



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**

FACULTAD DE FILOSOFÍA Y LETRAS  
DIVISIÓN DE ESTUDIOS DE POSGRADO

**EVALUACIÓN DE IMÁGENES DE BAJA RESOLUCIÓN ESPACIAL BAJO  
DIFERENTES METODOLOGÍAS PARA EL MONITOREO Y CARTOGRAFÍA DE  
LA COBERTURA DEL SUELO EN MÉXICO**

**T E S I S**

QUE PARA OBTENER EL GRADO ACADÉMICO DE  
DOCTOR EN GEOGRAFÍA

**P R E S E N T A**

**TZITZIKI JANIK GARCÍA MORA**

**DIRECTOR DE TESIS: JEAN FRANÇOIS MAS CAUSSEL**



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

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

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

# DEDICATORIA

A mis dos pilares

Eduardo Kaleb y Everett Alan

Sin ustedes no lo hubiera podido lograr,

...y si hubiera podido no habría valido la pena.

## AGRADECIMIENTOS

Al Centro de Investigaciones en Geografía Ambiental (CIGA) UNAM, que ha sido mi casa los últimos seis años y en donde he aprendido mucho de sus integrantes.

A mi asesor de Tesis, el Dr. Jean François Mas, a quien admiro y respeto profundamente. Agradezco su paciencia, enseñanzas, confianza, apoyo incondicional, su valioso tiempo, y gran profesionalismo.

A los integrantes de la mesa sinodal, Dr. Gerardo Bocco, Dr. Eleazar Carranza, Dr. Juan José Flores y Dr. Raúl Aguirre, que por medio de su punto de vista multidisciplinario y objetivo permitieron enriquecer sustancialmente este trabajo de tesis.

Al MSc. Antonio Navarrete y al MTI. Hugo Alejandro Zavala por su importante colaboración técnica en distintas partes de este proyecto. A la Dra. Yan Gao por facilitar los datos de campo del área de Tancítaro.

A la administración del CIGA, Dr. Manuel Bollo y Lic. Katia Méndez quienes manifestando siempre un gran interés y disposición, me brindaron apoyo total para llegar a la meta final. Gracias por su valiosa amistad. Al igual agradezco a la administración del Posgrado en Geografía: Dr. Javier Delgado Campos, Mtro. Macario Arredondo, Penélope Márquez y Sra. Vicky, por su disposición, amabilidad, organización y la gran labor que realizan para mantener en marcha este posgrado.

A Patrick Coronado del Centro Goddard NASA, por la oportunidad que me dio de conocer y aprender de la comunidad internacional MODIS y por sus palabras de aliento. A Brad Quayle, Bryan Schwind y Erick Neilsen del Servicio Forestal de Estados Unidos por compartir sus experiencias con MODIS, de igual manera agradezco al Dr. René Colditz y Geog. Isabel Cruz de CONABIO por sus importantes sugerencias en el segundo capítulo del presente manuscrito.

Al Director Nacional de Percepción Remota del Servicio Forestal de los Estados Unidos: Everett Alan Hinkley por sus recomendaciones, observaciones, ediciones, apoyo en campo y en la mejora de la calidad de las gráficas y cuadros de esta tesis, por su ejemplo de amor y compromiso al trabajo, su apoyo incondicional y por su confianza en mí. Gracias por compartir tu vida conmigo.

Al apoyo otorgado por CONACyT en la convocatoria 2007, que me permitió llevar a cabo mis estudios de doctorado. A la DGEP que me otorgó la beca de apoyo a la titulación para finalizar el trabajo de tesis.

Al Instituto Nacional de Estadística, Geografía e Informática (INEGI) por las licencias de uso para el continuo Nacional de Uso de Suelo y Vegetación serie III, escala 1:250,000 y los prototipos de los mapas de uso del suelo y vegetación, 1:50,000.

Al apoyo incondicional de mi familia Daniel, Nely, Leo y Erick, quienes han depositado toda su confianza en mí y en mis proyectos dándome todas las facilidades para lograr mis objetivos profesionales. A mi familia Hinkley, especialmente a Lawrence, Evelyn, Viv, Jan, así como a David y Marjorie Whipple, por su cariño y apoyo moral.

---

Este estudio se llevó a cabo en el ámbito del proyecto CONAFOR 2005-C02-14741 denominado "Evaluación del Sensor MODIS para el monitoreo Anual de la Vegetación Forestal de México" del Fondo Sectorial para la Investigación, el Desarrollo y la Innovación Tecnológica Forestal (CONAFOR-CONACYT).

## **RESUMEN**

En las últimas décadas, la agricultura, el pastoreo y la explotación forestal se han extendido considerablemente, afectando la sostenibilidad global y la biodiversidad. Para comprender estos cambios y gestionar con eficacia sus impactos es esencial monitorear el uso y cobertura del suelo mediante sensores de observación de la tierra con una alta frecuencia temporal y extensión espacial. Los datos del sensor MODIS proporcionan una opción atractiva, sin que hasta el momento se cuente con datos precisos de su fiabilidad a nivel regional. Este trabajo pretendió contribuir con un mejor entendimiento sobre los datos y metodologías adecuadas para monitorear el uso y cobertura del suelo, así como sobre la fiabilidad de su clasificación. Este conocimiento puede ser aplicado en la gestión y protección del medio ambiente. El análisis incluyó tres tipos de datos del sensor MODIS (índices de vegetación, compuestos de 8 días e imágenes diarias) durante un año para tres áreas de estudio ubicadas en los estados de Michoacán, Sonora y Quintana Roo. Se aplicaron dos técnicas de clasificación a cada conjunto de datos; se incorporaron datos auxiliares, y se evaluó la fiabilidad. El método de clasificación de Máxima Verosimilitud mostró resultados más consistentes que el método Perceptrón Multicapa. En la mayoría de los casos la incorporación de datos aumentó la fiabilidad de las clasificaciones. Los resultados permiten concluir que los mapas generados a partir de sensores de baja resolución muestran un gran potencial en el monitoreo de la cobertura terrestre a escala regional y nacional.

## **ABSTRACT**

In recent decades, agriculture, animal grazing, and forest harvesting have expanded greatly, impacting global sustainability and biodiversity. To understand these changes and manage their impacts effectively it is essential to monitor land use and land cover change closely, using earth observing sensors to obtain needed coverage with high temporal frequency, and large areal extent. The imagery derived from the MODIS sensor provides an attractive option, though there hasn't been sufficient analysis of its accuracy at regional levels for land cover mapping. This thesis work was carried out to provide progress toward a solution; specifically to contribute a better understanding of the appropriate data products and methodologies that are suitable for monitoring land use and land cover, and to gain a better understanding of their classification accuracies. This knowledge can then be applied toward the management and protection of the environment. The analysis included three types of data from the MODIS sensor spanning one year for three study areas located in the states of Michoacán, Sonora and Quintana Roo, Mexico. Two classification techniques were applied to each of these data sets; ancillary data were incorporated, and the accuracies were evaluated. The method of Maximum Likelihood classification showed more consistent results compared to the Multilayer Perceptron method. The accuracies were improved with the incorporation of ancillary data in most cases. Taken together with results based on other data analyses, we conclude that the maps generated from low-resolution sensors show great potential for land cover mapping at regional and national scales.

# ÍNDICE

<b>CAPÍTULO I. INTRODUCCIÓN .....</b>	<b>1</b>
1.1. Justificación del trabajo.....	1
1.2. Objetivos e hipótesis de trabajo.....	3
1.2.1.    Objetivo general .....	3
1.2.2.    Objetivos específicos .....	4
1.2.3.    Hipótesis de trabajo .....	4
1.3. Estructura de la tesis.....	4
<b>CAPÍTULO II. MARCO TEÓRICO .....</b>	<b>5</b>
2.1. Los datos de satélite en el monitoreo de La cobertura y el uso del suelo .....	5
2.1.1.    Desarrollo de los datos de sensores remotos .....	5
2.2. Especificaciones técnicas del sensor MODIS .....	9
2.2.1.    Deformaciones relacionadas con el barrido (ángulo de vista).....	12
2.2.2.    Transmisión y pre-procesamiento de los datos.....	13
2.2.3.    Niveles de procesamiento.....	14
2.3. Descripción de los Productos MODLAND.....	15
2.3.1.    Producción de los datos MODLAND.....	16
2.3.2.    Características Generales.....	18
1)    Las colecciones de datos MODIS.....	18
2)    Organización espacial.....	19
3)    Formato de Almacenamiento HDF.....	20
4)    Identificación de los productos.....	20
5)    Metadatos .....	21
2.3.3.    Descripción de los productos utilizados en esta tesis.....	22
1)    Reflectancia de la superficie MOD09 .....	22
2)    Índices de vegetación MOD13 .....	24

<b>CAPÍTULO III. ANTECEDENTES.....</b>	<b>28</b>
3.1. Selección de los datos.....	28
3.1.1. Estudios basados en índices de vegetación.....	28
3.1.2. Estudios basados en datos de reflectancia de la superficie.....	30
3.1.3. El uso de combinaciones de datos MODIS .....	30
3.1.4. Tres dimensiones de datos .....	31
3.1.5. Incorporación de datos auxiliares .....	31
3.2. Preprocesamiento de los datos .....	33
3.2.1. Métodos de selección de los datos.....	34
3.2.2. Método de extracción de datos .....	35
3.3. Filtrado y análisis de series de tiempo.....	36
3.3.1. Análisis de Fourier .....	37
3.4. Métodos de clasificación mas utilizados .....	37
3.5. Medición de la fiabilidad.....	38
3.6. Resultados de los estudios revisados.....	40
<b>CAPÍTULO IV. DESCRIPCIÓN DE LAS ÁREAS DE ESTUDIO .....</b>	<b>44</b>
4.1. Criterios de selección de las areas de estudio.....	44
4.2. Descripción de las áreas de estudio .....	46
4.2.1. Área del Tancítaro, Michoacán .....	46
4.2.2. Área de Mazatán, Sonora.....	47
4.2.3. Área Maya, Quintana Roo .....	47
<b>CAPÍTULO V. BASES DE DATOS Y MÉTODOS.....</b>	<b>49</b>
5.1. BASE DE DATOS .....	49
5.1.1. Datos MODIS .....	49
5.1.2. Datos de referencia .....	50
5.2. Metodología.....	51
5.2.1. Selección de los datos.....	51
5.2.2. Descarga, corte, reproyección y remuestreo de los datos.....	51
5.2.3. Análisis de la calidad de los datos .....	52
5.2.4. Análisis visual .....	53

5.2.5.	Selección del sistema clasificatorio .....	53
5.2.6.	Campos de entrenamiento .....	54
5.2.7.	Cálculo de la separabilidad.....	54
5.2.8.	Clasificación .....	55
5.2.9.	Clasificación con información auxiliar.....	59
5.2.10.	Evaluación de la fiabilidad .....	60
5.2.11.	Comparación de la fiabilidad obtenida de los diferentes métodos de clasificación - Prueba de McNemar .....	63
<b>CAPÍTULO VI. RESULTADOS .....</b>		<b>65</b>
6.1.	Análisis de la calidad de los datos.....	65
6.1.1.	Medición de la separabilidad (cálculo de la divergencia transformada) .....	66
6.1.2.	Selección espectral y por fechas .....	68
6.2.	Clasificaciones.....	69
6.2.1.	Máxima Verosimilitud (MV) y Perceptrón Multicapa (PMC).....	69
6.3.2	Datos auxiliares .....	71
6.3.3	Comparación entre los resultados de las tres áreas de estudio .....	73
<b>CAPÍTULO VII. DISCUSIÓN Y CONCLUSIONES .....</b>		<b>75</b>
7.1.	Discusión.....	75
7.2	Conclusiones .....	81
<b>CAPÍTULO VIII. LITERATURA CITADA .....</b>		<b>83</b>
<b>ANEXOS .....</b>		<b>92</b>

## ÍNDICE DE CUADROS Y FIGURAS

Cuadro 1 - Características generales de los sensores más utilizados .....	8
Cuadro 2 - Características de las bandas MODIS .....	11
Cuadro 3 - Referencias claves para el producto MOD09.....	22
Cuadro 4 - Productos de reflectancia MOD09.....	24
Cuadro 5 - Referencias claves para el producto MOD13.....	25
Cuadro 6 - Productos de índice de vegetación (Incluye EVI y NDVI).....	27
Cuadro 7 - Los diferentes esfuerzos de mapeo y datos de entrada revisados en esta tesis.....	32
Cuadro 8 - Tipos de ambiente, número de categorías y fiabilidades obtenidas en estudios con datos MODIS .....	41
Cuadro 9 - Bits seleccionados a partir de la capa de calidad MODIS (QA) para cada tipo de dato ....	53
Cuadro 10 - Definición cualitativa propuesta para la interpretación de los valores de separabilidad (divergencia transformada).....	55
Cuadro 11 - Método de interpretación de datos de campo .....	61
Cuadro 12 - Matriz de confusión.....	61
Cuadro 13 - Elementos de la matriz utilizada en la prueba de McNemar .....	63
Cuadro 14 - Reducción de datos a través de la capa de calidad .....	65
Cuadro 15 - Selección espectral y por fechas.....	68
Figura 1 - Esquema del cubrimiento MODIS (Tomado de MODIS Rapid Response Team.) <a href="http://earthobservatory.nasa.gov/Features/OrbitsManeuver/">http://earthobservatory.nasa.gov/Features/OrbitsManeuver/</a> .....	9
Figura 2 - Longitudes de onda de las 7 primeras bandas del sensor MODIS.....	10
Figura 3 - Deformaciones relacionadas con el ángulo de barrido. ....	12
Figura 4 - Diagrama de producción de los datos MODLAND. Algunos productos como MOD44 y MOD45 podrían considerarse como de nivel L4 (modificado de Roy et al. 2002) .....	17
Figura 5 - Organización de los productos MODLAND en tiles.....	19
Figura 6 - Nomenclatura utilizada para los nombres de los productos MODIS de nivel L2G o superior .....	20
Figura 7 - Localización de las áreas de estudio .....	45
Figura 8 - Área del Tancítaro Michoacán.....	46
Figura 9 - Área de Mazatán, Sonora.....	47
Figura 10 - Área Maya en Quintana Roo .....	48
Figura 11 - Red Neuronal Perceptrón Multicapa para clasificar una imagen multiespectral de siete bandas en 7 categorías .....	57
Figura 12 - Diagrama de flujo describiendo el procedimiento de análisis de los datos MODIS .....	60
Figura 13 - Qa Vs precipitación .....	66
Figura 14 - Separabilidad máxima alcanzada con 7 bandas por los diferentes productos en cada área de estudio.....	67
Figura 15 - Valores de fiabilidad global obtenidos por cada una de los tipos de datos analizados por las diferentes metodologías en el área de Tancítaro, Michoacán .....	69
Figura 16 - Valores de fiabilidad global obtenidos por cada una de los tipos de datos analizados por las diferentes metodologías en el área Maya, Quintana Roo.....	70
Figura 17 - Valores de fiabilidad global obtenidos por cada una de los tipos de datos analizados por las diferentes metodologías en el área de Mazatán, Sonora .....	71
Figura 18 - Distribución de diferentes tipos de coberturas (% del área total de la cobertura) en función de la elevación.....	72

# CAPÍTULO I. INTRODUCCIÓN

---

El presente trabajo de tesis surge de la necesidad de contar con información actualizada sobre el uso y la cobertura del suelo en México. Mapas que sean generados con datos de fácil obtención, analizados con herramientas adecuadas y evaluados de manera robusta. En esta sección se presenta la justificación, los objetivos, la hipótesis y la estructura del trabajo de tesis.

## 1.1. JUSTIFICACIÓN DEL TRABAJO.

La cobertura global del suelo ha sido modificada para proveer de alimento y abrigo a más de seis billones de personas. En décadas recientes, las áreas de cultivo, pastizales inducidos o cultivados y plantaciones se han expandido aumentando los requerimientos de energía, agua y fertilizantes, y ocasionado grandes pérdidas en la biodiversidad. Estos cambios en el uso de suelo han permitido al ser humano apropiarse de los recursos, pero al mismo tiempo pueden afectar la capacidad de los ecosistemas para sostener la producción de alimento, mantener los recursos forestales, regular el clima y afectar la disponibilidad del agua y la calidad del aire.

La Organización de las Naciones Unidas para la Agricultura y la Alimentación (FAO) estima que las tasas más altas de deforestación tropical a nivel mundial ocurren en México y Centroamérica, y son solo comparables con las del sureste de Asia y dos veces mayores a las de Sudamérica y África (FAO 1993).

Hoy en día tenemos la tarea de manejar los recursos de manera que sea posible mantener la capacidad de la biósfera para proveer bienes y servicios a largo plazo (Foley 2005). Es por lo tanto, indispensable contar con datos actualizados sobre la cobertura y el uso del suelo para entender y manejar todos estos procesos. En este contexto, la percepción remota es una herramienta que se ha vuelto indispensable en el mapeo de uso y cobertura de suelo.

La observación de la tierra desde el espacio, con fines civiles, comenzó en 1972 con el lanzamiento del satélite Landsat 1, en las cuatro décadas siguientes, se han puesto en marcha en promedio 5 satélites de observación de la tierra al año, muchos de ellos con énfasis en el monitoreo de vegetación. La observación desde los satélites ofrece diversos beneficios sobre la realizada desde los aviones, como poder monitorear la cobertura terrestre a escala mundial y

revisitar áreas de manera frecuente y predecible y a menor costo. Hoy en día, muchos satélites, con una variedad de capacidades de visualización de la Tierra, mapean la superficie terrestre de manera diaria. Existen sensores con una muy alta resolución espacial (del orden de los decímetros) que tienen aplicaciones en el análisis detallado urbano, catastral, mapeo de infraestructura para transporte, y planificación de servicios públicos. Por otro lado hay sensores con una muy alta resolución espectral, que cuentan con cientos de bandas y tienen aplicaciones principalmente en el campo de la geología identificando diferentes tipos de materiales, lo cual no es posible con otro tipo de imágenes. También existe una tendencia en la mejora de la resolución radiométrica o niveles de gris, lo que aumenta la capacidad de discernir los objetos de la imagen.

Las imágenes de alta o muy alta resolución espacial, como SPOT, Landsat ETM+, Ikonos o Quick Bird, presentan limitaciones para el mapeo de áreas extensas, la adquisición y el análisis de tales imágenes son costosos y lentos. Entre más alta es la resolución espacial, la resolución radiométrica y temporal (frecuencia de la toma de una imagen sobre la misma región) se reduce, lo que dificulta la obtención de imágenes libres de nube en particular en las regiones tropicales.

En el ámbito internacional se han realizado diferentes investigaciones utilizando percepción remota con imágenes de diferentes sensores, algunas de las investigaciones más relevantes son: el Programa VEGETATION con base en imágenes SPOT VGT, TREES-Project utilizando datos de ERS-1/2 y NOAA-AVHRR, Proyecto NASA Pathfinder GAC, el Programa Internacional de la Biósfera y Geósfera Data y sistemas de información (IGBP 1990).

A nivel nacional en el territorio mexicano, el INEGI a través del Departamento de Uso del Suelo elabora a partir de 1968 las diferentes versiones de la Carta de Uso del Suelo y Vegetación contando hasta el momento con tres mapas a nivel nacional. En 1997 el Instituto de Geografía de la Universidad Nacional Autónoma de México (UNAM) y la Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO), realizaron el proyecto "Obtención de cartas de vegetación a partir de imágenes AVHRR de alta resolución", cuyo objetivo principal fue la elaboración de mapas de vegetación anual y estacional para el año de 1996, en 1998 CONABIO, realizó un trabajo de monitoreo de incendios y su "afectabilidad" biológica. En 1999 y 2000 el Instituto de Geografía de la UNAM y CONABIO, realizaron el proyecto "Análisis multitemporal de las condiciones de la vegetación en México por medio de imágenes de satélite AVHRR de alta resolución".

Aunque las imágenes AVHRR fueron diseñadas para estudios meteorológicos, estas fueron ampliamente utilizadas para el monitoreo de las cubiertas del suelo debido a su alta resolución temporal, la gran superficie que abarca una sola imagen y su bajo costo. En México, se reportan diferentes ejercicios de mapeo de la cubierta vegetal con base en imágenes AVHRR (SARH 1992, Evans et al. 1992, Eggen-McIntosh et al. 1992, CONABIO).

Ante el éxito de AVHRR, se desarrollaron otros sensores con estas características de baja resolución espacial y alta resolución temporal, diseñados especialmente con el fin de monitorear las coberturas del suelo. Entre estos destacan SPOT-VEGETATION y MODIS (Moderate Resolution Imaging Spectro-Radiometer), que superan AVHRR en muchos aspectos (mejor calidad geométrica, bandas espectrales seleccionadas para la caracterización de las cubiertas terrestres, mejor calibración radiométrica, etc.).

MODIS presenta más potencial que SPOT-VEGETATION para monitorear las cubiertas forestales de México que se caracterizan por su alta diversidad y complejidad espacial. Debido a su mejor resolución espacial (hasta 250 m contra de 1150 m en SPOT-VGT) y espectral (7 bandas para monitoreo terrestre contra 3 de SPOT-VGT).

Este trabajo contribuye a conocer la efectividad de imágenes de baja resolución espacial (en el orden de los cientos de metros) e índices de vegetación derivados de las mismas bajo diferentes métodos de análisis para evaluar el nivel de discriminación alcanzable en tres regiones representativas de las coberturas presentes en el país.

## **1.2. OBJETIVOS E HIPÓTESIS DE TRABAJO.**

### **1.2.1. Objetivo general**

Evaluar la efectividad de diferentes tipos de datos (índices de vegetación, imágenes de reflectancia de la superficie diarias y compuestos de 8 días de reflectancia de la superficie) derivados de percepción remota con baja resolución espacial, analizados mediante diferentes metodologías de procesamiento (Máxima Verosimilitud y Redes Neuronales) en el mapeo de la cobertura y uso de suelo en México.

### **1.2.2. Objetivos específicos**

- 1) Comparar la efectividad de los tres tipos de datos de baja resolución espacial en la discriminación de clases de uso y cobertura de suelo en México.
- 2) Evaluar y comparar la fiabilidad alcanzada por un método paramétrico contra uno no paramétrico en datos de baja resolución espacial.
- 3) Determinar la aportación de diferentes tipos de datos auxiliares en las clasificaciones de los tres tipos de datos de baja resolución espacial.
- 4) Determinar si los diferentes tipos de datos y metodologías se comportan de manera consistente en diferentes tipos de ecosistemas mexicanos.

### **1.2.3. Hipótesis de trabajo**

Las imágenes de baja resolución espacial pueden ser analizadas bajo métodos de análisis que permitan producir cartografía actualizada obteniendo resultados confiables en el monitoreo y análisis del uso y cobertura de suelo en distintos tipos de paisajes de México.

## **1.3. ESTRUCTURA DE LA TESIS.**

El presente trabajo de tesis consta de siete capítulos. En el capítulo 1 se presenta la justificación del trabajo, los objetivos y la hipótesis. El capítulo 2 ofrece una descripción general de los sensores más utilizados en el mapeo del uso y la cobertura del suelo, así como una presentación detallada de las características del sensor MODIS y de los productos que fueron utilizados en este trabajo de tesis. El capítulo 3 es una revisión sobre los trabajos que se han hecho a partir de datos MODIS, las técnicas de procesamiento utilizadas y los resultados obtenidos en diferentes partes del mundo. En el capítulo 4 se presenta una descripción de los criterios de selección de las áreas, así como una caracterización general de cada una de estas. El capítulo 5 está conformado por una descripción de los datos utilizados en este trabajo de tesis, así como la descripción de los métodos de pre-procesamiento, análisis y evaluación de resultados utilizados para su análisis. El capítulo 6 presenta los resultados obtenidos para cada una de las áreas. En el capítulo 7 se presenta una discusión sobre la utilidad de los datos de baja resolución en el mapeo del uso y cobertura de suelo, y las conclusiones, consideraciones finales y recomendaciones para trabajos futuros en la misma línea.

## **CAPÍTULO II. MARCO TEÓRICO**

---

Este capítulo tiene como propósito presentar los conceptos que sirven como base a este trabajo de tesis. En la primera sección se da una breve introducción sobre los sensores más populares en el monitoreo del uso y la cobertura del suelo con el objeto de contextualizar el surgimiento y caracterización del sensor MODIS, cuyos datos son la base de este estudio. En la segunda sección se describen las especificaciones técnicas del sensor MODIS, que son de utilidad para conocer como se obtienen sus datos. En la tercera sección se hace una descripción general de los datos procesados o productos que genera MODIS, esto sirve como referencia para saber cómo y de donde se puede obtener, en esta misma sección se caracterizan con detalle los datos analizados en esta tesis.

### **2.1. LOS DATOS DE SATÉLITE EN EL MONITOREO DE LA COBERTURA Y EL USO DEL SUELO**

Gracias al desarrollo de las tecnologías aeroespacial y de análisis geoespacial, hoy en día es posible llevar a cabo el monitoreo remoto de la cobertura terrestre a través de imágenes de satélite. Esto representa una gran ventaja sobre todo en casos donde se requiere hacer estudios en áreas extensas, de difícil acceso, o bien en periodos de tiempo cortos (Chuvienco 2002). Debido a su importancia y potencial de aplicación, han surgido una gran variedad de sensores con diferentes enfoques.

#### **2.1.1. Desarrollo de los datos de sensores remotos**

Desde los 70's hasta el presente, los sensores remotos han venido evolucionando, con una tendencia a evolucionar principalmente en el incremento de la resolución.

La resolución se define como la capacidad de discernir objetos o detalles en las imágenes captadas. Una imagen de satélite se caracteriza por las siguientes modalidades de resolución: espacial, espectral, radiométrica y temporal.

**1) Resolución espacial:**

Este concepto designa al objeto más pequeño que se puede distinguir en la imagen. Está determinada por el campo instantáneo de vista, que es el ángulo sobre el cual el detector es sensible a la radiación. Este ángulo determina el tamaño del elemento más pequeño de la imagen (o pixel) y depende de: la altura del sensor con respecto a la Tierra, el ángulo de visión, la velocidad de escaneado y las características ópticas del sensor. Aunque los sensores han venido incrementando su resolución espacial, ello no sugiere que en todos los estudios de monitoreo se requiera siempre de una resolución espacial muy alta. Esto dependerá de la escala geográfica a la que deba recopilarse la información, y esto a su vez, en general, depende de la extensión del área de trabajo.

El término “muy alta resolución espacial” se refiere a una imagen que capta detalles finos, en algunos casos en el orden de los decímetros, mientras que una imagen de baja resolución, en general en el orden de más de la centena de metros, no tiene esta capacidad. Por ejemplo las imágenes Landsat TM, tienen una resolución espacial de 30 m en las bandas 1, 2, 3, 4, 5 y 7 y de 120 m en la 6 (térmica). El rango en tamaños de celda puede ser tan amplio como el del sensor AVHRR de 8 km captados en el pixel, Landsat MSS originalmente de 90 m, Landsat ETM de 30 m, SPOT de 10 m, o Ikonos de 1 m de resolución.

La limitación en el uso de imágenes de baja resolución espacial es que un píxel representa una extensión grande, donde el tamaño de la mayoría de los objetos en el suelo es menor que el área que cubre un píxel en la imagen. De este modo, la respuesta espectral en un píxel dado no necesariamente corresponderá a una clase de cobertura sino a una combinación de varias. Sin embargo, las imágenes de baja resolución representan una buena alternativa al abarcar áreas extensas, lo cual resulta muy útil en el mapeo de países de gran tamaño, como es el caso de México.

**2) Resolución temporal:**

La resolución temporal se refiere a la frecuencia con la que el sensor adquiere imágenes de la misma porción de la superficie terrestre (Chuvieco 1990). El ciclo de repetición de los Landsat-1 al Landsat -3 era de 17 días. A partir del Landsat 4 en 1984, el ciclo de repetición se redujo a 15 días. SPOT permite un ciclo de repetición de entre 3 y 26 días debido a la capacidad de tomar imágenes fuera del NADIR mediante una toma oblicua. En aplicaciones de mapeo de uso y cobertura terrestre, contar con datos frecuentes (diarios) de la vegetación permite tener

datos sobre su fenología y aumenta la probabilidad de disponer de imágenes libres de nubes. Los sensores SPOT VEGETATION (VGT) y MODIS brindan imágenes diarias que tienen el potencial para ser aprovechadas con este fin.

### 3) **Resolución espectral:**

La resolución espectral consiste en el número de canales espectrales (y su ancho de banda) que es capaz de captar un sensor. De acuerdo a la resolución espectral podemos clasificar a las imágenes como: multiespectrales, superespectrales e hiperespectrales.

Una imagen multiespectral consiste en unas pocas bandas, cada una representa una imagen adquirida en una longitud de onda particular. Por ejemplo, SPOT HRV es un sensor multiespectral ya que detecta radiaciones en tres bandas de longitud de onda, verde (500-590nm), roja (610-680nm) e infrarrojo (790-890nm). En este caso una sola escena SPOT multiespectral consiste en tres intensidades en tres longitudes de onda. En este caso cada pixel de cada escena tiene tres valores de intensidad correspondiendo a las tres bandas. Una imagen Ikonos tiene 4 bandas, y Landsat TM cuenta con 7.

Las imágenes superespectrales son captadas por una nueva generación de satélites capaces de adquirir escenas a muchas más longitudes de onda. Un ejemplo es el sensor MODIS el cual adquiere 36 bandas espectrales cubriendo regiones y rangos desde el visible, el infrarrojo cercano, infrarrojo de onda corta hasta el infrarrojo térmico.

Las imágenes hiperespectrales consisten en cientos de bandas espectrales contiguas. Hasta el momento las imágenes de satélite hiperespectrales no son de tipo comercial sino experimental (p. ej. el sensor Hyperion a bordo de EO1, el sensor CHRIS a bordo de PRABO-ESA), y surgen de la gran experiencia del mapeo principalmente con fines geológicos a partir de sensores como CASI y AVIRIS transportados a bordo de aviones. Las bandas en las imágenes hiperespectrales son más estrechas comparadas con las multiespectrales y superespectrales.

