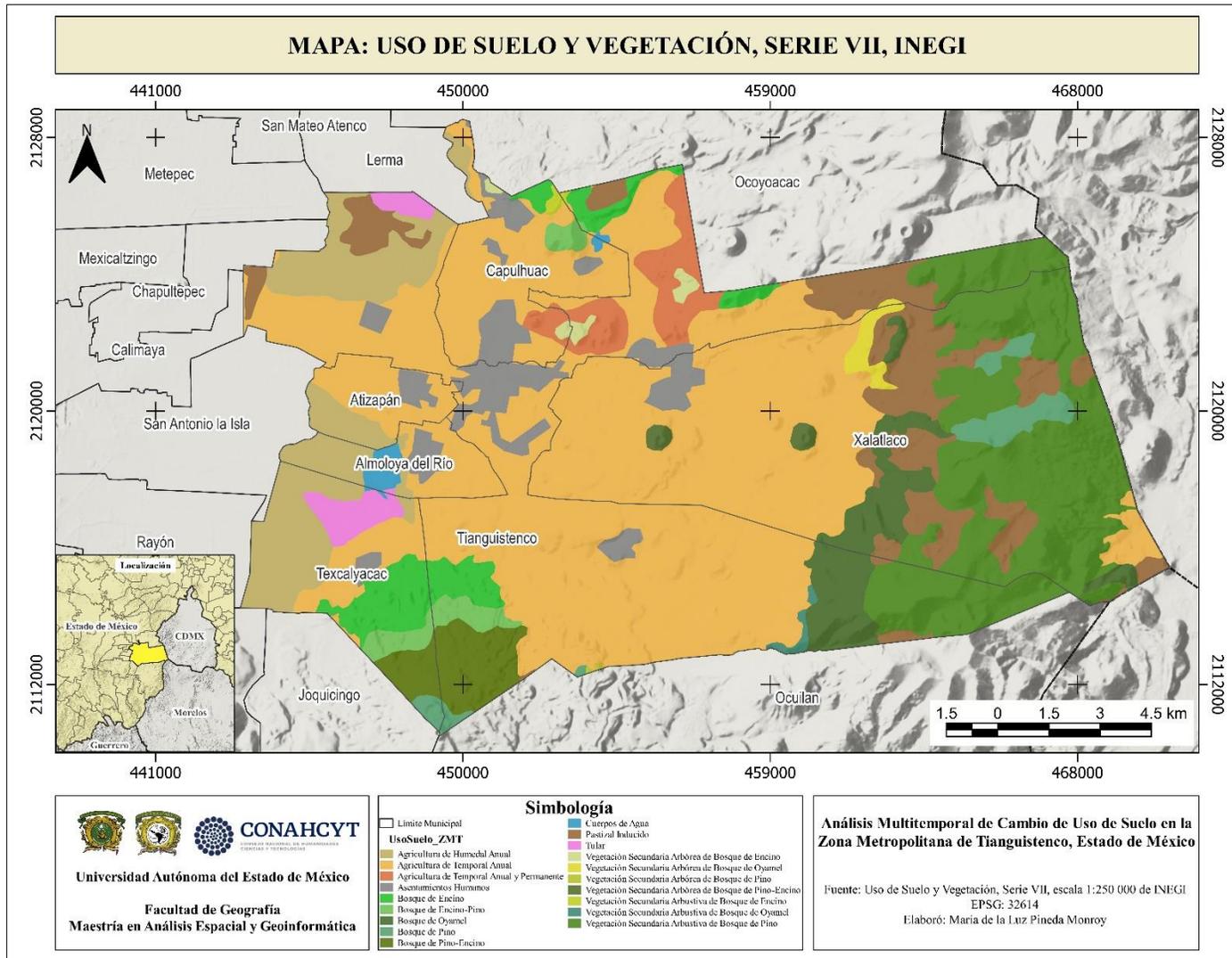
De acuerdo con la Serie VII de Uso de suelo y Vegetación, escala 1:250 000 de INEGI, es la agricultura de temporal anual la que ocupa el valor más alto de hectáreas dentro de la ZMT, seguido de la vegetación secundaria arbustiva de bosque de pino, y en el último sitio, la vegetación secundaria arbórea de bosque de pino (Tabla 2.1 y Figura 2.2).

Tabla 2.1. Uso de suelo y vegetación, serie VII, INEGI

Uso De Suelo	Hectáreas
Asentamientos Humanos	1,439.21
Bosque de Oyamel	1,198.51
Bosque de Pino	459.771
Bosque de Pino-Encino	749.82
Bosque de Encino	984.003
Bosque de Encino-Pino	393.032
Cuerpos de Agua	143.039
Agricultura de Humedal Anual	2,299.26
Pastizal Inducido	2,373.75
Agricultura de Temporal Anual	14,017.2
Agricultura de Temporal Anual y Permanente	789.798
Vegetación Secundaria Arbórea de Bosque de Oyamel	163.34
Vegetación Secundaria Arbustiva de Bosque de Oyamel	57.05
Vegetación Secundaria Arbustiva de Bosque de Pino	4,712.4
Vegetación Secundaria Arbórea de Bosque de Pino	7.14
Vegetación Secundaria Arbórea de Bosque de Pino-Encino	98.69
Vegetación Secundaria Arbustiva de Bosque de Encino	32.52
Vegetación Secundaria Arbórea de Bosque de Encino	110.22
Tular	371.2

Elaboración propia con base en INEGI

Figura 2.2. Mapa de Uso de Suelo y Vegetación, Serie VII, INEGI



Fuente: Elaboración propia con base en Uso de Suelo y Vegetación, Serie VII, INEGI

2.1.3 Ciénegas de Lerma

Las Ciénegas del Lerma poseen una gran riqueza hídrica, razón por la cual, en el 2004 se catalogaron por las ONU dentro de la Lista de Humedales de Importancia Internacional (Sitios RAMSAR). Este reconocimiento implica un compromiso del gobierno para implementar acciones que aseguren la preservación de las características ecológicas de estos humedales (Ramsar, 2014).

Las Ciénegas del Lerma forman parte de una de las 232 Áreas Naturales Protegidas de carácter federal del país, fue declarada como ANP en el 2002 debido a su importancia como unidad ecológica. Este espacio alberga diversos ecosistemas con una alta biodiversidad y una considerable fragilidad ambiental. Además, es uno de los pocos lugares en el país que aún conserva sistemas ecológicos bien desarrollados, considerados como los últimos vestigios de los extensos humedales que alguna vez existieron en el Altiplano Central. (SEMARNAT, 2018; Gobierno de México, 2024).

2.2 Método

A continuación, se presenta la metodología diseñada para cumplir con los objetivos de este estudio, el cual se ha dividido en cuatro fases (Figura 2.3 y Tabla 2.2):

La metodología empleada en este proyecto constó de una serie de procesos que, al complementarse entre sí, permiten avanzar por cada una de las fases, cumpliendo con cada uno de los objetivos específicos planteados. La selección de los procesos llevados a cabo en cada fase se fundamentó en una exhaustiva revisión bibliográfica especializada, que abarca técnicas y métodos empleados en la obtención de mapas de uso de suelo mediante teledetección, herramienta que ha experimentado un notable auge en los últimos años debido a su eficacia y precisión de resultados.

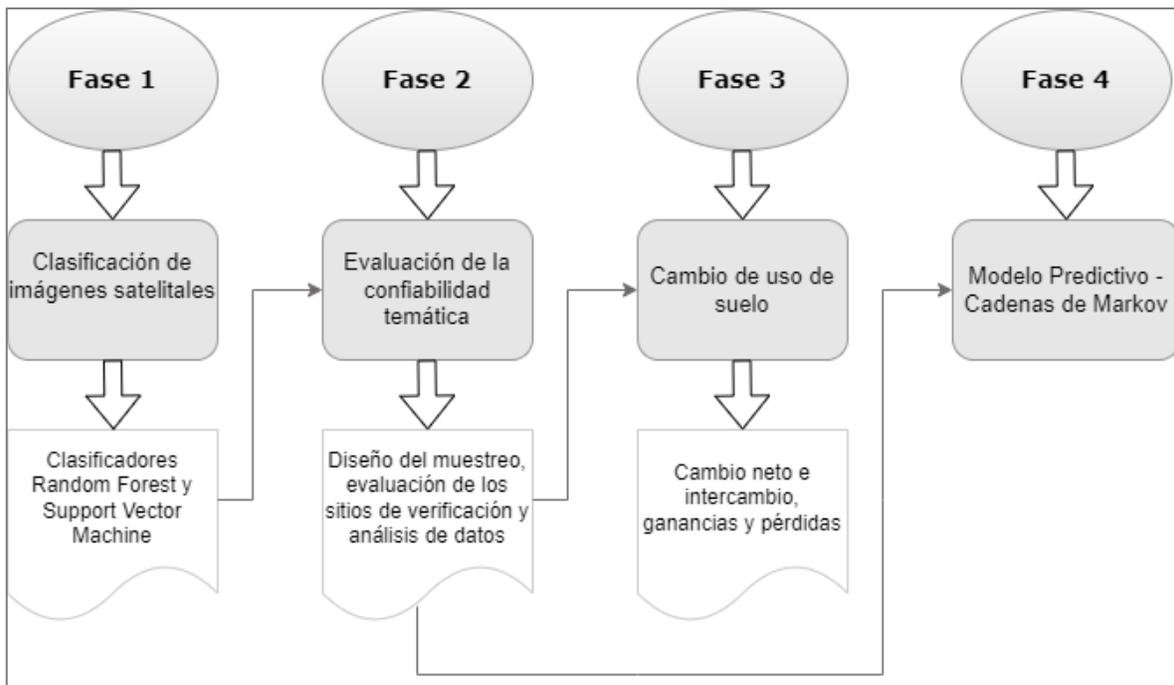
Para garantizar la máxima fiabilidad en las capas de uso de suelo, es necesario validar los mapas generados a partir de la clasificación en GEE. Con este propósito, se realizó la evaluación de la confiabilidad temática, utilizando una matriz de error o confusión. Esta matriz permite evaluar la precisión de la clasificación e identificar

errores de asignación en cada clase. Además, se calculó el coeficiente Kappa, el cual proporciona la concordancia de la clasificación.

Posteriormente el análisis de cambio de uso de suelo a partir de la matriz de tabulación cruzada, método ampliamente utilizado y explicado por Pontius y que ha sido utilizado por diversos autores para comparar los cambios en la cobertura del suelo en diferentes periodos de tiempo.

Por otra parte, el modelo predictivo Cadenas de Markov, es muy útil para la predicción de escenarios a futuro. Este modelo probabilístico proporciona una herramienta útil para la planificación y gestión del territorio.

Figura 2.3. Esquema metodológico



Fuente: Elaboración propia

Tabla 2.2. Descripción de los procesos para la aplicación de la metodología

Fase	Descripción
1. Clasificación de las imágenes satelitales	A través de herramientas geotecnológicas como la teledetección y dos algoritmos de clasificación: Random Forest (RF) y Support Vector Machine (SVM) se obtuvieron mapas clasificados de uso de

Fase	Descripción
	suelo para dos periodos de tiempo, el insumo principal son imágenes de satélite del sensor Landsat 8 y Sentinel-2.
2. Evaluación de la confiabilidad temática	El proceso consistió en comparar la información de la clasificación con una fuente de referencia considerada muy fiable. La evaluación de la confiabilidad temática se desarrolla en tres etapas: primero, el diseño de muestreo, contempla las unidades de muestreo, el método de muestreo y el tamaño de la muestra; segundo, la evaluación de los sitios de verificación, consiste en caracterizar estos sitios y asociarlos con una o varias clases del mapa que se está evaluando; y, por último, el análisis de datos, realizado mediante una matriz de confusión o de error comparando los sitios de verificación con la base cartográfica para calcular el coeficiente kappa y determinar la confiabilidad del mapa.
3. Cambio de uso de suelo	Se basó en el análisis de cambio de uso de suelo por medio de cálculo de pérdidas, ganancias, cambios netos e intercambios de categorías entre ambos periodos.
4. Modelo predictivo	Cadenas de Markov es un modelo probabilístico que estudia el comportamiento de variables aleatorias a lo largo del tiempo.

Fuente: Elaboración propia

2.3 Fase 1. Clasificación de las imágenes satelitales

A través de la herramienta de Google Earth Engine se realizó la clasificación de las imágenes de satélite del sensor Landsat 8 y Sentinel 2 para los periodos de 2016 y 2022 donde se aplicaron los algoritmos RF y SVM que, de acuerdo con la

bibliografía y comparados con otros algoritmos han mostrado mejores resultados para la obtención de mapas de uso de suelo. A continuación, se explica de manera teórica la aplicación de ambos clasificadores:

El algoritmo de Random Forest de acuerdo con Breiman, (2001) citando en Liaw y Wiener (2002):

Se extraen n_{tree} muestras de *bootstrap* de los datos originales

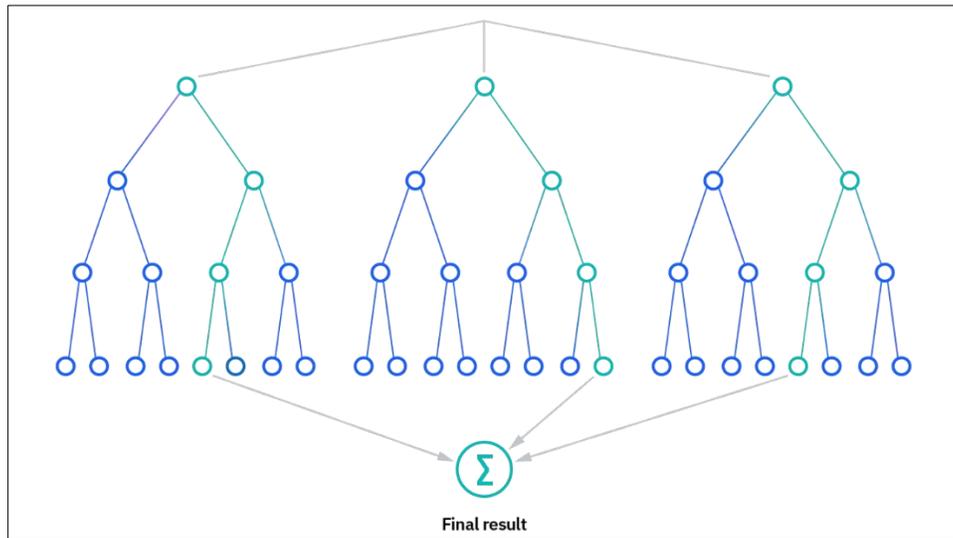
1. Para cada una de las muestras de *bootstrap* haga que crezca un árbol de clasificación o regresión *unpruned*, con la siguiente modificación: en cada nodo, en lugar de elegir la mejor división entre todos los predictores, muestrea aleatoriamente m_{try} de los predictores y elige la mejor división entre esas variables. (Bagging se puede considerar como el caso especial de un bosque aleatorio obtenido cuando $m_{try} = p$, el número de predictores).
2. Predicción de nuevos datos mediante la agregación de las predicciones de los árboles n_{tree} (es decir, votos mayoritarios para la clasificación, promedio de la regresión).

Es posible obtener una estimación de la tasa de error, basada en los datos de entrenamiento, mediante lo siguiente:

1. En cada iteración del *bootstrap*, prediga los datos que no están en la muestra del Bootstrap (el autor los llama datos *out-of-bag* u OOB) utilizando el árbol cultivado con la muestra *bootstrap*.
2. Se agregan las predicciones OOB. (En promedio, cada punto de datos estaría fuera de la bolsa alrededor del 36% de las veces, así que se agregan las predicciones). Se calcula la tasa de error y se le nombra estimación de la tasa de error OOB.

Random Forest se organiza en una estructura jerárquica similar a un árbol, compuesta por un nodo raíz, ramas, nodos internos y nodos hoja. En este sistema, cada árbol de decisión realiza una clasificación, emitiendo un "voto", y la clase seleccionada por el modelo final es aquella que recibe la mayoría de los votos, de manera gráfica se muestra en la Figura 2.4.

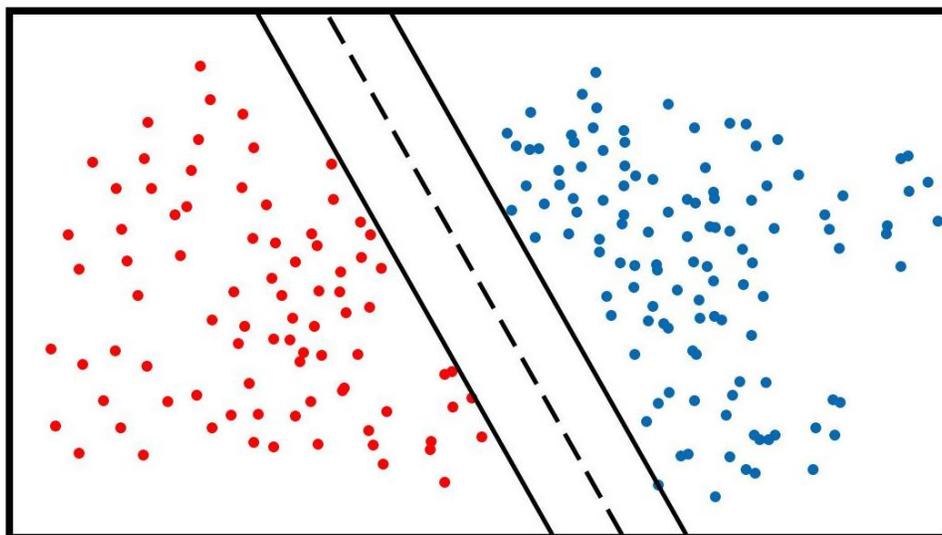
Figura 2.4. Random Forest



Fuente: IBM (s.f)

Support Vector Machine (SVM): El algoritmo necesita una función kernel adecuada para establecer los hiperplanos con precisión y minimizar los errores de clasificación. La parte esencial de la técnica SVM es el tipo de núcleo utilizado. La funcionalidad de la SVM depende principalmente del tamaño del kernel, y la similitud de una superficie lisa depende de la densidad del kernel más significativa (Talukdar *et al.*, 2020).

Figura 2.5. Support Vector Machine



Fuente: Robotic Electronics (2024)

SVM incluye una fase de entrenamiento con valores de entrada relevantes, al usar primero los datos de entrenamiento y los algoritmos de optimización, se obtienen los vectores de soporte para determinar el límite de decisión lineal óptimo y luego el vector lineal óptimo es obtenido para la separación de clases, como se muestra en la Figura 2.5.

Una vez clasificadas las imágenes de satélite se procede a continuar con la fase dos:

2.4 Fase 2. Evaluación de la Confiabilidad Temática

La evaluación de la confiabilidad temática consiste en confrontar la información de la clasificación con una fuente de referencia altamente confiable. Normalmente este proceso se basa en el muestreo de sitios de verificación, cuya clasificación se consigue a partir de observaciones de campo o mediante el análisis de imágenes de mayor resolución. Este procedimiento se divide en tres etapas según Stehman y Czaplewski (1998) citados en Mas *et al.*, (2003) y Mas y Couturier (2011):

2.4.1 Diseño de muestreo

Considera la determinación del tipo de unidades de muestreo, del método para seleccionarlas y la cantidad de unidades requeridas, es decir, el tamaño de la muestra (Mas *et al.*, 2003).

Unidades de muestreo

De acuerdo con Chuvieco (1995), las unidades de muestreo facilitan la vinculación entre la ubicación de la información del mapa y el terreno. Estas unidades pueden ser representadas por un punto, un píxel, un conjunto de píxeles, un polígono en el mapa, o una superficie con formas predefinidas.

El método de muestreo

Se utiliza para seleccionar una pequeña porción del área representada en el mapa, con el objetivo de garantizar que esta muestra sea representativa de toda la extensión cartografiada.

El tamaño de la muestra

Se refiere a la cantidad de sitios de verificación utilizados que se emplean para calcular la confiabilidad del mapa clasificado. Para reducir el número de estos sitios, se puede ajustar la precisión de acuerdo con los intereses del usuario. La Tabla 2.3 muestra el tamaño de la muestra requerido para distintos valores de fiabilidad y precisión de la evaluación. p hace referencia a la fiabilidad estimada y B al medio intervalo de confianza.

Tabla 2.3. Tamaño de la muestra por clase en función de p y B

p (%)	90	80	70	60	50
B (%)					
2.5	555	983	1291	1475	1537
5.0	138	246	323	369	384
10.0	35	61	81	92	96

Fuente: Elaboración propia con base en Mas y Couturier (2011)

2.4.2 Evaluación de los sitios de verificación

Consiste en la caracterización del sitio de verificación con el propósito de vincularlo a una o varias clases representadas en la leyenda del mapa que se está evaluando.

2.4.3 Análisis de datos

El proceso se lleva a cabo a través de una matriz de confusión o matriz de error (Tabla 2.4), la cual permite comparar la información de los sitios de verificación con los datos de la base cartográfica que se busca evaluar. En dicha matriz, las filas generalmente corresponden a las clases de referencia, mientras que las columnas representan las clases del mapa. La diagonal de la matriz indica el número de sitios de verificación en los que existe concordancia entre el mapa y los datos de referencia, mientras que los valores marginales señalan los errores de clasificación (Mas y Couturier, 2011).

Tabla 2.4. Matriz de confusión o matriz error

Mapa	C ₁	C ₂	...	C _j	...	C _q	Total
Sitios de Verificación							
C ₁	P_{11}	P_{12}	...	P_{1j}	...	P_{1q}	P_{1+}
C ₂	P_{21}	P_{22}	...	P_{2j}	...	P_{2q}	P_{2+}
...
C _i	P_{i1}	P_{i2}	...	P_{ij}	...	P_{iq}	P_{i+}
...
C _q	P_{q1}	P_{q2}	...	P_{qj}	...	P_{qq}	P_{q+}
Total	P_{+1}	P_{+2}	...	P_{+j}	...	P_{+q}	1

Fuente: Elaboración propia con base en Mas y Couturier, 2011

La Exactitud global se calcula sumando los valores de la diagonal de la matriz.

Existen dos tipos de errores que pueden identificarse según se analicen las filas o las columnas de la matriz. El primero es el error de comisión, que se refiere a la proporción de sitios que han sido clasificados en una determinada categoría en el mapa, pero que en realidad pertenecen a otra clase. El segundo es el error de omisión, el cual se refiere a los sitios de verificación correspondientes a una categoría que fue cartografiada en otra.

A partir de la relación entre los dos tipos de errores, se definen dos índices de precisión. El primero es la precisión del usuario, que se refiere a la probabilidad de que un sitio clasificado como “k” y seleccionado al azar corresponda realmente a la clase “k” en el terreno. El segundo índice es la precisión del productor, que indica la proporción de sitios de verificación de la clase “k” que están correctamente representados en el mapa o base de datos.

Los índices previamente mencionados no consideran los elementos fuera de la diagonal de la matriz, por lo que se utiliza el coeficiente kappa. Este coeficiente utiliza las sumas marginales de la matriz y refleja la contribución del azar en la confiabilidad del mapa.

El cálculo del coeficiente kappa se realiza mediante la ecuación 1:

$$K = \frac{P_o - P_c}{1 - P_c} \quad (1)$$

Donde:

K = Índice Kappa

P_o = proporción de área correctamente clasificada

P_c = confiabilidad resultante al azar

Para la interpretación del valor K se utiliza una escala que exprese cualitativamente la fuerza de concordancia (Tabla 2.5) (Landis y Koch, 1977).

Tabla 2.5. Coeficiente Kappa

Coeficiente Kappa	Fuerza de concordancia
< 0.00	Pobre (Poor)
0.01 – 0.20	Leve (Slight)
0.21 – 0.40	Aceptable (Fair)
0.41 – 0.60	Moderada (Moderate)
0.61 – 0.80	Considerable (Substantial)
0.81 – 1.00	Casi perfecta (Almost Perfect)

Fuente: Landis y Koch, 1977

2.5 Fase 3. Cambio de uso de suelo

Pontius *et al.* (2004) propone a la matriz de tabulación cruzada como una herramienta fundamental para el análisis de detección de cambios de uso de suelo, basada en dos pares de componentes: cambio neto e intercambio, así como ganancias y pérdidas brutas. A partir del mapa del tiempo 1 y mapa del tiempo 2 se construye una matriz de tabulación cruzada (Tabla 2.6), donde las filas representan las categorías del tiempo 1 y las columnas muestran las categorías del tiempo 2.

La ganancia (G_{ij}) se calcula a partir de la matriz, obteniéndose al restar la persistencia (P_{jj}) del total de la columna correspondiente al tiempo 2 (P_{+j}), tal como se muestra en la ecuación 2.

$$G_{ij} = P_{+j} - P_{jj} \quad (2)$$

La pérdida (Lij), tal como se indica en la ecuación 3, se obtiene al restar la persistencia (Pjj) del total de la fila correspondiente al tiempo 1 (Pj+).

$$L_{ij} = P_{j+} - P_{jj} \quad (3)$$

El intercambio entre categorías (Sj) se determina multiplicando por dos veces el valor mínimo de las ganancias y las pérdidas (Ecuación 4).

$$S_j = 2 \times \text{MIN} (P_{j+} - P_{jj}, P_{+j} - P_{jj}) \quad (4)$$

El cambio total para cada categoría (Cj) se establece como la suma del cambio neto (Dj) y el intercambio (Sj), o alternativamente, como la suma de las ganancias y pérdidas, según se presenta en la ecuación 5.

$$C_j = D_j + S_j \quad (5)$$

Tabla 2.6. Matriz de tabulación cruzada

	Tiempo 2				Total tiempo 1	Pérdidas
	Categoría 1	Categoría 2	Categoría 3	Categoría 4		
Tiempo 1						
Categoría 1	P_{11}	P_{12}	P_{13}	P_{14}	P_{1+}	$P_{1+} - P_{11}$
Categoría 2	P_{21}	P_{22}	P_{23}	P_{24}	P_{2+}	$P_{2+} - P_{22}$
Categoría 3	P_{31}	P_{32}	P_{33}	P_{34}	P_{3+}	$P_{3+} - P_{33}$
Categoría 4	P_{41}	P_{42}	P_{43}	P_{44}	P_{4+}	$P_{4+} - P_{44}$
Total tiempo 2	P_{+1}	P_{+2}	P_{+3}	P_{+4}	1	
Ganancias	$P_{+1} - P_{11}$	$P_{+2} - P_{22}$	$P_{+3} - P_{33}$	$P_{+4} - P_{44}$		

Fuente: Elaboración propia con base en Pontius (2004)

2.6 Fase 4. Modelo Predictivo – Cadena de Markov

Una Cadena de Markov es un modelo matemático utilizado para describir un proceso que cambia de estado a lo largo del tiempo, en el cual el estado futuro solo depende del estado presente, sin tener en cuenta los estados anteriores. Los cambios entre estados en una Cadena de Markov se conocen como transiciones, y cada transición tiene una probabilidad asociada. Una matriz de transición se construye al reunir todas estas posibilidades (Ibarra, 2009; López Hung y Joa Triay, 2017).

Este modelo es ampliamente utilizado en diversos campos debido a su capacidad para representar sistemas dinámicos que evolucionan con el tiempo de manera probabilística. En economía se aplica para analizar fluctuaciones de mercado, en biología para modelar el comportamiento de la población, así como predecir fenómenos naturales, predicción de patrones climáticos y en epidemiología para modelar la propagación de enfermedades.

CAPÍTULO 3.

