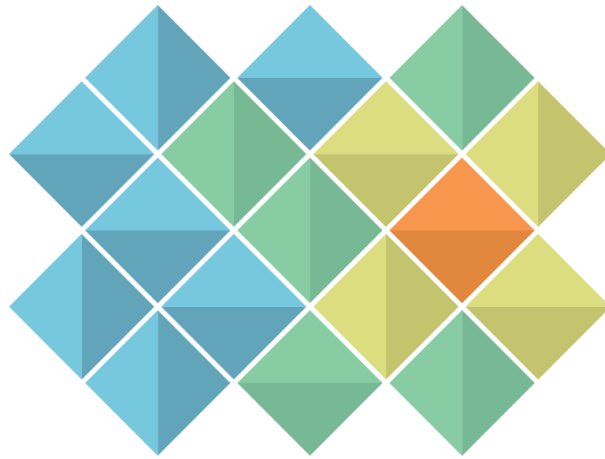




MAPBIOMAS MÉXICO

**Algorithm Theoretical Basis
Document (ATBD):
Urban Transversal Class
Handbook**

**Collection 1
Version 1.0**

Document Information

Item	Description
Document	Algorithm Theoretical Basis Document (ATBD)
Theme	Urban Transversal Class
Collection	1
Version	1.0
Authors	Auliz Ortiz, Daniel Martín: Methodology, Software, Validation, Formal Analysis, Writing – original draft, Writing – review & editing; Antonio Némiga, Xanat: Validation, Writing – original draft, Writing – review & editing; Quevedo Chacón, Alexander: Methodology, Software, Formal Analysis, Data curation, Validation; Hernández Zetina, Sandra Lucía: Project administration, Writing – review & editing; López Martínez, Jorge Omar: Validation, Writing – review & editing; Couturier, Stephane: Validation, Writing – review & editing; Osorno Covarrubias, Javier: Validation, Writing – review & editing; Mas Caussel, Jean François: Validation, Writing – review & editing; Castillo Santiago, Miguel Ángel: Validation, Writing – review & editing; Álvarez, Pérez, Andrea: Validation, Writing – review & editing, Validation; Flores Gatica, Mary Carmen: Writing – review & editing; Pérez Vega, Azucena: Validation; Torres Vivar, Juan Esteban: Validation; Gamboa Blanco, Eric Antonio: Validation.
Publication Date	July 2026
License	CC BY 4.0
Citation	MapBiomás México. (2026). Algorithm Theoretical Basis Document (ATBD): Urban Transversal Class Handbook, Collection 1, Version 1.0.

Table of Contents

1. Introduction.....	1
2. Defining the Urban Transversal Class.....	1
2.1 Conceptual Definition.....	1
2.2 Exclusion Criteria	1
3. Methodology	2
3.1 Mosaic Generation	3
3.2 Operational Regionalization and Grid-Based Partitioning	7
3.3 Training Sample Generation.....	8
3.4 Estimation of Urban-Class Probabilities	10
3.5 Probability Threshold Determination.....	11
3.6 Spatial Filtering.....	13
3.7 Temporal Filtering	14
3.8 Spatial Mask Application	17
4. Quality Control	17
4.1 Collaborative assessment.....	17
4.2 Limitations of the final product.....	19
5. Lessons Learned	19
References.....	20

1. Introduction

This document describes the methodological framework used to generate the urban area maps for Collection 1 of MapBiomass Mexico, covering the period 1985-2025. Its objective is to document the conceptual criteria, information sources, and procedures used to identify and map urban areas on a national scale consistently.

Urban mapping constitutes a transversal collection within MapBiomass Mexico; therefore, it was developed using a methodology independent of the land cover and land use collections, allowing for the incorporation of specific criteria and procedures to improve the detection of urbanized areas. Subsequently, the results were integrated into the national collection.

The methodology was based on Landsat imagery, spectral and textural variables, Random Forest models, spatial and temporal filters, and collaborative assessment processes. The following sections describe the conceptual definition, methodology, quality control, limitations, and lessons learned.

2. Defining the Urban Transversal Class

2.1 Conceptual Definition

For the Urban Transversal Class (UTC), all non-natural surfaces subject to human influence are considered urban. For classification purposes, this land cover includes four main components:

- **Built-up areas**
Residential buildings, commercial buildings, public facilities, and industrial facilities.
- **Functionally urban areas**
Industrial zones, airports, ports, industrial parks, and logistics hubs.
- **Undeveloped urban land**
Vacant lots, construction sites, and intra-urban open spaces (such as urban parks) that are part of a continuous urban patch.
- **Dispersed human settlements**
Settlements with an incipient urban structure that meet the minimum area and temporal persistence criteria.

2.2 Exclusion Criteria

To ensure the thematic consistency of the urban class, land cover types and features that are not part of the consolidated urban fabric are excluded, regardless of whether they exhibit spectral or textural characteristics like those of urbanized areas. These include:

- Agricultural land
- Natural vegetation
- Isolated bare soil (not contiguous with urban land use)
- Natural arid areas
- Beaches and dunes
- Mines and quarries
- Isolated linear infrastructure

In addition, the following spatial and temporal inclusion and exclusion criteria were established:

- **Spatial criteria:**

Minimum spatial connectivity (contiguity assessed within a 60-m radius).