### 4) **Resolución radiométrica:**

Se la llama también resolución dinámica, y se refiere a la cantidad de niveles de gris en que se divide la radiación recibida para ser almacenada y procesada posteriormente. Esto depende del convertidor analógico digital utilizado. Así por ejemplo Landsat MSS tiene una resolución espectral de  $2^6 = 64$  niveles de gris en el canal 6, y Landsat MSS en las bandas 4 a 7 de  $2^7 = 128$  niveles de gris, mientras que en Landsat TM es de  $2^8 = 256$ . Esto significa que tenemos una mejor

resolución dinámica en el sensor TM y podemos distinguir mejor las pequeñas diferencias de radiación.

Un resumen de las características de resolución de los satélites más populares se presenta en el Cuadro 1.

**Cuadro 1** - Características generales de los sensores más utilizados

<b>SENSOR</b>	<b>Resolución espacial (en metros)</b>	<b>Resolución espectral</b>	<b>Resolución temporal (Tiempo de revisita)</b>	<b>Cubrimiento o por escena (swath km)</b>	<b>Resolución radiométrica</b>	<b>Fecha de lanzamiento</b>
<b>AVHRR-2 AVHRR-3</b>	1.1km	5 bandas 6 bandas	12 horas	2580 km.	10 bits	1978
<b>LANDSAT TM</b>	30m	7 bandas	16 días	185km	8 bits	1982
<b>SPOT VGT</b>	1000M	3 bandas	1 día	2250km	8 bits	1998
<b>IKONOS</b>	PAN 1m MS 4m	5 bandas	3-5 días	11.3 km	11 bits	1999
<b>ASTER</b>	15m (bandas 1-3) 30m (bandas 4-9) 90m (bandas 10-14)	14 bandas	16 días	60 km	8 bits 8 bits 12 bits	1999
<b>MODIS</b>	250m (bandas 1-2) 500m (bandas 3-7) 1000m (bandas 8-36)	36 bandas	1-2 días	2330km	16 bits	2000 (terra) 2002 (agua)
<b>SPOT 5</b>	Pan: 5m (nadir) MS: 10m (nadir) SWI: 20m (nadir)	5 bandas	2-3 días	60m	8 bits	2002
<b>QUICK BIRD</b>	PAN 61 -72cm MS: 2.44- 2.88 m	5 bandas	1-3.5 días	16.5km	11 bits	2001

Pan.; Imagen pancromática, MS.; Imagen multiespectral.

Los datos utilizados en esta tesis, son derivados del sensor MODIS, el cual se describe en detalle en la siguiente sección.

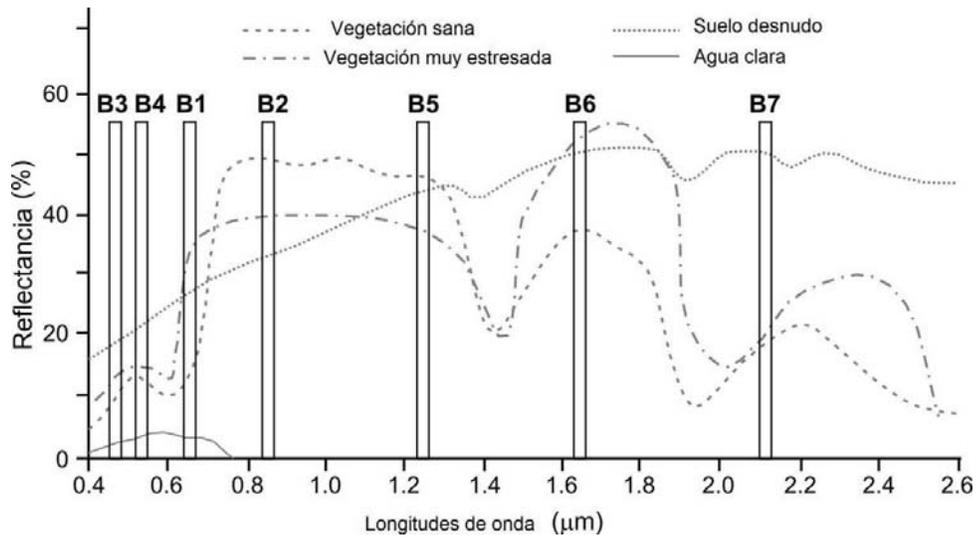
## 2.2. ESPECIFICACIONES TÉCNICAS DEL SENSOR MODIS

El sensor MODIS se encuentra a bordo de los satélites Terra o EOS-AM (lanzado en diciembre de 1999) y Aqua o EOS-PM (lanzado en mayo de 2002). La órbita de ambas plataformas es helio-sincrónica y cuasi-polar con una inclinación de 98.2° y 98° y una altitud media de 708 y 705 km respectivamente (NASA MODIS WEB 2008). Terra está programada para pasar de norte a sur cruzando el ecuador a las 10:30 de la mañana en su órbita descendente, mientras que Aqua pasa de sur a norte sobre el ecuador a las 1:30 de la tarde, un esquema del cubrimiento MODIS se presenta en la Figura 1. Las primeras imágenes captadas por Terra se obtuvieron en Febrero del año 2000. Ambas plataformas monitorean la totalidad de la superficie terrestre cada 1 o 2 días dependiendo de la latitud.



**Figura 1** - Esquema del cubrimiento MODIS (Tomado de MODIS Rapid Response Team.)  
<http://earthobservatory.nasa.gov/Features/OrbitsManeuver/>

El instrumento MODIS cuenta con una alta sensibilidad radiométrica (12 bit) en 36 bandas espectrales, en un rango de longitud de onda que va de los 0.4  $\mu\text{m}$  a los 14.4  $\mu\text{m}$  (ver Cuadro 2). Las primeras 19 bandas están posicionadas en la región del espectro electromagnético situado entre 0.405 nm y 2.155 nm. Las bandas de la 1 a la 7 son útiles para las aplicaciones terrestres (Figura 2); las bandas 8 a la 16 para las observaciones oceánicas y las bandas 17 a 19 para las mediciones atmosféricas. Las bandas 20 a la 36, cubren la porción del infrarrojo térmico del espectro de (3,660-14,385 nm) (NASA 2008).



**Figura 2** - Longitudes de onda de las 7 primeras bandas del sensor MODIS

El sensor MODIS es un explorador de barrido: con un espejo móvil que oscila perpendicularmente a la dirección de la trayectoria con un ángulo de  $\pm 55^\circ$  permite explorar una franja de terreno a ambos lados de la traza del satélite, cuyo ancho (swath) es de 2,330 km. El sistema óptico es un telescopio con dos espejos fuera de su eje focal que direccionan la radiación incidente a cuatro sistemas ópticos reflectantes, uno para cada región espectral (visible, infrarrojo cercano, medio y térmico). Para las bandas del visible y del infrarrojo cercano se utiliza una nueva tecnología de fotodiodos de silicio. Para el infrarrojo térmico se usan detectores de Telureto de Mercurio-Cadmio (HgCdTe).

**Cuadro 2 - Características de las bandas MODIS**

Uso principal	Banda	Ancho de banda <sup>1</sup>
Limites de Tierra/Nubes/aerosoles	1	620 - 670
	2	841 - 876
Propiedades de Tierra/Nubes/aerosoles	3	459 - 479
	4	545 - 565
	5	1230 - 1250
	6	1628 - 1652
	7	2105 - 2155
Color del océano /fitoplancton/bioquímica	8	405 - 420
	9	438 - 448
	10	483 - 493
	11	526 - 536
	12	546 - 556
	13	662 - 672
	14	673 - 683
	15	743 - 753
	16	862 - 877
Atmosférica Vapor de agua	17	890 - 920
	18	931 - 941
	19	915 - 965
Temperatura de la Superficie/nubes	20	3.660 - 3.840
	21	3.929 - 3.989
Temperatura Atmosférica	22	3.929 - 3.989
	23	4.020 - 4.080
	24	4.433 - 4.498
Nubes Cirrus Vapor de agua	25	4.482 - 4.549
	26	1.360 - 1.390
Propiedades de las nubes	27	6.535 - 6.895
	28	7.175 - 7.475
	29	8.400 - 8.700
Ozono	30	9.580 - 9.880
	31	10.780 - 11.280
Temperatura de la Superficie/nubes	32	11.770 - 12.270
	33	13.185 - 13.485
Altitud de las nubes	34	13.485 - 13.785
	35	13.785 - 14.085
	36	14.085 - 14.385

<sup>1</sup> Bandas de la 1 a la 19 se expresan en nm; Bandas de la 20 a la 36 en  $\mu\text{m}$

### 2.2.1. Deformaciones relacionadas con el barrido (ángulo de vista)

Debido al procedimiento de barrido mediante el cual el sensor MODIS captura los datos de la superficie terrestre, se produce una deformación de la imagen resultante. Tal como se muestra en la Figura 3, mientras que los píxeles que se encuentran en la vertical del sensor (al nadir en un ángulo cenital) presentan deformaciones mínimas, el campo instantáneo de vista del sensor aumenta en los extremos de la imagen llegando a tener una deformación de 5 veces el tamaño del píxel en el eje Y, y dos en el eje X.

Estas deformaciones se conocen con el efecto Bowtie, que consiste en la presencia de líneas repetidas cerca de los bordes de las imágenes MODIS. No obstante, existen programas para corregir este defecto, los píxeles localizados en los bordes de la imagen deben ser empleados con cautela.

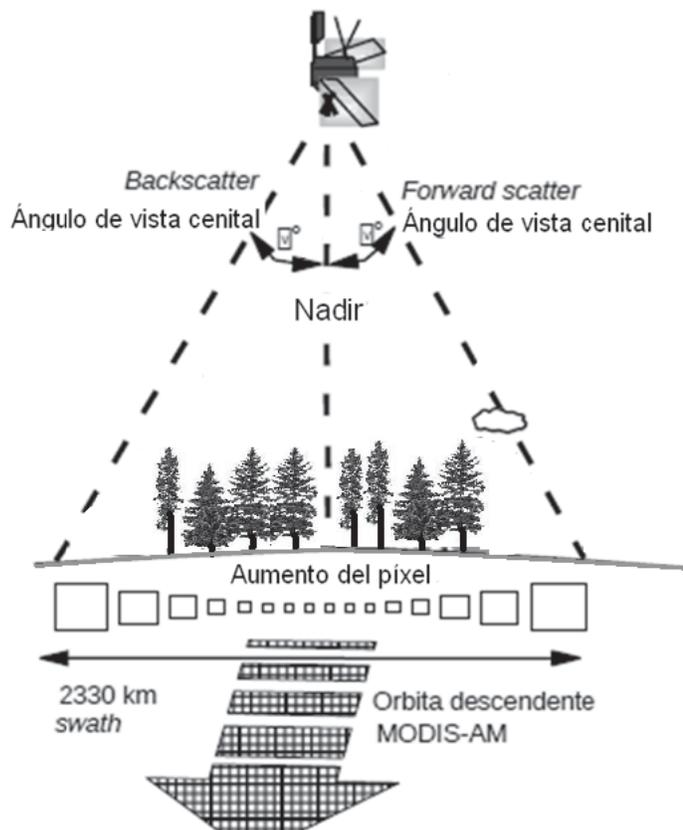


Figura 3 - Deformaciones relacionadas con el ángulo de barrido.

### **2.2.2. Transmisión y pre-procesamiento de los datos**

Junto con todos los datos de otros instrumentos a bordo de Terra y Aqua, los datos MODIS se transfieren a las estaciones en tierra en White Sands, Nuevo México, a través del seguimiento y retransmisión de datos desde el sistema de satélites (Tracking and Data Relay Satellite System - TDRSS).

Una de las características que distinguen a MODIS de cualquier otro instrumento es que no solo se puede tener acceso a los datos después de su captura y almacenamiento para la posterior descarga, además, tiene la capacidad de poder transmitir de manera directa los datos crudos que colecta a una base en tierra (Direct Broadcast). El instrumento de MODIS Terra fue uno de los primeros satélites en operar con una transmisión directa de datos, de manera constante y con disponibilidad de que cualquiera con el equipo y software adecuados pueda obtener estos datos gratuitamente.

Los datos son enviados al sistema de datos y operaciones del EOS Data and Operations System (EDOS) al centro espacial de vuelo Goddard. El nivel 1A, nivel 1B, georreferencia y los productos libres de nube, así como los de suelo y atmósfera de más alto nivel son producidos por un sistema de procesamiento adaptable (MODAPS) y posteriormente, son repartidos a tres centros activos de almacenamiento (DAACs por sus siglas en inglés) para su distribución.

Los distintos datos derivados de las observaciones de MODIS pueden obtenerse desde diferentes fuentes, MODIS nivel 1 y productos de atmósfera están disponibles en el portal de Nivel 1 y Sistema de Distribución de Archivos de la Atmósfera (LAADS por sus siglas en inglés) <http://ladsweb.nascom.nasa.gov/data/search.html>. Los productos para aplicaciones terrestres a través del centro de datos EROS de Estados Unidos en el portal <http://daac.gsfc.nasa.gov/>, los productos de cobertura de nieve y hielo se obtienen a través del portal del Goddard Space Flight Center (GSFC) <http://www.nasa.gov/centers/goddard/home/index.html>.

La obtención de imágenes satelitales de MODIS se estuvo realizando por un largo periodo de tiempo a través del portal del Earth Observing System Data Gateway <http://redhook.gsfc.nasa.gov/>. A partir de 2009 la nueva herramienta denominada Warehouse Inventory Search Tool (WIST) a través del portal <http://wist.echo.nasa.gov> reemplazó al primero.

En México, la CONABIO recibe diariamente imágenes del sensor MODIS de los satélites Terra-1 y Aqua-1. Las imágenes MODIS desde el mes de julio de 2004 y hasta la fecha se

encuentran a disposición del público, calibradas y rectificadas en el portal [http://www.conabio.gob.mx/informacion/geo\\_espanol/modis/modis\\_ftp.html](http://www.conabio.gob.mx/informacion/geo_espanol/modis/modis_ftp.html). (CONABIO 2009).

### **2.2.3. Niveles de procesamiento**

Los productos MODIS son divididos en cinco niveles (0 a 4) en función del grado de procesamiento realizado:

**Nivel 0:** Son los datos brutos sin ningún tratamiento, sólo los artefactos de comunicación han sido eliminados, como por ejemplo los paquetes de sincronización y los encabezados de comunicación entre el satélite y la estación de recepción.

**Nivel L1:** Son los datos de geolocalización (identificados como MOD03) que contiene coordenadas geodésicas, información sobre la elevación del terreno, máscara de tierra/agua, ángulo de elevación, cenit y azimut del satélite y del sol.

**Nivel 1A:** Son los productos utilizados para la geolocalización, la calibración y el procesamiento. Contienen los niveles de radiancia (MOD01) de las 36 bandas, junto con datos auxiliares del sensor y del satélite como coeficientes de calibración radiométrica y corrección geométrica, parámetros de georreferenciación (efemérides de la plataforma).

**Nivel 1B:** Son los productos que contienen las radiancias calibradas y con geolocalización (MOD02) para las 36 bandas generadas por el nivel 1A. Es por lo tanto una imagen radiométricamente corregida y calibrada a unidades físicas. Datos adicionales incluyen estimaciones de calidad, del error y datos de calibración.

**Nivel L2:** Son los productos que contienen variables geofísicas. Estos productos se generan a partir del producto nivel L1B aplicando correcciones atmosféricas y algoritmos bio-ópticos.

**Nivel L2G:** Son los productos que contienen variables geofísicas proyectadas en una malla uniforme. Los datos se adquieren durante un periodo fijo de tiempo (12 o 24 horas).

**Nivel 3:** Son productos de valor agregado que describen variables biofísicas. En algunos casos, se realiza un remuestreo a una menor resolución espacial que las bandas originales y se elaboran compuestos de intervalos de tiempo establecidos (1, 8, 16 y 30 días).

**Nivel 4:** Son los productos generados por la incorporación de los datos MODIS en modelos para así estimar variables biofísicas (NASA 2008). Estos productos se generan usando productos nivel L2 y datos auxiliares.

### **2.3. DESCRIPCIÓN DE LOS PRODUCTOS MODLAND**

Con base en los datos MODIS, se elabora una gran variedad de productos enfocados al estudio de los océanos (MOCEAN), la atmósfera (MODIS Atmosphere), la criósfera y las cubiertas terrestres (MODLAND). En esta sección se presenta, una descripción general de las características comunes de producción de los datos MODIS, y una caracterización más a detalle de los productos relacionados de la cobertura terrestre que se utilizaron en el presente trabajo de tesis. Junto con otros productos, estos son elaborados por un grupo de científicos que conforman el equipo MODLAND, encargado de producir los datos enfocados a la investigación del complejo terrestre y el cambio global (Justice et al. 2002).

De acuerdo a Morisette et al. (2002), los productos MODIS pueden dividirse en tres grandes grupos:

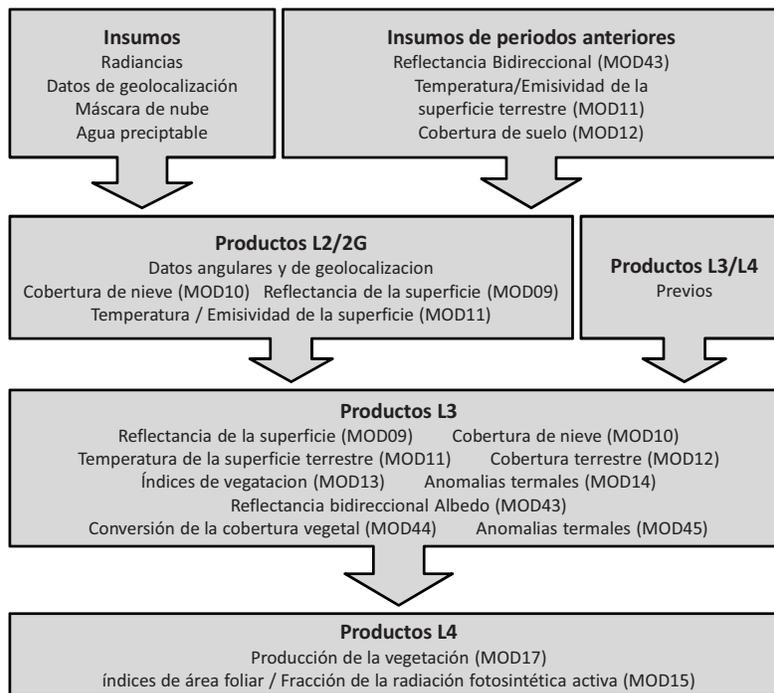
- Variables relacionadas con el balance de energía: estas incluyen la reflectancia de la superficie (MOD09), la cobertura de nieve (MOD10), la temperatura y la emisividad de la superficie terrestre (MOD11), el albedo y la función de la distribución de la reflectancia bi-direccional (MOD43).

- Variables biofísicas relacionadas con la vegetación que comprenden los índices de vegetación (MOD13), el índice de área foliar o LAI por sus siglas en inglés y la fracción de radiación activa fotosintética FPAR (MOD15) y la producción primaria (MOD17).

- Características de la cobertura terrestre: coberturas del suelo (MOD12), puntos de calor e incendios (MOD14), conversión de la cobertura vegetal y fracción de vegetación (MOD44) y finalmente áreas quemadas (MCD45).

### **2.3.1. Producción de los datos MODLAND**

La elaboración de los productos MODLAND sigue una cadena de procesamiento claramente definida y ampliamente documentada (Justice et al. 1998 y 2002a) (Figura 4). Los datos de nivel 1A (MOD01) sirven de insumo para generar MOD02 así como datos de geolocalización que incluyen elevación, ángulos de observación e inclinación solar. Los productos de nivel L2 necesitan información atmosférica como máscara de nube (MOD35), aerosoles (MOD04), vapor de agua (MOD05), perfiles atmosféricos (MOD07). La elaboración de ciertos productos como MOD09 (reflectancia de la superficie) o MOD11 (temperatura y emisividad) requiere de productos de más alto nivel producidos anteriormente (temperatura, parámetros de distribución de la función de reflectancia bidireccional). Datos adicionales del nivel L2 brindan información sobre los ángulos de observación y solares. El insumo más importante para los datos de nivel L3 y L4 es la reflectancia de la superficie. Los datos de nivel L3 están organizados en composiciones espaciales (mosaicos o “tiles”) y temporales (datos de 8, 16, 30 días por ejemplo). Finalmente, los productos de nivel L4 se derivan de los productos de nivel anterior y además de datos adicionales (información climática por ejemplo) y/o usan modelos para su generación.



**Figura 4** - Diagrama de producción de los datos MODLAND. Algunos productos como MOD44 y MOD45 podrían considerarse como de nivel L4 (modificado de Roy et al. 2002)

La evaluación de la calidad de los productos MODIS forma parte integral del proceso de producción y tiene el objetivo de estimar y documentar la calidad científica de los productos MODLAND con respecto a la calidad esperada (Roy et al. 2002). La calidad de los productos se brinda como metadata a nivel de imagen y en una capa adicional para cada píxel. Se generan diferentes índices sobre la calidad (Quality Assessment, QA) para cada píxel. Estos índices se almacenan en capas de información que forman parte del producto y brindan distintos tipos de información dependiendo del producto como: la fiabilidad de la estimación del valor del parámetro representado, el éxito de las correcciones, la presencia de nube, aerosol, vapor de agua, ozono o si el píxel corresponde a tierra o agua.

Los productos MODIS están sujetos a un proceso de validación que incluye comparaciones con datos tomados in situ en sitios de prueba, comparaciones con datos y productos de otro sensores como ASTER, AVHRR, Ikonos, MISR, TM/ETM+, y con datos de referencia obtenidos de manera independiente (Morissette et al. 2002 y 2003). El proceso de validación pasa por cuatro fases (Nightingale et al. 2009):

**Fase 1.** La fiabilidad del producto se estimó usando medidas independientes obtenidas de un pequeño número de sitios y periodos de tiempo.

**Fase 2.** La evaluación se hizo con base en un amplio grupo de sitios y fechas además de contar con muchos datos de campo.

**Fase 3.** La fiabilidad se obtuvo a través de medidas independientes hechas de manera sistemática y estadísticamente robusta que representa las condiciones globales. Por lo tanto, existen estimadores de la incertidumbre de estos productos. Se considera que, aunque puedan generarse versiones mejoradas, estos productos son idóneos para su uso en publicaciones científicas.

**Fase 4.** Los resultados de validación de la fase 3 se actualizan cuando se libera una nueva versión del producto.

### **2.3.2. Características Generales**

#### **1) Las colecciones de datos MODIS**

A medida que se perfeccionan o se corrigen los algoritmos de elaboración, se generan versiones (o colecciones) mejoradas. Todos los productos archivados se reprocesan, incluyendo los productos de fechas anteriores, por lo cual las colecciones más recientes brindan un mejor desempeño. Actualmente, la mayoría de los productos se encuentran en la versión 5, en la cual se mejoraron las máscaras de nube y sombra que ahora incluyen los píxeles adyacentes a las nubes y las capas sobre calidad, se refinó la corrección atmosférica, se mejoró la corrección del área de costa y se aumentó la congruencia entre los productos de diferente resolución espacial. En muchos casos, se eliminaron capas de información poco útiles o al contrario se incorporaron nuevas capas. Las principales mejoras de la colección 5 respecto a las anteriores se encuentran descritas a detalle en el portal de la NASA.

## 2) Organización espacial

Dependiendo del nivel de procesamiento, los productos MODIS presentan diferentes formas de organización espacial. Los que corresponden al nivel L2, corresponden a los datos colectados por el sensor, de tal forma que el tamaño de cada escena depende del ancho de la traza (aproximadamente 2330 km) y del tiempo de colecta: Los gránulos que corresponden a 5 minutos de colecta son de aproximadamente 2340 x 2330 km.

La mayoría de los productos de MODLAND son del nivel L2\_G, 3 y 4 y están organizados en una retícula de 1,200 x 1,200 km (10 por 10 grados al ecuador) en proyección sinusoidal (SIN). Cada cuadro, llamado tile, se localiza por un número de columna (h) y línea (v). Así por ejemplo la parte continental del territorio mexicano está cubierta en su totalidad por los mosaicos h08v06, h8v07, h09v06, h09v07. En la Figura 5 se presenta la retícula de mosaicos MODIS, donde 188 cuadros indicados en blanco fuera de la figura de proyección de la tierra no contienen información y los 134 cuadros en blanco dentro de la figura de proyección de la tierra contienen 100% de agua.

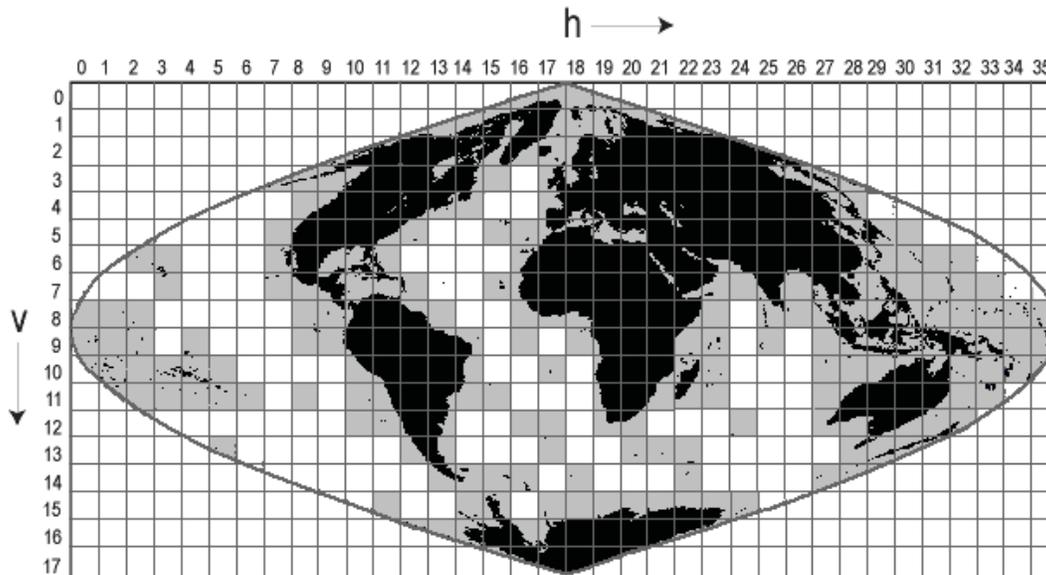


Figura 5 - Organización de los productos MODLAND en tiles

Finalmente, algunos productos, llamados Malla de Modelado Climático (Climate Modeling Grid, CMG por sus siglas en inglés), representan toda la superficie terrestre a una resolución espacial baja (0.05° que corresponden a 5600 m aproximadamente) en proyección geográfica.

### 3) Formato de Almacenamiento HDF

Los productos MODLAND se encuentran en el formato Hierarchical Data Format (HDF), desarrollado por el Centro Nacional de Aplicaciones de Supercomputación (NCSA). De acuerdo al NCSA, este formato de archivo físico, en el nivel más sencillo sirve para almacenar datos como cualquier otro formato, pero en su nivel de más complejidad, es una colección de aplicaciones y utilidades para manipular los datos (Fishtaler 1998). El formato HDF presenta diferencias con el formato EOS-HDF por lo que pueden surgir problemas en el procesamiento de datos. Por ejemplo el software LDOPE genera archivos HDF que no son compatible con la herramienta MRT que acepta solo datos en formato EOS-HDF.

### 4) Identificación de los productos

Los productos MODIS se nombran a través de la convención adoptada por la NASA. Las diferentes porciones del nombre del archivo permiten conocer la plataforma, el nombre del producto, su fecha de adquisición, versión y localización (Figura 6).

<b>Gránulo</b>				
<b>NOM.Aaaaaddd.hhmm.vv.aaaadddhss.hdf</b>				
1	2	3	4	5
MOD14	A2007364	.1805.005	.2009047020343	.hdf
<b>Tile</b>				
<b>NOM.Aaaaaddd.hxxxvyyy.vv.aaaadddhmm.hdf</b>				
1	2	3	4	5
MOD09GA	A2002006	.h08v07.005	.2008182172646	.hdf
MOD09CGM	A2007364	.005	.2008001133501	.hdf

**Figura 6** - Nomenclatura utilizada para los nombres de los productos MODIS de nivel L2G o superior

1. La primera parte del nombre largo del producto es su nombre corto. Las tres primeras letras hacen referencia a la plataforma: MOD para Terra, MYD para Aqua y MCD para datos combinados Terra/Aqua. El número se refiere al tipo de producto. CMG indica el producto global “Climate Modeling Grid”.

2. Después de la letra A aparece la fecha de adquisición (año y día juliano en los cuatro primeros y tres últimos dígitos respectivamente). Por ejemplo A2007364 indica el año 2007, día juliano 364 (29 de diciembre).

3a. En los productos en gránulos (nivel 2) sigue la hora (hora, minutos) del inicio de la colecta de datos.

3b. En los productos organizados en tiles aparece las coordenadas h, v del tile. Por ejemplo, h08v07 indica la retícula 08 es el horizontal y 07 en la vertical. Este campo no existe para los productos CMG.

4. Sigue la versión de procesamiento o colección, en los ejemplos de la Figura 6 corresponde a la versión cinco.

5. Finalmente aparece la fecha de procesamiento (año, día juliano, hora, minuto y segundo): 2008182172646 indica: Año 2008, día juliano 182 (30 de junio), horario 17 hrs. 26 min 46 seg.

## 5) Metadatos

Los productos MODIS tienen dos fuentes de metadatos. El metadato interno, encapsulado en el HDF, y el metadato externo ECS (“EOSDIS Core System”). El metadato HDF contiene información valiosa como los atributos globales y específicos del dato. El metadato externo es un documento en formato XML, que se entrega con el producto MODIS y contiene parte de la información del metadato HDF.

Algunas características importantes de ciertos atributos de los metadatos de MODIS son las siguientes:

- Xdim e Ydim representan respectivamente el número de filas y columnas.
- La mayoría de los productos de MODLAND usan la proyección sinusoidal con una esferoide de 6,371,007,181 metros.

- Los campos “UpperLeftPointMtrs” y “LowerRightMtrs” se expresan en el sistema de coordenadas de la proyección e identifican respectivamente las esquinas arriba y a la izquierda y abajo a la derecha de la imagen. No representan las coordenadas del centro de los píxeles de las esquinas como ocurre en otros sistemas de referencia.