RESULTADOS

En el siguiente capítulo se presentan los resultados obtenidos mediante la metodología aplicada para alcanzar los objetivos de este estudio. Los resultados muestran que, al aplicar los algoritmos RF y SVM a imágenes satelitales Landsat 8 y Sentinel-2, el algoritmo SVM obtuvo un mejor desempeño, alcanzando un coeficiente Kappa de 0.84 y 0.82 para las imágenes Sentinel-2 de 2016 y 2022, respectivamente.

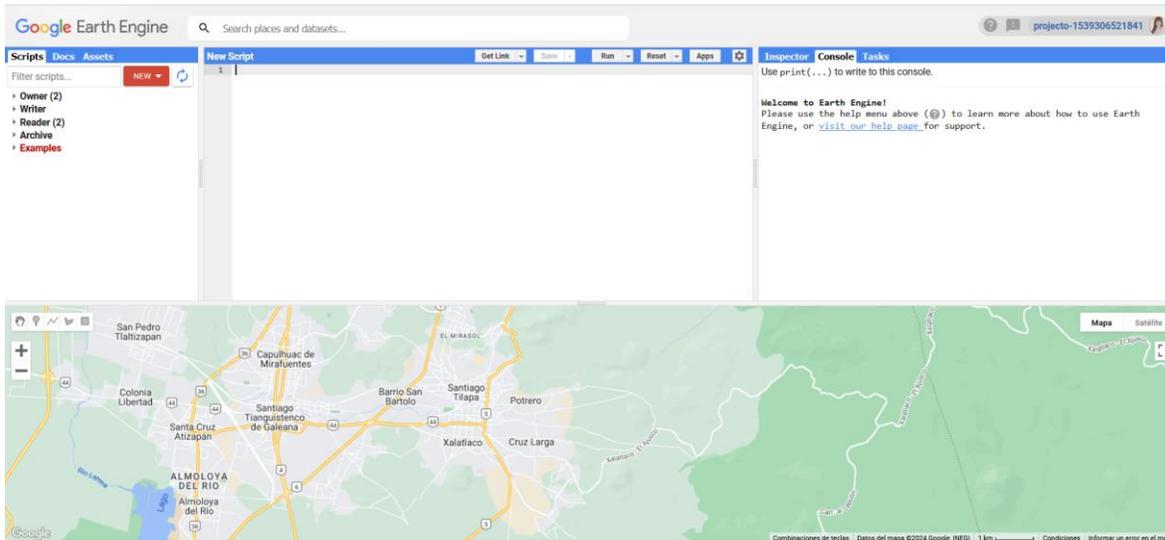
En cuanto a los cambios de uso de suelo, el bosque resultó la categoría con mayor reducción, perdiendo 2,246 hectáreas, debido principalmente a su conversión en pastizales y agricultura. Además, la zona urbana incrementó en 876 hectáreas durante el periodo de análisis.

Finalmente, el modelo predictivo Cadenas de Markov proyecta que, para el año 2028, se espera una disminución de 1,763 hectáreas en la cobertura forestal y un crecimiento de 844.14 hectáreas en la expansión urbana, resultados derivados de las dinámicas que han afectado a la Zona Metropolitana de Tianguistenco.

3.1 Clasificación de imágenes satelitales

Para obtener los mapas de uso de suelo de la ZMT, fue necesario contar con una cuenta de Google para acceder a la plataforma en la nube Google Earth Engine (Figura 3.1). Posteriormente, empleando código en JavaScript, se seleccionaron dos imágenes de la colección Sentinel-2 con preprocesamiento de nivel L2A, que incluye reflectancia de superficie corregida atmosféricamente y ortorectificación. Así mismo, se eligieron dos imágenes Landsat 8 con reflectancia de superficie y corrección atmosférica. A continuación, se presenta la nomenclatura de estas imágenes (Tabla 3.1):

Figura 3.1. Core Editor de Google Earth Engine



Fuente: Google Earth Engine (2024)

Tabla 3.1. Nomenclatura de imágenes de satélite del proyecto

Sensor	Nomenclatura
L8 2016	LANDSAT/LC08/C02/T1_L2/LC08_026047_20161206
L8 2022	LANDSAT/LC08/C02/T1_L2/LC08_026047_20221207
L2A 2016	COPERNICUS/S2_HARMONIZED/20161128T170642_20161128T17 1121_T14QMG
L2A 2022	COPERNICUS/S2_SR_HARMONIZED/20221202T170659_20221202 T170747_T14QMG

Fuente: Google Earth Engine, 2024

Se aplicó un filtro por región de interés utilizando el archivo shapefile de la zona de estudio y se estableció un rango de fechas, además de implementar una función para seleccionar las imágenes con la menor cobertura de nubes a partir de la siguiente sentencia: `.sort ('CLOUD COVER')`. La primera visualización de la imagen de interés se obtuvo recortando el shapefile de la zona de estudio, y se le añadió una composición RGB: B6-B5-B2 para Landsat 8 y B11-B8A-B2 para Sentinel-2A. Esta combinación de bandas se empleó para realizar un análisis orientado a la detección de zonas agrícolas y, a su vez, facilitó la diferenciación de las clases de uso de suelo.

Para el siguiente paso, es primordial considerar los usos del suelo al crear las muestras de entrenamiento. Se han seleccionado las siguientes categorías con base en la clasificación del INEGI de la cartografía de uso de suelo y vegetación (Serie VII, escala 1:250,000). En esta clasificación, varias subclases fueron consolidadas en una sola categoría, principalmente bosque y agricultura ya que las clases con mayor superficie en hectáreas correspondieron a la agricultura de temporal anual y vegetación secundaria arbórea de bosque de pino. Esta simplificación facilita el proceso de clasificación y análisis. Las clases resultantes son las siguientes:

1. Bosque
2. Agricultura
3. Zona Urbana
4. Cuerpos de agua
5. Tular
6. Pastizal

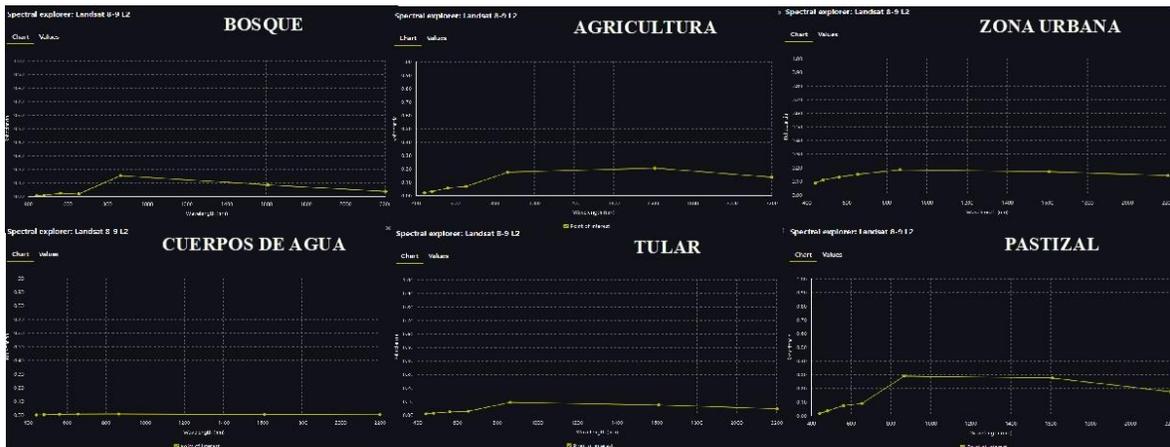
El muestreo se llevó a cabo utilizando polígonos generados en la composición en color verdadero de la ZMT, asignando a cada forma la clase de uso del suelo correspondiente. Con el objetivo de mejorar la precisión de los resultados, se resolvió un previo análisis de las firmas espectrales de las clases de uso del suelo. La firma espectral representa el comportamiento de la radiación reflejada por una superficie en función de la longitud de onda. Dicho análisis permite identificar y diferenciar las diferentes coberturas de suelo, ya que cada tipo de superficie, como la agricultura, cuerpos de agua o bosque, refleja la radiación de manera única en diferentes bandas espectrales. Al conocer las firmas espectrales, se logra una clasificación más precisa y confiable de los usos de suelo, ya que se puede distinguir mejor entre las diferentes categorías según su respuesta en las imágenes satelitales.

El procedimiento se llevó a cabo utilizando la plataforma EO Browser, herramienta de Sentinel Hub que permite visualizar imágenes satelitales de varias colecciones de datos. Facilita a los usuarios seleccionar áreas de interés, rangos de tiempo y

nivel de cobertura de nubes. Además, ofrece la posibilidad de crear visualizaciones personalizadas mediante el uso de bandas, índices de vegetación, generación de scripts, análisis estadístico, timelapses y firmas espectrales. Resulta útil para el análisis de información geoespacial.

En la Figura 3.2 se observan las firmas espectrales obtenidas de la imagen Landsat 8 del año 2022, la cual fue empleada para identificar el comportamiento espectral de las clases de uso de suelo en la ZMT para la clasificación. El bosque muestra una baja reflectancia en el espectro visible y un pico en el infrarrojo cercano (NIR), lo que sugiere vegetación saludable. En agricultura, los cultivos en fase de desarrollo presentan picos en el NIR, mientras que los ya cosechados muestran un comportamiento más uniforme entre el NIR e infrarrojo de onda corta (SWIR). Las zonas urbanas tienen una reflectancia más constante debido a la presencia de materiales artificiales. Los cuerpos de agua reflejan mínimamente en todas las bandas, facilitando su clasificación. El tular, vegetación de humedal, manifiesta un leve aumento en el NIR, menos acentuado que el de bosque. Por último, el pastizal refleja poco en el visible pero significativamente en NIR y SWIR.

Figura 3.2. Firmas espectrales en Landsat 8 de las clases de uso de suelo

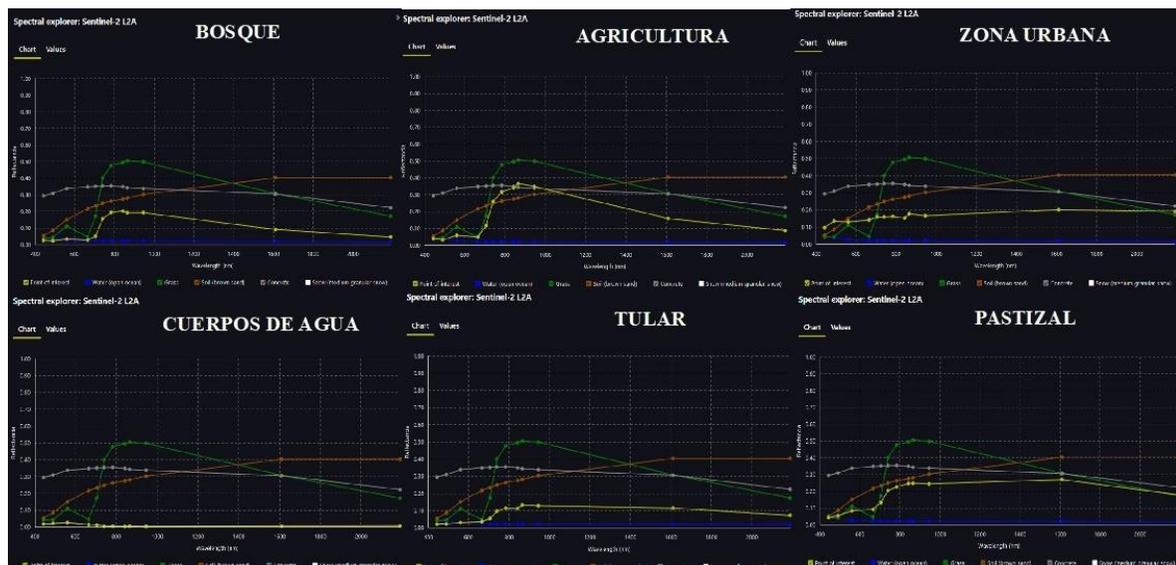


Fuente: EO Browser (2024)

Una de las ventajas de las firmas espectrales en imágenes Sentinel-2 es la posibilidad de compararlas con otras clases mediante EO Browser. En el caso del bosque, se observa baja reflectancia en el espectro visible y un pico en la banda 8 (NIR), típico de vegetación saludable, seguido de una disminución en el SWIR. La

agricultura muestra un comportamiento similar al bosque, especialmente en cultivos en desarrollo, con altos picos en la banda 8A, mientras que los cultivos cosechados destacan en las bandas 8 y 11. La zona urbana se encuentra estrechamente relacionada con la firma espectral del concreto, presentando el pico más alto en la banda 9. Los cuerpos de agua absorben y transmiten la mayoría de la radiación incidente, por lo que su reflexión al espacio es casi nula. El tular presenta un comportamiento semejante al bosque, pero con menor reflectancia, influenciado por la estación seca. Finalmente, el pastizal tiene su mayor reflectancia en la banda 11, siendo similar a la agricultura, lo que puede causar confusión en la clasificación (Figura 3.3).

Figura 3.3. Firmas espectrales en Sentinel-2 de las clases de uso de suelo



Fuente: EO Browser (2024)

Como se detalló en la metodología, los clasificadores empleados fueron RF y SVM; para RF se asignaron 50 árboles de decisión. Aunque se realizaron pruebas con diferentes cantidades de árboles para optimizar la clasificación, el mejor resultado se obtuvo precisamente con ese número, lo que garantiza un equilibrio entre precisión y eficiencia en el procesamiento. RF utilizó el clasificador *smileRandomForest()*.

Para la implementación del clasificador RF en GEE se emplearon las siguientes bandas espectrales (B2, B3, B4, B5, B6, B7 y B8), para Landsat 8 y para Sentinel-2 (B2, B3, B4, B5, B6, B7, B8, B8A, B11 y B12). En cuanto a SVM, se utilizó el clasificador *libsvm* (), que utiliza un tipo de kernel lineal. Este modelo fue entrenado de manera similar a RF, empleando el mismo conjunto de bandas y datos de entrenamiento. Ambos clasificadores se aplicaron sobre las imágenes recortadas de la zona de interés.

Un paso crucial tras la obtención de la clasificación de las imágenes satelitales es su exportación para procesarlas fuera de GEE. Para ello, fue necesario configurar ciertos parámetros, como la resolución de salida, que varía según el sensor: 10 metros para imágenes Sentinel-2 y 30 metros para Landsat 8. Además, se especificó una carpeta de destino en Google Drive para almacenar los archivos y definir el sistema de coordenadas adecuado.

Además de brindar acceso a grandes cantidades de datos satelitales, GEE ofrece una variedad de herramientas avanzadas para el preprocesamiento de imágenes. Esto es clave para mejorar la calidad de los análisis y reducir errores en los resultados. GEE incluye una amplia gama de algoritmos de clasificación supervisada, los cuales permiten generar categorías de diferentes usos de suelo. Al probar RF y SVM se permitió comparar los resultados obtenidos y seleccionar el algoritmo que mejor representara la realidad del área de estudio.

Una vez obtenida la clasificación de uso de suelo para ambos periodos en las imágenes Landsat 8 y Sentinel-2 se realizó la Fase 2, en la cual se llevó a cabo la Evaluación de la Confiabilidad Temática. Esta fase consta de tres etapas, comenzando con el diseño del muestreo, consiste en definir las unidades que relacionan el mapa con el terreno, para ello se seleccionó una muestra representativa del área clasificada y determinó el tamaño necesario de la muestra para evaluar la confiabilidad del mapa, ajustando esta cantidad de acuerdo al nivel de precisión requerido, en este sentido, para la obtención de una fiabilidad estimada de 80% y un intervalo medio de confianza del 5% se seleccionaron 246 sitios de verificación.

En la etapa correspondiente a la selección de unidades de muestreo, se eligieron píxeles generados de manera aleatoria dentro del área de estudio. Inicialmente, se definieron como puntos, los cuales fueron posteriormente transformados en píxeles. El tamaño del píxel dependió del mapa de uso de suelo de referencia utilizado en cada caso: para las imágenes de Landsat 8, se emplearon píxeles de 30 x 30 metros, mientras que para Sentinel-2 se utilizaron píxeles de 10 x 10 metros.

La información sobre el código JavaScript empleado para obtener las clasificaciones de uso de suelo en la Zona Metropolitana de Tianguistenco se detalla en el Anexo 1 y 2.

La segunda etapa de la evaluación de la confiabilidad temática es la evaluación de los sitios de verificación, la cual consistió en describir el sitio de verificación y vincularlo con una o varias clases de la leyenda del mapa que se evaluó. Para ello se realizó reconocimiento del área de estudio por medio de imágenes de satélite de alta resolución espacial, así como visitas a campo donde también se tomaron puntos de verificación de las categorías de suelo. Para el caso de las imágenes de satélite, la herramienta utilizada fue Google Earth Pro, misma que sirve para la visualización de imágenes de satélite de alta resolución, exploración de relieves, entre otras funciones.

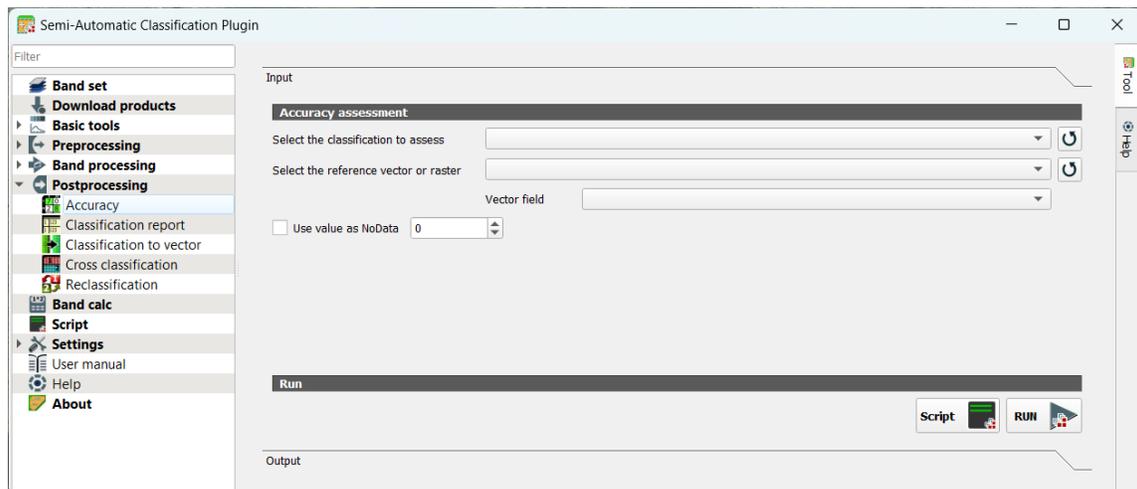
Las imágenes satelitales consultadas a través de la herramienta Google Earth Pro, que forman parte de la empresa Maxar y Airbus, cuentan con una resolución espacial de hasta 30 cm, lo que las convierte en un recurso valioso para esta etapa de la investigación. Una de las principales ventajas de usar Google Earth Pro es la posibilidad de acceder a imágenes de periodos pasados. En este proyecto, se consultaron imágenes de enero y abril de los años 2016 y 2017, todas proporcionadas por Maxar, con una cobertura mínima de nubes y una temporalidad de la vegetación similar a la clasificada previamente en GEE (Anexo 3).

Para la evaluación de los sitios de verificación correspondientes al año 2022, se utilizaron imágenes de Airbus y Maxar, capturadas en abril de 2022 y 2023. Además, se llevaron a cabo verificaciones en campo a través de puntos de control, con el objetivo de complementar los resultados obtenidos en el proyecto (Anexo 4 a 8).

La tercera etapa del proceso consistió en el análisis de los datos, el cual se llevó a cabo utilizando el software QGIS 3.22 y el complemento **Semi-Automatic Classification Plugin** (SCP). Este plugin permite realizar clasificación supervisada de imágenes satelitales. El objetivo de SCP es ofrecer un conjunto de herramientas integradas para el procesamiento de datos ráster, creando un flujo de trabajo automatizado que facilita la clasificación de la cobertura del suelo (Congedo, 2021).

Para la tercera etapa, se usa específicamente la etapa de postprocesamiento de SCP (Accuracy o precisión), en donde se requiere como insumos la imagen previamente clasificada en GEE y los sitios de verificación que anteriormente han sido verificados. Esta pestaña permite la validación de la clasificación donde se compara con los sitios de verificación. Accuracy permite calcular varias estadísticas, como la precisión general, precisión del usuario, precisión del productor y coeficiente Kappa (Figura 3.4).

Figura 3.4. Interfaz de Semi-Automatic Classification Plugin en QGIS



Fuente: Terrset 2020 Geospatial Monitoring and Modeling Software

A continuación, se muestran los resultados obtenidos para la Fase 2. Evaluación de la confiabilidad temática:

3.1.1 Random Forest en imagen Landsat 8 de 2016

El análisis de la evaluación de la confiabilidad temática de la imagen satelital Landsat 8 de 2016 mediante el algoritmo RF se realizó a través de un proceso de verificación en campo (imágenes con mayor resolución espacial), asociando los sitios de verificación con las distintas categorías de uso de suelo. A continuación, se presentan los resultados del cruce entre los sitios de verificación y el mapa clasificado. La información presentada a continuación corresponde a la Tabla 3.2.

Para la categoría de bosque se evaluaron 67 sitios, obteniéndose una precisión de 92.54% tanto para la precisión del productor (PA) como para la precisión del usuario (UA). Aunque la mayoría de los sitios fueron clasificados correctamente, se identificaron errores de omisión en 3 sitios clasificados como agricultura y 2 como tular, los cuales en realidad pertenecían a bosque. Asimismo, se registraron errores de comisión en 4 sitios clasificados como tular y 1 como pastizal.

En la clase agricultura, se verificaron 75 sitios, 70 se clasificaron correctamente, mientras que 3 se asignaron erróneamente a zona urbana, 1 a tular y 1 a pastizal. La PA fue de 93.33%, lo que indica que la proporción de sitios de verificación se representaron correctamente en el mapa. Sin embargo, la UA fue de 60.34%, lo que indica que otras clases se clasificaron como agricultura y pertenecen a otras clases, un ejemplo de ello es el pastizal con 21 sitios mal clasificados, a esto se le conoce como error de comisión.

Para zona urbana se evaluaron 49 sitios, de los cuales 40 fueron clasificados correctamente, con la principal confusión en la categoría agricultura. La PA fue de 81.63%, lo que indica que, en gran medida, los sitios urbanos correspondieron a esta clase en el mapa; sin embargo, 9 sitios se clasificaron como agricultura. La UA fue de 93.02%, lo cual significa que 3 sitios pertenecían a la clase de agricultura, pero se asignaron erróneamente a esta clase.

Se verificaron 14 sitios en la categoría de cuerpos de agua, 13 fueron correctamente clasificados, obteniéndose una PA del 92.86% y una UA de 100%. La clasificación de esta clase mostró una excelente precisión, sin errores significativos registrados.

Tabla 3.2. RF en imagen Landsat 8 de 2016

Sitios de verificación	Mapa						
	Bosque	Agricultura	Zona Urbana	Cuerpos de agua	Tular	Pastizal	Total
Bosque	62	0	0	0	4	1	67
Agricultura	3	70	9	1	12	21	116
Zona Urbana	0	3	40	0	0	0	43
Cuerpos de agua	0	0	0	13	0	0	13
Tular	2	1	0	0	3	0	6
Pastizal	0	1	0	0	0	12	13
Total	67	75	49	14	19	34	258
PA [%]	92.5373	93.3333	81.6327	92.8571	15.7895	35.2941	
UA [%]	92.5373	60.3448	93.0233	100	50	92.3077	
Precisión general							77.5194
Coeficiente Kappa							0.7042

Fuente: Elaboración propia

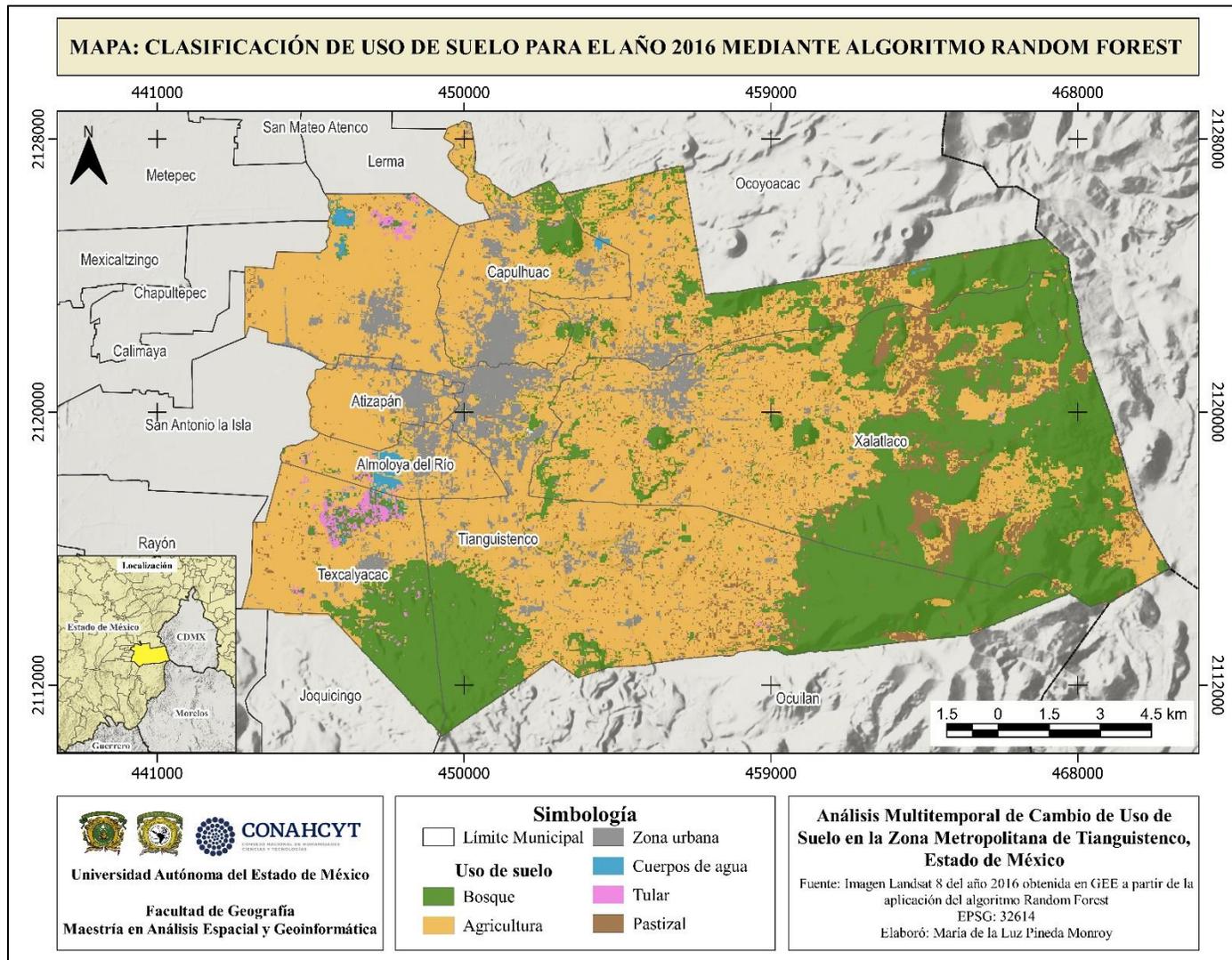
De los 19 sitios examinados, solo 3 fueron correctamente clasificados como tular. El resto de los sitios fue asignado erróneamente en otras categorías, principalmente en bosque (4) y agricultura (12). Esto resultó en una PA del 15.79% y una UA del 50%, lo que revela serios problemas de precisión en la clasificación de esta clase, con una alta confusión con otras categorías.

