Land use and land cover

MAPBIOMAS
MÉXICO

General manual

Algorithm Theoretical Basis Document (ATBD)

Collection 1.0

2026

ATBD Authors

Jean François Mas Caussel
Jonathan Vidal Solórzano Villegas
Daniel Martín Auliz Ortiz
Alexander Quevedo Chacón
Javier Osorno Covarrubias
Candelario Peralta Carreta

Directory of specialists & institutional support:

See website www.mexico.mapbiomas.org

Executive Summary

MapBiomas Mexico is an initiative of a collaborative network made up of universities, government institutions, and non-governmental organizations, organized with the objective of using quality and lower-cost technology to produce an annual series of land cover and land use maps (from 1985 onwards).

As part of the work of MapBiomas Mexico and MapBiomas International, we present the *Collection 1.0 of Annual Land Cover and Land Use Maps* of Mexico. Thanks to this collaborative effort, these maps will be available on an interactive platform for the country through the MapBiomas Mexico initiative.

The objective of this Algorithm Theoretical Basis Document (ATBD) is to provide users with an understanding of the methodological steps and computational algorithms for producing the *Coverage and Use Collection 1.0*, which includes the annual mapping of land cover and land use between 1985 and 2025.

Acronyms and Abbreviations

ATBD — Algorithm Theoretical Basis Document

AVHRR — Advanced Very High Resolution Radiometer (NOAA sensor)

AVINA — Fundación AVINA (private foundation; fiscal sponsor of MapBiomas México)

CCI — Climate Change Initiative (ESA)

CentroGeo — Centro de Investigación en Ciencias de Información Geoespacial

CICY — Centro de Investigación Científica de Yucatán

CLC — CORINE Land Cover

CONABIO — Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (National Commission for the Knowledge and Use of Biodiversity)

CONAFOR — Comisión Nacional Forestal (National Forestry Commission)

CONANP — Comisión Nacional de Áreas Naturales Protegidas (National Commission for Natural Protected Areas)

CORINE — Coordination of Information on the Environment

ECOSUR — El Colegio de la Frontera Sur

ESA — European Space Agency

ESRI — Environmental Systems Research Institute

ETM+ — Enhanced Thematic Mapper Plus (Landsat 7 sensor)

EVI2 — Two-Band Enhanced Vegetation Index

FAO — Food and Agriculture Organization of the United Nations

GEE — Google Earth Engine

GFW — Global Forest Watch

GHG — Greenhouse Gas

GLAD — Global Land Analysis & Discovery (University of Maryland laboratory)

GLC — Global Land Cover

GR — General Rules (temporal filter)

Ha — Hectares

INE — Instituto Nacional de Ecología (National Institute of Ecology)

INEGI — Instituto Nacional de Estadística y Geografía (National Institute of Statistics and Geography)

INIFAP — Instituto Nacional de Investigación Forestales, Agrícolas y Pecuarias

IUCN — International Union for Conservation of Nature

L4, L5, L7, L8, L9 — Landsat 4, 5, 7, 8, 9 satellites

LULC — Land Use and Land Cover

MMU — Minimum Mapping Unit

NALC — North American Landscape Characterization

NASA — National Aeronautics and Space Administration

NDC — Nationally Determined Contribution

NDVI — Normalized Difference Vegetation Index

NOAA — National Oceanic and Atmospheric Administration

OLI — Operational Land Imager (Landsat 8 sensor)

OLI-2 — Operational Land Imager 2 (Landsat 9 sensor)

REDD+ — Reducing Emissions from Deforestation and forest Degradation (UN Framework)

SAMOF — Sistema de Alerta y Monitoreo Forestal (Forest Alert and Monitoring System)

SEMARNAP — Secretaría de Medio Ambiente, Recursos Naturales y Pesca (former Mexican environmental ministry)

SEMARNAT — Secretaría de Medio Ambiente y Recursos Naturales (Ministry of Environment and Natural Resources)

SRTM — Shuttle Radar Topography Mission

SWIR — Shortwave Infrared

TDOM — Temporal Dark Outlier Mask (cloud shadow detection method)

TIRS — Thermal Infrared Sensor (Landsat 8 sensor)

TIRS-2 — Thermal Infrared Sensor 2 (Landsat 9 sensor)

TM — Thematic Mapper (Landsat 4/5 sensor)

UNAM — Universidad Nacional Autónoma de México (National Autonomous University of Mexico)

USGS — United States Geological Survey

WRI — World Resources Institute

Index

List of Tables

List of Figures

1. Introduction

1.1. Scope and Content of the Document

The objective of this document is to describe the theoretical basis, justification and methods applied to produce annual land cover and land use maps from 1985 to 2025 of Collection 1 of Annual Land Cover and Land Use Maps of MapBiomas Mexico.

This document covers the Landsat image classification methods (L4, L5, L7, L8, L9), the image processing architecture, and the approach to integrating the biomes and regions present in the country. It also presents historical context and background, as well as an overview of the satellite image dataset.

The specific procedures applied in each cross-cutting theme are found in the appendices (https://mexico.mapbiomas.org/en/download-of-atbds). The classification algorithms are available on the MapBiomas Mexico Github (https://github.com/mapbiomas/mexico-all-initiatives).

1.2. General Overview

The MapBiomas Mexico project began in July 2025 with the aim of supporting the understanding of land cover and land use (LULC) dynamics across the country. The project is made possible by: i) technological advancements that allow for the cloud-based processing of large amounts of spatial data using algorithms hosted on the Google Earth Engine platform; ii) the implementation of image processing methods focused on monitoring LULC; iii) the multidisciplinary technical team whose expertise enables the mapping of the territory; and iv) the support of visionary institutions and funders who back the project.

MapBiomas Mexico's products consist of annual thematic maps with a spatial resolution of 30 meters for the entire country. Their methodology uses annual mosaics of satellite images made up of information layers (spectral bands, derived indices, physical variables). They also obtain statistics derived from the maps by municipalities, states, ecophysiographic regions, watersheds, indigenous territories, conservation units, among others.

MapBiomas Mexico's mapping project has now released its first collection of annual maps; this collection will evolve in methodology, analysis period, detail of mapped land cover over time, and an improvement in map quality.

Collection 1: Mapping of land cover and land use between the years **1985 and 2025.** The methodology used was *machine learning (Random Forest)* with 16 layers of information (original Landsat bands, fractional and texture information derived from them, and indices). Landsat surface reflectance images were used, which present the most up-to-date version of the reprocessed Landsat files. The generated maps have a total of **16 classes**.

The MapBiomas collections aim to contribute to the development of a fast, reliable, collaborative, and low-cost method for processing large-scale datasets and generating historical time series of annual land cover and land use maps. All data, classification maps, statistical codes, and other analyses are freely accessible through the MapBiomas Mexico platform (http://mexico.mapbiomas.org)

The products in Collection 1.0 are as follows:

- Annual LULC maps for the entire continental territory of Mexico (1985–2025) at 30 m resolution.
    - Rasters of annual transitions between classes and years chosen by the user;
    - Preprocessed mosaics generated from Landsat file collections (Landsat 4, Landsat 5, Landsat 7, Landsat 8 and Landsat 9), stored as assets in Google Earth Engine.

- Image processing infrastructure and algorithms (scripts in Google Earth Engine and source code);
- LULC annual and transitional statistics with various units of analysis;
- Landsat mosaic quality assessment. Each scene may have a proportion of clouds and other interference. Thus, each pixel for a given year is rated according to the number of available observations, ranging from 0 to 40 observations per year. The Landsat mosaic quality assessment is available on the MapBiomas website.
- National infography on land cover and land use;
- Document of key findings in the analysis of the results;
- Technical documents for understanding the processes by topic.

1.3. Study Area

The scope of work is the entire territorial extension of Mexico, excluding only insular areas far from the coast. This encompasses an approximate area of 1,962,292 km^2. Within this area, five geographic regions defined by physiographic provinces were delimited: Northwest, North-Center, Northeast, Center-South and Southeast (Figure 1).

Figure 1. Regions and participating institutions

These five regions differ in several aspects, such as climate, vegetation, topography, hydrology, and the presence of human settlements.
For this purpose, technical criteria regarding ecosystem distribution were used, based on national reference maps and the experience of previous collections. The land cover types used were:

- Land use and land cover map (INEGI, 2021)
- Map of Ecoregions of Mexico (INEGI/CONABIO/INE, 2008)
- Physiographic Map of Mexico (INEGI, 2001)
- SRTM 30 m Digital Elevation Model (Farr et al., 2007)

- Mangrove Monitoring System (CONABIO, 2020)

Table 1. Classification regions and sub-regions of MapBiomas Mexico, Collection 1.

Region	Area (percentage)	Sub-regions (Level I Ecoregions)	Main characteristics
1. Northwest	27.92 Mha (14.2%)	North American Deserts (10), Tropical Dry Forests (14), Mediterranean California (11), Tropical Moist Forests (15), Southern Semi-Arid Highlands (12)	Arid and semi-arid zones of the northwest; deserts, shrublands, Mediterranean valleys
2. North-Center	63.55 Mha (32.4%)	North American Deserts (10), Temperate Sierras (13), Southern Semi-Arid Highlands (12), Tropical Dry Forests (14), Tropical Moist Forests (15)	High plains and ranges of the north-center; xeric shrublands and mid-elevation temperate forests
3. Northeast	37.29 Mha (19.0%)	North American Deserts (10), Great Plains (9), Temperate Sierras (13), Tropical Dry Forests (14), Tropical Moist Forests (15), Southern Semi-Arid Highlands (12)	North American Deserts and the Great Plains of North America; arid shrublands and grasslands transitioning to pine-oak forests in the Sierra Madre Oriental
4. Center-South	37.86 Mha (19.3%)	Temperate Sierras (13), Tropical Dry Forests (14), Southern Semi-Arid Highlands (12), Tropical Moist Forests (15)	Trans-Mexican Volcanic Belt, Southern Sierras; transition between temperate forests, dry and moist tropical forests
5. Southeast	29.61 Mha (15.1%)	Tropical Moist Forests (15), Tropical Dry Forests (14), Temperate Sierras (13)	Yucatan Peninsula, Chiapas; moist and dry tropical forests, mangroves

1.4. Applications

The annual LULC time series produced by MapBiomas Mexico has broad scientific, institutional, and policy applications. Key uses include:

- **Deforestation and land cover change monitoring**: Annual maps enable tracking of forest loss, degradation, and land conversion across Mexico's diverse ecosystems, supporting the work of CONAFOR, SEMARNAT, and state-level environmental agencies.

- **Greenhouse gas emissions accounting**: The time series provides a consistent spatial basis for estimating emissions and removals from land use change, directly supporting Mexico's National GHG Inventory and its Nationally Determined Contributions (NDCs) under the Paris Agreement.

- **REDD+ implementation and carbon markets**: The maps support the design, monitoring, and verification of REDD+ programs and forest carbon projects, particularly in areas of high deforestation pressure.