- Campos adicionales indican la latitud y longitud de la imagen en coordenadas geográficas.

- Los atributos de los metadatos contienen también información para cada capa de información como rango de los valores y factores de escala.

### 2.3.3. Descripción de los productos utilizados en esta tesis.

En esta tesis se utilizaron dos tipos de productos MODIS, el de reflectancia de la superficie o MOD09 y los índices de vegetación o MOD13. Estos se describen a continuación.

#### 1) Reflectancia de la superficie MOD09

El producto MOD09 es una evaluación de la reflectancia de las cubiertas terrestres, es decir la relación entre el flujo luminoso reflejado y el flujo luminoso incidente, tal y como hubieran sido medidos a nivel del suelo si no existiera absorción ni dispersión atmosférica. Los textos claves donde se detallan las características del producto MOD09 se presentan en el Cuadro 3.

**Cuadro 3** - Referencias claves para el producto MOD09

Clave del producto	MOD09
Liga y fecha del documento sobre el algoritmo de elaboración.	<a href="http://modis.gsfc.nasa.gov/data/atbd/atbd_mod08.pdf">http://modis.gsfc.nasa.gov/data/atbd/atbd_mod08.pdf</a> (1999)
Liga y fecha de la guía del usuario	<a href="http://modis-sr.ltdri.org/MAIN_SURFACE_PRODUCTAND%20USER%20GUIDE/MOD09_UserGuide.pdf">http://modis-sr.ltdri.org/MAIN_SURFACE_PRODUCTAND%20USER%20GUIDE/MOD09_UserGuide.pdf</a> (2008)
Artículo(s) clave(s) sobre la elaboración	Vermote et al. 1997 y 2002
Artículo(s) clave(s) adicional(es): nuevas versiones, aplicaciones	Levy et al. 2007; Vermote y Kotchenova 2008

La estimación de la reflectancia se lleva a cabo para todos los píxeles sin nubes para cada una de las siete primeras bandas del sensor MODIS. Para ello, se corrigen los efectos atmosféricos producidos por gases, aerosoles y nubes cirrus. El cálculo se realiza con base en la radiancia al nivel del sensor (MOD02, nivel 1B, utilizando un procedimiento de corrección basado en un modelo de transferencia radiativa desarrollado por Tanré et al. (1986) y simplificado por cuestiones operacionales. El modelo utiliza datos atmosféricos directamente derivados de los datos MODIS (aerosoles, vapor de agua) o, si estos no están disponibles, obtenidos por el Centro Nacional de Predicción Ambiental (NCEP por sus siglas en inglés). El espesor óptico de aerosoles, el parámetro más crítico para las bandas del visible y del infrarrojo próximo, se obtiene únicamente de los datos MODIS. El producto de reflectancia de la superficie es utilizado para generar varios productos MODLAND como son los índices de vegetación, la función de distribución de la reflectancia bi-direccional (BDRF), las coberturas del suelo, la cobertura de nieve, los puntos de calor así como el índice de área foliar y la fracción de radiación activa fotosintética (FPAR). Durante el proceso de validación de la versión 4, la evaluación de la exactitud de los valores de reflectancia de la superficie indica un buen ajuste con una desviación de menos de 5% (Liang et al. 2002).

Este producto se encuentra disponible en varias resoluciones temporales y espaciales como se indica en el Cuadro 4. Las capas de información difieren para cada producto. De manera general indican el número de observaciones, el valor de reflectancia en las diferentes bandas, los ángulos de observación y de inclinación solar. La capa de calidad indica la fiabilidad general, la presencia de nubes, la falla de algún detector, la ejecución o no de las correcciones atmosféricas. Además de los productos producidos por la NASA, la Universidad de Maryland produce un producto similar con resolución espacial de 500 m y temporal de 32 días, disponible a través del portal <http://glcf.umiacs.umd.edu/data/modis/>.

**Cuadro 4 - Productos de reflectancia MOD09.**

Clave	MOD09GQ	MOD09GA	MOD09CMG	MOD09A1	MOD09Q1
Plataforma	Terra, Aqua	Terra, Aqua	Terra, Aqua	Terra, Aqua	Terra, Aqua
Resolución temporal	Diaria	Diaria	Diaria	8 días	8 días
Resolución espacial	250 m	500 y 1000 m	0.05 grados	500 m	250 m
Resolución espectral	Bandas 1-2	Bandas 1-7	Bandas 1-7	Bandas 1-7	Bandas 1-2
Extensión espacial	Tile	Tile	CMG (Global)	Tile	Tile
Nivel	L2G	L2G	L3	L3	L3
Versión	5	5	5	5	5
Fase de validación	2	2	2	2	2
Tamaño del archivo (MB)	280	85	985	64	72
Número de capas	5	12	21	13	13
Capa(s) más importante(s)	250m Reflectancia de la superficie Banda 1-2	500m Reflectancia de la superficie Banda 1-7	Reflectancia de la superficie Reflectancia Banda 1-7 Resolución Brillo Temperatura Bandas 20,21,31,32	500m Reflectancia de la superficie Banda 1-7	250m Reflectancia de la superficie Banda 1-2

## 2) Índices de vegetación MOD13

El producto MOD13 presenta dos índices de vegetación (IV): El índice de vegetación de diferencia normalizada (NDVI) y el índice de vegetación mejorado (EVI). El NDVI se produce mediante el cálculo del cociente normalizado de la banda del infrarrojo cercano y la banda del rojo. Artículos interesantes con información de las características de este producto se presentan en el Cuadro 5.

$$NDVI = \frac{R_{IRC} - R_R}{R_{IRC} + R_R} \quad NDVI = \frac{R_{b2} - R_{b1}}{R_{b2} + R_{b1}} \quad (1)$$

Donde, RIRC y RR son respectivamente los valores de reflectancia bidireccional de la superficie para las bandas del infrarrojo cercano y rojo. Este índice da continuidad al NDVI obtenido con base en el sensor AVHRR para extender el registro de datos a mayor plazo en los estudios de monitoreo.

El índice mejorado de vegetación (EVI), se optimizó a partir del NDVI aumentando la sensibilidad en las regiones de alta biomasa y mejoró el monitoreo de la vegetación mediante la disociación de la señal del fondo del dosel y una reducción de la influencia atmosférica. La ecuación utilizada para su obtención se indica a continuación:

$$EVI = G \frac{R_{IRC} - R_R}{R_{IRC} + C_1 R_R - C_2 R_A + L} \quad EVI = G \frac{R_{B2} - R_{B1}}{R_{B2} + C_1 R_{B1} - C_2 R_{B3} + L} \quad (2)$$

Donde, RIRC, RR y RA son respectivamente los valores de reflectancia bidireccional de la superficie para las bandas del infrarrojo cercano, del rojo y del azul con una corrección de los efectos de la atmósfera (Absorción por ozono y difusión de Rayleigh). Hay que recordar que la banda azul se toma a una resolución de 500 m y se remuestrea a 250 m para coincidir con la banda del IRC y la del rojo, lo que pudiera representar una fuente adicional de error.

**L** es un ajuste del fondo del dosel que toma en cuenta la transferencia radiante diferencial del infrarrojo cercano y el rojo a través del dosel,

**G** es un factor de ganancia, y

**C1, C2** son los coeficientes de resistencia de aerosoles, que usan la banda azul para corregir la influencia del aerosol en la banda roja. Los coeficientes adoptados en el algoritmo EVI son L=1, C1=6, C2 = 7.5, y G = 2.5.

**Cuadro 5 - Referencias claves para el producto MOD13**

Clave del producto	MOD13
Liga y fecha del documento sobre el algoritmo de elaboración.	<a href="http://modis.gsfc.nasa.gov/data/atbd/atbd_mod13.pdf">http://modis.gsfc.nasa.gov/data/atbd/atbd_mod13.pdf</a> (1999)
Liga y fecha de la guía del usuario	<a href="http://tbrs.arizona.edu/project/MODIS/UserGuide/doc.php">http://tbrs.arizona.edu/project/MODIS/UserGuide/doc.php</a> (fecha no especificada)
Artículo(s) clave(s) sobre la elaboración	Huete <i>et al.</i> 1997 y 2002
Artículo(s) clave(s) adicional(es): nuevas versiones, aplicaciones	Glenn <i>et al.</i> 2008, Huete <i>et al.</i> en prensa.

El algoritmo para elaborar los compuestos de 16 días opera para cada píxel, y consiste en seleccionar las “mejores” observaciones dentro de un número determinado de tiempo. Debido a los traslapes de la órbita del sensor existe un número importante de observaciones en el ciclo de observación de 16 días. Sin embargo, debido a las distorsiones en el tamaño del píxel descritas en el apartado 2.2.1, el número de píxeles aceptable en un compuesto de 16 días generalmente se reduce a menos de 10 e incluso menos de 5 si consideramos una media global de cobertura de nube de 50-60%. Una vez colectados los datos, el algoritmo aplica un filtro basado en la calidad, presencia de nubes y geometría de observación y solo los píxeles de mayor calidad se toman en cuenta para la elaboración del compuesto (Van Leeuwen et al. 1999).

Existen dos métodos operacionales para elaborar el compuesto: 1) MVC: compuesto de valor máximo y 2) CV-MVC: filtro por ángulo de visión seguida de un compuesto de valor máximo. La técnica empleada depende del número y la calidad de las observaciones. El valor máximo es una metodología similar a la usada con los datos AVHRR, donde el valor que se toma es el del valor máximo presentado por el NDVI. Es eficiente con datos no corregidos por contaminación atmosférica (Holben 1986). El defecto del MVC es de favorecer la selección de píxeles con ángulos de observación fuera del nadir. En el caso de MODIS, los efectos de anisotropía de la superficie son más pronunciados ya que los valores de reflectancia están previamente corregidos antes de la composición y el cálculo del índice de vegetación, lo cual aumenta dramáticamente la selección de píxeles fuera del nadir sobretodo en dosel abierto que muestra mayor valor del NDVI cuando se ven de manera oblicua.

El método CV-MVC se diseñó para limitar las grandes variaciones angulares encontradas por el MVC. El CV-MVC compara los dos valores más grandes del NDVI y selecciona la observación más cercana a la del nadir para representar el ciclo del compuesto de 16 días. En la versión 5, el método CV-MCV utiliza también un criterio contextual para descartar los píxeles vecinos que presentan una importante variación temporal con el fin de aumentar la cohesión espacial de la imagen.

El producto MOD13 contiene además de los valores del NDVI y del EVI, las reflectancias de las bandas del rojo (banda 1), infrarrojo cercano (banda 2), infrarrojo medio (banda 6) y del azul (banda 3) así como el ángulo de visión y de inclinación solar, la fecha de toma de los píxeles seleccionados del compuesto y la capa de información sobre calidad. Esta capa brinda un índice

de calidad general y contiene información sobre la presencia de aerosoles, nubes, nieve así como la ejecución de correcciones atmosféricas. Este producto se encuentra disponible en varias resoluciones temporales y espaciales como se indica en el Cuadro 6.

**Cuadro 6 - Productos de índice de vegetación (Incluye EVI y NDVI)**

Clave	MOD13Q1	MOD13A1	MOD13A2	MOD13A3	MOD13C1	MOD13C2
<b>Plataforma</b>	Terra Aqua	Terra Aqua				
<b>Resolución temporal</b>	16-días	16 días	16 días	Mensual	16 días	Mensual
<b>Resolución espacial</b>	250m	500m	1 km	1 km	0.05 grados	0.05 grados
<b>Extensión espacial</b>	Tile	Tile	Tile	Tile	CMG	CMG
<b>Nivel</b>	L3	L3	L3	L3	L3	L3
<b>Fase de validación</b>	2	2	2	2	2	2
<b>Versión</b>	5	5	5	5	5	5
<b>Tamaño del archivo (MB)</b>	5-270	2-77	1-22	0.15-20	35.5	19
<b>Número de capas</b>	12	12	12	11	13	13
<b>Capa(s) más importante(s)</b>	NDVI, EVI	NDVI, EVI				
<b>Fecha de inicio de producción (Terra y Aqua)</b>	24/2/2000	24/2/2000	24/2/2000	18/2/2000	18/2/2000	18/2/2000
	14/7/2002	14/7/2002	14/7/2002	1/7/2002	14/7/2002	1/7/2002

En resumen podemos decir que en las últimas dos décadas ha aumentado significativamente el número y tipo de datos derivados de la percepción remota, buscando lograr representar fielmente a la superficie terrestre. Los productos del sensor MODIS enfocados al monitoreo terrestre representan muchas ventajas ya que surgen como resultado de la integración de conocimiento experto y multidisciplinario con una alta disponibilidad y el respaldo de una comunidad científica especializada.

Los datos MODIS han llamado la atención de científicos en diferentes partes del mundo. Estos realizaron estudios para evaluar su potencial para el monitoreo de las cubiertas del suelo a escalas desde globales hasta locales. Se presenta una revisión de estos estudios en el capítulo siguiente.

## **CAPÍTULO III. ANTECEDENTES**

---

Las características que presentan los datos derivados del sensor MODIS, sugieren un gran potencial para su implementación en el monitoreo de la cobertura terrestre. Razón por la cual se han realizado diversos estudios con el objetivo de probar las capacidades de los datos MODIS en diferentes regiones y analizados bajo diferentes metodologías.

En preparación para este estudio se realizó una revisión exhaustiva en revistas reconocidas en el campo de la percepción remota, principalmente en el *International Journal of Remote Sensing* y *Remote Sensing of Environment* de donde se seleccionaron los trabajos basados exclusivamente en datos MODIS en el monitoreo del uso y la cobertura terrestre

Con el objetivo de determinar y comprender los alcances del trabajo realizado anteriormente en este capítulo se presenta el tipo de datos MODIS que se usan más frecuentemente, como se preprocesan, que técnicas se usan para clasificarlos y medir su fiabilidad y cuáles los resultados que se han obtenido.

### **3.1. SELECCIÓN DE LOS DATOS**

#### **3.1.1. Estudios basados en índices de vegetación.**

El sensor MODIS cuenta con un gran número disponible de productos que pueden ser utilizados en el mapeo del uso y la cobertura del suelo, sin embargo, la mayoría de los autores se inclinan por el uso de series de tiempo basadas en el índice de vegetación normalizado (NDVI por sus siglas en inglés). El NDVI se utiliza de manera más frecuente que otros índices principalmente por la continuidad que da a los estudios basados con imágenes AVHRR, facilitando su comparación con otros estudios (Gitelson y Kaufman 1998).

La razón fundamental para utilizar series de tiempo es el aporte que representa la fenología característica de cada tipo de vegetación, con lo que aumenta la posibilidad de distinguirla espectralmente de otra en una clasificación. (Townshend et al. 1991, Defries et al. 1995, Ji y Peters 2007).

Los datos aportados por series de tiempo de NDVI MODIS han demostrado ser exitosos en estudios de discriminación de cultivos en el Norte América, Europa y Asia (Mingwei et al. 2008).

Mientras que en algunos estudios resulta favorable ampliar la serie de tiempo a más de un año de datos para retener suficiente información después de eliminar el ruido (Tottrup 2007), en algunos otros se usan series de tiempo parciales, restringiendo sus datos a una época particular de interés como puede ser la época de crecimiento, o de cosecha (Bagan et al. 2005, Sivanpillai y Latchininsky 2007), o periodos libres de nieve (Heiskanen y Kivinen 2008).

Una desventaja del NDVI es que la señal llega a saturarse en condiciones de alta biomasa (Huete 1988), por lo que el uso de otros índices de vegetación ha ido en aumento. Entre ellos el que más destaca es el índice de vegetación mejorado o EVI por sus siglas en inglés. EVI fue diseñado con el propósito de optimizar la señal de la vegetación en aquellas áreas susceptibles a saturar la señal del NDVI (Huete et al. 2002), además de ser más sensible a las variaciones en la cobertura de vegetación verde (Justice et al. 1998, Ferreira et al. 2003). Xavier et al. (2006) observaron este aumento en la sensibilidad que se debe particularmente a la fenología y al tipo de práctica de manejo de la tierra.

Diferentes estudios de la vegetación han demostrado que ambos índices pueden ser utilizados en conjunto para aprovechar las ventajas de cada uno, mejorando la detección de cambios así como la extracción de parámetros biofísicos del dosel (Huete et al. 2002).

También ha sido demostrado que el índice de vegetación ajustado al suelo o SAVI por sus siglas en inglés, tiene la capacidad de minimizar exitosamente las variaciones del suelo en el dosel de los pastizales y cultivos de algodón. Además cuenta con la ventaja de que puede ser refinado o calibrado para que las variaciones del sustrato se normalicen de una manera eficiente y así evitar un efecto adverso en las mediciones de la vegetación (Huete 1988).

Otros índices de vegetación de uso menos común pueden ser útiles en estudios específicos, Gitelson et al. 2005, sugirieron por primera vez el uso del índice verde o GI por sus siglas en inglés en datos MODIS. El GI es un índice simple calculado como:  $GI = \frac{pnir}{pgreen}$ . De acuerdo a los autores, el GI cuenta con una buena sensibilidad, al tiempo que evita la saturación (Gitelson et al. 2003). Ozdogan y Gutman et al. (2008) probaron que el GI obtiene mejores resultados en el monitoreo subpixel de cultivos de riego.

### **3.1.2. Estudios basados en datos de reflectancia de la superficie**

Los datos de reflectancia de la superficie se pueden considerar como los datos más puros, ya que representan solo la información captada por el sensor sin ninguna transformación, a diferencia de los índices de vegetación. A pesar de que son potencialmente útiles para el monitoreo terrestre, su uso es infrecuente y generalmente usado con el objetivo de efectuar comparaciones contra resultados obtenidos a partir de índices de vegetación. Se usa más frecuentemente como compuesto de 8 días (Carrão et al. 2008).

Los compuestos de 8 días son una nueva imagen generada a partir de los “mejores píxeles” de las 8 imágenes de una semana. De acuerdo a algunas investigaciones preliminares, la diferencia del ángulo de vista causa que pixeles adyacentes tengan diferencias radiométricas en el mismo tipo de cobertura, conduciendo a errores de clasificación (Cihlar 2000).

Para evitar estas dificultades, Wessels et al. (2004) Chen y Rao (2009) y Shimabukuro (2009) usaron datos diarios libres de nube para hacer estudios de mapeo regional, de la misma manera Sedano et al. (2005) usaron imágenes de un solo día en diferentes periodos de la época seca en África.

### **3.1.3. El uso de combinaciones de datos MODIS**

De acuerdo a algunos autores, la incorporación de diferentes tipos de datos MODIS puede aumentar la fiabilidad de los mapas. Por ejemplo, para generar el producto MOD12 de cobertura y dinámica del suelo, Friedl et al. (2010) agregaron a las bandas 1-7 la información del EVI y del producto de temperatura de la superficie de MODIS, usando además en el cálculo del EVI el producto Nadir BRDF-Adjusted Reflectance (NBAR).

En determinados tipos de vegetación, el uso de más de un tipo de dato es particularmente útil, al resaltar la respuesta espectral de la misma. Westra y De Wulf (2007) usaron información del infrarrojo medio, infrarrojo cercano y un índice basado en estos dos al que llamaron el “índice de diferencia normalizada del agua”. Combinándolo con el NDVI y el EVI obtuvieron mejores resultados que tomando en cuenta una sola variable.

Otra aproximación es combinar datos de más de un sensor, por ejemplo, Braswell et al. 2003 unieron los datos del sensor Multi-angle Imaging SpectroRadiometer (MISR) y las bandas múltiples de onda corta de MODIS. Mediante esta sinergia al aprovechar el muestreo simultáneo

de estos dos conjuntos de datos obtuvieron un aumento de 20% y 3% en el valor de correlación y de fiabilidad promedio respectivamente. Acerbi-Junior et al. (2006) realizaron una fusión entre imágenes MODIS y Landsat usando la transformada wavelet o transformada óndula que es un tipo especial de transformada de Fourier que representa una señal en términos de versiones trasladadas y dilatadas de una onda finita (denominada óndula madre).

#### **3.1.4. Tres dimensiones de datos**

Son pocos los autores que intentan comparar las tres dimensiones de datos (temporal, espectral y angular) obteniendo resultados distintos. Carrão et al. (2008) encontraron que la información espectral es más útil en la separabilidad de las clases que la información temporal. En su estudio observaron que la fiabilidad aumenta en la medida que se incorporan más fechas, pero que la contribución de este efecto desaparece al explotar la información de las 7 bandas. En cambio, aún teniendo información de todo un año de datos MODIS siempre fue necesario contar con por lo menos tres bandas espectrales para obtener una fiabilidad satisfactoria.

Por otro lado, Heiskanen and Kivinen (2008) demostraron que la información multitemporal y angular aumenta la fiabilidad en el mapeo forestal. Sus resultados sugieren que la estacionalidad afecta el desempeño del modelo, de tal manera, los resultados de finales de la primavera y principios del verano aportan resultados superiores en comparación con los obtenidos a mediados y finales del verano.

#### **3.1.5. Incorporación de datos auxiliares**

Las similitudes espectrales de muchos tipos de cobertura de suelo presentan un reto al intentar diferenciarlas usando solo información espectral. Este problema aumenta al bajar la resolución espectral del sensor, como sucede al usar datos MODIS.

Varios estudios han intentado usar otras fuentes de información para aumentar la fiabilidad final de los mapas. Para distinguir los tipos de vegetación con respuestas espectrales similares pero ubicadas en diferentes regiones, es útil tomar en cuenta las condiciones ecológicas. Los datos auxiliares más utilizados son aportados por el modelo digital de elevación, seguido por mapas de climas. La pendiente es otro dato útil y fácil de obtener a partir del modelo digital de elevación. Miettinen et al. (2008) usaron datos de elevación como datos auxiliares para separar

12 clases en el sureste de Asia. Friedl et al. (2010) utilizaron datos auxiliares y una combinación de datos MODIS para crear la colección 5 de MOD12 estos datos, incluyen la reflectancia, LST, BRDF y EVI, así como información derivada de mapas previos. Zhang et al. (2008) usaron datos

**Cuadro 7 - Los diferentes esfuerzos de mapeo y datos de entrada revisados en esta tesis**

<b>Autor(es) (fecha)</b>	<b>Tipo de mapeo</b>	<b>Productos MODIS</b>	<b>Datos auxiliares</b>	<b>Ubicación del área de estudio</b>
Mingwai <i>et al.</i> (2008)	Discriminación de cultivos	Serie de tiempo de NDVI	No	Norte de China
Tottrup (2007)	Fracciones de bosque	Serie de tiempo de NDVI	Datos de Elevación	Sureste de Asia
Bagan et al. (2005)	Cobertura de suelo	Serie de tiempo parcial EVI	No	Área árida y semiárida del norte de China
Heiskanen y Kivinen (2008)	Silvicultura	MOD12	No	Finlandia
Xavier et al. (2006)	Cultivos de azúcar	Serie de tiempo de EVI	No	Sao Paulo, Brasil
Ozdogan y Gutman (2008)	Mapeo de áreas de agricultura e riego.	Índice verde	Índices basados en Clima, mapas de humedad, mapas de áreas cultivadas.	USA
Carrão et al. (2008)	Cobertura de suelo	Compuestos de reflectancia	No	Portugal
Sedano et al. (2005)	Cobertura de suelo	Imágenes de reflectancia diarias	Mapas de cobertura, datos de elevación	Sabana arbolada de miombo, Zambesia, Mozambique.
Friedl et al. (2010)	Cobertura y dinámica de suelo	EVI, NDVI, y compuestos de reflectancia	LST BRDF NBAR	Global
Westra y De Wulf (2007)	Uso y cobertura de suelo	NDVI, EVI	Índice normalizado de humedad	Área semiárida de África
Zhang et al. (2008)	Cobertura de suelo	Compuestos de reflectancia, NDVI cubierta de nieve, LST BRDF	Datos de elevación Textura / mascara de agua	Planicie del norte de China
Braswell et al. (2003)	Cobertura de suelo	VIS NIR SWIR	MISR	Amazonas de Brasil
Matsuoka et al. (2007)	Cobertura de suelo	Compuestos de reflectancia	MOD10A2	Este de Asia
Miettinen et al. (2008)	Cobertura de suelo	Compuestos de reflectancia	Datos de elevación y mapas de turberas.	Península de Malasia e islas de Sumatra, Java y Borneo
Sivanpillai y Latchininsky (2007)	Mapeo de hábitats de langosta	Compuestos de reflectancia	No	Delta de Amudarya, Uzbekistan
Wessels, K. J., R. S. De Fries, et al. (2004).	Mapeo de cobertura regional	Imágenes de reflectancia diarias	No	Yellowstone
Shimabukuro et al. (2009)	Mapeo de áreas quemadas	Imágenes de reflectancia diarias.	No	Oeste del Amazonas Brasil.

de textura, elevación del terreno, índices de vegetación, cobertura de hielo y nieve, y temperatura de la superficie, superando los resultados obtenidos por la NASA en el desarrollo MOD12 con base en la colección 4. Ozdogan and Gutman (2008) usaron un mapa potencial de riego derivado de datos de clima y radiancia para mapear áreas irrigadas a partir de datos MODIS. En el Cuadro 7 se resumen los diferentes esfuerzos de mapeo y datos de entrada revisados en esta tesis.

### **3.2. PREPROCESAMIENTO DE LOS DATOS**

A pesar de que el preprocesamiento de los datos es poco mencionado en las investigaciones de uso y cobertura del suelo, este es un paso crítico en el análisis cuando se utilizan datos de satélite. La mayoría de los autores escogen metodologías que han sido probadas en estudios anteriores aun cuando al cambiar el área de estudio los resultados pueden diferir.

Por la alta resolución temporal así como el gran número de bandas y productos, los conjuntos de datos MODIS usados en el mapeo de la cobertura cuentan con una alta dimensionalidad, por lo cual están propensos a presentar el efecto Hughes (1968), que se define como un descenso en la fiabilidad de la clasificación cuando el número de bandas excede un número determinado para un tamaño de muestra fijo. Una reducción en el número de bandas minimiza este problema, aumentando así la fiabilidad de la clasificación (Hughes 1968, Landgrebe 2003).

Para disminuir los efectos de correlación así y de la alta dimensionalidad, es importante contar con una forma de reducción de la base de datos como una fase del pre-procesamiento. Un segundo objetivo del pre-procesamiento es reducir el ruido. La elaboración de compuestos de 8 o 16 días se puede considerar como una forma de reducción de ruido, donde se eliminan los píxeles contaminados con nube o con ángulos de visión oblicuos. Aun después de la reducción de ruido, pueden quedar datos con un alto grado de contaminación, en especial cuando se trata de áreas con lluvias frecuentes. En estos casos, se implementan procedimientos adicionales como el filtrado de series de tiempo. Los métodos de pre-procesamiento no pueden dividirse estrictamente en reducción de ruido o reducción dimensional, debido a que muchos de estos cumplen con ambos propósitos. Algunos métodos de análisis de series de tiempo como son la transformada ondulada y la de Fourier también pueden considerarse como métodos de extracción y reducción de ruido.

### 3.2.1. Métodos de selección de los datos

Una manera sencilla de discriminar bandas se hace con base en su información de calidad, la cual está incluida como una banda adicional en cada producto (QA). MODIS proporciona la información de la calidad tanto a nivel de producto como pixel a pixel. Un método de selección puede servir para descartar conjuntos completos de datos (por ej. la imagen completa) o grupos de pixeles específicos dentro de la escena. Un ejemplo es el trabajo de Sedano (2008), quien seleccionó imágenes MODIS libres de nube y humo reduciendo un año de datos a solo 5 fechas.

Algunos autores hacen máscaras para remover pixeles específicos bajo un umbral pre establecido en lugar de un filtro basado solamente en la información de QA. Como ejemplo, Xiao et al. (2005) observaron que ciertos pixeles con un alto contenido de nube no fueron descartados con base en el QA, por lo que aplicaron un filtro adicional descartando los pixeles con reflectancia en el azul equivalente a 0.2. El resultado es una imagen con espacios vacíos que después puede ser llenados mediante otros métodos como el filtrado de tiempo.

Cuando el resultado del filtrado de calidad aún contiene información redundante, se pueden utilizar métodos complementarios como son las medidas de separabilidad. Estas medidas permiten hacer una evaluación de la distancia estadística entre las muestras de las diferentes categorías de cobertura de suelo en un espacio definido por un determinado número de bandas.

El cálculo de las distancias estadísticas entre cada par de categorías para cada uno de los posibles subconjuntos de bandas permite determinar cuál es el número óptimo de bandas de entrada (el mínimo número de bandas del sub conjunto que permiten la máxima o mejor separabilidad), así como la mejor combinación de las mismas. Dentro de los índices de separabilidad más populares se encuentran las medidas de distancia de Bhattacharyya, Mahalanobis, y Jeffreys-Matusita, así como las medidas de divergencia y divergencia transformada (Landgrebe 2003).

Carrão et al. (2008) utilizaron la mediana de Mahalanobis como criterio para ordenar las características temporales y espectrales en su estudio de clasificación de cobertura de suelo en Portugal.

### 3.2.2. Método de extracción de datos

Cuando el conjunto de datos de entrada es muy grande para ser procesado de manera eficiente, y además se sospecha que pueda ser altamente redundante, debe de aplicarse un método de extracción de características. El resultado es una representación reducida del conjunto original de datos y se espera que sea de la información más relevante.

El método más común de extracción en percepción remota es el análisis de componentes principales (ACP) que permite la concentración de la información más útil en los primeros componentes que no están correlacionados. Ferreira et al. (2004) usaron ACP en datos similares a MODIS.

Cuando los tres primeros componentes que alcanzaron 99.5% del total de la varianza total, se consideraron de manera simultánea, aproximadamente el 91% del total del conjunto de datos se clasificó correctamente. Chen (2005) aplicó 3 ACP a datos MODIS de reflectancia de 250, 500 y 1000 m de resolución espacial respectivamente y encontró que: en los datos de 250 m más del 95% de la varianza se encontró en el primer componente, mientras que en los datos de 500 y 1000 m se puede reducir a dos o tres componentes con una pérdida del 10% de la varianza total en los diferentes sitios evaluados. Sedano et al. (2005) usaron ACP de una manera distinta, aplicándolo como una técnica de detección de cambios para realzar los cambios estacionales de diferentes tipos de vegetación en su área de estudio.