Finalmente, en la clase de pastizal se verificaron 34 sitios. De estos, 12 fueron correctamente clasificados como pastizal, pero se observó una fuerte confusión con agricultura, ya que 21 sitios de esta categoría fueron clasificados como pastizal. La PA fue del 35.29%, mientras que la UA fue del 92.31%, lo que indica que la mayoría de los sitios clasificados como pastizal provienen de esa clase, pero hubo una alta confusión desde otras clases, especialmente con agricultura.

La precisión global fue del 77.52%, lo que indica un nivel de precisión razonable en la clasificación. El coeficiente Kappa fue de 0.7042, lo que refleja una concordancia considerable. Si bien las clases de bosque, zona urbana y cuerpos de agua mostraron buena precisión, hubo confusión en las clases de agricultura, tular y pastizal, destacando áreas clave para mejorar en futuros procesos de clasificación a partir del algoritmo RF.

En la Figura 3.5 se puede observar el resultado de la clasificación de la imagen Landsat 8 del año 2016, utilizando el algoritmo Random Forest. Se destacan áreas en color azul y fucsia que representan la Ciénega del Lerma, donde algunos sectores fueron clasificados incorrectamente como bosque. Este error se debe, en gran medida, a la similitud en la reflectividad del infrarrojo cercano, lo cual provoca confusión entre clases y genera una sobreclasificación de la cobertura forestal sobre el tular. Otra clase con un sesgo considerable en la clasificación fue el pastizal, particularmente en la zona este, donde las condiciones orográficas limitan la viabilidad de la agricultura.

Figura 3.5. Clasificación de uso de suelo para el año 2016 mediante Random Forest



Fuente: Elaboración propia a partir de imagen Landsat 8 del año 2016 y algoritmo RF

3.1.2 Support Vector Machine en imagen Landsat 8 de 2016

Se estimó la confiabilidad temática de la clasificación de uso de suelo obtenida a partir de una imagen Landsat 8 de 2016 empleando el algoritmo SVM. Para ello, se compararon los resultados de la clasificación con datos de imágenes de satélite de mejor resolución espacial. Los resultados se presentan en la Tabla 3.3 y se detallan a continuación:

De los 67 sitios evaluados, 63 fueron correctamente clasificados como bosque. Se observaron errores de omisión en 3 sitios, que fueron mal clasificados como agricultura, y en 1 sitio identificado como zona urbana. Además, se registraron errores de comisión en las clases de agricultura, tular y pastizal. Los resultados mostraron una PA de 94.03% y una UA de 86.30%, lo que indica una buena precisión en la clasificación de esta categoría.

En la categoría de agricultura se verificaron 75 sitios en total. De estos, 68 fueron correctamente clasificados como agricultura, mientras que 2 fueron clasificados erróneamente en bosque, 2 en zona urbana, 2 como tular y 1 en pastizal. Se evidencia error de comisión en el pastizal. La PA fue de 90.67% y la UA de 76.40%, reflejando una buena concordancia desde el punto de vista del productor.

En la clase zona urbana, se verificaron 49 sitios, de los cuales 45 fueron correctamente clasificados. El error de omisión se presentó en agricultura con 4 sitios y error de comisión en bosque (1), agricultura (2) y pastizal (1). La PA fue de 91.84% y la UA de 91.84%, lo que muestra una alta precisión en dicha categoría.

Para cuerpos de agua, los 14 sitios de verificación fueron correctamente clasificados. Esto resultó en una PA del 100% y UA del 93.33%, lo que indica una clasificación muy precisa en esta categoría.

Tabla 3.3. SVM en imagen Landsat 8 de 2016

Mapa							
Sitios de verificación	Bosque	Agricultura	Zona Urbana	Cuerpos de agua	Tular	Pastizal	Total
Bosque	63	2	0	0	5	3	73
Agricultura	3	68	4	0	2	12	89
Zona Urbana	1	2	45	0	0	1	49
Cuerpos de agua	0	0	0	14	1	0	15
Tular	0	2	0	0	11	0	13
Pastizal	0	1	0	0	0	18	19
Total	67	75	49	14	19	34	258
PA [%]	94.0299	90.67	91.84	100	57.8947	52.9412	
UA [%]	86.3014	76.4045	91.84	93.3333	84.6154	94.74	
Precisión general							84.8837
Coeficiente Kappa							0.8043

Fuente: Elaboración propia

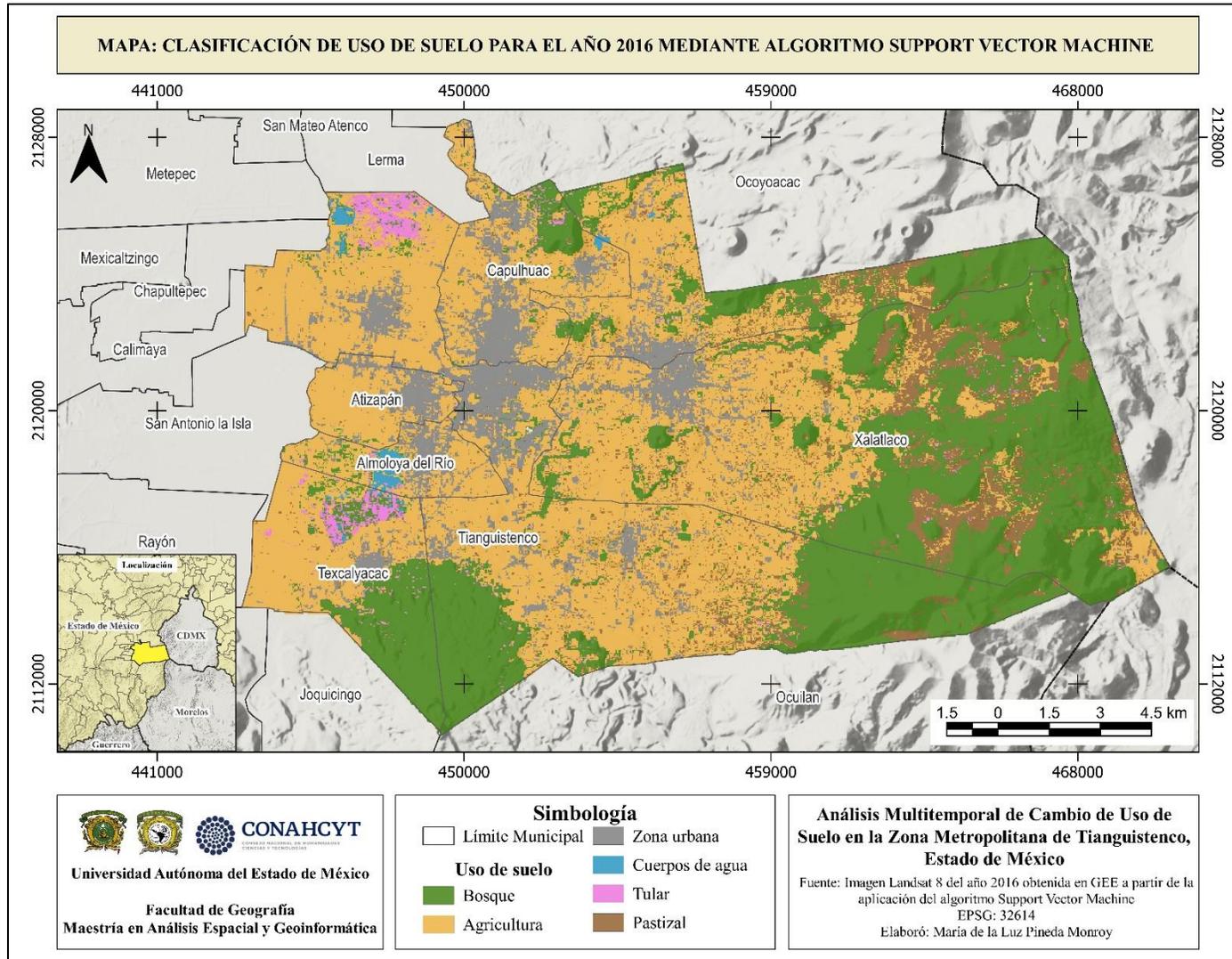
Continuando con la descripción de la Tabla 3.3, en la clase tular, se verificaron 19 sitios, de los cuales 11 fueron correctamente clasificados. Sin embargo, como error de omisión, 5 sitios se categorizaron como bosque, 2 sitios como agricultura y 1 como cuerpo de agua. Esto dio como resultado una PA de 57.89% y una UA de 84.62%, lo que sugiere una precisión moderada en la clasificación de tular y confusión hacia agricultura.

Finalmente, en la clase pastizal se verificaron 34 sitios, de los cuales 18 fueron correctamente clasificados como pastizal. Se observó confusión con la categoría de agricultura, ya que 12 sitios de esta clase fueron clasificados erróneamente, cuando realmente pertenecían a esta clase. La PA fue de 52.94% y la UA de 94.74%.

La precisión global fue del 84.88%, lo que indica un alto nivel de precisión en la clasificación general. El coeficiente Kappa fue de 0.8043, lo que refleja una concordancia considerable de acuerdo con Landis y Koch (1977). Los cuerpos de agua, la zona urbana y bosque obtuvieron los mejores resultados en cuanto a clasificación, siendo la clase tular y pastizal los que obtuvieron valores más bajos en cuanto a PA.

La Figura 3.6 muestra la clasificación de uso de suelo de la imagen Landsat 8 de 2016, obtenida mediante el algoritmo SVM. Al igual que en la clasificación con RF, el bosque presenta una sobreclasificación respecto al tular, un efecto que se observa especialmente cerca de los cuerpos de agua y áreas de tular en las Ciénegas del Lerma, ubicadas en los municipios de Almoloya del Río y Texcalyacac, que forman parte de la ZMT. Por otro lado, el pastizal, aunque mostró problemas en la PA, alcanzó una buena UA, lo cual se refleja en el mapa, especialmente en la zona montañosa del municipio de Xalatlaco.

Figura 3.6. Clasificación de uso de suelo para el año 2016 mediante SVM



Fuente: Elaboración propia a partir de imagen Landsat 8 del año 2016 y algoritmo SVM

3.1.3 Random Forest en imagen Landsat 8 de 2022

El análisis de la imagen Landsat 8 para el año 2022, utilizando el algoritmo RF, se realizó a partir de la creación de una matriz de confusión para evaluar la precisión de la clasificación en las clases de uso de suelo. A continuación, se presenta la descripción de cada clase, considerando el total de sitios verificados por categoría, los datos presentados a continuación se encuentran en la Tabla 3.4.

Para la clase de bosque, se verificaron 61 sitios, de los cuales 54 fueron correctamente clasificados, mientras que 7 se asignaron erróneamente a agricultura, lo cual representa al error de omisión. Por otro lado, se registraron errores de comisión, con 8 sitios clasificados como tular, 3 en agricultura, y 1 como cuerpo de agua y pastizal, respectivamente. Esto resultó en una PA de 88.52%, lo que indica una alta concordancia entre los sitios clasificados como bosque en el mapa y la realidad. Sin embargo, la UA fue del 80.60%.

En la categoría de agricultura, se verificaron 77 sitios. De estos, 72 fueron correctamente clasificados en agricultura. Sin embargo, hubo confusión con otras clases, 3 sitios designados como bosque, 1 en asentamientos humanos y 1 como pastizal. Esto resultó en una PA de 93.51%, lo que refleja una alta precisión para el productor. Sin embargo, la UA fue de 58.54%, mostrando una considerable confusión de otros sitios, a lo que se conoce como error de comisión (especialmente pastizales y asentamientos humanos).

Para la clase de asentamientos humanos se verificaron 46 sitios, de estos, 28 fueron correctamente clasificados, la principal confusión se encontró con agricultura con 18 sitios mal clasificados. La PA fue de 60.87%, lo que refleja una precisión moderada para la categoría. Por otro lado, la UA fue muy alta, alcanzando un 96.55%, lo que significa que la mayoría de los sitios clasificados como asentamientos humanos eran correctos.

Se verificaron 13 sitios de cuerpos de agua, 9 se clasificaron correctamente, el error de omisión se encontró en agricultura (2), bosque (1) y tular (1).

Tabla 3.4. RF en imagen Landsat 8 de 2022

Mapa							
Sitios de verificación	Bosque	Agricultura	Asentamientos humanos	Cuerpos de agua	Tular	Pastizal	Total
Bosque	54	3	0	1	8	1	67
Agricultura	7	72	18	2	2	22	123
Asentamientos humanos	0	1	28	0	0	0	29
Cuerpos de agua	0	0	0	9	1	0	10
Tular	0	0	0	1	5	0	6
Pastizal	0	1	0	0	0	25	26
Total	61	77	46	13	16	48	261
PA [%]	88.5246	93.51	60.87	69.2308	31.25	52.0833	
UA [%]	80.60	58.54	96.5517	90	83.3333	96.1538	
Precisión general							73.95
Coeficiente Kappa							0.6570

Fuente: Elaboración propia.

La PA para cuerpos de agua fue de 69.23%, lo que indica un desempeño moderado en esta categoría. Sin embargo, la UA fue del 90%, lo que significa que la mayoría de los sitios clasificados como cuerpos de agua eran correctos, aunque se encontró confusión con tular (Tabla 3.4).

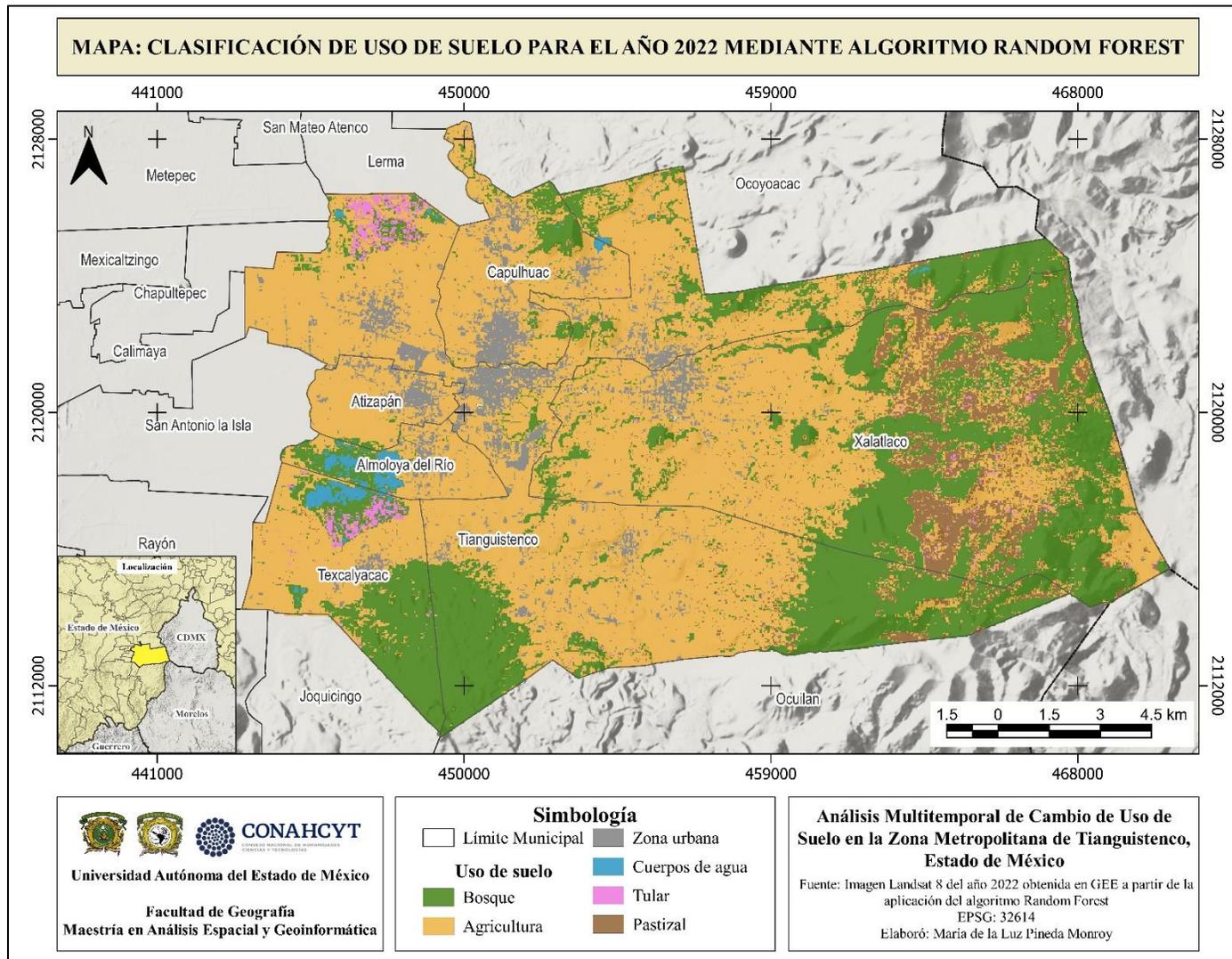
En la clase de tular, se verificaron 16 sitios, de los cuales solo 5 fueron correctamente clasificados. Hubo una confusión significativa con varios sitios clasificados erróneamente en otras clases, como bosque, agricultura y cuerpos de agua. La PA fue de 31.25%, lo que refleja un bajo nivel de precisión en esta categoría. Sin embargo, la UA fue del 83.33%, lo que indica que la mayoría de los sitios clasificados como tular eran correctos, salvo una excepción con cuerpos de agua (error de comisión).

Se verificaron un total de 48 sitios en la clase de pastizal, de los cuales 25 se clasificaron correctamente. Sin embargo, hubo una notable confusión con otros sitios, principalmente con agricultura (22 sitios clasificados erróneamente como pastizal). La PA fue de 52.08%, lo que sugiere un rendimiento bajo para el productor mientras que la UA fue muy alta, con un 96.15%, lo que indica que la mayoría de los sitios clasificados como pastizal eran correctos, aunque hubo confusión desde otra clase.

La precisión general de la clasificación fue del 73.95%, lo que indica un desempeño aceptable del modelo Random Forest para esta clasificación. El coeficiente Kappa fue de 0.6570, lo que sugiere una concordancia moderada entre la clasificación del mapa y los datos de verificación de campo. Las clases con mejor rendimiento en PA fueron el bosque y agricultura y UA en asentamientos humanos y pastizal.

La Figura 3.7 muestra los resultados de la clasificación de la imagen Landsat 8 de 2022 empleando el algoritmo Random Forest. El desempeño del clasificador es similar al observado en la imagen de 2016, destacándose que el bosque se asignó en exceso sobre áreas de tular. De igual forma, la clase de agricultura fue etiquetada sobre áreas urbanas, lo cual puede confirmarse mediante el análisis de los sitios de verificación.

Figura 3.7. Clasificación de uso de suelo para el año 2022 mediante RF



Fuente: Elaboración propia a partir de imagen Landsat 8 del año 2022 y algoritmo RF

3.1.4 Support Vector Machine en imagen Landsat 8 de 2022

El análisis de la imagen Landsat 8 de 2022, clasificada mediante el algoritmo SVM, se evaluó utilizando una matriz de confusión que compara los sitios de verificación con la clasificación del mapa, junto con el coeficiente Kappa. A continuación en la Tabla 3.5 se detalla el desempeño de cada categoría, considerando el total de sitios verificados en cada clase.

En total, se evaluaron 61 sitios en la categoría bosque. De estos, 58 se asignaron correctamente, mientras que 4 se etiquetaron erróneamente en agricultura, 1 en tular y 2 en pastizal, lo cual constituye un error de omisión. Este resultado produjo una PA de 88.52%, reflejando una alta concordancia entre los sitios clasificados como bosque en el mapa y la realidad. Sin embargo, la UA fue de 73.97%, lo que indica que una cantidad significativa de sitios fueron clasificados en una clase a la que no pertenecen realmente.

En la clase de agricultura, se verificaron 77 sitios, de los cuales 66 fueron correctamente clasificados. Sin embargo, 5 sitios se asignaron erróneamente como bosque, 3 como zona urbana y 3 como pastizal. Esto resultó en una PA de 85.71%, reflejando una precisión relativamente alta para el productor, mientras que la UA fue de 72.53%, lo que muestra una notable confusión con otras clases, principalmente zona urbana y pastizal.

En la categoría de zona urbana, se verificaron 46 sitios en total, de los que 35 fueron clasificados correctamente. Hubo confusión con 10 sitios asignados a agricultura y 1 a bosque. La PA fue de 76.09%, lo que sugiere una precisión moderada, mientras que la UA fue de 92.11%, donde 3 sitios se clasificaron en agricultura (error de comisión).

Para la clase de cuerpos de agua, se verificaron 13 sitios, de los cuales 10 fueron correctamente clasificados. Hubo cierta confusión con 2 sitios clasificados como bosque y 1 sitio de tular clasificado erróneamente como cuerpos de agua. La PA fue de 76.92%, mientras que la UA alcanzó el 90.91%, lo que refleja una clasificación precisa, con una pequeña confusión hacia la clase de tular.

Tabla 3.5. SVM en imagen Landsat 8 de 2022

Sitios de verificación	Mapa						
	Bosque	Agricultura	Zona Urbana	Cuerpos de agua	Tular	Pastizal	Total
Bosque	54	5	1	2	6	5	73
Agricultura	4	66	10	0	0	11	91
Zona Urbana	0	3	35	0	0	0	38
Cuerpos de agua	0	0	0	10	1	0	11
Tular	1	0	0	1	9	1	12
Pastizal	2	3	0	0	0	31	36
Total	61	77	46	13	16	48	261
PA [%]	88.5246	85.7143	76.09	76.9231	56.25	64.5833	
UA [%]	73.9726	72.5275	92.11	90.91	75	86.1111	
Precisión general							78.5441
Coeficiente Kappa							0.7234

Fuente: Elaboración propia

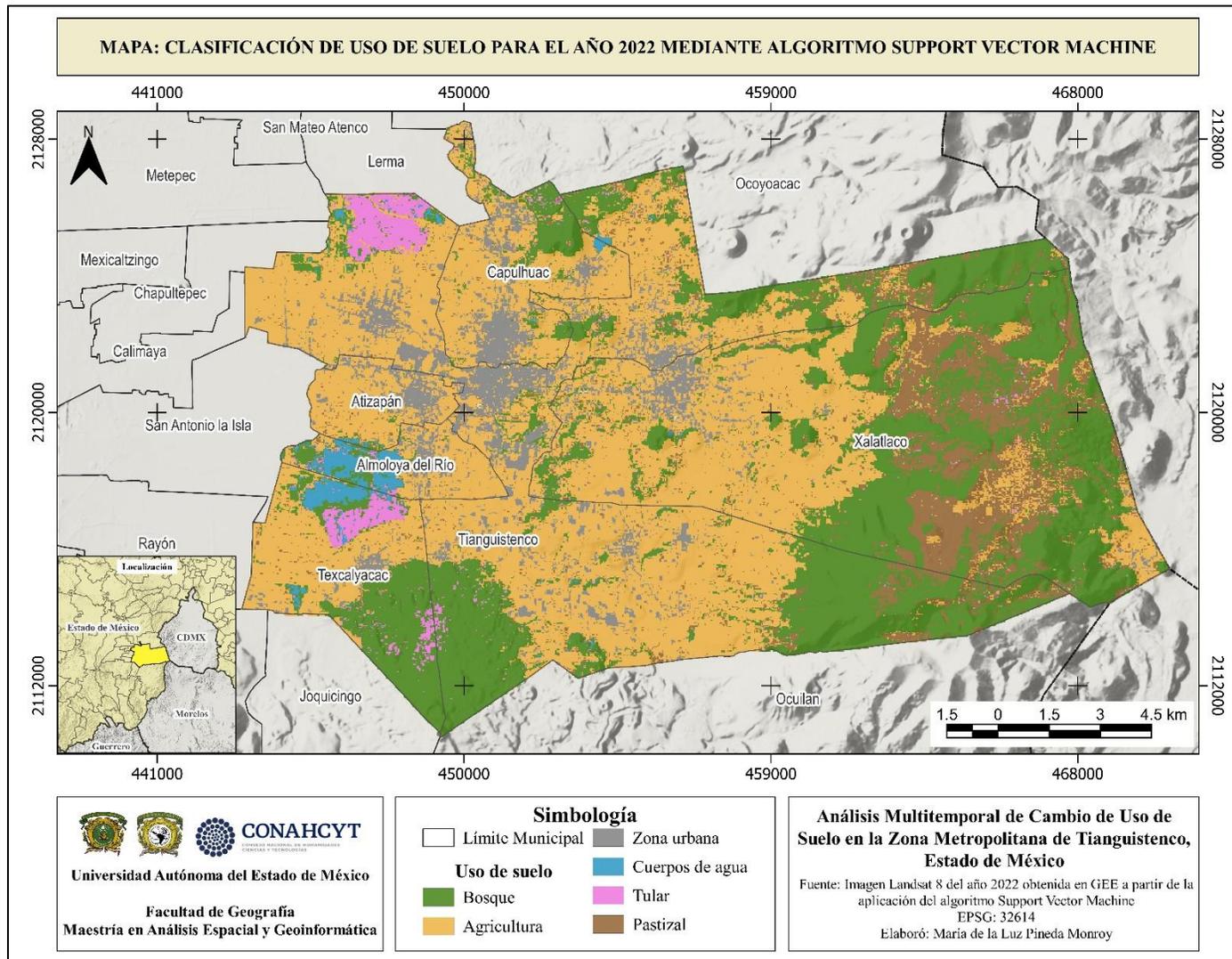
Continuando con la descripción de la Tabla 3.5. De los 16 sitios evaluados como tular, solo 9 fueron clasificados correctamente. Se observó confusión con el bosque (6 sitios) y con cuerpos de agua (1 sitio). La PA fue de 56.25%, indicando un rendimiento bajo del productor. Sin embargo, la UA fue de 75%, lo que sugiere que la mayoría de los sitios clasificados como tular eran correctos.

Se comprobaron un total de 48 sitios en la clase pastizal, de los cuales 31 fueron correctamente clasificados. Hubo una confusión significativa con 11 sitios en agricultura y bosque (5). La PA fue de 64.58%, mientras que la UA fue de 86.11%, lo que indica que, aunque la mayoría de los sitios clasificados como pastizal eran correctos, hubo confusión desde otras clases, especialmente con agricultura, tanto en el error de omisión y comisión.

La precisión general de la clasificación fue del 78.54%, lo que indica un rendimiento aceptable del modelo SVM en esta imagen. El coeficiente Kappa fue de 0.7234, lo que refleja una concordancia considerable entre la clasificación del mapa y los datos de verificación de campo. Las clases con mejor rendimiento fueron zona urbana y cuerpos de agua, mientras que tular y pastizal presentaron mayores confusiones, particularmente con las clases de agricultura y bosque. Esto sugiere que, aunque la clasificación en general es precisa, hay áreas específicas que necesitan mejoras para reducir la confusión entre clases de uso del suelo.

La Figura 3.8 muestra el resultado de la clasificación de la imagen Landsat 8 del año 2022 utilizando el algoritmo Support Vector Machine. Con base en la matriz de tabulación cruzada y el mapa de uso de suelo, se observa una sobreasignación de bosque en el área de la Ciénega del Lerma, ubicada en el municipio de Almoloya del Río. Además, en la zona boscosa de Texcalyacac se identificó tular, una vegetación que no es propia de los bosques en esta región. Asimismo, se detectó una sobreclasificación de agricultura en la zona urbana.