- **Biodiversity and ecosystem monitoring**: As one of the world's megadiverse countries — hosting approximately 12% of global biodiversity — Mexico requires robust land cover data to assess habitat loss, fragmentation, and ecosystem integrity. MapBiomas Mexico's products are directly applicable to monitoring commitments under the Kunming-Montreal Global Biodiversity Framework.

- **Indigenous and community territory management**: A large share of Mexico's forests are under indigenous or community stewardship. The annual maps provide these communities and their supporting institutions with tools to document and respond to land use pressures within their territories.

- **Water resources management**: Changes in surface water extent, watershed coverage, and vegetation in riparian zones can be tracked annually, informing integrated water resource management.

- **Urban and agricultural expansion**: The time series documents the spatial dynamics of urban growth and agricultural frontier expansion, supporting regional planning by municipal and state governments.

- **Protected area management**: Annual maps allow CONANP and conservation organizations to monitor land cover change inside and adjacent to Mexico's natural protected areas and priority conservation regions identified by CONABIO.

- **Scientific research**: The 41-year time series supports academic research on land use dynamics, secondary vegetation recovery, ecosystem resilience, and human-environment interactions across Mexico's physiographic regions.

Added Value of MapBiomas Mexico

MapBiomas Mexico complements existing national and global land cover mapping efforts, and offers several distinct advantages:

1. **Temporal depth and consistency**: Maps cover 41 years (1985–2025) with annual frequency, using a single consistent methodology and legend throughout, enabling robust trend analysis and long-term comparisons that are not possible with other datasets.

2. **National coverage at high resolution**: The entire continental territory is mapped at 30-meter spatial resolution, allowing analysis at local, regional, and national scales.

3. **Regional comparability**: Shared methodology, legend, and spatial resolution with other MapBiomas country nodes in Latin America allow cross-border and regional comparisons.

4. **Local expertise**: Maps are produced by specialists with deep knowledge of Mexico's ecosystems, ensuring ecological validity and relevance for national applications.

5. **Open access**: All methods, algorithms, and products are publicly available through the MapBiomas Mexico platform and GitHub repository, enabling transparent and reproducible use by any institution or researcher.

1.5. Products of the Collection

The products of MapBiomas Mexico Collection 1.0 include: (1) annual land cover and land use maps for the continental territory of Mexico (1985–2025) at 30 m resolution; (2) preprocessed Landsat image mosaics (L4, L5, L7, L8, L9) stored in Google Earth Engine; (3) image processing algorithms and scripts (GEE/GitHub); (4) annual and transitional LULC area statistics by country, state, municipality, river basin, indigenous territories, and protected areas; (5) Landsat mosaic quality assessment layers; (6) national land cover infographics and key findings documents; (7) cross-cutting theme technical documents (appendices).

2. Basic Information and Background

2.1. Global MapBiomas Network

MapBiomas is a global, multi-institutional network of NGOs, universities, and technology companies founded in Brazil in 2015. Its mission is to reveal territorial transformations through science — with precision, agility, and quality — and to make knowledge about land cover and land use accessible, in order to promote the conservation and sustainable management of natural resources as a means of combating climate change (Souza et al., 2020).

The network was born out of a practical challenge: how to produce annual land use and land cover maps for all of Brazil faster, cheaper, and more frequently than existing methods allowed. The answer was a technical collaboration with Google using the Google Earth Engine (GEE) platform as the computational backbone, combined with a distributed network of regional specialists contributing local ecological knowledge. This model proved scalable, and what began as a Brazilian initiative has grown into a global effort. By 2025 — the network's tenth anniversary — MapBiomas was active in 14 countries across South America and Indonesia, with a stated goal of mapping 70% of tropical forests in 20 countries by 2030.

Current country and regional nodes include Brazil, the Amazon basin, the Gran Chaco, the Trinational Atlantic Forest, the Trinational Pampas, Argentina, Bolivia, Chile, Colombia, Ecuador, Indonesia, Mexico, Paraguay, Peru, Uruguay, and Venezuela, each operated by local institutions with expertise in their national ecosystems.

The core product of MapBiomas is the annual Land Use and Land Cover (LULC) map series, which extends from 1985 to the present at 30-meter spatial resolution. Beyond LULC mapping, the global network has developed specialized thematic modules including deforestation alerts (MapBiomas Alert), surface water dynamics (MapBiomas Water), fire occurrence (MapBiomas Fire), soil mapping (MapBiomas Soil), forest degradation, secondary vegetation recovery, and climate risk assessment. Not all modules are active in every country node; their development follows each initiative's priorities and capacities.

All data, methods, and source code produced by MapBiomas are freely available to the public through each initiative's online platform (https://mapbiomas.org). The network operates according to the following shared principles across all its initiatives:

- **Open and free data**: All maps, statistics, and methods are publicly accessible without restriction.
- **Methodological transparency**: Algorithms, scripts, and validation procedures are fully documented and publicly available on GitHub.
- **Local expertise**: Each node is led by institutions with deep knowledge of national ecosystems, ensuring ecological validity and relevance for local applications.
- **Cloud-based, automated processing**: All classification is performed on Google Earth Engine, enabling reproducible, large-scale, and cost-effective map production.
- **Continuous improvement**: Products are released in versioned collections, allowing systematic methodological advances and quality improvements over time.
- **Scientific and technical rigor**: Accuracy assessments will be carried out to guide improvement of the next collection, and uncertainty reporting are integral to each collection release.
- **Capacity building**: The network actively trains new specialists and promotes the adoption of open remote sensing methods across institutions and countries.

2.2. Institutional Context

MapBiomas Mexico is carried out thanks to a national network of 22 institutions spanning government agencies, academic institutions, and civil society organizations. Partner institutions include CentroGeo, CONABIO, CONAFOR, INEGI, ECOSUR, INIFAP/CICY, Universidad Nacional Autónoma de México (Centro de Investigaciones en Geografía Ambiental, Instituto de Biología, Instituto de Ecología, Instituto de Geografía), Universidad

Autónoma de Ciudad Juárez, Universidad Autónoma de Baja California, Universidad Autónoma de Chihuahua, Universidad Autónoma de Nayarit, Universidad Autónoma de Sinaloa, Universidad Autónoma del Estado de México, Universidad de Guanajuato, Iniciativa Climática de México (ICM), The Nature Conservancy (TNC), Reforestamos México, WRI México, FAO/CONANP, among others

Formally established in 2025, the Mexican network currently comprises 54 members from these 22 institutions. MapBiomas Mexico operates as the national node of the MapBiomas global initiative — a collaborative, multi-institutional network founded in Brazil in 2015 that brings together non-governmental organizations, research institutions, universities, and technology companies to produce open-access, annual, fine-scale land cover and land use time series. Within this network, MapBiomas Mexico shares the common methodology, legend, and cloud-based processing platform (Google Earth Engine) while adapting them to the country's specific ecological and territorial conditions. Its mandate is to articulate technical capacities and territorial knowledge to generate transparent, freely available, and regularly updated land cover and land use information useful for public, social, community, and private decision-makers, in support of sustainable land management and natural resource conservation in a context of climate change.

MapBiomas México received financial support from Avina, the **Iniciativa Climática de México** (ICM), and the **Secretaría de Ciencia, Humanidades, Tecnología e Innovación** (SECIHTI, Project **CBF-2025-G-828**).

2.3. Historical Background

Mexico has a long tradition in the production of land cover and land use maps, dating back to the 1970s and 1980s, when cartography was produced primarily through visual interpretation of aerial photographs and intensive fieldwork. The vegetation map by Rzedowski (1978) was a pioneering reference, followed by the INEGI Land Use and Vegetation Series, whose first edition was based on aerial photographs at a scale of 1:250,000, with a classification system distinguishing more than 50 categories inspired by the work of prominent Mexican botanists such as Miranda and Hernández-X. This cartographic effort evolved over time, incorporating Landsat satellite imagery as the primary input for updating successive series (Series II, III, IV, V, VI, and VII), consistently maintaining visual interpretation as the central method and a working scale of 1:250,000.

In the early 1990s, the Instituto de Geografía at UNAM produced the Inventario Forestal de Gran Visión (1991) using NOAA AVHRR imagery and NDVI indices, as well as contributing to the North American NALC project, which unsupervisedly classified a Landsat image triplet from the 1970s, 1980s, and 1990s. The Inventario Forestal Nacional Periódico of 1994, developed by the Instituto de Geografía in collaboration with SEMARNAP, adopted a classification system compatible with FAO standards and produced digital cartography in vector format. The Inventario Forestal Nacional 2000 represented a significant qualitative leap: in addition to updating the cartography through visual interpretation of Landsat ETM+ imagery at 1:250,000, it introduced for the first time a statistically robust accuracy assessment using digital aerial photographs and the calculation of confusion matrices, and provided the first map-based estimate of land cover and land use change — a methodological milestone subsequently adopted by both INEGI and CONAFOR.

The current official monitoring system, jointly operated by CONAFOR and INEGI, updates the cartography every three to four years through visual interpretation at 1:250,000, coupled with an intensive field sampling campaign of approximately 24,000 sampling clusters. Its main limitation remains the minimum mapping unit of 25 to 50 hectares, which prevents the detection of fine-scale changes.

In parallel, CONABIO developed the MadMex pilot system, a cloud-based automated mapping platform initially using RapidEye imagery and later Landsat, with change detection capabilities but no visual verification component, and a reported overall accuracy of 76%. This system was subsequently taken up by CONAFOR through the Sistema de Alerta y Monitoreo

Forestal (SAMOF), focused on automated monitoring of forest change. These latest developments mark the transition toward an era of large-scale satellite data processing on platforms such as Google Earth Engine and the use of machine learning algorithms such as Random Forest, which have progressively displaced traditional supervised classification methods.

2.4. Formation of the National MapBiomas Mexico Network

MapBiomas Mexico, launched in July 2025 as part of the MapBiomas global initiative, builds on this cartographic tradition while also benefiting from the accumulated experience of South American countries. The MapBiomas global network originated in Brazil in 2015 as a collaborative, multi-institutional annual mapping effort — bringing together non-governmental organizations, research institutions, universities, and technology companies — with the goal of producing fine-scale, open-access time series of land cover and land use to support decision-making, sustainable land management, and natural resource conservation in a context of climate change. The Mexican network was formally established in 2025 and currently comprises 54 members from 22 institutions spanning government agencies, academic institutions, and civil society organizations. Its distinctive value lies in articulating technical capacities and territorial knowledge to adapt MapBiomas methodologies to the specific needs of the country and to produce information that is useful for public, social, community, and private decision-makers. The network's first major milestone is the production and publication of a complete land cover and land use collection spanning the period 1985–2025, to be launched on August 4, 2026, in Mexico City. This collection is unprecedented in Mexico: no equivalent initiative has previously combined historical map series, open and freely available data, annual updates, transparent methodologies, and a multisectoral collaborative process. Beyond this first collection, MapBiomas Mexico aims to generate annual updates, develop thematic maps on fire scars and surface water dynamics, and promote the use of its products through training workshops and outreach activities targeting decision-makers and strategic actors across the country.