Continuous urban patch consisting of 5 to 9 connected pixels (approximately 0.45–0.8 ha at 30-m spatial resolution).

- **Temporal criterion:**

Persistence of urban conditions for 2–3 consecutive years or spectral stability verified through multitemporal analysis.

3. Methodology

The methodological workflow was organized into eight stages: (1) generation and validation of synthetic image mosaics; (2) operational regionalization; (3) generation of training samples; (4) estimation of urban-class probabilities; (5) definition of probability thresholds to assign pixels to the urban class in the final map; (6) application of spatial filters; (7) application of temporal filters; and (8) application of a spatial mask to improve the quality of the final map product (Figure 1).

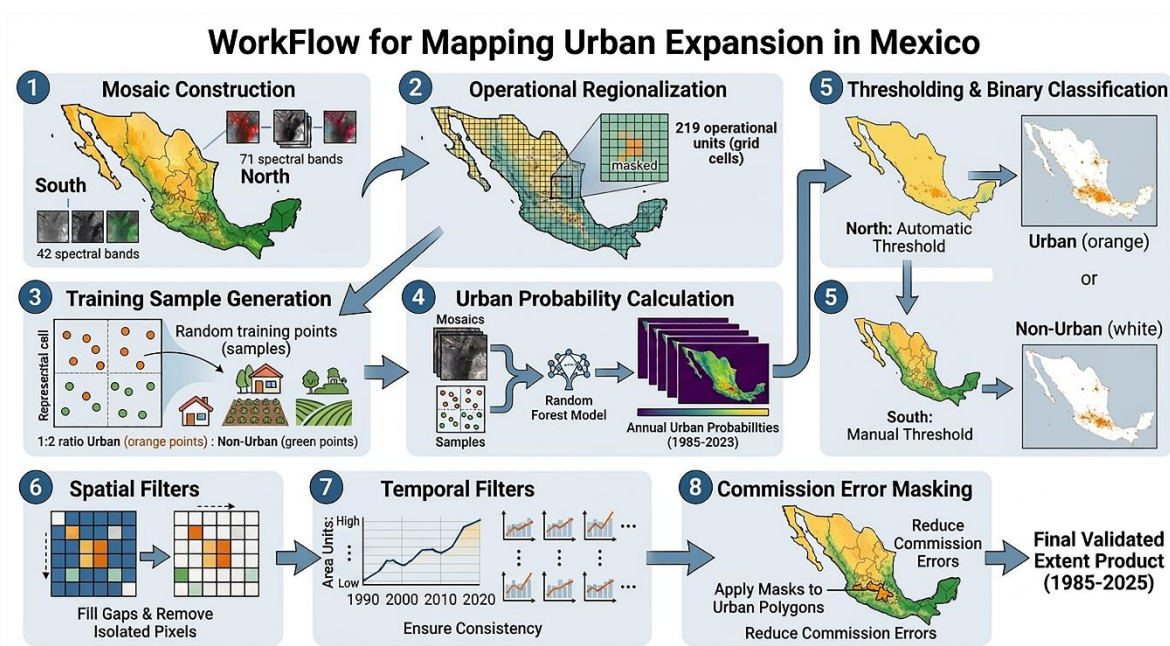


Figure 1. General workflow of the proposed methodology.

Monitoring built-up areas over time is essential for understanding the dynamics of the national territory (Couturier et al., 2020). The UTC was developed using the official cartographic products generated by the National Institute of Statistics and Geography (INEGI). The delineation of urban localities and the Land Use and Vegetation Series corresponding to the different periods analyzed were used as reference datasets. These products provided the cartographic baseline for identifying and characterizing urban areas throughout the study time series.

As an auxiliary dataset, the Global Human Settlement Layer (GHSL), developed by the Joint Research Centre (JRC) of the European Commission, was incorporated. This dataset

provides multitemporal information on the distribution and evolution of human settlements at the global scale, derived from satellite imagery and census data (Pesaresi and Politis, 2023). In particular, the built-up surface and human settlement layers were used to support the identification and validation of urbanized areas.

The National Geostatistical Framework, developed by the National Institute of Statistics and Geography (INEGI, 2025), was also used. Specifically, the vector boundaries of the Basic Geostatistical Areas (AGEBs) served as a reference for delineating and verifying urban areas.

3.1 Mosaic Generation

3.1.1 Spectral Data Sources

The classification of the UTC was based on imagery acquired by the Landsat missions and their respective sensors (Table 1). These satellites provide multispectral radiometric data for characterizing the Earth's surface at medium spatial resolution (on the order of tens of meters). The sensors used to generate image mosaics for different periods across Mexico are listed below.

Table 1

Spectral data sources used for the classification of the UTC.