Figura 3.8. Clasificación de uso de suelo para el año 2022 mediante SVM



Fuente: Elaboración propia a partir de imagen Landsat 8 del año 2022 y algoritmo SVM

3.1.5 Random Forest en imagen Sentinel-2 de 2016

Producto del análisis de la imagen satelital Sentinel-2 para el año 2016 mediante el algoritmo RF, se evaluó la matriz de confusión para las categorías de uso de suelo. A continuación, se describen los resultados para cada categoría, información contenida en la Tabla 3.6.

De los 70 sitios identificados como bosque en el mapa, 68 coincidieron con la clasificación de campo (imágenes de mayor resolución espacial), alcanzando una PA de 97.14%. Sin embargo, se detectaron 2 confusiones con tular. La UA fue de 88.31%, ya que 7 sitios fueron clasificados erróneamente en esta clase cuando en realidad correspondían a tular y 2 a agricultura, lo que se conoce como error de comisión.

Para la clase de agricultura, se verificaron un total de 73 sitios clasificados como agricultura en el mapa. De estos, 69 fueron clasificados correctamente, mientras que 2 sitios se clasificaron de manera incorrecta como bosque y 2 en pastizal. Esto dio lugar a una PA de 94.52%, lo que indica una alta precisión del productor. Sin embargo, la UA fue del 69.70%, lo que revela una considerable confusión de otros sitios, especialmente de pastizal, que fueron mal clasificados como agricultura.

Se evaluaron 47 sitios para la clase urbana, de los cuales 43 fueron clasificados correctamente, mientras que 4 se asignaron a agricultura. Esto resultó en una PA de 91.49% y una UA de 100%, lo que indica que la clasificación de zonas urbanas fue precisa en cuanto al error de comisión.

Para la categoría de cuerpos de agua, se evaluaron 15 sitios clasificados como tales en el mapa. Todos fueron clasificados correctamente, lo que resultó en una PA del 100%, reflejando la ausencia de error de omisión. La UA fue de 88.24%, con cierta confusión hacia la clase de tular.

Se verificaron un total de 19 sitios clasificados como tular en el mapa. De estos, 7 fueron correctamente clasificados, mientras que 7 sitios se identificaron incorrectamente como bosque, 3 como agricultura y 2 como cuerpos de agua.

Tabla 3.6. RF en imagen Sentinel-2 de 2016

Mapa							
Sitios de verificación	Bosque	Agricultura	Zona Urbana	Cuerpos de agua	Tular	Pastizal	Total
Bosque	68	2	0	0	7	0	77
Agricultura	0	69	4	0	3	23	99
Zona Urbana	0	0	43	0	0	0	43
Cuerpos de agua	0	0	0	15	2	0	17
Tular	2	0	0	0	7	1	10
Pastizal	0	2	0	0	0	11	13
Total	70	73	47	15	19	35	259
PA [%]	97.1429	94.5205	91.49	100	36.8421	31.4286	
UA [%]	88.3117	69.70	100	88.24	70	84.62	
Precisión general							82.2394
Coeficiente Kappa							0.7687

Fuente: Elaboración propia

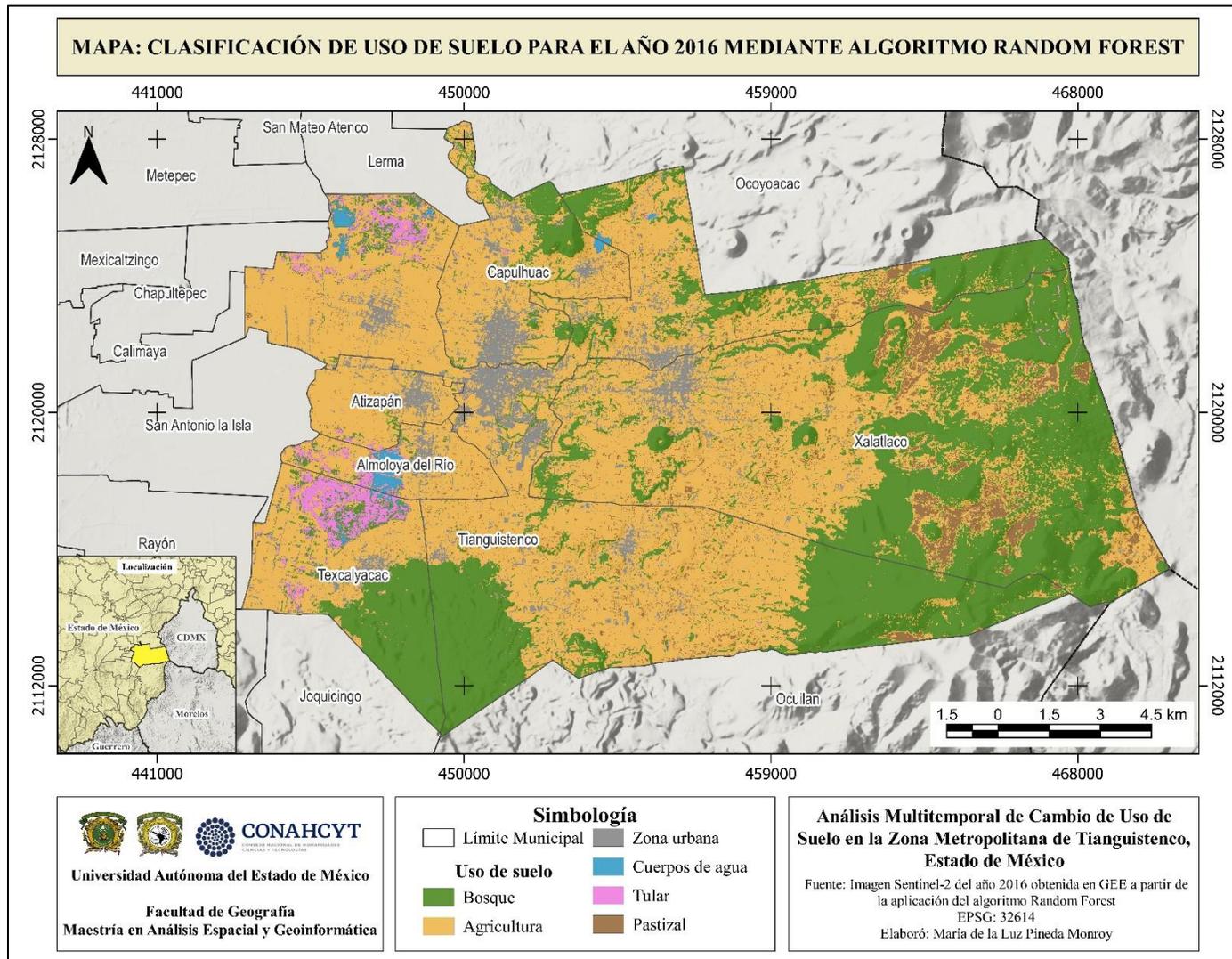
Continuando con la descripción de la Tabla 3.6 y la descripción de tular. Esto dio lugar a una PA del 36.84%, lo que refleja una baja precisión en la clasificación de tular (error de omisión). La UA fue del 70%, lo que indica que la mayoría de los sitios clasificados como tular eran correctos, aunque hubo una confusión considerable con otras clases, como bosque y pastizal.

En la clase de pastizal, se verificaron un total de 35 sitios en el mapa. De estos, 11 fueron correctamente clasificados, mientras que 23 sitios fueron mal clasificados como agricultura. Esto resultó en una PA de 31.43%, lo que indica una baja precisión para el productor. La UA fue del 84.62%, lo que refleja que la mayoría de los sitios clasificados como pastizal en el mapa eran en realidad agricultura, lo que se conoce como error de comisión.

La precisión general de la clasificación fue del 82.24%, lo que indica un buen rendimiento del algoritmo Random Forest en esta imagen Sentinel-2 de 2016. El coeficiente Kappa fue de 0.7687, lo que refleja una concordancia considerable entre la clasificación del mapa y los datos de verificación de campo. Las clases con mejor rendimiento fueron zona urbana, bosque y cuerpos de agua, mientras que tular y pastizal presentaron mayor confusión, especialmente con agricultura.

La Figura 3.9 muestra el resultado de la clasificación de la imagen Sentinel-2 de 2016 empleando el algoritmo Random Forest. Al igual que en los mapas anteriores, se observa una sobreclasificación del bosque sobre tular en la Ciénega del Lerma, específicamente en la zona comprendida entre los municipios de Almoloya del Río y Atizapán. Los cuerpos de agua y las zonas urbanas se encuentran entre las clases mejor representadas en el mapa; estas últimas se identifican principalmente en la porción central del área de estudio.

Figura 3.9. Clasificación de uso de suelo para el año 2016 mediante RF



Fuente: Elaboración propia a partir de imagen Sentinel-2 del año 2016 y algoritmo RF

3.1.6 Support Vector Machine en imagen Sentinel-2 de 2016

El análisis de la imagen Sentinel-2 del año 2016, clasificada con el algoritmo SVM, se evaluó mediante la matriz de confusión y el coeficiente Kappa. A continuación, se detalla el rendimiento de cada clase de uso de suelo, considerando la PA y UA, como se muestra en la Tabla 3.7.

Se evaluaron un total de 70 sitios clasificados como bosque en el mapa. De estos, 66 fueron asignados correctamente a esta clase, mientras que 3 se clasificaron erróneamente como agricultura y 1 como tular. Esto dio como resultado una PA de 94.29%, lo que indica una alta concordancia entre los sitios clasificados como bosque en el mapa y los verificados en campo. La UA fue de 91.67%, reflejando una alta probabilidad de que los sitios de esta clase en el mapa correspondan con su categoría real en el terreno.

Para la clase de agricultura, se verificaron 73 sitios en total. De estos, 70 fueron correctamente clasificados, mientras que 3 se asignaron mal a la clase pastizal. Esto resultó en una PA de 95.89%, reflejando una alta precisión en la clasificación de agricultura. La UA fue del 81.40%, lo que indica una notable confusión con pastizal y, en menor medida, con zona urbana.

En la categoría de zona urbana, se verificaron 47 sitios, de los cuales 44 fueron correctamente clasificados como zona urbana. De los cuales, 3 fueron clasificados como agricultura. Lo anterior dio como resultado una PA de 93.62% y una UA de 100%, indicando la ausencia de error de comisión.

En la clase de cuerpos de agua, se verificaron 15 sitios, de los cuales 15 fueron correctamente clasificados como cuerpos de agua. Esto resultó en una PA de 100%, lo que indica que todos los sitios clasificados como cuerpos de agua en el mapa coinciden con los datos de verificación de campo. La UA fue del 93.75%, donde encontró confusión con la clase tular.

Tabla 3.7. SVM en imagen Sentinel-2 de 2016

Sitios de verificación	Mapa						
	Bosque	Agricultura	Zona Urbana	Cuerpos de agua	Tular	Pastizal	Total
Bosque	66	0	0	0	4	2	72
Agricultura	3	70	3	0	1	9	86
Zona Urbana	0	0	44	0	0	0	44
Cuerpos de agua	0	0	0	15	1	0	16
Tular	1	0	0	0	13	0	14
Pastizal	0	3	0	0	0	24	27
Total	70	73	47	15	19	35	259
PA [%]	94.29	95.8904	93.62	100	68.4211	68.5714	
UA [%]	91.67	81.40	100	93.75	92.86	88.89	
Precisión general							89.58
Coeficiente Kappa							0.8654

Fuente: Elaboración propia

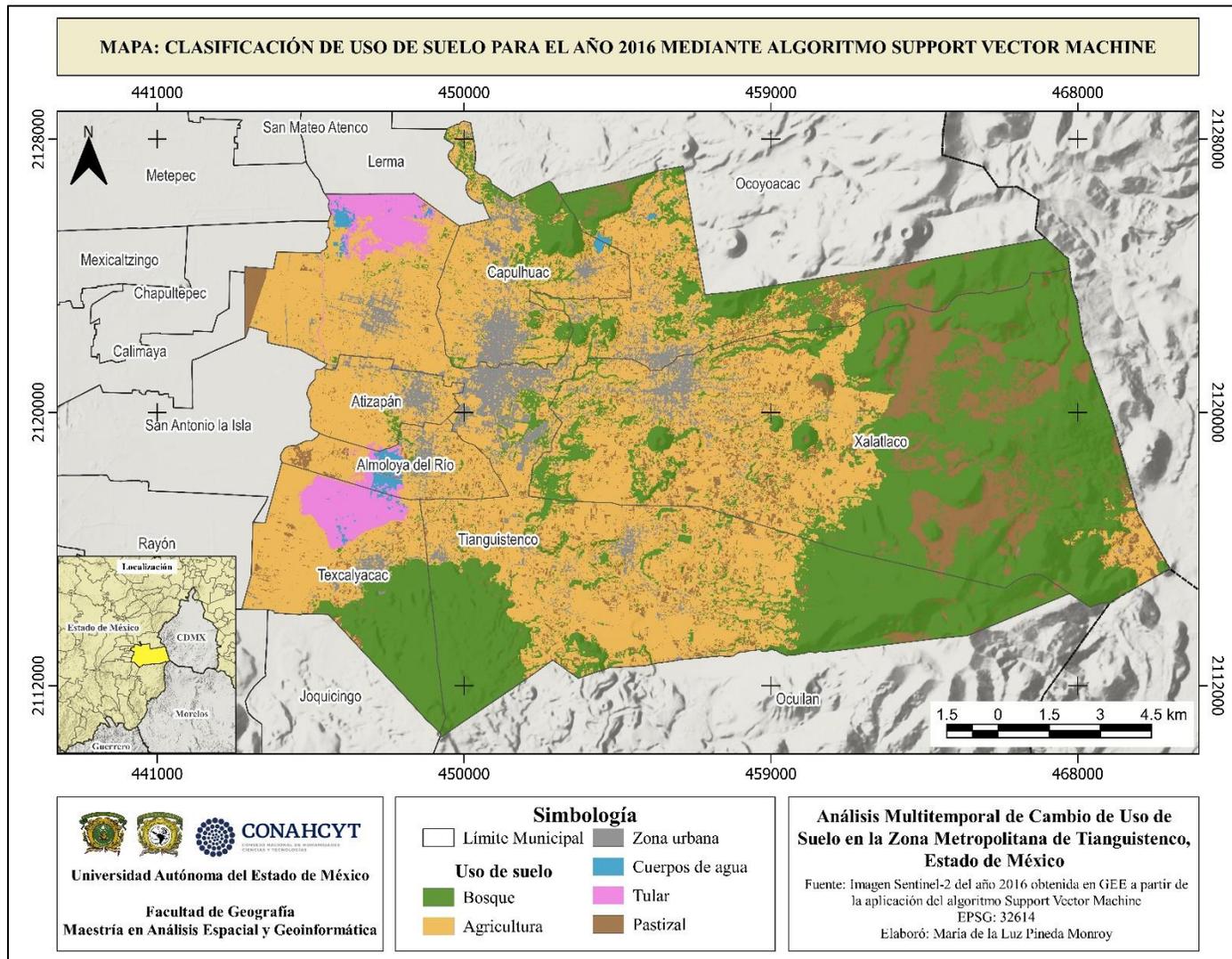
Continuando con la descripción de la Tabla 3.7. 19 sitios fueron evaluados para la categoría tular, en ellos, 13 fueron adecuadamente clasificados como tular, mientras que 4 sitios se asignaron a bosque, 1 a agricultura y 1 como zona urbana. Esto resultó en una PA de 68.42%, lo que refleja una precisión moderada para el productor. Sin embargo, la UA fue del 92.86%, lo que indica que la mayoría de los sitios clasificados como tular en el mapa fueron correctos.

Para la categoría de pastizal se verificaron 35 sitios en total. De estos, 24 fueron correctamente clasificados, mientras que 9 se asignaron incorrectamente a agricultura y 2 a bosque. Esto resultó en una PA de 68.57%, indicando una precisión moderada para el productor. La UA fue del 88.89%, reflejando una alta probabilidad de que los sitios clasificados como pastizal en el mapa coincidan con la realidad en el terreno.

La precisión general de la clasificación fue del 89.58%, lo que indica un alto rendimiento del modelo SVM en esta imagen. El coeficiente Kappa fue de 0.8654, lo que refleja una concordancia considerable entre la clasificación del mapa y los datos de verificación de campo. Las clases con mejor rendimiento fueron zona urbana, bosque y cuerpos de agua, mientras que tular y pastizal presentaron confusión, particularmente con agricultura.

La Figura 3.10 muestra los resultados de la clasificación de la imagen Sentinel-2 de 2016 mediante el algoritmo Support Vector Machine. La clasificación obtenida es de alta calidad, permitiendo una clara diferenciación entre las clases de bosque (verde), pastizal (café) y cuerpos de agua (azul).

Figura 3.10. Clasificación de uso de suelo para el año 2016 mediante SVM



Fuente: Elaboración propia a partir de imagen Sentinel-2 del año 2016 y algoritmo SVM

3.1.7 Random Forest en imagen Sentinel-2 de 2022

El análisis de la imagen Sentinel-2 del año 2022, clasificada mediante el algoritmo RF, se evaluó utilizando una matriz de confusión que comparó los sitios de verificación con la clasificación del mapa. A continuación, la Tabla 3.8 presenta el rendimiento de cada clase de uso de suelo, destacando la precisión del productor, la precisión del usuario y el coeficiente Kappa.

Se evaluaron un total de 64 sitios clasificados como bosque en el mapa. De estos, 62 fueron correctamente clasificados, mientras que 2 se asignaron incorrectamente a la clase de agricultura. Esto resultó en una PA de 96.88%, lo que indica una alta concordancia entre los sitios clasificados como bosque en el mapa y los datos de verificación en campo. La UA fue del 78.48%, reflejando cierta confusión debido al error de comisión; es decir, varios sitios de otras clases, como tular (11), agricultura (4), zona urbana y cuerpos de agua (1), se asignaron incorrectamente a bosque.

Para la clase de agricultura se evaluaron un total de 77 sitios, 72 fueron correctamente clasificados, mientras que 4 se asignaron erróneamente a bosque y 1 a pastizal. Esto resultó en una PA de 93.51%, lo que indica una buena representación de la clase en el mapa en comparación con los datos de campo. La UA fue del 80%, con una notable confusión hacia la clase de pastizal.

En la categoría de zona urbana se evaluaron 46 sitios en total. De estos, 40 fueron correctamente clasificados, mientras que 5 se asignaron erróneamente a agricultura y 1 a bosque. Esto resultó en una PA de 86.96% y una UA de 100%, lo que indica la ausencia de sitios clasificados en una clase a la que no pertenecen (error de comisión).

Para la clase de cuerpos de agua, se validaron 13 sitios en total. De estos, 12 fueron correctamente clasificados como cuerpos de agua, mientras que 1 sitio fue clasificado erróneamente como bosque. Esto dio lugar a una PA de 92.31%, lo que refleja una alta concordancia entre los sitios clasificados como cuerpos de agua en el mapa y los datos de verificación. La UA fue también del 92.31%, con un sitio mal asignado a tular.

Tabla 3.8. RF en imagen Sentinel-2 de 2022

Sitios de verificación	Mapa						
	Bosque	Agricultura	Zona Urbana	Cuerpos de agua	Tular	Pastizal	Total
Bosque	62	4	1	1	11	0	79
Agricultura	2	72	5	0	0	11	90
Zona Urbana	0	0	40	0	0	0	40
Cuerpos de agua	0	0	0	12	1	0	13
Tular	0	0	0	0	5	0	5
Pastizal	0	1	0	0	1	36	38
Total	64	77	46	13	18	47	265
PA [%]	96.88	93.51	86.96	92.31	27.7778	76.60	
UA [%]	78.481	80	100	92.31	100	94.74	
Precisión general							85.6604
Coeficiente Kappa							0.8151

Fuente: Elaboración propia

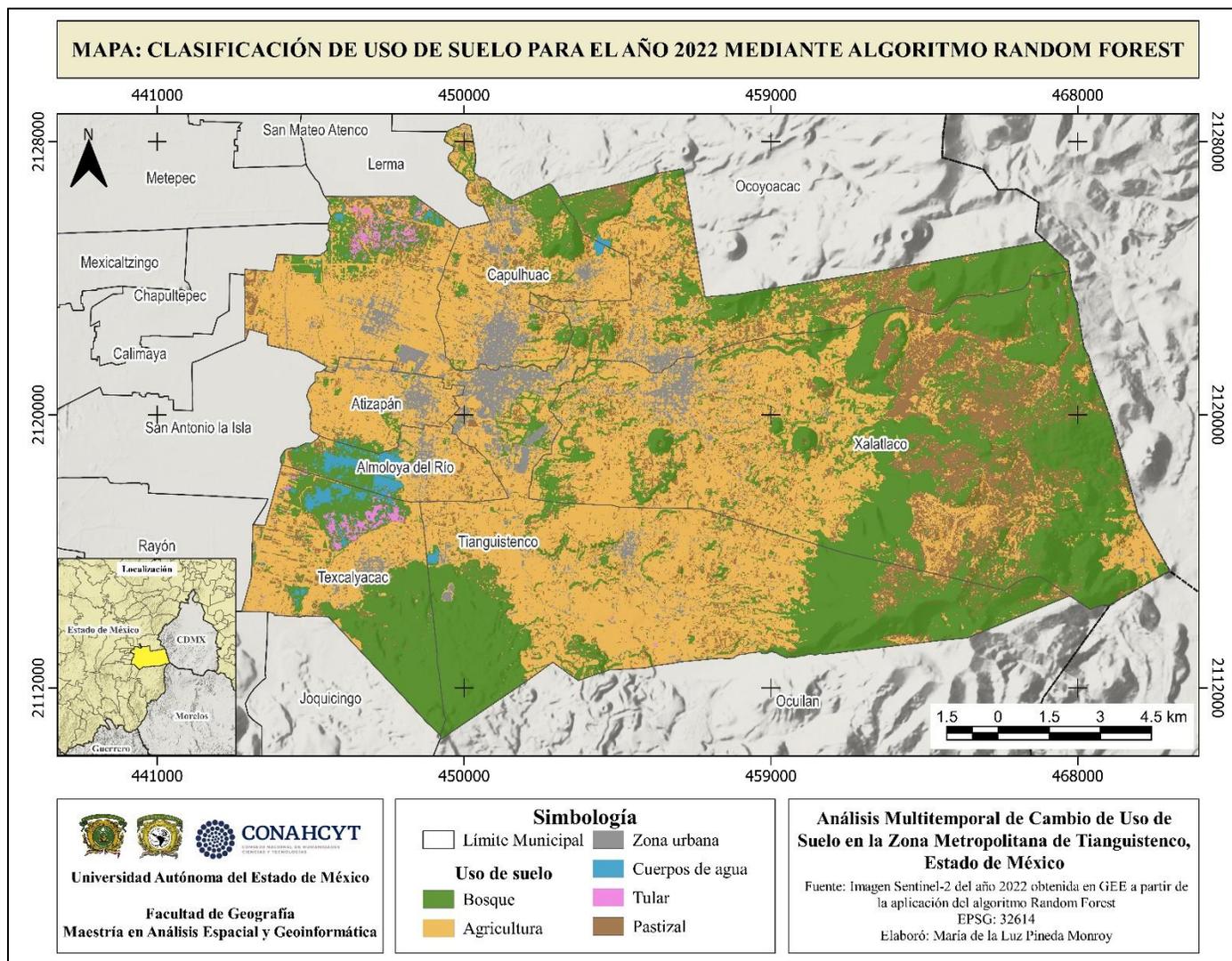
Continuando con la descripción de los resultados de la Tabla 3.8; se verificaron un total de 18 sitios clasificados como tular en el mapa. De estos, solo 5 fueron correctamente clasificados, mientras que 11 sitios se clasificaron erróneamente como bosque y 1 como cuerpos de agua. Esto resultó en una PA de 27.78%, lo que refleja una baja precisión en la clasificación de tular. Sin embargo, la UA fue del 100%.

En la clase de pastizal, se verificaron un total de 47 sitios. De estos, 36 fueron correctamente clasificados como pastizal, mientras que 11 sitios fueron clasificados erróneamente como agricultura. Esto dio lugar a una PA de 76.60%, lo que refleja una precisión aceptable para el productor. La UA fue del 94.74%, lo que indica una buena precisión para los sitios clasificados como pastizal, con confusión con agricultura.

La precisión general de la clasificación fue del 85.66%, lo que indica un buen rendimiento del modelo Random Forest en esta imagen Sentinel-2 de 2022. El coeficiente Kappa fue de 0.8151, lo que refleja una concordancia considerable entre la clasificación del mapa y los datos de verificación de campo. Las clases con mejor rendimiento fueron zona urbana y cuerpos de agua, mientras que tular presentó una mayor confusión, especialmente con la clase de bosque.

La Figura 3.11 presenta la clasificación de la imagen Sentinel-2 de 2022 utilizando el algoritmo Random Forest. Aunque la clasificación general indica una buena precisión, se observa un sesgo en la diferenciación entre tular y bosque, especialmente en la zona del cuerpo de agua ubicado entre Almoloya del Río y Texcalyacac. Además, la matriz de confusión indica una sobreclasificación de la cobertura agrícola sobre el pastizal en el área forestal de Xalatlaco.

Figura 3.11. Clasificación de uso de suelo para el año 2022 mediante RF



Fuente: Elaboración propia a partir de imagen Sentinel-2 del año 2022 y algoritmo RF

3.1.8 Support Vector Machine en imagen Sentinel-2 de 2022

El análisis de la imagen Sentinel-2 del año 2022, clasificada con el algoritmo SVM, se evaluó mediante una matriz de confusión que compara los sitios de verificación con la clasificación del mapa. A continuación, la Tabla 3.9 detalla el rendimiento de cada categoría y el coeficiente Kappa.

Para la clase de bosque, se evaluaron 64 sitios, entre ellos 63 fueron correctamente clasificados y 1 fue asignado incorrectamente a agricultura. Esto resultó en una PA de 98.44%, reflejando una alta correspondencia entre los sitios clasificados como bosque en el mapa y los datos de campo. La UA fue del 82.89%, lo que señala una confusión considerable con la clase tular, indicando error de comisión.

En la categoría de agricultura, se verificaron un total de 77 sitios clasificados como tal en el mapa. De estos, 69 fueron correctamente asignados, mientras que 4 se asignaron mal en bosque, 1 en zona urbana, y 3 como pastizal. Esto resultó en una PA de 89.61%, lo que refleja una alta precisión en la clasificación para el productor. La UA fue del 87.34%, mostrando cierta confusión hacía clases que se clasificaron en esta categoría y pertenecían a otra, particularmente pastizal y zona urbana.

En la clase de zona urbana, se evaluaron 46 sitios en total. De estos, 41 fueron correctamente clasificados como zona urbana, mientras que 1 sitio fue mal clasificado como bosque y 4 como agricultura. Esto dio lugar a una PA de 89.13% y una UA de 97.62%, lo que indica una muy alta precisión tanto para el productor como para el usuario, encontrando confusión con la clase agricultura.

De 13 sitios clasificados como cuerpos de agua, 12 se asignaron correctamente, mientras que 1 sitio fue erróneamente categorizado como tular. Esto dio lugar a una PA de 92.31%, lo que indica una alta concordancia entre los sitios clasificados como cuerpos de agua y los datos de verificación. La UA también fue del 92.31%, mostrando una precisión robusta, con una mínima confusión con tular.

