



UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MÉXICO
INSTITUTO INTERAMERICANO DE TECNOLOGÍA Y CIENCIAS DEL AGUA



*“Clasificación de cubiertas vegetales por sensores remotos con ACP en
Distritos de Riego del Altiplano Mexicano”*

Tesis que para obtener el grado de

MAESTRA EN CIENCIAS DEL AGUA

Presenta:

Lic. Marlen Lopez Gutierrez

TUTOR ACADÉMICO:

Dr. Carlos Roberto Fonseca Ortiz

TUTORAS ADJUNTAS:

Dra. Marivel Hernández Téllez

Dra. Rocio Becerril Piña

Toluca, México, noviembre 2022

RESUMEN

Derivado del constante cambio de uso /cobertura del suelo a nivel global, principalmente aquel que afecta a la vegetación, la evaluación de impactos sobre recursos naturales e hídricos ha generado una gran necesidad de identificar temporal y territorialmente coberturas vegetales para la toma de decisiones relacionados con producción agrícola, gestión de recursos hídricos y zonas de conservación ambiental, entre otros.

El uso de sensores remotos, tales como imágenes satelitales, ha permitido el monitoreo y evaluación de estos cambios de uso de suelo. Para ello, es usual encontrar diversas técnicas de clasificación supervisada con resultados satisfactorios. Esta supervisión se refiere a la asignación *a priori* de clases para lugares específicos, conocidos como sitios de entrenamiento. Por lo tanto, la clasificación supervisada, aunque proporciona resultados precisos, se encuentra limitada a la experiencia y disponibilidad de información sobre la región a clasificar en un periodo de tiempo en concreto.

El presente proyecto tiene como finalidad proporcionar las bases para una clasificación no supervisada de la cubierta vegetal, a partir de la alimentación de un modelo con información validada tanto en el espacio como en el tiempo. La información que alimenta a este modelo es conformada por índices espectrales de vegetación, características espectrales propias de la vegetación (clorofila y carotenoides) y variables ambientales como la precipitación y temperatura. Esta información es sintetizada por medio de un análisis de componentes principales y validada desde diferentes criterios de clasificación considerados por la FAO. Si bien, la clasificación de cubiertas vegetales, basada en el método de máxima verosimilitud, requiere de una validación con información de sitios de entrenamiento, el modelo conserva la distribución y dispersión de los componentes principales de las escenas para que sean consideradas en aquellas regiones sin información disponible. En efecto, el presente modelo genera una clase “sin identificar” para aquellas celdas cuyas características no correspondan a alguna de las alimentadas en el proceso.

Este modelo fue desarrollado con base en la información de distritos de riego en el Altiplano Mexicano y con tres criterios mencionados por la FAO: cubierta vegetal y no vegetal, ambiente (acuático y terrestre) y práctica (vegetación natural y regulada). Los resultados mostraron clasificaciones aceptables para el primer y tercer criterio, mientras que el segundo mostró limitaciones muy marcadas. Con ello, el objetivo del proyecto fue alcanzado en términos de sentar las bases de una metodología para una clasificación no

supervisada con base en algunos criterios de clasificación de vegetación. Sin embargo, se recomienda extender el experimento hacia clasificaciones más detalladas de vegetación y validarlas en zonas adicionales a las consideraciones en el modelo.

Palabras clave: Identificación de vegetación, sensores remotos, clasificación no supervisada.

ABSTRACT

Given the constant change in land use/cover at a global level, mainly that which affects vegetation, the assessment of impacts on natural and water resources has generated a need to temporarily and territorially identifying vegetation covers for decision-making in agricultural production, management of water resources and environmental conservation areas, among others.

The use of remote sensors, such as satellite images, has allowed the monitoring and evaluation of these changes in land use. For this, it is usual to find several supervised classification techniques with satisfactory results. This supervision refers to the *a priori* assignment of classes to specific places, known as training sites. Therefore, the supervised classification, although it provides accurate results, is limited to the experience and availability of information on the region to be classified in a specific period.

The purpose of this project is to provide the basis for an unsupervised classification of vegetation cover, based on feeding a model with information validated both in space and in time. The information that feeds this model is made up of spectral indices of vegetation, characteristics of the vegetation (chlorophyll and carotenoids) and environmental characteristics such as precipitation and temperature. This information is synthesized through a principal components analysis and validated from different classification criteria considered by the FAO. Although the classification of vegetation covers, based on the maximum likelihood method, requires validation with information from training sites, the model preserves the distribution and dispersion of the principal components of the scenes so that they could be considered in those regions without available information. Indeed, the present model generates an "unidentified" class for those cells whose characteristics do not correspond to any of those fed into the process.

This model was developed based on information from irrigation districts in the Mexican Highlands and with three of the criteria mentioned by the FAO: vegetal and non-vegetable cover, environment (aquatic and terrestrial) and practice (natural and regulated

vegetation). The results showed acceptable classifications for the first and third criteria, while the second showed very marked limitations. With this, the objective of the project was achieved in terms of laying the basis of a methodology for an unsupervised classification based on some vegetation classification criteria. However, it is recommended to extend the experiment towards more detailed vegetation classifications and validate them in additional zones to those regarded in the model.

Key words: Vegetation identification, remote sensing, unsupervised classification

ÍNDICE

RESUMEN	III
ÍNDICE DE FIGURAS	VII
ÍNDICE DE TABLAS	VIII
NOMENCLATURA, ACRÓNIMOS Y UNIDADES	IX
CAPÍTULO 1. INTRODUCCIÓN	1
1.1 JUSTIFICACIÓN CONTEXTUAL Y CIENTÍFICA	2
1.2 HIPÓTESIS	4
1.3 OBJETIVOS	4
1.3.1 <i>Objetivo General</i>	4
1.3.2 <i>Objetivos Específicos</i>	4
CAPÍTULO 2. ANTECEDENTES DE LA IDENTIFICACIÓN DE CUBIERTAS VEGETALES CON SENSORES REMOTOS	5
2.1 CONTEXTO INTERNACIONAL	5
2.2 CONTEXTO NACIONAL	9
CAPITULO 3. MARCO TEÓRICO	11
3.1 CRITERIOS DE DELIMITACIÓN PARA LA DETERMINACIÓN DE LA VEGETACIÓN	11
3.2 ÍNDICES DE VEGETACIÓN	13
3.3 ANÁLISIS DE COMPONENTES PRINCIPALES (ACP)	15
3.3.1 <i>Aplicaciones del análisis de componentes principales</i>	15
3.3.2 <i>Cálculo de las componentes principales</i>	16
3.3.3 <i>Interpretación geométrica del ACP</i>	17
CAPÍTULO 4. METODOLOGÍA	19
1.1 FASE I. DETERMINACIÓN DE ESTRUCTURAS JERÁRQUICAS DE LOS CULTIVOS	19
5.1 FASE II. DESARROLLO DEL MODELO	26
5.2 FASE III. VALIDACIÓN DE RESULTADOS	29
CAPITULO 5. RESULTADOS Y DISCUSION	30
CONCLUSIONES	43
RECOMENDACIONES	45
CAPITULO 7. REFERENCIAS	46
ANEXOS	49

Índice de figuras

FIGURA 1. REPRESENTACIÓN GEOMÉTRICA DE COMPONENTES PRINCIPALES	18
FIGURA 2. DIAGRAMA DE LA METODOLOGÍA	19
FIGURA 3. CLASIFICACIÓN DEL TIPO DE CULTIVOS	21
FIGURA 4. MAPA DE UBICACIÓN DEL ÁREA DE ESTUDIO. COMPUESTO FALSO COLOR (345) A PARTIR DE LA IMAGEN LANDSAT 8 (2020-11-22 Y 2021-01-02)	22
FIGURA 5A Y 5B. CLASIFICACIONES PUNTUALES RESULTANTES	29
FIGURA 6. DIAGRAMAS CON CRITERIOS PARA LA CLASIFICACIÓN DE CUBIERTAS VEGETALES/CULTIVOS A) CLASIFICACIÓN DE ACUERDO CON SU ETAPA FENOLÓGICA; B) CLASIFICACIÓN POR CULTIVO O VEGETACIÓN.	31
FIGURA 7. PUNTOS DE MUESTREO SELECCIONADOS PARA EL AÑO 2013	31
FIGURA 8. RESULTADOS DE ÍNDICES DE VEGETACIÓN 2013-04-18	33
FIGURA 9.A) GRÁFICA DE LOS COEFICIENTES KAPPA (PORCENTAJE) 2013-2016, PARA LAS TRES CLASES; B) FRECUENCIA DE LOS VALORES DE KAPPA PARA ÁREAS CON VEGETACIÓN; C) FRECUENCIA DE COEFICIENTE KAPPA PARA ZONAS SIN VEGETACIÓN; D) KAPPA PARA ÁREAS NO IDENTIFICADAS.....	35
FIGURA 10. EJEMPLO DE IMÁGENES CLASIFICADAS EN EL PRIMER NIVEL (ÁREAS CON VEGETACIÓN Y SIN VEGETACIÓN).....	37
FIGURA 11. COEFICIENTES KAPPA TOTAL Y POR CLASE ENTRE 2013-2016	39
FIGURA 12. RESULTADOS DE LA CLASIFICACIÓN PARA ÁREAS CON VEGETACIÓN TERRESTRE Y ACUÁTICA	39
FIGURA 13. COEFICIENTES KAPPA TOTAL Y POR CLASE PARA VEGETACIÓN REGULADA Y NATURAL.....	41
FIGURA 14. CLASIFICACIÓN REGULADA Y NATURAL	42

Índice de tablas

TABLA 1. CULTIVOS DEL DR 003 TULA HIDALGO (CONAGUA, 2015, 2017).....	23
TABLA 2. CULTIVOS DEL DR 023 SAN JUAN DEL RÍO (CONAGUA, 2015, 2017)	24
TABLA 3. PRINCIPALES CULTIVOS DEL DR 033 ESTADO DE MÉXICO (CONAGUA, 2015, 2017).....	25
TABLA 4. PRINCIPALES CULTIVOS DEL DR 044 JILOTEPEC (CONAGUA, 2015, 2017)	25
TABLA 5. PRINCIPALES CULTIVOS DEL DR 096 ARROYOZARCO (CONAGUA, 2015, 2017)	26
TABLA 6. CARACTERÍSTICAS GENERALES DEL SENSOR A UTILIZAR. BASADO EN: EARTH EXPLORER, 2021.....	26
TABLA 7. FÓRMULAS DE ÍNDICES DE VEGETACIÓN ESPECTRAL (IV) (BASADO EN: GAO, 1995; HUETE, 1988; HUETE, 2002; DAUGHTRY ET AL., 2000).....	28
TABLA 8. DISPONIBILIDAD DE IMÁGENES LANDSAT DEL 2013 - 2016	32
TABLA 9. VALORES DEL ÍNDICE KAPPA DEL 2013 – 2016	34
TABLA 10. ERRORES TOTALES PONDERADOS POR AÑO.....	36
TABLA 11. VALORES DE LOS EIGEN VECTOR PARA ÁREAS TERRESTRES Y ACUÁTICAS	38
TABLA 12. DATOS RESULTANTES PARA LAS FIRMAS ESPECTRALES	38
TABLA 13. VALORES DE LOS EIGEN VECTOR PARA ÁREAS NATURALES Y REGULADAS	40
TABLA 14. DATOS RESULTANTES PARA LAS FIRMAS ESPECTRALES	40

Nomenclatura, acrónimos y unidades

Símbolo	Descripción
ACP	Análisis de Componentes Principales
APAR	Radiación absorbida fotosintéticamente por la planta (por sus siglas en inglés)
ARVI	Atmospherically Resistant Vegetation Index (adimensional)
ASTER	Advanced Spaceborne Thermal Emission and Reflection
AUC	Area under the curve
Bm	Biomasa
CONABIO	Comisión Nacional para el Conocimiento y Uso de la Biodiversidad
CONAGUA	Comisión Nacional del Agua
CP	Componente principal
DR	Distritos de Riego
ET	Lámina de evapotranspiración (mm)
ETM+	Enhanced Thematic Mapper Plus
ETP	Lámina de evapotranspiración potencial (mm)
FAO	Organización de las Naciones Unidas para la Agricultura y la Alimentación
GPS	Global Positioning System
IITCA	Instituto Interamericano de Tecnología y Ciencias del Agua
INEGI	Instituto Nacional de Estadística y Geografía
IV	Índice de vegetación
k_c	Coefficiente de cultivo (adimensional)
IVIS	índice de vegetación iso-suelo (adimensional)
LAI	Índice de área foliar (por sus siglas en inglés) (adimensional)
m	Metros
MCARI	Índice de relación de absorción de clorofila modificado (adimensional)
MODIS	Moderate Resolution Imaging Spectroradiometer
NDVI	Normalized Difference Vegetation Index (adimensional)
ROC	Receiver Operating Characteristic
SAVI	Soil Adjusted Vegetation Index (adimensional)
SIAP	Servicio de Información Agroalimentaria y Pesquera
SIG	Sistemas de información geográfica
SINA	Sistema Nacional de Información del Agua
SMN	Servicio Meteorológico Nacional
SO	Suroeste
TM	Thematic Mapper
USGS	United States Geological Survey

CAPÍTULO 1. INTRODUCCIÓN

Actualmente el cambio de uso de suelo ha incrementado de manera importante en todo el mundo, principalmente de dos formas: como zona habitacional, debido al crecimiento exponencial de la población (Rogan & chen, 2004), y la segunda forma se centra en áreas para cultivo, que se ha vuelto un factor importante tanto en economía local como regional (Marini et al., 2007). A partir de lo anterior, se han implementado técnicas que permiten monitorear estos cambios de forma más exacta y puntual.

La gestión eficiente de los recursos hídricos para riego, la identificación de los cultivos en parcelas específicas y la evaluación de las prácticas de manejo del suelo son solo algunas de las razones más relevantes que hacen importante el monitoreo del cambio uso/ cobertura (LUCC) del suelo, principalmente en aquella correspondiente a la vegetación (Peña-Barragán et al., 2011; Pôças et al., 2020).

Una de las técnicas para monitoreo de la cobertura vegetal es la teledetección, ya que permite la observación remota de la superficie terrestre mediante el análisis de patrones de cobertura y su discriminación a través de los sistemas de información geográfica (SIG). La información obtenida de los SIG y de la teledetección brinda ventajas en los costos de los insumos y menor tiempo en su preparación (Rojas Núñez & Rodríguez Jiménez, 2020). La teledetección a través de sensores remotos, constituyen una herramienta de gran ayuda para la implementación de programas de monitoreo de la humedad de los suelos con distintos tipos de cultivos, ya que permite identificar esas zonas de cultivos o con otros tipos de vegetación (Arango, 2020) y, permite hacer un uso eficiente o controlado de los recursos hídricos para evitar o reducir el deterioro ambiental o del recurso (Gómez & Fonseca, 2013).

En la presente investigación, se propone desarrollar una metodología para identificar distintos tipos de vegetación (bosques, cultivos de riego o temporal, pastizales, entre otros), y, con ello, clasificarlas en diferentes cubiertas vegetales de manera automatizada, en la Altiplanicie Mexicana, abarcando los distritos de riego 003, 0023, 033, 044 y 096, utilizando, no solo información de sensores remotos que miden longitudes de onda de luz absorbidas y reflejadas por la vegetación, sino datos estimados sobre características vegetales (como el índice de área foliar y carotenoides) y ambientales (como precipitación y temperatura).

En este sentido, de acuerdo con Ayala-Carcedo & Olcina Cantos (2002) a partir de una imagen de satélite se pueden obtener parámetros que indican el estado de la vegetación, por ejemplo, ciertos pigmentos en las hojas de la planta absorben fuertemente longitudes de onda de luz visible. Las hojas en sí mismas reflejan fuertemente longitudes de onda de luz infrarroja cercana, que es invisible para el ojo humano (USGS, 2020). También pueden indicar la cantidad de biomasa existente, el grado de cobertura vegetal, el índice de área foliar (LAI), la radiación absorbida fotosintéticamente por la planta (APAR) y componentes relacionados con la evapotranspiración (ETP), esto, mediante la integración de índices espectrales de vegetación (USGS, 2020). Con las estimación de estos elementos, se puede llevar a cabo un Análisis de Componentes Principales (ACP), que permite disminuir las dimensiones del modelo y generar clúster para la identificación y delimitación de la vegetación adecuada.

Para la identificación de los distintos tipos de vegetación se propusieron estructuras jerárquicas basadas en criterios proporcionados por la FAO (2000). La base de estas estructuras inicio con la delimitación de zonas sin vegetación y con vegetación. Posteriormente, se manejaron niveles de identificación considerando el ambiente de crecimiento de la vegetación, la existencia de alguna práctica de regulación, entre otros.

1.1 Justificación contextual y científica

La capacidad de los recursos naturales del mundo para mantener el desarrollo de la población es una cuestión fundamental para la comunidad internacional. Debido a que la población mundial continúa creciendo un 1.6% por año, superando el 3% anual en muchos de los países menos desarrollados (FAO, 1997). Al mismo tiempo, la disponibilidad de los recursos naturales esenciales, tales como tierras y aguas, están disminuyendo en cantidad y calidad. Es decir, el problema básico es el aumento de la presión ejercida sobre estos recursos naturales, por lo cual se hace evidente la necesidad de una evaluación y monitoreo de su estado.

Para el conocimiento actualizado de la tierra, García & Martínez (2010) manifiestan que el uso de imágenes de sensores remotos en la agricultura, se ha planteado como una alternativa para obtener datos cuantitativos que permiten ubicar y caracterizar los cultivos, en lo que se ha denominado agricultura de precisión. Esto, con el fin de tomar decisiones para su planificación y manejo, abarcando grandes extensiones de terreno y facilitando información tanto histórica como en zonas de difícil acceso (Acevedo & Hernández, 2013).

Además, en los últimos años, debido al gran avance en el desarrollo de los sensores remotos en cuanto al número de sistemas disponibles, la resolución espacial, espectral y temporal y con la tendencia a disminuir el costo, han surgido un gran número de aplicaciones. Algunas de ellas son la determinación de su estado nutricional, el manejo del agua (ya que se puede conocer la demanda de agua, su zonificación y con ello su distribución), manejo sanitario, predicción de cosechas, recursos genéticos, zonificación de aptitud de tierras, hasta la agricultura de precisión (García & Martínez, 2010).

En el presente estudio se pretendió desarrollar un método para la identificación y clasificación de cubiertas vegetales (incluyó bosques y áreas reguladas como lo son los cultivos), basado en análisis de componentes principales (ACP) y clúster. A diferencia de los árboles de decisión, que se trabajan con todo un conjunto de variables, el ACP permitió disminuir las dimensiones de los datos, minimizando la dispersión de la varianza en las variables.

El método propuesto, es integrador, dado que las agrupaciones de elementos clasificados crecen conforme otras zonas con cubiertas vegetales diferentes sean agregados al modelo. El estudio basado en sensores remotos, se pretendió realizar en primera instancia para la altiplanicie mexicana, como área principal, con la finalidad de identificar diferentes tipos de vegetación y teniendo como base el conocimiento previo de lo que hay en la zona de estudio, para posteriormente ser aplicado a diferentes áreas.

1.2 Hipótesis

La estimación de componentes principales, basados en índices espectrales, variables climatológicas y características vegetales (principalmente pigmentos), por medio de sensores remotos permite la identificación de la estructura jerárquica más adecuada en la clasificación de diferentes tipos de vegetación en la zona central del altiplano mexicano.

1.3 Objetivos

1.3.1 Objetivo General

Desarrollar un modelo basado en Análisis de Componentes Principales (ACP) para la identificación de diferentes tipos de vegetación por medio de sensores remotos, tomando como referencia Distritos de Riego del Altiplano Central Mexicano.

1.3.2 Objetivos Específicos

- ❖ Analizar diferentes criterios de clasificación de vegetación mostrados por la FAO para la identificación de posibles estructuras jerárquicas.
- ❖ Estimar índices espectrales, propiedades espectrales de la vegetación (pigmentos), y obtener variables climatológicas, desde información obtenida por sensores remotos para la realización de un análisis de componentes principales en la zona de los Distritos de Riego 003, 0023, 033, 044 y 096
- ❖ Obtener información histórica de cultivos en los distritos de riego en estudio para la determinación de sitios de entrenamiento del modelo.
- ❖ Validar las estructuras jerárquicas en relación con una clasificación por máxima verosimilitud para la obtención de la más adecuada en la clasificación.

CAPÍTULO 2. ANTECEDENTES DE LA IDENTIFICACIÓN DE CUBIERTAS VEGETALES CON SENSORES REMOTOS

El número de aplicaciones de la teledetección está en constante crecimiento, gracias a las innovaciones en los sensores y en los métodos de tratamiento (Chuvienco, 2002). Ejemplo de esto es que la teledetección posibilita a través de la clasificación de imágenes satelitales, conocer el uso de los suelos y monitorear sus cambios, que suceden muy rápidamente (Nolasco et al., 2015). El análisis multitemporal permite detectar cambios entre diferentes fechas de referencia, deduciendo la evolución del medio natural o las repercusiones de la acción humana sobre el medio (Ruiz et al., 2013).

El estudio, inventario y conocimiento del uso de la tierra y la superficie cultivada cobra significativa importancia para la planificación agrícola y económica de cualquier territorio. Esto se debe a que de este tipo de estudios puede depender del cálculo de la producción agrícola (Nolasco et al., 2015).

Las diversas cubiertas vegetales están relacionadas con el clima, las condiciones físico-geográficas, condiciones de suelo y la disponibilidad del agua, lo que permite identificar las potencialidades y limitaciones de cada zona. Los componentes del clima, suelo y agua se caracterizan por poseer una variación espacio-temporal, condicionada por distintos factores fisiográficos, como la altitud, la latitud, el relieve y la posición geográfica (Alva & Beraun, 2013).