<i>Sensor</i>	<i>Mission</i>	<i>Sensor Type</i>	<i>Number of Bands</i>	<i>Spatial Resolution</i>	<i>Image Availability</i>
OLI / OLI-2	Landsat 8 and Landsat 9	Multispectral (Push-broom)	9 bands (Coastal/Aerosol, Visible, NIR, SWIR, Cirrus, and Panchromatic)	30 m (15 m panchromatic)	February 2013 (L8) / September 2021 (L9) – Present
TIRS / TIRS-2	Landsat 8 and Landsat 9	Thermal Infrared (QWIP)	2 thermal bands	100 m (resampled to 30 m)	February 2013 (L8) / September 2021 (L9) – Present
ETM+	Landsat 7	Multispectral Scanner (Whisk-broom)	8 bands (Visible, NIR, 2 SWIR, 1 Thermal, and 1 Panchromatic)	30 m (60 m thermal, 15 m panchromatic)	April 1999 – April 2022 (Extended mission phase thereafter)
TM	Landsat 4 and Landsat 5	Multispectral Scanner	7 bands (Visible, NIR, 2 SWIR, and 1 Thermal)	30 m (120 m thermal)	July 1982 – November 2011

Source: United States Geological Survey (USGS). Landsat Satellite Missions, <https://www.usgs.gov/landsat-missions/landsat-satellite-missions>

The Landsat series provides a consistent foundation for generating multitemporal mosaics, owing to the availability of spectrally comparable bands across sensors and observation periods. This characteristic enables the integration of imagery acquired at different dates, the application of radiometric harmonization procedures, and quality control processes designed to reduce noise and other artifacts present in the original imagery.

3.1.2 Urban Classification Variables and Indicators (Urban Indices, Thermal Variables, and Texture Metrics)

The classification of the UTC was based on a set of spectral bands, spectral indices, thermal variables, and texture metrics derived from Landsat imagery (Tables 2 and 3). Owing to the environmental contrasts between northern and southern Mexico, particularly in precipitation regimes and land cover patterns, region-specific combinations of variables were defined to improve classification performance and enhance the discrimination of the urban class.

Table 2

Spectral bands and indices used as input variables for the classification of urban areas in southern Mexico.

Variable	Full Name	Formula	Source
BLUE, GREEN, RED, Original Reflectance Bands NIR, SWIR1, SWIR2		Untransformed surface reflectance values	---
NDVI	Normalized Difference Vegetation Index	$(NIR - RED) / (NIR + RED)$	Rouse et al. (1974)
EVI	Enhanced Vegetation Index	$2.5 \cdot (NIR - RED) / (NIR + 6 \cdot RED - 7.5 \cdot BLUE + 1)$	Huete et al. (2002)
EVI2	2-band Enhanced Vegetation Index	$2.5 \cdot (NIR - RED) / (NIR + 2.4 \cdot RED + 1)$	Jiang et al. (2008)
SAVI	Soil Adjusted Vegetation Index	$[(NIR - RED) / (NIR + RED + 0.5)] \cdot 1.5$	Huete (1988), L=0.5
MNDWI	Modified Normalized Difference Water Index	$(GREEN - SWIR1) / (GREEN + SWIR1)$	Xu (2006)
NDWIm	Normalized Difference Water Index (McFeeters)	$(GREEN - NIR) / (GREEN + NIR)$	McFeeters (1996)
AWEIsh	Automated Water Extraction Index (shadow)	$BLUE + 2.5 \cdot GREEN - 1.5 \cdot (NIR + SWIR1) - 0.25 \cdot SWIR2$	Feyisa et al. (2014)
NDBI	Normalized Difference Built-up Index	$(SWIR1 - NIR) / (SWIR1 + NIR)$	Zha et al. (2003)
UI	Urban Index	$(SWIR2 - NIR) / (SWIR2 + NIR)$	Kawamura et al. (1996)
NDMI	Normalized Difference Moisture Index	$(NIR - SWIR1) / (NIR + SWIR1)$	Bo-cai Gao (1996)
BSI	Bare Soil Index	$[(SWIR1 + RED) - (NIR + BLUE)] / [(SWIR1 + RED) + (NIR + BLUE)]$	Rikimaru (1996)
BU	Built-up Index	$NDBI - NDVI$	Zha et al. (2003)
EBBI	Enhanced Built-Up and Bareness Index	$(MIR - NIR) / [10 \cdot V(MIR + TIR)]$	As-syakur et al. (2012)
NBR	Normalized Burn Ratio	$(NIR - SWIR2) / (NIR + SWIR2)$	Key & Benson (2006)
NDRI	Normalized Difference — Urban Features	$(RED - BLUE) / (RED + BLUE)$	De Pinho, Ummus & Novack (2011)