2.5. Google Earth Engine and MapBiomas Mexico Applications

All image processing, classification, and post-classification steps are performed on the Google Earth Engine (GEE) cloud computing platform (Gorelick et al., 2017; Google, 2019a). GEE provides access to the Landsat data catalog (Google, 2019b), standard algorithms, and scalable cloud processing infrastructure. MapBiomas Mexico scripts are publicly available at: https://github.com/mapbiomas/mexico-all-initiatives

The mapping data for Collection 1 were obtained from satellite images for the period 1985 to 2025, acquired by the following Landsat sensors: the Thematic Mapper (TM), aboard Landsat 4 (L4, used to fill information gaps at the beginning of the series, mainly in 1988) and Landsat 5 (L5, for the years 1985–2012); the Enhanced Thematic Mapper Plus (ETM+), aboard Landsat 7 (L7, for the years 2000–2021); the Operational Land Imager and Thermal Infrared Sensor (OLI/TIRS), aboard Landsat 8 (L8, from 2013 onwards); and the Operational Land Imager 2 and Thermal Infrared Sensor 2 (OLI-2/TIRS-2), aboard Landsat 9 (L9, from 2021 onwards).

The surface reflectance images are part of the Landsat data catalog Level 2 Collection 2 Tier 1. These images underwent radiometric calibration, orthorectification based on ground control points, and digital elevation modeling to ensure pixel-level recording and atmospheric correction. The catalog of Landsat images with a spatial resolution of 30 meters was accessed through the Google Earth Engine platform, provided by NASA and the United States Geological Survey (USGS).

2.6. Other Mapping Initiatives

In recent years, various tools have been developed for globally mapping land cover and land use with increasingly robust approaches. They all share a common interest in contributing to one of today's most pressing issues: understanding the current state of land cover types and monitoring changes in them. These initiatives have contributed to detecting deforestation, monitoring terrestrial and aquatic ecosystems, and strengthening the detection of forest degradation, conservation efforts, hotspots, and more. Below, we list the most relevant initiatives.

---

[1] https://developers.google.com/earth-engine/datasets/catalog/landsat/
[2] https://developers.google.com/earth-engine/
[3] code.earthengine.google.com

2.6.1. Global Sources

- GLC 2000 - Global Land Cover mapping for the year 2000, the project was an international partnership of some 30 research groups coordinated by the European Commission's Joint Research Centre, with the aim of producing a global land cover database for the year 2000. The database contains land cover maps with detailed and regionally relevant map legends and a global product that combines all regional classes into a coherent legend.
- Intact Forest Landscapes (IFL) is a global spatial database at a scale of 1:1,000,000, showing the extent of intact forest landscapes (IFL) for the years 2000, 2013, and 2016. The first global IFL map, for the year 2000, was prepared in 2005–2006 under the leadership of Greenpeace, with contributions from the Biodiversity Conservation Center, the International Socio-Ecological Union, and Transparent World (Russia), Luonto Liitto (Finnish Nature League), Forest Watch Indonesia, and Global Forest Watch, a network initiated by the World Resources Institute. The 2013 version was subsequently generated, and the map was finally updated to 2016 with support from the University of Maryland, the Wildlife Conservation Society, Greenpeace, and Transparent World.
- Global Forest Watch (GFW) – A collaboration between the GLAD (Global Land Analysis & Discovery) laboratory at the University of Maryland, Google, USGS, and NASA, measures areas of tree cover loss across the Earth (excluding Antarctica and other Arctic islands) at a 30 × 30 meter resolution. Their project focuses on developing global tree cover change data products based on Landsat satellite imagery, available on the Global Forest Watch 3.0 web platform. It includes annual forest cover change (gains and losses) from 2000 to 2020.
- GlobeLand30, is an initiative of the National Geomatics Center of China. This collection comprises spatial datasets compiled at a resolution of 30 meters. It considers ten land cover types, including forests, artificial surfaces, and wetlands, for the years 2000 and 2010. These datasets were extracted from more than 20,000 satellite images from Landsat and the Chinese HJ-1 satellite.
- ESA CCI Land Cover, a joint project of the European Space Agency (ESA) and the Climate Change Initiative (CCI), provides annual global land cover maps, which describe the Earth's surface in 22 classes. The series of annual global land cover maps spans the period from 1992 to 2018.
- The CORINE Land Cover (CLC) inventory was initiated in 1985 (reference year 1990). Updates were carried out in 2000, 2006, 2012, and 2018. It consists of a land cover inventory in 44 classes. CLC uses a Minimum Mapping Unit (MMU) of 25 hectares (ha) for area features and a minimum width of 100 m for linear features. Time series data are supplemented with change layers, which highlight land cover changes with an MMU of 5 ha.
- ESRI 2020 Global Land Use/Land Cover from Sentinel-2: This layer displays a global land use/land cover (LULC) map for 2020. The map is derived from ESA Sentinel-2 imagery with a 10-meter resolution and contains 10 classes. This map was produced by a deep learning model trained on over 5 billion hand-labeled Sentinel-2 pixels sampled from more than 20,000 sites distributed across the world's major biomes.
- ESA WorldCover 2020 and 2021: This is a reference global land cover product with a spatial resolution of 10 m, generated from Sentinel-2 and Sentinel-1 imagery with 10 land cover classes and an overall accuracy of 75%. The legend includes 11 generic classes that adequately describe the land surface: "Tree cover", "Shrubland",

"Grassland", "Cropland", "Built-in", "Bare/sparse vegetation", "Snow and ice", "Permanent water bodies", "Herbaceous wetland", "Mangroves", and "Mosses and lichens".

- Dynamic World is a near real-time, 10-meter resolution global land cover dataset generated from Sentinel-2 imagery using deep learning. It is freely available and openly licensed. The legend presents the probabilities per pixel for nine land cover classes: Water, Forest, Shrub and Bush, Grassland, Floodplain Vegetation, Cropland, Buildings, Bare Soil, Snow, and Ice. This data is the result of a partnership between Google and the World Resources Institute to produce a dynamic dataset of the physical material on Earth's surface.

2.6.2. National Sources

- Land use and vegetation (7th series) is the official land use land cover map made by the INEGI at 1:250,000. The dataset was made using field data and photointerpretation. The legend presents 183 land uses or vegetation types, which include different types of agriculture (e.g., temporal, semipermanent, permanent), secondary vegetation in different stages (e.g., herbaceous, shrubland and woodland) and natural vegetation with different phenological and stand attributes.
- MADMex (Monitoring Activity Data for the Mexican REDD+ Program) is a nationally automated approach to construct land use land cover maps in support of REDD+ program. This initiative is led by CONABIO and makes use of INFyS and INEGI data, as well as Landsat, Sentinel-2 and RapidEye images.
- SAMOF is a satellite monitoring system led by CONAFOR to monitor forest cover, deforestation, forest degradation, reforestation, afforestation, among other land use changes. This initiative uses mainly Landsat and very-high-resolution imagery for validation and sample interpretation.

3. Methodology

The processing chain adopted for the generation of Collection 1 of MapBiomas Mexico is summarized in Figure 2 and is organized into nine stages: generation of annual mosaics (3.1), identification of stable pixels (3.2), generation of samples per year (3.3), classification (3.4), integration (3.5), post-classification (3.6), statistics (3.7), transition maps (3.8), and qualitative quality assessment (3.9).

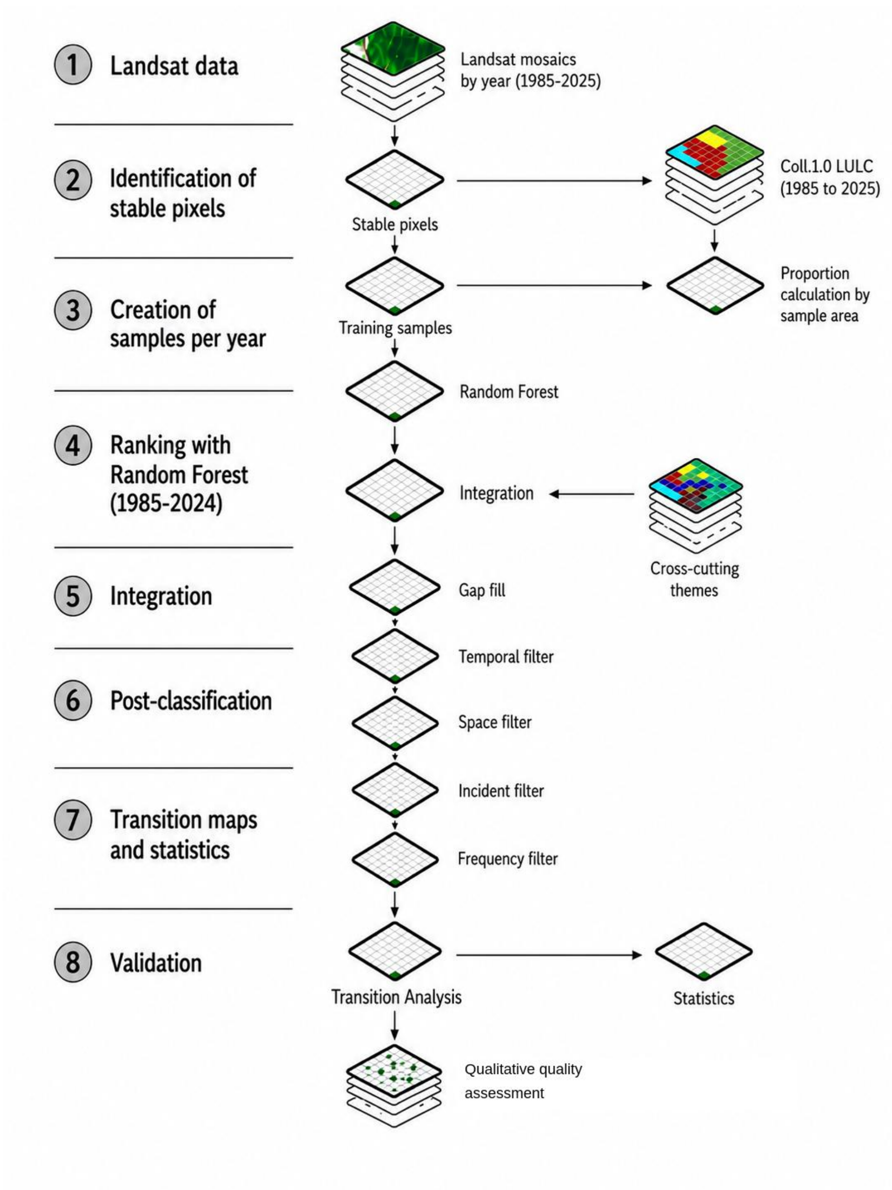

Figure 2. Methodological flowchart of Collection 1.0

3.1. Generation of Annual Mosaics

3.1.1. Division of the Analysis Space into Regions

As mentioned in section 1.3, the territory was divided into five main regions based on the Physiographic Provinces of Mexico (INEGI, 2001). The five analysis regions used in MapBiomas Mexico were defined by grouping the physiographic provinces of INEGI according to their geographic position within the country. Each region aggregates several contiguous provinces, as follows:

- Northwest comprises the Baja California Peninsula, the Sonoran Plain, and the Pacific Coastal Plain.
- North-Center is made up of the Sierra Madre Occidental, the Central Mesa (Mesa del

Centro), and the Northern Sierras and Plains.
- Northeast groups the Sierra Madre Oriental, the Northern Gulf Coastal Plain, and the Great Plains of North America.
- Center-South brings together the Trans-Mexican Volcanic Belt (Neovolcanic Axis) and the Sierra Madre del Sur.
- Southeast encompasses the Southern Gulf Coastal Plain, the Yucatán Peninsula, the Sierras of Chiapas and Guatemala, and the Central American Cordillera.