2.1 Contexto internacional

Marini et al. (2007), realizaron un estudio multitemporal en la zona de Guaminí, Buenos Aires, Argentina con imágenes satelitales Landsat 5 TM (resolución espacial de 30 m), seleccionaron siete fechas representativas de los distintos estados fenológicos de cada cultivo en el periodo 2004 - 2005. En cada una de ellas se obtuvo el índice de Vegetación de Diferencia Normalizado (Normalized Difference Vegetation Index - NDVI). De acuerdo con la evolución de dicho parámetro, se establecieron cinco categorías de uso de la tierra (además de los cuerpos de agua y las localidades) mediante la elaboración de un árbol de decisión. Los cultivos de invierno y de verano (girasol, maíz y pasturas) así como la soya, pudieron discriminarse óptimamente.

Araque & Jiménez (2009) dieron seguimiento a una problemática de sanidad en una plantación de palma de aceite a escala regional y temporal en la Zona Central palmera ubicadas en el área de Puerto Wilches, Colombia. El estudio lo hicieron con imágenes del sensor MODIS (resolución espacial de 250, 500 y 1000 m), para la caracterización de los patrones de las curvas espectrales, identificaron los valores de las bandas espectrales asociados con la obtención de índices de vegetación, analizando series temporales. Esto permitió identificar los eventos de defoliación e interpretar los resultados en diferentes niveles.

Para la identificación de cultivos de arroz en Colombia (Cundinamarca y Tolima), García & Martínez (2010) utilizaron imágenes Landsat 7 y una imagen ASTER. Para ellos los autores realizaron un pretratamiento de las imágenes mediante el algoritmo del vecino más cercano, obtuvieron el NDVI para caracterizar las firmas espectrales de los cultivos y máscaras de textura promedio (la textura promedio es el valor promedio de los píxeles que comprenden una región dentro de un arreglo de píxeles usado para limitar una operación a un grupo de píxeles), aplicaron un análisis de componentes principales que ayudó a complementar el NDVI, y se realizó una extracción por grupos para realizar la reclasificación de las coberturas. Encontraron algunas limitaciones en los resultados por la temporalidad de las imágenes y por ende de la fenología de los cultivos. Para validarlo, tomaron puntos GPS de los cultivos en campo. Los resultados arrojaron que el método permitió identificar los cultivos de arroz con una exactitud superior al 70%.

Peña-Barragán et al. (2011), identificaron los principales tipos de cultivos en el condado Yolo, en California. Primero discriminaron entre cereales de invierno, prados, cultivos de verano y cultivos permanentes, para la generación de políticas agrarias. Utilizaron árboles de decisión para clasificar el área de tierras de cultivo, y técnicas de análisis de imágenes basadas en objetos, que incorporan nuevas características espectrales, texturales y jerárquicas, obtenidas, principalmente de los índices de vegetación espectral, para la identificación y mapeo de cultivos. El estudio lo realizaron para dos y tres periodos de tiempo, obteniendo un rango de error de 12-16% para el primer caso y del 9% para el segundo. Los árboles de decisión crean modelos de clasificación basado en diagramas de flujo. Clasifican casos en grupos o pronostican valores de una variable dependiente (criterio) basada en valores de variables independientes (predictoras) (Berlanga et al., 2013).

Bohn et al. (2012) elaboraron una investigación en donde el objetivo principal fue analizar y cuantificar la variabilidad del estado y densidad de la vegetación mediante el uso de sensores remotos e índices de vegetación en algunos ambientes del suroeste de la provincia de Buenos Aires, Argentina. Para ese análisis, se trabajó con imágenes satelitales MODIS (resolución espacial de 250 m) y con ellas obtuvieron valores de NDVI de 12 sitios correspondientes a distintas coberturas de suelo, obteniendo valores medios de NDVI asociados a vegetaciones sanas y en crecimiento, y para el caso de los periodos de verano e invierno se observó un deterioro de la vegetación.

El propósito del estudio de Ruiz et al., (2013) fue evaluar el cambio de uso del suelo en el Paisaje Terrestre Miraflor Moropotente en Nicaragua, para el período 1993-2011, a través de imágenes satelitales, a fin de determinar el estado de fragmentación del paisaje. Los cambios de usos de suelo fueron derivados de la clasificación de tres imágenes Landsat 5 TM, con una resolución espacial de 30 m de febrero 1993, abril 2000 y enero 2011. Se realizó una verificación en campo para la identificación de coberturas de suelo y la corroboración con las imágenes satelitales. La fragmentación se realizó con el cálculo de métricas e índices de fragmentación a nivel del paisaje. Los principales resultados muestran que los cambios de uso de suelo están determinados por la degradación antrópica, principalmente en la conversión de la vegetación nativa a espacios agrícolas y la expansión de la ganadería.

Nolasco et al. (2015) desarrollaron su trabajo en la región central de la provincia de Córdoba, identificaron los principales cultivos de soya y maíz de la región con imágenes Landsat 5 TM. El estudio lo hicieron en tres fechas representativas (en el 2010) de los distintos estados fenológicos de los cultivos. Realizaron clasificación supervisada con el algoritmo de Máxima Verosimilitud, levantamiento en campo de los diferentes usos de suelo. Evaluaron su precisión con una matriz de confusión que compara el mapa obtenido de la clasificación de las imágenes satelitales con la base de datos original recogida durante el levantamiento en campo. Obtuvieron que los resultados fueron mejores en las etapas más avanzadas.

Para la identificación de cultivos de frijol en la provincia de Sumapaz, Colombia, Ávila, et al. (2020) realizaron una toma de puntos en campo para realizar zonas de muestreo y con ello poder realizar clasificación supervisada para dos etapas de crecimiento distintas dentro de una misma imagen compuesta y calibrada en Sentinel 2A (resolución espacial de 10, 20 y 60 m). Extrajeron los valores de reflectancia para cada una de las 13 bandas y realizaron una reclasificación de cada banda para obtener las áreas de cultivo. Posteriormente se sumaron los resultantes para poder obtener las zonas de cultivo de frijol y se contrastó la información obtenida para las dos etapas fenológicas de los cultivos (etapa de vainas y etapa de madurez). Con ello se obtuvo la cantidad de hojas en crecimiento y también evaluaron el estrés hídrico de la planta.

Le et al. (2020), proponen un método que utiliza una combinación de características filtradas extraídas por operadores combinados de patrones binarios locales y características extraídas por máscaras de contorno de planta-hoja para mejorar la tasa de discriminación entre plantas de hoja ancha (canola – cultivo) y rábano silvestre (maleza). Este método alcanzó una precisión de clasificación del 98.63% con 4 clases, esta metodología la propusieron porque las malezas son una de las principales causas de la baja productividad agrícola y algunas malezas tienen características morfológicas similares a las de los cultivos, lo que dificulta su discriminación.

La FAO (1997) ha dedicado una atención considerable al desarrollo de procedimientos para inventariar, evaluar y planificar los recursos de tierras, tanto a nivel global, como en regiones y países concretos a través de sus programas de campo. Creó un sistema de clasificación con una serie de clases, generando dos fases principales, en el que define áreas con vegetación y zonas sin vegetación. Posterior a eso generó una fase modular-jerárquica para definir las clases de coberturas más finas (FAO, 2000). Este tipo de clasificación se puede tomar como base para la identificación y definición de las cubiertas vegetales de este estudio.

2.2 Contexto nacional

Sánchez et al. (2006) describen un algoritmo de clasificación genérico, basado en la forma de la firma espectral de 12 objetos presentes en las imágenes del sensor ETM+ del satélite Landsat 7. El algoritmo es un árbol de decisión que no requiere de campos de entrenamiento para llevar a cabo la clasificación, ya que se fundamenta en operaciones Booleanas que identifican los objetos por formas características de su firma espectral. El algoritmo ha sido utilizado con éxito en la estimación de superficies sembradas de cultivos e identificación de cuerpos de agua en imágenes que cubren casi la totalidad de la República Mexicana.

Para el caso de la estimación de biomasa (B_m), del área asociada al crecimiento de cinco cultivos agrícolas, previamente medidos en campo en la zona del Valle del Yaqui en Sonora, Pascual-Ramírez et al. (2012) utilizaron sensores remotos para la obtención del índice de vegetación iso-suelo (IVIS). Este índice resultó óptimo para este caso, ya que al comparar la cantidad generada de biomasa de acuerdo con la etapa fenológica de los cultivos y su rendimiento tienden a ser similares.

Castañeda-Ibáñez et al. (2015) con el uso de imágenes de satélite estimaron índices de vegetación (IV), el de vegetación de diferencias normalizadas (NDVI) y el de vegetación ajustado para suelo (SAVI), para monitorear el desarrollo de los cultivos y estimar con precisión la variabilidad espacial y temporal del coeficiente de cultivo (K_c). El objetivo del estudio fue validar el uso de IV para estimar K_c y ET en el Distrito de Riego del Río Yaqui, Sonora, México, con ocho imágenes Landsat 5 TM y 7 ETM+ de los satélites en el cultivo de trigo, para el año 2008. Los valores estimados mostraron grado alto de correspondencia (0.69 mm d^{-1}) y un índice de 0.92 para el SAVI en estimaciones diarias, por ello concluyeron que los IV permiten estimar con precisión la variabilidad espacial y temporal de los K_c y la ET en grandes regiones agrícolas.

De León et al. (2019), elaboraron una metodología para la estimación del estrés hídrico que presentan los cultivos de los acuíferos de la región noroeste del Estado de Chihuahua. Este estudio abarcó de 1985-2018, así mismo se hizo una investigación sobre los tipos de cultivos característicos de la zona y las necesidades agro-hidrológicas de cada uno, con esto se determinó la cantidad del recurso hídrico necesario por campo.

Con el procesamiento digital de imágenes aéreas tomadas en un cultivo de cebolla en Zacatecas se pretendió encontrar las problemáticas que afectan el crecimiento de la planta, por ejemplo, la maleza, la densidad de población vegetal y el exceso o la falta de humedad de los cultivos. Esto lo realizaron mediante la segmentación de imágenes y una clasificación manual y una automatizada con un algoritmo que se desarrolló para la extracción de las anomalías antes mencionadas y su índice de vegetación, se denominó algoritmo grabCut (Espinoza García et al., 2019).

CAPITULO 3. MARCO TEÓRICO

3.1 Criterios de delimitación para la determinación de la vegetación

Los estudios fenológicos con teledetección utilizan datos recopilados por sensores satelitales que miden longitudes de onda de luz absorbidas y reflejadas por las plantas (USGS, 2020). La vegetación interactúa con la radiación solar de manera diferente a otros materiales naturales, como suelos y cuerpos de agua. La absorción y reflexión de la radiación solar es el resultado de muchas interacciones con diferentes materiales vegetales, que varía considerablemente según la longitud de onda. El agua, los pigmentos, los nutrientes y el carbono se expresan en el espectro óptico reflejado de 400 nm a 2500 nm, con comportamientos de reflectancia a menudo superpuestos, pero espectralmente distintos (Harris Geospatial Solutions, 2020)

Ciertos pigmentos en las hojas de la planta absorben fuertemente longitudes de onda de luz visible (roja). Las hojas en sí mismas reflejan fuertemente longitudes de onda de luz infrarroja cercana, que es invisible para los ojos humanos. A medida que el dosel de la planta cambia del crecimiento de principios de la primavera a la madurez y senescencia a finales de temporada, estas propiedades de reflectancia también cambian (USGS, 2020).

Según Zhang et al. (2003) se determinan cuatro fechas de transición entre estados fenológicos distintos:

- 1) Aumento en el verdor asociado a una alta actividad fotosintética
- 2) Madurez, asociado al máximo índice de área foliar
- 3) Senescencia, fecha en la cual la actividad fenológica disminuye drásticamente
- 4) Latencia, cuando las actividades fisiológicas son mínimas, la radiación solar es baja, la vegetación puede estar cubierta por nieve y el acceso al agua en estado líquido es muy limitado

La FAO (2000), considera ciertos criterios de diseño para clasificar la cubierta vegetal denominado Sistema de Clasificación de la Cubierta Terrestre (LCCS), dividido en 2 fases principales: Áreas con vegetación y sin vegetación, y la “Fase Modular-Jerárquica (criterios)”, en la que las clases de cobertura del suelo se crean mediante la combinación de conjuntos de clasificadores predefinidos:

- La forma de vida se refiere a un grupo de plantas que tienen ciertas características morfológicas en común.
- Distribución espacial, se define como la distribución espacial horizontal de la vegetación en un área determinada. Describe la disposición espacial de tipos específicos de vegetación estructural (por ejemplo, bosque cerrado, arbustos cerrados).
- Los cultivos combinados: hace referencia a si hay más de un cultivo en la zona a estudiar, los cultivos presentes pueden especificarse junto con detalles de la posible superposición en el período de crecimiento entre los cultivos principal y secundario o el tipo de entorno en el que crecen los cultivos¹
- Las prácticas culturales: puede hacer referencia a la forma en que suministran agua a los cultivos, es decir, la forma de riego y también se puede referir a las prácticas que realizan para barbechar (después de la cosecha).
- Litología/ Suelo: La litología se puede describir en función del material geológico y su edad. Se puede describir primero el aspecto superficial del suelo, seguido de una descripción detallada del perfil del suelo.
- Clima: la Duración del Período de Crecimiento (LGP-por sus siglas en inglés) reconoce las limitaciones climáticas relevantes en cualquier región importante del mundo. La combinación de clases térmicas y clases de humedad da el clima.
- Altitud: este atributo se puede utilizar en todos los principales tipos de cobertura terrestre.
- Erosión: se hace hincapié en la erosión acelerada o inducida por el hombre. Esta última es, a menudo, el resultado de un uso irracional y una mala gestión, como prácticas agrícolas incorrectas, sobrepastoreo o sobreexplotación de la vegetación (semi)natural. Estas prácticas dan como resultado un tipo de cubierta con características específicas.
- Cobertura: esta puede considerarse como la presencia de un área particular del suelo, sustrato o superficie de agua cubierta por una capa de plantas consideradas en el mayor nivel perimetral horizontal de cada planta en la capa. Se hace una distinción entre cerrado (> 60 – 70 %), abierto (entre 60 - 70 y 10 - 20 %) y escaso (por debajo 10 - 20 %, pero >1 %).
- La altura de una determinada capa se mide desde el suelo hasta la parte superior promedio de la forma de vida que se está examinando. Se clasifica como: Árboles >30 - 3 m; Arbustos 5 - 0.3 m; y Herbáceas 3 m - 0.03 m.

Otro elemento que ayuda en la identificación de los cultivos es el uso de los índices de Vegetación, donde algunos autores proponen obtener variables de textura para discriminar mejor entre los diferentes tipos de cultivos (García & Martínez, 2010; Peña-Barragán et al., 2011). Por ejemplo, para el estudio de Peña-Barragán et al. (2011) las variables de textura fueron necesarias para discriminar entre la mayoría de los campos de cultivo permanentes (huertos, viñedos, alfalfa y pradera).

3.2 Índices de Vegetación

De acuerdo con Ayala-Carcedo & Olcina Cantos (2002) a partir de una imagen de satélite se pueden obtener parámetros que indican el estado de la vegetación, la cantidad de biomasa existente, así como otros parámetros directamente relacionados como el grado de cobertura vegetal, el índice de área foliar (LAI por sus siglas en inglés), la evapotranspiración potencial (ETP) o la radiación absorbida fotosintéticamente por la planta (APAR). El contenido de humedad de la vegetación es un parámetro difícil de estimar, sin embargo, el estrés hídrico se puede obtener relativamente fácil a través del análisis de series multitemporales del índice de Vegetación de Diferencias Normalizadas (NDVI). El incremento del NDVI está asociado al LAI y el APAR, por lo que es un buen indicador del estado óptimo de la vegetación. Por el contrario, los acumulativos decrecimientos de los valores NDVI en cada píxel a lo largo de la temporada estival indican que la vegetación ha sufrido un incendio, la USGS (2020) menciona que los cambios de la vegetación también podrían ser provocados por la deforestación o cambios en la etapa fenológica de las plantas.

Un índice de vegetación es un indicador que describe el verdor —la densidad relativa y la salud de la vegetación— para cada elemento de imagen, o píxel, en una imagen de satélite (USGS, 2020) son adimensionales, siendo los más utilizados el NDVI, el SAVI y el ARVI, entre otros. El índice NDVI (ratio normalizada de reflectancia terrestre en bandas del rojo e infrarrojo cercano) puede utilizarse para descubrir la cobertura vegetal, ya que está íntimamente relacionado con el LAI (índice de área foliar), el APAR (índice de vegetación resistente al ambiente) y el grado de cobertura vegetal de la vegetación. Se obtiene mediante la ecuación 1 (Ayala-Carcedo & Olcina Cantos, 2002):

$$NDVI = (NIR - Red)/(NIR + Red) \quad (1)$$

Siendo:

Red = Reflectancia en el rojo

NIR = Reflectancia en el infrarrojo cercano

Los valores NDVI oscilan entre -1.0 y +1.0. Las áreas de roca, arena o nieve estériles suelen mostrar valores NDVI muy bajos (por ejemplo, 0.1 o menos). La vegetación escasa, como arbustos y pastizales o cultivos senosante, puede dar lugar a valores moderados de NDVI (aproximadamente de 0.2 a 0.5). Los valores altos de NDVI (aproximadamente 0.6 a 0.9) corresponden a vegetación densa como la que se encuentra en los bosques templados y tropicales o cultivos en su etapa de crecimiento máximo.

Los valores de NDVI se pueden promediar a lo largo del tiempo para establecer condiciones de crecimiento "normales" en una región durante una época determinada del año.

El índice SAVI (ecuación 2) se obtiene a partir del NDVI para corregir la contaminación provocada por la reflectancia de la superficie asociada a variaciones en la luminosidad del terreno. La efectividad del índice ha sido comprobada en diversos estudios de pastizales en los que el terreno, la atmósfera y medidas tomadas a partir de imágenes satélite han sido tenidas en cuenta. La expresión del índice es la siguiente:

$$SAVI = (NIR - Red)/(NIR + Red + L) \quad (2)$$

Siendo: L = 0.5 (constante)

El índice ARVI (ecuación 3), índice de vegetación de corrección atmosférica, minimiza las variaciones de la atmosférica sobre el índice de vegetación, siendo la expresión:

$$ARVI = (NIR - Blue)/(NIR + Blue) \quad (3)$$

Siendo: Blue = reflectancia en el azul

Una mezcla de ambos índices, el SAVI y el ARVI es el SARVI (ecuación 4), que minimiza tanto los errores producidos por la atmósfera, así como los del terreno, resultando un índice más estable y completo. Su expresión es:

$$SARVI = (NIR - Blue)/(NIR + Blue + L) \quad (4)$$

Uno de los inconvenientes de las imágenes satélite son los errores recogidos por los sensores a causa de la existencia de nubes, la dirección de reflectancia terrestre o las condiciones atmosféricas adversas, pudiendo afectar el valor del NDVI. Existen métodos que permiten corregir estos errores, como son el método MVC., que interpola el máximo valor de NDVI en un periodo (semanal o cada dos semanas) y lo asigna a la celda correspondiente (Ayala-Carcedo & Olcina Cantos, 2002).

El índice NDVI es muy útil para señalar áreas cubiertas por vegetación en estado óptimo (vegetación verde), pero no permite distinguir con tanta facilidad aquellas zonas donde la vegetación sufre sequía. Por ello es importante combinar los índices de vegetación con mapas reales de vegetación para poder contrastar y comprobar el estado de la cubierta vegetal en todo momento (Ayala Carcedo & Olcina Cantos 2002).

3.3 Análisis de Componentes Principales (ACP)

El Análisis de Componentes Principales pertenece al grupo de técnicas estadísticas multivariantes (Martín et al., 1994), y tiene por objeto transformar un conjunto de variables, a las que se denomina originales, en un nuevo conjunto de variables denominadas componentes principales (De la Fuente, 2011).

En el ACP existe la opción de usar la matriz de correlaciones o bien, la matriz de covarianzas. Las nuevas variables (componentes principales) son obtenidas como combinaciones lineales de las variables originales. Los componentes se ordenan en función del porcentaje de varianza explicada. En este sentido, el primer componente será el más importante por ser el que explica mayor porcentaje de la covarianza de los datos. Queda a criterio del investigador decidir cuántos componentes se elegirán en el estudio (De la Fuente, 2011; Martín et al., 1994).

3.3.1 Aplicaciones del análisis de componentes principales

Según Martín et al. (1994), entre los usos más frecuentes están:

1. Como técnica de análisis exploratorio que permite descubrir interrelaciones entre los datos y de acuerdo con los resultados, proponer los análisis estadísticos más apropiados.
2. Reducir la dimensionalidad de la matriz de datos con el fin de evitar redundancias y destacar relaciones. En la mayoría de los casos, tomando sólo los primeros

componentes, se puede explicar la mayor parte de la variación total contenida en los datos originales.

3. Construir variables no observables (componentes) a partir de variables observables.
4. Bajo ciertas circunstancias, es de gran utilidad usar estos componentes incorrelacionados, como datos de entrada para otros análisis. Por ejemplo, en el caso de la regresión múltiple cuando las variables independientes presentan alta colinealidad es preferible hacer la regresión sobre los componentes principales en lugar de usar las variables originales.

El método de PCA permite por lo tanto “condensar” la información aportada por múltiples variables en solo unas pocas componentes, sin olvidar que se necesitan conocer los valores de las variables originales para calcular las componentes.