Variable	Full Name	Formula	Source
BAI	Blue/NIR Normalized Difference Index	$(\text{BLUE} - \text{NIR}) / (\text{BLUE} + \text{NIR})$	Santos et al. (2022), Remote Sensing 14(3):704
NDFI	Normalized Difference Fraction Index	Normalized Difference (GVS, NPV+SOIL), rescaled to byte $(\text{NDFI} \cdot 100 + 100)$	Souza et al. (2005)
GV, NPV, SOIL, CLOUD	Spectral Mixture Analysis (SMA) Fraction Images — Four Endmembers	Outputs of linear spectral unmixing using fixed reflectance endmembers (BLUE–SWIR2)	Souza et al. (2005)
SHADE	Shade (spectral unmixing residual)	$ 100 - (\text{GV} + \text{NPV} + \text{SOIL}) $	Souza et al. (2005)
GVS	Shade-Normalized Green Vegetation (GV)	$\text{GV} / (\text{GV} + \text{NPV} + \text{SOIL}) \cdot 100$	Souza et al. (2005)
SUBS, VEG, DARK	Fractions from an Independent Three-Endmember Spectral Unmixing Model (Substrate, Vegetation, and Dark)	Outputs of linear spectral unmixing using a distinct set of reflectance endmembers GV/NPV/SOIL/CLOUD	No published reference identified
EVI_p10, EVI2_p10, EVI_p90, EVI2_p90	10th and 90th Temporal Percentiles	Temporal statistic computed from the pixel time series	---
EVI_dif9010, EVI2_dif9010	90th–10th Percentile Difference	Temporal variability of the index $(\text{P90} - \text{P10})$	---
EBBI_p25, EBBI_p75, EBBI_p90, EBBI_dif7525	Temporal Percentile of EBBI	Temporal variability of the EBBI (p25, p75, p90, and P75–P25 difference)	---

Table 3

Additional spectral bands and indices are used as input variables for classifying urban areas in northern Mexico.

Variable	Full Name	Formula	Source
BLFEI	Built-up Land Features Extraction Index	$((\text{G} + \text{R} + \text{S2}) / 3 - \text{S1}) / ((\text{G} + \text{R} + \text{S2}) / 3 + \text{S1})$	Bouhennache et al. (2019), Geocarto Int. 34(14):1531–1551
BRBA	Band Ratio for Built-up Area	$\text{RED} / \text{SWIR1}$	Waqar et al. (2012)
NBAI	Normalized Built-up Area Index	$(\text{S2} - \text{S1}/\text{G}) / (\text{S2} + \text{S1}/\text{G})$	Waqar et al. (2012)
PISI	Perpendicular Impervious Surface Index	$0.8192 \cdot \text{BLUE} - 0.5735 \cdot \text{NIR} + 0.0750$	Lin & Wang (2011), cit. en Harrak et al. (2025)
VIBI	Vegetation Index Built-up Index	$\text{NDVI} / (\text{NDVI} + \text{NDBI})$	Awesome Spectral Indices for GEE (Montero Loaiza)
VgNIRBI	Visible Green-Based Built-Up Index	$(\text{GREEN} - \text{NIR}) / (\text{GREEN} + \text{NIR})$	Estoque & Murayama (2015)
VrNIRBI	Visible Red-Based Built-Up Index	$(\text{RED} - \text{NIR}) / (\text{RED} + \text{NIR})$	Estoque & Murayama (2015)
IBI	Index-based Built-up Index	$(\text{NDBI} - \text{partA}) / (\text{NDBI} + \text{partA}), \text{partA} = [\text{SAVI} + (\text{GREEN} - \text{SWIR1}) / (\text{GREEN} + \text{SWIR1})] / 2$	Xu (2008), Int. J. Remote Sensing 29(14)

Variable	Full Name	Formula	Source
IISI	Isoline Impervious Surface Index	$1e8 / \sqrt{[(NIR/OPT_SCALE)^2 + (1 - SWIR1/OPT_SCALE)^2]}$	Martin, Blesius & Liu (2026)
MBAI	Modified Built-up Area Index	$[NIR + 1.57 \cdot GREEN + 2.4 \cdot SWIR1] / [1 + NIR] \times 1e-6$	Benkouider et al. (2019), cit. en Martin et al. (2026)
BAEM	Built-up Area Extraction Method	$NDBI - (SAVI + MNDWI) / 2$, where $MNDWI = (GREEN - SWIR1) / (GREEN + SWIR1)$ recalculada	Bhatti & Tripathi (2014)
DBSI	Dry Bare-Soil Index	$(SWIR1 - GREEN) / (SWIR1 + GREEN) - NDVI$	Rasul et al. (2018)
MBI	Modified Bare Soil Index	$(SWIR1 - SWIR2 - NIR) / (SWIR1 + SWIR2 + NIR) + 0.5$	Nguyen et al. (2021)
MNDBaI	Modified Normalized Difference Bareness Index	$(SWIR2 - SWIR1) / (SWIR2 + SWIR1)$	No published reference identified
DULI1	Desert Urban Landscape Index (v1)	$(BLUEn - TIR) / (BLUEn + TIR) - NDBI - NDVI$, where $BLUEn = BLUE / OPT_SCALE$	Martin, Blesius & Liu (2026)
DULI2	Desert Urban Landscape Index (v2)	$(BLUEn - TIR) / (BLUEn + TIR) - NDBI$	Martin, Blesius & Liu (2026)
DULI3	Desert Urban Landscape Index (v3)	$DULI1 - (SWIR1 - SWIR2) / (SWIR1 + SWIR2)$	Martin, Blesius & Liu (2026)
DBI	Dry Built-up Index	$(BLUEn - TIR) / (BLUEn + TIR) - NDVI$	Rasul et al. (2018), cit. en Martin et al. (2026)
NBLI	Normalized Difference Bare Land Index	$(REDn - TIR) / (REDn + TIR)$, where $REDn = RED / OPT_SCALE$	Remote Sensing 9(3):249 (2017)
NIR_asm_5 / _11	Angular Second Moment (NIR)	GLCM texture metric ¹	Haralick et al. (1973)
NIR_contrast_5 / _11	Contraste (NIR)		
NIR_var_5 / _11	Varianza (NIR)		
NIR_idm_5 / _11	Inverse Difference Moment (NIR)		
NIR_ent_5 / _11	Entropía (NIR)		