In each region, a local team carried out the training of the classifier. Each region was further subdivided into sub-regions defined by Mexico's Level I Ecoregions (INEGI/CONABIO/INE, 2008). This division facilitates classifier parameterization and the collection of training samples adapted to local ecological conditions. Table 2 shows the ecoregions present in each region. As we can observe in Figure 3, Mexico is covered by 184 Landsat images.

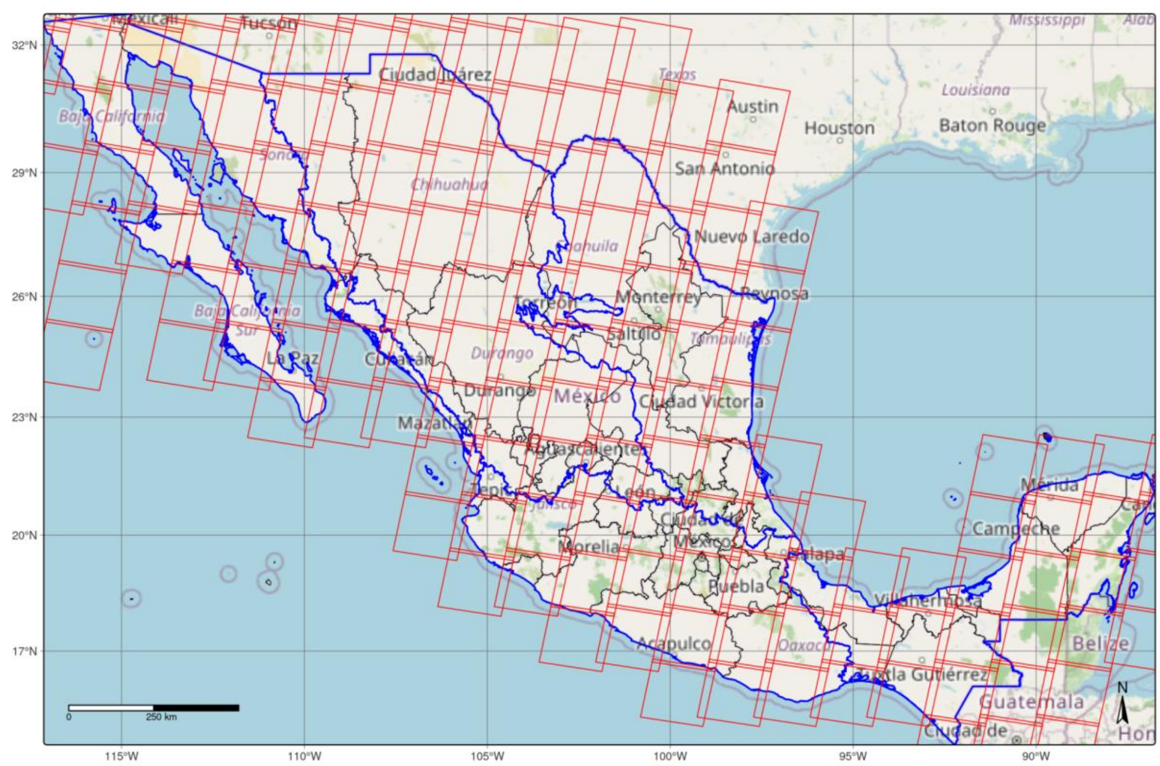

Figure 3. Landsat image grid (red), regions (blue) and Mexican states boundaries (black).

Table 2. Summary of the number of MapBiomas Mexico maps by ecoregion, with the area (rounded km²) of each ecoregion within the region.

Region	Ecoregions that it encompasses (area km²)	Cards / Mosaic Region	Total tiles (× 41 years)
Northwest	1. North American Deserts (209,160 km²) 2. Tropical Dry Forests (38,980 km²) 3. Mediterranean California (25,036 km²) 4. Tropical Moist Forests (4,626 km²) 5. Southern Semi-Arid Highlands (1,320 km²) 6. Not classified (29 km²)	41	1,681
North-Center	1. North American Deserts (202,382 km²) 2. Temperate Sierras (178,866 km²) 3. Southern Semi-Arid Highlands (172,016 km²) 4. Tropical Dry Forests (78,790 km²) 5. Tropical Moist Forests (3,450 km²)	49	2,009
Northeast	1. North American Deserts (147,642 km²) 2. Great Plains (107,550 km²) 3. Temperate Sierras (47,702 km²) 4. Tropical Dry Forests (43,464 km²) 5. Tropical Moist Forests (26,012 km²) 6. Southern Semi-Arid Highlands (542 km²)	32	1,312
Center-South	1. Temperate Sierras (179,568 km²) 2. Tropical Dry Forests (129,043 km²) 3. Southern Semi-Arid Highlands (55,963 km²) 4. Tropical Moist Forests (14,023 km²) 5. Not classified (3 km²)	30	1,230
Southeast	1. Tropical Moist Forests (233,554 km²) 2. Tropical Dry Forests (33,416 km²) 3. Temperate Sierras (27,788 km²) 4. Not classified (1,367 km²)	32	1,312
Total	**1,962,292**	**184**	**7,544**

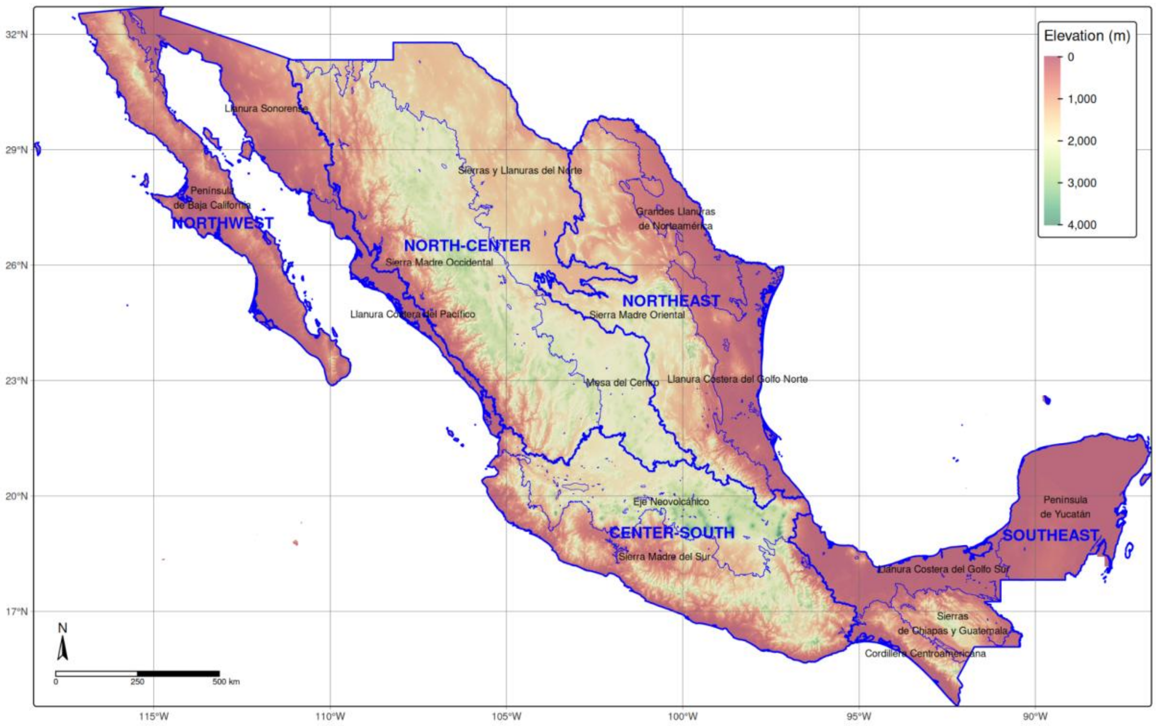

Figure 4. Map of the five regions and the 15 physiographic provinces for Mexico.

3.1.2. Parameterization of Annual Mosaics

An annual mosaic is the aggregation of pixels from several Landsat images from which a representative mosaic for a year is generated, constructed from the following parameters:

- ID: Unique identifier of the letter-region unit
- Year: Year of the series (1985 to 2025) to which the mosaic corresponds.
- Letter: Letter identifier code
- Start Date/ End Date: Period of the year (start and end date) for the selection of images from the Google Earth Engine Landsat image data catalog.
- Sensor: The satellite and its respective sensor: Landsat 4 TM, Landsat 5 TM, Landsat 7 ETM+, Landsat 8 OLI, or a combination of Landsat 5 and Landsat 7, Landsat 9 OLI-2 TIRS-2.
- Cloudiness: The maximum percentage of cloud cover accepted for each Landsat image that will be used to construct the image mosaic. This data comes from the Landsat image metadata.
- Probability of clouds:
- shadowSum: This is a TDOM parameter for cloud shadow detection. A lower number masks fewer pixels with clouds.
- cloudThresh: This is a CloudScore parameter for cloud detection. Lower numbers increase masking, while higher numbers decrease it. Decreased masking can be useful in areas of bare ground where high reflectance leads to false cloud detection.
- Blacklist: Images that, due to their quality, are excluded from the mosaic construction. Some quality issues are related to snow pixels in non-glacial regions and clouds that have not been masked.

The parameters for constructing annual mosaics are defined by the interpreter and represent the criteria for selecting images available in the Landsat data catalog from which the annual mosaic is constructed. The images selected for each year were *reduced* to an individual image, or annual mosaic, using operators called *reducers* existing in Google Earth Engine, as illustrated in Figure 5.

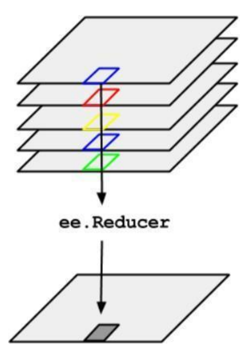

Figure 5. Schematic of the application of a reducer to a set of images (Google, 2020)[5].

When parameterizing the mosaics, it was considered that higher accuracy values could be achieved by using satellite image mosaics with minimal noise. Therefore, each mosaic was designed to have the least possible cloud cover and interference, as well as the largest possible coverage of available Landsat data within the defined period. In exceptional cases where images were unavailable for the selected period, the image search period was extended.