Otra metodología que vale la pena mencionar es la transformación Tasseled Cap (TCT). Esta metodología inicialmente se desarrolló para imágenes MSS, inspirada en ACP, presentando varias ventajas respecto a este, TCT proporciona índices físicos, concretos y fácilmente interpretables (como son el brillo, el verdor y la humedad). Es importante tomar en cuenta que los coeficientes se determinan de manera empírica, y son válidos solo para un tipo de sensor. TCT se utilizó en principio con Imágenes Landsat y posteriormente, Lobser and Cohen (2007) los adaptaron para MODIS, demostrando que TCT tiene un gran potencial en el análisis en los datos MODIS para propósitos de evaluación de la cobertura del suelo.

### 3.3. FILTRADO Y ANÁLISIS DE SERIES DE TIEMPO

La alta resolución temporal del sensor MODIS, es una característica interesante que permite producir series de tiempo, estas se hacen generalmente a partir de índices de vegetación. El ordenamiento de las series de tiempo, permiten un análisis distinto al de otro tipo de datos (p.ej. bandas espectrales) donde se pueden aplicar herramientas que consideren las variaciones estacionales. Este tipo de técnicas han sido aplicadas en series derivadas de datos AVHRR y se basan principalmente en filtros de tiempo.

Los filtros de tiempo consisten en suavizar los valores aberrantes de fechas específicas tomando en cuenta los valores previos y posteriores de la serie. Este suavizado puede hacerse tan simple como mediante el cálculo de la media, hasta métodos muy sofisticados de ajuste y se pueden aplicar a todos los datos o solamente a observaciones puntuales detectadas como ruido. Los métodos más usados de ajuste de las series de tiempo son el filtro de media móvil, la gaussiana asimétrica, la doble logística y el filtro adaptado de Savitzky–Golay.

El filtro de media móvil consiste en el cálculo de un valor suavizado de una observación promediando los valores anteriores y posteriores. El tamaño de la ventana (número de observaciones previas y posteriores que se toman en cuenta) determina el efecto de suavizado. Una alternativa es usar la mediana en lugar del promedio. El filtro adaptado de Savitzky–Golay utiliza funciones locales polinomiales en el ajuste, éste es un filtro paso bajo, sin embargo no reduce tanto los valores extremos como ocurre en el método de los promedios, lo cual, por un lado permite capturar cambios sutiles y rápidos en la serie de tiempo, pero también lo hace muy sensible al ruido. Los métodos tanto gaussiana asimétrica y doble logística son semi-locales, esto provoca que sean menos sensibles al ruido y den una mejor descripción en el inicio y final de las estaciones (Jonsson y Eklundh 2002).

Con base en datos MODIS, Boschetti et al. (2009) usaron el método de Savitzky–Golay para suavizar series de tiempo de NDVI. Mingwei et al. (2008) aplicaron el mismo método para remover pixeles marcados como nube en la capa de calidad MODIS. Los resultados muestran que, aunque este filtro compensa hasta cierto punto la pérdida de información de los espacios de la imagen, los valores generados no son equivalentes a la información original. Para evitar sesgos causados por contaminación en la imagen, Van Dijk et al. (1987) sugirieron que un filtrado promedio de ventana móvil es superior comparado contra los mejores ajustes de otras técnicas de

suavizado (como los filtros lineares o polinomiales), ya que remueven medidas aberrantes de la serie de tiempo. Esta técnica fue empleada también por Carrão et al. (2008) con datos MODIS.

### **3.3.1. Análisis de Fourier**

Un método usado frecuentemente es el análisis de Fourier, en el cuál se toma en cuenta la amplitud y ángulo de las imágenes (Jakubauskas et al. 2001). Este método emplea una variedad de aproximaciones que incluyen; perfiles temporales de fenología manifestados en el NDVI (Defries et al. 1995, Reed et al. 1994), clasificaciones de datos multitemporales (Brown et al. 1993, Loveland et al. 1995), análisis de componentes principales (Townshend et al. 1987, Tucher et al. 1985), y análisis de series de tiempo de un perfil temporal de NDVI estandarizado como componente principal (Eastman and Fulk 1993).

## **3.4. MÉTODOS DE CLASIFICACIÓN MAS UTILIZADOS**

Los tipos de clasificación pueden dividirse en dos grandes grupos: clasificaciones “duras” y “difusas”. Este último enfoque es particularmente interesante para intentar compensar la generalización de la información de las imágenes de baja resolución espacial. Esta aproximación ha sido utilizada con datos MODIS por Ozdogan y Gutman (2008), Braswell et al. (2003) y Totrup et al. (2007).

A pesar de los avances en las metodologías de análisis de datos en las últimas décadas, algunos autores todavía utilizan métodos convencionales, por ejemplo Mingwei et al. (2008) utilizan la Máxima Verosimilitud para mapear diferentes tipo de cultivos. Un buen número de autores utilizan el método de Máxima Verosimilitud solo como referencia para comparar los resultados obtenidos a partir de otros clasificadores. (Bagan et al. 2005, Westra y De Wulf 2007, Chen y Rao 2009).

Otro tipo de clasificadores cuyo uso se ha popularizado en el análisis de datos son diferentes tipos de redes neuronales. Algunos de estos métodos probados en datos MODIS incluyen self-organizing map (SOM) (Bagan et al. 2005), Fuzzy ARTMAP (Borak y Strahler 1999, Sedano et al. 2005), LNNS, simulador de redes neuronales (Westra 2007. Bagan et al.

2005, Borak y Strahler 1999). Estos autores reportan un aumento en la fiabilidad final de los mapas en comparación con la Máxima Verosimilitud.

Los árboles de decisión son otro algoritmo que también se ha probado con datos MODIS. Algunos ejemplos son: Matsuoka et al. (2007) quien usó este método para distinguir 11 tipos de cobertura en el este de Asia y Totrupp (2007) quienes usaron un algoritmo de regresión para generar un modelo que predice una sola variable de respuesta. De acuerdo a algunos autores los árboles de decisión permiten obtener mejores resultados que otras metodologías. Chen y Rao (2009) reportaron que los árboles de decisión superan los resultados obtenidos a partir de Máxima Verosimilitud (con valores de fiabilidad global de 64.5% contra 48.9% respectivamente) al clasificar la cobertura de suelo a partir de datos de reflectancia MODIS.

Algunos autores utilizan métodos menos convencionales, Heiskanen and Kivinen (2008) utilizaron un modelo de regresión binomial generalizada (GLM) para estimar cobertura de dosel. Shimabukuro et al. (2009) utilizaron segmentación y agrupamiento mediante una clasificación no supervisada. Gonçalves et al. (2006) y Carrão et al. (2008) utilizaron el método Support Vector Machine.

### **3.5. MEDICIÓN DE LA FIABILIDAD**

El cálculo de fiabilidad es una tarea esencial en la fase final de la clasificación. Sin embargo, está muy limitada por los recursos disponibles, y sigue siendo un reto especialmente para los productos globales y regionales (Defries et al. 1998, Cihlar 2000). Es importante contar con un conjunto de datos de referencia robusto para el cálculo de la fiabilidad del mapa final.

El método más común para evaluar la fiabilidad final obtenida de los mapas MODIS es la comparación contra clasificaciones obtenidas a partir de datos de mejor resolución espacial (Friedl et al. 2002, Latifovic et al. 2004), otros mapas de cobertura (Liu et al. 2003, Boles et al. 2004, Han et al. 2004, Giri et al. 2005) o a un subconjunto de los datos de entrenamiento (validación cruzada) (Friedl et al. 2010). El uso de censos y trabajo de campo en la validación de clasificaciones MODIS es menos común debido, por un lado, a la escasez de este tipo de datos y por otro a la dificultad de comparar datos detallados con imágenes de baja resolución. Para superar este problema de escala, algunos autores optan por el remuestreo de los datos de referencia para hacerlos comparables de alguna manera. En el caso de trabajo de campo, donde

se manejan datos puntuales, una alternativa es utilizar una aproximación difusa, en la cual, cuando el área sea heterogénea o se encuentre fragmentada, el dato de campo considere más de una clase que puede ser tomada como correcta en la evaluación final del mapa.

Otra consideración importante para evitar sesgos en la evaluación de la fiabilidad de los mapas, es asegurar que los datos de entrenamiento sean independientes de los datos de referencia, especialmente cuando se utiliza la validación cruzada. Con esta finalidad Friedl et al. (2010) estratificaron su base de datos de entrenamiento en diez sub conjuntos únicos (en lugar de píxeles individuales) para evitar la correlación espacial entre los datos de entrenamiento y los datos de corroboración, con este enfoque se reduce hasta cierto punto la predicción de la fiabilidad, sin embargo, como los sitios de entrenamiento están determinados por polígonos derivados de imágenes de mejor resolución, aún existe un sesgo en las áreas homogéneas, que tienen una mayor probabilidad de resultar calificadas correctamente (Jung et al. 2006).

En el caso de las clasificaciones difusas, tanto los datos de referencia como los del mapa se expresan como variables continuas (p.ej. valores difusos de proporciones estimadas de cobertura contra proporciones verdaderas) y los índices más comúnmente utilizados para evaluar las diferencias entre los valores estimados y reales son: el coeficiente de correlación, la raíz del error medio cuadrático (ambos utilizados por Braswell et al. 2003), y la evaluación del sesgo (Muukkonen y Heiskanen 2007).

En las clasificaciones duras, los índices más utilizados son la fiabilidad global y el coeficiente de Kappa. Unos pocos autores como Cohen et al. (2003) y Heiskanen y Kivinen (2008) proporcionan índices de fiabilidad por clase, tales como la fiabilidad de usuario y productor o la matriz de confusión. La fiabilidad por clase es un dato importante porque el rango de precisión dentro de una clase específica es generalmente amplio (véase, por ejemplo Friedl et al. 2010). Cuando se presenta la matriz de confusión completa, se brinda la oportunidad de aprovechar esta información para la generación de nuevos índices de fiabilidad (Stehman y Czaplewski 1998).

### 3.6. RESULTADOS DE LOS ESTUDIOS REVISADOS

Dependiendo del problema a resolver en cada estudio, se determina el número de categorías de cobertura. Para algunos casos es suficiente con discriminar solo dos categorías, como puede ser; bosque / no bosque (Westra y De Wulf 2007) o agricultura riego / no riego (Ozdogan y Gutman 2008). Otros estudios tienen un número reducido de clases y son raros aquellos que incluyen más de diez. Por ejemplo, Matsuoka et al. (2007) y Borak y Strahler (1999) tomaron en cuenta 14 y 11 categorías respectivamente. Los mapas globales de cobertura terrestre generalmente están basados en el sistema de clasificación de la IGBP con 17 categorías a nivel mundial (Bagan et al. 2005, Friedl et al. 2010, Borak y Strahler 1999).

Como se muestra en el Cuadro 8, la fiabilidad obtenida a partir de los estudios hechos con MODIS varía de un caso a otro. Depende mucho del tipo de clasificación, el número de categorías, y la evaluación de la fiabilidad efectuada. En la mayoría de los casos si se utiliza una clasificación con clases muy generales los resultados varían desde un 74% hasta 91%.

Estos resultados parecen suficientes para la mayoría de las aplicaciones cartográficas. Sin embargo, debido a que se utilizan diferentes bases de datos y enfoques, hay que subrayar que las medidas reportadas de fiabilidad no son comparables y no deben considerarse como estimaciones cuantitativas robustas. Algunos métodos se basan en esquemas de muestreo, mientras que otros en validaciones cruzadas que suelen tener resultados más optimistas. Una alternativa promisoría en la medición de la fiabilidad es la cuantificación de pertenencias o proporciones a través de una lógica difusa (Couturier et al. 2009). Este método fue utilizado para evaluar el mapa MOD12 IGBP de las áreas de bosque en México MOD12 IGBP con una muestra de datos de campo (más de 24,000 conglomerados en 4 sitios de observación) y un mapa de cobertura de suelo (Couturier en prensa). Con base en la opción más estricta (sin tolerancia difusa) y la más flexible, la fiabilidad del mapa es de 54.6% y 65% respectivamente.

**Cuadro 8 -** Tipos de ambiente, número de categorías y fiabilidades obtenidas en estudios con datos MODIS

Autor(es) (fecha)	Esquema de clasificación (número de categorías)	Categorías	Datos de referencia usados en la evaluación de fiabilidad	Fiabilidad obtenida	Área de estudio
Bagan et al. (2005)	IGBP (9)	1. Perenne hojas aciculares 2. Perenne de hojas anchas. 3. Caducifolio de hojas anchas 4. Pastizales 5. Humedales Permanentes 6. Cultivos 7. Urbano 8. Suelo desnudo / ralo 9. Agua	Mapa de cobertura generado de datos Landsat y ASTER	FG = 91.1% K =0.8998 // fiabilidad por clase de 75 a 100%	Varias regiones de China
Borak y Strahler (1999) *	IGBP (11)	1. Perenne hojas aciculares 2. Perenne de hojas anchas. 3. Caducifolio de hojas anchas. 4. Bosque mixto 5. Sabana, 6. Pastizal, Matorral cerrado, 7. Matorral abierto, 8. Cultivos 9. Urbano 10. Suelo desnudo 11. Rocoso	Mapa generado del programa de análisis Gap (GAP)	FG = 78.6%, //fiabilidad por clase de 28 a 100%	Región semi-árida de Arizona USA
Braswell et al. (2003)	Diseñadas para el estudio (3 a 4)	1. Bosque 2. Área clareada 3. Vegetación secundaria	Clasificación Landsat TM	RMSE de 0.01 a 0.27	Región del Amazonas en Brasil
Carrão et al. (2008)	CORINE Cobertura de suelo (9)	1. Áreas artificiales 2. Cultivos irrigados 3. Cultivos de temporal 4. Bosque de hojas anchas 5. Bosque de hojas aciculares 6. Pastizal natural 7. Matorral 8. Suelo desnudo 9. Agua	Mapa de cobertura de CORINE + Landsat y fotos aéreas	FG =90%	Portugal
Chen and Rao (2009)	Diseñadas para el estudio (4)	1. Fango, 2. Pastizal degradado, 3. Cultivos 4. Agua	Interpretación De datos Landsat validada con trabajo de campo	FG= 64.5 K =0.45	
Matsuoka et al. 2007	IGBP (14)	1. Perenne hojas aciculares 2. Perenne hojas anchas 3. Caducifolia de hojas aciculares 4. Caducifolia de hojas anchas, 5. Matorral abierto 6. Pastizales 7. Cultivos (incluye arrozales) 8. Cultivos (sin arrozales) 9. Doble cultivo (incluye arrozales) 10. Doble cultivo (sin arrozales) 11. Campos irrigados 12. Urbano 13. Nieve y hielo. 14. Suelo desnudo. 15. Agua	Comparación basada en pixel comparada con MOD12Q1 Comparación basada en provincias por medio de un mapa de cobertura de suelo a 1 km de resolución y censos.	Fiabilidad por clase de 14-84% Los resultados numéricos contra el mapa de cobertura y censos no se presentan.	Este de Asia

**Cuadro 8 - Continuación**

Autor(es) (fecha)	Esquema de clasificación (número de categorías)	Categorías	Datos de referencia usados en la evaluación de fiabilidad	Fiabilidad obtenida	Área de estudio
Miettinen et al. (2008)	Diseñadas para el estudio (11)	1. Agua 2. Manglar 3. Vegetación hidrófila 5. Bosque bajo 6. Bosque de baja montaña 7. Bosque de alta montaña 8. Plantaciones 9. Mosaico montano 10. Abierto en tierras bajas 11. Urbano	Muestreo aleatorio estratificado contra clasificación SPOT	FG= 82% / K= 0.77	Parte oeste del sureste insular en Asia
Ozdogan y Gutman (2008)	Diseñadas para el estudio (2)	1. Riego 2. No riego	Base de datos de irrigación	Correlación 0.88 entre las áreas estimadas contra las zonas de irrigación reportadas	EUA
Shimabukuro et al. (2009)	Diseñadas para el estudio (1)	Área quemada	Imágenes Landsat clasificación, observaciones de campo	Correspondencia visual, no se presentan datos numéricos	Oeste de Brasil en el Amazonas
Sivanpillai y Latchininsky (2007)	Diseñadas para el estudio (6)	1. Caña 2. Caña y mezcla de arbustos 3. Arbustos 4. Vegetación dispersa, 5. Suelo arenoso 6. Agua.	Fotografía digital datos de campo Imágenes Landsat entrevistas	OA =74% K =0.67	Amurdaya River Delta Uzbekistan
Tottrup (2007)	Diseñadas para el estudio (3)	1. Bosque maduro 2. Bosque secundario 3. No-bosque	Clasificaciones de SPOT, ASTER e IKONOS	Desviación media absoluta entre 14.6% y 21.6%	Sureste de Asia, tierras altas de Vietnam, Laos, Camboya, y Tailandia
Westra y De Wulf (2007)	Diseñadas para el estudio (5)	1. Tierras áridas 2. Área inundada 3. arroz 4. cultivos, 5. cuerpos de agua permanentes	Validación cruzada	K = 0.98	Llanuras de inundación Camerún
Zhang et al. (2008)	Diseñadas para el estudio (8)	1. Cultivos, 2. pastizales densos, 3. pastizales dispersos 4. Bosque de aciculares 5. Bosque de hojas anchas 6. Arbustos, 7. Urbano 8. Agua	Clasificación Landsat	OA 75.5% PCA 35.4- 96.3%	Planicie del norte de China

Como se puede ver en el Cuadro 8, existen diversas maneras de analizar los datos MODIS. Algunas de las metodologías implementadas en los estudios revisados son las mismas que se han utilizado con otros sensores, otras técnicas han sido adecuadas a las características del sensor MODIS y algunas otras son innovadoras. Los resultados y recomendaciones que se presentan en las investigaciones objeto de esta revisión se tomaron en cuenta para seleccionar el conjunto de datos idóneo y conformar la metodología seguida en este trabajo de tesis, considerando no solo aquellas que obtuvieron las mayores fiabilidades, sino también la complejidad de implementación de las técnicas, el número y tipo de clases así como la robustez en la medida de fiabilidad y la factibilidad de comparación con otros estudios.

Es necesario considerar que los resultados obtenidos no dependen únicamente del tipo de dato y la metodología empleada, las condiciones ambientales varían enormemente en diferentes regiones del planeta, lo que añade complejidad a los esfuerzos de mapeo. Es por esta razón indispensable tomar en consideración la información previa sobre cuál ha sido el desempeño de los datos MODIS en regiones altamente diversas como son muchas partes de México. El siguiente capítulo presenta la descripción de las áreas objeto de este trabajo de tesis.

## **CAPÍTULO IV. DESCRIPCIÓN DE LAS ÁREAS DE ESTUDIO**

---

México está considerado por la Organización de las Naciones Unidas como un país megadiverso, ya que aquí se encuentra representada cerca del 10% de la diversidad mundial (Mittermeier y Goettsch 1992).

Esto se debe por un lado, a la gran extensión territorial que lo hacen el decimocuarto país más grande del mundo con un área de 1,972,550 km<sup>2</sup>. México se encuentra en el límite entre dos regiones biogeográficas: la neártica y la neotropical, lo que contribuye a su gran riqueza natural. Además, el paso del Trópico de Cáncer por el país le permite contar con dos zonas climáticas (templada y tropical). Se le suma una topografía altamente compleja determinada por dos grandes sistemas montañosos; la Sierra Madre Oriental y la Sierra Madre Occidental.

Sin embargo, México es también uno de los países que presenta las mayores tasas de deforestación a nivel mundial. Esta situación evidencia la importancia del adecuado manejo, planeación e implementación de estrategias de conservación de los ecosistemas.

### **4.1. CRITERIOS DE SELECCIÓN DE LAS AREAS DE ESTUDIO**

En este estudio se eligieron tres áreas representativas de los ecosistemas de México que incluye, los tipos de vegetación de mayor superficie en el país (Figura 7). Los bosques templados (18% de la cobertura nacional) se encuentran representados en el área de Tancítaro en Michoacán, las selvas húmedas y subúmedas que en su conjunto suman el 16% de la cobertura nacional se representan en el área de Quintana Roo; mientras que los matorrales (26%) se encuentran representados en el área de Mazatán, Sonora. Además en cada una de las áreas encontramos diferentes tipos de coberturas agrícolas y pecuarias (22%).

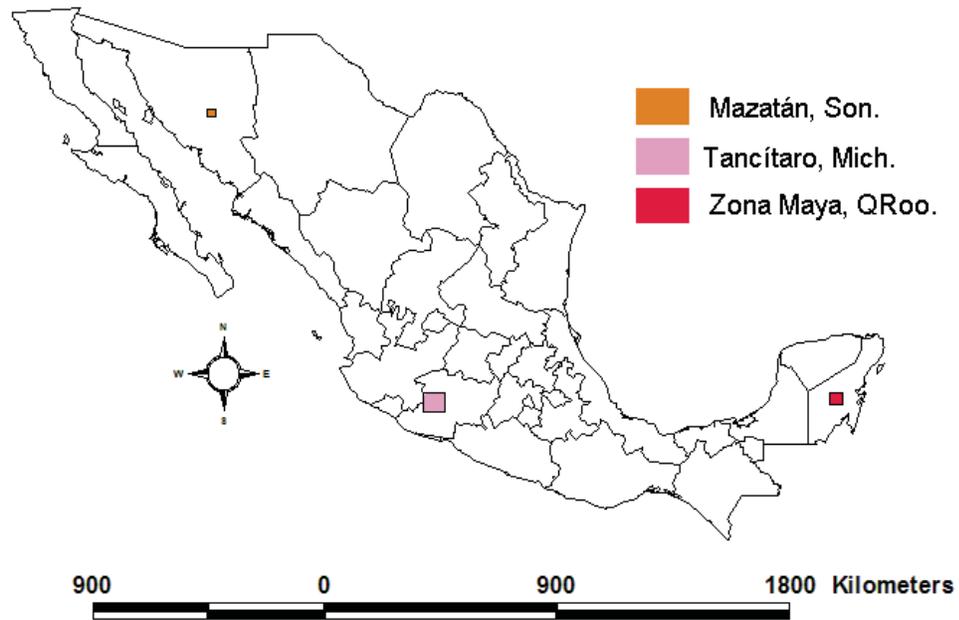


Figura 7 - Localización de las áreas de estudio

Las tres áreas cuentan con distintos grados de manejo y conservación. De acuerdo a INEGI, Michoacán se encuentra dentro de los estados con más cambios en la cobertura vegetal, con menos de 35% de la vegetación original. El área de Mazatán, Sonora, cuenta con un grado de perturbación moderado, mientras que el área Maya es una de las áreas más conservadas del país.

A continuación se presenta una descripción de las condiciones específicas de cada una de las áreas de estudio de acuerdo a los mapas temáticos de INEGI.

## 4.2. DESCRIPCIÓN DE LAS ÁREAS DE ESTUDIO

### 4.2.1. Área del Tancítaro, Michoacán

Esta área de estudio cuenta con una superficie de 4,650 km<sup>2</sup>, se ubica en la región centro-oriental de México (Figura 7), encierra áreas que corresponden a dos provincias fisiográficas del país, la Sierra Madre del Sur y el Eje Neovolcánico Transversal.

Cuenta con un amplio rango altitudinal (600 a 4000 m), con climas que varían desde templado-húmedo, semicálido subhúmedo, semicálido húmedo hasta cálido subhúmedo. La temperatura media anual varía desde los 10°C en el pico de la montaña del Tancítaro hasta los 24°C en las partes bajas mientras que la precipitación media oscila entre los 1000 a los 1500 mm

La cobertura vegetal incluye Selva Baja Caducifolia, Bosque de Pino y de Pino-Encino, además de extensas áreas de agricultura de temporal, agricultura de riego, pastizales y huertas de aguacate (Figura 8). Hay también un área sin vegetación (colada de lava del volcán Parícutín que hizo erupción entre 1943 y 1952) así como ciudades de gran importancia en el estado entre las que destaca Uruapan “la capital mundial del aguacate” con 279,229 habitantes y Apatzingán con 114,837 habitantes.

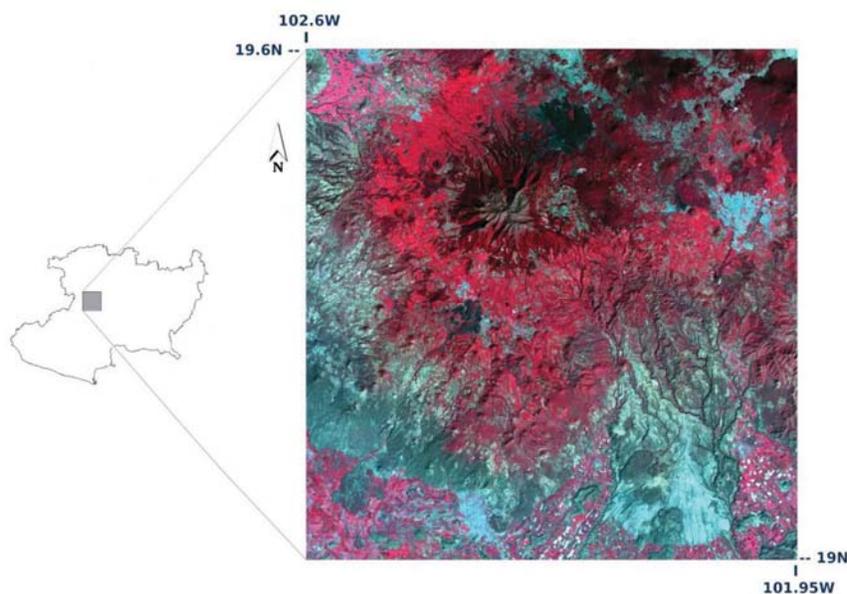


Figura 8 - Área del Tancítaro Michoacán

#### 4.2.2. Área de Mazatán, Sonora

Esta área cuenta con una extensión de 892 km<sup>2</sup>, se encuentra ubicada al noroeste del país en el estado de Sonora (Figura 7), en la provincia fisiográfica denominada llanura sonorenses, y específicamente en la subprovincia sierras y llanuras sonorenses.

El rango altitudinal va desde los 450 m en las zonas bajas hasta los 1,550 m en la parte más alta del cerro “Los buzones”, con climas que varían de semiseco templado a seco semicálido, y temperaturas que oscilan de los 18°C a los 22°C, presentando una precipitación media anual de 400 a 500 mm por año (Figura 9).

El tipo de vegetación predominante en esta área de estudio es el matorral subtropical y mezquital, en la zona de la montaña se encuentra una porción de bosque de encino, hay también algunos pastizales y agricultura de riego. Dentro del área se encuentran pequeños poblados dentro de los cuales el de mayor extensión es Mazatán con 1,743 habitantes.

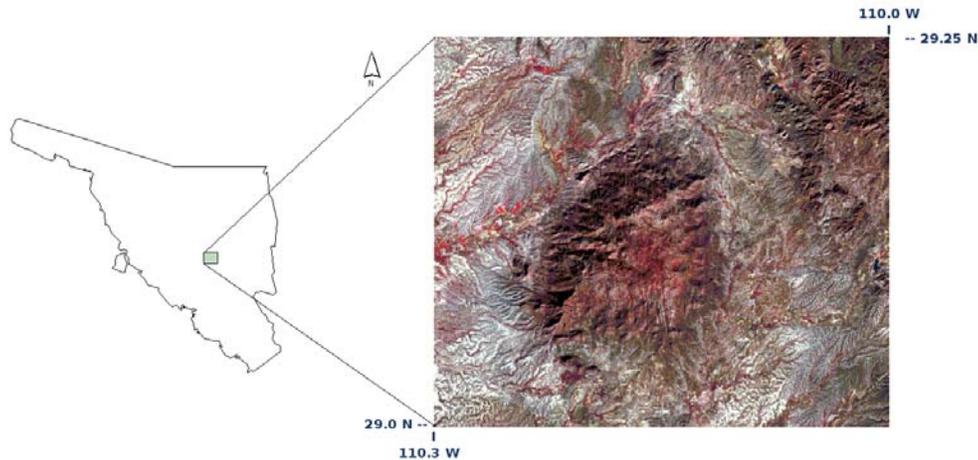
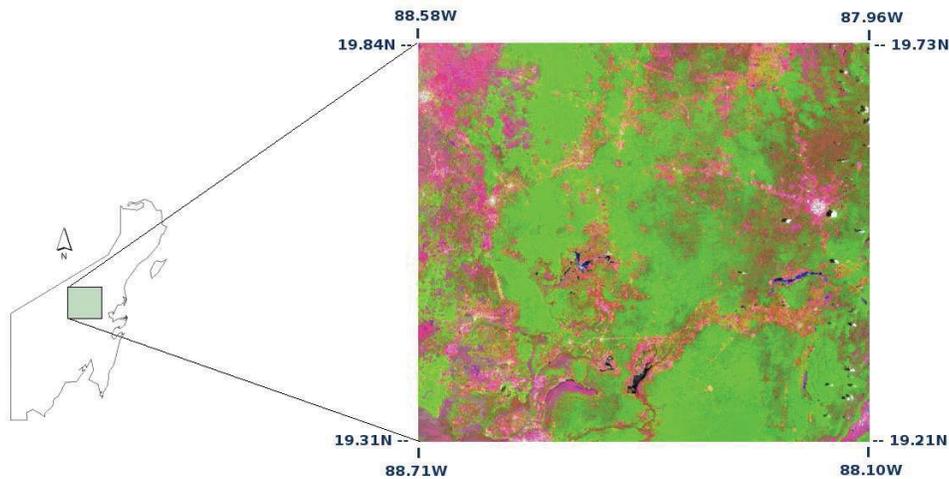


Figura 9 - Área de Mazatán, Sonora

#### 4.2.3. Área Maya, Quintana Roo

Esta área se localiza al sureste de México (Figura 7), abarca una extensión de 3,600 km<sup>2</sup>, se encuentra sobre una planicie de origen tectónico, el relieve es prácticamente plano con algunas colinas de tamaño pequeño y numerosas hondonadas; la altitud promedio es de 10 metros sobre el nivel del mar (msnm).