3.1.3 Image Mosaic Generation and Validation

For southern Mexico, image mosaics composed of 42 bands were generated for the 1985–2025 period using an adaptation of the processing workflow and variable definitions originally developed by MapBiomás Brazil (Table 2). For northern Mexico, characterized by predominantly arid and semi-arid environments, the same set of variables was supplemented with additional spectral indices specifically designed to improve the characterization of dryland landscapes (Table 3). As a result, the mosaics generated for the northern region comprised 71 bands.

Image mosaics were generated from Landsat imagery with 30-m spatial resolution covering the 1985–2025 period. Mosaic generation prioritized images that provided the most complete spatial coverage while minimizing missing pixels, radiometric noise, cloud

¹ All GLCM texture metrics were computed from integer-valued NIR data using 5 × 5 (size = 2) and 11 × 11 (size = 5) moving windows.

contamination, and other artifacts that could adversely affect classification performance (Figure 2). Additionally, image selection emphasized radiometric characteristics, including tonal depth and contrast, to maximize the spectral separability of highly reflective urban surfaces, vegetation, and water bodies.

To reduce the volume of data to be processed while maximizing the availability of high-quality imagery for mosaic generation, the analysis was organized into multi-year compositing periods, with their duration adjusted according to the availability of suitable Landsat observations. Because suitable Landsat imagery for 1985 was insufficient, the first compositing period (1986–1990) was used to represent both 1985 and the years within that interval. The resulting compositing periods were 1986–1990, 1991–1995, 1996–2000, 2001–2005, 2006–2010, 2011–2015, 2016–2020, and 2021–2025.



Figure 2. Example of a Landsat image mosaic for the southern region of Mexico year 2025.

3.2 Operational Regionalization and Grid-Based Partitioning

To reduce the amount of data processed simultaneously, improve classification performance, and optimize computational resources, the Mexican territory was partitioned into 219 equidistant grid cells, each measuring 100×100 km, providing complete national coverage (Figure 3). This partitioning was based on the Level 4 National Geostatistical Grid developed by INEGI (2023). The resulting regionalization follows the same operational strategy as other MapBiomass initiatives, including those in Brazil and Argentina, enabling more efficient model execution while improving computational performance and resource management.

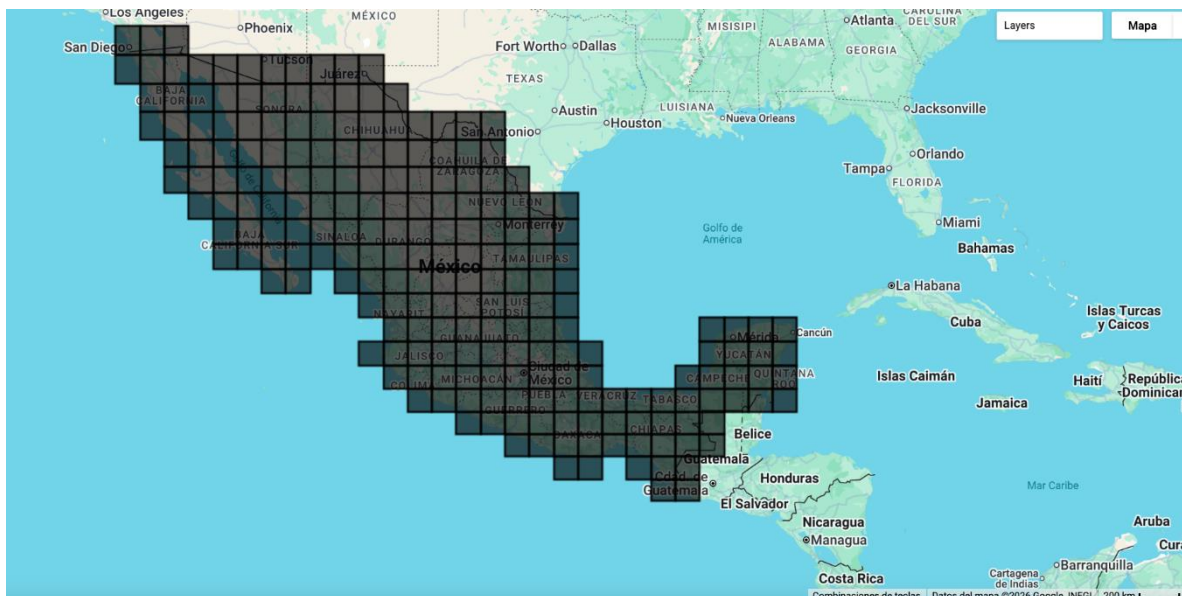


Figure 3. Partitioning of the Mexican territory into equidistant 100 × 100 km grid cells.