3.3.2 Cálculo de las componentes principales

De acuerdo con Rodrigo (2017) cada componente principal (Z_i) se obtiene por combinación lineal de las variables originales. Se pueden entender como nuevas variables obtenidas al combinar de una determinada forma las variables originales. La primera componente principal de un grupo de variables (X_1, X_2, \dots, X_p) es la combinación lineal normalizada de dichas variables que tiene mayor varianza:

$$Z_{1i} = u_{11}X_{1i} + u_{12}X_{2i} + \dots + u_{1p}X_{pi}$$

Que la combinación lineal sea normalizada implica que:

$$\sum_{j=1}^p u_{j1}^2 = 1$$

Los términos u_{11}, \dots, u_{1p} reciben el nombre de *loadings* y son los que definen a la componente. U_{11} es el *loading* de la variable x_1 de la primera componente principal. Los *loadings* pueden interpretarse como el peso/importancia que tiene cada variable en cada componente y, por lo tanto, ayudan a conocer qué tipo de información recoge cada una de las componentes, estimadas de la siguiente manera:

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

- Se resuelve un problema de optimización para encontrar el valor de los *loadings* con los que se maximiza la varianza. Una forma de resolver esta optimización es mediante el cálculo de eigenvector-eigenvalue de la matriz de covarianzas.

Por un lado, los eigenvectors son un caso particular de multiplicación entre una matriz y un vector. El vector resultante de la multiplicación es un múltiplo entero del vector original. Los eigenvectors de una matriz son todos aquellos vectores que, al multiplicarlos por dicha matriz, resultan en el mismo vector o en un múltiplo entero del mismo. Los eigenvectors tienen una serie de propiedades matemáticas específicas:

- Los eigenvectors solo existen para matrices cuadradas y no para todas. En el caso de que una matriz $n \times n$ tengan eigenvectors, el número de ellos es n .
- Si se escala un eigenvector antes de multiplicarlo por la matriz, se obtiene un múltiplo de este eigenvector. Esto se debe a que, si se escala un vector multiplicándolo por cierta cantidad, lo único que se consigue es cambiar su longitud, pero la dirección es la misma.
- Todos los eigenvectors de una matriz son perpendiculares (ortogonales) entre ellos, independientemente de las dimensiones que tengan.

Dada la propiedad de que multiplicar un eigenvector solo cambia su longitud, pero no su naturaleza de eigenvector, es frecuente escalarlos de tal forma que su longitud sea 1. De este modo se consigue que todos ellos estén estandarizados. Por otro lado, cuando se multiplica una matriz por alguno de sus eigenvectors se obtiene un múltiplo del vector original, es decir, el resultado es ese mismo vector multiplicado por un número. Al valor por el que se multiplica el eigenvector resultante se le conoce como eigenvalue. A todo eigenvector le corresponde un eigenvalue y viceversa.

3.3.3 Interpretación geométrica del ACP

Una forma intuitiva de entender el proceso de ACP consiste en interpretar las componentes principales desde un punto de vista geométrico. Supóngase un conjunto de observaciones para las que se dispone de dos variables (X_1, X_2). El vector que define la primera componente principal (Z_1) sigue la dirección en la que las observaciones varían en mayor medida (línea roja de la figura 1). La proyección de cada observación sobre esa dirección equivale al valor de la primera componente para dicha observación.

La segunda componente (Z_2 Figura 1) sigue la segunda dirección en la que los datos muestran mayor varianza y que no está correlacionada con la primera componente. La condición de no correlación entre componentes principales equivale a decir que sus direcciones son perpendiculares/ortogonales.

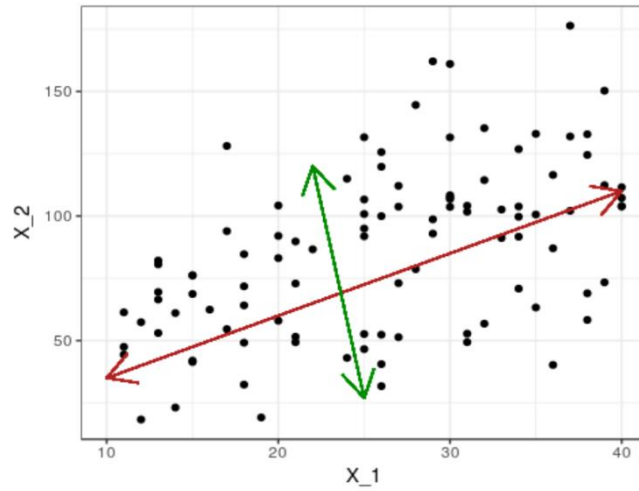


Figura 1. Representación geométrica de componentes principales

CAPÍTULO 4. METODOLOGÍA

La presente metodología muestra una propuesta para identificar y clasificar cubiertas vegetales bajo el criterio más adecuado por medio de análisis de componentes principales (ACP; Figura 2) Este procedimiento fue aplicado en parte Central de la Altiplanicie Mexicana y consta de tres fases generales: determinación de estructuras jerárquicas de vegetación, desarrollo del modelo y la validación de resultados por medio del método de máxima verosimilitud.

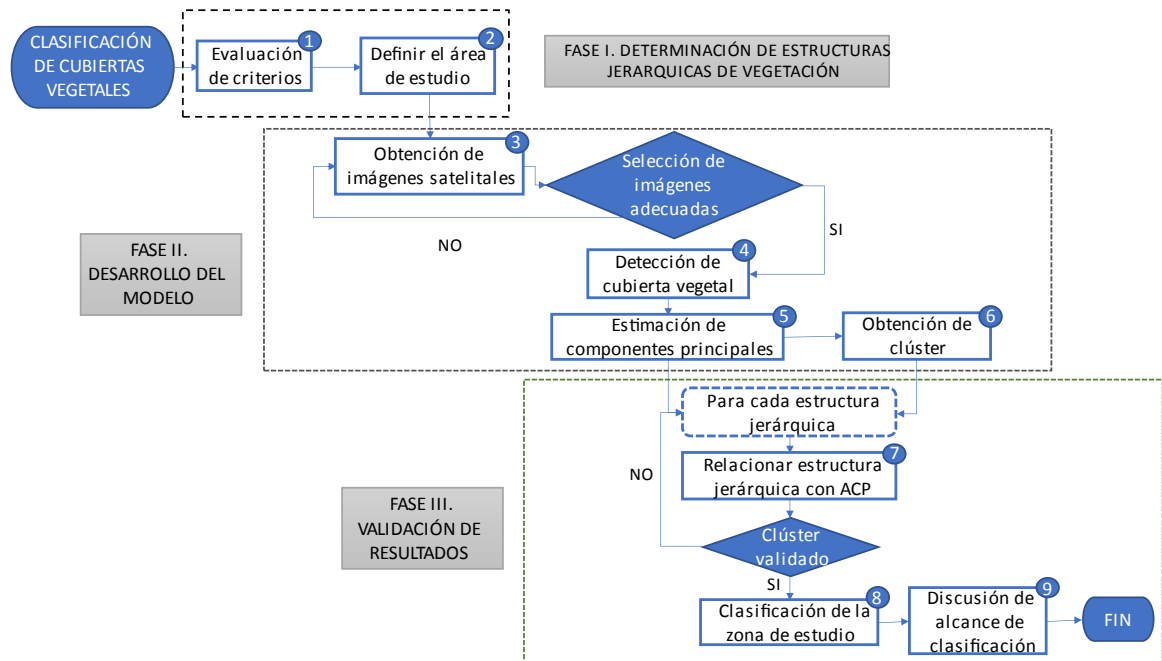


Figura 2. Diagrama de la metodología

1.1 Fase I. Determinación de estructuras jerárquicas de los cultivos

En la primera etapa y de acuerdo con los diversos criterios de clasificación de cubiertas vegetales mencionadas en los antecedentes, para la presente investigación se sugiere tomar la estructura base mostrada en la Figura 3. En ella, se muestran 3 niveles básicos: zonas con vegetación y sin vegetación; las áreas con vegetación se dividen en acuática y terrestre y; cada una de ellas especificadas posteriormente en sitios naturales y regulados (cultivos). A partir de esta estructura base se plantearán subniveles de clasificación basados en los criterios identificados por la FAO (2000):

1. Presencia de vegetación:

- Condición edáfica:

1.1 Áreas terrestres

- Áreas de cultivo → Tipo de cultivo (Clase)

- Forma de vida
- Aspectos espaciales
- Cultivos combinados
- Prácticas culturales
- Forma de la tierra
- Litología/ Suelo
- Clima
- Altitud
- Erosión
- Cobertura

CRITERIOS

- Vegetación natural (Clase)

- Forma de vida y cobertura
- Altura
- Patrón
- Tipo de hoja
- Fenología de la hoja
- Estratificación
- Forma de la tierra
- Litología/ Suelo
- Clima
- Altitud
- Erosión
- Aspectos florísticos

CRITERIOS

- Áreas no identificadas (Clase)

1.2 Áreas acuáticas o inundadas

2. Zonas sin vegetación:

- Condición edáfica: Áreas terrestres o acuáticas

Clasificación base

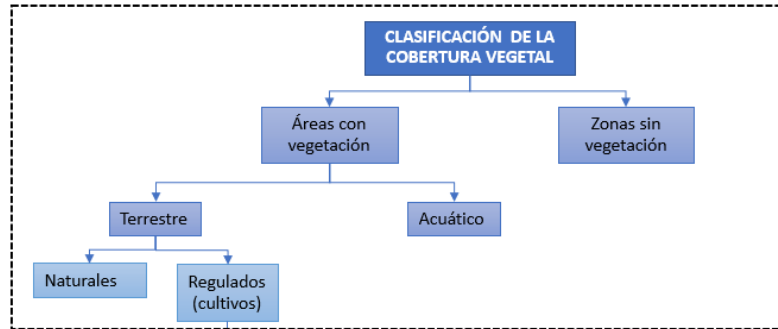


Figura 3. Clasificación del tipo de cultivos

El modelo propuesto en esta investigación se ha concebido como un modelo integrado, es decir, la clasificación se puede expandir en función de la superficie de la o las zonas de estudio que se pueden considerar. Por lo tanto, para efectos del presente proyecto y como parte de la etapa 2, el área de estudio se delimitó con base en la disponibilidad de información, tanto de producción agrícola como de imágenes satelitales en la Zona Central de la Altiplanicie Mexicana.

En esta zona se pueden apreciar varios Distritos de Riego (DR), entre estos se encuentran: 003, 023, 033, 096, 044, pertenecientes a los estados de Querétaro, Hidalgo, Michoacán y Estado de México (Figura 4; SINA, 2021). Para estos distritos de riego se cuentan con estadísticas y los principales cultivos que se practican en la zona a partir de la base de datos del Servicio de Información Agroalimentaria y Pesquera (SIAP, 2022). Es importante destacar que dicha información es un referente de los usos de suelo y prácticas agrícolas que se llevan a cabo en la región, por lo que es referencia de las clases o categorías a identificar en el área de estudio. A continuación, se presentan los cultivos de los principales cultivos presentes en la zona de estudio.

ÁREA DE ESTUDIO

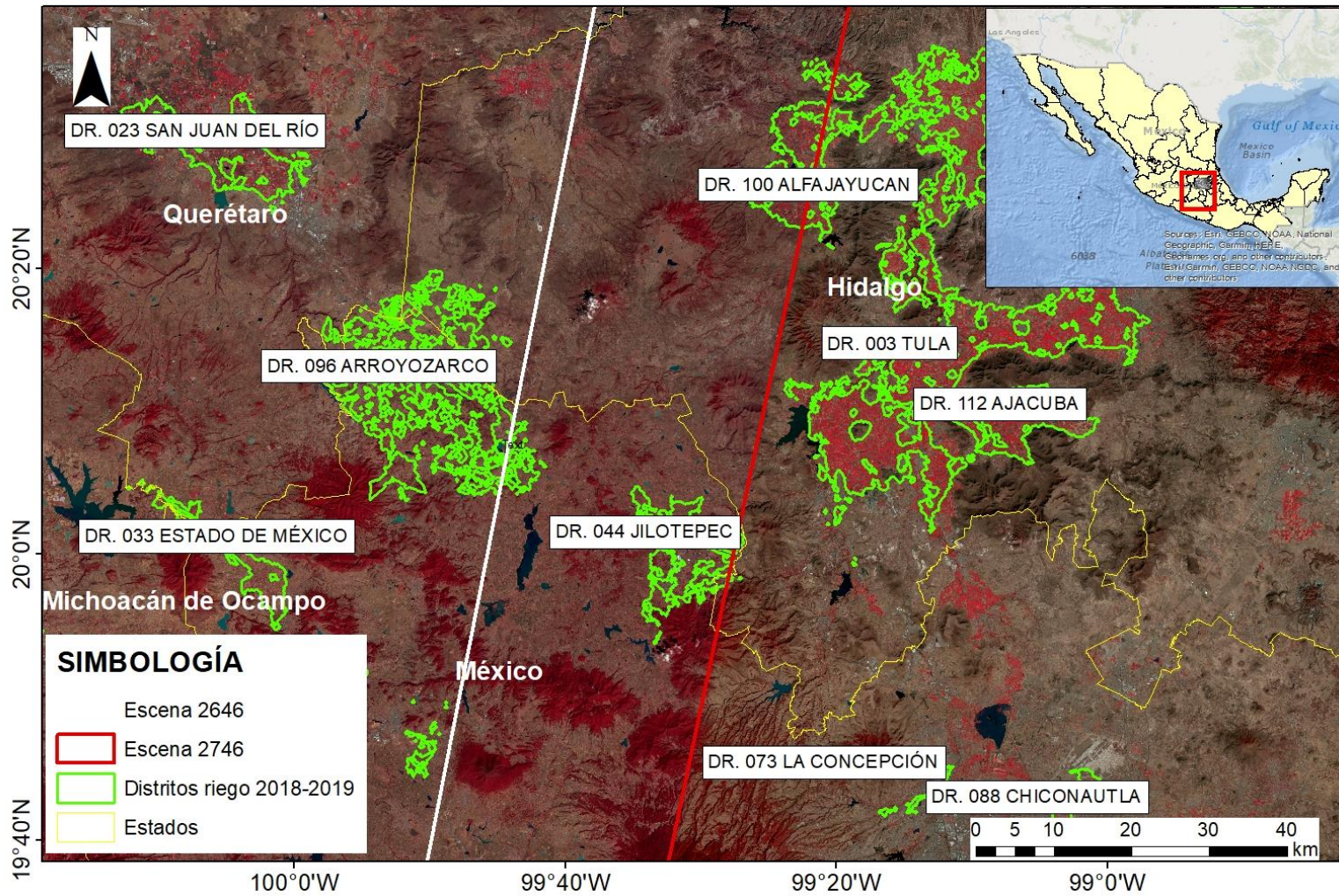


Figura 4. Mapa de ubicación del área de estudio. Compuesto falso color (345) a partir de la imagen Landsat 8 (2020-11-22 y 2021-01-02)

El DR 003 Tula (tabla 1), se ubica en el Estado de Hidalgo, pertenece a la Región Hidrológica Pánuco y a la Región Hidrológico-administrativo Valle de México, en su mayoría es abastecido por agua superficial (tabla 1; SINA, 2021).

Tabla 1. Cultivos del DR 003 Tula Hidalgo (CONAGUA, 2015, 2017)

DISTRITO DE RIEGO	CICLO (2013-2016)	MODALIDAD (2013-2016)	CULTIVOS (2013-2016)	SUPERFICIE (Ha -2013-2016)
003 TULA, HIDALGO	Otoño - Invierno	RIEGO	Avena Forrajera verde	54, 812 Ha. (2013-2014) 52,597 Ha. (2015-2016)
			Brócoli	
			Cebada Forrajera Verde	
			Coliflor	
			Haba	
			Nabo	
			Otros Pastos (verde)	
			Trigo Grano	
			Triticale	
	Primavera - Verano		Avena Forrajera verde	
			Calabacita (Calabacín)	
			Cebada Forrajera Verde	
			Chile Verde	
			Coliflor	
			Frijol (Alubia)	
			Haba	
			Maíz Grano	
			Nabo	
			Otros Cultivos	
	Perennes		Tomate de la Cáscara (Tomatillo)	
	Segundos Cultivos		Alfalfa	
			Bermuda (Zacate) Verde	
			Avena Forrajera verde	
			Brócoli	
			Calabacita (Calabacín)	
			Cebada Forrajera Verde	
			Chile Verde	
			Coliflor	
			Frijol (Alubia)	
			Maíz Grano	
Nabo				
Otras Flores				
Otros Cultivos				
Tomate de la Cáscara (Tomatillo)				

El DR 023 San Juan del Río se encuentra ubicado entre los municipios de San Juan del Río y Pedro Escobedo en el estado de Querétaro, México. El agua para el riego se obtiene de las presas San Ildefonso, Constitución de 1917 y La Llave, así como del bordo La Venta, y de 54 pozos profundos (Tabla 2; Morales-Durán et al., 2018).

Tabla 2. Cultivos del DR 023 San Juan del Río (CONAGUA, 2015, 2017)

DISTRITO DE RIEGO	CICLO (2013-2014)	MODALIDAD (2013-2014)	CULTIVOS (2013-2014)	SUPERFICIE (Ha -2013-2014)	CULTIVOS (2015-2016)	SUPERFICIE (Ha -2015-2016)
023 SAN JUAN DEL RÍO	Otoño Invierno	RIEGO	Avena Forrajera Verde	8025 Ha.	Avena Forrajera Verde	8983 Ha.
			Cebada		Cebada	
			Chile Verde		Hortalizas Invernadero	
			Otros Cultivos		Otras Hortalizas	
			Zanahoria		Zanahoria	
	Primavera - Verano		Chile Verde		Chile seco	
			Frijol (Alubia)		Frijol Asociado	
			Maíz Grano		Maíz Grano	
			Otros Cultivos		Otras Hortalizas	
			Sorgo Grano		Sorgo Grano	
	Perennes		Alfalfa		Alfalfa	
			Frutales Asociados		Frutales Asociados	
			Otros Frutales		Otros cultivos	
			Rosal		Rosal	
			Vid Mesa		Vid Mesa	
	Segundos Cultivos		Maíz Grano		Maíz Grano	
			Sorgo Grano		Sorgo Grano	
			Zanahoria			

El DR 033, denominado “Estado de México”, está ubicado en la región noreste del Estado de México, se localiza en las coordenadas 19° 30´, latitud norte y 99° 70´ longitud oeste. Comprende cuatro municipios del Estado de México (Acambay, Atlacomulco, Jocotitlán, Temascalcingo), uno de Querétaro (Amealco), y uno de Michoacán (Contepec). Pertenece a la región hidrológico- administrativa VIII (Lerma-Santiago-Pacífico). En la

Tabla 3, se muestran los cultivos que se practican en ese DR (Enríquez et al., 2016).

Tabla 3. Principales cultivos del DR 033 Estado de México (CONAGUA, 2015, 2017)

DISTRITO DE RIEGO	CICLO (2013-2016)	MODALIDAD (2013-2016)	CULTIVOS (2013-2016)	SUPERFICIE (Ha -2013-2016)
033 ESTADO DE MÉXICO	Otoño - Invierno	RIEGO	Avena Forrajera verde	6,789 Ha. (2013-2014) 6,288 Ha. (2015-2016)
			Rosal	
	Trigo Grano			
	Jitomate Invernadero			
	Maíz Grano			
Primavera - Verano		Rye Grass (Zacate) verde		
Perennes				

El Distrito de Riego 044 Jilotepec se extiende principalmente en el Norte del Estado de México, y parcialmente en el Estado de Hidalgo. Al igual que los otros Distritos, este también es abastecido por agua superficial (Tabla 4)

Tabla 4. Principales cultivos del DR 044 Jilotepec (CONAGUA, 2015, 2017)

DISTRITO DE RIEGO	CICLO (2013-2016)	MODALIDAD (2013-2016)	CULTIVOS (2013-2016)	SUPERFICIE (Ha -2013-2016)
044 JILOTEPEC	Otoño - Invierno	RIEGO	Avena Forrajera verde	2,015 Ha. (2013-2014) 2,414 Ha. (2015-2016)
	Primavera - Verano		Jitomate Invernadero	
			Maíz Grano	
	Perennes		Rye Grass (Zacate) verde	
			Frutales asociados	

Arroyozarco es el DR 096, se encuentra en su mayor parte en el Estado de México, y pequeña porción de este en los Estados de Hidalgo y Querétaro, pertenece a la Región Hidrológica Pánuco y a la Región Hidrológica Administrativa Valle de México (Tabla 5; SINA, 2021).

Tabla 5. Principales cultivos del DR 096 Arroyozarco (CONAGUA, 2015, 2017)

DISTRITO DE RIEGO	CICLO (2013-2016)	MODALIDAD (2013-2016)	CULTIVOS (2013-2016)	SUPERFICIE (Ha. -2013-2016)
096 ARROYO ZARCO	Otoño - Invierno	RIEGO	Avena Forrajera verde	5,383 Ha. (2013-2014) 4,967 Ha. (2015-2016)
	Primavera - Verano		Jitomate Invernadero	
	Perennes		Maíz Grano Rye Grass (Zacate) verde	

5.1 Fase II. Desarrollo del modelo

El desarrollo del modelo se basó en el análisis de imágenes de satélite, por lo que la primera fase contempla la obtención de imágenes satelitales (etapa 3). Para ello, se realizó la búsqueda y descarga de imágenes de calidad, es decir, sin más del 20% de nubosidad dentro de las escenas para no afectar el resultado de las clasificaciones (Tabla 6). Las imágenes satelitales que se procesaron fueron del satélite Landsat 8 OLI. Obtenidas del portal <https://earthexplorer.usgs.gov/> del Servicio Geológico de los Estados Unidos (USGS, por sus siglas en inglés). Con las imágenes se construyó una base de datos del periodo 2013 al 2016, esto, considerando la disponibilidad de información de cultivos, registrados en las Estadísticas Agrícolas de los Distritos de Riego (CONAGUA, 2013; SIAP, 2021).