Clouds and cloud shadows are pre-masked so that only pixels free of clouds and cloud shadows are selected from the available images. The cloud and cloud shadow masking methods used were Cfmask (Zhu & Woodcock, 2012) and CloudScore. The annual mosaics were built using all the images of each year (i.e., Jan 1 to Dec 31).

[5]Taken from: https://developers.google.com/earth-engine/guides/reducers_image_collection

3.1.3. Classification Variables (Feature Space)

Variables were calculated (*feature space*) from the annual mosaic representing the inputs of the classification process. The Landsat bands, along with the classification variables, are consolidated into raster files composed of 16 bands in total.[6] which include: spectral Landsat bands, spectral indices, fractional and texture information derived from them, and indices of spectral fractions.

Additionally, 1 static variable was used: slope, which helped to classify classes that are spectrally very similar but are differentiated by these topographic aspects.

The calculation was applied to the images available each year. Statistical reducers are used to generate the values for each pixel. The process is represented in Figure 6. These reducers are:

- Median: Median[7]of all available values in the annual mosaic for that location (pixel).
- Median dry season: Calculation of the statistical median applied to the pixels of the 25th percentile (with the lowest values) of NDVI (proxy of dry season).
- Median wet season: Calculation of statistical median applied to the pixels of the 75th percentile (with the highest values) of NDVI (proxy of rainy season).
- Standard deviation: Standard deviation of the values of all available pixels in the annual mosaic for a given location.

---

[6] Available for download on the MapBiomas Mexico platform.
[7] The median is the value that separates the upper half from the lower half of a data sample or population. Documentation of the tool in Google Earth Engine.

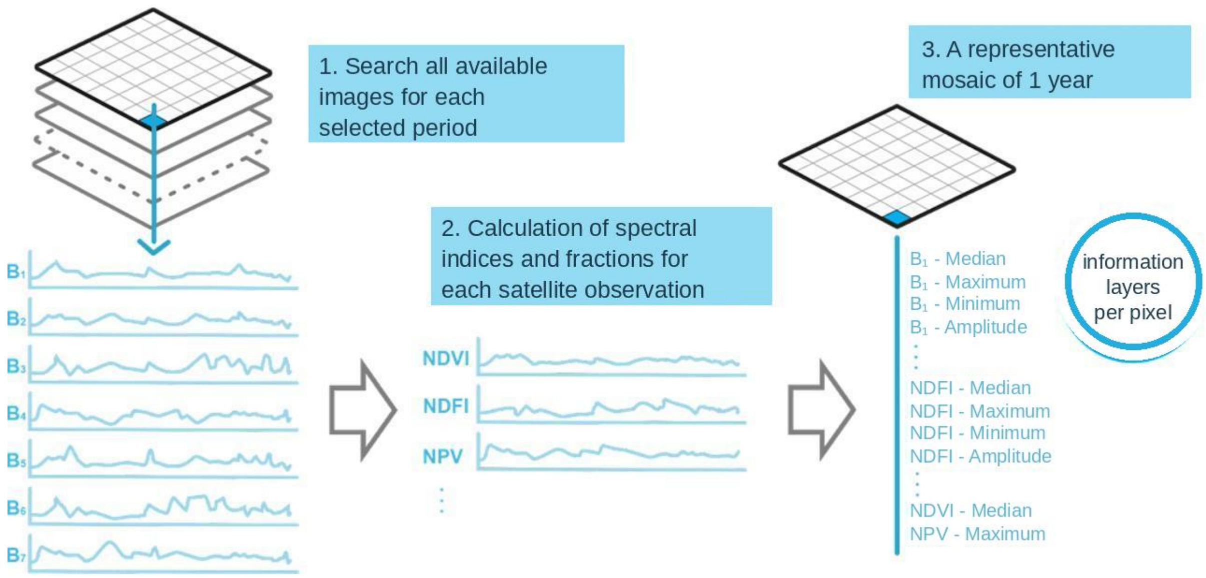

Figure 6. Process of calculating bands that make up the annual Landsat image mosaics.

Table 3 shows the complete list of bands for the final mosaics or feature *space.* Each band represents a training variable of the classifier.

Table 3. Description of bands and variables used for Collection 1.0.

Type	Name	Formula	Description	Reducer[8]				
				Median	Median_dry	Median_wet	amp	stdDev
Band	blue	B1 (L5 y L7); B2 (L8 y L9)	Blue visible spectrum	X				
	green	B2 (L5 y L7); B3 (L8 y L9)	Green visible spectrum	X				
	red	B3 (L5 y L7); B4 (L8 y L9)	Red visible spectrum	X				
	nir	B4 (L5 y L7); B5 (L8 y L9)	Near infrared	X				
	swir1	B5 (L5 y L7); B6 (L8 y L9)	Shortwave infrared 1	X				
	swir2	B7 (L5 - L9)	Shortwave infrared 2	X				
Indices	NDVI	(nir - red)/(nir + red)	Normalized Difference Vegetation Index		X	X		X

	EVI	2.5 * (nir - red)/(nir + 6 * red - 7.5 * blue + 1)	Enhanced vegetation index (EVI)		X	X		X
	EVI2	2.5 * (nir - red)/(nir + 2.4 * red + 1)	Index (EVI) that only uses NIR and Red, ignoring the blue band.		X	X		X
	NDWI_GAO	(nir - swir)/(nir + swir)	Normalized difference water index (GAO)		X	X		
Static and/or topographic variables	slope			X				

3.2. Identification of Stable Pixels

Stable pixels are defined as pixels that maintain the same land cover class consistently throughout the time series. A preliminary classification is generated for the full 41-year period (1985–2025), and pixels with stable class assignments across all years are identified. These stable pixels serve as the primary training sample pool for the final classification (Section 3.4). This approach reduces the dependence on manual sample collection and leverages the temporal consistency of the Landsat time series.

3.3. Generation of Samples per Year

Training samples are generated for each year of the time series based on the stable pixel pool identified in Section 3.2. The sample strategy balances class representation across classification regions and sub-regions. Details of the sample collection methodology and the classification regions used in Mexico are presented in the subsections below.

3.4. Classification

3.4.1. Legend

The MapBiomas Mexico Collection 1 classification includes 16 LULC classes, defined based on the ecological characteristics of the Mexican territory and compatible with the general legend of the MapBiomas Network. Table 4 presents the classes, their numeric identifiers, and the colors used for cartographic representation.

Table 4. General legend of land cover and land use in Collection 1

ID	Class	Group	Color
88	Temperate forest	1. Forest	
89	Tropical dry forest	1. Forest	
3	Tropical humid forest	1. Forest	
5	Mangrove	1. Forest	
66	Shrublands	2. Natural herbaceous and shrub vegetation	
45	Savanna and natural grassland	2. Natural herbaceous and shrub vegetation	
11	Flooded	2. Natural herbaceous and shrub	

ID	Class	Group	Color
		vegetation	
15	Cultivated grassland	3. Farming	
36	Perennial crop	3. Farming	
19	Annual crop	3. Farming	
9	Forest plantation	3. Farming	
21	Mosaic of land uses	3. Farming	
24	Urban and Built-up Area	4. Non vegetated area	
25	Other non-vegetated areas	4. Non vegetated area	
33	River, lake or ocean	5. Water body	
34	Ice and permanent snow	5. Water body	
27	Not observed	6. Not observed	

As part of the process of describing the land cover classes, a table was created to correlate the MapBiomas legend classes with the official Mexican land use land cover classes (INEGI, 2021). The following descriptive report and interpretation criteria (Table 5) for each class in the MapBiomas Mexico legend, within the context of Mexico's land use land cover, are presented.

Table 5. Descriptive report for the MapBiomas Mexico legend

Class	ID	Hexacode	Color	Class description	INEGI classification
1. Forest	**1**	**#1f8d49**			
1.1 Temperate forest	88	#329c5a		A type of forest typical of regions with temperate or cold climates. It often consists primarily of conifers and oaks. It is defined as an area of at least 1 ha with tree cover of more than 30%. It includes disturbed or secondary forests.	Pine forest, Mixed oak and pine forest, Juniperus forest, Abies forest, Cedar forest, Pseudotsuga forest, Montane cloud forest, Conifer shrubland, and Riparian forest.
1.2 Tropical dry forest	89	#6bd46c		A tree community typical of a tropical climate, with a distinctly seasonal rainy season. Deciduous trees predominate, shedding their leaves during the dry season.	Tropical deciduous forest, tropical semi-deciduous forest, subtropical shrubland, tropical mezquite shrubland.
1.3 Tropical humid forest	3	#1f8d49		A forest community in a tropical climate with high precipitation. Evergreen trees predominate.	Tropical humid forest, Riparian forest.
1.4 Mangrove	5	#04381d		A woody community of trees or shrubs, ranging from 1 to 30 m in height, composed of one or more mangrove species, with virtually no herbaceous plants or climbers. The species that make up this community are evergreen and somewhat succulent. It grows in estuarine environments, along the shores of lagoons, and near the mouths of coastal rivers. These plants have specialized physiological mechanisms that allow them to thrive in oxygen-poor, muddy soils.	Mangrove
2. Natural herbaceous and shrub vegetation	**10**	**#d6bc74**			
2.1 Shrublands	66	#a89358		A plant community dominated by shrubby plants (which branch from the base), generally less than 4 m tall. These are predominant in arid and semiarid regions.	Microphyllous shrubland, Rosette-forming shrubland, Succulent shrubland, Chaparral, Thorny shrubland, Sand desert vegetation, Gypsophilous vegetation, Xerophyllous vegetation.
2.2 Savanna and natural grassland	45	#807a40		A plant community dominated by grasses or natural pastures, with or without scattered trees.	Natural grassland, Halophilous grassland, Hypsophilous grassland, highland grassland, Savanna, Coastal dunes, palm-dominated areas.
2.3 Flooded	11	#519799		Vegetation that is permanently or temporally flooded.	Hydrophilous vegetation, Peten vegetation.
3. Farming	**14**	**#ffefc3**			
3.1 Cultivated grassland	15	#edde8e		Areas with pastures cultivated for livestock, including pastures resulting from frequent disturbances.	Cultivated and induced grassland.
3.2 Perennial crop	36	#d082de		Crops whose production spans more than one year, generally woody plants. Aimed at producing food products.	Orchards and fruit plantations, including, oil palm, avocado, cacao, coffee, mango, lemon, nut, among others.
3.3 Annual crop	19	#C27BA0		Crops grown for food that are harvested annually, generally consisting of non-woody plants.	Rainfed crops, such as maize and bean. Irrigated crops, such as sugar cane and banana.
3.4 Forest plantation	9	#7a5900		Planted trees for timber or resin production.	Forest plantations, such as Teak, Cedar, Rubber, Mahogany, Melina.
3.5 Mosaic of land uses	21	#ffefc3		Mosaic of natural and anthropogenic vegetation.	
4. Non vegetated area	**22**	**#d4271e**			
4.1 Urban and Built-up Area	24	#d4271e		Urban human settlements characterized by a high density of buildings, impervious surfaces, and roads. According to the Urban ATBD, “An urban area is defined as an area that exhibits a combination of built-up coverage, infrastructure, and spatial organization associated with human settlements, and that can be identified through spectral and structural evidence.”	Human settlements.
4.2 Other non-vegetated areas	25	#db4d4f		Areas with no apparent vegetation cover; bare ground.	Non-vegetated areas.
5. Water body	**26**	**#2532e4**			

Class	ID	Hexacode	Color	Class description	INEGI classification
5.1 River, lake or ocean	33	#2532e4		Natural and artificial bodies of water.	Water.
5.2 Ice and permanent snow	34	#93dfe6		Regions with permanent ice and snow.	Non-vegetated areas.
6. Not observed	**27**	**#ffffff**			

3.4.2. Random Forest Classifier

Random forest is a classification method that uses a machine learning algorithm and which reports high accuracy values, even in complex scenarios due to their heterogeneity (Breiman, 2001). The conceptual basis of Random Forest is based on what Tumer and Ghosh (1996) found when demonstrating that the product resulting from the combination of multiple classifiers achieves high accuracy. Random Forest uses training data to build multiple decision trees from which a class is assigned to each pixel. Random Forest has gained importance in recent years due to its robustness against noise and outliers. Thanks to the accuracy of its classifications, it has been widely embraced by the remote sensing community (Belgiu & Drăguţ, 2016). The Random Forest algorithm is part of the classifier package of *machine learning* available on the Google EE platform. The methodology applies a pixel-based classification criterion.