Tabla 3.9. SVM en imagen Sentinel-2 de 2022

Sitios de verificación	Mapa						
	Bosque	Agricultura	Zona Urbana	Cuerpos de agua	Tular	Pastizal	Total
Bosque	63	4	1	1	7	0	76
Agricultura	1	69	4	0	1	4	79
Zona Urbana	0	1	41	0	0	0	42
Cuerpos de agua	0	0	0	12	1	0	13
Tular	0	0	0	0	9	1	10
Pastizal	0	3	0	0	0	42	45
Total	64	77	46	13	18	47	265
PA [%]	98.44	89.6104	89.1304	92.31	50	89.3617	
UA [%]	82.8947	87.3418	97.62	92.31	90	93.3333	
Precisión general							89.06
Coeficiente Kappa							0.8602

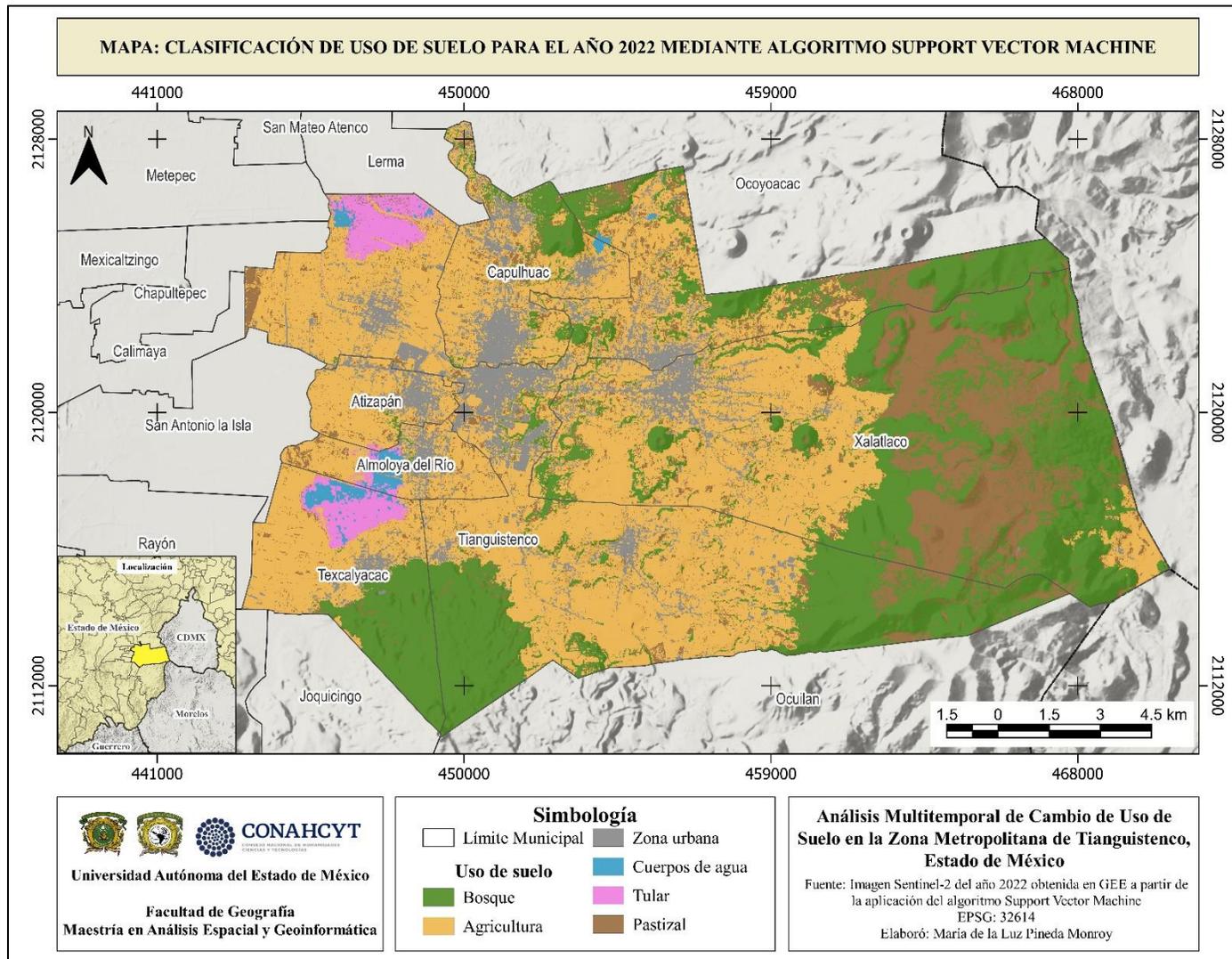
Fuente: Elaboración propia

Continuando con la descripción de la matriz de confusión de la Tabla 3.9, se verificaron un total de 18 sitios clasificados como tular en el mapa. De estos, solo 9 fueron correctos, mientras que 7 sitios fueron clasificadas como bosque, 1 como agricultura y 1 sitio más fue clasificado como pastizal. Esto dio lugar a una PA de 50%, lo que indica una precisión moderada en la clasificación de tular. Sin embargo, la UA fue del 90%, lo que significa que la mayoría de los sitios clasificados como tular eran correctos, pese a encontrar confusión con pastizal.

En la clase de pastizal, se evaluaron un total de 47 sitios. Entre ellos, 42 fueron correctamente clasificados como pastizal, mientras que 4 sitios fueron mal clasificados en agricultura y 1 como tular. Esto resultó en una PA de 89.36%, lo que indica una alta precisión del productor. La UA fue del 93.33%, lo que refleja que algunos sitios se asignaron a una clase que realmente no correspondía.

La precisión general fue del 89.06%, lo que indica un buen rendimiento del modelo Support Vector Machine en esta imagen Sentinel-2 de 2022. El coeficiente Kappa fue de 0.8602, lo que refleja una concordancia considerable entre la clasificación del mapa y los datos de verificación de campo. Las clases con mejor rendimiento fueron zona urbana y pastizal, mientras que tular presentó mayor confusión, especialmente con la clase de pastizal. La Figura 3.12 presenta la clasificación de la imagen Sentinel-2 de 2022 utilizando el algoritmo Support Vector Machine.

Figura 3.12. Clasificación de uso de suelo para el año 2022 mediante SVM



Fuente: Elaboración propia a partir de imagen Sentinel-2 del año 2022 y algoritmo SVM

Conforme a los resultados de la clasificación de imágenes satelitales Landsat 8 y Sentinel-2 realizadas por medio de GEE, y la posterior evaluación de la confiabilidad temática en QGIS 3.22.10, se concluye que el algoritmo SVM aplicado a las imágenes Sentinel-2 de 2016 y 2022 mostró el mejor desempeño. La Tabla 3.10 presenta los datos obtenidos en cada una de las clasificaciones.

Tabla 3.10. Precisión general y Coeficiente Kappa en Landsat 8 y Sentinel-2

LANDSAT 8	Precisión General	Coeficiente Kappa
Random Forest 2016	77.51	0.7
Support Vector Machine 2016	84.88	0.8
Random Forest 2022	73.94	0.65
Support Vector Machine 2022	78.54	0.72
SENTINEL-2	Precisión General	Coeficiente Kappa
Random Forest 2016	82.23	0.76
Support Vector Machine 2016	87.69	0.84
Random Forest 2022	85.66	0.81
Support Vector Machine 2022	85.93	0.82

Fuente: Elaboración propia

De esta manera, se cumplió con el primer objetivo específico, que consistió en evaluar la eficiencia de los algoritmos Random Forest y Support Vector Machine a través del uso de herramientas geotecnológicas para la adquisición de las capas de uso de suelo.

3.2 Cambio de uso de suelo en la Zona Metropolitana de Tianguistenco

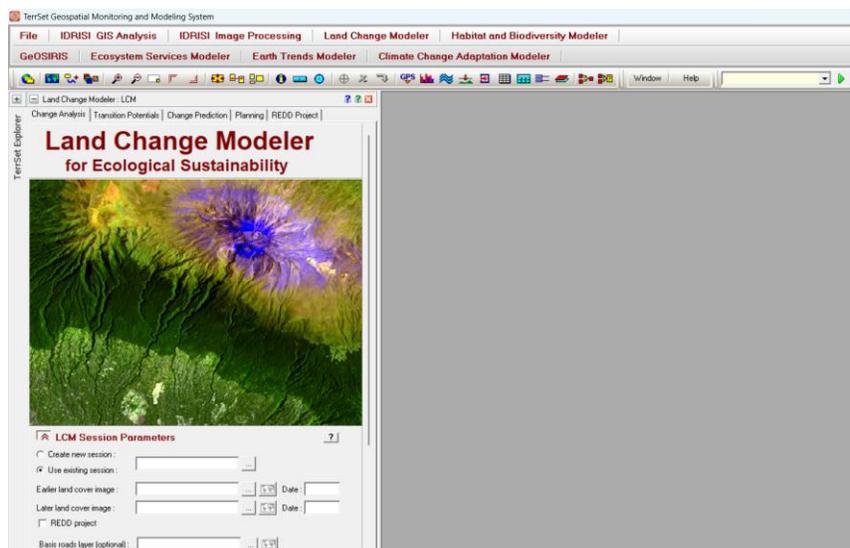
La Fase 3 se centró en determinar el cambio de uso de suelo entre 2016 y 2022, a partir de la clasificación de imágenes de satélite (Fase 1) y la evaluación de la confiabilidad temática (Fase 2). Tal como se indicó anteriormente, los resultados indicaron que el algoritmo SVM empleado a las imágenes de Sentinel-2 mostró un

mejor desempeño. En consecuencia, a partir de estas dos clasificaciones, se identificó el cambio de uso de suelo en la ZMT.

Es importante resaltar que en la Fase 3 del proyecto se empleó el software Terrset 2020 Geospatial Monitoring and Modeling Software, un programa integrado y diseñado para el monitoreo y modelado de sistemas terrestres orientado al desarrollo sostenible. Terrset 2020 se destaca por su amplia gama de herramientas geospaciales, ideales para abordar múltiples áreas del análisis ambiental y territorial.

Entre sus herramientas más relevantes se encuentra Land Change Modeler (LCM), un poderoso instrumento destinado a la planificación territorial y la toma de decisiones. LCM facilita el análisis y la predicción de cambios en el uso de suelo, ofreciendo una plataforma que permite modelar transiciones, generar tendencias y proyecciones a futuro. Terrset se ha consolidado como una herramienta fundamental en la planificación y conservación de los recursos naturales (Figura 3.13).

Figura 3.13. Interfaz de Terrset y la herramienta Land Change Modeler



Fuente: TerrSet, 2020

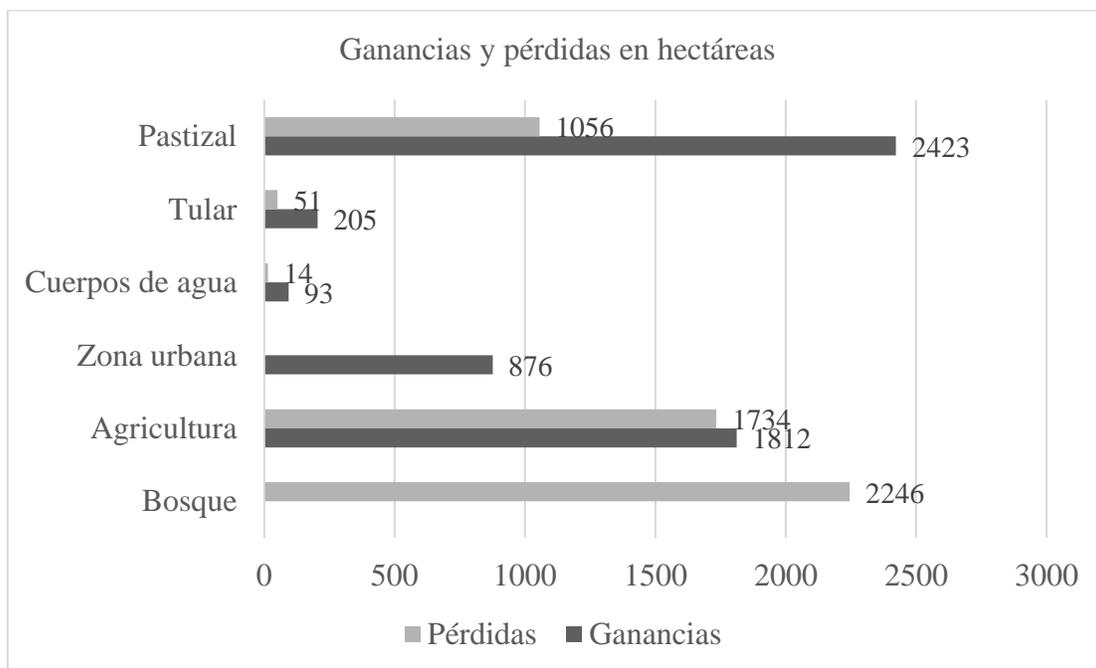
El cambio de uso de suelo hace referencia a la modificación que el ser humano realiza en el uso o manejo de la tierra, lo que genera alteraciones en su cobertura terrestre. Este proceso implica la evolución o transformación de la vegetación

original hacia otro tipo de uso, o bien la degradación de su calidad debido a procesos físicos, químicos o biológicos. Las consecuencias de estos cambios pueden afectar negativamente la capacidad del suelo para proporcionar servicios ecológicos y productivos, contribuyendo a problemas como la degradación del suelo y la pérdida de biodiversidad.

3.2.1 Ganancias y pérdidas

A partir de la información obtenida de la aplicación del algoritmo SVM en imágenes de satélite (mapas de uso de suelo del año 2016 y 2022) y el software Terrset 2020, se identificó el cambio de uso de suelo en la ZMT. A continuación, se describen los resultados sobre las ganancias y pérdidas de cada clase de uso de suelo (Gráfica 3.1).

Gráfica 3.1. Ganancias y pérdidas en hectáreas durante el periodo 2016 a 2022



Fuente: Elaboración propia

El bosque experimentó la mayor pérdida de cobertura con 2,246 hectáreas. Es importante señalar que la ZMT ha sido fuertemente impactada por la tala ilegal. Aunque los municipios han trabajado en la recuperación de áreas forestales mediante jornadas de reforestación, estos esfuerzos han resultado insuficientes

ante la presión de la tala, los incendios forestales y la proliferación de plagas que afectan actualmente la región.

La agricultura, por su parte, sufrió una pérdida de 1,734 ha, compensada por una ganancia de 1,812 ha. Es importante resaltar que la ZMT cuenta con una extensa área dedicada a actividades agrícolas, siendo el cultivo de temporal el predominante. Esto indica la relevancia de la agricultura en la región, pese a los cambios y transformaciones en el uso de suelo experimentados entre los años 2016 y 2022.

La zona urbana experimentó una notable expansión, registrando una ganancia de 876 ha. Cabe destacar que, a partir de 2016, los municipios de Atizapán, Almoloya del Río, Tianguistenco, San Mateo Texcalyacac, Capulhuac y Xalatlaco se integraron a la más reciente zona metropolitana creada en el Estado de México. Este hecho ha impulsado el crecimiento urbano en la región, reflejado claramente en el aumento del área urbana observado durante el periodo, como se muestran en la Gráfica 3.1.

En cuanto a la cobertura de cuerpos de agua, entre 2016 y 2022 se registró una pérdida de 14 ha, a su vez, se observó una ganancia de 93 ha. Esta variación puede deberse a diversos factores, como el incremento en las precipitaciones durante 2022 en comparación con 2016, o a la presencia de tular en áreas previamente identificadas por el algoritmo como cuerpos de agua o agricultura. Estas coberturas fueron altamente variables entre periodos, lo que explica los cambios observados en la extensión de los cuerpos de agua.

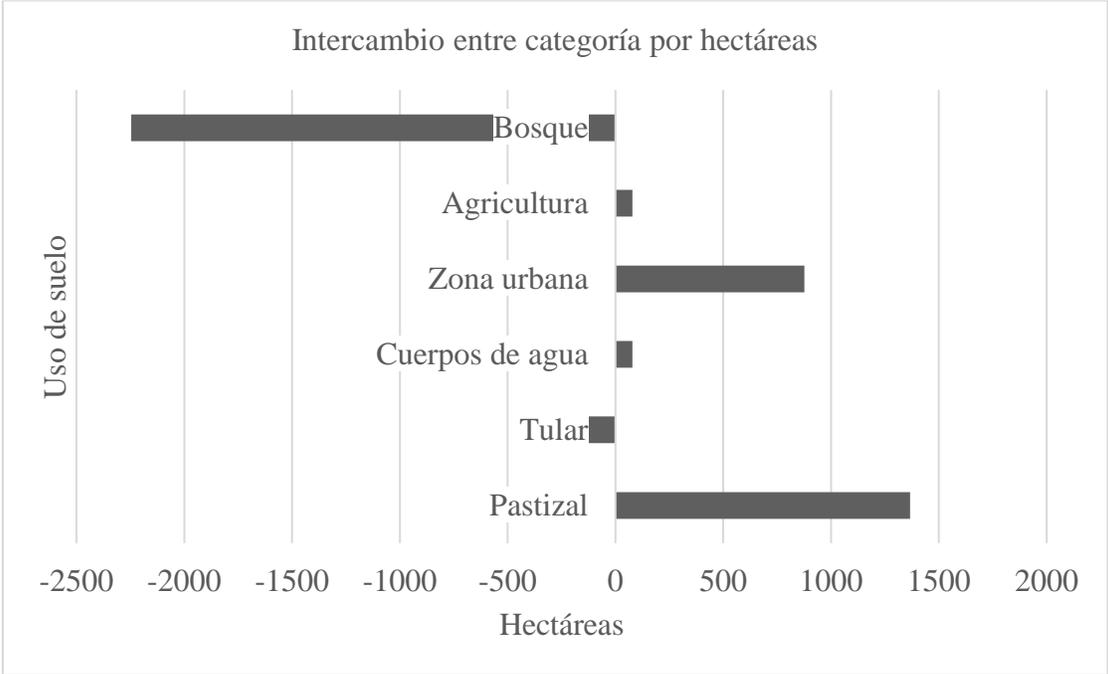
El tular registró una pérdida significativa de 205 ha, aunque se evidenció una ligera recuperación con una ganancia de 51 ha. Los principales factores de esta disminución son la expansión y la actividad agrícola, que presentan variaciones entre periodos que previamente se encontraban cubiertos por tular. Además, las fluctuaciones en la extensión de los cuerpos de agua, directamente relacionadas con las variaciones en el nivel hídrico, influyeron significativamente en esta cobertura.

Por último, el pastizal resultó una de las coberturas que presentó importantes pérdidas, registrando una disminución de 1,056 ha. Sin embargo, también fue una de las clases que más se expandió, con un incremento notable de 2,423 ha. Esta dinámica está motivada en gran medida por la tala ilegal de los bosques, que ha provocado la conversión de áreas forestales a pastizal en la ZMT.

3.2.2 Intercambio entre categorías

De acuerdo con la Gráfica 3.2 relacionada con el intercambio entre categorías de uso de suelo en la ZMT para los años 2016 a 2022, se detalla lo siguiente:

Gráfica 3.2. Intercambio entre categoría por hectáreas durante el periodo 2016 a 2022



Fuente: Elaboración propia

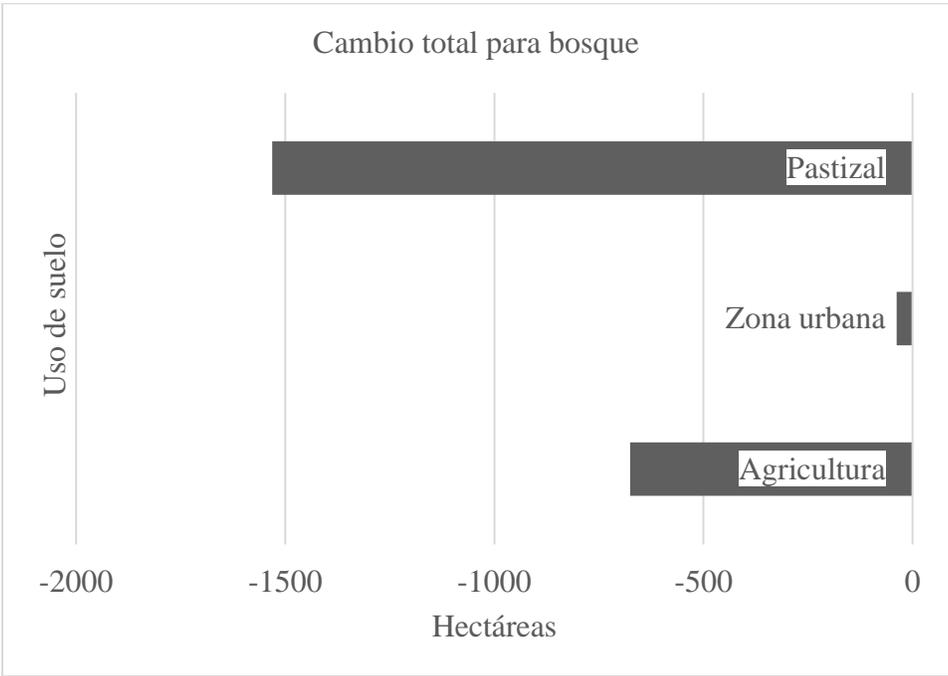
Como se ha mencionado anteriormente, el bosque sufrió la mayor pérdida de cobertura con una disminución significativa de 2,246 ha. Por su parte, la agricultura ha mantenido una estabilidad entre ambos periodos con un incremento de 78 ha. Lo anterior se debe al intercambio con otras coberturas, como pastizales y tulares. Como se citó, la agricultura forma parte de las actividades primarias más importantes de la región.

Con base en la Gráfica 3.2, la zona urbana presentó un aumento de 876 hectáreas. La urbanización sigue siendo una causa importante del cambio de uso de suelo en la ZMT. Los cuerpos de agua registraron un aumento de 79 ha, lo que sugiere una relativa estabilidad a pesar de las presiones ambientales. Sin embargo, el tular experimentó una disminución de 154 hectáreas.

Finalmente, el pastizal fue la categoría con mayor expansión, con un crecimiento de 1,367 ha. Este aumento está directamente relacionado con la pérdida de áreas forestales y, en menor medida, la cobertura agrícola y tular, lo que refleja un cambio significativo hacia actividades ganaderas y el uso extensivo del suelo.

3.2.3 Cambio total por categoría

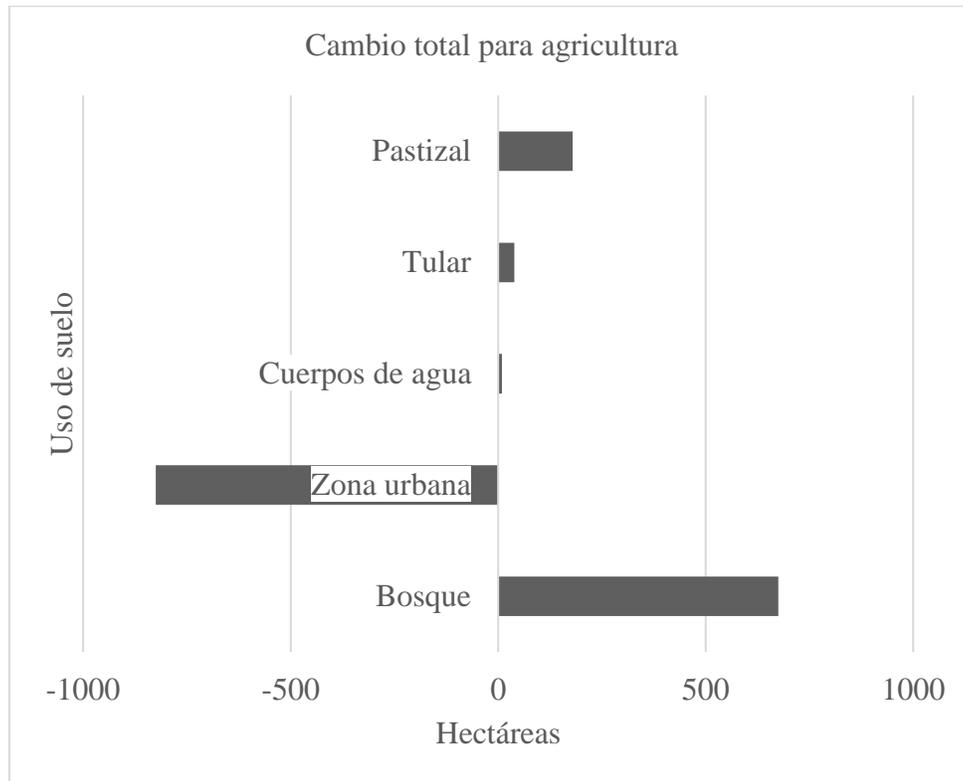
Gráfica 3.3. Cambio total para bosque



Fuente: Elaboración propia

El cambio total en la categoría de bosque muestra que entre 2016 y 2022 se perdieron 1,531 hectáreas convertidas en pastizal (Anexo 10) y 675 ha que se transformaron en tierras agrícolas. Además, 38 ha de bosque fueron reemplazadas por áreas urbanas, lo que refleja una contribución menor de la expansión urbana a la pérdida forestal (Gráfica 3.3).

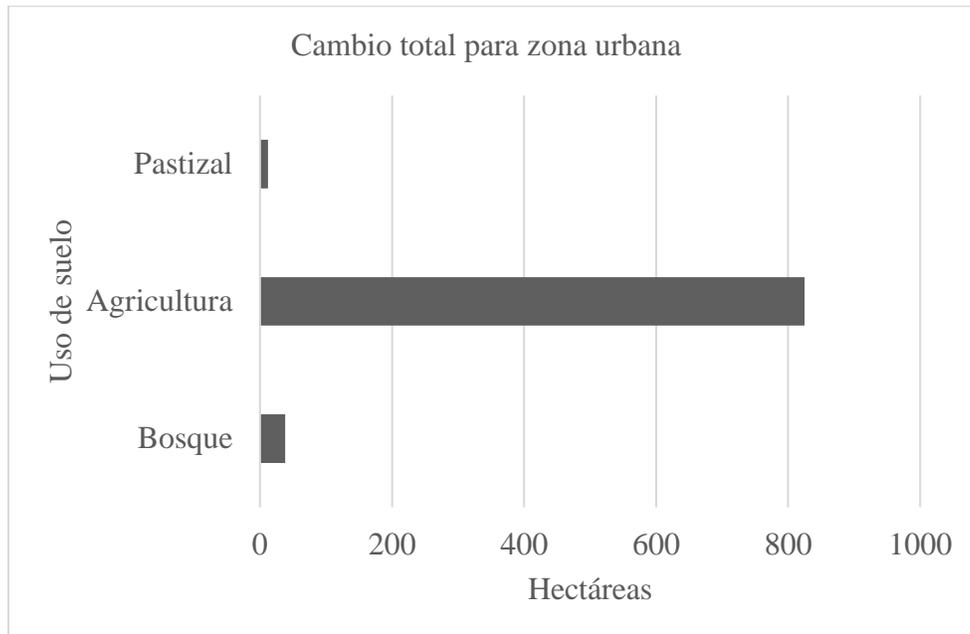
Gráfica 3.4. Cambio total para agricultura



Fuente: Elaboración propia

Como se observa en la Gráfica 3.4, el cambio total para la categoría de agricultura entre 2016 y 2022 muestra un aumento de 675 hectáreas provenientes de áreas forestales, mientras que 825 ha de tierras agrícolas se convirtieron en zonas urbanas, lo que indica un significativo crecimiento urbano. Además, 179 ha se convirtieron a pastizal y 39 ha a tular y 9 ha en cuerpos de agua.