Tabla 6. Características generales del sensor a utilizar. Basado en: Earth Explorer, 2021

Sensor	Fecha	Resolución	Path y Row
Landsat 8 OLI/TIRS	2013 – 2016	30 m (en la mayoría de sus bandas)	027 y 046

En la etapa de selección y descarga de imágenes satelitales se consideró que la escena no presentará más del 15 o 20% de nubosidad, principalmente en el área de interés, esto con el objetivo de evitar incluir píxeles de baja calidad en el proceso de construcción del modelo. Posteriormente, estas imágenes fueron tratadas en el Software TerrSet (Ronald, 2020), para ello se procesaron con el módulo de *Landsat data archive*. El módulo Landsat permite convertir los números digitales (datos originales de las imágenes) a datos de reflectancia y, realiza una corrección de la reflectancia por la sustracción del objeto más oscuro.

A partir de los datos de reflectancia se pueden estimar tres índices de vegetación por escena (tabla 7): NDWI (Gao, 1995), SAVI (Huete, 1988) y EVI (Huete, 2002). El fin de este procedimiento es resaltar algunas propiedades de la vegetación, por ejemplo, la condición de la cubierta vegetal, el contenido de humedad, identificación de cuerpos de agua, entre otros (etapa 4). Con ellos, se realizó una primera clasificación general, por medio del método de máxima verosimilitud y considerando los sitios de entrenamiento detectados en la zona de estudio.

Es importante mencionar que para los niveles de clasificación se seleccionaron 100 sitios de entrenamiento para cada imagen (36 imágenes). Esto se realizó para tres criterios; a) cobertura (vegetación y no vegetación); b) ambiente (vegetación acuática y terrestre) y; tipo de práctica (natural y regulada). A todas estas clases, se le agrega por default una clase “Sin identificar” para aquellas celdas que pertenezcan al fondo o “background” de la escena, o bien, no correspondan a ninguna agrupación por verosimilitud. Para los criterios de ambiente y tipo de práctica se debe verificar, además, que los sitios de entrenamiento correspondan a la clasificación de áreas con vegetación del primer nivel.

A partir de las escenas segmentadas, se propuso extraer una muestra estratificada distribuida tanto en ambas escenas como en todo el periodo de análisis (15 imágenes) con el módulo Multiclass data organizer (Fonseca et al, 2021) para su procesamiento en el paquete informático SPSS Statistics (IBM corporation, 2021). Con los datos obtenidos de índices de vegetación, características vegetales y climáticos se elaboró una base de datos por cada tipo de vegetación para realizar un análisis de Componentes Principales (ACP), (De la Fuente, 2011; González et al., 1994), como parte de la etapa 5.

Las variables consideradas en el ACP fueron; a) para índices espectrales (NDWI, SAVI, y EVI), (Tabla 7,– Gae, 1995; Huete, 1988; Huete, 2002), b) para características vegetales clorofila, índice LAI y carotenoides (tabla 7), y; c) para variables climatológicas, temperatura media mensual, humedad relativa media mensual y precipitación media mensual (Alva & Beraun, 2013; Daymet, 2021; Thornton et al., 2020).

Para el caso de las variables climatológicas de temperatura promedio y precipitación se integraron bases de datos obtenidas de Daymet (Daymet, 2021) y, para el caso de la humedad relativa se estimó a partir de la ecuación 5 en el programa terrset(FAO, 2021).

$$HR = \frac{e_a}{e^o(T)} * 100 \quad (5)$$

Siendo:

- HR = Humedad relativa
- e^a = Presión real de vapor derivada de la temperatura del punto de rocío (temperatura mínima diaria)
- e_o = Presión media de vapor de la saturación.

Tabla 7. Fórmulas de índices de vegetación espectral (IV) (Basado en: Gao, 1995; Huete, 1988; Huete, 2002; Daughtry et al., 2000)

Índice de Vegetación	Formula	Significado de las variables
Índice diferencial de agua normalizado (Gao, 1995)	$NDWI = \frac{NIR - SWIR1}{NIR + SWIR1}$	NIR = Reflectancia en el infrarrojo cercano SWIR1 = Infrarrojo de onda corta
Índice de vegetación ajustada al suelo (Huete, 1988)	$SAVI = (NIR - R)/(NIR + R + L)$	R = Reflectancia en el rojo L = 0.5 (constante)
Índice de vegetación mejorada (Huete, 2002)	$EVI = 2.5 * \frac{(NIR - R)}{NIR + 6 * R - 7.5 * B + 1}$	B = Reflectancia en el azul
Índice de clorofila verde (Gitelson et al, 2003)	$GCI = \left(\frac{NIR}{G}\right) - 1$	G = Reflectancia en el verde

Índice de reflectancia carotenoides 2 (Gitelson et al, 2002)	$CRI2 = \frac{1}{B} - \frac{1}{R}$	
---	------------------------------------	--

De esta manera, se obtuvieron 3 componentes principales, considerando un procesamiento sin rotación para la estimación de componentes en todas las imágenes desde los eigenvector.

Con el conjunto de imágenes de componentes principales se estima la matriz de varianzas y covarianzas, su media, mínima y máxima, se hizo para cada clase, y generar una firma espectral utilizada en la clasificación con máxima verosimilitud (etapa 6)

5.2 Fase III. Validación de resultados

Para la validación del modelo (etapa 7) correspondiente a cada nivel jerárquico de clasificación se propuso estimar la precisión del error global y por clase, así como el índice kappa.

La estructura que muestre una mejor aproximación puede ser utilizada para clasificar las escenas completas (etapa 8) e identificar aquel criterio que muestre mejor correspondencia entre los componentes principales y sus clases (etapa 9). Por ejemplo, de manera gráfica, la figura 5a muestra una mejor agrupación de los pixeles a diferencia de la figura 5b, en que los sitios de entrenamiento pueden llegar a confundirse entre sí, y por ende la agrupación (clasificación) no es la más adecuada.

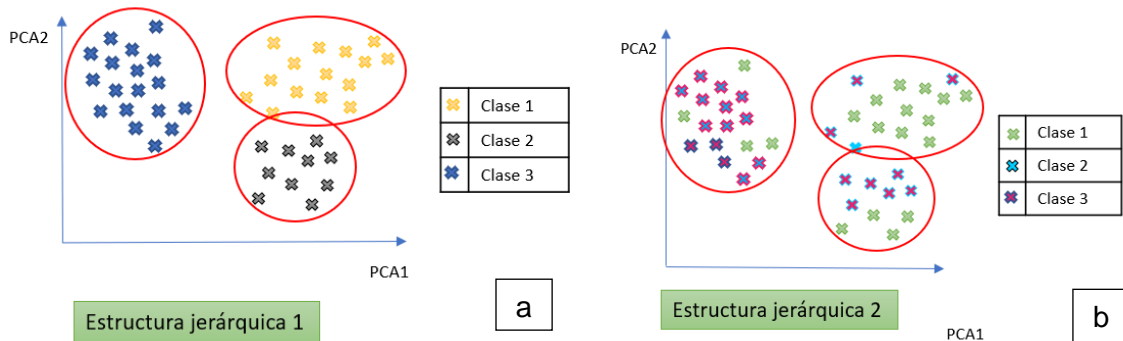
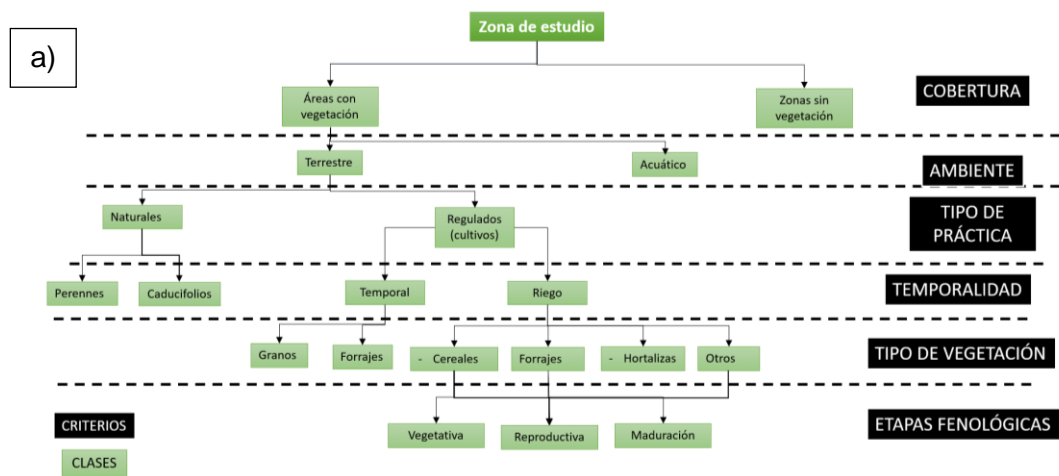


Figura 5a y 5b. Clasificaciones puntuales resultantes

CAPITULO 5. RESULTADOS Y DISCUSION

Entre los 19 criterios para la identificación y clasificación de la vegetación que fueron mencionados en la metodología, algunos de ellos se consideraron como parte de las variables en el ACP. Este fue el caso de las variables climáticas y algunas propiedades espectrales de la vegetación, principalmente la clorofila y los carotenoides. Por lo tanto y para complementar la clasificación de la etapa 1, los subsecuentes niveles de clasificación propuestos consideran los criterios de temporalidad y de una manera paralela dos alternativas diferenciadas por el tipo de desarrollo de la vegetación (anexo 1).

La primera de ellas (tipo de vegetación y desarrollo fenológico; Figura 6a), pretende identificar el grado de crecimiento en la cubierta vegetal sin importar el cultivo o vegetación del cual se trate, mientras que la segunda (grupo y tipo de vegetación figura 6b) excluye el desarrollo fenológico para determinar el cultivo o vegetación presente en el área de estudio. Posteriormente se espera verificar la idoneidad de alguna de las dos alternativas para identificar de manera correcta el tipo de vegetación como siguiente nivel de clasificación. Dentro del alcance del presente trabajo, se abordaron los niveles de cobertura, ambiente y tipo de práctica.



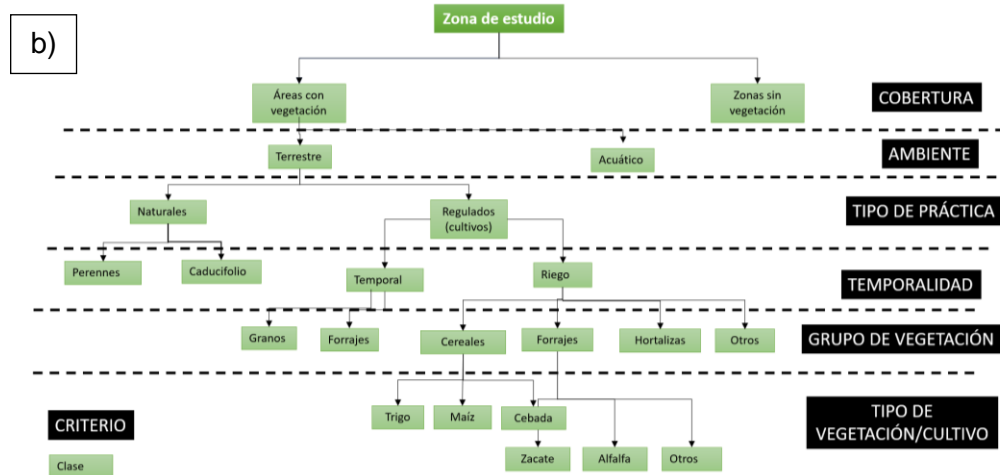


Figura 6. Diagramas con criterios para la clasificación de cubiertas vegetales/cultivos a) clasificación de acuerdo con su etapa fenológica; b) Clasificación por cultivo o vegetación.

Los criterios y clases seleccionados de la Figura 6 para la clasificación de cubiertas vegetales fue aplicado para la zona centro del Altiplano Central Mexicano, abarcando principalmente 6 Distritos de Riego. Para la clasificación de esta zona se seleccionaron aproximadamente 100 sitios de entrenamiento (Figura 7) por escena (2646 y 2746), es decir, los sitios de entrenamiento entre escenas mostraron algunas diferencias esperadas, en su distribución espacial. Para ello, fue utilizada una segmentación con un factor 30 de tolerancia en similitud de 30.

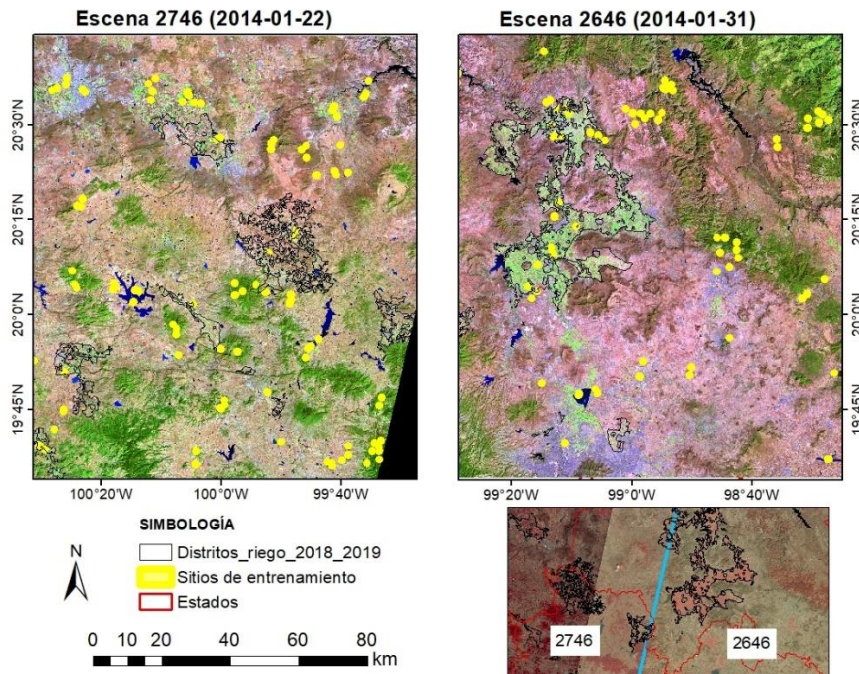


Figura 7. Puntos de muestreo seleccionados para el año 2013

En relación con la disponibilidad de información el área de estudio abarcó 2 imágenes Landsat (2746 y 2646, Tabla 8). En el pretratamiento de las imágenes se verificó un porcentaje de nubosidad no mayor al 30%, y se realizó una corrección atmosférica y radiométrica en TerrSet. En total se contó con 36 imágenes Landsat a lo largo de los 4 años de estudio.

Tabla 8. Disponibilidad de imágenes Landsat del 2013 - 2016

	2013												2014												2015												2016												
CLAVE IMAGEN	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12	
2646																																																	
2746																																																	

Simbología
■ 1 imagen
■ 2 imágenes

Considerando las imágenes por separado, se realizó la clasificación base (áreas con y sin vegetación; etapa 4). Esto con la finalidad de verificar los sitios de entrenamiento de niveles de clasificación subsecuentes principalmente. Para ello, en primera instancia se estimaron y trabajaron en conjunto los índices de vegetación NDWI, SAVI y EVI. Por ejemplo, la Figura 8 muestra el resultado de los índices para el 31 de enero de 2014, donde se puede observar como el DR 003 Tula es acotado por rangos de índices aproximados entre 0.08 y 0.8, -0.26 a 0.5 y 0.08 a 0.6 aproximadamente, para los índices EVI, NDWI y SAVI respectivamente, todo eso asociado a la vegetación

El resultado de la clasificación para los cuatro años analizados por máxima verosimilitud (tabla 9) mostró una precisión global (OA por el término en inglés Overall Accuracy) entre el 70% y el 92% con una precisión media (AA del término Average Accuracy) entre 66% y 81%, respecto con las tres clases consideradas en la primera clasificación: Áreas con vegetación, sin vegetación y no identificado. Esta última clase permitió agrupar celdas correspondientes a nubes, el background de la imagen, y para los niveles subsecuentes, las clases descartadas de niveles anteriores. En términos generales el coeficiente Kappa K (Tao & Mughees, 2021) obtenido de manera ponderada por todas las imágenes y para las tres clases fue de 0.73 (figura 9a). Por clase, este índice se encontró entre 0.66 y 0.80, representando una concordancia adecuada de acuerdo con Russel & Kass (1999). La clase de vegetación mostró una mayor frecuencia en las imágenes para valores de Kappa mayores a 0.60 (figura 9b). La clase sin vegetación fue la mejor representada con el grueso de imágenes en valores de kappa mayores a 0.70 (figura 9c) y la clase “no identificada” mostró una distribución similar de frecuencia para las imágenes con valores

kappa menores a 0.70 (figura 9d). Los datos por imagen de los resultados de máxima verosimilitud se encuentran en el anexo 2.