As noted previously, to improve classification performance and account for the pronounced environmental and territorial contrasts across Mexico—particularly the more arid conditions in the north relative to the south—the grid cells were grouped into two major operational regions during mosaic generation: Southern Mexico (83 grid cells) and northern Mexico (136 grid cells).

3.3 Training Sample Generation

Training samples for the urban classification were generated using a stratified random sampling approach. As ancillary information for sample generation, the Global Human Settlement Layer (GHSL) built-up surface dataset (Pesaresi and Politis, 2023) was used. This dataset provides annual estimates of built-up surface from 1975 to the present at a spatial resolution of 1 ha.

For each cell of the national geostatistical grid, 3,000 training points were generated using a stratified random sampling scheme, comprising 1,500 urban and 1,500 non-urban samples. Urban areas were identified using a threshold value of 3,000 or greater in the GHSL built-up surface layer. This threshold is consistent with the sampling methodology adopted by the MapBiomias initiatives in Brazil and Argentina.

To improve the quality of the training samples, the values of four spectral indices commonly used to characterize urban areas (NDBI, UI, NDUI, and EBBI) were transformed using Principal Component Analysis (PCA) to generate a composite band corresponding to the first principal component. Training samples with values exceeding ± 2 standard deviations from the mean of the first principal component were subsequently discarded. This procedure was intended to remove outliers from the training dataset, thereby improving the consistency of the training samples.

A preliminary analysis indicated that the optimal number of urban training samples to achieve satisfactory classification accuracy (overall accuracy > 0.85) ranged from 200 to 300. Increasing the number of training samples beyond this range did not result in significant improvements in classification performance. Consequently, this sampling strategy was

adopted because it reduced computational complexity while maintaining high classification accuracy (Figure 4).

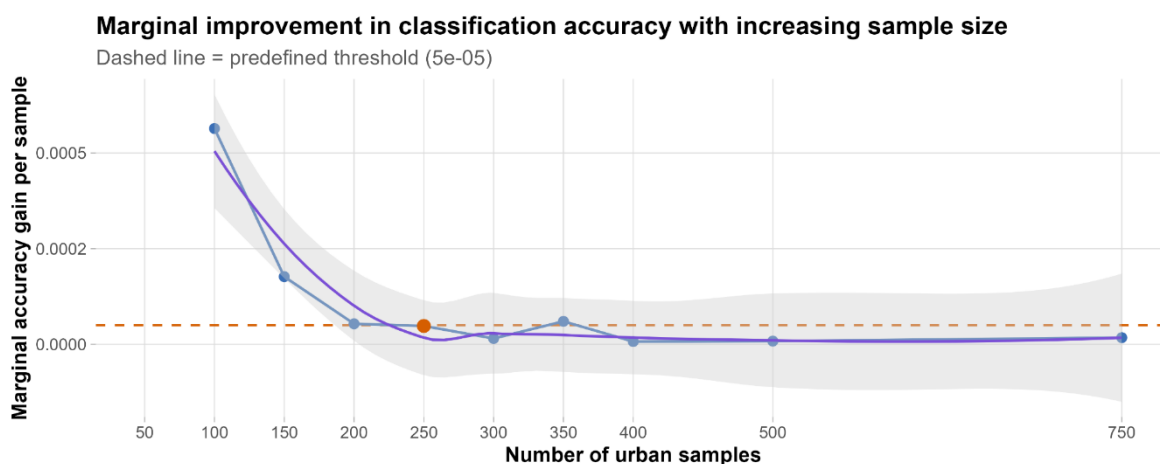


Figure 4. Marginal variation in overall classification accuracy as a function of the number of urban training samples. Beyond approximately 200 samples, additional training points resulted in only marginal improvements in overall accuracy.

After this filtering process, 230 urban training samples were randomly selected for 1985 within each grid cell. In addition, a sufficient number of non-urban samples was randomly selected to maintain an urban-to-non-urban ratio of 1:2, following the sampling strategy adopted by MapBiomias Brazil. This resulted in 460 non-urban training samples per grid cell for 1985.

In subsequent years, the number of urban training samples was allowed to increase in proportion to the urban expansion observed in each grid cell. An additional 70 urban samples were distributed across the eight multi-year compositing periods spanning 1985–2025 (1985–1990, 1991–1995, ..., 2021–2025) in proportion to the share of each cell's total urban growth occurring during each period. Consequently, periods experiencing greater urban expansion received more additional training samples than periods with little or no growth, up to a maximum of 300 urban samples per grid cell (230 baseline samples plus 70 growth-dependent samples). To preserve the 1:2 urban-to-non-urban sampling ratio, up to 140 additional non-urban samples were incorporated, resulting in a maximum of 600 training samples per grid cell. Unlike the urban samples, these additional non-urban samples were drawn from the remaining pool of stable non-urban samples identified for 1985 after PCA-based filtering and were not allocated among compositing periods.