**Figura 10 - Área Maya en Quintana Roo**

El clima es cálido subhúmedo con lluvias en verano y una temperatura media anual de 26° C. La precipitación pluvial anual oscila entre los 1,300 y los 1,500 mm, con estación de lluvia de marzo a octubre. El clima se ve afectado por los ciclones, que aumentan la precipitación sobre todo en verano.

La vegetación se conforma de Selva Mediana Subperennifolia, particularmente valiosa para la explotación forestal debido a la presencia de maderas preciosas como la caoba y el cedro, así como Selva Baja Subperennifolia. Dentro del área encontramos áreas urbanas pequeñas, siendo la mayor de estas Chunhuhub con 3,928 habitantes (Figura 10).

No obstante la alta riqueza específica, la gran biomasa y la elevada producción primaria bruta, en términos agronómicos convencionales las selvas son improductivas, de ahí la tendencia a destruirlas o modificarlas en favor de agroecosistemas monoculturales simples (Chargoy 1999). Sin embargo, de los tres estados que conforman la península de Yucatán, en Quintana Roo se encuentran las áreas naturales mejor conservadas. Esto se debe al importante y adecuado manejo comunitario y a la existencia de áreas naturales protegidas (Chiappy y Gama 2004).

Se consideró que las tres áreas elegidas para este estudio son lo suficientemente distintas entre sí y que representan adecuadamente la variabilidad de los ecosistemas mexicanos, dándonos confianza en que los datos y métodos con los que se analizaron estas áreas pueden ser aplicables con resultados consistentes en el mapeo nacional. La descripción de los datos y métodos aplicados a los datos que corresponden a estas áreas se presenta a continuación.

## CAPÍTULO V. BASES DE DATOS Y MÉTODOS

---

Este capítulo describe los diferentes tipos de datos utilizados en esta investigación. Se conformó una extensa base de datos que fue analizada de manera cuidadosa en tres pasos principales: preparación de los datos, análisis y evaluación. Cada uno de los apartados se describe de manera desglosada a continuación.

### 5.1. BASE DE DATOS

La base de datos se integró con tres tipos de información: 1) los datos a evaluar derivados del sensor MODIS, 2) los datos auxiliares utilizados para mejorar las clasificaciones y, 3) los datos para medir la fiabilidad de los mapas obtenidos.

#### 5.1.1. Datos MODIS

Existen diferentes productos obtenidos a partir de las imágenes MODIS. En este estudio utilizamos algunos de los productos enfocados al estudio de las cubiertas terrestres (MODLAND) obtenidos con el satélite Terra. La descripción desarrollada de los productos se encuentra en el Capítulo 2 de esta tesis. Los datos MODIS utilizados en el presente estudio fueron, para cada una de las áreas de estudio:

1) Reflectancia de la superficie (MOD09) De este producto se derivaron dos conjuntos de datos evaluados de manera independiente. El primero en su forma de compuesto de 8 días y el segundo como imágenes diarias. En ambos casos, se utilizaron las siete primeras bandas y las bandas 3-5 fueron remuestreadas para coincidir con la resolución de las bandas 1-2 a 250m.

2) Índices de vegetación (MOD13) Se utilizó el índice de vegetación de diferencia normalizada (NDVI) y el índice de vegetación mejorado (EVI) ambos a una resolución espacial de 250 m en compuestos de 16 días. Estos dos índices integraron un solo conjunto de datos analizándose de manera conjunta.

3) De manera adicional se integro una base de datos donde todas las imágenes (Índices de vegetación EVI, NDVI, compuestos de reflectancia de la superficie e imágenes diarias de reflectancia de la superficie) se analizaron de manera conjunta.

Todos los datos cubren un año completo de observaciones (en el área de Tancítaro, y área Maya del año 2002 y en el área de Mazatán del año de 2003). En cada área se tuvieron: 24 composiciones de 16 días de ambos índices de vegetación, 365 imágenes diarias con 7 bandas cada una y, 48 compuestos multiespectrales de 8 días también con 7 bandas cada uno.

Todos los datos MODIS utilizados fueron obtenidos en la última versión disponible (colección 5) de manera gratuita por medio del portal Warehouse Inventory Search Tool (WIST) incluyendo la capa de información que indica, píxel a píxel, la calidad de los datos.

### **5.1.2. Datos de referencia**

Se obtuvo información pre existente para cada una de las áreas de estudio, la cual se utilizó en diferentes etapas y con diferentes propósitos en el proceso de análisis tal como se describe en la metodología. Estos datos incluyeron mapas previos, imágenes de otros sensores, datos de campo e información no espectral.

#### **1. Mapas previos**

Se contó con el mapa del Inventario Nacional Forestal 2000, así como el Mapa de Uso del Suelo y Vegetación del Instituto Nacional de Estadística, Geografía e Informática (INEGI) de 2002, ambos a escala 1 : 250,000.

En el caso de la región Tancítaro, se contó con una clasificación supervisada basada en una imagen Landsat ETM (30 m) del año 2002, mientras que en Mazatán y Quintana Roo se contó clasificaciones visuales muy detalladas basadas en imágenes SPOT del año 2003 y 2002 respectivamente.

#### **2. Datos de campo**

En todas las áreas se contó con las observaciones de campo utilizadas en la verificación del IFN, así como datos de campo de la Comisión Nacional Forestal (CONAFOR).

En el caso de la región de Tancítaro se llevó a cabo la interpretación de orto-fotografías del INEGI y visitas de campo para elaborar una base de datos con 600 puntos de verificación, la cual se utilizó para evaluar la fiabilidad de los mapas obtenidos con MODIS. En el área Maya y Mazatán el trabajo de campo se realizó únicamente para validar la clasificación visual mencionada en el punto anterior.

### **3. Datos no espectrales**

Se utilizó un mapa edafológico digital de INEGI escala 1:250,000 así como un modelo digital de elevación del Shuttle Radar Topography Mission (SRTM) a una resolución de 90 m, que fue remuestreado por medio del método del vecino más cercano a 250 m.

## **5.2. METODOLOGÍA**

A continuación se presenta una descripción de cada uno de los pasos que se siguieron para el análisis de los datos, un resumen general de los procedimientos se presenta en la Figura 12.

### **5.2.1. Selección de los datos**

Se seleccionaron los datos más populares utilizados en los trabajos revisados en los antecedentes de esta tesis. Estos datos, son; los índices de vegetación (MOD13), y los datos de reflectancia de la superficie (MOD09).

### **5.2.2. Descarga, corte, reproyección y remuestreo de los datos**

Los productos MODIS se cortaron para coincidir con el tamaño de los datos de referencia mediante la herramienta Modis Reprojection Tool (MRT). Con esta misma herramienta todos los datos MODIS se proyectaron de un sistema de coordenadas global (latitud longitud) a un sistema de coordenadas más utilizado a nivel local, este es el Universal Transversal de Mercator (UTM) a las zonas que corresponde; Tancítaro (UTM 14), Quintana Roo (UTM 16) y Mazatán (UTM 12). Las bandas 3-5 fueron remuestreadas por medio del método del vecino más cercano a 250 m para coincidir con las bandas 1-5.

### 5.2.3. Análisis de la calidad de los datos

A partir de la capa de calidad Quality Assessment (QA) se seleccionaron y extrajeron los bits de utilidad que permitieron discriminar las imágenes de mejor calidad para cada conjunto de datos. La extracción de bits se hizo por medio de la herramienta Land Data Operational Product Evaluation (*LDOPE*).

En los índices de vegetación, la información sobre la utilidad la imagen está codificada en los bits 2-5 y varía de 0 (más alta calidad) a 7 (imagen inutilizable). En los productos de reflectancia se encuentra en los bits 0-1 y varía de 0 (corregido a calidad ideal en todas las bandas) a 3 (no producido). Además de la información de utilidad, en las imágenes diarias se consideraron los bits 2-3 donde se encuentra codificada la información sobre el estado de nebulosidad con valores de 0 (píxel sin nube), 1 (píxel con mezcla de nube) y 2 (píxel con nube). Una descripción completa de los índices de calidad se encuentra en [http://modis-250m.nascom.nasa.gov/cgi-bin/QA\\_WWW/newPage.cgi?fileName=MODLAND\\_guide](http://modis-250m.nascom.nasa.gov/cgi-bin/QA_WWW/newPage.cgi?fileName=MODLAND_guide).

Con base en el índice de calidad o utilidad se establecieron parámetros para descartar las escenas con baja calidad. En el caso de las áreas de Tancítaro y Mazatán se seleccionaron los compuestos de reflectancia de 8 días que presentaron la más alta calidad reportada en todos sus píxeles (imágenes cuyo índice de calidad promedio es igual a 0), en el caso del área Maya no existieron imágenes con calidad ideal en todos los píxeles, por lo cual se permitió considerar imágenes con 10% de píxeles con calidad no ideal siempre y cuando no rebasaran un valor de 1.

En el caso de los índices de vegetación en el área Tancítaro y Maya se consideraron como imágenes de buena calidad las que tuvieron un promedio de valor de utilidad de 1.5 o menor, con no más del 5% de píxeles con calidad no ideal siempre y cuando estos no rebasaran un valor de 3.

En las imágenes diarias de todas las áreas antes de evaluar su utilidad, se descartaron las escenas que presentaron píxeles reportados con nube, se obtuvieron imágenes sin nube en las áreas de Tancítaro y Mazatán e imágenes con menos del 10% de nube en el área Maya. Las imágenes diarias que pasaron el filtro de nube fueron evaluadas en cuando a su índice de calidad con los mismos parámetros establecidos de los compuestos de reflectancia de 8 días. Los parámetros finales se muestran en el Cuadro 9.

**Cuadro 9** - Bits seleccionados a partir de la capa de calidad MODIS (QA) para cada tipo de dato

PRODUCTO	TIPO DE DATO	BITS SELECCIONADOS DE LA CAPA DE CALIDAD	PARÁMETROS		
			Descripción	Valores en las áreas Mazatán y Tancítaro	Valores en el área Maya
MOD09 Reflectancia de la superficie	Imágenes diarias	2-3 (estado de nube)	Porcentaje de nube	0	0.1
		0-1 (calidad)	Media de la imagen	0	0.5
			QA muy bajo	--	1
	Compuestos de 8 días	0-1 (calidad)	Umbral de proporción	0	0.1
			Media de la imagen	0	0.5
			QA muy bajo	--	1
			Umbral de proporción	0	0.1
MOD 13 Índices de Vegetación	Compuestos de 16 días	2-5 (utilidad)	Media de la imagen	1.5	2
			Qa muy bajo	3	3
			Umbral de proporción	0.05	0.1

#### 5.2.4. Análisis visual

Se hicieron compuestos de color de las imágenes de reflectancia y estos fueron revisados de manera visual, al igual que las imágenes de los índices de vegetación. Se hizo una lista con el nombre de cada una de las imágenes y se asignaron valores del 0 al 4 de acuerdo a su calidad visual, siendo 0 la imagen sin datos (aparece la escena en negro) 1 con nube o ruido en más del 30% de la imagen, 2 con nube entre 10-30% de nube o ruido, 3 con menos de 10% de nube o ruido y 4 sin nube.

Estas listas se compararon con los resultados del análisis de calidad para comprobar la correspondencia visual con la del algoritmo. En una segunda fase, el análisis visual sirvió para descartar imágenes con nube o ruido que permanecieron después de efectuar el análisis automatizado de la capa de calidad.

#### 5.2.5. Selección del sistema clasificatorio

En un primer paso, se consideró un sistema clasificatorio detallado de acuerdo a cada uno de los mapas de referencia de cada una de las zonas. Debido a que en ningún área se logró separar estas clases eficientemente, se hizo una agrupación a un nivel clasificatorio más general basado en el sistema clasificatorio del INEGI. La leyenda final es más detallada que las

elaboradas usualmente con este tipo de datos como la del Programa internacional de la Biósfera – Geósfera (IGBP por sus siglas en inglés).

Para el área de Tancítaro las categorías finales fueron: 1) Bosque de Pino-Encino, 2) Bosque de Pino, 3) Selva Baja Caducifolia, 4) Agricultura de Riego 5) Huertas de Aguacate, 6) Agricultura de Temporal 7) Pastizales, 8) Áreas sin Vegetación y 9) Áreas Urbanas.

En el área Maya: 1) Selva Mediana Sub-Perennifolia, 2) Selva Baja Sub-Perennifolia, 3) Vegetación Hidrófila, 4) Cuerpo de Agua, 5) Mosaico Agrícola, y 6) Áreas Urbanas.

En el área de Mazatán, 1) Matorral (incluye Mezquital y Huizachal), 2) Agricultura, 3) Bosque de Encino y 4) Área Urbana.

#### **5.2.6. Campos de entrenamiento**

Los campos de entrenamiento se trazaron en la imagen MODIS, tomando en cuenta como referencia imágenes Landsat TM y SPOT, así como una clasificación no supervisada y las clases del IFN. Se tuvo la precaución de representar de manera homogénea cada clase y sobre toda la extensión de la imagen. Se realizaron clasificaciones no supervisadas que fueron sobrepuestas en los mapas preexistentes con el objetivo de determinar la discriminación espectral de las clases de interés.

#### **5.2.7. Cálculo de la separabilidad**

Generalmente, el desempeño del clasificador aumenta al aumentar el número de bandas de entrada. Sin embargo, a partir de un cierto número de bandas, el desempeño ya no aumenta, e inclusive disminuye. Este efecto de saturación, conocido como el efecto Hughes (Hughes 1968), afecta a los clasificadores de Máxima Verosimilitud y a las redes neuronales. Para identificar el número óptimo de bandas y el conjunto de bandas más eficiente para diferenciar las categorías que se pretenden cartografiar, se puede evaluar la separabilidad entre pares de categorías a través del cálculo de la distancia entre las muestras (campo de entrenamiento) en el espacio multidimensional definido por un cierto conjunto de variables de entrada. Este enfoque permite identificar que categorías se confunden entre sí (aquellas que presentan poca separabilidad) y tomar en cuenta los valores de separabilidad de todos los pares de categorías, para definir la combinación de bandas de entrada que permita la mejor separabilidad general (Landgrebe 2003).

También permite identificar las categorías no discernibles con los datos y por consecuencia brinda elementos para adaptar (simplificar) el sistema clasificatorio. Con el objeto de reducir la dimensionalidad de la base de datos se evaluó la separabilidad, a través del cálculo de la divergencia transformada (DT). Esta medida de separabilidad permite prever el desempeño del clasificador de Máxima Verosimilitud (Maussel et al. 1990), indicando cuantas y cuales bandas son suficientes para discriminar las clases de interés. Los valores de la DT fluctúan entre 0 y 2000, donde el valor de 2000 representa la separabilidad máxima entre los pares de clases. El cálculo de la separabilidad se hizo primero por cada tipo de producto MODIS (compuestos multiespectrales, imágenes diarias e índices de vegetación), y en una segunda fase para combinaciones de productos. Para evaluar cada conjunto de bandas, se tomó en cuenta el valor mínimo de separabilidad, correspondiente el par de categorías que más tienden a confundirse. Se adoptaron las reglas mostradas en el Cuadro 10.

**Cuadro 10** - Definición cualitativa propuesta para la interpretación de los valores de separabilidad (divergencia transformada)

VALOR DT	GRADO DE SEPARABILIDAD
0 – 1000	Separabilidad muy pobre
1000 - 1900	Separabilidad pobre
1900 - 2000	Buena separabilidad

(Tomado de PCI Geomatics 2007)

### 5.2.8. Clasificación

Para conducir las clasificaciones se eligieron dos tipos de clasificadores, uno paramétrico (Máxima Verosimilitud) y uno no paramétrico (Perceptrón Multicapa). Estos algoritmos se describen a continuación.

#### 1) Clasificador de *Máxima Verosimilitud* (MVS)

El clasificador de Máxima Verosimilitud (o máxima probabilidad) es el algoritmo más comúnmente utilizado en percepción remota, ya que es sencillo de aplicar, así como de interpretar los resultados que arroja, además de considerarse uno de los procedimientos de discriminación más acertados y eficientes, esto siempre que los datos sigan una función de distribución normal (Maselli et al. 1992). En la fase de aprendizaje, el algoritmo elabora el

“patrón espectral” de cada categoría con base en la media y de la varianza/covarianza de un conjunto de sitios de entrenamiento localizados en la imagen. En la fase de clasificación, se calcula, para cada píxel, la probabilidad de pertenencia a cada categoría con base en su respuesta espectral (ecuación 1). El píxel se asigna finalmente a la clase a la cual es más probable que pertenezca de acuerdo a la información espectral.

$$P(H/E) = \frac{p(E/H)p(H)}{p(E)} \quad (1)$$

Donde:

$P(H/E)$  = es la probabilidad condicionada de la hipótesis  $H$  dado evidencia  $E$ , es decir la probabilidad que el píxel pertenezca a cierta categoría hipótesis  $H$  tomando en cuenta su respuesta espectral (Evidencia  $E$ ),

$p(H)$  es la probabilidad *a priori* de la hipótesis  $H$ , es decir la probabilidad que un píxel pertenezca a la categoría considerada en la hipótesis  $H$  sin tomar en cuenta su información espectral.

$p(E)$  es la probabilidad de la evidencia  $E$ ,

$p(E/H)$  es la probabilidad condicionada de la evidencia  $E$  dada la hipótesis  $H$ .

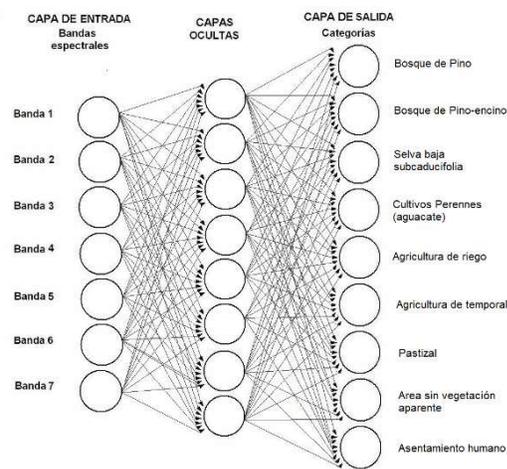
En ausencia de información sobre la probabilidad *a priori* de encontrar cierta categoría en determinado sitio, se asume que las probabilidades *a priori*  $p(H)$  son iguales para todas las clases y en toda el área de la imagen, es decir que si se consideran  $n$  categorías, esta probabilidad es igual a  $1/n$  (ecuación 2).

$$P(H/E) = \frac{p(E/H)(1/n)}{p(E)} \quad (2)$$

## 2) Clasificador de *Perceptrón Multicapa* (PMC)

El clasificador de Perceptrón Multicapa es el algoritmo de redes neuronales más utilizado debido a su capacidad de tolerancia ante información incompleta o contaminada con ruido (Mas

y Flores 2008). Consiste en un modelo estadístico no paramétrico de regresión no lineal (Sarle 1994) y se describe como una red de alimentación hacia adelante compuesta por una capa de unidades (neuronas) de entrada (en este caso una unidad por banda), otra capa de unidades de salida (una unidad para cada categoría) y un número determinado de capas intermedias de unidades de procesamiento, también llamadas capas ocultas porque no tienen conexiones con el exterior. Las neuronas de cada capa están conectadas a todas las unidades de la capa siguiente (Figura 11). Con esta red se pretende establecer una correspondencia entre un conjunto de entrada (respuesta espectral por ejemplo) y un conjunto de salidas deseadas (categorías de cobertura del suelo). Generalmente, el PMC usa una sola capa oculta para dividir completamente el espacio espectral por medio de hiperplanos a lo largo del cual el nivel de activación de las unidades ocultas es constante (Foody 2000).



**Figura 11 - Red Neuronal Perceptrón Multicapa para clasificar una imagen multiespectral de siete bandas en 7 categorías**

En el PMC, la señal que se transmite es un valor numérico. Cada conexión entre dos neuronas está asociada a un peso de ponderación que controla la fuerza de una interconexión. La neurona recibe como señal de entrada la suma de las señales de salida de cada neurona de la capa anterior, ponderada por su respectivo peso. Esta señal de entrada es transformada por una función de activación, generalmente sigmoideal, que permite generar una señal de salida que se transmite a las neuronas de las capas siguientes. En un PMC cuyo objetivo es clasificar, cada neurona de salida está asociada a una categoría, la categoría final corresponde a la neurona que presentan la

señal, o valor de activación, más grande. El valor de activación de cada categoría puede interpretarse con un valor de pertenencia difusa ya que estos valores no se pueden considerar como probabilidad en sentido estricto como en el caso del método de la Máxima Verosimilitud.

En el entrenamiento de la red neuronal los pesos están elegidos de manera aleatoria al inicio, y la respuesta del PMC es por lo tanto totalmente errónea. Con base en los datos de entrenamiento, se modifican los pesos de ponderación para obtener las respuestas correctas. El procedimiento de aprendizaje es sencillo: si la red da la respuesta equivocada, los pesos se corrigen de tal manera que el error disminuye. Los datos de entrenamiento se presentan iterativamente para ajustar los pesos y obtener el mejor ajuste entre valores esperado y obtenido por la red. El método de entrenamiento más conocido es el de retro-propagación (back propagation). En este algoritmo, se compara, con base en los datos de entrenamiento, la señal de salida de la red con el valor esperado para calcular el error. El error es entonces retroalimentado por la red y los pesos de las conexiones se alteran para minimizar este error (Bishop 1995).

Existen varios parámetros que debe definir el usuario, uno es el número de píxeles para efectuar el entrenamiento y el número de píxeles de prueba para validar los resultados. El segundo parámetro es el número de neuronas ocultas, el tercero determina la forma de la función de activación (lineal o sigmoideal). El cuarto parámetro es la tasa de aprendizaje, aquí debe de tenerse cuidado de no establecer un valor muy pequeño que puede resultar en un aprendizaje muy lento, por el contrario si el valor es muy alto los resultados pueden ser caóticos en lugar de converger obteniendo resultados mediocres. El quinto parámetro es el factor momentum, este se utiliza para acelerar el procedimiento de convergencia. Los últimos tres parámetros controlan la terminación del procedimiento antes de entrar en una fase de sobreentrenamiento en la cual la red neural es capaz de clasificar muy bien los datos de entrenamiento pero falla en clasificar los datos “no vistos” al momento de clasificar la imagen. Se basan en el porcentaje aceptable del error, el número de iteraciones, y la tasa de fiabilidad, al alcanzar cualquiera de estos valores el entrenamiento se detiene. (Eastman 2003). Una revisión sobre redes neurales puede encontrarse en Mas (2005) y Mas y Flores (2008).

En este trabajo se hicieron varias pruebas para elegir los parámetros óptimos en el entrenamiento de la red neural, se definieron finalmente un valor para la función sigmoideal de 1.0 con una tasa de fiabilidad de 70 y 500 iteraciones. El número de neuronas en la capa oculta se definió de acuerdo a las recomendaciones de Kavzoglu y Mather (2003) como el promedio de las

capas de entrada y las de salida siendo finalmente 8 en Tancítaro, 7 en área Maya y 6 en Mazatán.

### **5.2.9. Clasificación con información auxiliar**

La incorporación de datos auxiliares en la clasificación ha mostrado aumentar la fiabilidad de los mapas obtenidos (Hutchinson 1982), por lo que en este estudio se probaron dos formas de incorporar datos auxiliares.

Las características del método de Máxima Verosimilitud lo hacen ideal para incorporar datos auxiliares para aumentar su desempeño conservando las ventajas del método original (Maselli et al. 1995; Wulder et al. 2006). Se generaron mapas de probabilidades *a priori*  $p(H)$  para cada una de las categorías, aplicando un filtro paso-bajo de 2 x 2 km en el mapa de cobertura / uso del suelo de INEGI. En términos prácticos, para cada píxel, estos mapas de probabilidad *a priori* disminuyen o anula la probabilidad que el clasificador escoja una categoría que no esté presente en los alrededores del píxel en el mapa de referencia (ver ecuación 1).

En el caso de las redes neuronales, se usaron los datos de elevación y un mapa de suelos de INEGI, para generar mapas de “probabilidad” *a priori* (Mas 2004) que se combinaron con los valores de activación del PMC entrenado con los datos de MODIS. Para ello, el modelo digital de elevación fue clasificado por otro PMC usando el mapa de uso/cobertura del suelo de INEGI como campos de entrenamiento y reteniendo los valores de activación de la clasificación. Este mapa expresa por lo tanto, para cada píxel, la posibilidad (no se puede hablar de probabilidad en el sentido estricto) de encontrar cada una de las categorías de cobertura tomando en cuenta la elevación. El mismo mapa de INEGI se traslapó con el mapa de suelos para calcular la probabilidad condicional de cada categoría de cobertura para cada tipo de suelo. Estos mapas de probabilidad se combinaron, categoría por categoría, con el mapa de valores de activación de la clasificación con datos MODIS tomando en cuenta el valor mínimo de los tres valores (derivado de la elevación, del suelo y de los datos MODIS), lo cual corresponde al operador *AND* en lógica difusa. Finalmente, el píxel se clasificó en la categoría que presentaba el valor de activación final máximo.

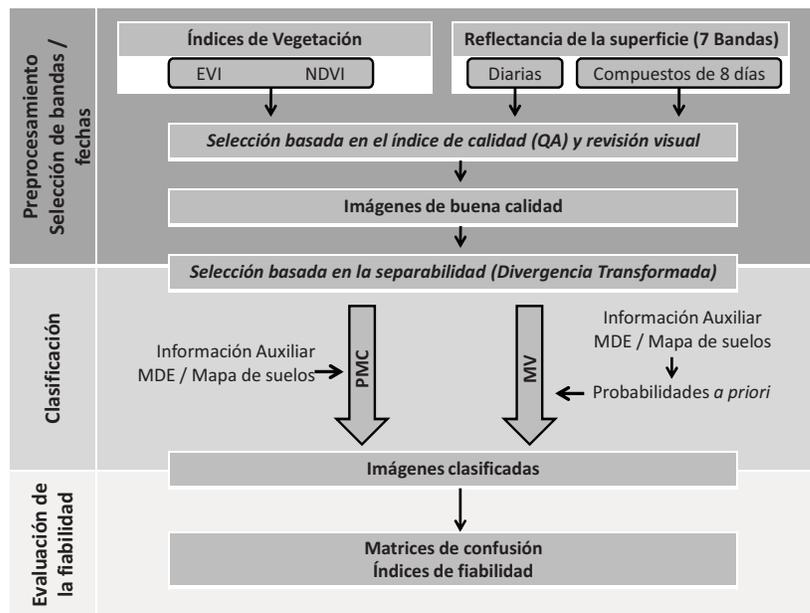


Figura 12 - Diagrama de flujo describiendo el procedimiento de análisis de los datos MODIS

### 5.2.10. Evaluación de la fiabilidad

La evaluación de la fiabilidad se llevó a cabo de manera distinta en cada área de acuerdo a los datos disponibles para cada una. Los procesos se describen a continuación.

#### 1) Unidades de verificación

En el área de Tancítaro los resultados de las clasificaciones se compararon contra las observaciones de campo en un muestreo aleatorio estratificado. La interpretación de datos de campo se hizo mediante un enfoque difuso adaptado del propuesto por Woodcock y Gopal (2000). Como se observa en el Cuadro 11, en el caso de que un sitio estuviera situado en el límite de dos categorías o en un área con clases mezcladas se registraron dos observaciones (Obs.1 y Obs. 2), con esta información, al calificar el mapa podía ser considerada como correcta cualquiera de las dos observaciones (ver punto de control 3). Cuando el punto de control se situó en un área homogénea se consideró una sola clase (ver puntos de control 1 y 2).

**Cuadro 11 - Método de interpretación de datos de campo**

Punto de control	Obs. 1	Obs. 2	MAPA
1	A	A	B
2	C	C	C
3	D	C	C

En el caso del área de Mazatán y Quintana Roo, los resultados de las clasificaciones se compararon contra los mapas detallados del programa piloto de INEGI y de una tesis de doctorado del CIGA-UNAM respectivamente, mediante una validación cruzada.

**2) Matrices de confusión y cálculo de índices de fiabilidad**

Para conocer la proporción de error obtenido en los mapas se hicieron matrices de confusión, las cuales permiten confrontar la información de los sitios de muestreo y la imagen clasificada. A esta matriz se le aplicó el método propuesto por Card (1982), incorporando los valores de proporción (áreas relativas) de cada una de las categorías consideradas para compensar el sesgo debido al muestreo estratificado. De esta manera, una clase con mayor área tendrá más peso que una de área pequeña al momento de calcular la fiabilidad.

El Cuadro 12 es una representación de una matriz de confusión, en donde se cruzan los valores obtenidos por el mapa (clasificación) con la información de referencia (verdad). Se calcula el total (sumatoria) de cada una de las columnas, así como la proporción en de cada una de las clases en el mapa. El total y la proporción se usan hacer la corrección de Card (1992) como se muestra en la ecuación 2.