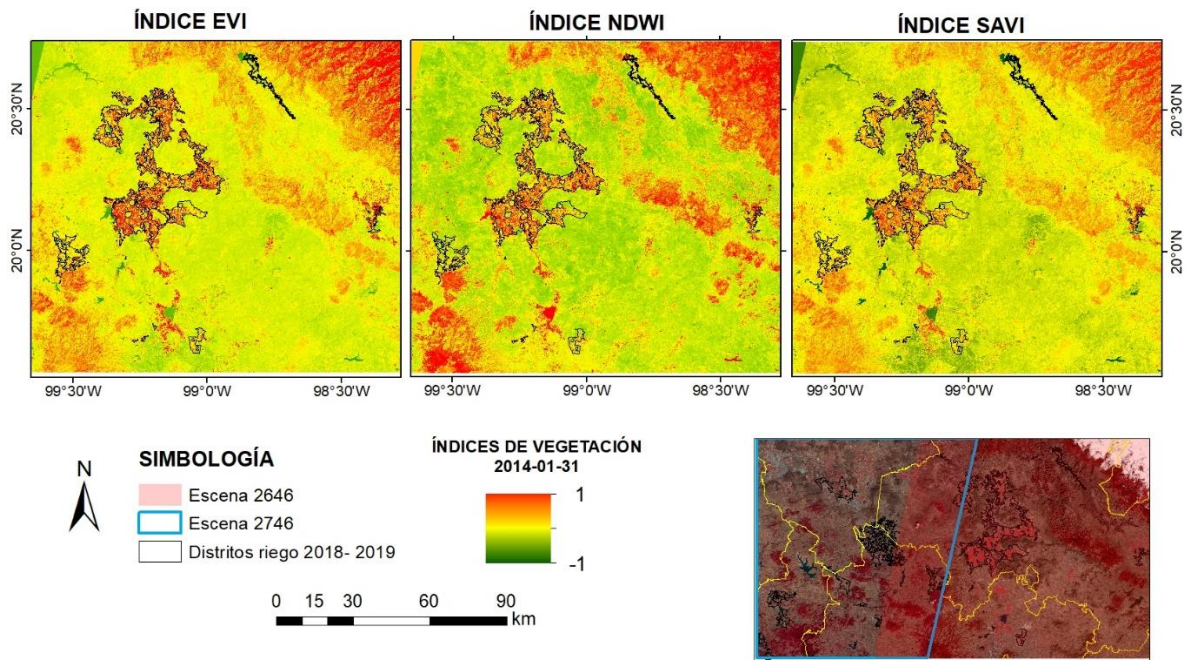


Figura 8. Resultados de índices de vegetación 2013-04-18

Tabla 9. Valores del índice Kappa del 2013 – 2016

ID Fecha	Clase 1		Clase 2		Clase 3		Total de pixeles	Kappa total	kappa (porcentaje)	Precisión	PONDERACIÓN			
	Total de pixeles	Kappa	Total de pixeles	Kappa	Total de pixeles	Kappa					CLASE 1	CLASE 2	CLASE 3	TOTAL
1	3692	0.600248	2420	0.844333	0	0	6112	0.701669	70.1669	29.8331	2216.116	2043.286	0	4288.601
2	2810	0.844195	2453	0.909549	343	0.563883	5606	0.839146	83.9146	16.0854	2372.188	2231.124	193.412	4704.252
3	3981	0.561977	1497	0.87154	805	0.972924	6283	0.726724	72.6724	27.3276	2237.230	1304.695	783.204	4566.007
4	3911	0.673986	1809	0.897882	1112	0.833285	6832	0.775299	77.5299	22.4701	2635.959	1624.269	926.613	5296.843
5	4453	0.601274	1477	0.872664	943	0.689598	6873	0.689902	68.9902	31.0098	2677.473	1288.925	650.291	4741.696
6	3160	0.850413	3171	0.54006	445	0.846791	6776	0.689472	68.9472	31.0528	2687.305	1712.530	376.822	4671.862
7	3471	0.682525	2704	0.857317	0	0	6175	0.76	76	24	2369.044	2318.185	0	4693.000
8	3454	0.774739	3840	0.847626	0	0	7294	0.809545	80.9545	19.0455	2675.949	3254.884	0	5904.821
9	3070	0.724226	4186	0.786658	1644	0.83441	8900	0.775221	77.5221	22.4779	2223.374	3292.950	1371.770	6899.467
10	3013	0.755812	2793	0.816409	427	0.938924	6233	0.801437	80.1437	19.8563	2277.262	2280.230	400.921	4995.357
11	3189	0.931611	1834	0.816035	776	0.674077	5799	0.834999	83.4999	16.5001	2970.907	1496.608	523.084	4842.159
12	3582	0.637931	2220	0.82723	643	0.826008	6445	0.734348	73.4348	26.5652	2285.069	1836.451	531.123	4732.873
13	3035	0.607506	2529	0.840784	0	0	5564	0.705358	70.5358	29.4642	1843.781	2126.343	0	3924.612
14	2870	0.387241	1962	0.792855	1119	0.873482	5951	0.607024	60.7024	39.2976	1111.382	1555.582	977.426	3612.400
15	3691	0.363833	2465	0.755104	0	0	6156	0.491058	49.1058	50.8942	1342.908	1861.331	0	3022.953
16	2845	0.590825	2264	0.775923	551	0.524972	5660	0.644362	64.4362	35.5638	1680.897	1756.690	289.260	3647.089
17	3038	0.537410	2422	0.802356	609	0.706473	6069	0.650225	65.0225	34.9775	1632.652	1943.306	430.242	3946.216
18	3070	0.732798	2847	0.80666	975	0.953401	6892	0.804777	80.4777	19.5223	2249.690	2296.561	929.566	5546.523
19	4050	0.782458	1796	0.628885	575	0.583486	6421	0.686785	68.6785	31.3215	3168.955	1129.477	335.504	4409.846
20	3629	0.539263	1706	0.924377	4143	0.847084	9478	0.733511	73.3511	26.6489	1956.985	1576.987	3509.469	6952.217
21	3918	0.773048	1906	0.836343	933	0.784104	6757	0.79655	79.655	20.345	3028.802	1594.070	731.569	5382.288
22	3730	0.917003	1820	0.885401	823	0.701699	6373	0.860871	86.0871	13.9129	3420.421	1611.430	577.498	5486.331
23	3305	0.842197	1416	0.92181	3390	0.834666	8111	0.855741	85.5741	14.4259	2783.461	1305.283	2829.518	6940.915
24	4065	0.905030	5998	0.76008	681	0.664792	10744	0.809457	80.9457	19.0543	3678.947	4558.960	452.723	8696.806
25	4148	0.579125	1472	0.754398	1069	0.693283	6689	0.650864	65.0864	34.9136	2402.211	1110.474	741.120	4353.629
26	4459	0.627229	2120	0.749668	618	0.915318	7197	0.709377	70.9377	29.0623	2796.814	1589.296	565.667	5105.386
27	4554	0.447300	1838	0.790032	681	0.951216	7073	0.620402	62.0402	37.9598	2037.004	1452.079	647.778	4388.103
28	3828	0.422680	1994	0.665715	554	0.862352	6376	0.547958	54.7958	45.2042	1618.019	1327.436	477.743	3493.780
29	4359	0.647966	1998	0.61607	763	0.895362	7120	0.682016	68.2016	31.7984	2824.484	1230.908	683.161	4855.954
30	4317	0.547467	1947	0.806516	1123	0.629568	7387	0.633071	63.3071	36.6929	2363.415	1570.287	707.005	4676.495
31	5203	0.905546	5517	0.721175	0	0	10720	0.802912	80.2912	19.7088	4711.556	3978.722	0	8607.217
32	4333	0.602411	1546	0.865195	855	0.856055	6734	0.725687	72.5687	27.4313	2610.247	1337.591	731.927	4886.776
33	4410	0.529640	1346	0.94087	488	0.950372	6244	0.702199	70.2199	29.7801	2335.712	1266.411	463.782	4384.531
34	4625	0.843311	2524	0.867012	699	0.662991	7848	0.824819	82.4819	17.5181	3900.313	2188.338	463.431	6473.180
35	4102	0.609671	2161	0.824098	1822	0.911586	8085	0.751721	75.1721	24.8279	2500.870	1780.876	1660.910	6077.664
36	5065	0.476803	2514	0.684014	0	0	7579	0.562521	56.2521	43.7479	2415.007	1719.611	0	4263.347
SUMA TOTAL DE PÍXELES	136435		86512		29609		252556			SUMA	90042.409	68552.186	23962.537	183471.198
									MEDIA PONDERADA		0.6600	0.7924	0.8093	0.7265

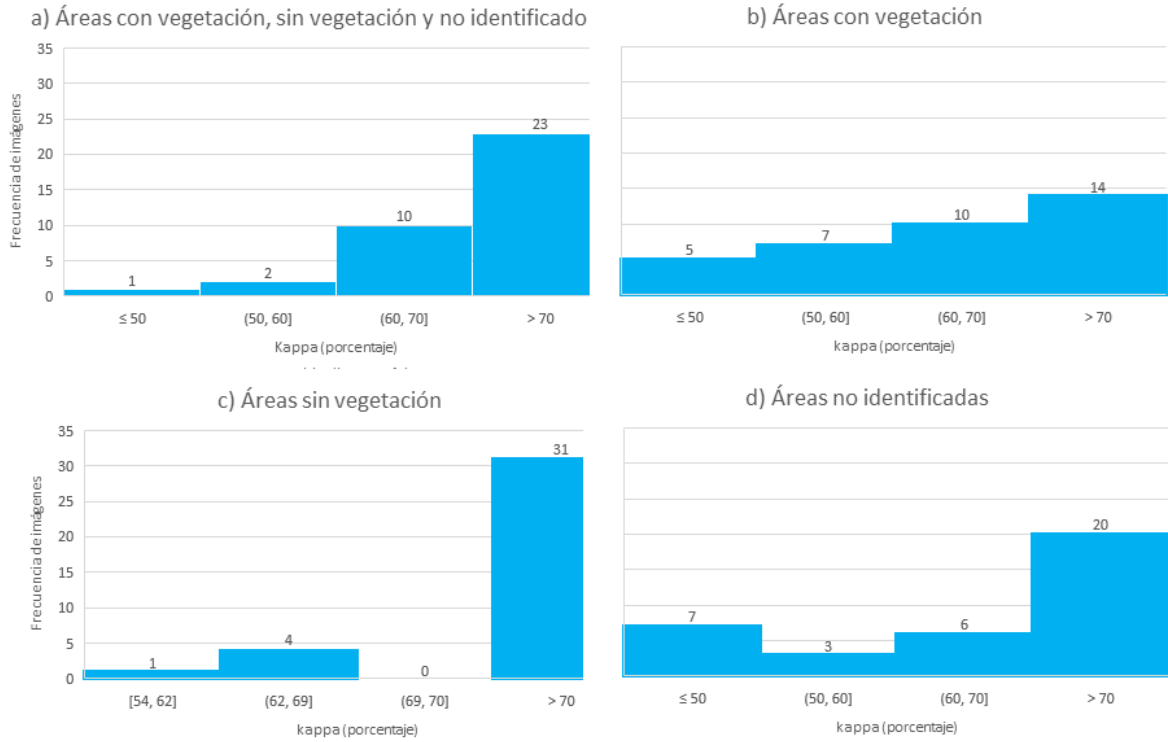


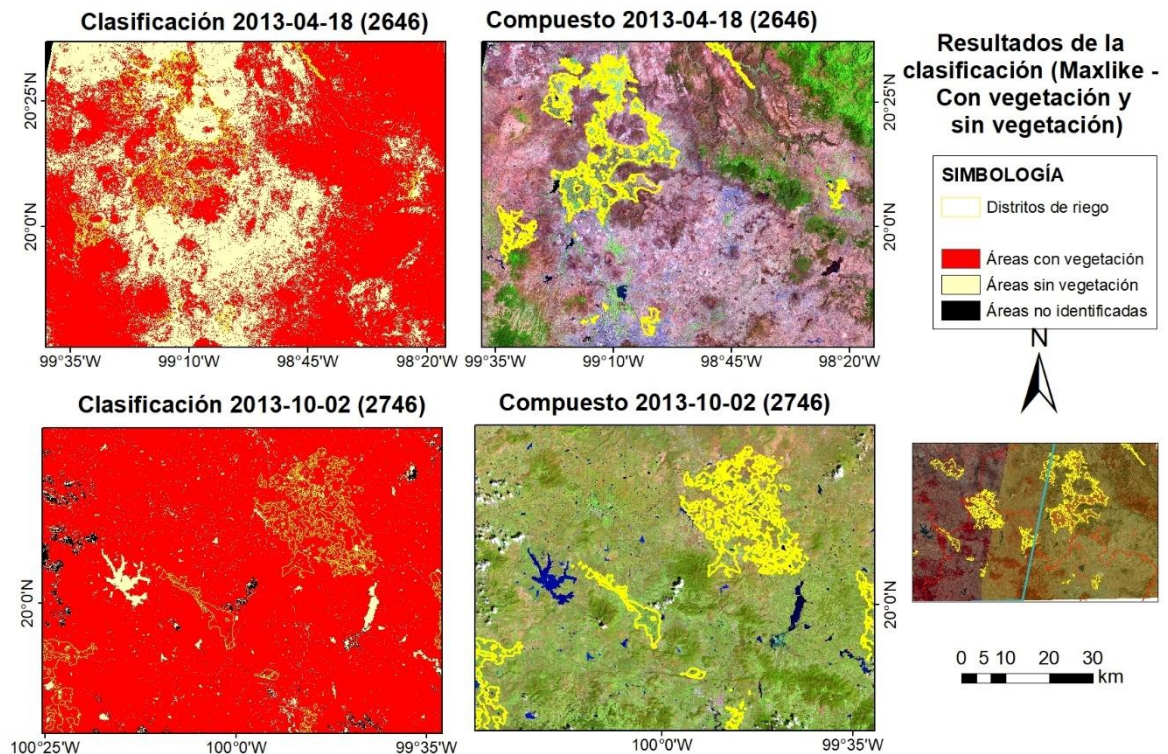
Figura 9.a) Gráfica de los coeficientes kappa (porcentaje) 2013-2016, para las tres clases; b) frecuencia de los valores de kappa para áreas con vegetación; c) frecuencia de coeficiente kappa para zonas sin vegetación; d) kappa para áreas no identificadas.

De manera general y considerando los 4 años del estudio, el 2016 mostró mejores resultados, es decir, el error total ponderado fue menor (tabla 10).

Tabla 10. Errores totales ponderados por año

AÑOS	2013	2014	2015	2016
Error Total Ponderado	0.147	0.156	0.16	0.14

A partir de las imágenes clasificadas (figura 10), se encontró coherencia con la información consultada de la cobertura de los distritos de riego. Con ello, fue posible estimar coberturas de vegetación entre 1,800,000 ha y 400,000 ha en las imágenes, representa en términos medios un 80 hasta 50 % de la superficie.



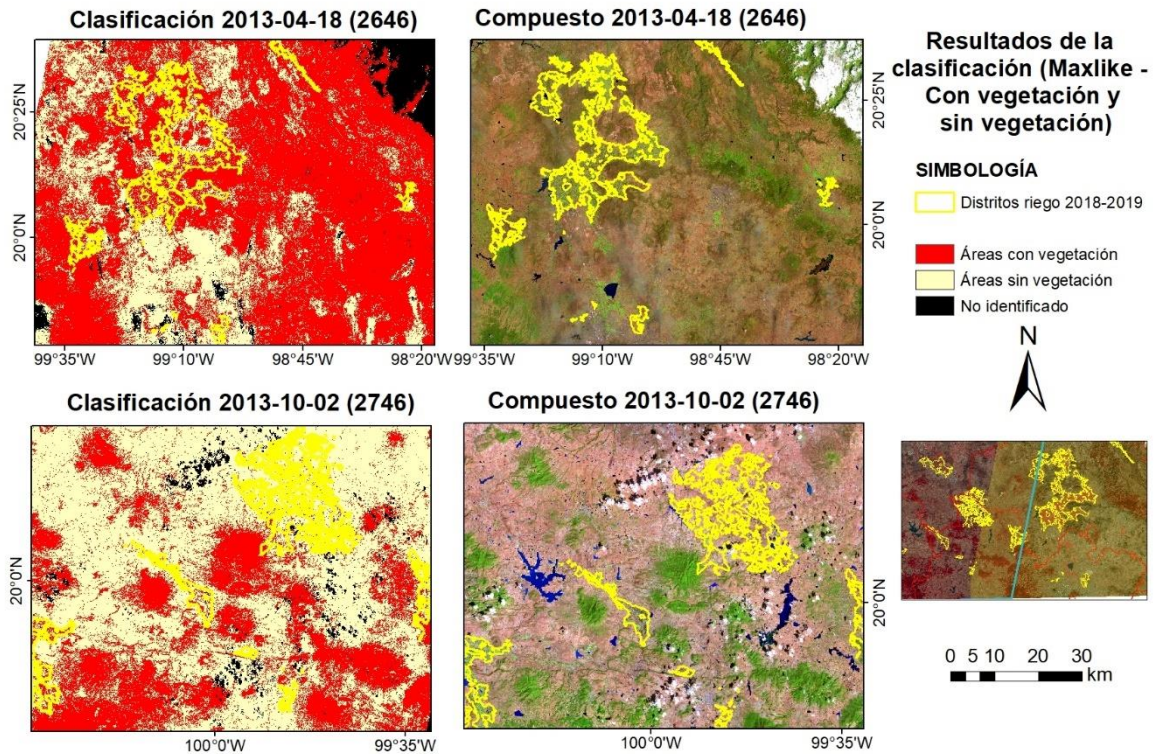


Figura 10. Ejemplo de imágenes clasificadas en el primer nivel (áreas con vegetación y sin vegetación)

Con la obtención del primer nivel de clasificación, fue posible estimar las variables adicionales a los índices espectrales requeridas para los componentes principales (CP) por fecha. Estas variables fueron la clorofila, carotenoides, índice de área foliar, y variables climáticas como la precipitación, la temperatura y la humedad relativa, en total se consideraron nueve variables para los componentes principales. A partir de estos datos fue extraída una muestra estratificada.

Los datos adquiridos se trabajaron de manera conjunta en SPSS para obtener los valores de los eigen vector que se consideraron al momento de estimar los CP en TerrSet (tabla 11). Los datos se procesaron de manera separada por clases, pero con todas sus variables, con la finalidad de generar las firmas espectrales de las nuevas clases a partir del Análisis de Componentes Principales (ACP) llevado a cabo en SPSS. La varianza explicada de los 3 componentes fue de 90.61, 5 y 3.69 respectivamente. Además, con base en los coeficientes estandarizados se puede observar que las variables precipitación, EVI y LAI, influyen más en la capacidad explicativa de la varianza del componente 1. El componente 2 es representado en mayor medida por los carotenoides, mientras que el componente 3 por la humedad relativa, la clorofila y la temperatura.

Tabla 11. Valores de los eigen vector para áreas terrestres y acuáticas

Variables	Componente puro			Componente reescalado		
	1	2	3	1	2	3
EVI	0.048	-0.027	-0.012	0.267	-0.149	-0.068
NDWI	0.041	0.020	-0.016	0.202	0.097	-0.078
SAVI	0.027	-0.005	-0.009	0.196	-0.035	-0.063
Carotenoides	-1.151	8.271	0.139	-0.138	0.989	0.017
Clorofila	-0.107	0.674	-0.318	-0.052	0.329	-0.155
LAI	0.172	-0.096	-0.044	0.268	-0.150	-0.068
Precipitación mensual	35.681	0.292	-0.400	1.000	0.008	-0.011
Humedad relativa	1.933	0.032	7.156	0.261	0.004	0.965
Temperatura mensual	0.693	-1.271	0.829	0.243	-0.445	0.290

A partir de las imágenes de componentes principales se estimaron por clase los valores mínimos, máximos y la matriz de covarianzas, requeridos para generar las firmas espectrales. La tabla 12 muestra las matrices para el segundo nivel de clasificación, y es posible observar que, aunque los valores medios de los componentes pueden ser similares entre ambas clases, los valores altos de varianza para la clase de vegetación acuática pueden generar una mayor confusión en la clasificación.

Tabla 12. Datos resultantes para las firmas espectrales

	VEGETACIÓN ACUÁTICA			VEGETACIÓN TERRESTRE		
	CP1	CP2	CP3	CP1	CP2	CP3
Media	1247.19	55.89	395.63	1111.28	50.75	406.66
Mínima	37.92	-21.23	293.59	7.27	-29.04	271.43
Máxima	5074.25	348.60	501.62	4987.38	882.00	559.92
Matrix de covarianzas						
	CP1	CP2	CP3	CP1	CP2	CP3
CP1	2298415.93			1473793.83		
CP2	-13393.81	3283.15		3048.02	5377.56	
CP3	-5471.87	118.07	2396.72	1587.89	-17.17	2776.90

En efecto, las firmas espectrales mostraron valores similares entre la clase terrestre y acuática, por lo que el coeficiente kappa se encontró por debajo del 50% en la mayoría de los casos (figura 11a). Una de las razones de este resultado puede ser por la similitud de la variable precipitación, en ambas clases, principalmente con la vegetación con riego, pero, además, por la baja cantidad de vegetación acuática encontrada en el área de estudio.

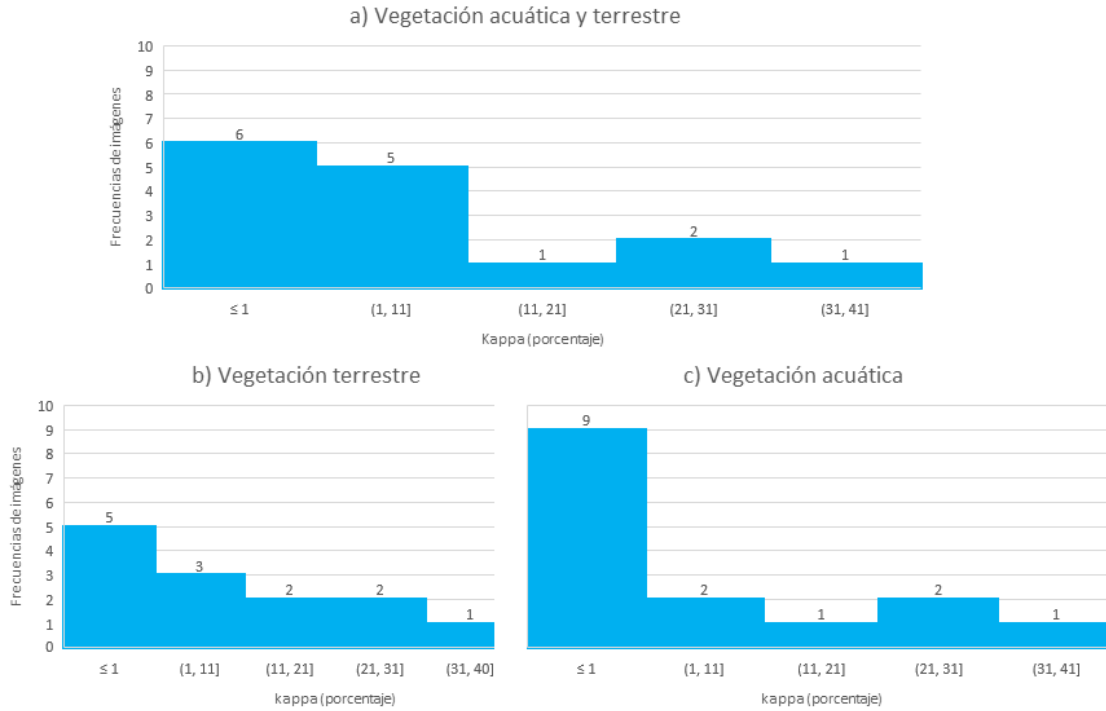


Figura 11. Coeficientes kappa total y por clase entre 2013-2016

En la figura 12, se muestran algunos ejemplos de los resultados de este nivel de clasificación comparándolo con una imagen en falso color, comprobando con ello que el resultado como lo mostró el índice kappa no corresponde con la vegetación presente en el área de estudio.

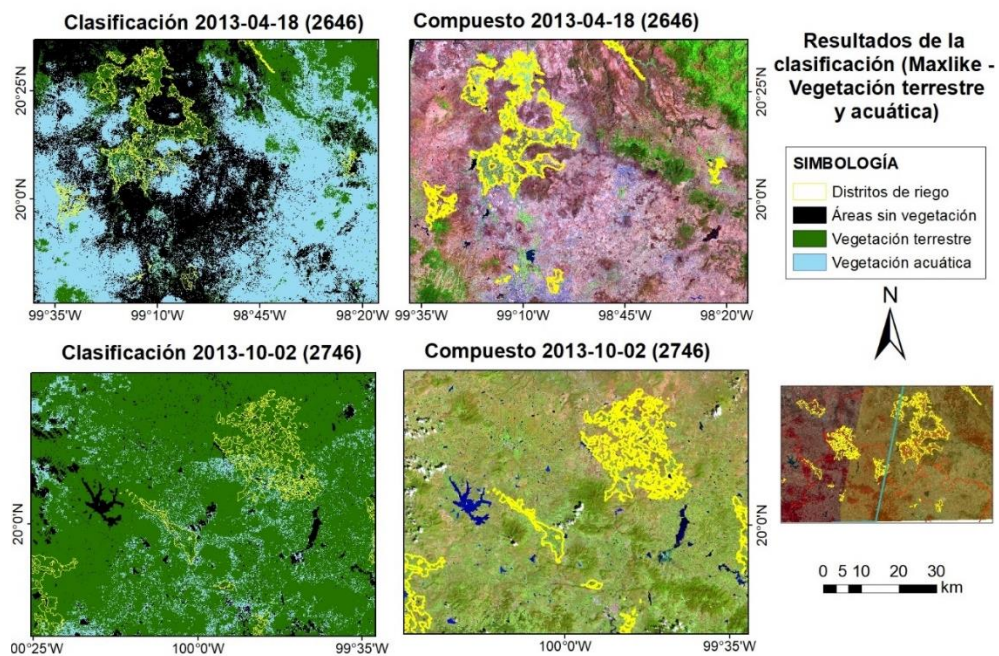


Figura 12. Resultados de la clasificación para áreas con vegetación terrestre y acuática

Bajo un proceso análogo, el siguiente nivel de clasificación (en función de la práctica regulado y natural), mostró mejores resultados y datos diferentes entre sí para generar las firmas espectrales correspondientes. Para este caso la varianza explicada de los 3 componentes fue de 86.81, 8.81 y 3.73, en este caso también la precipitación es la variable con mayor peso para el caso del primer componente, como también el índice de área foliar y la humedad relativa. El componente 2 está representado por los carotenoides, la clorofila y el índice SAVI, mientras que el componente 3 por la humedad relativa, la temperatura mensual y la clorofila.