One of the parameters that Random Forest requires is a defined number of trees. It also requires a list of variables (see the "Variables" section) and training data (see the "Spectral Collection" section).

3.4.3. Sample Collection Strategy

The classification method used was random forest and used different sample sources depending on the region to be classified.

A collection of stable samples (same class throughout the time series) was carried out for each of the categories of the legend that was being sought. A preliminary classification for the 41-year period was then generated. Stable samples (for 41 years) were extracted from this data to be used as input for a new classification. The interpreter contributes by excluding and/or including samples to improve the classification.

The random selection of points was balanced in relation to the extent of each class within each classification. The maximum number of samples by year was set to 3000, and the number of samples of each land use land cover was calculated as the proportion of that cover in INEGI (2018) land use land cover data. Optionally, additional samples were manually included and designated supplementary *samples,* using the geometry creation tools directly in GEE. The random forest used to perform the classification used 100 trees, a bag fraction of 0.5 and the square root of the number of predictive variables per split (i.e., 16 predictive variables)

3.4.4. Cross-Cutting Themes

The various classification tests revealed a limitation in differentiating certain classes, so the decision was made to classify them separately. These classes are called Cross-Cutting Themes and are mapped using binary classification algorithms (class of interest and "Not Observed" class).

This strategy was applied, within the context of Mexico, for the following classes:

- Mangrove (ID = 5);
- Urban (ID = 24).

The details of each methodology can be found in the appendices of each cross-cutting theme.

*Cross-cutting classes that have been mapped in other initiatives such as MapBiomas Agua

3.5. Integration

The results of the annual classifications obtained for each of the classification regions constitute sectors of the general base map, which need to be integrated into a single annual regional map with the cross-cutting themes; following rules of prevalence or order of integration that define the prevalence of classes where overlap of different values occurs.

The integration for Mexico was carried out using the probability of the classification to select the LULC class in the final map (in areas where two or more regions overlapped). This probability was generated by the same random forest algorithm used for the classification. This was performed to construct the LULC map for general classes, while the transversal classes were given priority over the other maps. The priority order to ensemble the final classification, from first to last was: Mangrove, Urban, LULC.

3.6. Post-Classification

The preliminary classification result was subjected to the application of a sequence of filters in order to reduce temporal inconsistencies and reduce classification noise smaller than the minimum mapping unit.[12]and fill information gaps. The post-classification process in this collection does not follow a specific order in the application of filters; their sequence was determined by the needs of each classification region, and repeated use of filters was even facilitated. Exceptions were also included within each filter, such as excluding certain years and classes. All these tools were implemented on the Google EE platform using scripts written in JavaScript. Each of these tools is presented in greater detail below. A description of how these filters were adapted by country can be found in the national appendices.

[12] 6 pixels = approximately half a hectare.

3.6.1. Gap Fill

Certain areas of Mexican territory are characterized by atmospheric and climatic conditions that result in an almost permanent presence of clouds throughout the year in various parts of the territory. As a result, the composition of annual mosaics contains pixels without observations or data (*No data*).

The gap fill filter has the ability to reduce these residual gaps by assigning values to pixels without data due to absences of satellite observation (“gaps”).

Pixels with no data ("gaps") in a classification are filled with the nearest temporally available valid value. The Gap Fill filter processes the time series in two passes. In the first pass it scans forward (from the start to the end of the series) and assigns each no-data pixel the value of the immediately preceding year that has valid data. In a second pass it handles any gaps that still remain — typically those at the beginning of the series, for which no preceding value exists — by scanning backward and filling them with the value of the immediately following year (Figure 7). For each pixel whose value is filled by this filter, the change is recorded in a metadata file that stores the year from which the value was taken (its history).

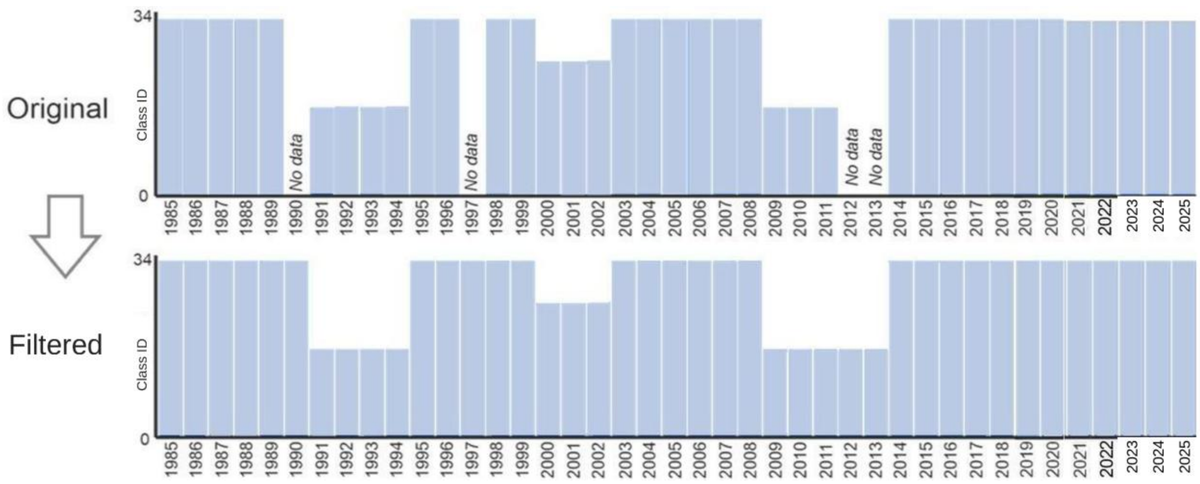

Figure 7. Functionality of the Gap Fill filter

3.6.2. Temporal Filter

The temporal filter examines the value of each classified pixel in relation to its value in consecutive temporal classifications. It does this using a one-way moving window that considers classification sequences of 3 years and identifies disallowed temporal transitions. The temporal filter is applied to every pixel in every year of the collection. The same process was applied to the classification, as well as the probability images. Depending on the year the rule will modify, there are three types of rules:

- General Rules (GR). Applied to pixels in intermediate positions within 3-year sequences. This rule applies only in cases of temporal inconsistency; for example, when sequences of consecutive years have identical values except for the central pixel. In these cases, the filter will modify the value of the central pixel to maintain consistency with the preceding and following pixels. For 3-year sequences, there is only one central position or intermediate year option. This rule modifies the classification values for the years 1986 to 2024.
- First-year rules (PR). Applied only to the first year of the time series. This rule modifies the classification values for the year 1985.
- Last-Year Rules (RU). Applied to the last year of the ranking. This rule modifies the ranking values for the year 2025.

In this way, temporal filters reduce information gaps and temporal inconsistencies or changes that are not possible or permitted (Figure 8). For example, if in three consecutive years a pixel has the following values: Forest Formation > Non-Vegetated Area > Forest Formation, the filter will correct the intervening year. This case is a typical classification error due to the presence of cloud haze in the mosaic of the intervening year.

The decision to choose the size of the time window was made by each country according to the needs and characteristics of its land cover and land use by subregion and/or cross-cutting theme. See more details in the respective country appendices.

Figure 8. Functionality of the time filter

3.6.3. Spatial Filter

The spatial filter is based on the GEE's native "connectedPixelCount" function. This function locates connected pixels (neighbors) that share the same value using a moving window. Only pixels that do not share a connection with a predefined number of identical neighbors are considered isolated pixels. The minimum mapping unit was defined as 0.54 ha (6 pixels). Consequently, at least six connected pixels were required to meet the minimum connection criterion. Thus, the spatial filter smooths out local differences by eliminating isolated or edge pixels smaller than 0.54 ha, increasing the spatial consistency of the ratings (Figure 9). For the probability image produced by the random forest classifier, the substituted values were calculated as the mean of the neighbours.

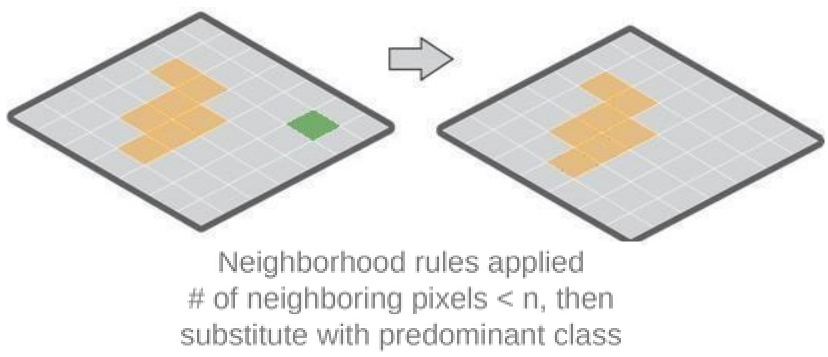

Figure 9. Functionality of the spatial filter

3.6.4. Frequency Filter

This filter takes into account the frequency of occurrence of natural classes throughout the time series. Therefore, classes with occurrences below a percentage defined by the interpreter are replaced by the value of the most frequent class. This mechanism helps reduce the temporal oscillation associated with a natural class, decreasing the frequency of false positives and preserving established trends (Figure 10). This same process was applied to the probability images generated by the random forest classifier.