**Cuadro 12 - Matriz de confusión**

Mapa \ Referencia	Mapa				
	1	2	...	j	Total
1	$P_{11}$	$P_{12}$	...	$P_{1j}$	$P_{1+}$
2	$P_{21}$	$P_{22}$	...	$P_{2j}$	$P_{2+}$
.....	...	...	...	...	...
i	$P_{q1}$	$P_{q2}$	...	$P_{ij}$	$P_{i+}$
Total	$P_{+1}$	$P_{+2}$	...	$P_{+j}$	
Proporción de mapa	$\pi_1$	$\pi_2$		$\pi_j$	1

(Tomado de Mas et al. 2003)

Suponiendo una matriz con  $j$  columnas e  $i$  líneas (Cuadro 12), Card (1992) calcula cada elemento  $P_{ij}$  de la matriz corregida aplicando:

$$P_{ij} = \frac{\pi_j n_{ij}}{n_{+j}} \quad (3)$$

Donde:

$n_{ij}$  es el número en cada celda

$n_{+j}$  es la suma de las celdas de la columna  $j$

$\pi_j$  es la proporción en el mapa de la categoría  $j$

Se calcularon los valores de la fiabilidad del usuario y fiabilidad de productor. La fiabilidad de usuario mide el desempeño de la clasificación en el campo por categoría (líneas). Esta medida de fiabilidad toma en cuenta los errores de comisión, que resultan cuando un píxel se asigna a una clase incorrecta (Fórmula 4). La fiabilidad de productor mide la certidumbre con la cual se clasificó la imagen por categoría (columnas). La fiabilidad de productor evidencia los errores de omisión que son los que resultan cuando un píxel se clasifica en otra categoría, es decir se omite de su categoría correcta (Fórmula 5).

$$FU = \frac{P_{jj}}{\pi_j} \quad (4) \quad FP = \frac{P_{ii}}{\sum_{j=1}^{n_c} P_{ij}} \quad (5)$$

Donde:

$FU$ = Fiabilidad de usuario

$FP$ = Fiabilidad de productor

$P_{ij}$  es el valor de cada celda después de aplicar la corrección de Card

$\sum P_i$  es la sumatoria de las celdas de la fila  $i$

$\pi_j$  es la proporción en el mapa de la categoría  $j$

**5.2.11. Comparación de la fiabilidad obtenida de los diferentes métodos de clasificación - Prueba de McNemar**

Los valores de fiabilidad de los mapas nos permiten compararlos en términos de cual obtuvo resultados mayores, por ejemplo un mapa de 90% tiene un porcentaje de fiabilidad mayor que uno de 86%. Sin embargo, esto no es suficiente para concluir que uno sea más confiable que otro, ya que para esto es necesario determinar si las diferencias son estadísticamente significativas. Con el objetivo de contar con pruebas de evaluación y comparación robustas se utilizó la metodología propuesta por Foody (2004), que permite evaluar la significancia estadística de las diferencias en la fiabilidad entre dos clasificaciones, estimada con base en los mismos datos de verificación. Esta es una prueba no paramétrica basada en una matriz de confusión de dos por dos que expresa las relaciones entre las confusiones de las dos clasificaciones en comparación (Cuadro 13).

**Cuadro 13 - Elementos de la matriz utilizada en la prueba de McNemar**

		Clasificación 2	
		Correcto	Incorrecto
Clasificación 1	Correcto	$f_{11}$	$f_{12}$
	Incorrecto	$f_{21}$	$f_{22}$

(Tomado de Mas 2005)

El índice de la prueba  $Z^2$  expresa la diferencia en los errores cometidos en las dos clasificaciones. Esta variable sigue una distribución de probabilidad chi cuadrada con un grado de libertad y por lo tanto, se compara en la tabla de valores críticos de chi cuadrado para determinar si esta diferencia en la fiabilidad es estadísticamente significativa.

$$Z^2 = \frac{(\ell_{12} - \ell_{21})^2}{\ell_{12} + \ell_{21}} \quad (6)$$

El estadístico  $Z^2$  mide la diferencia de la fiabilidad en comparaciones pareadas de los diferentes clasificadores. La prueba de  $Z^2$  es así una prueba no dirigida (prueba de planteamiento

bilateral), que indica si existe o no una diferencia significativa entre las fiabilidades alcanzadas por pares de clasificaciones (Foody 2004).

Los resultados alcanzados mediante la aplicación de los métodos descritos se presentan a continuación.

---

## CAPÍTULO VI. RESULTADOS

---

En este capítulo se presentan los resultados alcanzados en cada una de las fases de procesamiento para cada una de las áreas de estudio.

### 6.1. ANÁLISIS DE LA CALIDAD DE LOS DATOS

El filtrado de las imágenes basado en los valores de la capa de calidad (QA) resultó en una conveniente reducción del número de datos para cada una de las áreas de estudio como se muestra en el Cuadro 14.

**Cuadro 14 - Reducción de datos a través de la capa de calidad**

Tipo de dato	Número original	Tancítaro Michoacán	Área Maya Quintana Roo	Mazatán Sonora
IV	22	12	5	12
Diarias	365	81	30	300
C8	48	21	4	44

Se observó una muy buena correspondencia entre la inspección visual de las imágenes y los valores QA en los índices de vegetación, las imágenes IV que pasaron el filtro QA mediante el análisis automatizado con los parámetros propuestos coincidieron en todos los casos con las marcadas con los valores más altos de calidad visual.

En el caso de los datos de reflectancia (compuestos de 8 días e imágenes diarias), mediante el análisis visual se detectaron imágenes donde el análisis de calidad no descartó la totalidad de las imágenes con nebulosidad o ruido a pesar que el QA indica el máximo valor de calidad (0), por lo que fue necesario descartarlas de manera manual.

Como se observa en la figura 13 que corresponde al área de Tancítaro, la mayoría de las imágenes descartadas corresponden a la época de lluvias: el valor promedio de QA aumenta durante la época de lluvias, lo cual indica imágenes de baja calidad.

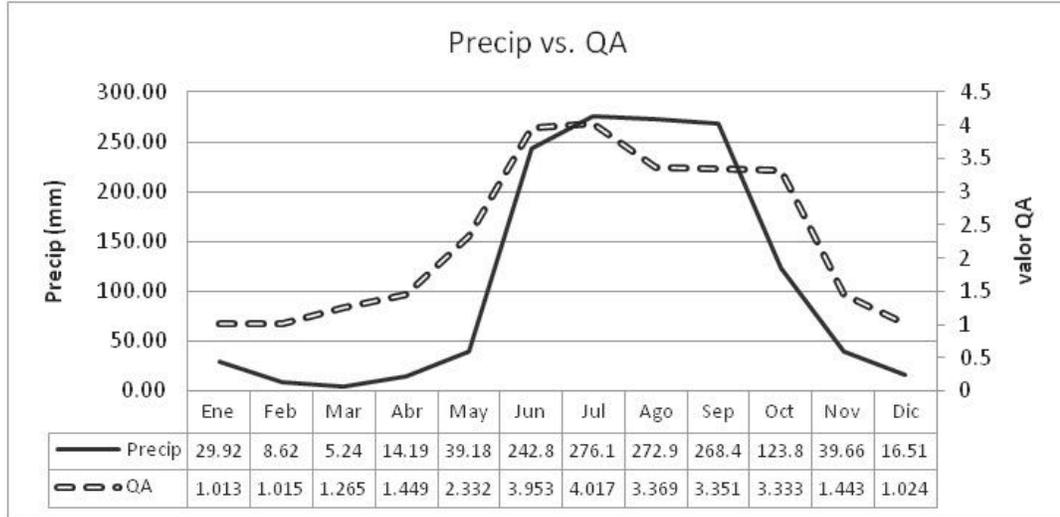


Figura 13 - Qa Vs precipitación

### 6.1.1. Medición de la separabilidad (cálculo de la divergencia transformada)

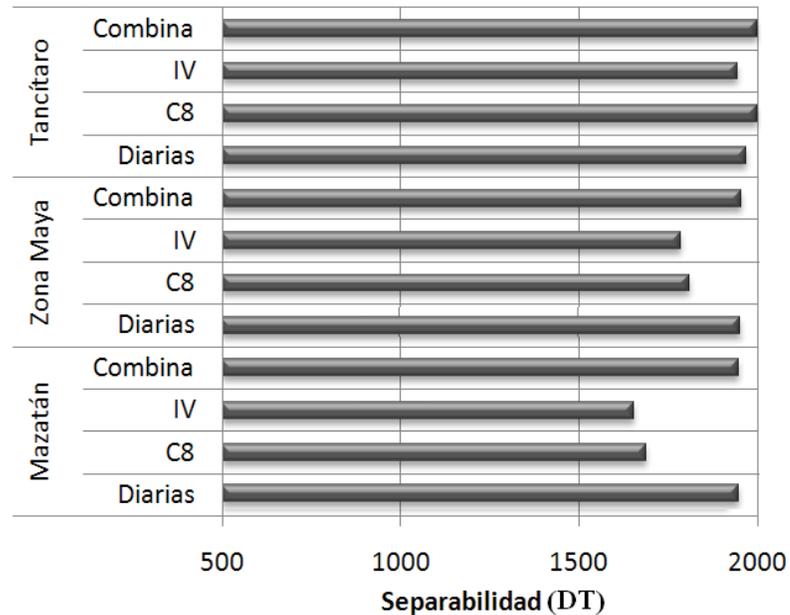
Los resultados del análisis de separabilidad por medio de la divergencia transformada para 10 clases en Tancítaro, 7 clases en Mazatán y 11 clases en el área Maya propuestos en los mapas de referencia no obtuvieron valores satisfactorios. Aun con la mejora de los campos de entrenamiento fue imposible distinguir las clases que presentaron confusión por lo cual éstas se fusionaron contando finalmente con una leyenda más general.

En Tancítaro se agrupó el Bosque de Pino con el Bosque de Pino-Encino con lo que resultaron 9 clases finales. En Mazatán fue imposible separar espectralmente el Matorral, Mezquital, Huizachal y Matorral Desértico Micrófilo, inclusive en campo no fue clara la diferenciación entre estas clases, por lo cual a partir de estas se conformó una sola clase a la que se denominó Matorral coincidiendo así con la leyenda en Inventario Nacional Forestal y teniendo al final una leyenda de 4 clases en esta área.

En el área de Quintana Roo el mapa de referencia generado con base en una imagen pancromática SPOT mediante clasificación visual, delimita áreas de diferente tipo de cultivos de superficie pequeña, así como pequeñas áreas de desmonte en distinto estado de regeneración. Fue imposible separar espectralmente estas clases, por lo cual se integraron en una sola clase a la

cual denominamos Mosaico Agrícola. El resultado final fue de seis clases con las cuales se obtuvo una separabilidad satisfactoria con todos los tipos de datos.

Como se observa en la figura 14 fue posible alcanzar una separabilidad casi óptima con el número final de clases, con todos los tipos de datos en todas las áreas de estudio contando con un número reducido de bandas.



**Figura 14 - Separabilidad máxima alcanzada con 7 bandas por los diferentes productos en cada área de estudio**

De acuerdo a los valores de separabilidad en las tres áreas los datos que permiten una mayor discriminación de las categorías son los de reflectancia de la superficie, mientras que los índices de vegetación presentaron la menor separabilidad en todos los casos.

La separabilidad promedio obtenida con los datos combinados (conjunto de datos que incluye índices de vegetación, compuestos de 8 días e imágenes diarias) fue ligeramente mayor con respecto a los resultados obtenidos por medio de un solo tipo de dato de superficie de la reflectancia en el área de Tancitaro mientras que para las áreas Maya y Mazatán la mejor combinación de datos resultó estar compuesta solo por las imágenes diarias (cuadro 15).

### 6.1.2. Selección espectral y por fechas

Los resultados de la separabilidad pueden ser un buen indicador sobre el tipo de información que resulta más útil para separar las clases de interés en cada una de las áreas.

En el cuadro 15 se presentan las fechas, bandas espectrales y/o tipo de índice de vegetación que fue seleccionado a través del análisis de separabilidad para distinguir las clases propuestas en cada área.

**Cuadro 15 - Selección espectral y por fechas**

		<b>Tancítaro, Mich.</b>	<b>Área Maya. QRoo</b>	<b>Mazatán, Son.</b>
IV	<b>Fecha</b>	ENE FEB NOV DIC	FEB, MAR, NOV, DIC.	ENE, FEB, SEP, NOV
	<b>Tipo IV</b>	3 NDVI, 4 EVI	4 NDVI, 3 EVI	4 NDVI, 3 EVI
C8	<b>Fecha</b>	ENE FEB MAR	ENE, DIC.	ENE, FEB, JUL, DIC
	<b>Banda</b>	2,3,4,5,6	1,2,4,6,7	2,4,5,6,7
Diarias	<b>Fecha</b>	ENE, MARZO	ENE, DIC.	ENE, MAR, MAY, JUN, AGO, OCT
	<b>Banda</b>	3,4,5,6	1,2,3,4,5,6,7	2,3,5,7
Combina	<b>Fecha</b>	ENE FEB MAR MAY	ENE, DIC.	ENE, MAR, MAY, JUN, AGO, OCT
	<b>Banda</b>	C8(2,4,6), EVI, D-6	1,2,4,5,6,7	2,3,5,7

Como se puede observar, a partir del conjunto de datos compuesto por los índices de vegetación, en todas las áreas resultaron seleccionados por igual datos del EVI y el NDVI, esto se puede deber a que como se indica en la sección 2.6.3 cada índice tiene sensibilidad a distintas características de la vegetación y resultan complementarios.

En cuanto a los datos de reflectancia la tendencia es la selección de datos de la mayor parte del rango espectral, en los compuestos de 8 días de contar con al menos 5 bandas y en las imágenes diarias al menos 4 bandas de diferente rango espectral.

En cuanto a la selección por fechas, el área que contó con datos más distribuidos durante el año fue Mazatán, lo cual se explica al ser esta área la que contó con más datos de calidad disponibles a lo largo del año.

## 6.2. CLASIFICACIONES

### 6.2.1. Máxima Verosimilitud (MV) y Perceptrón Multicapa (PMC)

La fiabilidad global alcanzada por los diferentes grupos de insumos y métodos de clasificación indican que de manera general en las tres áreas de estudio, la MV tuvo un mejor desempeño que el PMC. Eso puede deberse a la diferencia en el tamaño de los campos (número de píxeles) de entrenamiento de las categorías. En la MV el clasificador utiliza estadísticas derivadas de los campos de entrenamiento, el tamaño del campo no tiene mucha influencia siempre y cuando sea suficientemente grande para tener representatividad estadística. En el caso del PMC, el algoritmo de aprendizaje que busca disminuir el error global por lo que tiende a subestimar las categorías poco representadas en los campos de entrenamiento.

En el área de Tancítaro, las fiabilidades más altas se alcanzan con el algoritmo de la Máxima Verosimilitud en los datos de reflectancia compuestos de 8 días y datos combinados obteniendo una fiabilidad global de 74%, y 76% respectivamente (Figura 15), y sin diferencias significativas entre estas dos clasificaciones de acuerdo a la prueba de de Mc Nemmar.

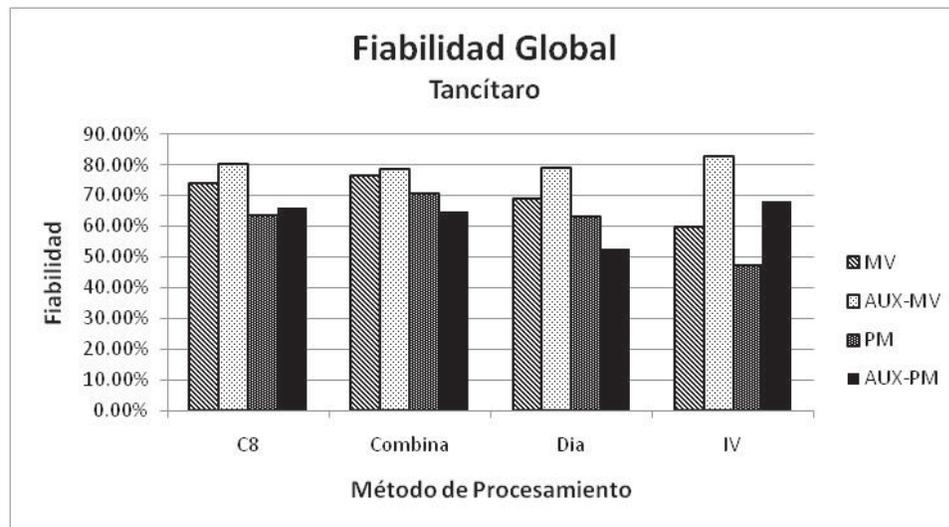


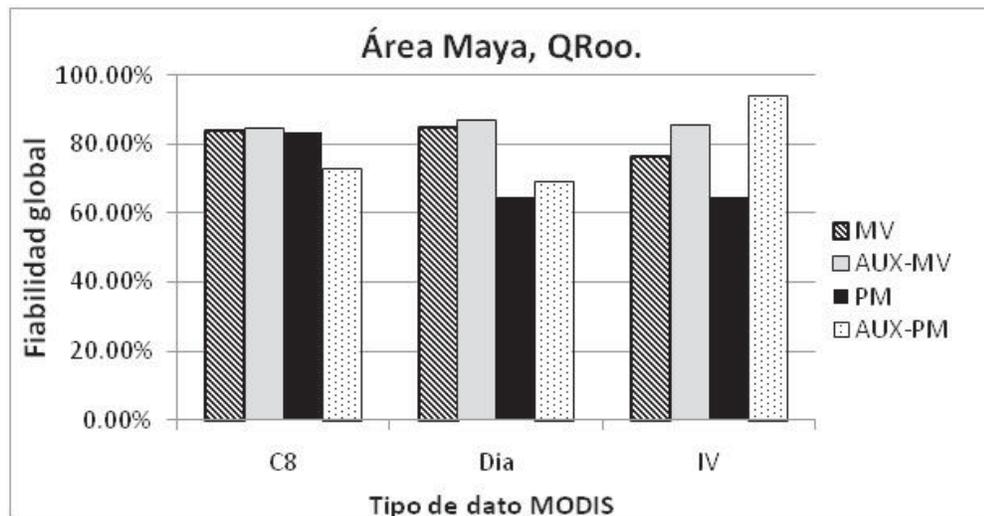
Figura 15 - Valores de fiabilidad global obtenidos por cada una de los tipos de datos analizados por las diferentes metodologías en el área de Tancítaro, Michoacán

En el análisis por clase del área del Tancitaro, observamos que la categoría de plantaciones de aguacate se clasifica satisfactoriamente con todos los tipos de datos y alcanza fiabilidades de hasta 92%. Mientras que la clase de agricultura de temporal no logró ser clasificada con niveles de fiabilidad confiables mediante ningún tipo de dato. Las clases de selva baja subcaducifolia, agricultura de riego, área urbana, pastizal y bosque de pino encino resultan clasificadas correctamente con valores de por lo menos 70% usando datos combinados.

Los resultados de fiabilidad global obtenidos con el clasificador Perceptrón Multicapa son en general menores y son confiables sólo para las categorías de Aguacate y Selva Baja Subcaducifolia (ver anexo 1).

En el área Maya el método de la Máxima Verosimilitud clasifica de manera correcta 84% de los píxeles con los datos de reflectancia diarios, 84% con los compuestos de 8 días y 76% con los índices de vegetación (Figura 16). Sin embargo, la Selva Mediana Subperennifolia representa la mayoría de la superficie en el área de estudio. Esta se clasifica satisfactoriamente independientemente del tipo de dato y método de clasificación utilizado.

El método PMC únicamente clasifica satisfactoriamente con datos compuestos de 8 días e índices de vegetación a la Selva Mediana Subperennifolia, mientras que para el resto de las clases las fiabilidades fueron muy bajas (ver anexo 2).



**Figura 16** - Valores de fiabilidad global obtenidos por cada una de los tipos de datos analizados por las diferentes metodologías en el área Maya, Quintana Roo

En el área de Mazatán, los datos que permitieron una mejor clasificación fueron los índices de vegetación con una fiabilidad global de 69%, seguido de los compuestos de 8 días con 66%, mientras que los datos diarios alcanzaron 63% (Figura 17). Todos estos resultados están por debajo de la fiabilidad aceptable (70%). En el análisis por clase observamos que con base en las imágenes diarias analizadas con MV, las clases; Agricultura, Área Urbana y Bosque de Encino, obtienen fiabilidades por encima del 70%. El Matorral presenta las fiabilidades más bajas 47% a 50%, mientras que la clasificación con base en los índices de vegetación clasifican correctamente con fiabilidades de 74% al matorral pero fiabilidades bajas para todas las demás categorías (ver anexo 3).

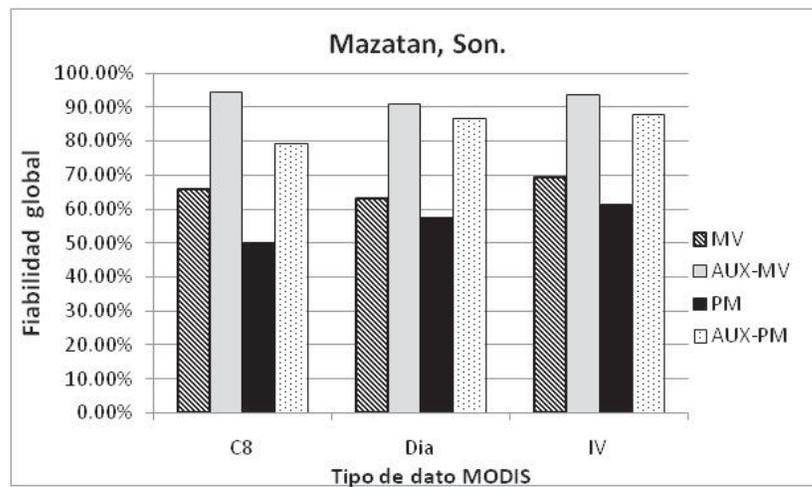


Figura 17 - Valores de fiabilidad global obtenidos por cada una de los tipos de datos analizados por las diferentes metodologías en el área de Mazatán, Sonora

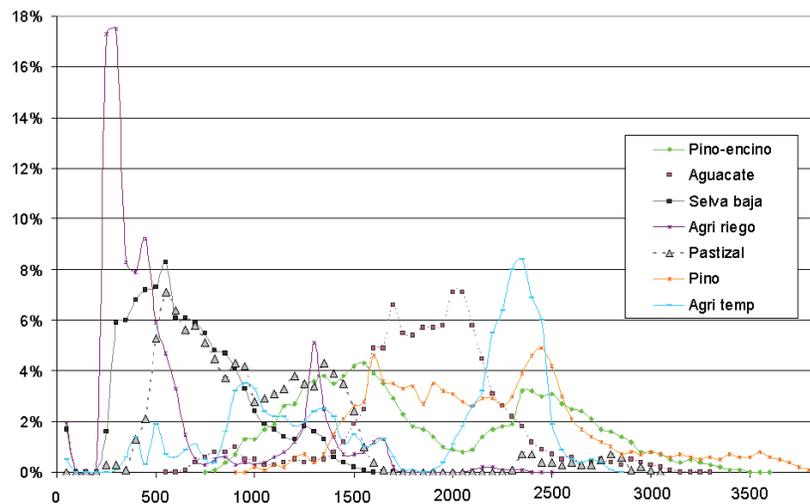
### 6.3.2 Datos auxiliares

La incorporación de datos auxiliares como mapas de probabilidad *a priori* derivados de otro mapa de uso/cobertura del suelo permitió mejorar consistentemente las clasificaciones de la MV en todas las áreas, con aumentos en la fiabilidad de hasta un 30% a través de la restricción de las posibilidades de clasificación.

La incorporación de información auxiliar no espectral en el clasificador de PMC fue más variable. En el caso del área de Tancítaro muchas categorías tienen una distribución altitudinal

muy amplia que se traslapa con otros tipos de cobertura lo que explica que el Modelo Digital de Elevación no tuviera un desempeño muy eficiente en el intento de elevar los resultados de la fiabilidad.

En el área de Mazatán los datos de elevación incorporados como datos auxiliares en la clasificación por medio de PMC incrementaron las fiabilidades desde 27 y hasta 30% obteniendo fiabilidades globales de hasta 88% con los índices de vegetación.



**Figura 18** - Distribución de diferentes tipos de coberturas (% del área total de la cobertura) en función de la elevación

En el área Maya, la incorporación de la información de suelos al clasificar con PMC obtuvo resultados variables. Se observa que la fiabilidad disminuye cuando se clasifican compuestos de reflectancia de 8 días, aumenta ligeramente si se trata de datos diarios, y aumenta significativamente (29%) al clasificar índices de vegetación alcanzando con estos últimos los mejores resultados de fiabilidad global (94%) en comparación con cualquier otra combinación de datos y métodos.

### 6.3.3 Comparación entre los resultados de las tres áreas de estudio

El desempeño de los diferentes tipos de datos MODIS en cada una de las áreas de estudio fue variable dependiendo del método de análisis utilizado.

Cuando se clasifica mediante el método de Máxima Verosimilitud, en el área de Tancítaro, los mejores datos son los compuestos de 8 días, en el área Maya las imágenes diarias y en Mazatán los índices de vegetación. Al incorporar datos auxiliares al método MV, en el Tancítaro no hay diferencia significativa al usar cualquier tipo de dato, en el área Maya las imágenes diarias obtienen mejores resultados, en Mazatán son los compuestos de 8 días y los índices de vegetación sin diferencias significativas entre estos.

Cuando se clasificó mediante un enfoque no paramétrico con el algoritmo Perceptrón Multicapa, en Tancítaro se obtienen mejores resultados con los compuestos de 8 días, al igual que en el área Maya, mientras que para Mazatán son los índices de vegetación. Al incorporar datos auxiliares al PMC en Tancítaro se obtienen mejores resultados con los compuestos de 8 días, en el área Maya los índices de vegetación mientras que en Mazatán son las imágenes diarias y los índices de vegetación sin diferencia significativa entre estos dos tipos de datos.

Los resultados obtenidos por el método de la Máxima Verosimilitud en las diferentes áreas y datos fueron en general más consistentes en comparación con el método de Perceptron Multicapa. Sin embargo, es importante considerar que en las redes neuronales existe dificultad para encontrar los parámetros adecuados para llevar a cabo el entrenamiento.

Los resultados alcanzados mediante la incorporación de información auxiliar por medio de mapas de probabilidad *a priori* aumentan las fiabilidades de las clasificaciones consistentemente, sin embargo es importante tomar en cuenta que mediante este método se corre el riesgo de imponer la información de los datos auxiliares sobre la respuesta puramente espectral. Los datos no espectrales utilizados como información auxiliar tales como mapas de suelos, elevación, climas, o censos agrícolas entre otros se encuentran limitados en su aplicación principalmente por su ausencia o restricción, la diferencia de resoluciones con respecto a las imágenes a clasificar así como la frecuente omisión de sus metadatos. Por último es conveniente recalcar que se debe de contar con un muy buen conocimiento del comportamiento de las clases de uso y cobertura de suelo a separar para elegir el tipo ideal de información auxiliar que sirva a este propósito.

En comparación con otros estudios reportados en la literatura observamos que los resultados de esta tesis son superiores en cuanto a la fiabilidad. Por ejemplo, basado en datos MODIS, Carrão et al. (2008) encuentran una fiabilidad media de usuario de 64% y de productor igual a 46% tomando en cuenta sólo tres clases en su estudio realizado en Portugal, en este estudio en el área de Mazatán con cuatro clases se obtienen fiabilidades de 64% y hasta 72% con el método MV sin datos auxiliares.

También se comparó el producto MOD12 contra los resultados obtenidos en este trabajo. El MOD12 es un mapa de cobertura del suelo con datos MODIS basado en el sistema de clasificación del IGBP. De la comparación resultó que este producto presenta errores de hasta el 44% de la superficie total. Sin embargo, se debe considerar que el producto MOD12 se calibra para todo el continente americano.

## CAPÍTULO VII. DISCUSIÓN Y CONCLUSIONES

---

### 7.1. DISCUSIÓN

Un punto clave en la elaboración de cartografía a partir de datos de percepción remota es la elección del sistema clasificatorio. Este debe de ser lo suficientemente detallado para dar información aplicable a programas de manejo y al mismo tiempo debe adaptarse a las capacidades del sensor y ser compatible con mapas preexistentes para, por un lado, dar continuidad a otros trabajos y, por otro, tomar ventaja de la información preexistente. Como se muestra en la revisión de trabajos previos realizados a partir de datos MODIS, la mayoría de los estudios están basados en el sistema clasificatorio de la IGBP o incluso en una generalización de este y sólo en unos pocos se elabora una leyenda más detallada (Matsuoka et al. 2007, Miettinen 2008 y Zhang et al. 2008). En este trabajo se evaluó el nivel de detalle del sistema clasificatorio alcanzable con MODIS comenzando el análisis con una leyenda detallada, basada en mapas locales desarrollados con datos de alta resolución. Sin embargo, de acuerdo al análisis de separabilidad no fue posible discriminar las categorías este nivel de detalle con datos MODIS, por lo que la leyenda se generalizó alcanzando a tener un nivel de detalle un poco mayor al propuesto por el IGBP y compatible con la cartografía generada a nivel nacional por el Instituto Nacional de Geografía, Estadística e Informática (INEGI). En el caso de México, parece por lo tanto factible un sistema clasificatorio de una docena de categorías, que este entre el nivel tipo y formación del sistema del IFN 2000 incluyendo las categorías 1) Agricultura de riego, 2) Agricultura de temporal, 3) Bosque de Coníferas y Coníferas-latifoliadas, 4) Bosque de latifoliadas, 5) Selva perennifolia y Subperennifolia, 6) Selva Caducifolia y Subcaducifolia, 7) Matorrales, 8) Pastizal, 9) Vegetación Hidrófila, 10) Área sin vegetación aparente, 11) Áreas urbanas y 12) Cuerpo de agua.