Tabla 13. Valores de los eigen vector para áreas naturales y reguladas

MATRIZ DE COMPONENTES PRINCIPALES						
Variable	Componente puro			Componente reescalado		
	1	2	3	1	2	3
EVI	0.041	-0.046	-0.002	0.233	-0.257	-0.010
NDWI	0.039	0.014	0.002	0.226	0.083	0.009
SAVI	0.021	-0.015	0.001	0.160	-0.112	0.007
Carotenoides	-1.463	9.352	0.129	-0.154	0.987	0.014
Clorofila	-0.149	0.607	-0.292	-0.072	0.291	-0.140
LAI	0.153	-0.144	0.002	0.245	-0.230	0.003
Precipitación mensual	36.040	0.404	-0.374	1.000	0.011	-0.010
Humedad relativa	1.804	0.054	7.205	0.243	0.007	0.969
Temperatura mensual	0.655	-1.322	0.961	0.220	-0.444	0.323

Se obtuvieron mejores resultados para este nivel de clasificación (tabla 14), es decir, la diferencia entre las medias de cada componente fue más dispersas, lo que ayudo a generar las firmas espectrales con mayor varianza, esto mejoró los resultados en este nivel de clasificación.

Tabla 14. Datos resultantes para las firmas espectrales

	VEGETACIÓN REGULADA			VEGETACIÓN NATURAL		
	CP1	CP2	CP3	CP1	CP2	CP3
Media	1142.80	47.12	419.84	1221.36	97.45	401.82
Mínima	65.31	-21.59	312.00	-29.29	-28.16	244.18
Máxima	6023.27	202.84	515.30	5227.43	970.50	567.31
Matriz de Covarianzas						
	CP1	CP2	CP3	CP1	CP2	CP3
CP1	1473121.17			1924331.79		
CP2	2932.80	1325.54		-4841.28	13416.04	
CP3	2504.72	26.33	2138.02	-1759.69	423.52	3323.45

La precisión general se encontró entre 40% a 94%, predominando las imágenes con una precisión entre el 20 y 50% (figura 13a). El valor de kappa mostró un ponderado del 52% para áreas con vegetación regulada (figura 13b) y del 28 % para zonas con vegetación natural (figura 13c). Anexo 3.

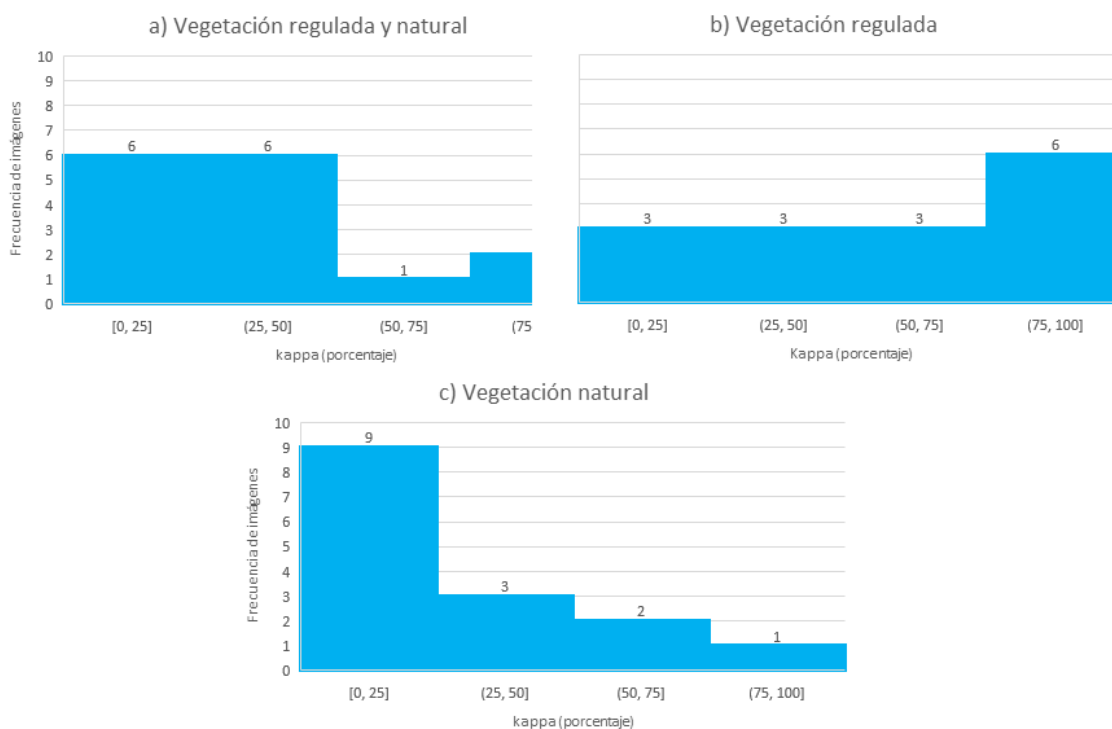


Figura 13. Coeficientes kappa total y por clase para vegetación regulada y natural

Los resultados de la clasificación para este nivel tuvieron una mayor concordancia con la realidad, en la figura 14 se muestran los resultados comparándolos con su compuesto en falso color, por ejemplo, para el caso de la vegetación natural las tonalidades en el compuesto son verdes claro, al compararlas con los resultados se observa que el resultado concuerda en gran medida.

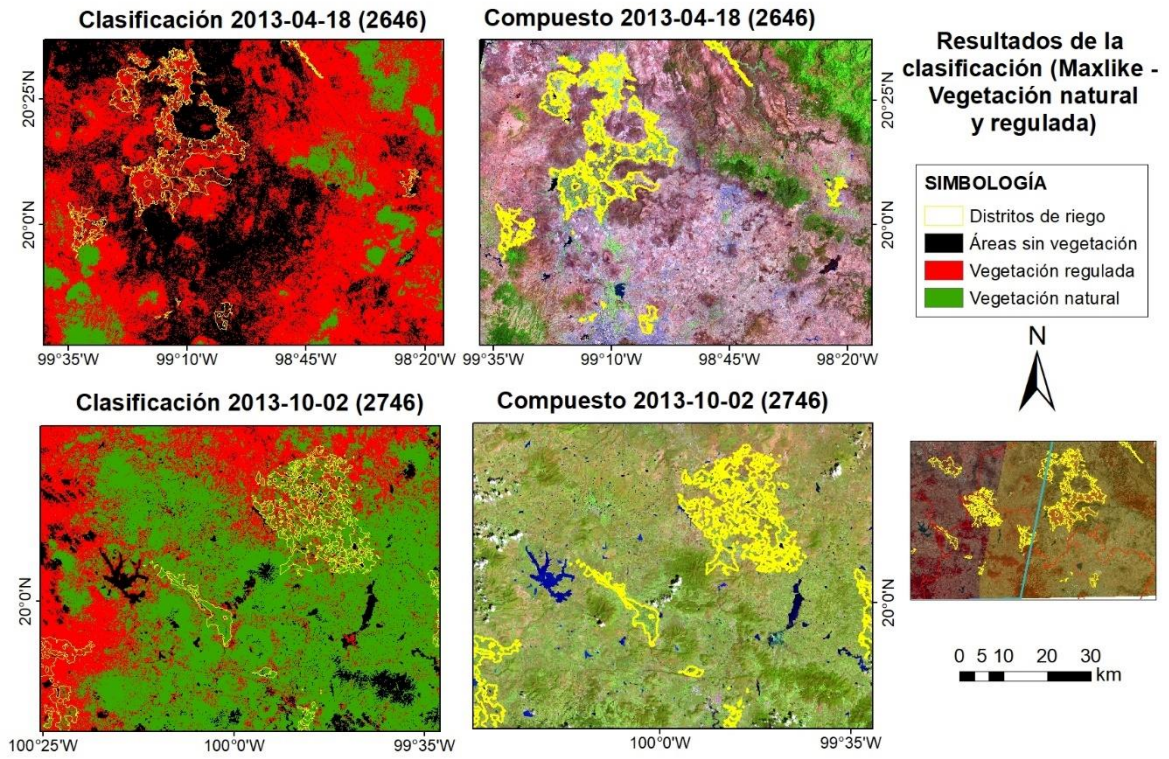


Figura 14. Clasificación regulada y natural

CONCLUSIONES

Considerando el objetivo de desarrollar un modelo basado en Análisis de Componentes Principales para la identificación de diferentes tipos de vegetación por medio de sensores remotos, se estimaron 3 componentes principales por fecha, a partir de 9 variables: tres índices espectrales (SAVI, EVI y NDWI), tres variables climatológicas (precipitación media, temperatura promedio y humedad relativa) y, tres características vegetales (clorofila, carotenoides e índice de área foliar). Como zona de estudio se consideró la región que abarca principalmente los distritos de riego 003, 0023, 033, 044 y 096.

Los tipos de vegetación clasificados fueron obtenidos de un análisis de varios criterios identificados por la FAO, considerando el primer objetivo específico. En este sentido, se propuso una jerarquización de niveles de clasificación para encontrar el más adecuado. Sin embargo, el alcance del presente proyecto permitió la evaluación de solo tres niveles: cobertura vegetal, ambiente y práctica. Cada uno de estos criterios proporcionó dos clases (omitiendo la clase por default "No identificado"): áreas con y sin vegetación para el primer nivel, vegetación acuática y terrestre para el segundo nivel y, vegetación natural y regulada para el tercer nivel. El primer nivel de clasificación fue realizado con base en máxima verosimilitud a partir de los índices espectrales mencionados. Los siguientes niveles requirieron del análisis de componentes principales

Los componentes principales fueron estimados, con base en el segundo objetivo específico, en la zona de estudio en el periodo de 2013 a 2016. Con ellos, fue posible evaluar su correspondencia, bajo un criterio de máxima verosimilitud, con los siguientes niveles de clasificación de cobertura vegetal a partir de la información encontrada para 100 sitios de entrenamiento por imagen.

Estos sitios de entrenamiento fueron identificados, como se esperó en el tercer objetivo específico, a partir de información encontrada en literatura, información proporcionada por autoridades regionales y bajo propia experiencia y conocimiento de la zona de estudio.

La correspondencia de los componentes principales con los niveles de clasificación conformando el cuarto objetivo específico, fue evaluada a partir de la precisión global y local, así como con el índice Kappa. Sin embargo, se recomienda para estudios posteriores el análisis con el resto de los niveles de clasificación propuestos en la jerarquización. La clasificación requirió una matriz de varianzas y covarianzas entre los valores de componentes principales obtenidos desde una muestra estratificada y aleatoria

entre todas las imágenes disponibles para el estudio. Estos elementos obtenidos de manera conjunta permitieron evaluar las clases consideradas en imágenes, incluso sin sitios de entrenamiento. Sin embargo, hay que tomar en cuenta las limitaciones presentadas como la disminución de precisión y el acotamiento del número de clases consideradas en cada nivel. Este acotamiento de clases significa que, para las imágenes evaluadas se puede presentar la clase “no identificado” cuando no se tiene la información suficiente en el nivel de clasificación.

El primer nivel, con las clases áreas con vegetación, sin vegetación y zonas de identificadas, mostró una clasificación adecuada considerando únicamente los índices espectrales. La precisión media fue de entre 66% y 81% y un índice Kappa medio de 0.73. Sin embargo, el nivel correspondiente al ambiente mostró una pobre correspondencia para las clases terrestre y acuática con precisión e índice Kappa medios entre 33 al 70% y <50% respectivamente. El nivel correspondiente al tipo de práctica mostró una mejoría en la correspondencia en relación con el nivel anterior incrementando la precisión y el índice Kappa entre 40 a 94 % y del 52% respectivamente.

Lo anterior demuestra que, en función del tipo de clases considerado y de las variables incluidas en el estudio, la clasificación de coberturas vegetales puede ser llevada a cabo mediante componentes principales. Pero, probablemente más importante, este procedimiento proporciona indicios de una identificación de coberturas vegetales mediante una clasificación no supervisada. Además de que al obtener información en diferentes temporadas del año toma características y datos diversos de los distintos tipos de vegetación, por lo que ayuda a mejorar la clasificación de la información.

RECOMENDACIONES

Para el caso de la primera clasificación (base) se podrían considerar más sitios de entrenamiento, para seleccionar más áreas con vegetación diferentes y con ello mejorar los resultados de los niveles posteriores.

La clasificación de los niveles posteriores al nivel base se realizó con los componentes principales, estos incluyeron 9 variables, una de ellas fue la precipitación que como se observó presentó un valor alto en los componentes, por tanto, fue un factor que influyó mucho en la clasificación de las áreas con vegetación acuática y terrestre, por ende los resultados no fueron los más favorables como lo demostró el coeficiente kappa, por ello se realiza la observación de no considerar esa variable en estudios posteriores para que la clasificación mejore.

Realizar los otros niveles de clasificación, para verificar que tan bueno es el modelo desarrollado, además de comparar este modelo con otros métodos de clasificación, para verificar que tan apropiado es la aplicación del modelo, por el tiempo de procesamiento que se requiere respecto de otros métodos ya desarrollados.

CAPITULO 7. REFERENCIAS

- Acevedo, J. A. A., & Hernández, G. M. V. (2013). Fenología de ambientes tropicales en el marco de la teledetección. *GeoFocus. Revista Internacional de Ciencia y Tecnología de la Información Geográfica*, 13_2, 195-211.
- Araque, L., & Jiménez, A. (2009). Caracterización de firma espectral a partir de sensores remotos para el manejo de sanidad vegetal en el cultivo de palma de aceite. *Revista Palmas*, 30(3), 63-79.
- Ayala-Carcedo, F. J., & Olcina Cantos, J. (2002). *Riesgos Naturales* (1°). Ariel.
- Bohn, V. Y., Fornerón, C., & Piccolo, M. C. (2012). Análisis de cubiertas vegetales en ambientes naturales del sudoeste de la provincia de Buenos Aires (Argentina). *JORNADAS NACIONALES DE GEOGRAFÍA FÍSICA*, 9, 1-12.
- Castañeda-Ibáñez, C. R., Martínez-Menes, M., Pascual-Ramírez, F., Flores-Magdaleno, H., Fernández-Reynoso, D. S., & Esparza-Govea, S. (2015). Estimación de coeficientes de cultivo mediante sensores remotos en el distrito de riego río Yaqui, Sonora, México. *Agrociencia*, 49(2), 221-232.
- Chaparro, D. M. G., & Quintero, S. B. (2003). Modelamiento difuso con técnicas de Clustering. *Ingeniería*, 8(1), 86-94.
- De la Fuente, F. (2011). Santiago (2011b).“. *Análisis Componentes Principales-ACP*”, *Material de apoyo docente, Santiago de la Fuente Fernández, Fac. Económica y de Ciencias Empresariales, Univ. Autónoma de Madrid (UAM)*.
- Espinoza García, J. G., Olvera Olvera, C. A., Espinoza García, G., Moreno Lucio, M., Araiza Esquivel, M. A., Villagrana Barraza, S., & López Monteagudo, F. E. (2019). *Identificación de parámetros comunes en un cultivo de cebolla (Allium cepa) mediante PDI*.
- Fonseca, C. R., Cruz-Retana, A., Gómez-Albórez, M. A., & Becerril-Piña, R. (2021). *Multiclass Data Organizer-TerrSet* por IITCA-UAEM. Toluca, México: Instituto Interamericano de Tecnología y Ciencias del Agua IITCA-UAEM.
- García, S. A., & Martínez, L. J. (2010). Método para identificación de cultivos de arroz (*Oryza sativa* L.) con base en imágenes de satélite. *Agronomía Colombiana*, 38(2), 281-290.
- Harris Geospatial Solutions. (2020). Vegetation Indices [Científica]. *L3HARRIS GEOSPATIAL*. <https://www.l3harrisgeospatial.com/docs/vegetationindices.html>
- IBM Corporation, 2021. IBM SPSS Statistics. United States. <https://www.ibm.com/downloads/cas/ERYRD6G5>

- Le, V. N. T., Ahderom, S., Apopei, B., & Alameh, K. (2020). A novel method for detecting morphologically similar crops and weeds based on the combination of contour masks and filtered Local Binary Pattern operators. *GigaScience*, 9(3), g1aa017.
- Marini, M. F., Vergara, M. F., & Krüger, H. (2007). Determinación del uso de la tierra en el partido de Guamini (Argentina) mediante un estudio multitemporal con imágenes Landsat. *Revista de Teledetección*, 27, 80-88.
- Martín, P. G., de Pascual, A. D., Lezama, E. T., & Olmos, E. G. (1994). Una aplicación del análisis de componentes principales en el área educativa. *Economía*, 19(9), 55-72.
- Nolasco, M., Willington, E., & Bocco, M. (2015). Uso del suelo agrícola: Comparación entre series temporales e imágenes satelitales individuales para su clasificación. *Revista de Investigaciones de la Facultad de Ciencias Agrarias-UNR*, 26, 017-021.
- Pascual-Ramírez, F., Paz-Pellat, F., & Bolaños-González, M. (2012). Estimación de biomasa aérea en cultivos con sensores remotos. *Terra Latinoamericana*, 30(1), 17-28.
- Peña-Barragán, J. M., Ngugi, M. K., Plant, R. E., & Six, J. (2011). Object-based crop identification using multiple vegetation indices, textural features and crop phenology. *Remote Sensing of Environment*, 115(6), 1301-1316.
- Rodrigo, J. A. (2017). Análisis de Componentes Principales (Principal component Analysis, PCA) y t-SNE. *GirHub*. https://github.com/JoaquinAmatRodrigo/Estadistica-con-R/blob/master/PDF_format/35_Principal_Component_Analysis.pdf
- Rogan, J., & chen, D. (2004). *Remote sensing technology for mapping and monitoring land-cover and land use-change*. 61, 301-325.
- Ronald, E. (2020). Manual IDRISI. United States: Universidad de Clarck.
- Ruiz, V., Savé, R., & Herrera, A. (2013). *Análisis multitemporal del cambio de uso del suelo, en el Paisaje Terrestre Protegido Miraflores Moropotente Nicaragua, 1993-2011*. https://rua.ua.es/dspace/bitstream/10045/34856/1/Ecosistemas_22_3_16.pdf
- Russel, G., & Kass, G. (1999). *Assesing the Accuracy of Remotely Sensed Data: Principles and Practices*. Lewis Publishers. <https://books.google.com.mx/books?id=THm9VJ40-iQC&pg=PA45&dq=what+is+overall+accuracy+in+classification&hl=es-419&sa=X&ved=2ahUKEwia3YvDurr3AhV8IUQIHx97BqYQ6AF6BAgCEAI#v=onepage&q=what%20is%20overall%20accuracy%20in%20classification&f=false>

- Sánchez, L. A. P., Pellat, F. P., Mota, J. L. O., Sandoval, B. F., Menes, M. M., Solorio, C. A. O., & García, A. E. (2006). Clasificador genérico de objetos en imágenes ETM+. *Agrociencia*, 40(5), 613-626.
- Satélite Landsat 8 (2013-2016). Centro de la República Mexicana (.tif). Earth Explorer. [EarthExplorer \(usgs.gov\)](https://earthexplorer.usgs.gov)
- Saxena, A., Prasad, M., Gupta, A., Bharill, N., Patel, O. P., Tiwari, A., Er, M. J., Ding, W., & Lin, C.-T. (2017). A review of clustering techniques and developments. *Neurocomputing*, 267, 664-681.
- Tao, L., & Mughees, A. (2021). *Deep Learning for Hyperspectral Image Analysis and Classification*. Springer.
[https://books.google.com.mx/books?id=30wfEAAAQBAJ&pg=PA25&lpg=PA25&dq=The+Kappa+Coefficient+\(K\),+Average+Accuracy+\(AA\)+and+Overall+Accuracy&source=bl&ots=nNIAvQjAo_&sig=ACfU3U0KRfeqoYkARANot0sFsxhyr2UfnA&hl=es-419&sa=X&ved=2ahUKEwjpi8DM7sLzAhX_m2oFHeS-BdlQ6AF6BAgZEAM#v=onepage&q=The%20Kappa%20Coefficient%20\(K\)%2C%20Average%20Accuracy%20\(AA\)%20and%20Overall%20Accuracy&f=false](https://books.google.com.mx/books?id=30wfEAAAQBAJ&pg=PA25&lpg=PA25&dq=The+Kappa+Coefficient+(K),+Average+Accuracy+(AA)+and+Overall+Accuracy&source=bl&ots=nNIAvQjAo_&sig=ACfU3U0KRfeqoYkARANot0sFsxhyr2UfnA&hl=es-419&sa=X&ved=2ahUKEwjpi8DM7sLzAhX_m2oFHeS-BdlQ6AF6BAgZEAM#v=onepage&q=The%20Kappa%20Coefficient%20(K)%2C%20Average%20Accuracy%20(AA)%20and%20Overall%20Accuracy&f=false)
- USGS. (2020). NDVI, the Foundation for Remote Sensing Phenology [Científica]. USGS. https://www.usgs.gov/core-science-systems/eros/phenology/science/ndvi-foundation-remote-sensing-phenology?qt-science_center_objects=0#qt-science_center_objects
- Zhang, X., Friedl, M. A., Schaaf, C. B., Strahler, A. H., Hodges, J. C., Gao, F., Reed, B. C., & Huete, A. (2003). Monitoring vegetation phenology using MODIS. *Remote sensing of environment*, 84(3), 471-475. [https://doi.org/10.1016/S0034-4257\(02\)00135-9](https://doi.org/10.1016/S0034-4257(02)00135-9)

ANEXOS

Anexo 1. Criterios y clases considerados en el modelo propuesto

CRITERIO	CLASE PRECEDENTE	CLASES
Cobertura	Ninguno	1. Vegetación 2. Sin vegetación
Ambiente	1. Vegetación	1.1. Terrestre 1.2. Acuático
Tipo de práctica	1.1. Terrestre	1.1.1. No regulado 1.1.2. Regulado
Temporalidad	1.1.1. No regulado	1.1.2.1. Perennes
	1.1.2. Regulado	1.1.2.1. Perennes 1.1.2.2. Anuales
Grupo de vegetación	1.1.2.1. Perennes	1.1.2.1.1. Pastizales 1.1.2.1.2. Caducifolios y arbustos
	1.1.2.1. Perennes	1.1.2.1.1. Hortalizas
	1.1.2.2. Anuales	1.1.2.2.1. Cereales 1.1.2.2.2. Forrajes 1.1.2.2.3 Leguminosas 1.1.2.2.n. Otros
Etapas fenológicas ^a	1.1.2.2.1. Cereales 1.1.2.2.2. Forrajes 1.1.2.2.3 Leguminosas 1.1.2.2.n. Otros (a)	1.1.2.2.1.1 Vegetativa 1.1.2.2.1.2 Reproductiva 1.1.2.2.1.3 Maduración
Familia/especie ^b	1.1.2.2.1. Cereales 1.1.2.2.2. Forrajes 1.1.2.2.3 Leguminosas 1.1.2.2.n. Otros (b)	1.1.2.2.1.1 Trigo 1.1.2.2.1.2 Maíz 1.1.2.2.1.3 Cebada 1.1.2.2.1.4 Avena Forrajera Verde 1.1.2.2.1.5 Sorgo Grano 1.1.2.2.2.1 Zacate 1.1.2.2.2.2 Alfalfa 1.1.2.2.2.3 Bermuda

		1.1.2.2.3.1 Frijol (Alubia) 1.1.2.2.3.2 Haba
--	--	---

Anexo 2. Resultados de la matriz de confusión por imagen clasificada con identificador para el primer nivel de clasificación (áreas con vegetación y sin vegetación).