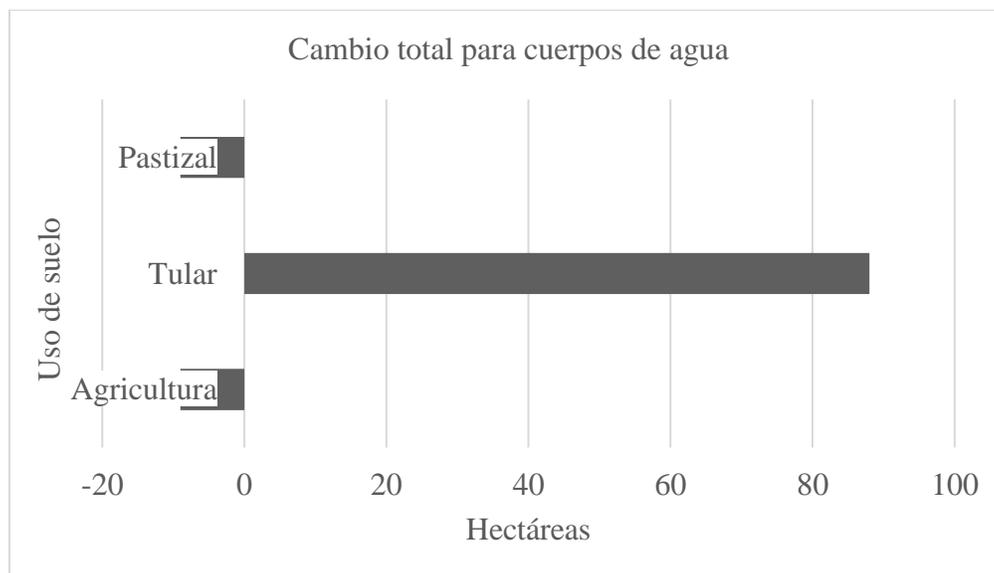
Gráfica 3.5. Cambio total para zona urbana



Fuente: Elaboración propia

El cambio total para la zona urbana entre 2016 y 2022 muestra una expansión significativa, con 825 hectáreas provenientes de áreas agrícolas, 38 ha de bosque y 12 ha de pastizales convertidas en suelo urbano. Esto refleja el gran crecimiento urbano, principalmente debido a la pérdida de la cubierta agrícola y, en menor medida, del bosque y pastizal (Gráfica 3.5 y Anexo 9 y 10).

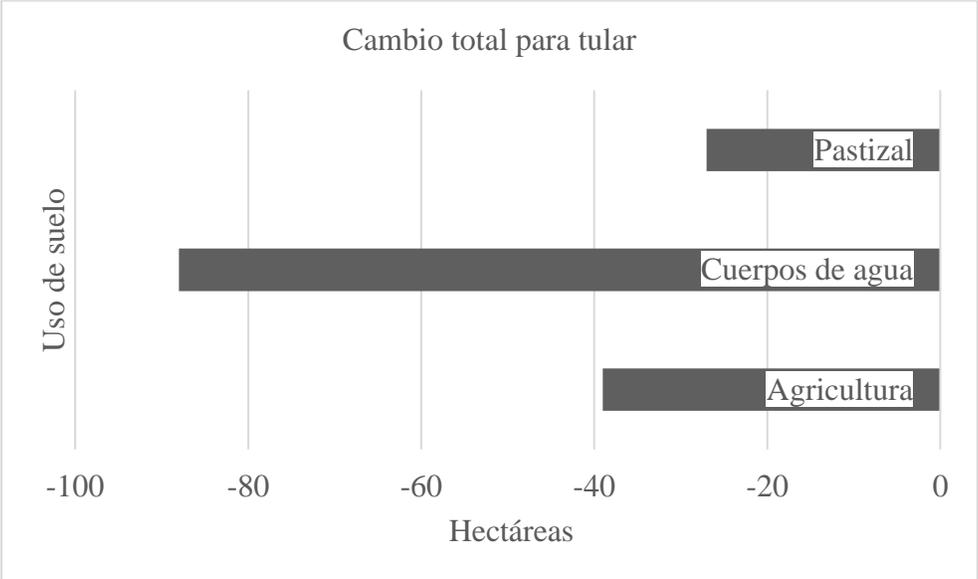
Gráfica 3.6. Cambio total para cuerpos de agua



Fuente: Elaboración propia

El cambio total para cuerpos de agua entre los años 2016 y 2022 muestra un notable aumento de 88 hectáreas provenientes de tular, lo que sugiere una expansión en áreas previamente ocupadas por dicha cobertura. Sin embargo, los cuerpos de agua ocuparon 9 ha de uso agrícola y de pastizales, respectivamente (Gráfica 3.6).

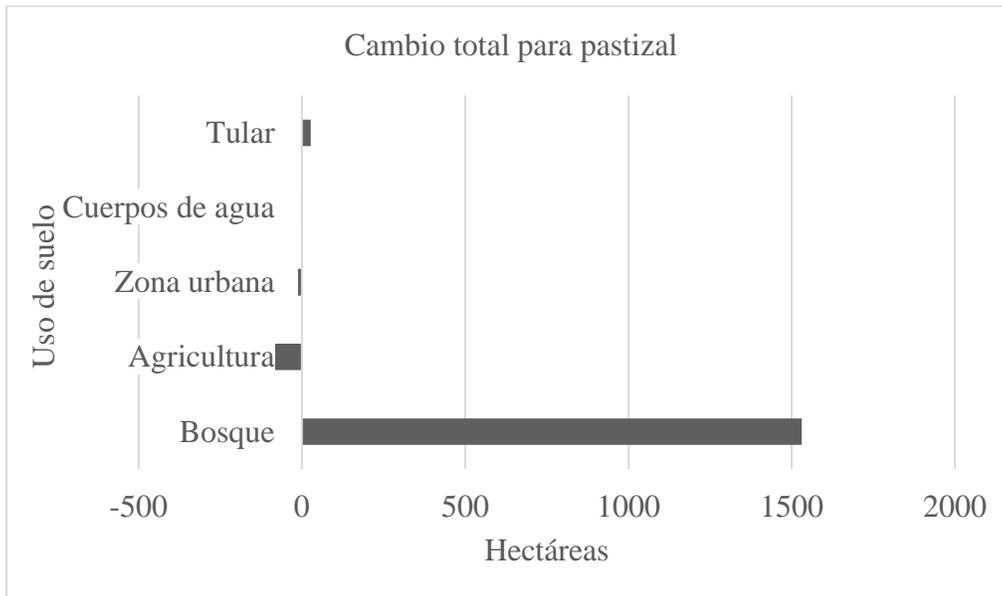
Gráfica 3.7. Cambio total para tular



Fuente: Elaboración propia

El cambio total para el tular entre 2016 y 2022 muestra una disminución de 39 hectáreas convertidas en tierras agrícolas, 88 ha que pasaron a cuerpos de agua y 27 ha que fueron transformadas en pastizales. Esto refleja una reducción significativa en la cobertura de tular, con pérdidas importantes hacia otros usos de suelo, principalmente cuerpos de agua y agricultura (Gráfica 3.7).

Gráfica 3.8. Cambio total para pastizal



Fuente: Elaboración propia

El cambio total para el pastizal entre 2016 y 2022 muestra una conversión de 1,531 hectáreas de bosque a pastizales, lo que refleja una pérdida significativa de cobertura forestal. Además, se incorporaron 27 ha provenientes del tular, mientras que se perdieron 179 ha que fueron convertidas en agricultura y 12 ha que pasaron a ser zona urbana. También hubo una ligera ganancia de 1 ha proveniente de cuerpos de agua. Estos cambios resaltan la expansión de los pastizales, mayormente a expensas del bosque, lo que indica una transformación significativa en el uso del suelo (Gráfica 3.8).

3.2.4 Cambio de uso de suelo durante el periodo 2016 a 2022

A continuación se describen los cambios de uso de suelo durante el periodo 2016 a 2022, para las coberturas más importantes para el estudio, lo siguiente se encuentra en la Tabla 3.11 y Figura 3.14.

Tabla 3.11. Cambios de uso de suelo representativos en la ZMT

Cambio	Hectáreas
Bosque a Agricultura	675.22
Tular a agricultura	86.017

Cambio	Hectáreas
Pastizal a Agricultura	1,041.39
Bosque a Zona Urbana	38.403
Agricultura a Zona urbana	825.182
Pastizal a Zona Urbana	12.2
Bosque a Pastizal	1,531.45
Agricultura a Pastizal	861.904
Tular a Pastizal	28.132

Fuente: Elaboración propia

El cambio más significativo es de bosque a pastizal con 1,531.45 hectáreas. Esta transición se puede observar claramente en la porción este de la zona de estudio en color café. Las áreas afectadas se encuentran en los municipios de Xalatlaco y Tianguistenco, colindantes con el municipio de Ocuilan y la Ciudad de México. Este cambio se debe principalmente a la tala ilegal y la conversión del bosque para realizar actividades de pastoreo (Anexo 10).

El segundo cambio más significativo es de pastizal a agricultura, con 1,041.39 ha, representado en color naranja. Es importante destacar la presencia de este cambio en la porción sureste, donde una pequeña localidad ha experimentado un incremento en las actividades agrícolas. Sin embargo, esta transición no es propia de un solo sitio; se encuentra presente en gran parte del área de estudio.

En tercer lugar, se encuentra el cambio de agricultura a pastizal, con 861.904 ha, representado en color amarillo. Esta transformación se encuentra dispersa en el mapa, pero particularmente visible en la periferia central del área urbana, la agricultura es más predominante. Una posible causa de esta transición es la práctica de agricultura de temporal en la zona. Para la fecha de las imágenes satelitales, los cultivos ya habían sido cosechados, y la firma espectral muestra una similitud con la de los pastizales.

El cambio de bosque a agricultura abarca 675.22 hectáreas, representado en color naranja claro y se observa disperso en la zona de estudio. Esto responde a la

necesidad de aumentar la producción agrícola a partir de coberturas de importancia para la conservación de áreas forestales (Anexo 10).

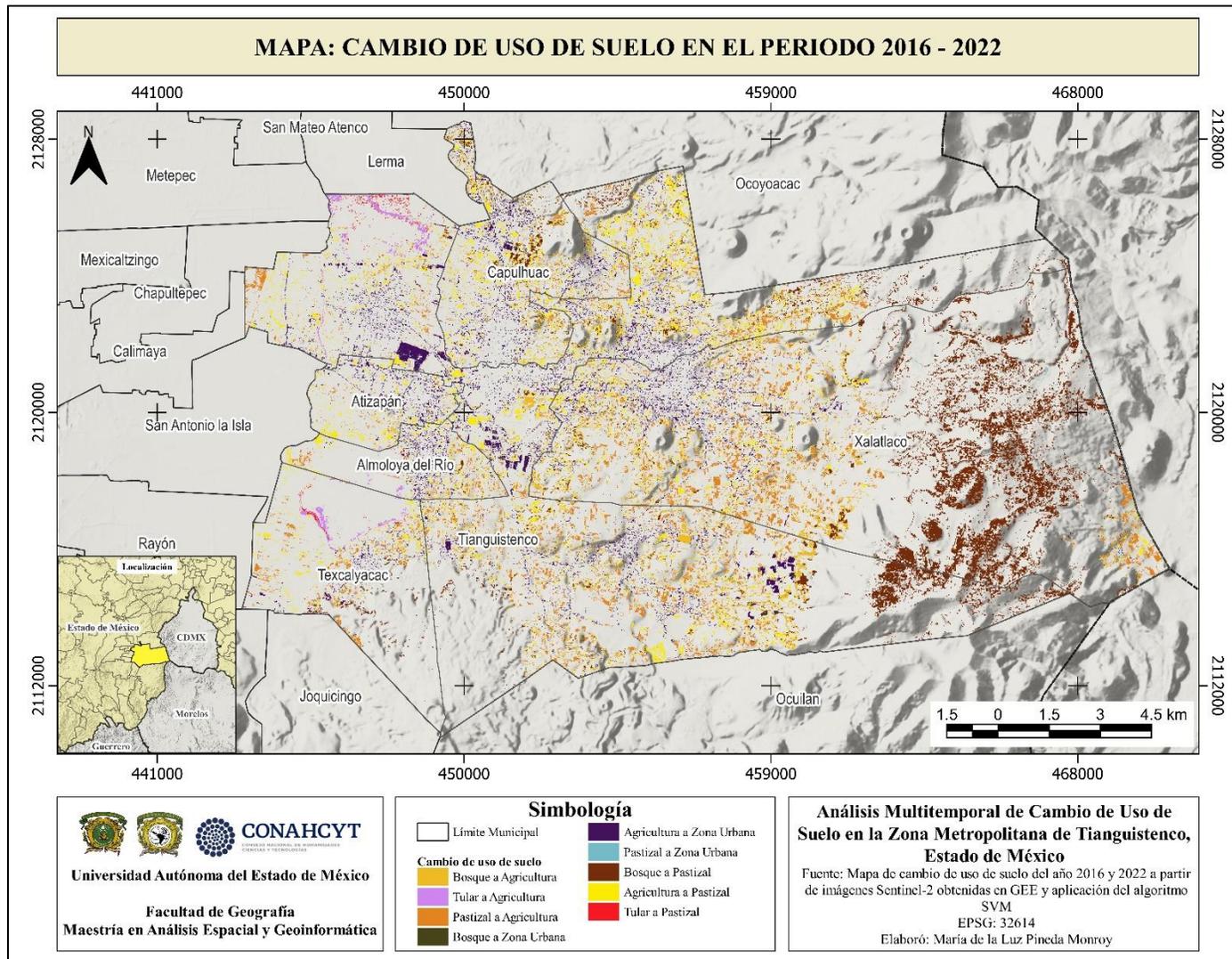
Con 825.182 hectáreas se presentó el cambio de agricultura a zona urbana, representada en color morado. Este cambio se observó principalmente en las periferias áreas urbanizadas. Cabe destacar que, entre 2016 y 2022 se creó una importante zona habitacional denominada “El Buen Suceso”, ubicada en el municipio de Tianguistenco, colindante con los límites de Atizapán, así como el crecimiento de la zona industrial, la cual se evidencia por las formas muy marcadas de lo que son naves industriales en el centro-oeste de la ZMT (Anexo 9).

El cambio de bosque a zona urbana abarcó 38.403 hectáreas y se encuentra representado en color verde militar. Este cambio se observa en áreas dispersas en el mapa. La conversión de bosque a zona urbana es el resultado de la expansión urbana, que afecta áreas forestales. Otro fenómeno muy visible en zonas anteriormente ocupadas por bosque, particularmente en Xalatlaco, es la minería a cielo abierto, la cual se consideró como parte de las características de las zonas urbanas (Anexo 9).

Con 86.017 hectáreas se encuentra el tular que cambió a agricultura, representado en color fucsia, y el tular a pastizal con 28.132 ha en color rojo. Estos cambios se encuentran particularmente alrededor de las Ciénegas del Lerma. Como se mencionó anteriormente, estos cambios pueden estar relacionados con la modificación del sistema hidrológico.

El cambio de pastizal a zona urbana abarcó 12.2 hectáreas, representado en azul rey, y se observa en áreas urbanas dispersas de menor escala. Esta conversión es el resultado del crecimiento de infraestructura habitacional en la región (Anexo 9).

Figura 3.14. Mapa de cambio de uso de suelo en la ZMT



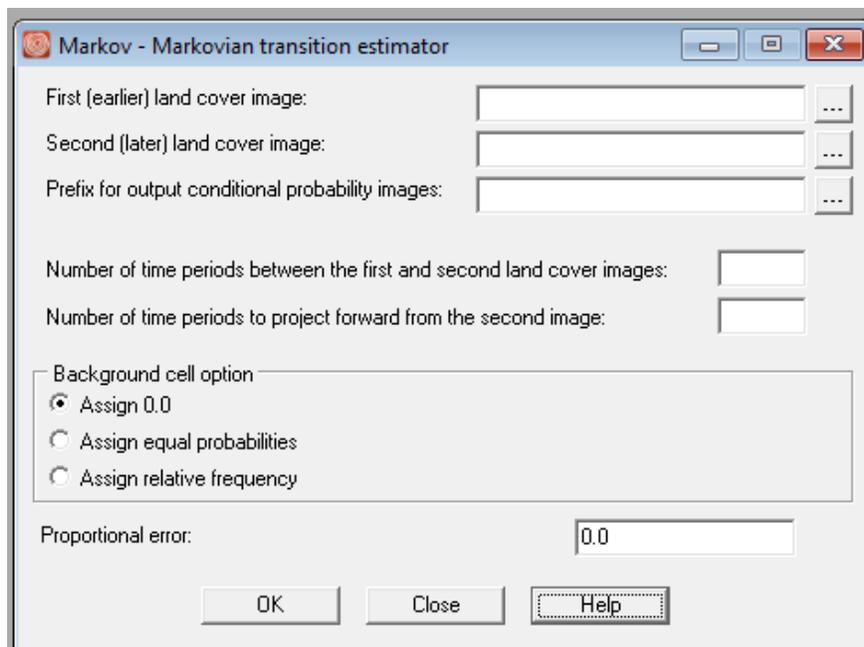
Fuente: Elaboración propia a partir de mapa de uso de suelo del año 2016 y 2022 a partir de imágenes Sentinel-2

3.3 Modelo predictivo de uso de suelo para 2028 mediante Cadenas de Markov

Para obtener el modelo predictivo de uso de suelo para el año 2028 en la Zona Metropolitana de Tianguistenco se utilizó el software Terrset 2020. La primera herramienta empleada fue el módulo Markov (Figura 3.15), cuya función es analizar el cambio en la cobertura de suelo entre dos imágenes y generar tres tipos de resultados: una matriz de probabilidad de transición, una matriz de áreas de transición y, por último, imágenes de probabilidad condicional (Takada *et al.*, 2010).

El procedimiento en el módulo Markov implica varios pasos, como la carga de las imágenes de entrada (las clasificaciones de 2016 y 2022), la definición de periodos entre dichas imágenes, y el número de periodos a proyectar (6 años para este análisis). Además, se configura el tratamiento de los píxeles de fondo para poder optimizar el modelo. De acuerdo con la bibliografía consultada el error proporcional es de 0.15.

Figura 3.15. Herramienta Markov en Terrset



Fuente: Terrset (2024)

El resultado del módulo Markov es una tabla que muestra la probabilidad de cambio para cada clase de uso de suelo. Por ejemplo, la probabilidad de que la clase 1

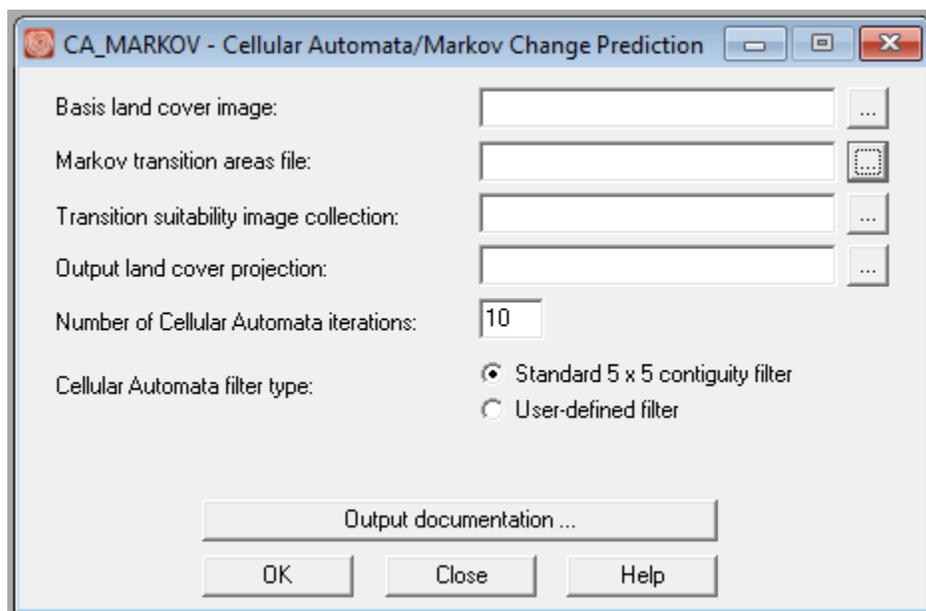
(bosque) permanezca sin cambios es de 0.6790, mientras que la probabilidad de que se transforme en clase 2 (agricultura) es de 0.0965 (Anexo 11). Adicionalmente esta información se representa de forma cartográfica en el Anexo 12, facilitando una mejor interpretación visual de las transiciones entre clases.

El segundo módulo que fue utilizado es CA_Markov, el cual es un procedimiento de predicción de la cobertura de uso de suelo que combina autómatas celulares, cadenas de Markov y asignación multicriterio.

CA_Markov involucra varios pasos, incluyendo la entrada de una imagen de cobertura de suelo base (mapa de uso de suelo 2022), un archivo de áreas de transición generado por el módulo Markov, y mapas de idoneidad para cada clase de suelo, que indican que tan apto es un píxel para cambiar a una clase específica. Los mapas de idoneidad se ajustan mediante un filtro de contigüidad que da preferencia a áreas cercanas a zonas ya existentes de una misma clase. El resultado final es una proyección de la cobertura de suelo basada en la dinámica de cambio y la idoneidad espacial (Kamusoko *et al.*, 2009; Memarian *et al.*, 2012) (Figura 3.16).

Es importante señalar que, en el apartado del número de iteraciones (periodos de tiempo), se hace referencia a la fecha de predicción futura especificada en el módulo anterior. El número de iteraciones divide el periodo total de intervalos de igual duración. Por ejemplo, las áreas se basarán en una predicción de 6 años y se asignaron 6 iteraciones, el modelo se estructura en incrementos anuales (Anexo 13).

Figura 3.16. Módulo CA_MARKOV para predicciones de cambio



Fuente: Terrset (2024)

Como resultado del proceso anterior, se obtuvo un mapa del modelo predictivo a partir de Cadenas de Markov, es importante aclarar que el producto de salida es una predicción, es decir, una proyección a futuro, una estimación de cómo se podrían transformar las diferentes clases de uso de suelo en función de patrones observados en el pasado. A continuación, se detalla el comportamiento para cada categoría de uso de suelo a partir de la Tabla 3.12 y Figura 3.17.

Tabla 3.12. Predicción de los usos de suelo al año 2028

Uso de suelo	2022	2028	Pérdida o ganancia
Bosque	8,924.76	7,161.43	-1,763.33
Agricultura	13,421.7	13,789.7	368
Zona urbana	2,588.65	3,432.79	844.14
Cuerpos de agua	234.052	288.125	54.07
Tular	662.516	540.188	-122.33
Pastizal	4616.23	5236.59	620.36

Fuente: Elaboración propia

El bosque proyecta una notable disminución en su cobertura forestal, de 8,924.76 a 7,161.43 hectáreas, lo que representa una pérdida de 1,763 ha. Esta reducción sugiere una presión considerable sobre las áreas de bosque, posiblemente ocasionada por la tala ilegal que transforma estas tierras en pastizales. La pérdida de bosque podría ser crítica, ya que esta cobertura es fundamental por los servicios ecosistémicos que ofrece, como la captación de agua, la biodiversidad, regulación del clima, entre otros.

Para la agricultura se espera un aumento de la superficie agrícola de 368 hectáreas, dado que, en el año 2022, se contaban con 13,421.7 ha. Esta expansión posiblemente sea impulsada por la demanda de producción alimentaria o el intercambio entre otras categorías. Sin embargo, esta expansión conlleva riesgos si no se acompañan de prácticas sostenibles, ya que podría contribuir a la degradación del suelo, la pérdida de biodiversidad y coberturas de importancia ambiental.

Se proyecta un incremento en la superficie agrícola de 368 hectáreas, pasando de 13,421.7 ha en 2022 a 13,789.7 ha en 2028. Esta expansión puede depender de la creciente demanda de productos agrícolas o a la conversión de otras categorías de uso de suelo. Sin embargo, este crecimiento implica riesgos si no se realiza mediante prácticas sostenibles, ya que podría conducir a la degradación del suelo, la pérdida de biodiversidad y reducción de coberturas de importancia ambiental.

La expansión urbana es una de las tendencias más marcadas, con un incremento de 2,588.65 hectáreas a 3,432.79 ha, lo que representa un aumento de 844.14 ha. Esto refleja la presión de la urbanización en la región, posiblemente impulsada por el crecimiento poblacional y el desarrollo de infraestructura. Cabe destacar que, al ser una zona metropolitana, la tendencia al crecimiento es evidente e inevitable. Sin embargo, la expansión de las áreas urbanas suele estar asociada con la pérdida de espacios naturales y agrícolas, así como con un aumento en la demanda de servicios y una mayor presión sobre el medioambiente.

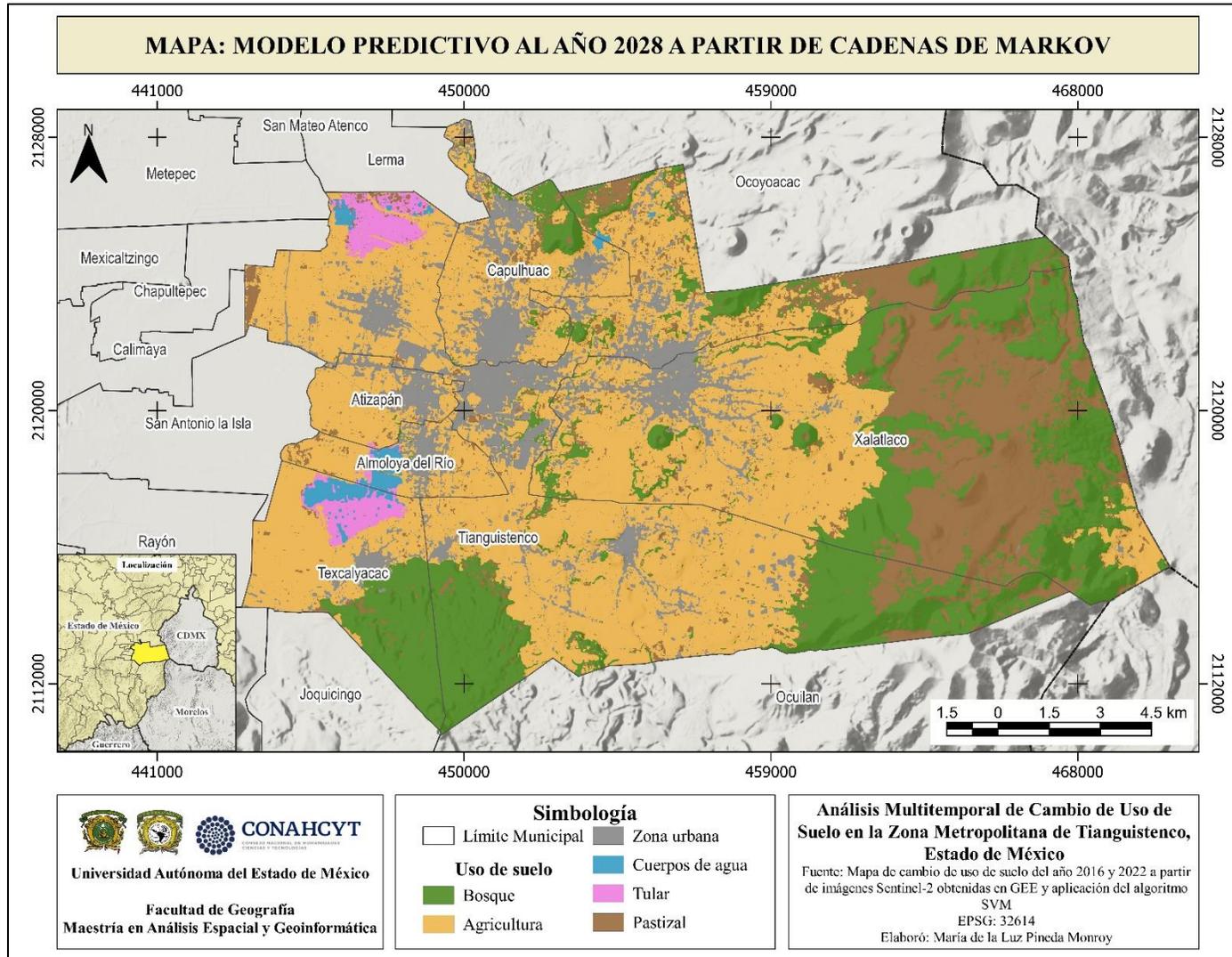
Aunque los cuerpos de agua aumentan ligeramente de 234.052 a 288.125 hectáreas, este crecimiento es relativamente menor en comparación con otras categorías. Sin embargo, esta estabilidad y leve incremento resultan positivos, ya

que los cuerpos de agua son esenciales para la regulación hídrica y el mantenimiento de la biodiversidad acuática. Esta cobertura es especialmente importante en la región, ya que incluye a las Ciénegas del Lerma, una relevante Área de Protección de Flora y Fauna. Es trascendental señalar que, al tratarse de una predicción, no se han considerado factores adicionales, como la posibilidad de sequías en los próximos años, lo cual podría modificar significativamente esta proyección.