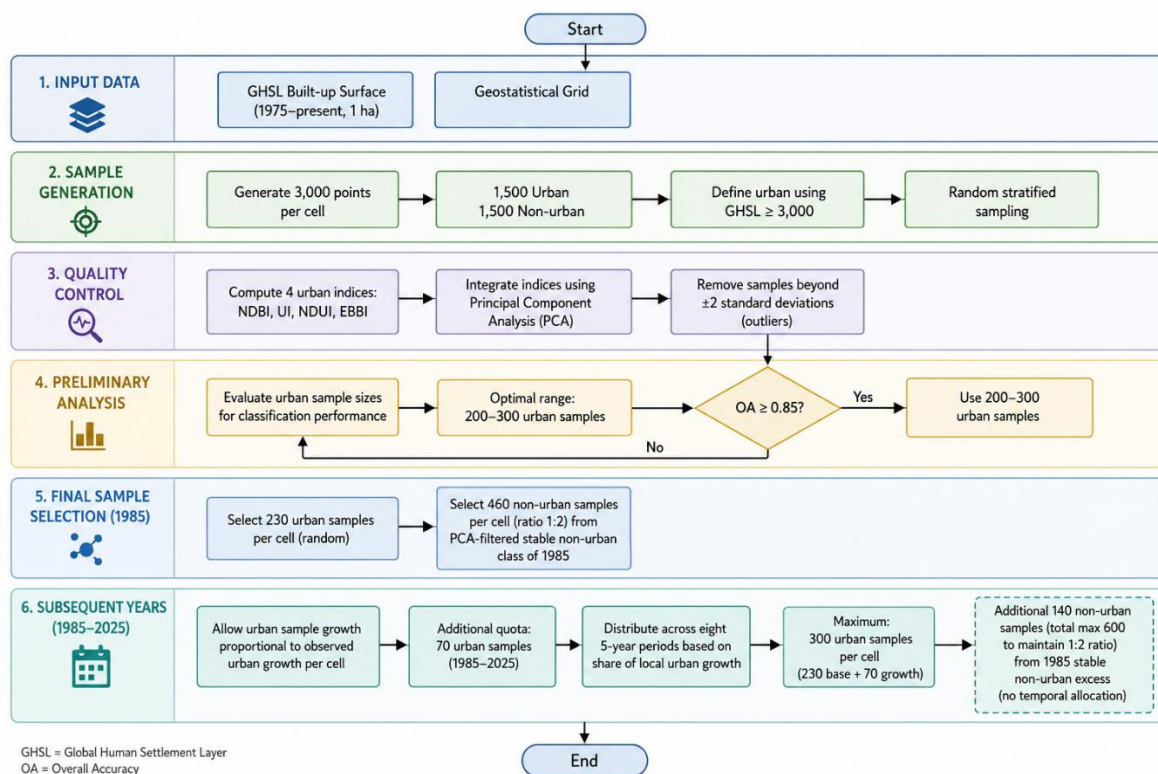


Figure 5. Workflow illustrating the generation of training samples for the classification of urban areas.

3.4 Estimation of Urban-Class Probabilities

As a preliminary step before classifying the land surface into urban and non-urban classes, the Random Forest algorithm was used to generate pixel-wise probability layers indicating the likelihood that each pixel belonged to the urban class.

Random Forest is a supervised machine learning classification algorithm that constructs an ensemble of decision trees using randomly selected subsets of training samples and predictor variables (Belgiu and Drăguț, 2016). Owing to its computational efficiency, robustness, and high classification accuracy, it has become one of the most widely used algorithms for remote sensing image classification. As noted by Pal (2005), "constructing an ensemble of decision trees and allowing them to vote for the most popular class" results in a substantial improvement in land-use and land-cover classification accuracy. Additional advantages of Random Forest include its non-parametric nature, its ability to assess variable importance, its robustness to noisy data and overfitting, its capability to estimate missing values, and its flexibility for a wide range of analytical tasks, including classification, regression, survival analysis, and unsupervised learning. Overall, Random Forest consistently outperforms a single decision tree in terms of predictive accuracy (Rodríguez-Galiano et al., 2012).

Urban-class probabilities were estimated using Random Forest models trained independently for each of the eight multi-year compositing periods defined between 1985 and 2025, with 100 decision trees per model. For each grid cell, the previously generated training samples were used to extract the values of all available predictor bands corresponding to each year within the compositing period. To reduce potential biases arising from insufficient training samples per cell, the training dataset was augmented with samples

from neighboring grid cells within a 3-km buffer around the target cell. However, key model parameters—including the neighborhood buffer distance and the urban-to-non-urban sampling ratio—were defined dynamically, allowing them to be adjusted for specific grid cells where the default configuration did not produce satisfactory classification results.