ID 1.- 18 DE ABRIL DE 2013

	1	2	Total	ErrorC
1	2984	196	3180	0.061635
2	708	2224	2932	0.241473
Total	3692	2420	6112	
ErrorO	0.191766	0.080992		0.147906

- ErrorO = Errores de omisión (expresado en proporciones)
- ErrorC = Errores de Comision (expresado en proporciones)
- 90% Intervalo de confianza = +/- 0.007470 (0.140436 - 0.155376)
- 95% Intervalo de confianza = +/- 0.008900 (0.139006 - 0.156806)
- 99% Intervalo de confianza = +/- 0.011716 (0.136190 - 0.159621)

Indice kappa de concordancia (KIA – por sus siglas en inglés)

Usando el Maxlike	
Category	KIA
1	0.844333
2	0.600248

Utilizando la segmentación	
Category	KIA
1	0.600248
2	0.844333

Kappa promedio : 0.701669

ID 2. 20 DE MAYO DE 2013

	1	2	3	Total	ErrorC
1	2576	57	0	2633	0.02164831
2	154	2333	110	2597	0.101655757
3	80	63	233	376	0.167553191
Total	2810	2453	343	5606	
ErrorO	0.083274021	0.023236853	0.320699708		0.082768462

- 90% Intervalo de confianza = +/- 0.006287 (0.084590 - 0.097164)
- 95% Intervalo de confianza = +/- 0.007491 (0.083386 - 0.098368)
- 99% Intervalo de confianza = +/- 0.009861 (0.081016 - 0.100738)

Indice kappa de concordancia (KIA)

Usando el Maxlike	
Category	KIA
1	0.956977
2	0.820492
3	0.591282

Utilizando la segmentación	
Category	KIA
1	0.844195
2	0.909549
3	0.563883

Kappa promedio: 0.839146

ID 3.- 05 DE JUNIO DE 2013

	1	2	3	Total	ErrorC
1	3147	116	15	3278	0.039963
2	823	1372	4	2199	0.37608
3	11	9	786	806	0.024814
Total	3981	1497	805	6283	
Error0	0.209495	0.0835	0.023602		0.155658

- 90% Intervalo de confianza = +/- 0.007524 (0.148134 - 0.163182)
- 95% Intervalo de confianza = +/- 0.008964 (0.146694 - 0.164622)
- 99% Intervalo de confianza = +/- 0.011800 (0.143858 - 0.167458)

Indice kappa de concordancia (KIA)

Usando el Maxlike	
Category	KIA
1	0.890925
2	0.506287
3	0.97154

Utilizando la segmentación	
Category	KIA
1	0.561977
2	0.87154
3	0.972924

Kappa promedio: 0.726724

ID 4.- 28 DE NOVIEMBRE DE 2013

	1	2	3	Total	ErrorC
1	3269	75	48	3392	0.036262
2	561	1688	108	2357	0.283835
3	81	46	956	1083	0.117267
Total	3911	1809	1112	6832	
Error0	0.164152	0.066888	0.140288		0.134514

- 90% Intervalo de confianza = +/- 0.006791 (0.127723 - 0.141305)
- 95% Intervalo de confianza = +/- 0.008091 (0.126423 - 0.142605)
- 99% Intervalo de confianza = +/- 0.010650 (0.123864 - 0.145164)

Indice kappa de concordancia (KIA)

Usando el Maxlike	
Category	KIA
1	0.915186
2	0.613943
3	0.859936

Utilizando la segmentación	
Category	KIA
1	0.673986
2	0.897882
3	0.833285

Kappa promedio: 0.775299

ID 5.- 14 DE DICIEMBRE DE 2013

	1	2	3	Total	ErrorC
1	3647	92	14	3753	0.028244
2	709	1352	244	2305	0.413449
3	97	33	685	815	0.159509
Total	4453	1477	943	6873	
Error0	0.181002	0.084631	0.273595		0.172996

- 90% Intervalo de confianza = +/- 0.007505 (0.165491 - 0.180501)
- 95% Intervalo de confianza = +/- 0.008942 (0.164053 - 0.181938)
- 99% Intervalo de confianza = +/- 0.011771 (0.161225 - 0.184767)

Indice kappa de concordancia (KIA)

Usando el Maxlike	
Category	KIA
1	0.919785
2	0.473381
3	0.815125

Utilizando la segmentación	
Category	KIA
1	0.601274
2	0.872664
3	0.689598

Kappa promedio: 0.689902

ID 6.- 02 DE OCTUBRE DE 2013

	1	2	3	Total	ErrorC
1	2942	702	7	3651	0.194193
2	203	2253	55	2511	0.102748
3	15	216	383	614	0.376221
Total	3160	3171	445	6776	
Error0	0.068987	0.289499	0.139326		0.1768

- 90% Intervalo de confianza = +/- 0.007624 (0.169177 - 0.184424)
- 95% Intervalo de confianza = +/- 0.009084 (0.167717 - 0.185884)
- 99% Intervalo de confianza = +/- 0.011957 (0.164843 - 0.188758)

Indice kappa de concordancia (KIA)

Usando el Maxlike	
Category	KIA
1	0.636102
2	0.806874
3	0.597334

Utilizando la segmentación	
Category	KIA
1	0.850413
2	0.54006
3	0.846791

Kappa promedio: 0.689472

ID 7.- 31 DE ENERO DE 2014

	1	2	Total	ErrorC
1	2926	195	3121	0.06248
2	545	2509	3054	0.178454
Total	3471	2704	6175	
Error0	0.157015	0.072115		0.119838

- 90% Intervalo de confianza = +/- 0.006799 (0.113039 - 0.126637)
- 95% Intervalo de confianza = +/- 0.008101 (0.111737 - 0.127939)
- 99% Intervalo de confianza = +/- 0.010663 (0.109175 - 0.130501)

Indice kappa de concordancia (KIA)

Usando el Maxlike	
Category	KIA
1	0.857317
2	0.682525

Utilizando la segmentación	
Category	KIA
1	0.682525
2	0.857317

Kappa promedio: 0.76

ID 8.- 16 DE FEBRERO DE 2014

	1	2	Total	ErrorC
1	3027	264	3291	0.080219
2	427	3576	4003	0.10667
Total	3454	3840	7294	
Error0	0.123625	0.06875		0.094735

- 90% Intervalo de confianza = +/- 0.005641 (0.089095 - 0.100376)
- 95% Intervalo de confianza = +/- 0.006721 (0.088015 - 0.101456)
- 99% Intervalo de confianza = +/- 0.008847 (0.085889 - 0.103582)

Indice kappa de concordancia (KIA)

Usando el Maxlike	
Category	KIA
1	0.847626
2	0.774739

Utilizando la segmentación	
Category	KIA
1	0.774739
2	0.847626

Kappa promedio: 0.809545

ID 8.- 04 DE MARZO DE 2014

	1	2	3	Total	ErrorC
1	2502	409	18	2929	0.145784
2	568	3747	210	4525	0.171934
3	0	30	1416	1446	0.020747
Total	3070	4186	1644	8900	
Error0	0.185016	0.104873	0.138686		0.138764

- 90% Intervalo de confianza = +/- 0.006028 (0.132736 - 0.144792)
- 95% Intervalo de confianza = +/- 0.007182 (0.131582 - 0.145946)
- 99% Intervalo de confianza = +/- 0.009454 (0.129310 - 0.148218)

Indice kappa de concordancia (KIA)

Usando el Maxlike	
Category	KIA
1	0.777449
2	0.67539
3	0.974552

Utilizando la segmentación	
Category	KIA
1	0.724226
2	0.786658
3	0.83441

Kappa promedio: 0.775221

ID 10.- 26 DE JULIO DE 2014

	1	2	3	Total	ErrorC
1	2606	176	3	2785	0.064273
2	407	2523	21	2951	0.145036
3	0	94	403	497	0.189135
Total	3013	2793	427	6233	
Error0	0.135081	0.09667	0.056206		0.112466

- 90% Intervalo de confianza = +/- 0.006583 (0.105883 - 0.119049)
- 95% Intervalo de confianza = +/- 0.007844 (0.104622 - 0.120309)
- 99% Intervalo de confianza = +/- 0.010325 (0.102141 - 0.122791)

Indice kappa de concordancia (KIA)

Usando el Maxlike	
Category	KIA
1	0.875586
2	0.737207
3	0.796955

Utilizando la segmentación	
Category	KIA
1	0.755812
2	0.816409
3	0.938924

Kappa promedio: 0.801437

ID 11.- 27 DE AGOSTO DE 2014

	1	2	3	Total	ErrorC
1	3094	151	28	3273	0.05469
2	62	1605	196	1863	0.138486
3	33	78	552	663	0.167421
Total	3189	1834	776	5799	
Error0	0.02979	0.124864	0.28866		0.094499

- 90% Intervalo de confianza = +/- 0.006319 (0.088180 - 0.100818)
- 95% Intervalo de confianza = +/- 0.007529 (0.086970 - 0.102028)

- 99% Intervalo de confianza = +/- 0.009911 (0.084588 - 0.104410)

Indice kappa de concordancia (KIA)

Usando el Maxlike	
Category	KIA
1	0.878488
2	0.797457
3	0.806714

Utilizando la segmentación	
Category	KIA
1	0.931611
2	0.816035
3	0.674077

Kappa promedio: 0.834999

ID 12.- 01 DE DICIEMBRE DE 2014

	1	2	3	Total	ErrorC
1	2928	205	62	3195	0.083568
2	645	1996	40	2681	0.255502
3	9	19	541	569	0.049209
Total	3582	2220	643	6445	
Error0	0.18258	0.100901	0.158631		0.152056

- 90% Intervalo de confianza = +/- 0.007358 (0.144698 - 0.159414)
- 95% Intervalo de confianza = +/- 0.008767 (0.143289 - 0.160822)
- 99% Intervalo de confianza = +/- 0.011540 (0.140516 - 0.163596)

Indice kappa de concordancia (KIA)

Usando el Maxlike	
Category	KIA
1	0.811877
2	0.610247
3	0.945337

Utilizando la segmentación	
Category	KIA
1	0.637931
2	0.82723
3	0.826008

Kappa promedio: 0.734348

ID 13.- 22 DE ENERO DE 2014

	1	2	Total	ErrorC
1	2397	187	2584	0.072368
2	638	2342	2980	0.214094
Total	3035	2529	5564	
Error0	0.210214	0.073942		0.148275

- 90% Intervalo de confianza = +/- 0.007837 (0.140438 - 0.156112)
- 95% Intervalo de confianza = +/- 0.009338 (0.138937 - 0.157612)
- 99% Intervalo de confianza = +/- 0.012292 (0.135983 - 0.160566)

Indice kappa de concordancia (KIA)

Usando el Maxlike	
Category	KIA
1	0.840784
2	0.607506

Utilizando la segmentación	
Category	KIA
1	0.607506
2	0.840784

Kappa promedio: 0.705358

ID 14.- 07 DE FEBRERO DE 2014

	1	2	3	Total	ErrorC
1	1677	152	85	1914	0.123824
2	1135	1755	30	2920	0.398973
3	58	55	1004	1117	0.101164
Total	2870	1962	1119	5951	
Error0	0.415679	0.105505	0.10277		0.254579

- 90% Intervalo de confianza = +/- 0.009289 (0.245290 - 0.263868)
- 95% Intervalo de confianza = +/- 0.011068 (0.243511 - 0.265647)
- 99% Intervalo de confianza = +/- 0.014569 (0.240010 - 0.269148)

Indice kappa de concordancia (KIA)

Usando el Maxlike	
Category	KIA
1	0.760831
2	0.404792
3	0.875409

Utilizando la segmentación	
Category	KIA
1	0.387241
2	0.792855
3	0.873482

Kappa promedio: 0.607024

ID 15.- 11 DE MARZO DE 2014

	1	2	Total	ErrorC
1	2327	253	2580	0.098062
2	1364	2212	3576	0.381432
Total	3691	2465	6156	
Error0	0.369548	0.102637		0.262671

- 90% Intervalo de confianza = +/- 0.009227 (0.253444 - 0.271897)
- 95% Intervalo de confianza = +/- 0.010994 (0.251677 - 0.273664)
- 99% Intervalo de confianza = +/- 0.014471 (0.248199 - 0.277142)

Indice kappa de concordancia (KIA)

Usando el Maxlike	
Category	KIA
1	0.755104
2	0.363833

Utilizando la segmentación	
Category	KIA
1	0.363833
2	0.755104

Kappa promedio: 0.491058

ID 16.- 27 DE MARZO DE 2014

	1	2	3	Total	ErrorC
1	2156	102	52	2310	0.066667
2	621	2009	185	2815	0.286323
3	68	153	314	535	0.413084
Total	2845	2264	551	5660	
Error0	0.242179	0.112633	0.430127		0.208657

- 90% Intervalo de confianza = +/- 0.008885 (0.199772 - 0.217542)
- 95% Intervalo de confianza = +/- 0.010586 (0.198071 - 0.219244)
- 99% Intervalo de confianza = +/- 0.013935 (0.194722 - 0.222592)

Indice kappa de concordancia (KIA)

Usando el Maxlike	
Category	KIA
1	0.865956
2	0.522795
3	0.542365

Utilizando la segmentación	
Category	KIA
1	0.590825
2	0.775923
3	0.524972

Kappa promedio: 0.644362

ID 17.-28 DE ABRIL DE 2014

	1	2	3	Total	ErrorC
1	2184	162	35	2381	0.082738
2	833	2192	128	3153	0.304789
3	21	68	446	535	0.166355
Total	3038	2422	609	6069	
Error0	0.281106	0.094963	0.267652		0.20547

- 90% Intervalo de confianza = +/- 0.008532 (0.196939 - 0.214002)
- 95% Intervalo de confianza = +/- 0.010165 (0.195305 - 0.215636)
- 99% Intervalo de confianza = +/- 0.013381 (0.192089 - 0.218851)

Indice kappa de concordancia (KIA)

Usando el Maxlike	
Category	KIA
1	0.834332
2	0.492798
3	0.81509

Utilizando la segmentación	
Category	KIA
1	0.53741
2	0.802356
3	0.706473

Kappa promedio: 0.650225

ID 18.- 15 DE JUNIO DE 2014

	1	2	3	Total	ErrorC
1	2597	303	18	2918	0.110007
2	441	2536	21	2998	0.154103
3	32	8	936	976	0.040984
Total	3070	2847	975	6892	
Error0	0.154072	0.109238	0.04		0.119414

- 90% Intervalo de confianza = +/- 0.006425 (0.112988 - 0.125839)
- 95% Intervalo de confianza = +/- 0.007656 (0.111758 - 0.127070)
- 99% Intervalo de confianza = +/- 0.010078 (0.109336 - 0.129491)

Indice kappa de concordancia (KIA)

Usando el Maxlike	
Category	KIA
1	0.801631
2	0.737435
3	0.952263

Utilizando la segmentación	
Category	KIA
1	0.732798
2	0.80666
3	0.953401

Kappa promedio: 0.804777

ID 19.- 19 DE SEPTIEMBRE DE 2014

	1	2	3	Total	ErrorC
1	3717	375	6	4098	0.092972182
2	313	1304	108	1725	0.244057971
3	20	117	461	598	0.22909699
Total	4050	1796	575	6421	
Error0	0.082222222	0.273942094	0.19826087		0.146238904

- 90% Intervalo de confianza = +/- 0.007575 (0.160306 - 0.175456)
- 95% Intervalo de confianza = +/- 0.009026 (0.158855 - 0.176907)
- 99% Intervalo de confianza = +/- 0.011881 (0.156000 - 0.179762)

Indice kappa de concordancia (KIA)

Usando el Maxlike	
Category	KIA
1	0.758668
2	0.664471
3	0.741825

Utilizando la segmentación	
Category	KIA
1	0.782458
2	0.628885
3	0.583486

Kappa promedio: 0.686785

ID 20.- 19 DE FEBRERO DE 2015

	1	2	3	Total	ErrorC
1	2431	86	170	2687	0.095274
2	1184	1618	211	3013	0.462994
3	14	2	3762	3778	0.004235
Total	3629	1706	4143	9478	
Error0	0.330118	0.051583	0.091962		0.175881

- 90% Intervalo de confianza = +/- 0.006433 (0.169448 - 0.182314)
- 95% Intervalo de confianza = +/- 0.007665 (0.168216 - 0.183546)
- 99% Intervalo de confianza = +/- 0.010089 (0.165792 - 0.185970)

Indice kappa de concordancia (KIA)

Usando el Maxlike	
Category	KIA
1	0.845614
2	0.435376
3	0.992476

Utilizando la segmentación	
Category	KIA
1	0.539263
2	0.924377
3	0.847084

Kappa promedio: 0.733511

ID 21.- 08 DE ABRIL DE 2015

	1	2	3	Total	ErrorC
1	3524	202	37	3763	0.063513
2	371	1696	141	2208	0.231884
3	23	8	755	786	0.03944
Total	3918	1906	933	6757	
Error0	0.100562	0.110178	0.190782		0.115732

- 90% Intervalo de confianza = +/- 0.006402 (0.109330 - 0.122134)
- 95% Intervalo de confianza = +/- 0.007628 (0.108104 - 0.123360)
- 99% Intervalo de confianza = +/- 0.010041 (0.105691 - 0.125772)

Indice kappa de concordancia (KIA)

Usando el Maxlike	
Category	KIA
1	0.848835
2	0.677007
3	0.954242

Utilizando la segmentación	
Category	KIA
1	0.773048
2	0.836343
3	0.784104

Kappa promedio: 0.796550

ID 22.- 11 DE JUNIO DE 2015

	1	2	3	Total	ErrorC
1	3602	103	33	3738	0.036383
2	110	1676	187	1973	0.150532
3	18	41	603	662	0.089124
Total	3730	1820	823	6373	
Error0	0.034316	0.079121	0.267315		0.077201

- 90% Intervalo de confianza = +/- 0.005500 (0.071701 - 0.082701)
- 95% Intervalo de confianza = +/- 0.006553 (0.070648 - 0.083754)
- 99% Intervalo de confianza = +/- 0.008626 (0.068575 - 0.085827)

Indice kappa de concordancia (KIA)

Usando el Maxlike	
Category	KIA
1	0.91227
2	0.789295
3	0.89766

Utilizando la segmentación	
Category	KIA
1	0.917003
2	0.885401
3	0.701699

Kappa promedio: 0.860871

ID 23.- 13 DE JULIO DE 2015

	1	2	3	Total	ErrorC
1	2988	83	110	3181	0.060673
2	313	1331	240	1884	0.293524
3	4	2	3040	3046	0.00197
Total	3305	1416	3390	8111	
Error0	0.095915	0.060028	0.103245		0.092714

- 90% Intervalo de confianza = +/- 0.005298 (0.087416 - 0.098011)
- 95% Intervalo de confianza = +/- 0.006312 (0.086402 - 0.099026)
- 99% Intervalo de confianza = +/- 0.008309 (0.084405 - 0.101022)

Indice kappa de concordancia (KIA)

Usando el Maxlike	
Category	KIA
1	0.897604
2	0.644395
3	0.996616

Utilizando la segmentación	
Category	KIA
1	0.842197
2	0.92181
3	0.834666

Kappa promedio: 0.855741

ID 24.- 29 DE JULIO DE 2015

	1	2	3	Total	ErrorC
1	3835	442	66	4343	0.11697
2	215	5320	147	5682	0.06371
3	15	236	468	719	0.349096
Total	4065	5998	681	10744	
Error0	0.056581	0.113038	0.312775		0.104337