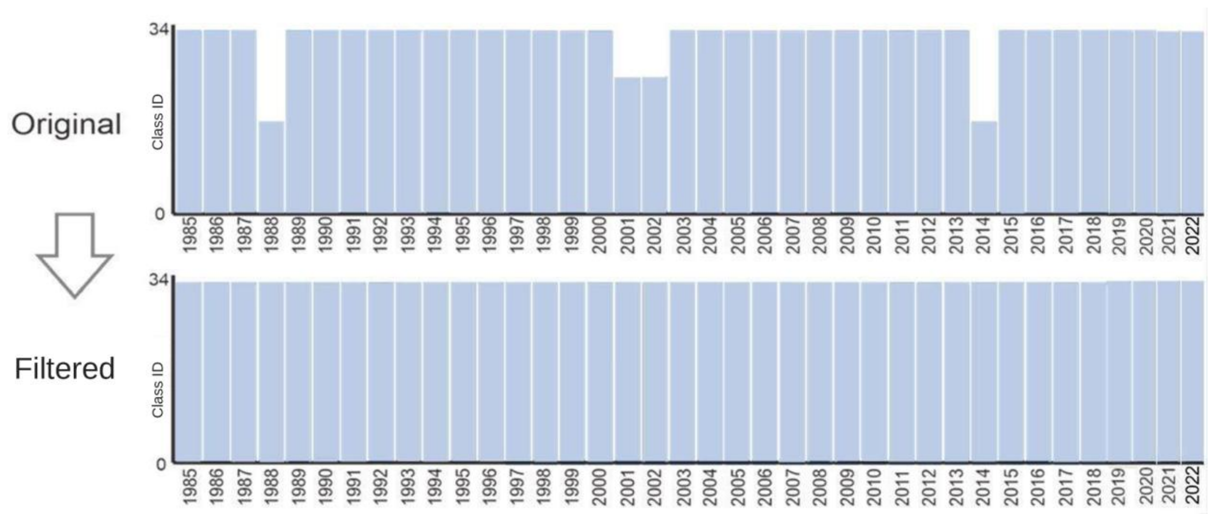

Figure 10. Functionality of the frequency filter

3.6.5. Modal Filter

Finally, it was observed that some pixels showed different natural land cover types over the course of the time series—transitions that were attributed to spectral confusion between categories and to the lower quality of the mosaic images, particularly those from the 1980s. This assumption rests on a well-established ecological principle: the type of natural vegetation at a site is determined by underlying climatic and edaphic variables—temperature, precipitation regime, seasonality, and soil properties—that remain essentially constant over time scales of decades (Holdridge, 1967; Whittaker, 1975; Woodward, 1987; Prentice et al., 1992). Consequently, a real transition between climatically distinct vegetation types would require changes in the thermal or moisture regime that operate over scales of centuries to millennia; and even following disturbance, vegetation tends to recover toward the type determined by the local climate through a slow, directional succession, with lags extending over decades to centuries (Odum, 1969; Svenning & Sandel, 2013). Therefore, within the study period, the natural cover of a pixel can be considered practically invariant, and a "change" between climatically incompatible natural categories is more parsimoniously interpreted as a classification error than as real change. Such apparent changes are better explained by the spectral confusion that can arise between vegetation types with similar spectral responses, accentuated by intra- and inter-annual phenological variability, the presence of mixed pixels, and the heterogeneous quality of the mosaics across the series.

Because this reasoning indicates that an error exists but does not by itself identify in which year or period it occurs, the correction was resolved by means of a decision rule based on the temporal consistency of the series—an approach widely used to remove spurious or "illogical" transitions in remotely sensed land cover time series (Abercrombie & Friedl, 2016; Gong et al., 2017). To this end, the most frequent category was computed by weighting the years so as to assign less weight to those with lower-quality mosaics, and the discrepant natural categories were replaced with that most frequent category. It should be noted that this procedure was applied only to confusion between natural categories and not to transitions of anthropogenic origin (e.g., deforestation or conversion to agriculture), which do constitute real, frequent, and potentially rapid changes, and which were therefore retained.

3.7. Statistics

Based on the integrated annual land cover and land use maps, area statistics in hectares (ha) are calculated for both zonal and annual areas of the mapped classes. The spatial units considered for the calculation of the statistics are:

- State boundaries

- Municipal boundaries
- National Urban System 2018
- Terrestrial Ecoregions Level 1
- Federal Protected Natural Areas
- State and Municipal Protected Natural Areas
- Priority Sites for Biodiversity Conservation in Mexico (2016)
- Aquifer boundaries
- Administrative Hydrological Regions
- Hydrological Basins
- Hydrological Sub-basins
- Internationally Important Wetlands (Ramsar Sites)
- IUCN Key Biodiversity Areas

3.8. Transition Maps

Based on the integrated annual land cover and land use maps, transitions are calculated. These represent the changes between pairs (2) of maps, that is, during a period of time defined by two dates. The results are available on the MapBiomas México platform. Transitions are calculated for different periods, such as:

- (A) consecutive years, annual (for example, from 2001 to 2002, or from 2013 to 2014, etc.)
- (B) five-year periods (for example, 2000-2005)
- (C) ten-year periods (e.g., 2000-2010)
- (D) complete time series (1985 - 2025)
- (E) special periods (for example, 2000-2025)

3.9. Qualitative Quality Assessment

For this first collection, a rigorous statistical accuracy assessment following methods such as those proposed by Olofsson et al. (2014) was not yet carried out. Instead, the quality of the data and products was evaluated in a more qualitative way, focusing on the availability of imagery throughout the mapping period (1985-2025) and its influence on mosaic quality and on the most evident errors and inconsistencies found in the resulting maps. The following case, concerning the Isthmus of Tehuantepec and the Yucatán Peninsula, illustrates this type of assessment for the regions most affected by long-term image scarcity.

During the first period, 30-meter remote sensing relied almost entirely on Landsat 5 TM, a satellite with no onboard solid-state recorder that depended on real-time downlink via TDRS or line-of-sight ground stations, at a time when Mexico did not yet have an operational ERMEXS receiving station (Fig. 11). This situation was compounded by the privatization of the Landsat program under EOSAT, whose commercial logic led to sensors over southern Mexico being switched off in the absence of a prior purchase request from an institution (Fosnight et al., 2016). This explains the lower number of Landsat images available for central and southern Mexico, compared to the northern part (Fig. 12). For the 2004-2011 period, data gaps worsened due to the Scan Line Corrector failure (SLC-off) of Landsat 7 ETM+ in 2003—which caused a recurring loss of about 22% of data in systematic stripes—compounded by the aging and degradation of the Landsat 5 transmission system (Fig. 11). In both periods, this severe scarcity of acquisitions coincided with the region's persistent convective cloud cover and with the dense aerosol and smoke loading from agricultural burning during the short dry-season windows (Fig. 12).

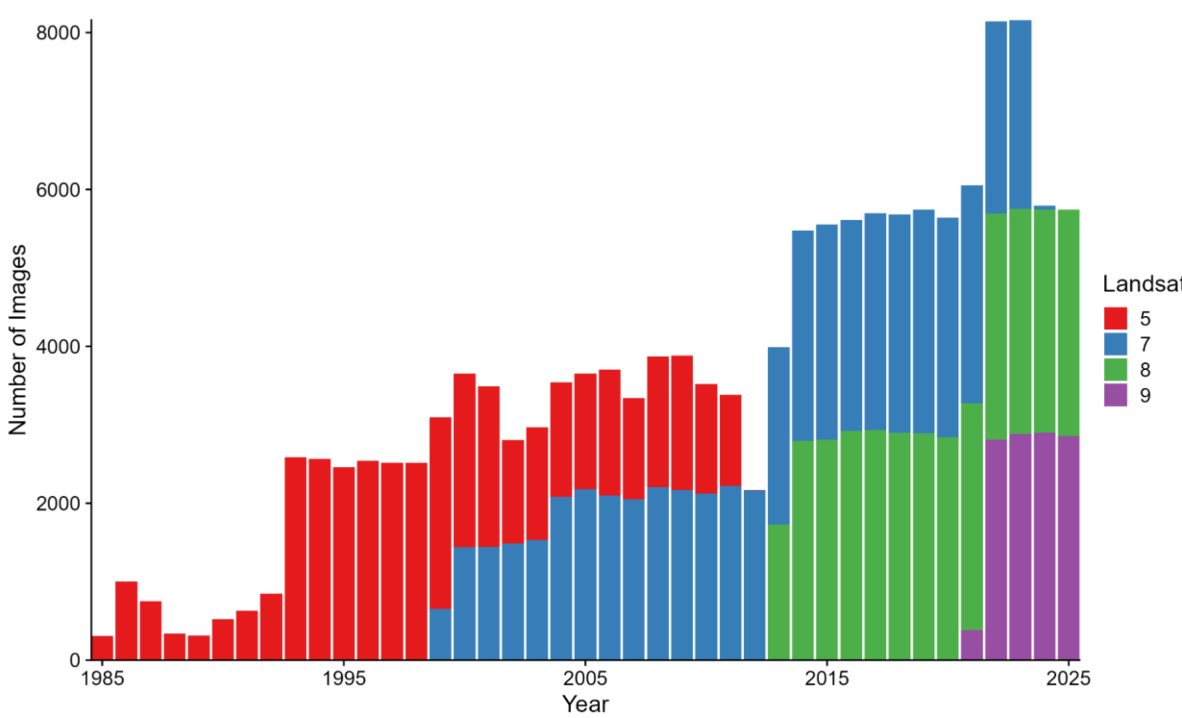

Figure 11. Number of available Landsat images for Mexico from 1985 to 2025.

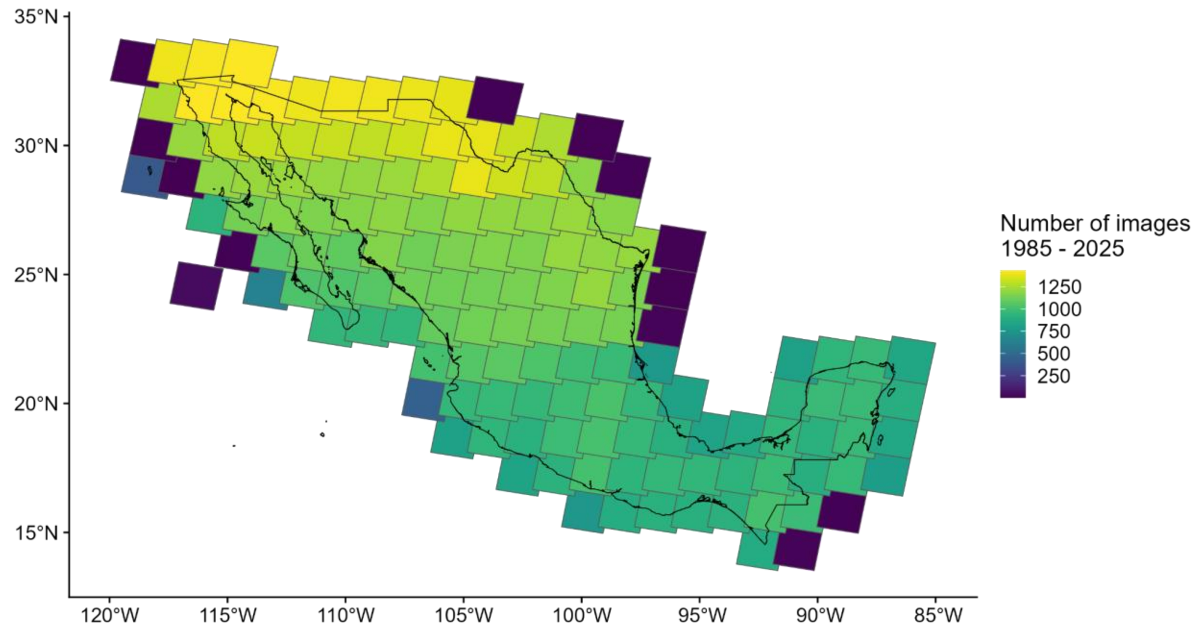

Figure 12. Number of available Landsat images by Path/Row for Mexico from 1985 to 2025.