Otro aspecto abordado en la tesis es una evaluación de los diferentes insumos MODIS (reflectancia diarias o en compuestos, índices de vegetación) y de la disponibilidad de datos para diferentes fechas. Siempre es deseable contar con una representación temporal de los datos además de la información espectral. La disponibilidad frecuente de datos aumenta la posibilidad de contar con imágenes sin ruido o nube y representa una ventaja al separar las clases de acuerdo

a su fenología. Por ejemplo Matsuoka et al. (2007) separa diferentes tipos de cultivos utilizando las curvas fenológicas relacionadas con los ciclos de cultivo. De manera más general, parece factible separar algunas clases que se confunden espectralmente en ciertas fechas utilizando datos en otros periodos del año. Existen solo dos trabajos encontrados en la literatura que evalúan y comparan la efectividad de la información temporal en datos MODIS y los resultados no demuestran una superioridad significativa al incorporar distintas fechas cuando hay una buena representación espectral. En contraste, aún teniendo información de todo un año de datos MODIS es necesario contar con por lo menos tres bandas espectrales para obtener una fiabilidad satisfactoria (Carrão et al. 2005). Sin embargo, la estacionalidad sí tiene un efecto en el desempeño de los modelos de clasificación, por lo cual los datos de ciertas estaciones aportan resultados superiores con respecto a otras. (Heiskanen y Kivinen 2008). Sin embargo las condiciones meteorológicas muchas veces impiden contar con imágenes de calidad especialmente en la época de lluvias llegando a tener periodos completos sin información. Los métodos de filtrado que buscan completar esta información faltante tomando en cuenta los valores anteriores y posteriores corren el riesgo de tergiversar las curvas de respuesta espectral de las clases de uso y cobertura de suelo entre mayores sean los periodos sin datos. En este estudio se observa que cuando hay una buena representación de datos durante el año, como es el caso del área de Mazatán en Sonora, las clases se separan mejor con datos en diferentes estaciones. La disponibilidad de datos para fechas diferentes parece por lo tanto importante para regiones diversas como las intertropicales. Sin embargo, para estas mismas regiones, la cobertura de nube durante una parte del año es una limitante fuerte

Respecto al tipo de dato más idóneo para efectuar la clasificación, se observó que la mayoría de los trabajos se basan en índices de vegetación, estos datos son muy populares debido a que toman en cuenta los rangos en los que la vegetación absorbe y refleja más y porque hay una gran experiencia del uso de los datos que se han empleado desde hace décadas con imágenes AVHRR. Algunos trabajos han discutido el desempeño del índice más difundido; el NDVI, el cual tiende a saturarse en condiciones de alta biomasa, y poco a poco se han ido introduciendo otros índices dentro de los cuales destaca el EVI. Se ha sugerido que un índice no es mejor que otro, su desempeño depende de las condiciones de la vegetación por lo cual son más bien complementarios. En este estudio se analizaron de manera conjunta y derivado del análisis de la separabilidad, se observó que la mejor separabilidad de las clases se obtiene cuando se utilizan en

conjunto. Sin embargo, los datos de reflectancia obtuvieron resultados significativamente mayores que la combinación de índices de vegetación. La única excepción se presentó al clasificar el matorral en el área de Mazatán donde una combinación de 4 NVDVI y 3 EVI obtuvieron resultados mayores. Al comparar los resultados de las clasificaciones derivadas de las imágenes diarias de la reflectancia contra las de los compuestos de reflectancia de 8 días no se evidencia una ventaja significativa al utilizar las imágenes diarias y no compensa el esfuerzo y tiempo de procesamiento.

Los métodos de clasificación utilizados para examinar los datos de percepción remota son variados e incluyen desde aproximaciones basadas en cálculos de la distancia hasta complejos análisis que requieren de una variedad de parámetros definidos por el usuario. Dentro de los métodos más simples destaca el método de la Máxima Verosimilitud (MV) que aunque tiene ciertas limitaciones (principalmente que asume una distribución normal de los datos) es uno de los métodos más consistentes y más utilizados hasta la fecha por su sencilla aplicación, disponibilidad en la mayoría de los paquetes y fácil interpretación de los resultados. En contraste un método que fue recibido con entusiasmo y que de acuerdo a muchas publicaciones, supera al método de la MV es el de las redes neurales. Éste método se creó tomando como base el funcionamiento cerebral y después del método de la MV ha sido uno de los más aplicados debido a su comprobada habilidad para aprender patrones complejos, generalizar en ambientes con ruido, incorporar conocimiento previo y diferente tipo de datos, ya que no asume una distribución normal de los datos de entrada. Otra ventaja de esta técnica es su capacidad para desarrollar clasificaciones supervisadas utilizando menos datos de entrenamiento y tomando en cuenta las características de entrenamiento de todas las clases. A pesar de las ventajas reportadas de las redes neurales, existen varias limitaciones que hasta la fecha no han podido resolverse totalmente (Mas y Flores 2008). El funcionamiento de las redes neurales se ha comparado con una “caja negra” por su naturaleza compleja especialmente por parte de usuarios nuevos. El método más común de entrenamiento de las redes neurales es el de retro propagación, el cuál es el único que se encuentra disponible en la mayoría de los paquetes de procesamiento a pesar de que existen otros con mejor potencial y con menos requerimientos para el entrenamiento. El entrenamiento de la red neural, por lo tanto, es un proceso delicado en el cual hay que determinar los valores óptimos de varios parámetros (número de neuronas en la capa oculta, tasa de aprendizaje, momentum). Debido a todas estas limitaciones el equipo MODLAND dejó de

utilizar este método para elaborar el producto MOD12 a partir de la colección 5 (Mark Friedl, comunicación personal). En esta tesis el método Perceptrón Multicapa presentó en general fiabilidades más bajas que las obtenidas mediante el algoritmo de la Máxima Verosimilitud, esto puede deberse a que en las áreas de estudio se contó con varias categorías de superficie pequeña, y no fue posible contar con un número de píxeles de entrenamiento suficiente lo que se tradujo en un sobre entrenamiento de la red neural que impidió obtener resultados más satisfactorios a partir de éste método.

En fechas recientes se han comenzado a utilizar cada vez más métodos alternativos dentro de los cuales destacan los Árboles de Decisión y Support Vector Machine. En estudios comparativos, estos nuevos métodos obtienen fiabilidades superiores en comparación con el método de la Máxima Verosimilitud, sin embargo, es importante destacar que existe una tendencia a sobreestimar los resultados obtenidos a partir de nuevas tecnologías generalmente debido a que los autores que proponen un nuevo método están más familiarizados con éste que con el método de control, de igual manera, un método que funciona mejor que los convencionales es más susceptible a ser publicado (Mas y Flores 2008). Debido a que la complejidad del tipo de ambiente que se pretende mapear tiene un efecto fundamental en el desempeño del algoritmo de clasificación no es posible esperar que las metodologías utilizadas en otros estudios tengan resultados repetibles al cambiar el tipo de ambiente. Esto es especialmente cierto cuando se pretende mapear un territorio tan diverso como lo es el mexicano.

La incorporación de datos auxiliares en la clasificación es generalmente muy útil al aumentar la fiabilidad final de los mapas. Este método no se utiliza muy frecuentemente principalmente por la falta de datos en algunas regiones del mundo. En el caso de que los datos auxiliares estén representados por información de otros mapas, es indispensable asegurar que el mapa auxiliar sea confiable, esto restringe un poco las opciones pues en muchos casos los mapas no incluyen metadatos o la evaluación de su fiabilidad no fue robusta. En el caso de incorporar información no espectral, en el caso de MODIS se puede contar con los datos de elevación del Shuttle Radar Topographic Mission disponibles para todo el globo y provistos de manera gratuita a una resolución de 90 m. Los mapas de suelos a nivel nacional no se encuentran actualizados y el nivel de fiabilidad generalmente no es conocido. Sin embargo, la disponibilidad de más mapas temáticos y de mejor calidad se incrementará en el futuro permitiendo mayor facilidad para su incorporación en el mapeo de la cobertura y uso del suelo. En este estudio se obtuvo una mejora

importante y consistente a partir del uso de probabilidades a priori derivadas de mapas previos de cobertura cuando se clasificó mediante el algoritmo de la Máxima Verosimilitud. En el caso de la incorporación de datos de elevación los resultados son más variables, pero estos fueron analizados mediante un método de redes neurales, razón por la cual es difícil definir si los resultados finales del análisis se atribuyen a los datos auxiliares o a la propia red neural. Sin embargo los mapas de elevación presentan una dificultad para incorporarse a métodos como MV ya que no se distribuyen de manera normal. Otro factor a considerar es que muchas de las clases, además de compartir características espectrales similares, comparten también distribución altitudinal y/o tipo de suelo, en estos casos esta información no permite mejorar el desempeño de la clasificación en comparación basada únicamente en la información espectral.

La fiabilidad de los mapas es un aspecto imprescindible para determinar cómo y en qué aplicaciones pueden emplearse los datos generados. A pesar de ser un elemento crítico, muchas veces los métodos de evaluación de los mapas no son lo suficientemente robustos o no toman en cuenta las características del sensor. Son varios los elementos que intervienen en una sólida medición de la fiabilidad, un aspecto es contar con una base de datos de referencia confiable. En la mayoría de los casos la fuente de referencia es una clasificación con mayor resolución por lo que es importante esta cuente con la mayor fiabilidad posible y que exista independencia entre los datos de entrenamiento. Otro factor clave se encuentra en la delimitación de los campos de entrenamiento, los cuales se trazan en áreas homogéneas de la imagen, esto produce que en general el resto de las áreas homogéneas se clasifiquen bien pero las áreas de transición o mezcladas no tendrán la misma suerte. En este estudio las clasificaciones que sirvieron como referencia para validar los mapas del área Maya y Mazatán fueron generadas con una alto detalle y fiabilidad. Cuando la referencia son datos de campo se debe considerar la resolución del sensor, en este estudio la validación de la clasificación en el área de Tancítaro se hizo mediante un enfoque difuso que dio más flexibilidad a la evaluación de la fiabilidad evitando subestimar el desempeño de la clasificación. En la mayoría de los casos de estimación de la fiabilidad reportada en la literatura, los resultados que se presentan están probablemente optimistamente sesgados. En particular, muchos estudios están evaluados por métodos de validación cruzada que no garantizan totalmente la independencia de los datos de entrenamiento y validación y/o con base en datos localizados únicamente en áreas homogéneas. En otros casos las medidas de fiabilidad reportadas

no se presentan por clase o con detalle, e incluso existen trabajos donde la descripción del método de evaluación de la calidad final de los mapas no es del todo clara.

Se observa que cuando la evaluación de la fiabilidad es robusta los mapas derivados de percepción remota no alcanzan valores tan altos incluso con imágenes de mejor resolución espectral como Landsat.

Tomando en cuenta resultados basados en otro tipo de datos podemos decir que aun sin el uso de información auxiliar, se lograron generar mapas con una fiabilidad satisfactoria, sobretodo tomando en cuenta que la medición de la fiabilidad fue por demás rigurosa. Por ejemplo, Couturier et al. (2009) evalúan la confiabilidad de una clasificación basada en Landsat en una región en el estado de Michoacán, encontrando una fiabilidad promedio de usuario de 76% y de productor de 73%, mientras que en este estudio encontramos 80% y 74% respectivamente con datos combinados sin el uso de información auxiliar en el mismo estado. En Estados Unidos, Zhu et al. (2000) y Franklin et al. (2000) reportan que la fiabilidad de mapas obtenidos a través de clasificaciones digitales de imágenes Landsat con base en un sistema clasificatoria similar al del IGBP es del orden de 70% con algunas categorías presentando una gran cantidad de errores (fiabilidad < 40%).

De acuerdo a las consideraciones presentadas, podemos decir que es factible utilizar la información que provee el sensor MODIS para contar con un monitoreo frecuente (anual) del país, brindando información entre dos esfuerzos de mapeo con imágenes de mayor resolución espacial. También es posible utilizar los datos MODIS combinados con datos Landsat o SPOT haciendo posible analizarlos con otro enfoque que tome en cuenta las ventajas de cada tipo de sensor o para rellenar espacios con nube o ruido.

Trabajos futuros deben enfocarse al desarrollo de una combinación de productos y métodos de preprocesamiento y análisis que provean información confiable, consistente y reproducible que alcancen o excedan los límites mínimos de fiabilidad para programas operacionales. Es fundamental contar con equipos de trabajo multidisciplinarios que aprovechen los conocimientos derivados principalmente de temas biológicos físicos y matemáticos que permitan desarrollar metodologías y algoritmos específicos para el análisis de datos en áreas diversas.

La utilización del sensor MODIS ha dejado una gran experiencia y conocimiento en la comunidad científica, y debido a que el final de sus operaciones se acerca ya se prepara el sensor que dará continuidad a este tipo de datos. Este sensor es VIIRS, siglas en inglés de Visible

Infrared Imager Radiometer Suite. Los algoritmos de VIIRS surgen del aprovechamiento y mejora de la herencia de MODIS, con el objetivo de dar continuidad a estos datos con coherencia.

Por las ventajas que tiene contar con una alta resolución tanto espacial como temporal se comienza a avanzar en sistemas que brinden ambos tipos de información, esto a través de “constelaciones de satélites” es decir, varios sensores similares montados en diferentes satélites.

Por ejemplo la constelación *Disaster Monitoring Constellation (DMC)* brinda imágenes con una resolución de 22 m y un amplio barrido de más de 600 km. El uso combinado de los satélites de esta constelación proporciona una capacidad única de observación de la Tierra, con más de una visita diaria a cualquier lugar del globo (<http://www.deimos-imaging.com/sobre-dmi>). La experiencia adquirida en el análisis de datos de alta resolución temporal como MODIS aportará sin duda elementos para analizar este tipo de datos.

## 7.2 CONCLUSIONES

El presente trabajo de tesis se llevó a cabo pretendiendo aportar elementos enfocados a resolver una problemática fundamental; aportar información sobre datos y metodologías que puedan ser implementadas en el monitoreo del uso y la cobertura terrestre que puedan ser aplicadas en pro del manejo y protección del ambiente.

Las características de MODIS tienen un gran potencial de aplicación en áreas extensas como es el territorio mexicano, siendo además de acceso gratuito y frecuencia temporal diaria, además representan una alternativa para mapear el uso y la cobertura de suelo a nivel regional ya que cuenta con una diversidad de productos que permiten desarrollar una variedad de aproximaciones de procesamiento y aplicaciones.

En este trabajo se presenta una propuesta de análisis de la capa de calidad provista en los datos MODIS para filtrar y reducir los datos de manera efectiva. Se estableció que puede alcanzarse un nivel de fiabilidad aceptable considerando categorías a un nivel más detallado que el IGBP y compatible con otros esfuerzos de monitoreo nacionales.

Se determinó que los mejores datos para el proceso de clasificación son los compuestos espectrales de 8 días con fiabilidades significativamente mayores a los índices de vegetación y similares a los de las imágenes diarias de reflectancia de la superficie pero de más fácil proceso.

Se destacaron las ventajas y fallas de los dos métodos de clasificación más populares, uno en la línea paramétrica y otro en la no paramétrica, de los cuales el desempeño del método paramétrico (MV) es superior y más consistente.

Se logró incorporar información adicional derivada de mapas de probabilidad *a priori* logrando aumentos significativos en la fiabilidad de los mapas generados a través del método de Máxima Verosimilitud. Se exploraron las capacidades de la incorporación de información temática además de la espectral en una clasificación de redes neurales buscando los mejores parámetros.

Se definieron métodos sólidos y robustos para medir la fiabilidad en dos aproximaciones distintas; contra datos de campo mediante una aproximación difusa y por medio de la validación utilizando conjuntos de datos independientes en la fase de entrenamiento con respecto a la fase de verificación.

Se determinó que tomando en cuenta el contexto y el alcance logrado con este tipo de datos MODIS cuenta con el potencial para ser implementado en el mapeo a nivel nacional para complementar otros mapeos con imágenes de más alta resolución.

Por último se concluye que toda la experiencia ganada a partir del estudio de los datos MODIS podrá ser implementada con sus sucesores, por un lado VIIRS, que surge directamente del mismo equipo científico que respalda a MODIS, así como en datos provenientes de constelaciones de sensores que en conjunto brindan información con una alta resolución tanto espacial como temporal.

Las posibilidades son ilimitadas y el reto es cada vez más grande para la comunidad científica a nivel mundial ya que es urgente contar con información confiable y expedita que sirva como base a planes de manejo y conservación de los ecosistemas a nivel local, regional, nacional y mundial.

## CAPÍTULO VIII. LITERATURA CITADA

---

- Acerbi-Junior, F. W., Clevers, J. G., & Schaepman, M. E. (2006). The assessment of multi-sensor image fusion using wavelet transforms for mapping the Brazilian Savanna. (J. C. F.W. Acerbi-Junior, Ed.) *International Journal of Applied Earth Observation and Geoinformation* , 8 (4), 278-288.
- Atkinson, P. M., Dash, J., & Jeganathan, C. (2009). Analysing the effect of different geocomputational techniques on estimating phenology in India. In B. G. Lees, & S. W. Laffan (Ed.), *10th International Conference on GeoComputation, UNSW, Sydney*. Sydney.
- Bagan, H., Wang, Q., Watanabe, M., Yang, Y., & J, M. (2005). Land cover classification from MODIS EVI times-series data using SOM neuronal network. *International Journal of Remote Sensing* , 26, 4999 – 5012.
- Barnes, W. L., Pagano, T. S., & Salomonson, V. V. (1998). Prelaunch characteristics of the Moderate Resolution Imaging Spectrometer (MODIS) on EOS-AM/1. *IEEE Transactions on Geoscience and Remote Sensing* , 36, 1088-1100.
- Belward, A. S., Estes, J. E., & Klin, K. D. (n.d.). The IGBP-DIS Global 1-km Land-Cover Data Set DISCover: A Project Overview. *Photogrammetric Engineering and Remote Sensing* .
- Bishop, C. M. (1995). *Neuronal Networks for Pattern Recognition*. Oxford: University Press.
- Boles, S. H., Xiao, X., Liu, J., Zhang, Q., Munkhtuya, S., Chen, S., et al. (2004). Land cover characterization of Temperate East Asia using multi-temporal VEGETATION sensor data. *Remote Sensing of Environment* , 90, 477-489.
- Borak, J. S., & Strahler, A. H. (1999). Feature selection and land cover classification of a MODIS-like data set for a semiarid environment. *International Journal of Remote Sensing* , 20, 919-938.
- Boschetti, Stroppiana, M. D., Brivio, P. A., & Bocchi, S. (2009). Multi-year monitoring of rice crop phenology through time series analysis of MODIS images. *International Journal of Remote Sensing* , 30 (18), 4643-4662.
- Braswell, B. H., Hagen, S. C., Salas, W. A., & Frolking, S. E. (2003). A multivariable approach for mapping sub-pixel land cover distributions using MISR and MODIS: An application in the Brazilian Amazon. *Remote Sensing Environment* , 87, 243– 256.
- Brown, J. F., Loveland, T. R., Merchant, J. W., Reed, B. C., & Ohlen, D. O. (1993). Using multi-source data in global land cover characterization: concepts, requirements and methods. *Photogrammetric Engineering & Remote Sensing* , 59, 977-987.
- Card, H. D. (1982). Using known map category marginal frequencies to improve estimates of thematic map accuracy. *Photogrammetric Engineering and Remote Sensing* , 48, 431-439.

- Carrão, H., Gonçalves, P., & Caetano, M. (2008). Contribution of multispectral and multitemporal information from MODIS images to land cover classification. *Remote Sensing of Environment*, 112, 986–997.
- Chargoy, C. (1999). La selva agrícola tropical: un ejemplo en el sur de Quintana Roo, México. *Red de Gestión de Recursos Naturales*, 2 (14), 22-29.
- Chen, S. (2005). Initial assessment of terra-MODIS reflectance data structure for land surface applications in northeast Asia. *Proceedings of the Geoscience and Remote Sensing Symposium*. 6, pp. 3899-3901. IEEE International.
- Chen, S., & Rao, P. (2009). Regional Land Degradation Mapping Using MODIS Data and Decision Tree (DT) Classification in a Transition Zone between Grassland and Cropland of Northeast China. *2009 First IEEE International Conference on Information Science and Engineering*, (pp. 1395-1398).
- Chiappy, C., & Gama, L. (2004). Modificaciones y Fragmentacion de los Geocompejos Tropicales de la Peninsula de Yucatan. *Univerisidad y Ciencia*, Numero Especial I, 17-25.
- Chuvienco, E. (1990). *Fundamentos de Teledetección Espacial*. Madrid, Spain: Rialp.
- Chuvienco E. 2002. Teledetección Ambiental: La observación de la Tierra desde el Espacio. Ariel Ciencia. Barcelona España. Pp. 586.
- Chuvienco, E., Riaño, D., Aguado, I., & Cocero, D. (2002). Estimation of fuel moisture content from multitemporal analysis of Landsat Thematic Mapper reflectance data: applications in fire danger assessment. *International Journal of Remote Sensing*, 23 (11), 2145–2162.
- Cihlar, J. (2000). Land cover mapping of large areas from satellites: status and research priorities. *International Journal of Remote Sensing*, 21 (6-7), 1093-1114.
- Cohen, W. B., Maierberger, T. K., Yang, Z., Gower, S. T., Turner, D. P., Ritts, W. D., et al. (2003). Comparisons of land cover and LAI estimates derived from ETM+. *Remote Sensing of the Environment*, 88, 233–255.
- Colditz, R. R. (2008). *Time series generation and classification of MODIS data for land cover mapping*. Tesis de Doctorado, Universidad de Würzburg, Würzburg, Alemania.
- Conabio. (2009). *Imágenes MODIS*. Retrieved 2009, from Comisión Nacional para el Conocimiento: [http://www.conabio.gob.mx/informacion/geo\\_espanol/modis/modis\\_ftp.html](http://www.conabio.gob.mx/informacion/geo_espanol/modis/modis_ftp.html)
- Couturier, S. (En prensa). Evaluación de la cobertura vegetal en México: aciertos y sesgos del mapa derivado del sensor MODIS. In J. F. Mas (Ed.), *Aplicaciones del sensor MODIS para el monitoreo del territorio*. Distrito Federal, DF, Mexico: INE-UNAM.
- Couturier, S., Mas, J. F., Cuevas, G., Benítez, J., Vega, A., & Tapia, V. (2009). A thematic-focused accuracy assessment of land cover maps for highly biodiverse regions. *Photogrammetric Engineering and Remote Sensing*, 75 (7), 789-806.

- Defries, R. S., Field, C. B., & Fung, I. (1995). Mapping the land surface for global atmosphere-biosphere models: toward continuous distributions of vegetations's functional properties. *Journal of Geophysical Research* , 100 (20), 867-882.
- Defries, R. S., Hansen, M., Townshend, J. R., & Sohlberg, R. (1998). Global land cover classification at 8 km spatial resolution: the use of training data derived from Landsat imagery in decision tree classifiers. *International Journal of Remote Sensing* , 19, 3141–3168.
- Eastman, J. R., & Fulk, M. (1993). Long sequence time series evaluation using standardized principal components. *Photogrammetric Engineering and Remote Sensing* , 59 (8), 1307-1312.
- Eastman, J.R., (2003). Idrisi Kilimanjaro (Worcester, MA: Clark University).
- Eggen-McIntosh, S., Borbolla Munoz, E., Ornelas de Anda, J. L., Zhu, Z., & Evans, D. L. (1992). Forest cover mapping of Mexico using AVHRR imagery. *Mapping and Monitoring Global Change: ASPRS/ACSM/RT 1992 Technical Papers, Washington, D.C. American Society for Photogrammetry and Remote Sensing* (pp. 410-416). Bethesda, MD: ASPRS/ACSM.
- Evans, D. L., Zhu, Z., Eggen-McIntosh, S., Garcia Mayoral, P., & Omelas de Anda, J. L. (1992). *Mapping Mexico's forest lands With Advanced Very High Resolution Radiometer*. New Orleans: USDA Forest Service, Southern Forest Experiment Station.
- FAO 1997 Progress Report. (1997). *FAO 1997 Progress Report, Integrated approach to the planning and management of land resources*.
- Ferreira, L. G., Yoshioka, H., Huete, A., & Sano, E. E. (2004). Optical characterization of the Brazilian Savanna physiognomies for improved land cover monitoring of the cerrado biome: preliminary assessments from an airborne campaign over an LBA core site. *Journal of Arid Environments* , 56 (3), 425-447.
- Ferreira, L. G., Yoshioka, H., Huete, A., & Sano, E. E. (2003). Seasonal Landscape and Spectral Vegetation Index Dynamics in the Brazilian Cerrado: An Analysis within The Large-Scale Biosphere-Atmosphere Experiment in Amazonia (LBA). *Remote Sensing of Environment* , 87, 534-550.
- Foley, J. A., DeFries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., et al. (2005). Global Consequences of Land Use. *Science* , 309, 570-574.
- Foody, G. M. (2000). Mapping Land Cover from Remotely Sensed Data with a Softened Feedforward Neuronal Network Classification. *Journal of Intelligent and Robotic Systems* , 29, 433-449.
- Foody, G. M. (2004). Supervised image classification by MLP and RBF neuronal networks with and without an exhaustively defined set of classes. *International Journal of Remote Sensing* , 25, 3091-3104.
- Friedl, M. A., McIver, D. K., Hodges, J. C., Xhang, X. Y., Muchoney, D., Strahler, A. H., et al. (2002). Global land cover mapping from MODIS: algorithms and early results. *Remote Sensing of Environment* , 83, 287–302.

- Friedl, M. A., Menashe, D. S., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., et al. (2010). MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sensing of Environment*, 114, 168-182.
- Giri, C., Zhu, Z., & Reed, B. (2005). A comparative analysis of the Global Land Cover 2000 and MODIS land cover datasets. *Remote Sensing of Environment*, 94, 123-132.
- Gitelson, A. A., Viña, A., Arkebauer, T. J., Rundquist, D. C., Keydan, G., & Leavitt, B. (2003). Remote estimation of leaf area index and green leaf biomass in maize canopies. *Geophysical Research Letters*, 30, 1248.
- Gitelson, A. A., Viña, A., Ciganda, V., Rundquist, D. C., & Arkebauer, T. J. (2005). Remote estimation of canopy chlorophyll content in crops. *Geophysical Research Letters*, 32, L08403.
- Gitelson, A., & Kaufman, Y. (1998). MODIS NDVI Optimization to Fit the AVHRR Data Series - Spectral Considerations. *Remote Sensing of Environment*, Vol. 66, No. 3, 66 (3), 343-350.
- Glenn, E. P., Huete, A. R., Nagler, P. L., & Nelson, S. G. (2008). Relationship between sensed vegetation indices, canopy attributes and plant physiological processes: What vegetation indices can and cannot tell us about the landscape. *Sensors*, 8, 2136-2160.
- Goddard Earth Sciences Data and Information Services Center. (2009). Retrieved December 2009, from <http://daac.gsfc.nasa.gov/>
- Goddard Space Flight Center. (2009). Retrieved December 2009, from Goddard Space Flight Center: <http://www.nasa.gov/centers/goddard/home/index.html>
- Gonçalves, P., Carrão, H., Pinheiro, A., & Caetano, M. (2006). Land cover classification with Support Vector Machine applied to MODIS imagery. In A. Marçal (Ed.), *Global Developments in Environmental Earth Observation from Space* (pp. 517-526). Rotterdam: Millpress.
- Heiskanen, J., & Kivinen, S. (2008). Assessment of multispectral, temporal and angular MODIS data for tree cover mapping in the tundra-taiga transition zone. *Remote Sensing of Environment*, 112, 2367-2380.
- Holben, B. N. (1986). Characteristics of maximum-value composite images from temporal AVHRR data. *International Journal of Remote Sensing*, 7 (11), 1417.
- Huete, A. (1988). A Soil-Adjusted Vegetation Index (SAVI). *Remote Sensing of Environment*, 25, 295-309.
- Huete, A. (en prensa). NEED CHAPTER TITLE. In J. F. Mas (Ed.), *Aplicaciones del sensor MODIS para el monitoreo del territorio*. Distrito Federal, DF, Mexico: INE-UNAM.
- Huete, A. R., Liu, H. Q., Batchily, K., & Van Leeuwen, W. J. (1997). A comparison of vegetation indices over a global set of TM images for EOS-MODIS. *Remote Sensing of Environment*, 59 (3), 440-451.

- Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS Vegetation indices. *Remote Sensing of Environment*, 83, 195–213.
- Hughes, G. F. (1968). On the mean accuracy of statistical pattern recognizers. *IEEE Transactions on Information Theory*, IT-14, 55-63.
- Hutchinson, C. F. (1982). Techniques for Combining Landsat and Ancillary Data for Digital Classification Improvement. *Photogrammetric Engineering and Remote Sensing*, 48, 123-130.
- Jakubauskas, M. E., Legates, D. R., & Kastens, J. H. (2001). Harmonic analysis of time-series AVHRR NDVI data. *Photogrammetric Engineering & Remote Sensing*, 67 (4), 461–470.
- Ji, L., & Peters, A. J. (2007). Performance evaluation of spectral vegetation indices using a statistical sensitivity function. *Remote Sensing of Environment*, 106 (1), 59-65 .
- Jonsson, P., & Eklundh, L. (2002). Seasonality extraction by function fitting to time-series of satellite sensor data. *IEEE Transactions on Geoscience and Remote Sensing*, 40 (8), 1824–1832.
- Jung, M., Henkel, K., Herold, M., & Churkina, G. (2006). Exploiting synergies of global land cover products for carbon cycle modeling. *Remote Sensing of Environment*, 101, 534-553.
- Justice, C. L., Giglio, L., Korontzi, S., Owens, J., Morisette, J. T., Roy, D., et al. (2002). The MODIS Fire Products. *Remote Sensing of Environment*, 83, 244-262.
- Justice, C. O., Townshend, E. F., Vermote, E., Masuoka, E., Wolfe, R. E., Saleous, N., et al. (2002a). An overview of MODIS Land data processing and product status. *Remote Sensing of Environment*, 83, 3-15.
- Justice, C., Vermote, E., Townshend, J. R., Defries, R., Roy, D. P., Hall, D. K., et al. (1998). The Moderate Resolution Imaging Spectroradiometer (MODIS): Land remote sensing for global change research. *IEEE Transactions on Geoscience and Remote Sensing*, 36 (4), 1228-1249.
- Landgrebe, D. (2003). *Signal theory methods in multispectral remote sensing*. Hoboken, New Jersey: Wiley.
- Latifovic, R., Zhu, Z., Cihlar, J., Giri, C., & Olthof, I. (2004). Land cover mapping of North and Central America—Global Land Cover 2000. *Remote Sensing of Environment*, 89, 116–127.
- Levy, R. C., Remer, L., Mattoo, S., Vermote, E., & Kaufman, Y. J. (2007). Second-generation algorithm for retrieving aerosol properties over land from MODIS spectral reflectance. *Journal of Geophysical Research*, 13211.
- Liang, S., Fang, H., Chen, M., Shuey, C. J., Walthall, C. L., Daughtry, C. S., et al. (2002). Validating MODIS land surface reflectance and albedo products: methods and preliminary results. *Remote Sensing of Environment*, 83 (1-2), 149-162.
- Liu, J., Liu, M., Zhang, D., & Deng, X. (2003). Study on spatial patterns of land use change in China during 1995–2000. *Sci. China*, 46, 373-384.