- 90% Intervalo de confianza = +/- 0.004851 (0.099486 - 0.109189)
- 95% Intervalo de confianza = +/- 0.005780 (0.098557 - 0.110118)
- 99% Intervalo de confianza = +/- 0.007609 (0.096728 - 0.111946)

Indice kappa de concordancia (KIA)

Usando el Maxlike	
Category	KIA
1	0.81184
2	0.855773
3	0.627279

Utilizando la segmentación	
Category	KIA
1	0.90503
2	0.76008
3	0.664792

Kappa promedio: 0.809457

ID 25.- 14 DE AGOSTO DE 2015

	1	2	3	Total	ErrorC
1	3329	195	27	3551	0.062518
2	809	1235	260	2304	0.463976
3	10	42	782	834	0.06235
Total	4148	1472	1069	6689	
Error0	0.197445	0.161005	0.268475		0.200777

- 90% Intervalo de confianza = +/- 0.008057 (0.192720 - 0.208834)
- 95% Intervalo de confianza = +/- 0.009600 (0.191177 - 0.210377)
- 99% Intervalo de confianza = +/- 0.012637 (0.188141 - 0.213414)

Indice kappa de concordancia (KIA)

Usando el Maxlike	
Category	KIA
1	0.835427
2	0.405111
3	0.92579

Utilizando la segmentación	
Category	KIA
1	0.579125
2	0.754398
3	0.693283

Kappa promedio: 0.650864

ID 26.- 18 DE NOVIEMBRE DE 2015

	1	2	3	Total	ErrorC
1	3695	158	36	3889	0.049884
2	658	1769	10	2437	0.274108
3	106	193	572	871	0.343284
Total	4459	2120	618	7197	
Error0	0.171339	0.165566	0.074434		0.161317

- 90% Intervalo de confianza = +/- 0.007132 (0.154185 - 0.168450)
- 95% Intervalo de confianza = +/- 0.008498 (0.152819 - 0.169815)
- 99% Intervalo de confianza = +/- 0.011186 (0.150131 - 0.172503)

Indice kappa de concordancia (KIA)

Usando el Maxlike	
Category	KIA
1	0.868876
2	0.611434
3	0.62447

Utilizando la segmentación	
Category	KIA
1	0.627229
2	0.749668
3	0.915318

Kappa promedio: 0.709377

ID 27.- 20 DE DICIEMBRE DE 2015

	1	2	3	Total	ErrorC
1	3306	230	30	3566	0.072911
2	1215	1606	0	2821	0.430698
3	33	2	651	686	0.05102
Total	4554	1838	681	7073	
Error0	0.274045	0.126224	0.044053		0.213488

- 90% Intervalo de confianza = +/- 0.008015 (0.205473 - 0.221503)
- 95% Intervalo de confianza = +/- 0.009550 (0.203938 - 0.223038)
- 99% Intervalo de confianza = +/- 0.012571 (0.200917 - 0.226059)

Indice kappa de concordancia (KIA)

Usando el Maxlike	
Category	KIA
1	0.795277
2	0.418084
3	0.943544

Utilizando la segmentación	
Category	KIA
1	0.4473
2	0.790032
3	0.951216

Kappa promedio: 0.620402

ID 28.- 26 DE FEBRERO DE 2015

	1	2	3	Total	ErrorC
1	2540	105	15	2660	0.045113
2	1255	1635	52	2942	0.444256
3	33	254	487	774	0.370801
Total	3828	1994	554	6376	
Error0	0.336468	0.18004	0.120939		0.268821

- 90% Intervalo de confianza = +/- 0.009133 (0.259687 - 0.277954)
- 95% Intervalo de confianza = +/- 0.010882 (0.257938 - 0.279703)
- 99% Intervalo de confianza = +/- 0.014325 (0.254496 - 0.283145)

Indice kappa de concordancia (KIA)

Usando el Maxlike	
Category	KIA
1	0.887112
2	0.353589
3	0.593915

Utilizando la segmentación	
Category	KIA
1	0.42268
2	0.665715
3	0.862352

Kappa promedio: 0.547958

ID 29.- 25 DE NOVIEMBRE DE 2015

	1	2	3	Total	ErrorC
1	3793	565	3	4361	0.130245357
2	465	1432	2	1899	0.245918905
3	101	1	758	860	0.118604651
Total	4359	1998	763	7120	
Error0	0.1298463	0.283283283	0.00655308		0.159691011

- 90% Intervalo de confianza = +/- 0.007329 (0.166747 - 0.181404)
- 95% Intervalo de confianza = +/- 0.008732 (0.165343 - 0.182807)
- 99% Intervalo de confianza = +/- 0.011494 (0.162581 - 0.185569)

Indice kappa de concordancia (KIA)

Usando el Maxlike	
Category	KIA
1	0.666962
2	0.66042
3	0.865943

Utilizando la segmentación	
Category	KIA
1	0.647966
2	0.61607
3	0.895362

Kappa promedio: 0.682016

ID 30.- 27 DE DICIEMBRE DE 2015

	1	2	3	Total	ErrorC
1	3293	175	47	3515	0.063158
2	1022	1727	324	3073	0.438008
3	2	45	752	799	0.058824
Total	4317	1947	1123	7387	
Error0	0.237202	0.112994	0.330365		0.218627

- 90% Intervalo de confianza = +/- 0.007911 (0.210717 - 0.226538)
- 95% Intervalo de confianza = +/- 0.009425 (0.209202 - 0.228053)
- 99% Intervalo de confianza = +/- 0.012407 (0.206220 - 0.231034)

Indice kappa de concordancia (KIA)

Usando el Maxlike	
Category	KIA
1	0.84803
2	0.405226
3	0.930631

Utilizando la segmentación	
Category	KIA
1	0.547467
2	0.806516
3	0.629568

Kappa promedio: 0.633071

ID 31.- 21 DE ENERO DE 2016

	1	2	Total	ErrorC
1	4978	834	5812	0.143496
2	225	4683	4908	0.045844
Total	5203	5517	10720	
Error0	0.043244	0.151169		0.098787

- 90% Intervalo de confianza = +/- 0.004741 (0.094047 - 0.103528)
- 95% Intervalo de confianza = +/- 0.005648 (0.093139 - 0.104436)
- 99% Intervalo de confianza = +/- 0.007435 (0.091352 - 0.106222)

Índice kappa de concordancia (KIA)

Usando el Maxlike	
Category	KIA
1	0.721175
2	0.905546

Utilizando la segmentación	
Category	KIA
1	0.905546
2	0.721175

Kappa promedio: 0.802912

ID 32.- 25 DE MARZO DE 2016

	1	2	3	Total	ErrorC
1	3554	118	17	3689	0.036595
2	775	1408	92	2275	0.381099
3	4	20	746	770	0.031169
Total	4333	1546	855	6734	
Error0	0.179783	0.089263	0.127485		0.152361

- 90% Intervalo de confianza = +/- 0.007204 (0.145157 - 0.159565)
- 95% Intervalo de confianza = +/- 0.008583 (0.143778 - 0.160945)
- 99% Intervalo de confianza = +/- 0.011299 (0.141063 - 0.163660)

Índice kappa de concordancia (KIA)

Usando el Maxlike	
Category	KIA
1	0.897362
2	0.505335
3	0.964298

Utilizando la segmentación	
Category	KIA
1	0.602411
2	0.865195
3	0.856055

Kappa promedio: 0.725687

ID 33.- 28 DE MAYO DE 2016

	1	2	3	Total	ErrorC
1	3527	50	9	3586	0.016453
2	780	1293	13	2086	0.380153
3	103	3	466	572	0.185315
Total	4410	1346	488	6244	
Error0	0.200227	0.039376	0.045082		0.153427

- 90% Intervalo de confianza = +/- 0.007503 (0.145925 - 0.160930)
- 95% Intervalo de confianza = +/- 0.008939 (0.144488 - 0.162367)
- 99% Intervalo de confianza = +/- 0.011767 (0.141660 - 0.165194)

Índice kappa de concordancia (KIA)

Usando el Maxlike	
Category	KIA
1	0.943985
2	0.515378
3	0.798974

Utilizando la segmentación	
Category	KIA
1	0.52964
2	0.94087
3	0.950372

Kappa promedio: 0.702199

ID 34.- 22 DE DICIEMBRE DE 2016

	1	2	3	Total	ErrorC
1	4320	201	24	4545	0.049505
2	250	2306	195	2751	0.161759
3	55	17	480	552	0.130435
Total	4625	2524	699	7848	
Error0	0.065946	0.086371	0.313305		0.094546

- 90% Intervalo de confianza = +/- 0.005433 (0.089113 - 0.099979)
- 95% Intervalo de confianza = +/- 0.006473 (0.088073 - 0.101020)
- 99% Intervalo de confianza = +/- 0.008521 (0.086025 - 0.103067)

Índice kappa de concordancia (KIA)

Usando el Maxlike	
Category	KIA
1	0.879456
2	0.761554
3	0.856812

Utilizando la segmentación	
Category	KIA
1	0.843311
2	0.867012
3	0.662991

Kappa promedio: 0.824819

ID 35.- 16 DE MARZO DE 2016

	1	2	3	Total	ErrorC
1	3194	229	77	3500	0.087429
2	908	1916	50	2874	0.333333
3	0	16	1695	1711	0.009351
Total	4102	2161	1822	8085	
Error0	0.221355	0.113373	0.069704		0.158318

- 90% Intervalo de confianza = +/- 0.006678 (0.151640 - 0.164996)
- 95% Intervalo de confianza = +/- 0.007957 (0.150361 - 0.166275)
- 99% Intervalo de confianza = +/- 0.010474 (0.147844 - 0.168792)

Índice kappa de concordancia (KIA)

Usando el Maxlike	
Category	KIA
1	0.822531
2	0.545071
3	0.987928

Utilizando la segmentación	
Category	KIA
1	0.609671
2	0.824098
3	0.911586

Kappa promedio: 0.751721

ID 36.- 13 DE DICIEMBRE DE 2016

	1	2	3	Total	ErrorC
1	3999	470	0	4469	0.105168942
2	1066	2044	0	3110	0.342765273
3	0	0	0	0	0
Total	5065	2514	0	7579	
ErrorO	0.210463968	0.186953063	0		0.202665259

- 90% Intervalo de confianza = +/- 0.007631 (0.198908 - 0.214170)
- 95% Intervalo de confianza = +/- 0.009092 (0.197447 - 0.215631)
- 99% Intervalo de confianza = +/- 0.011968 (0.194571 - 0.218507)

Índice kappa de concordancia (KIA)

Usando el Maxlike	
Category	KIA
1	0.681398
2	0.488338

Utilizando la segmentación	
Category	KIA
1	0.476803
2	0.684014

Kappa promedio: 0.562521

Anexo 3. Resultados de la matriz de confusión por imagen clasificada para áreas con vegetación natural y regulada.

ID 1.- 18 DE ABRIL DE 2013

	1	2	Total	ErrorC
1	1068	1113	2181	0.510316
2	0	634	634	0
Total	1068	1747	2815	
ErrorO	0	0.637092		0.395382

- ErrorO = Errores de omisión (expresado en proporciones)
- ErrorC = Errores de Comision (expresado en proporciones)
- 90% Intervalo de confianza = +/- 0.015159 (0.380223 - 0.410541)
- 95% Intervalo de confianza = +/- 0.018062 (0.377320 - 0.413444)
- 99% Intervalo de confianza = +/- 0.023775 (0.371606 - 0.419157)

Índice kappa de concordancia (KIA)

Usando el maxlike	
Category	KIA
1	0.17771
2	1

Utilizando la segmentación	
Category	KIA
1	1
2	0.17771

Kappa promedio: 0.301789

ID 4.- 28 DE NOVIEMBRE DE 2013

	1	2	Total	ErrorC
1	1875	677	2552	0.265282
2	350	518	868	0.403226
Total	2225	1195	3420	
Error0	0.157303	0.566527		0.300292

- 90% Intervalo de confianza = +/- 0.012894 (0.287399 - 0.313186)
- 95% Intervalo de confianza = +/- 0.015363 (0.284929 - 0.315655)
- 99% Intervalo de confianza = +/- 0.020223 (0.280070 - 0.320515)

Índice kappa de concordancia (KIA)

Usando el maxlike	
Category	KIA
1	0.240783
2	0.38021

Utilizando la segmentación	
Category	KIA
1	0.38021
2	0.240783

Kappa promedio: 0.294844

ID 6.- 02 DE OCTUBRE DE 2013

	1	2	Total	ErrorC
1	955	531	1486	0.357335128
2	473	734	1207	0.391880696
Total	1428	1265	2693	
Error0	0.331232493	0.419762846		0.372818418

- 90% Intervalo de confianza = +/- 0.014402 (0.467022 - 0.495827)
- 95% Intervalo de confianza = +/- 0.017160 (0.464265 - 0.498585)
- 99% Intervalo de confianza = +/- 0.022588 (0.458836 - 0.504013)

Índice kappa de concordancia (KIA)

Usando el maxlike	
Category	KIA
0	0
1	0.363674
2	0.106194

Utilizando la segmentación	
Category	KIA
1	0.390839
2	0.048816

Kappa promedio: 0.186585

ID 8.- 16 DE FEBRERO DE 2014

	1	2	Total	ErrorC
1	939	691	1630	0.423926
2	183	1131	1314	0.139269
Total	1122	1822	2944	
Error0	0.163102	0.379254		0.296875

- 90% Intervalo de confianza = +/- 0.013852 (0.283023 - 0.310727)
- 95% Intervalo de confianza = +/- 0.016504 (0.280371 - 0.313379)

- 99% Intervalo de confianza = +/- 0.021725 (0.275150 - 0.318600)

Índice kappa de concordancia (KIA)

Usando el maxlike	
Category	KIA
1	0.315017
2	0.634573

Utilizando la segmentación	
Category	KIA
1	0.634573
2	0.315017

Kappa promedio: 0.421026

ID 11.- 27 DE AGOSTO DE 2014

	1	2	Total	ErrorC
1	765	1594	2359	0.67571
2	154	351	505	0.30495
Total	919	1945	2864	
Error0	0.167573	0.819537		0.610335

- 90% Intervalo de confianza = +/- 0.014990 (0.595345 - 0.625325)
- 95% Intervalo de confianza = +/- 0.017861 (0.592474 - 0.628196)
- 99% Intervalo de confianza = +/- 0.023511 (0.586825 - 0.633846)

Índice kappa de concordancia (KIA)

Usando el maxlike	
Category	KIA
1	0.005021
2	0.049643

Utilizando la segmentación	
Category	KIA
1	0.049643
2	0.005021

Kappa promedio: 0.009120

ID 15.- 11 DE MARZO DE 2014

	1	2	Total	ErrorC
1	533	69	602	0.114618
2	66	1596	1662	0.039711
Total	599	1665	2264	
Error0	0.110184	0.041441		0.059629

- 90% Intervalo de confianza = +/- 0.008187 (0.051442 - 0.067816)
- 95% Intervalo de confianza = +/- 0.009754 (0.049875 - 0.069383)
- 99% Intervalo de confianza = +/- 0.012840 (0.046789 - 0.072469)

Índice kappa de concordancia (KIA)

Usando el maxlike	
Category	KIA
1	0.844147
2	0.849906

Utilizando la segmentación	
Category	KIA
1	0.849906
2	0.844147

Kappa promedio: 0.847017

ID 19.- 19 DE SEPTIEMBRE DE 2014

	1	2	Total	ErrorC
1	1096	348	1444	0.240997
2	363	1745	2108	0.172201
Total	1459	2093	3552	
Error0	0.248801	0.166269		0.200169

- 90% Intervalo de confianza = +/- 0.011044 (0.189125 - 0.211213)
- 95% Intervalo de confianza = +/- 0.013159 (0.187010 - 0.213328)
- 99% Intervalo de confianza = +/- 0.017321 (0.182848 - 0.217490)

Índice kappa de concordancia (KIA)

Usando el maxlike	
Category	KIA
1	0.591007
2	0.580769

Utilizando la segmentación	
Category	KIA
1	0.580769
2	0.591007

Kappa promedio: 0.585843

ID 20.- 19 DE FEBRERO DE 2015

	1	2	Total	ErrorC
1	1116	453	1569	0.288719
2	329	731	1060	0.310377
Total	1445	1184	2629	
Error0	0.227682	0.382601		0.297452

- 90% Intervalo de confianza = +/- 0.014666 (0.282785 - 0.312118)
- 95% Intervalo de confianza = +/- 0.017475 (0.279977 - 0.314926)
- 99% Intervalo de confianza = +/- 0.023002 (0.274449 - 0.320454)

Índice kappa de concordancia (KIA)

Usando el maxlike	
Category	KIA
1	0.358917
2	0.435307

Utilizando la segmentación	
Category	KIA
1	0.435307
2	0.358917

Kappa promedio: 0.393438

ID 25.- 14 DE AGOSTO DE 2015

	1	2	Total	ErrorC
1	2289	1020	3309	0.30825
2	0	0	0	
Total	2289	1020	3309	
Error0	0	1		0.30825

- 90% Intervalo de confianza = +/- 0.013205 (0.295045 - 0.321455)
- 95% Intervalo de confianza = +/- 0.015734 (0.292516 - 0.323984)
- 99% Intervalo de confianza = +/- 0.020711 (0.287539 - 0.328961)

Índice kappa de concordancia (KIA)

Usando el maxlike	
Category	KIA
1	0
2	0

Utilizando la segmentación	
Category	KIA
1	0
2	0

Kappa promedio: 0

ID 28.- 26 DE FEBRERO DE 2015

	1	2	Total	ErrorC
1	988	220	1208	0.182119
2	28	1149	1177	0.023789
Total	1016	1369	2385	
Error0	0.027559	0.160701		0.103983

- 90% Intervalo de confianza = +/- 0.010282 (0.093702 - 0.114265)
- 95% Intervalo de confianza = +/- 0.012250 (0.091733 - 0.116234)
- 99% Intervalo de confianza = +/- 0.016126 (0.087858 - 0.120109)

Índice kappa de concordancia (KIA)

Usando el maxlike	
Category	KIA
1	0.682721
2	0.944156

Utilizando la segmentación	
Category	KIA
1	0.944156
2	0.682721

Kappa promedio: 0.792433

ID 29.- 25 DE NOVIEMBRE DE 2015

	1	2	Total	ErrorC
1	2669	990	3659	0.270566
2	13	81	94	0.138298
Total	2682	1071	3753	
Error0	0.004847	0.92437		0.267253

- 90% Intervalo de confianza = +/- 0.011883 (0.255370 - 0.279136)

- 95% Intervalo de confianza = +/- 0.014158 (0.253095 - 0.281411)
- 99% Intervalo de confianza = +/- 0.018637 (0.248616 - 0.285890)

Índice kappa de concordancia (KIA)

Usando el maxlike	
Category	KIA
1	0.051883
2	0.806476

Utilizando la segmentación	
Category	KIA
1	0.806476
2	0.051883

Kappa promedio: 0.097494

ID 32.- 25 DE MARZO DE 2016

	1	2	Total	ErrorC
1	2222	1299	3521	0.368929
2	0	0	0	
Total	2222	1299	3521	
Error0	0	1		0.368929

- 90% Intervalo de confianza = +/- 0.013377 (0.355553 - 0.382306)
- 95% Intervalo de confianza = +/- 0.015938 (0.352991 - 0.384867)
- 99% Intervalo de confianza = +/- 0.020980 (0.347950 - 0.389909)

Índice kappa de concordancia (KIA)

Usando el maxlike	
Category	KIA
1	0
2	0

Utilizando la segmentación	
Category	KIA
1	0
2	0

Kappa promedio: 0

ID 33.- 28 DE MAYO DE 2016

	1	2	Total	ErrorC
1	1385	1344	2729	0.492488
2	0	329	329	0
Total	1385	1673	3058	
Error0	0	0.803347		0.439503

- 90% Intervalo de confianza = +/- 0.014764 (0.424739 - 0.454267)
- 95% Intervalo de confianza = +/- 0.017592 (0.421911 - 0.457095)
- 99% Intervalo de confianza = +/- 0.023156 (0.416347 - 0.462659)

Índice kappa de concordancia (KIA)

Usando el maxlike	
Category	KIA
1	0.099804
2	1

Utilizando la segmentación	
Category	KIA
1	1
2	0.099804

Kappa promedio: 0.181493

ID 35.- 16 DE MARZO DE 2016

	1	2	Total	ErrorC
1	2030	879	2909	0.302166
2	53	851	904	0.058628
Total	2083	1730	3813	
Error0	0.025444	0.508092		0.244427

- 90% Intervalo de confianza = +/- 0.011448 (0.232979 - 0.255875)
- 95% Intervalo de confianza = +/- 0.013641 (0.230786 - 0.258068)
- 99% Intervalo de confianza = +/- 0.017956 (0.226471 - 0.262383)

Índice kappa de concordancia (KIA)

Usando el maxlike	
Category	KIA
1	0.334013
2	0.892679

Utilizando la segmentación	
Category	KIA
1	0.892679
2	0.334013

Kappa promedio: 0.486131

ID 36.- 13 DE DICIEMBRE DE 2016

	1	2	Total	ErrorC
1	2301	666	2967	0.224469
2	141	253	394	0.357868
Total	2442	919	3361	
Error0	0.05774	0.724701		0.240107

- 90% Intervalo de confianza = +/- 0.012120 (0.227987 - 0.252227)
- 95% Intervalo de confianza = +/- 0.014441 (0.225666 - 0.254548)
- 99% Intervalo de confianza = +/- 0.019009 (0.221098 - 0.259116)

Índice kappa de concordancia (KIA)

Usando el maxlike	
Category	KIA
1	0.179063
2	0.507455

Utilizando la segmentación	
Category	KIA
1	0.507455
2	0.179063

Kappa promedio: 0.264717