Se proyecta una reducción en los terrenos de tular, de 662.516 a 540.188 hectáreas, lo que implica una pérdida de alrededor de 122 hectáreas. Este ecosistema de humedales juega un papel importante en la filtración de agua, protección contra inundaciones y como hábitat para diversas especies del humedal. La disminución de la cobertura podría indicar una presión de cambio hacia otros usos, como la agricultura, los pastizales e incluso la zona urbana.

Por último, la superficie ocupada por pastizal incrementó de 4,616.23 a 5,236.59 hectáreas, es decir, un aumento de 620.36 ha. Lo anterior, podría responder principalmente al pronóstico de pérdida de cobertura forestal, como se puede apreciar en la Figura 3.17. Sin embargo, existen otras posibles causales, como la necesidad de espacios para ganadería o agricultura.

Figura 3.17. Mapa de modelo predictivo al año 2028



Fuente: Elaboración propia basada en mapa de cambio de uso de suelo 2016 y 2022 obtenidas mediante imágenes Sentinel-2

Conclusiones

El presente estudio ha demostrado que el análisis del cambio de uso de suelo constituye una herramienta fundamental para comprender las dinámicas territoriales y sus implicaciones socioambientales. Los resultados obtenidos contribuyen a ampliar el conocimiento sobre los procesos de conversión del uso de la tierra en la región y ofrecen una base sólida para la toma de decisiones informadas en materia de planificación y gestión ambiental.

La aplicación de geotecnologías, especialmente el uso de imágenes Sentinel-2 en GEE, ha sido fundamental para el desarrollo de este estudio, dado que han contribuido a cumplir con los objetivos de la investigación. Al comparar algoritmos de clasificación, se observó que Support Vector Machine superó a Random Forest en precisión con valores de coeficiente Kappa de 0.84 y 0.82 para SVM y 0.76 y 0.81 para RF, en los años 2016 y 2022, respectivamente. Los valores de precisión y coeficiente Kappa obtenidos con SVM respaldan su elección como el modelo más adecuado para cartografiar los cambios en la cobertura terrestre de la Zona Metropolitana de Tianguistenco. Estos resultados, junto con la disponibilidad gratuita y la frecuencia temporal de las imágenes Sentinel-2, demuestran el potencial de GEE para monitorear de manera eficiente la dinámica del paisaje.

El análisis de cambio de uso de suelo en la ZMT entre 2016 y 2022, combinado con la proyección del modelo predictivo para 2028, revela una dinámica territorial caracterizada por la pérdida de la cobertura forestal y la expansión de la frontera agrícola y urbana. La conversión de 1,531 ha de bosque a pastizal entre 2016 y 2022, impulsada principalmente por actividades agrícolas y tala ilegal, evidencia una presión creciente sobre los ecosistemas naturales. A su vez, el crecimiento urbano, con un aumento de 876 ha en el mismo periodo, refleja los procesos de urbanización y desarrollo asociados a la designación de metrópoli.

El modelo predictivo proyecta una intensificación de estas tendencias para 2028, con una pérdida adicional de 1,763 ha de bosque y un crecimiento urbano de 844 ha. La expansión de la agricultura, aunque presenta una dinámica más compleja, también contribuirá a la transformación del paisaje. La disminución proyectada de

tular y el aumento de los cuerpos de agua, si bien puede parecer beneficioso a primera vista, puede representar una amenaza para el humedal, ecosistema de gran importancia ecológica de la región. La pérdida de la vegetación acuática puede afectar la calidad del agua, la biodiversidad y los servicios ecosistémicos que el humedal proporciona.

Como último punto, los resultados obtenidos acentúan la necesidad urgente de implementar políticas públicas y estrategias de gestión territorial que promuevan la conservación de los ecosistemas naturales, la restauración de áreas degradadas y un uso más sostenible del suelo. Es fundamental fortalecer los instrumentos de planificación y regulación de uso del suelo, promover prácticas agrícolas sostenibles y fomentar la participación de la sociedad en la toma de decisiones.

Recomendaciones

A partir de los resultados obtenidos en este estudio, se proponen las siguientes recomendaciones.

La selección del algoritmo de clasificación adecuado es crucial para obtener mapas de uso de suelo precisos y confiables. Si bien los algoritmos Random Forest y Support Vector Machine son ampliamente utilizados, la bibliografía consultada sugiere que otros métodos, como Maximum Likelihood Classification (MLC), K-Means, Redes Neuronales Convolucionales (CNN) y K-Nearest Neighbors (K-NN), podrían ofrecer ventajas específicas dependiendo de las características de los datos y los objetivos del estudio.

El modelo predictivo Cadenas de Markov proyecta una pérdida significativa de cobertura forestal y un aumento en la expansión urbana para 2028. Integrar estos modelos en el proceso de planeación territorial ayudaría a anticipar y mitigar impactos ambientales, permitiendo ajustar políticas y estrategias de uso de suelo en función de escenarios futuros. A su vez, la incorporación de insumos en los modelos predictivos permite obtener resultados más robustos y confiables.

Se sugiere explorar y comparar diferentes modelos predictivos, así como utilizar otras herramientas geoinformáticas para generar escenarios futuros en la Zona Metropolitana de Tianguistenco. Un ejemplo de ello es la aplicación de regresión logística geográficamente ponderada utilizando R.

A partir de los resultados obtenidos en este trabajo y posterior a demostrar que la cobertura forestal y el humedal se han visto afectados por las dinámicas de cambio del suelo es de suma importancia fortalecer la gobernanza ambiental en la Zona Metropolitana de Tianguistenco a partir de la colaboración con actores políticos, autoridades locales, organizaciones no gubernamentales y la población en general para promover políticas públicas orientadas a la conservación y restauración de las áreas forestales y las Ciénegas del Lerma.

Referencias

- Adame Martínez, S., Sánchez Nájera, R. M. y Hoyos Castillo, G. d. (2020). Factores socioterritoriales de cambio de uso de suelo en el centro de México. Caso oriente de la Zona Metropolitana de Toluca, México. *Revista Universitaria de Geografía*, 29(1), 153-183. <http://www.scielo.org.ar/pdf/reuge/v29n1/v29n1a07.pdf>
- Aguilera Benavente, F. (2006). Predicción del crecimiento urbano mediante sistemas de información geográfica y modelos basados en autómatas celulares. *Revista Internacional de Ciencia y Tecnología de la Información Geográfica*(6), 81-112. <https://www.geofocus.org/index.php/geofocus/article/view/90>
- Alonso, C., Moreno, V., y Rodríguez, E. (1999). Determinación experimental de la firma espectral de la vegetación. Una sencilla práctica de introducción a la teledetección. *TELEDETECCIÓN. Avances y Aplicaciones. VIII Congreso Nacional de Teledetección.*, 429-432. https://www.researchgate.net/profile/Carmelo-Alonso-Jimenez/publication/237515103_DETERMINACION_EXPERIMENTAL_DE_LA_FIRMA_ESPECTRAL_DE_LA_VEGETACION_UNA_SENCILLA_PRACTICA_DE_INTRODUCCION_A_LA_TELEDETECCION/links/5a5b3557a6fdcc3bfb606a24/DETERMINACION-EXPER
- Araque Torres, L. y Jiménez, A. (2009). Caracterización de firma espectral a partir de sensores remotos para el manejo de sanidad vegetal en el cultivo de palma de aceite. *PALMAS*, 30(3), 63-79. Obtenido de <https://publicaciones.fedepalma.org/index.php/palmas/article/view/1455/1455>
- Basheer, S., Wang, X., Farooque, A. A., Nawas, R. A., Liu, K., Adekanmbi, T. y Liu, S. (2022). Comparison of Land Use Land Cover Classifiers Using Different Satellite Imagery and Machine Learning Techniques. *Remote Sensing*, 14(19), 1-18. doi:10.3390/rs14194978
- Buzai, G. D. (2015). Conceptos fundamentales del Análisis Espacial que sustentan la investigación científica basada en Geotecnologías. En M. Fuenzalida, G. D. Buzai, A. Moreno Jiménez, & A. García de León, *Geografía, Geotecnología y Análisis Espacial: tendencias, métodos y aplicaciones* (Primera ed., págs. 55-72). Santiago de Chile: Editorial Triángulo.
- Buzai, G. D. y Baxendale, C. A. (2010). Análisis Espacial con Sistemas de Información Geográfica. Aportes de la Geografía para la Elaboración del Diagnóstico en el Ordenamiento Territorial. En M. C. Pineda, & J. Bosque

Sendra , *Actas I Congreso Internacional en Ordenamiento y Tecnologías de la Información Geográfica* (págs. 319-341). Alcalá de Henares, España: Editorial Universidad de Alcalá.

Buzai, G. D. y Baxendale, C. A. (2013). Aportes del análisis geográfico con Sistemas de Información Geográfica como herramienta teórica, metodológica y tecnológica para la práctica del ordenamiento territorial. *Persona y Sociedad*, XXVII(2), 113-141.
https://ri.conicet.gov.ar/bitstream/handle/11336/3836/CONICET_Digital_Nro_5009_A.pdf?sequence=2

Caballero Martínez, Y. (2017). *Análisis del Intercambio Comercial Local desde la Perspectiva de la Economía Solidaria, 2016. "Caso de estudio: Santiago Tianguistenco [Tesis de Licenciatura, Universidad Autónoma del Estado de México]*. <http://hdl.handle.net/20.500.11799/68025>

Camacho Sanabria, J. M. (Enero de 2016). Cambios de Cobertura y Uso del Suelo en el Surponiente del Estado de México. Modelos Dinámicos y Prospectivos. Aportación Teórica-Metodológica en las Ciencias Ambientales [Tesis de Doctorado no publicada]. Universidad Autónoma del Estado de México.

Camacho-Sanabria, R., Camacho-Sanabria, J. M., Balderas-Plata, M. Á. y Sánchez-López, M. (2017). Cambio de cobertura y uso de suelo: estudio de caso en Progreso Hidalgo, Estado de México. *Madera y Bosques*, 23(3), 39-60. doi:10.21829/myb.2017.2331516

Chuvienco, E. (1995). *Fundamentos de Teledetección Espacial* (Segunda ed.). Madrid: Ediciones Rialp, S.A.

COESPO. (2017). *Zona Metropolitana de Tianguistenco*. <https://coespo.edomex.gob.mx/sites/coespo.edomex.gob.mx/files/files/coespopdfzms17.pdf>

Congedo, L. (2021). Semi-Automatic Classification Plugin: A Python tool for the download and processing of remote sensing images in QGIS. *Journal of Open Source Software*, 6(64), 1-6. doi:10.21105/joss.03172

El Naqa, I. y Murphy, M. J. (2015). What Is Machine Learning? En I. El Naqa, R. Li, & M. J. Murphy, *Machine Learning in Radiation Oncology. Theory and Applications* (págs. 3-13). Switzerland: Springer. doi:10.1007/978-3-319-18305-3_1

EOS Data Analytics. (2023). *Landsat 8*. Obtenido de EOS Data Analytics, Inc: <https://eos.com/es/find-satellite/landsat-8/>

- ESA. (2015). *Sentinel 2*. Obtenido de The European Space Agency: https://www.esa.int/Space_in_Member_States/Spain/SENTINEL_2
- ESRI. (2016). *ArcGIS for Desktop*. Obtenido de ¿Qué son los datos ráster?: <https://desktop.arcgis.com/es/arcmap/10.3/manage-data/raster-and-images/what-is-raster-data.htm>
- Estupiñan Ricardo, J., Leyva Vázquez, M. Y., Peñafiel Palacios, A. J. y El Assafari Ojeda, Y. (2021). Inteligencia artificial y propiedad intelectual. *Revista Universidad y Sociedad*, 13(S3), 362-368. <https://rus.ucf.edu.cu/index.php/rus/article/view/2490/2445>
- European Space Agency [ESA]. (2024). *Firmas espectrales*. https://www.esa.int/SPECIALS/Eduspace_ES/SEM6DYD3GXF_0.html
- FAO. (2023). *Portal de Suelos de la FAO*. <https://www.fao.org/soils-portal/soil-biodiversity/conservacion-del-suelo-y-agricultura/es/>
- Feizizadeh, B., Omarzadeh, D., Garajeh, M. K., Lakes, T. y Blaschke, T. (2021). Machine learning data-driven approaches for land use/cover mapping and trend analysis using Google Earth Engine. *Journal of Environmental Planning and Management*, 33. doi:10.1080/09640568.2021.2001317
- Galicia Sarmiento, L., García Romero, A., Gómez Mendoza, L. y Ramírez, M. I. (2007). Cambio de uso del suelo y degradación ambiental. *ciencia. Revista de la Academia Mexicana de Ciencias*, 58(4), 50-60. https://www.amc.edu.mx/revistaciencia/images/revista/58_4/PDF/09-470-50-59.pdf
- García Orozco, J. A., Cedillo Gutiérrez, J. G., Juan Perez, J. I. y Balderas Plata, M. (2012). Procesos de Cambio en el Uso del Suelo de una Microcuenca en el Altiplano Mexicano. El caso del Río San José en el Estado de México. *Papeles de Geografía*(55-56), 63-73. <https://www.redalyc.org/articulo.oa?id=40726731009>
- Gobierno de México. (30 de septiembre de 2024). *El gobierno de México ahora cuida 232 Áreas Naturales Protegidas federales (CONANP)*. <https://www.gob.mx/conanp/prensa/el-gobierno-de-mexico-ahora-cuida-232-areas-naturales-protegidas-federales>
- Gobierno del Estado de México. (25 de noviembre de 2016). Decreto número 159.- Por el que se aprueba la declaratoria de Zona Metropolitana de Santiago Tianguistenco integrada por los municipios de Almoloya del Río, Atizapán, Capulhuac, Texcalyacac, Tianguistenco y Xalatlaco. *Gaceta del Gobierno*.

<http://www.secretariadeasuntosparlamentarios.gob.mx/mainstream/Actividad/Decretos/LIX/DECRETO%20159.pdf>

Henríquez Ruiz, C. y Azócar García, G. (1 de agosto de 2007). Propuesta de modelos predictivos en la planificación territorial y evaluación de impacto ambiental. *Scripta Nova. Revista Electrónica de Geografía y Ciencias Sociales*, XI(245), 1-14. [https://www.researchgate.net/profile/Cristian-Henriquez-](https://www.researchgate.net/profile/Cristian-Henriquez-2/publication/28185031_Propuesta_de_modelos_predictivos_en_la_planificacion_territorial_y_evaluacion_de_impacto_ambiental/links/606f95bd92851c8a7bb2dd4f/Propuesta-de-modelos-predictivos-en-la-planifi)

[2/publication/28185031_Propuesta_de_modelos_predictivos_en_la_planificacion_territorial_y_evaluacion_de_impacto_ambiental/links/606f95bd92851c8a7bb2dd4f/Propuesta-de-modelos-predictivos-en-la-planifi](https://www.researchgate.net/profile/Cristian-Henriquez-2/publication/28185031_Propuesta_de_modelos_predictivos_en_la_planificacion_territorial_y_evaluacion_de_impacto_ambiental/links/606f95bd92851c8a7bb2dd4f/Propuesta-de-modelos-predictivos-en-la-planifi)

Higueras Arnal, A. M. (2003). *Teoría y Método de la Geografía: Introducción al análisis geográfico regional*. Zaragoza, España: Prensas Universitarias de Zaragoza.

Ibarra, L. F. (2009). Predicciones de Markov Aplicadas en el Programa de Ingeniería Industrial de la Universidad Nacional Experimental del Táchira (UNET). *Ingeniería Industrial. Actualidad y Nuevas Tendencias*, 1(2), 39-51. <https://www.redalyc.org/pdf/2150/215016874004.pdf>

IBM. (s.f). *What is the random forest?* Obtenido de IBM: <https://www.ibm.com/topics/random-forest>

INEGI. (2014). Sistema de Información Geográfica. <https://www.inegi.org.mx/inegi/spc/doc/internet/sistemainformaciongeografica.pdf>

IPCC. (2012). Glossary of Terms. En C. B. Field, V. Barros, T. F. Stocker, Q. Dahe, D. J. Dokken, K. L. Ebi, . . . P. M. Midgley, *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation* (págs. 555-564). Cambridge and New York, UK, USA: Cambridge University Press.

Jiménez López, E. (2019). Cadenas de Markov espaciales para simular el crecimiento del Área Metropolitana de Toluca, 2017-2031. *Economía, Sociedad y Territorio*, XIX(60), 109-140. doi:10.22136/est20191324

Jiménez-Moreno, M. J., González-Guillen, M. d., Escalona-Maurice, M., Valdez-Lazalde, J. R. y Aguirre-Salado, C. A. (2011). Comparación de métodos espaciales para detectar cambios en el uso del suelo urbano. *Revista Chapingo Serie Ciencias Forestales y del Ambiente*, 17(3), 389-406. <https://www.scielo.org.mx/pdf/rcscfa/v17n3/v17n3a9.pdf>

Kamusoko, C., Aniya, M., Adi, B. y Manjoro, M. (2009). Rural sustainability under threat in Zimbabwe - Simulation of future land use/cover changes in the

- Bindura district based on the Markov-cellular automata model. *Applied Geography*, 29, 435-447.
- Kumar, L. y Mutanga, O. (2019). Google Earth Engine Applications. *Remote Sensing*, 11(5). doi:<https://www.mdpi.com/2072-4292/11/5/591>
- Labrador García, M., Évora Brondo, J., Arbelo Pérez, M. (2012). *Satélites de Teledetección para la Gestión del Territorio*. Canarias : Litografía Romero.
- Lambin, E. F., Turner, B L., Geist, H. J., Agbola, S. B., Angelsen, A., Bruce, J. W., Coomes, O. T., Dirzo, R., Fisher, G., Folke, C., George, P. S., Homewood, K., Imbernon, J., Leemans, R., Li, X., Moran, E. F., Mortimore, M., Ramakrishnan, P. S., Richards, J. F.,... Xu, J. (2001). The causes of land-use and land-cover change: moving beyond the myths. *Global Environmental Change*, 11(4), 261-269. doi:10.1016/S0959-3780(01)00007-3
- Landis, J. R. y Koch, G. G. (1977). The Measurement of Agreement for Categorical Data. *Biometrics*, 33(1), 159-174. doi:10.2307/2529310
- Liaw, A. y Wiener, M. (2002). Classification and Regression by randomForest. *R News*, 2, 18-22.
- López Hung, E. y Joa Triay, L. G. (2017). Cadenas de markov aplicadas al análisis de la ejecución de proyectos de investigación/Markov chains applied to the analyzis of the progress of research projects. *Revista Cubana de Informática Médica*, 9(1), 44-51. <https://www.medigraphic.com/pdfs/revcubinmed/cim-2017/cim171e.pdf>
- López Portillo, V. (24 de Junio de 2019). *Reporta Global Forest Watch máximos en pérdida de cobertura forestal desde 2019*. <https://wrimexico.org/bloga/reporta-global-forest-watch-m%C3%A1ximos-en-p%C3%A9rdida-de-cobertura-forestal-desde-2001>
- López Trigal, L. (2015). *Diccionario de Geografía aplicada y profesional: Terminología de análisis, planificación y gestión del territorio*. León: Universidad de León.
- López Vazquez, V. H., Balderas Plata, M. Á., Chávez Mejía, M. C., Juan Pérez, J. I. y Gutiérrez Cedillo, J. G. (2015). Cambio de uso de suelo e implicaciones socioeconómicas en un área mazahua del altiplano mexicano. *CIENCIA ergo sum*(22), 136-144. Obtenido de <https://www.redalyc.org/articulo.oa?id=10439327004>
- Mas, J. F. y Couturier, S. (2011). Evaluación de Bases de Datos Cartográficas. En F. Bautista Zúñiga, J. L. Palacio Prieto, H. Delfín González, R. Paéz Bistrain,

- E. Carmona Jiménez, & M. D. Delgado Carranza, *Técnicas de muestreo para manejadores de recursos naturales* (Segunda ed., págs. 675-703).
- Mas, J. F., Díaz Gallegos, J. R. y Pérez Vega, A. (2003). Evaluación de la confiabilidad temática de mapas o de imágenes clasificadas: una revisión. *Investigaciones Geográficas, Boletín del Instituto de Geografía, UNAM*(51), 53-72.
- Memarian, H., Balasundram, S. K., Jamal, B. T. y Teh, C. B. (2012). Validation of CA-Markov for Simulation of Land Use and Cover Change in the Langat Basin, Malaysia. *Journal of Geographic Information System*, 4, 542-554.
- Montero Mora, A. y Viales Hurtado, R. (2015). La teoría del cambio en el paisaje a partir del cambio del uso de la tierra y la cobertura del suelo (enfoque lucc). Su utilidad para la historia ambiental. *Reflexiones*, 94(2), 25-33. <https://www.redalyc.org/journal/729/72946471002/html/#B1>
- Morales Sánchez, D. (2016). *Huella Ecológica del Intercambio de Leña en el Mercado de Trueque de Tianguistenco, Estado de México [Tesis de Licenciatura, Universidad Autónoma del Estado de México]*. <http://hdl.handle.net/20.500.11799/49201>
- Muhammad, R., Zhang, W., Abbas, Z., Guo, F. y Gwiazdzinski, L. (2022). Spatiotemporal Change Analysis and Prediction of Future Land Use and Land Cover Changes Using QGIS MOLUSCE Plugin and Remote Sensing Big Data: A Case Study of Linyi, China. *Land*, 11(3). doi:10.3390/land11030419
- My GIS Notebook. (3 de marzo de 2019). *Ráster vs vectorial*. <https://mygisnotebook.blog/2019/03/03/raster-vs-vectorial/>
- NASA Earth Observatory. (s.f). *Measuring Vegetation (NDVI & EVI)*. https://earthobservatory.nasa.gov/features/MeasuringVegetation/measuring_vegetation_2.php
- Olaya, V. (2014). *Sistemas de Información Geográfica*. Obtenido de https://www.icog.es/TyT/files/Libro_SIG.pdf
- Oracle. (2023). *¿Qué es el aprendizaje automático*. <https://www.oracle.com/mx/artificial-intelligence/machine-learning/what-is-machine-learning/>
- Perilla, G. A. y Mas, J-F. (Abril de 2020). Google Earth Engine (GEE): una poderosa herramienta que vincula el potencial de los datos masivos y la eficacia del procesamiento en la nube. *Investigaciones Geográficas*(101), 1-6. doi:10.14350/rig.59929

- Pimenta, F. M., Speroto, A. T., Costa, M. H. y Dionizio, E. A. (2021). Historical Changes in Land Use and Suitability for Future Agriculture Expansion in Western Bahia, Brazil. *Remote sensing*, 13(6), 30. doi:10.3390/rs13061088
- Pineda Monroy, M. d. (2019). *Análisis Multitemporal de Cambio de Uso de Suelo en la Zona Metropolitana de Tianguistenco, Estado de México, Mediante Teledetección y SIG [Tesina de Especialidad no publicada]*. Universidad Autónoma del Estado de México.
- Pontius, R. G., Shusas, E. y McEachern, M. (2004). Detecting important categorical land changes while accounting for persistence. *Agriculture, Ecosystems and Environment* 101, 251-268. doi:10.1016/j.agee.2003.09.008
- PROFEPA. (2020). *Importancia de los Ecosistemas Forestales; Especies de los Bosques y Selvas*. Obtenido de Gobierno de México: <https://www.gob.mx/profepa/articulos/importancia-de-los-ecosistemas-forestales-especies-de-los-bosques-y-selvas?idiom=es>
- Protectora de Bosques del Estado de México. (31 de enero de 2024). *Convocatoria a participar en los Programas: Pago por Servicios Ambientales Hidrológicos del Estado de México, Restauración Ecológica de Ecosistemas Forestales, Capturando Carbono, Plantaciones Sustentables y Manejo Forestal Sustentable*. <https://probosque.edomex.gob.mx/sites/probosque.edomex.gob.mx/files/files/convocatorias/2024/convocatoria2024.pdf>
- Rahnama, M. R. (2021). Forecasting land use changes in Mashhad Metropolitan area using Cellular Automata and Markov chain model for 2016-2030. *Sustainable Cities and Society*, 64, 11. doi:10.1016/j.scs.2020.102548
- Ramsar. (2014). *Humedales de Importancia Internacional (Los Sitios RAMSAR)*. Obtenido de Ramsar: <https://www.ramsar.org/es/acerca-de/humedales-de-importancia-internacional-los-sitios-ramsar-0>
- Robotic Electronics. (2024). *Support Vector Machine*. Obtenido de Robotic Electronics: <https://roboticelectronics.in/svm/>
- Sacristán Romero, F. (2007). La teledetección satelital y los sistemas de protección ambiental. *Scripta Ethnologica*, XXIX, 79-90. <http://www.redalyc.org/articulo.oa?id=14802905>
- Salinas-Castillo, W. E., Terrazas-Ruiz, M., Mora-Olivo, A. y Paredes-Hernández, C. U. (2020). Análisis multitemporal de cambios de uso de la tierra en San Fernando, Tamaulipas. durante el periodo 1987 a 2017. *CienciaUAT*, 14(2), 160-173. doi:10.29059/cienciauat.v14i2.1298