- Lobser, S. E., & Cohen, W. B. (2007). The MODIS tasselled cap: land cover characteristics expressed through transformed MODIS data. *International Journal of Remote Sensing*, *28*, 5079-5101.
- Loveland, T. R., Merchant, J. W., Brown, J. F., Ohlen, D. O., Reed, B. C., Olson, P., et al. (1995). Seasonal land-cover regions of the United States. *Annals of the Association of American Geographers*, *85*, 339–355.
- Loveland, T. R., Reed, B. C., Brown, J. F., Ohlen, D. O., Zhu, Z., Yang, L., et al. (2000). Development of a global land cover characteristics database and IGBP DISCover from 1-km AVHRR data. *International Journal of Remote Sensing*, *21*, 1303–1330.
- Mas J. F., J. R. Díaz-Gallegos, A. Pérez-Vega. 2003. Evaluación de la confiabilidad temática de mapas o de imágenes clasificadas: una revisión. Investigaciones geográficas. Boletín del Instituto de Geografía, UNAM, No. 51, Pp 53-72.
- Mas, J. F. (2004). Mapping land use/cover in a tropical coastal area using satellite sensor data, GIS and artificial neuronal networks: Estuarine. *Coastal and Shelf Science*, *59*, 219-230.
- Mas, J.F., (2005). Clasificación de imágenes de teledetección por redes neurales artificiales: una revisión. *Revista SELPER*, *22*(2):16-24.
- Mas, J. F., & Flores, J. J. (2008). The application of artificial neuronal networks to the analysis of remotely sensed data. *International Journal of Remote Sensing*, *29*, 617-663.
- Maselli, F., Conese, C., De Filippis, T., & Romani, M. (1995). Integration of ancillary data into a maximum-likelihood classifier with nonparametric priors. *ISPRS Journal of Photogrammetry and Remote Sensing*, *50*, 2-11.
- Maselli, F., Conese, C., Petkov, L., & Resti, R. (1992). Inclusion of prior probabilities derived from a nonparametric process into the maximum-likelihood classifier. *Photogrammetric Engineering and Remote Sensing*, *58*, 201-207.
- Matsuoka, M., Hayasaka, T., Fukushima, Y., & Honda, Y. (2007). Land Cover in East Asia Classified Using Terra MODIS and DMSP OLS Products. *International Journal of Remote Sensing*, *28* (2), 221-248.
- Maussel, P., Kramber, J., & Lee, J. (1990). Optimum band selection for supervised classification of multispectral data. *Photogrammetric Engineering & Remote Sensing*, *56*, 55-60.
- Miettinen, J., Wong, C. M., & Liew, S. C. (2008). New 500 m spatial resolution land cover map of the western insular Southeast Asia region. *International Journal of Applied Earth Observations and Geoinformation*, *29* (20), 6075 - 6081.
- Mingwei, Z., & Qingbo, Z. (2008). Crop discrimination in Northern China with double cropping systems using Fourier analysis of time-series MODIS data. *International Journal of Applied Earth Observations and Geoinformation*, *10* (4), 476-485.

- Mittermeier, R., & Goettsch, C. (1992). La importancia de la diversidad biológica de México, pp. 57-62 en: . In *México ante los retos de la biodiversidad*. Mexico City, Mexico: Conabio.
- Morissette, J. T., Nickeson, J. E., Davis, P., Wang, Y., Tian, Y., Woodcock, C. E., et al. (2003). High spatial resolution satellite observations for validation of MODIS land products: IKONOS observations acquired under the NASA Scientific Data Purchase. *Remote Sensing of Environment* , 88 (1-2), 100-110.
- Morissette, J., Privette, J., & Justice, C. (2002). A framework for the validation of MODIS Land products. *Remote Sensing of Environment* , 83, 77-96.
- Muukkonen, P., & Heiskanen, J. (2007). Biomass estimation over a large area based on standwise forest inventory data, ASTER and MODIS satellite data: a possibility to verify carbon inventories. *Remote Sensing of Environment* , 107 (4), 617–624.
- NASA - Warehouse Inventory Search Tool*. (n.d.). Retrieved December 2009, from <http://redhook.gsfc.nasa.gov/>
- NASA MODIS Web - Components of MODIS*. (2008). Retrieved 2008, from The Moderate Resolution Imaging Spectroradiometer website that houses all central information on the MODIS project: <http://modis.gsfc.nasa.gov/about/specifications.php>
- Nightingale, J. M., Coops, N. C., Waring, R. H., & Hargrove, W. W. (2007). Comparison of MODIS Gross Primary Production estimates for forests across the U.S.A. with those generated by a simple process model. *Remote Sensing of Environment* , 109, 500-509.
- Nightingale, J., Nickeson, J. E., Justice, C. O., Baret, F., Garrigues, S., Wolfe, R., et al. (2009). Global validation of EOS land products, lessons learned and future challenges: A MODIS case study. *33rd Proc. of the International Symposium on Remote Sensing of Environment*. Stresa, Italy.
- Ozdogan, M., & Gutman, G. (2008). Comparisons of land cover and LAI estimates derived from ETM+. *Remote Sensing of Environment* , 112, 3520–3537.
- PGIGeomatics - Signature Separability*. (2007). Retrieved December 2009, from ImageWorks Multispectral: <http://www.pcigeomatics.com/cgi-bin/pcihlp/CLWORKS|Signature+Separability>
- Reed, B., Brown, J., VanderZee, D., Loveland, T., Merchant, J., & Ohlen, D. O. (1994). Measuring phenological variability from satellite imagery. *Journal of Vegetation Science* , 5, 703-714.
- Roerink, G. J., Menenti, M., & Verhoef, W. (2000). Reconstructing cloudfree NDVI composites using Fourier analysis of time series. *International Journal of Remote Sensing* , 21 (9), 1911-1917.
- Roy, D. P., Borak, J. S., Devadiga, S., Wolfe, R. E., Zheng, M., & Descloitres, J. (2002). The MODIS Land product quality assessment approach. *Remote Sensing of Environment* , 83, 62–76.
- SARH. (1992). *Inventario Forestal Nacional de Gran Visión*. Mexico: Subsecretaría Forestal y de Fauna Silvestre, Secretaría de Agricultura y Recursos Hidráulicos (SARH).

- Sarle, W. S. (1994). Neuronal networks and statistical models (resumen). *Proceedings of the 19th Annual SAS Group Conference*, (pp. 1538-1550). Cary, NC.
- Sedano, F., Gómez, D., Gong, P., & Biging, G. S. (2008). Tree density estimation in a tropical woodland ecosystem with multiangular MISR and MODIS data. *Remote Sensing of Environment*, *112* (5), 2523-2537.
- Sedano, F., Gong, P., & Ferrao, M. (2005). Land cover assessment with MODIS imagery in southern African Miombo ecosystems. *Remote Sensing of Environment*, *98*, 429 – 441.
- Shimabukuro, Y. E., Arai, E., Duarte, V., Freitas, R. M., Lima, A., Valeriano, D. M., et al. (2009). Fraction images derived from Terra Modis data for mapping burnt areas in Brazilian Amazonia. *International Journal of Remote Sensing*, *30* (6), 1537-1546.
- Sivanpillai, R., & Latchininsky, A. V. (2007). Mapping locust habitats in the Amudarya River Delta, Uzbekistan with multi-temporal MODIS imagery. *Environmental Management*, *39* (6), 876-886.
- Stehman, S. V., & Czaplewski, R. L. (1998). Design and analysis for thematic map accuracy assessment: fundamental principles. *64*, 331–344.
- Tanré, D., Deroo, C., Duhaut, P., Herman, M., Morcrette, J., Perbos, J., et al. (1986). *Simulation of the Satellite Signal in the Solar Spectrum (5S)*. Lille, France: Laboratoire d'Optique Atmosphérique, Université des Sciences et Techniques de Lille.
- Tottrup, C. (2007). Forest and Land Cover Mapping in a Tropical Highland Region. *Photogrammetric Engineering & Remote Sensing*, *73* (9), 1057-1065.
- Townshend, J. R., Justice, C. O., & Kalb, V. (1987). Characterization and Classification of South American Land Cover Types Using Satellite Data. *International Journal of Remote Sensing*, *8* (8), 1189-1207.
- Townshend, J., Justice, C., Li, W., Gurney, C., & McManus, J. (1991). Global Land cover classification by remote sensing: Present capabilities and future possibilities. *Remote Sensing of the Environment*, *35*, 243—255.
- Tucher, C. T., Townshend, J. R., & Goff, T. E. (1985). African land cover classification using satellite data. *Science* *227*, *59*, 369–375.
- Van Dijk, A., Callis, S. L., Sakamoto, C. M., & Decker, W. L. (1987). Smoothing vegetation index profiles: An alternative method for reducing radiometric disturbances in NOAA/AVHRR data. *Photogrammetric Engineering and Remote Sensing*, *53*, 1059-1067.
- Van Leeuwen, W. J., Huete, A. R., & Laing, T. W. (1999). MODIS Vegetation Index Compositing approach: a prototype with AVHRR data. *Remote Sensing of Environment*, *69*, 264-280.
- Vermote, E. F., & Kotchenova, S. (2008). Atmospheric correction for the monitoring of land surfaces. *Journal of Geophysical Research*, *113*, D23S90, doi:10.1029/2007JD009662.

- Vermote, E. F., El Saleous, N. Z., Justice, C. O., Kaufman, Y. J., Privette, J., Remer, L., et al. (1997). Atmospheric correction of visible to middle infrared EOS-MODIS data over land surface, background, operational algorithm and validation. *Journal of Geophysical Research*, Vol. , 102., d14,17,131-17,141., d14,17,131-17,141.
- Vermote, E., El Salleous, N., & Justice, C. O. (2002). Atmospheric correction of MODIS data in the visible to near infrared: first results. *Remote Sensing of the Environment* , 83 (1-2), 97-111.
- Wessels, K. J., De Fries, R. S., Dempewolf, J., Anderson, L. O., Hansen, A. J., & Powel, S. L. (2004). Mapping regional land cover with MODIS data for biological conservation: Examples from the Great Yellowstone ecosystem, USA and Pará State, Brazil. *Remote Sensing of Environment*, 92, , 92, 67–83.
- Westra, T., & De Wulf, R. R. (2007). Monitoring Sahelian floodplains using Fourier analysis of MODIS time-series data and artificial neuronal networks. *International Journal of Remote Sensing* , 28 (7), 1595 - 1610.
- Woodcock, C., & Gopal, S. (2000). Fuzzy set theory and thematic maps: accuracy assessment and area estimation. *International Journal of Geographical Information Science* , 14, 153-172.
- Wulder, M. A., Franklin, S. E., White, J. C., Linke, J., & Magnussen, S. (2006). An accuracy assessment framework for large-area land-cover classification products derived from medium-resolution satellite data. *International Journal of Remote Sensing* , 27, 663–683.
- Xavier, A. C., Rudorff, B. F., Shimabukuro, Y. E., Berka, L. M., & Moreira, M. A. (2006). Multi-temporal analysis of MODIS data to classify sugarcane crop. *International Journal of Remote Sensing* , 27 (4), 755-768.
- Xiao, X., Boles, S., Liu, J., Zhuang, D., Frohling, S., Li, C., et al. (2005). Mapping paddy rice agriculture in southern China using multi-temporal MODIS images. *Remote Sensing of Environment* , 95, , 480–492.
- Zhang, X., Sun, R., Zhang, B., & Tong, Q. (2008). Land cover classification of the North China Plain using. *ISPRS Journal of Photogrammetry & Remote Sensing* , 63, 476–484.

## ANEXOS

ANEXO1 - Fiabilidades de usuario y productor en el área del Tancitaro Michoacán

<b>MV</b>	<b>MV-C8</b>		<b>MV-DIA</b>		<b>MV-IV</b>		<b>MV-COMB</b>	
	<b>FIA US</b>	<b>FIA PROD</b>	<b>FIA US</b>	<b>FIA PROD</b>	<b>FIA US</b>	<b>FIA PROD</b>	<b>FIA US</b>	<b>FIA PROD</b>
Pino	55%	64%	47%	52%	44%	70%	57%	73%
Aguacate	92%	80%	80%	77%	86%	81%	92%	81%
Sin Veg	94%	59%	96%	61%	94%	47%	98%	63%
SBC	65%	88%	74%	81%	63%	64%	81%	84%
Ag Riego	84%	81%	67%	84%	68%	70%	71%	78%
Área urbana	86%	77%	89%	63%	38%	33%	89%	72%
pastizal	86%	72%	89%	72%	38%	67%	89%	79%
BPQ	67%	71%	73%	73%	63%	75%	75%	77%
Ag Temporal	59%	54%	38%	33%	18%	16%	65%	57%

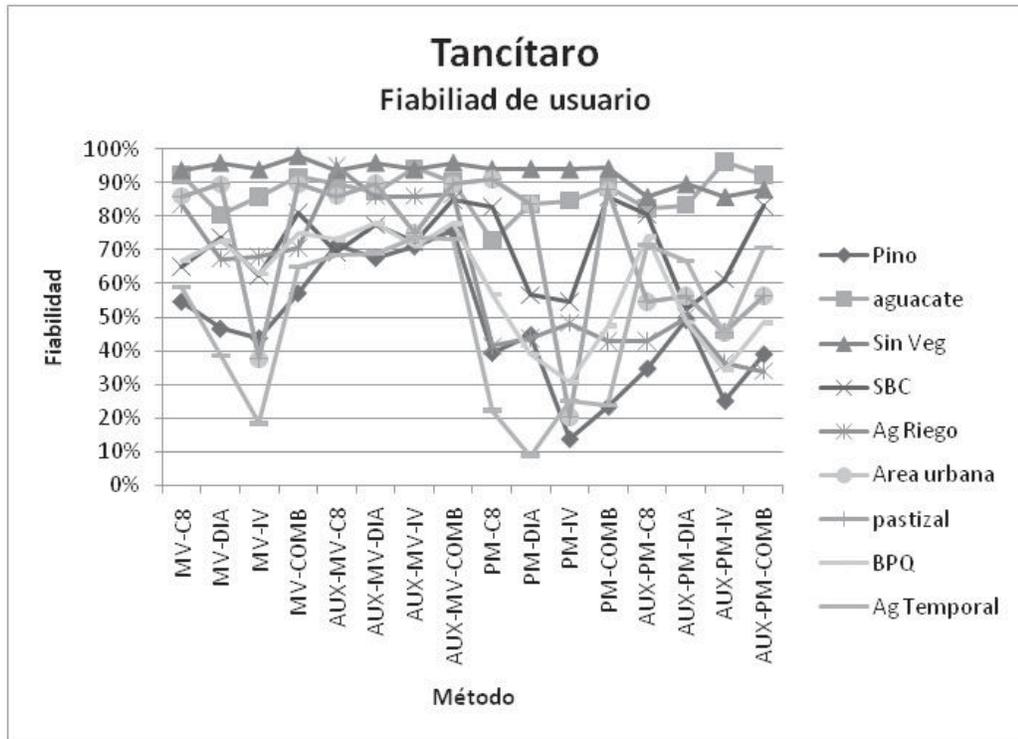
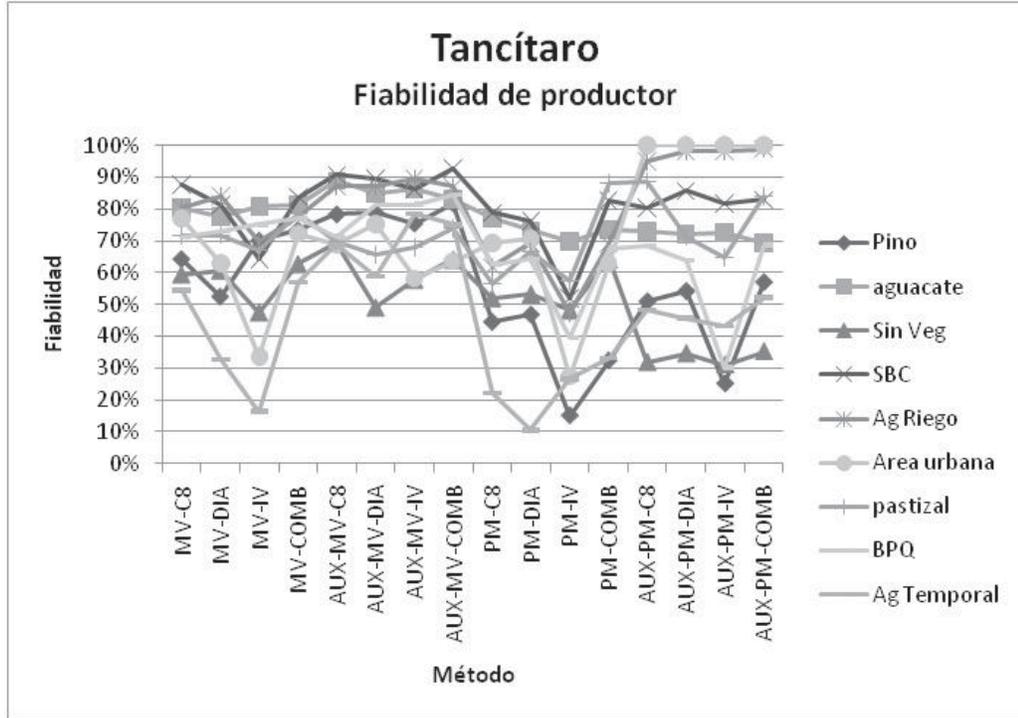
<b>AUX-MV</b>	<b>ANC-MV-C8</b>		<b>ANC-MV-DIA</b>		<b>ANC-MV-IV</b>		<b>ANC-MV-COMB</b>	
	<b>FIA US</b>	<b>FIA PROD</b>	<b>FIA US</b>	<b>FIA PROD</b>	<b>FIA US</b>	<b>FIA PROD</b>	<b>FIA US</b>	<b>FIA PROD</b>
Pino	71%	78%	68%	79%	71%	75%	76%	82%
Aguacate	90%	89%	87%	85%	94%	86%	90%	83%
Sin Veg	94%	70%	96%	49%	94%	58%	96%	64%
SBC	69%	91%	77%	90%	72%	86%	85%	93%
Ag Riego	95%	87%	86%	87%	86%	89%	86%	87%
Área urbana	86%	69%	90%	75%	75%	58%	89%	64%
pastizal	86%	70%	90%	66%	75%	68%	89%	74%
BPQ	73%	71%	78%	81%	71%	81%	78%	85%
Ag Temporal	68%	69%	69%	59%	74%	79%	73%	75%

<b>PMC</b>	<b>PMC-C8</b>		<b>PMC-DIA</b>		<b>PMC-IV</b>		<b>PMC-COMB</b>	
	<b>FIA US</b>	<b>FIA PROD</b>	<b>FIA US</b>	<b>FIA PROD</b>	<b>FIA US</b>	<b>FIA PROD</b>	<b>FIA US</b>	<b>FIA PROD</b>
Pino	39%	44%	45%	47%	14%	15%	24%	32%
Aguacate	73%	77%	84%	73%	85%	70%	89%	73%
Sin Veg	94%	52%	94%	53%	94%	48%	94%	64%
SBC	83%	79%	57%	76%	55%	52%	86%	83%
Ag Riego	41%	62%	44%	69%	48%	47%	43%	69%
Area urbana	91%	69%	84%	71%	20%	27%	88%	63%
pastizal	91%	57%	84%	67%	20%	58%	88%	88%
BPQ	57%	63%	39%	64%	31%	40%	47%	68%
Ag Temporal	22%	22%	9%	10%	25%	27%	24%	33%

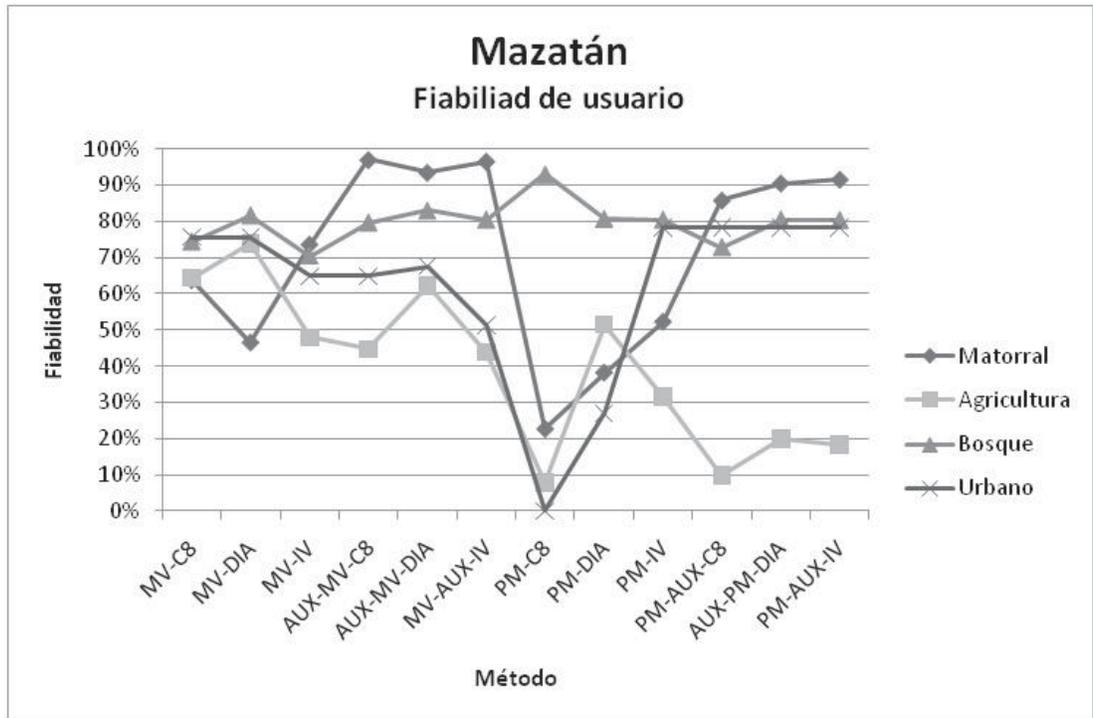
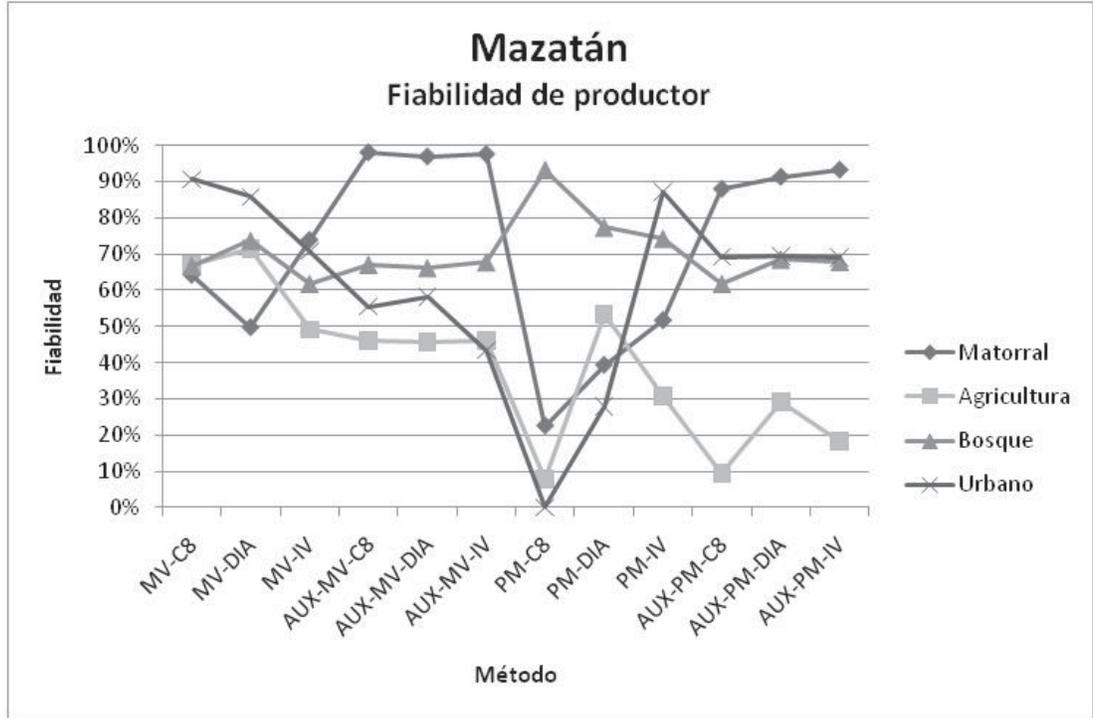
  

<b>AUX-PMC</b>	<b>ANC-PMC-C8</b>		<b>ANC-PMC-DIA</b>		<b>ANC-PMC-IV</b>		<b>ANC-PMC-COMB</b>	
	<b>FIA US</b>	<b>FIA PROD</b>	<b>FIA US</b>	<b>FIA PROD</b>	<b>FIA US</b>	<b>FIA PROD</b>	<b>FIA US</b>	<b>FIA PROD</b>
Pino	35%	51%	49%	54%	25%	25%	39%	57%
Aguacate	82%	73%	83%	72%	96%	73%	92%	69%
Sin Veg	86%	32%	90%	35%	86%	31%	88%	35%
SBC	81%	80%	52%	86%	61%	82%	83%	83%
Ag Riego	43%	95%	50%	98%	36%	98%	34%	99%
Area urbana	55%	100%	56%	100%	45%	100%	56%	100%
pastizal	55%	89%	56%	71%	45%	65%	56%	84%
BPQ	74%	69%	48%	64%	34%	30%	48%	69%
Ag Temporal	71%	48%	67%	46%	44%	43%	71%	52%



ANEXO2 - Fiabilidades de usuario y productor en el área de Mazatán, Sonora

<i>MV</i>	<i>MV-C8</i>		<i>MV-DIA</i>		<i>MV-IV</i>	
	<i>FIA US</i>	<i>FIA PROD</i>	<i>FIA US</i>	<i>FIA PROD</i>	<i>FIA US</i>	<i>FIA PROD</i>
<b>Matorral</b>	64%	64%	47%	50%	74%	74%
<b>Agricultura</b>	64%	67%	74%	72%	48%	49%
<b>Bosque</b>	74%	67%	82%	74%	70%	62%
<b>Urbano</b>	76%	91%	76%	86%	65%	71%
<i>AUX-MV</i>	<i>AUX-MV-C8</i>		<i>AUX-MV-DIA</i>		<i>AUX-MV-IV</i>	
	<i>FIA US</i>	<i>FIA PROD</i>	<i>FIA US</i>	<i>FIA PROD</i>	<i>FIA US</i>	<i>FIA PROD</i>
<b>Matorral</b>	97%	98%	93%	97%	97%	98%
<b>Agricultura</b>	45%	46%	62%	46%	44%	46%
<b>Bosque</b>	80%	67%	83%	66%	80%	68%
<b>Urbano</b>	65%	56%	68%	58%	51%	43%
<i>PMC</i>	<i>PMC-C8</i>		<i>PMC-DIA</i>		<i>PMC-IV</i>	
	<i>FIA US</i>	<i>FIA PROD</i>	<i>FIA US</i>	<i>FIA PROD</i>	<i>FIA US</i>	<i>FIA PROD</i>
<b>Matorral</b>	23%	23%	38%	39%	52%	52%
<b>Agricultura</b>	8%	8%	52%	53%	32%	31%
<b>Bosque</b>	93%	93%	81%	77%	80%	74%
<b>Urbano</b>	0%	0%	27%	28%	78%	87%
<i>AUX-PMC</i>	<i>AUX-PMC-C8</i>		<i>AUX-PMC-DIA</i>		<i>AUX-PMC-IV</i>	
	<i>FIA US</i>	<i>FIA PROD</i>	<i>FIA US</i>	<i>FIA PROD</i>	<i>FIA US</i>	<i>FIA PROD</i>
<b>Matorral</b>	86%	88%	91%	91%	92%	93%
<b>Agricultura</b>	10%	9%	20%	29%	18%	18%
<b>Bosque</b>	73%	62%	80%	69%	80%	68%
<b>Urbano</b>	78%	69%	78%	69%	78%	69%



ANEXO 3 – Fiabilidades de usuario y productor en el área Maya de Quintana Roo

<i>MV</i>	MV-C8		MV-DIA		MV-IV	
	FIA US	FIA PROD	FIA US	FIA PROD	FIA US	FIA PROD
<b>SMSP</b>	92%	92%	94%	92%	85%	82%
<b>SBSP</b>	62%	56%	50%	55%	70%	61%
<b>HIDRO</b>	4%	21%	19%	45%	25%	64%
<b>AGUA</b>	82%	49%	82%	65%	72%	86%
<b>AGRI</b>	14%	16%	13%	16%	4%	11%
<b>URBANO</b>	54%	80%	60%	78%	47%	49%

<i>AUX-MV</i>	AUX-MV-C8		AUX-MV-DIA		AUX-MV-IV	
	FIA US	FIA PROD	FIA US	FIA PROD	FIA US	FIA PROD
<b>SMSP</b>	95%	91%	96%	93%	96%	90%
<b>SBSP</b>	58%	78%	52%	61%	58%	67%
<b>HIDRO</b>	0%	0%	0%	0%	0%	0%
<b>AGUA</b>	80%	49%	80%	43%	72%	34%
<b>AGRI</b>	22%	23%	19%	24%	10%	20%
<b>URBANO</b>	32%	32%	43%	68%	0%	1%

<i>PMC</i>	PMC-C8		PMC-DIA		PMC-IV	
	FIA US	FIA PROD	FIA US	FIA PROD	FIA US	FIA PROD
<b>SMSP</b>	94%	90%	69%	71%	80%	74%
<b>SBSP</b>	50%	51%	40%	60%	40%	41%
<b>HIDRO</b>	12%	40%	5%	10%	34%	88%
<b>AGUA</b>	70%	41%	64%	0%	61%	36%
<b>AGRI</b>	27%	33%	67%	61%	15%	20%
<b>URBANO</b>	54%	96%	64%	73%	23%	42%

<i>AUX-PMC</i>	AUX-PMC-C8		AUX-PMC-DIA		PMC-AUX-IV	
	FIA US	FIA PROD	FIA US	FIA PROD	FIA US	FIA PROD
<b>SMSP</b>	88%	78%	76%	78%	98%	97%
<b>SBSP</b>	57%	61%	45%	48%	57%	61%
<b>HIDRO</b>	47%	86%	6%	12%	36%	75%
<b>AGUA</b>	31%	16%	40%	22%	26%	13%
<b>AGRI</b>	25%	36%	63%	57%	1%	2%
<b>URBANO</b>	69%	97%	48%	52%	1%	2%