These critical gaps in the time series affect the methodological robustness and stability of the land cover and vegetation classifiers. With few valid intra-annual observations available, the mapping algorithms are forced to widen their search windows to incorporate pixels from adjacent years, diluting the strict annual resolution of the resulting map. Likewise, the lack of periodic observations distorts the phenological metrics of the vegetation cycle, generating spectral instability or cartographic “flickering.” This translates into erratic jumps and spurious transitions between classes in consecutive years (such as unrealistic alternations between forest and grassland with no change on the ground), as well as omission and commission errors in dynamic land covers such as secondary vegetation (“acahuales”) or seasonal agriculture. At the same time, the lack of clean data produces a detection lag, causing deforestation or land-cover change to be mapped years after it actually occurred, with artificial peaks of vegetation loss appearing only when atmospheric conditions finally allow a clear observation (Figure 13). For example, the scarcity of satellite data over the Isthmus of Tehuantepec and the Yucatán Peninsula during 1985-1992 and 2004-2011 reflects a combination of structural sensor limitations and extreme geo-atmospheric conditions.

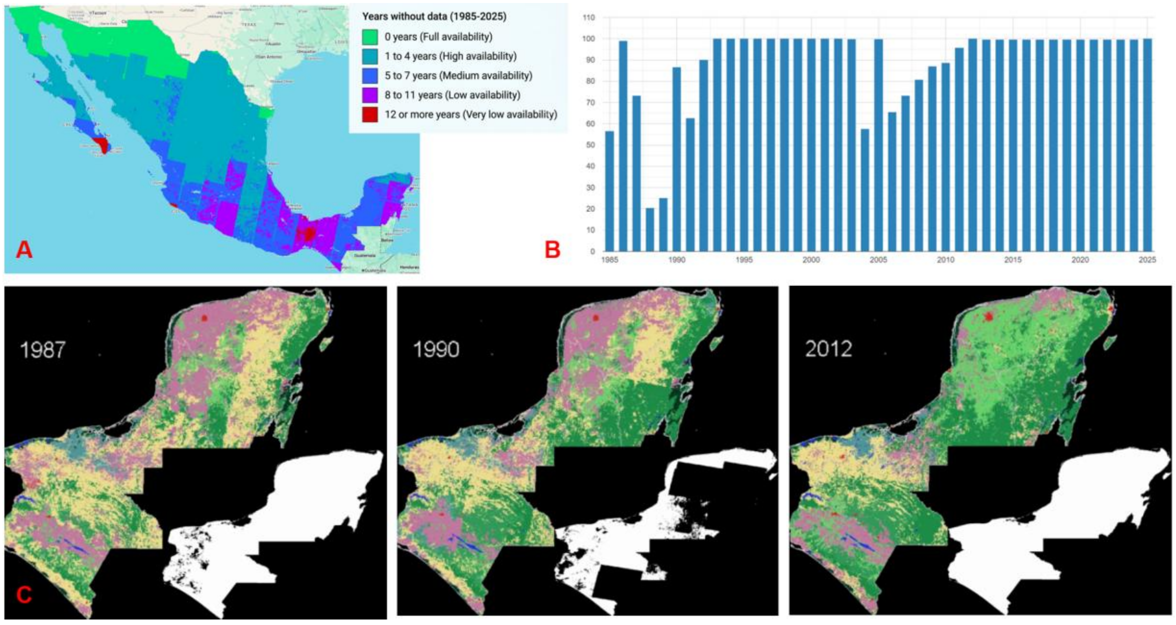

Figure 13. Anomalous LULC classification patterns driven by valid pixel scarcity: A) Spatial distribution of the cumulative number of years lacking data. B) Annual percentage of available pixels. C) Between 1985 and 1994, classification artifacts aligned with Landsat scene boundaries (path/row) are evident in southeastern Mexico. The 1987 and 1990 classifications clearly illustrate this issue. From 2012 onwards, these anomalies attenuate or disappear.

To ensure analytical reliability and mitigate these biases when interpreting the cartographic products for these regions and periods, specific methodological criteria must be applied. The post-classification Temporal, Frequency, and Modal Filters described in Sections 3.6.2, 3.6.4, and 3.6.5 already reduce part of this instability by enforcing temporal consistency and removing implausible transitions; however, they were not designed to specifically target the historically data-poor windows identified here, and some residual noise is expected to persist in these areas. Building on that existing framework, for the Isthmus of Tehuantepec and the Yucatán Peninsula during 1985-1992 and 2004-2011 it is recommended to prioritize the multi-year transition options already defined in Section 3.8 (five-year or ten-year periods) over strictly annual transitions, so that the detection lag and pixel gaps are more robustly absorbed by the regional statistics. It is likewise recommended to make more systematic use of the QA_PIXEL band and of the observation-count layers already available in the mosaic collections, making the associated uncertainty explicit in technical reports and area statistics whenever they cover these regions and periods.

4. Practical Considerations and Challenges

Despite the anomalous classification patterns identified in specific regions and periods (Section 3.9), the 1.0 collection of annual land cover and land use maps for Mexico is a strategic monitoring tool that reflects the country's history over more than four decades. This multi-year information can support applications for estimating trends in land cover change and understanding the factors that modify land cover dynamics, provided the usual precautions are taken when interpreting data derived from remote sensing.

For the development of this project, with its unprecedented spatial and temporal scope, a standardized methodology was used that can be replicated in other areas of the world. The use of Google Earth Engine's cloud-based work platforms and open-source technology has shown promise for large-scale data accessibility and processing.

Through the learning and experience gained from producing this first collection, along with the exchange of ideas with the teams from the various MapBiomas initiatives, it was possible to achieve greater efficiency in terms of time and processes. Through the collaborative and networked work of a multidisciplinary team, it was possible to develop a methodology tailored to the specific needs of each region.

The use of the Random Forest algorithm as a classifier, combined with a flexible mapping protocol, allows each country in the various initiatives to define its feature space and samples.

Applying post-classification filters reduced the effects associated with the low quality and limited availability of satellite imagery, which were most prevalent at the beginning of the time series. Furthermore, the inclusion of new, hierarchically integrated cross-cutting themes made it possible to provide greater thematic detail in the land use and land cover maps.

The next step in this project is to expand the level of detail in the legend, increase the accuracy of the mapping, and use ancillary data, new remote sensing technologies and tools to obtain a higher quality product.

5. References

Abercrombie, S. P., & Friedl, M. A. (2016). Improving the consistency of multitemporal land cover maps using a hidden Markov model. IEEE Transactions on Geoscience and Remote Sensing, 54(2), 703–713.

Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. ISPRS Journal of Photogrammetry and Remote Sensing, 114, 24–31.

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324

CONABIO. (2020). Sistema de Monitoreo de los Manglares de México (SMMM). Comisión Nacional para el Conocimiento y Uso de la Biodiversidad, México.

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., et al. (2007). The Shuttle Radar Topography Mission. Reviews of Geophysics, 45(2), RG2004. https://doi.org/10.1029/2005RG000183

Fosnight, E. A., Shaw, J., Masek, J. G., & Roy, D. P. (2016). The global Landsat archive: Status, consolidation, and direction. Remote Sensing of Environment, 185, 271–283.

Gong, W., Fang, S., Yang, G., & Ge, M. (2017). Using a hidden Markov model for improving the spatial-temporal consistency of time series land cover classification. ISPRS International Journal of Geo-Information, 6(10), 292.

Google. (2019a). Google Earth Engine API. https://developers.google.com/earth-engine/

Google. (2019b). Landsat Collections. Google Earth Engine Data Catalog. https://developers.google.com/earth-engine/datasets/catalog/landsat/

Google. (2020). Reducers: ImageCollection reductions. Google Earth Engine Guides. https://developers.google.com/earth-engine/guides/reducers_image_collection

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment, 202, 18–27.

Holdridge, L. R. (1967). Life Zone Ecology. San José, Costa Rica: Tropical Science Center.

INEGI. (2001). Provincias fisiográficas de México (escala 1:1 000 000). Instituto Nacional de Estadística y Geografía, México.

INEGI, CONABIO, & INE. (2008). Ecorregiones terrestres de México (escala 1:1 000 000). México.

INEGI. (2018). Uso del suelo y vegetación, Serie VII (escala 1:250 000). Instituto Nacional de Estadística y Geografía, México.

INEGI. (2021). Conjunto nacional de uso del suelo y vegetación (escala 1:250 000). Instituto Nacional de Estadística y Geografía, México.

Odum, E. P. (1969). The strategy of ecosystem development. Science, 164(3877), 262–270.

Olofsson, P., Foody, G. M., Herold, M., Stehman, S. V., Woodcock, C. E., & Wagner, W. (2014). Good practices for estimating area and assessing accuracy of land change. Remote Sensing of Environment, 148, 42–57. https://doi.org/10.1016/j.rse.2014.02.015

Prentice, I. C., Cramer, W., Harrison, S. P., Leemans, R., Monserud, R. A., & Solomon, A. M. (1992). A global biome model based on plant physiology and dominance, soil properties and climate. Journal of Biogeography, 19(2), 117–134.

Rzedowski, J. (1978). Vegetación de México. Limusa, México. 432 pp.
Souza, C. M., Shimbo, J. Z., Rosa, M. R., Parente, L. L., et al. (2020). Reconstructing three decades of land use and land cover changes in Brazilian biomes with Landsat archive and Earth Engine. Remote Sensing, 12(17), 2735. https://doi.org/10.3390/rs12172735
Svenning, J.-C., & Sandel, B. (2013). Disequilibrium vegetation dynamics under future climate change. American Journal of Botany, 100(7), 1266–1286.
Tumer, K., & Ghosh, J. (1996). Analysis of decision boundaries in linearly combined neural classifiers. Pattern Recognition, 29(2), 341–348.
Whittaker, R. H. (1975). Communities and Ecosystems (2nd ed.). New York: Macmillan.
Woodward, F. I. (1987). Climate and Plant Distribution. Cambridge: Cambridge University Press.
Zhu, Z., & Woodcock, C. E. (2012). Object-based cloud and cloud shadow detection in Landsat imagery. Remote Sensing of Environment, 118, 83–94. https://doi.org/10.1016/j.rse.2011.10.028

6. Appendices

Specific procedures for each cross-cutting theme (Mangrove, Urban Infrastructure) are documented in the appendices, available at: https://mexico.mapbiomas.org/en/download-of-atbds