- Sandoval-García, R., González-Cubas, R. y Jiménez-Pérez, J. (2021). Análisis multitemporal del cambio en la cobertura del suelo en la Mixteca Alta Oaxaqueña. *Revista Mexicana de Ciencias Forestales*, 12(66), 96-121. doi:10.29298/rmcf.v12i66.816
- Sarmiento, V. C. y Morales Medina, L. (2021). Inteligencia artificial aplicada al estudio de imágenes satelitales. *Revista TIA - Tecnología, Investigación y Academia*, 115-121. <https://revistas.udistrital.edu.co/index.php/tia/article/view/15720/17017>
- SEMARNAT. (5 de septiembre de 2018). ACUERDO por el que se da a conocer el resumen del Programa de Manejo del Área Natural Protegida con categoría de Área de Protección de Flora y Fauna la zona conocida como Ciénegas de Lerma.
- SEMARNAT. (2022). *Ordenamientos Ecológicos Expedidos*. Obtenido de Secretaría de Medio Ambiente y Recursos Naturales: <https://www.semarnat.gob.mx/gobmx/ordenamiento.html>
- SIGSA. (2022). SIGSA. <https://www.sigsa.info/es-mx/what-is-gis/overview>
- Solórzano Villegas, J. V. y Perilla Suárez, G. (2022). *Cómo usar Google Earth Engine y no fallar en el Intento* (Primera ed.). Morelia, Bogotá, México, Colombia: UNAM.
- Takada, T., Miyamoto, A. y Hasegawa, S. F. (2010). Derivation of a yearly transition probability matrix for land-use dynamics and its applications. *Landscape Ecology*, 25, 561-572.
- Talukdar, S., Singha, P., Mahato, S., Shahfahad, Pal, S., Liou, Y.-A. y Rahman, A. (2020). Land-Use Land-Cover Classification by Machine Learning Classifiers for Satellite Observations-A review. *Remote Sensing*, 12(7), 1-24. doi:10.3390/rs12071135
- Tempfli, K., Kerle, N., Huurneman, G. C. y Janssen, L. L. (2009). *Principles of Remote Sensing*. Enschede, The Netherlands: The International Institute for Geo-Information Science and Earth Observation.
- Trucíos Caciano, R., Estrada Ávalos, J., Cerano Paredes, J. y Rivera González, M. (2011). Interpretación del cambio en vegetación y uso de suelo. *Terra Latinoamericana*, 29(4), 359-367. http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0187-57792011000400359&lng=es&tlng=es
- USGS. (s.f). *Landsat 8*. Obtenido de USGS: <https://www.usgs.gov/landsat-missions/landsat-8>

- Vázquez Jiménez, R., Romero Calcerrada, R., Ramos Bernal, R. N., Arrogante Funes, P. y Novillo, C. J. (2021). An Alternative Method for the Generation of Consistent Mapping to Monitoring Land Cover Change: A Case Study of Guerrero State in Mexico. *Land*, 10(7), 24. doi:10.3390/land10070731
- Wang, S. W., Munkhnasan, L. y Lee, W.-K. (2021). Land use and land cover change detection and prediction in Bhutan's high altitude city of Thimphu, using cellular automata and Markov chain. *Environmental Challenges*, 2. Obtenido de <https://www.sciencedirect.com/science/article/pii/S2667010020300172>
- Ye, J., Hu, Y., Zhen, L., Wang, H. y Zhang, Y. (2021). Analysis on Land-Use Change and Its Driving Mechanism in Xilingol, China, during 2000-2020 Using the Google Earth Engine. *Remote sensing*, 23. doi:10.3390/rs13245134

Anexos

Anexo 1. Código en Google Earth Engine con Landsat 8

```
Landsat2022 Get Link Save Run Reset
Imports (8 entries)
var ZMT: Table projects/ee-luzpineda/assets/limite
var bosque: FeatureCollection (26 elements)
var agricultura: FeatureCollection (9 elements)
var asentamientos: FeatureCollection (16 elements)
var agua: FeatureCollection (5 elements)
var tular: FeatureCollection (4 elements)
var pastizal: FeatureCollection (17 elements)
var imageVisParam: SR_B6, SR_B5 and SR_B2 from 7367.1 to 19275.9

1 var L8_2022 = ee.ImageCollection("LANDSAT/LC08/C02/T1_L2");//colección de imagenes landsat 8
2   .filterBounds(ZMT)//Filtro para recortar el área de estudio
3   .filterDate('2022-12-01', '2022-12-31')//Filtro de fecha
4   .sort('CLOUD_COVER');//Ordenar una colección por la cobertura de nubes
5
6 var L8_22 = L8_2022.first ();//Devuelve la primera entrada de una colección determinada.
7 print (L8_22);//Imprime la información de la imagen satelital
8
9 var clipL8_22 = L8_22.clip(ZMT); //Recorte del área de estudio
10
11 //Variable de muestras de entrenando - clases para uso de suelo
12 var muestras = bosque.merge(agricultura).merge(asentamientos).merge(agua).merge(tular).merge(pastizal);
13
14 /*Variable de entrenamiento - Convierte cada píxel de una imagen (a una escala determinada)
15 que intersecta una o más regiones en una característica y las devuelve como una colección de características.
16 Cada característica de salida tendrá una propiedad por banda de la imagen de entrada, así como cualquier propiedad
17 especificada copiada de la característica de entrada*/
18 var entrenamiento = clipL8_22.sampleRegions(muestras, ['clase'], 30);//Escala de imagenes Landsat
19 print (entrenamiento.size())
20
21 //Variable de bandas en Landsat 8
22 var bandas = ['SR_B2', 'SR_B3', 'SR_B4', 'SR_B5', 'SR_B6', 'SR_B7', 'ST_B10'];
23
24 //clasificación Random Forest
25 var clasificacion_rf = ee.Classifier.smileRandomForest(10).train(entrenamiento, 'clase', bandas);
26 var imagen_rf2022 = clipL8_22.classify(clasificacion_rf);
27
28 //Clasificación Support Vector Machine
29 var clasificacion_svm = ee.Classifier.libsvm().train (entrenamiento, 'clase', bandas);
30 var imagen_svm2022 = clipL8_22.classify(clasificacion_svm);
31
32 //Agregar el límite del área de estudio
33 Map.addLayer (ZMT, {color: "black"}, 'ZMT');
34 //Agregar la composición RGB del área de estudio
35 Map.addLayer (clipL8_22, {bands:['SR_B6', 'SR_B5', 'SR_B2'], min: 7367, max:19275}, 'zmt_corte');
36 //Agregar la clasificación obtenida a partir de Random Forest
37 Map.addLayer(imagen_rf2022,{min:1, max:6, palette:['#196f3d', '#f7dc6f', '#99a3a4', '#3498db', '#c39bd3', '#eb984e']},
38 'RF');
39 //Agregar la clasificación obtenida a partir de Support Vector Machine
40 Map.addLayer(imagen_svm2022,{min:1, max:6, palette:['#196f3d', '#f7dc6f', '#99a3a4', '#3498db', '#c39bd3', '#eb984e']},
41 'SVM');
42
43 // Estima la matriz de confusión que represente la precisión de la clasificación para RF.
44 print('RF error matrix: ', clasificacion_rf.confusionMatrix());
45 print('RF accuracy: ', clasificacion_rf.confusionMatrix().accuracy());
46
47 // Estima la matriz de confusión que represente la precisión de la clasificación para SVM
48 print('SVM error matrix: ', clasificacion_svm.confusionMatrix());
49 print('SVM accuracy: ', clasificacion_svm.confusionMatrix().accuracy());
50
51
52 //Exportar la imagen de Random Forest
53 Export.image.toDrive({
54   image: imagen_rf2022,
55   description: 'L8_classRF_2022',
56   scale: 30, // Ajusta la escala según tus necesidades
57   region: ZMT.geometry(),
58   folder: 'GEE_exportar', // Cambia la carpeta según tus necesidades
59   maxPixels: 1e13, // Ajusta según tus necesidades
60   crs: 'EPSG:4326'
61 });
62 //Exportar la imagen Support Vector Machine
63 Export.image.toDrive({
64   image: imagen_svm2022,
65   description: 'L8_classSVM_2022',
66   scale: 30, // Ajusta la escala según tus necesidades
67   region: ZMT.geometry(),
68   folder: 'GEE_exportar', // Cambia la carpeta según tus necesidades
69   maxPixels: 1e13, // Ajusta según tus necesidades
70   crs: 'EPSG:4326'
71 });
```

Elaboración propia en GEE

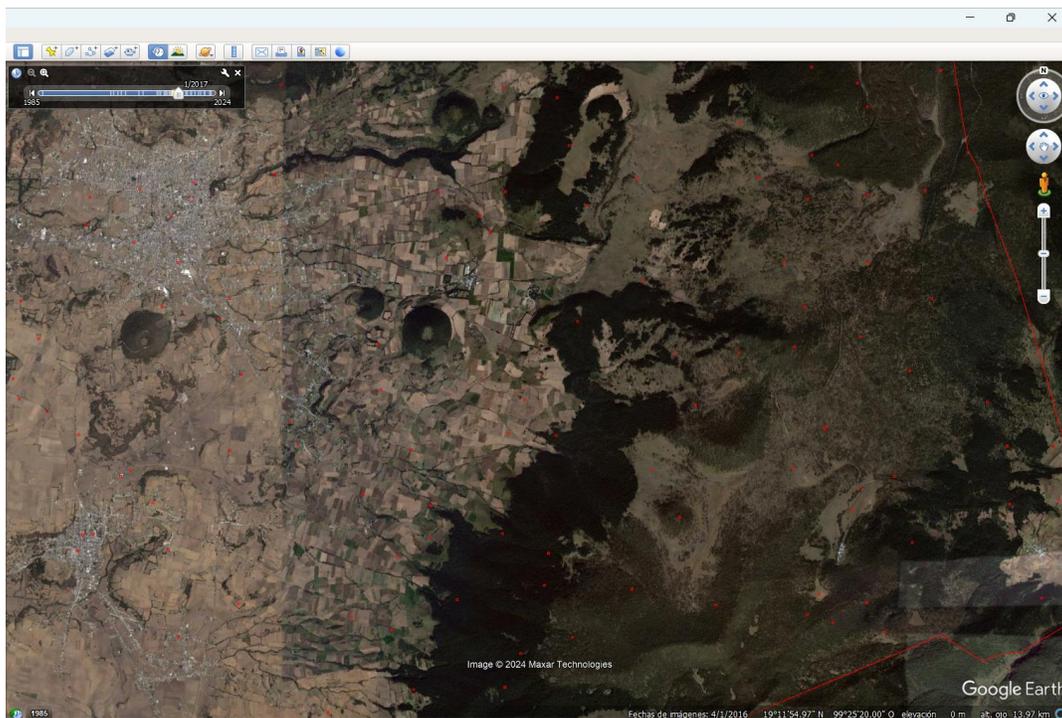
Anexo 2. Código en Google Earth Engine con Sentinel-2

```
Sentinel2_2022 * Get Link Save Run
Imports (8 entries)
  var imageVisParam: B11, B8A and B2 from 214.94 to 5396.06
  var bosque: FeatureCollection (31 elements)
  var agricultura: FeatureCollection (23 elements)
  var asentamientos: FeatureCollection (18 elements)
  var agua: FeatureCollection (4 elements)
  var tular: FeatureCollection (7 elements)
  var pastizal: FeatureCollection (27 elements)
  var ZMT: Table projects/ee-luzpineda/assets/limite

1 var S2_H = ee.ImageCollection("COPERNICUS/S2_SR_HARMONIZED");//Colección de imagenes Sentinel-2
2   .filterBounds(ZMT)//Filtro para recortar el área de estudio
3   .filterDate('2022-12-01', '2022-12-15')//Filtro de fecha
4   .sort('CLOUD_COVER');//Ordenar una colección por la cobertura de nubes
5
6 var S2 = S2_H.first();//Devuelve la primera entrada de una colección determinada.
7 print(S2);//Imprime la información de la imagen satelital
8
9 var clipSE = S2.clip(ZMT); //Recorte del área de estudio
10
11 //Agregar composición RGB del área de estudio
12 Map.addLayer(clipSE, {bands: ['B11', 'B8A', 'B2'], min: 0, max: 3000}, 'clipSE');
13
14 /*Variable de entrenamiento - Convierte cada píxel de una imagen (a una escala determinada)
15 que intersecta una o más regiones en una característica y las devuelve como una colección de características.
16 Cada característica de salida tendrá una propiedad por banda de la imagen de entrada, así como cualquier propiedad
17 especificada copiada de la característica de entrada*/
18 var muestras = bosque.merge(agricultura).merge(asentamientos).merge(agua).merge(tular).merge(pastizal);
19 var entrenamiento = clipSE.sampleRegions(muestras, ['clase'], 10); //Escala de imágenes Sentinel-2
20 print(entrenamiento.size());
21
22 //Variable de bandas en Sentinel-2
23 var bandas = ['B2', 'B3', 'B4', 'B5', 'B6', 'B7', 'B8', 'B8A', 'B11', 'B12'];
24 //Clasificación Random Forest
25 var clasificacion_rf = ee.Classifier.smileRandomForest(50).train(entrenamiento, 'clase', bandas);
26 var imagens2_rf2022 = clipSE.select(bandas).classify(clasificacion_rf);
27 //Clasificación Support Vector Machine
28 var clasificacion_svm = ee.Classifier.libsvm().train(entrenamiento, 'clase', bandas);
29 var imagens2_svm2022 = clipSE.select(bandas).classify(clasificacion_svm);
30
31 //Agregar la clasificación obtenida a partir de Random Forest
32 Map.addLayer(imagens2_rf2022, {min:1, max:6, palette:['#196f3d', '#f7dc6f', '#99a3a4', '#3498db', '#c39bd3', '#eb984e']},
33 'RF_S2_2022');
34 //Agregar la clasificación obtenida a partir de Support Vector Machine
35 Map.addLayer(imagens2_svm2022, {min:1, max:6, palette:['#196f3d', '#f7dc6f', '#99a3a4', '#3498db', '#c39bd3', '#eb984e']},
36 'SVM_S2_2022');
37
38 // Estima la matriz de confusión que represente la precisión de la clasificación para RF.
39 print('RF error matrix: ', clasificacion_rf.confusionMatrix());
40 print('RF accuracy: ', clasificacion_rf.confusionMatrix().accuracy());
41
42 // Estima la matriz de confusión que represente la precisión de la clasificación para SVM
43 print('SVM error matrix: ', clasificacion_svm.confusionMatrix());
44 print('SVM accuracy: ', clasificacion_svm.confusionMatrix().accuracy());
45
46 //Exportar la imagen de Random Forest
47 Export.image.toDrive({
48   image: imagens2_rf2022,
49   description: 'S2_classRF_2022',
50   scale: 10, // Ajusta la escala según tus necesidades
51   region: ZMT.geometry(),
52   folder: 'GEE_exportar', // Cambia la carpeta según tus necesidades
53   maxPixels: 1e13, // Ajusta según tus necesidades
54   crs: 'EPSG:4326'
55 });
56
57 //Exportar la imagen Support Vector Machine
58 Export.image.toDrive({
59   image: imagens2_svm2022,
60   description: 'S2_classSVM_2022',
61   scale: 10, // Ajusta la escala según tus necesidades
62   region: ZMT.geometry(),
63   folder: 'GEE_exportar', // Cambia la carpeta según tus necesidades
64   maxPixels: 1e13, // Ajusta según tus necesidades
65   crs: 'EPSG:4326'
66 });
67
68
```

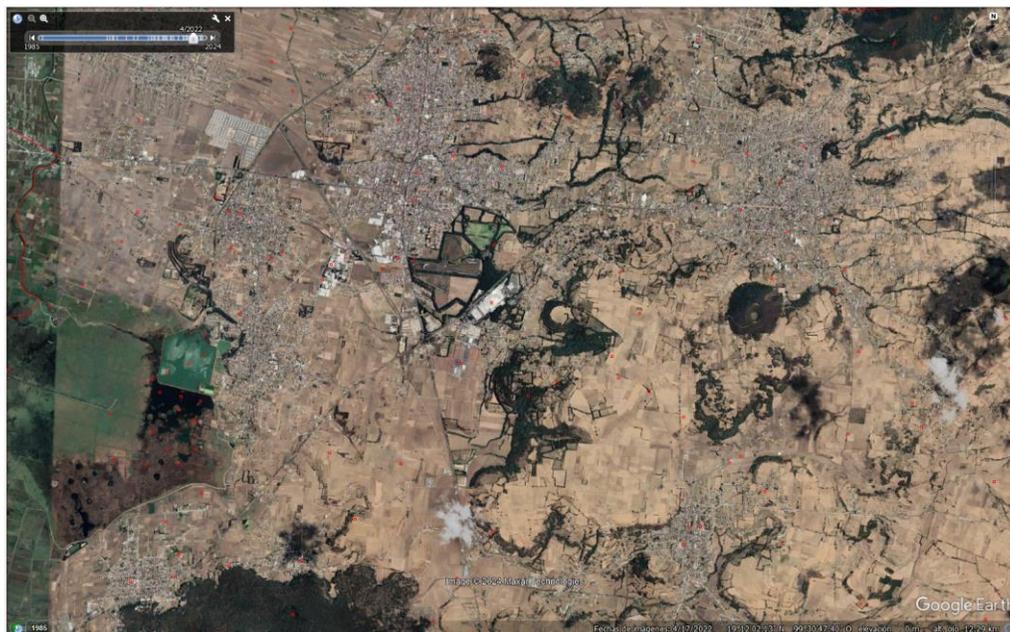
Elaboración propia en GEE

Anexo 3. Verificación en imagen Maxar para validar año 2016



Obtenido de Google Earth Pro

Anexo 4. Verificación en imagen Maxar para validar año 2022



Obtenido de Google Earth Pro

Anexo 5. Verificación en campo en área afectada por incendios forestales
(Municipio de Xalatlaco)



Anexo 6. Verificación en campo (de fondo Volcán Negro con presencia de plaga)



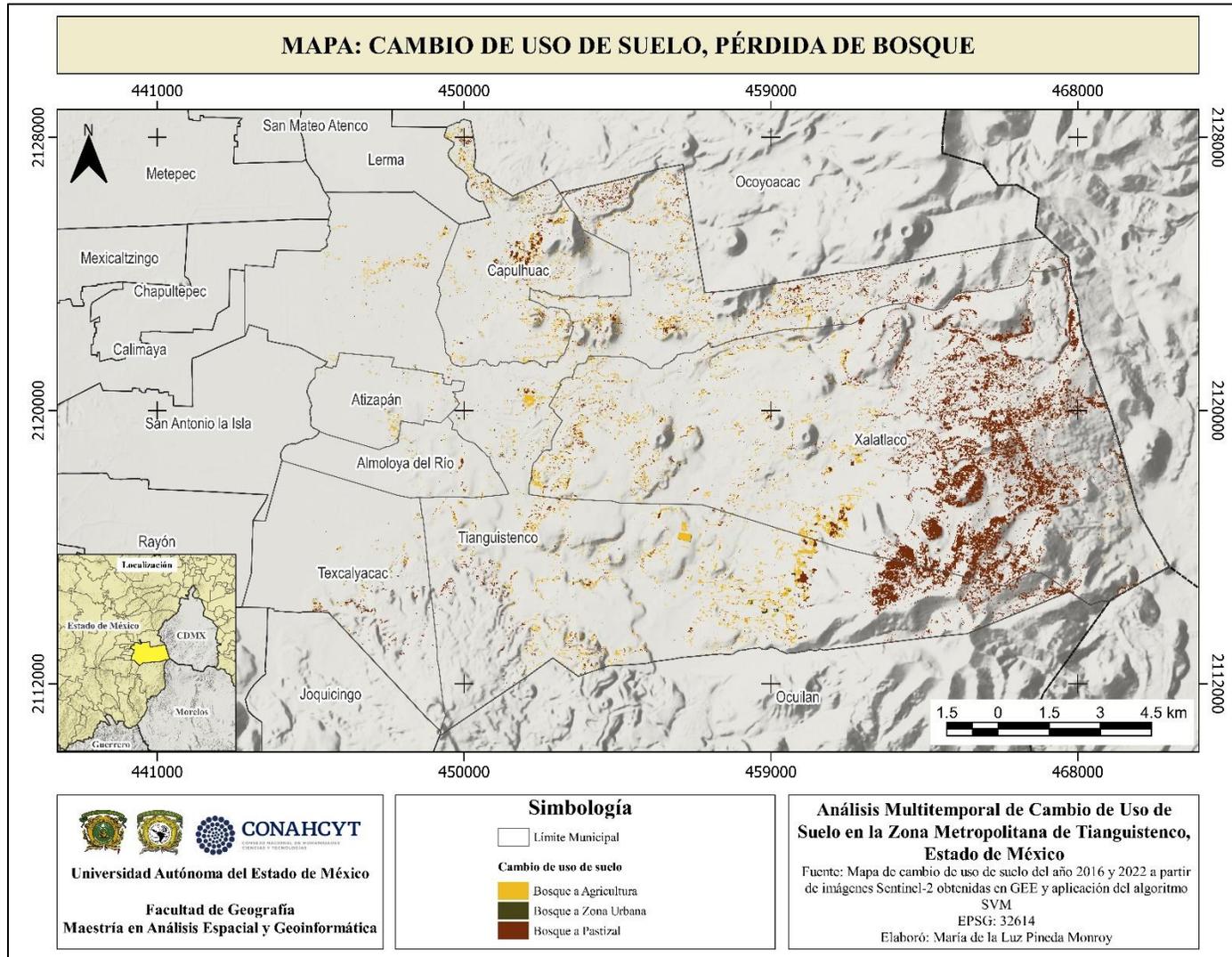
Anexo 7. Ciénegas del Lerma (Laguna Chignahuapan - Almoloya del Río)



Anexo 8. Verificación en campo en el municipio de Tianguistenco (en área agrícola)

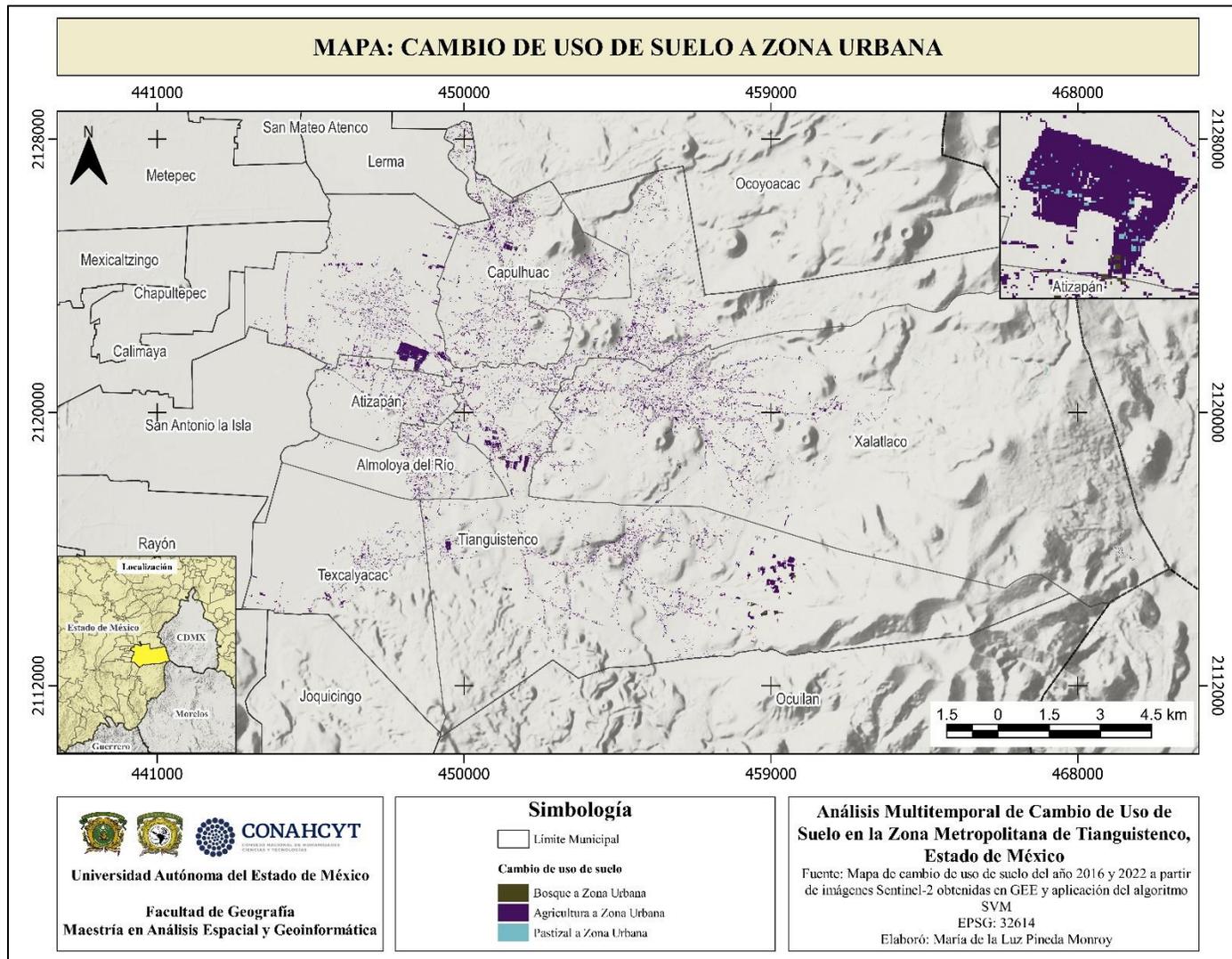


Anexo 9. Cambios de bosque a otros usos de suelo



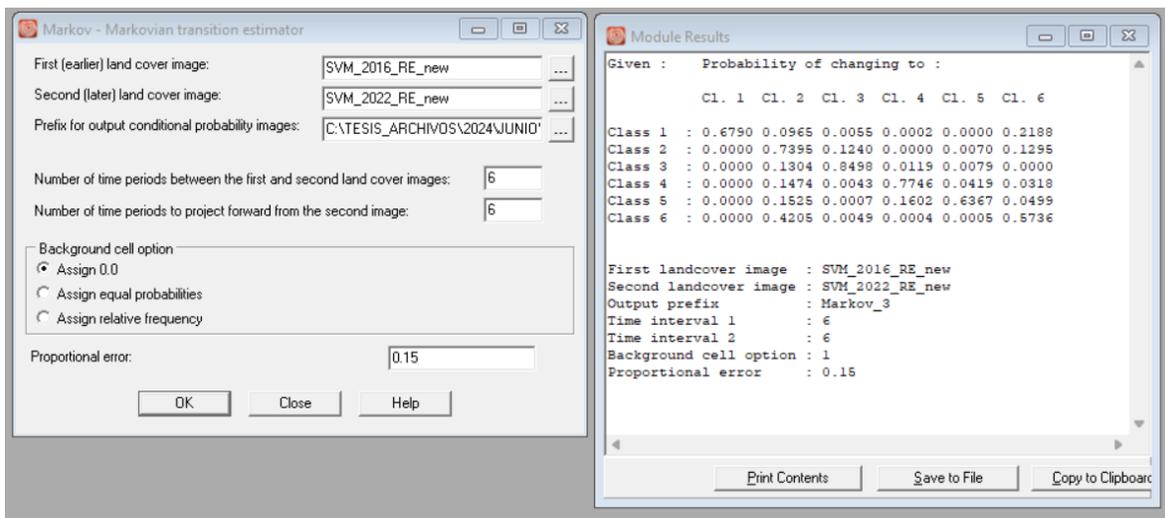
Elaboración propia a partir de mapa de uso de suelo del año 2016 y 2022 a partir de imágenes Sentinel-2

Anexo 10. Cambios a zona urbana de 2016 a 2022



Elaboración propia a partir de mapa de uso de suelo del año 2016 y 2022 a partir de imágenes Sentinel-2

Anexo 11. Módulo Markov



Anexo 12. Representación cartográfica de la probabilidad de cambio



Anexo 13. CA_MARKOV