The resulting models produced a probability map in which each pixel was assigned a probability of belonging to the urban class. This probability corresponds to the proportion of decision trees within the Random Forest ensemble that voted to classify the pixel as urban.

Figure 6 presents the urban-class probability map generated for grid cell 168, corresponding to the Toluca Valley Metropolitan Area.

To improve the temporal consistency of the probabilities generated by the classification models, a temporal harmonization procedure was applied independently to each grid cell within the processing framework. For each year in the 1985–2025 time series, the harmonized probability was computed as the mean of the original probabilities within a ± 2 -year moving window (up to five years for the central years of the series, with progressively smaller windows at the beginning and end of the time series). This was implemented through a temporal self-join on the annual collection of probability images. The resulting output was a multiband image of temporally smoothed urban-class probabilities, with each band corresponding to a single year in the 1985–2025 time series.

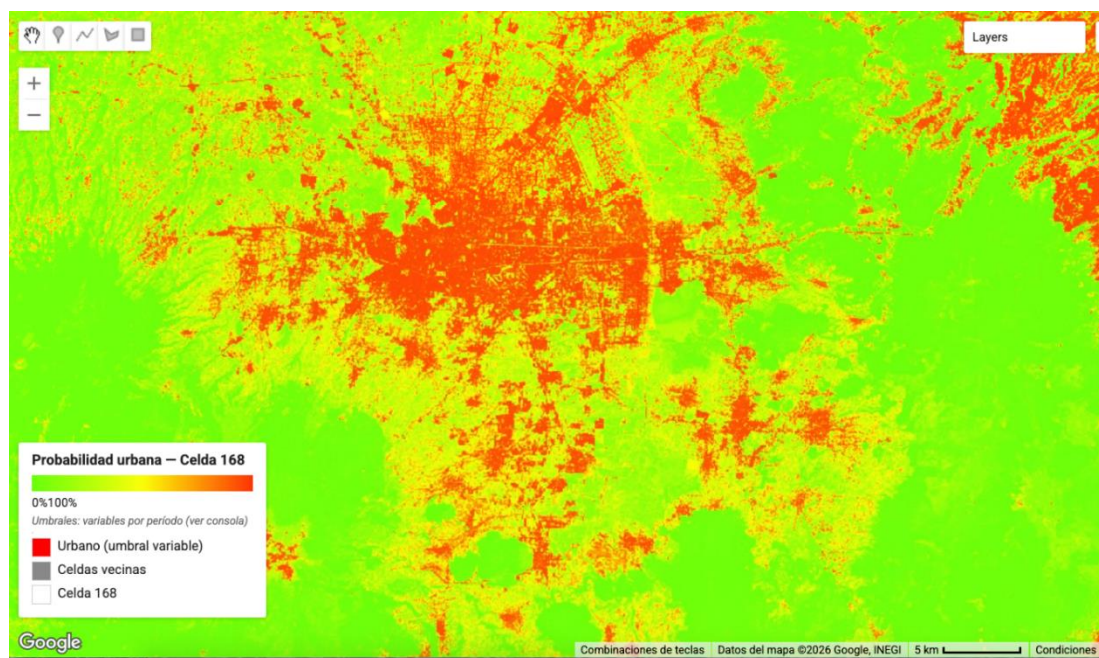


Figure 6. Urban-class probability map.

3.5 Probability Threshold Determination

Two complementary strategies were adopted to determine the probability thresholds for distinguishing urban from non-urban areas using the urban-class probability maps. In the Southern Mexico region, threshold values were determined through cell-by-cell visual interpretation for each of the previously defined multi-year compositing periods. Each grid cell was examined to identify the probability threshold that most accurately represented the shape and spatial extent of the urban footprint. For the historical periods, this interpretation was supported by official geospatial datasets and reference products developed in previous

studies, including the Urban Footprint of Mexico (Couturier et al., 2020), available at https://terradigitalis.igg.unam.mx/html/ojs3/index.php/terra_digitalis/article/view/70. For the Northern Mexico region, the threshold was determined automatically through an optimization procedure. This approach exploited the large number of training samples that had not been selected during model training. From this pool of unused samples, 460 independent urban and non-urban samples were retained for threshold optimization. Urban-class probabilities corresponding to the reference year of each compositing period were then extracted for these independent samples and used to determine the optimal probability threshold.

For the urban reference samples, a threshold denoted as t_{urb} , was computed as the percentile corresponding to the lower tail of the probability distribution (by default, the 5th percentile). This threshold represents the probability cutoff that allows no more than the specified proportion of true urban pixels to be excluded because of low predicted probabilities, thereby controlling omission errors. Similarly, for the non-urban reference samples, a threshold denoted as t_{nonurb} , was computed as the percentile corresponding to the upper tail of the probability distribution (by default, the 99th percentile). This threshold represents the probability cutoff that ensures that no more than 1% of the reference non-urban pixels are incorrectly classified as urban, thereby controlling commission errors. Both percentile thresholds were implemented as dynamic parameters in the classification workflow, allowing their values to be adjusted on a cell-by-cell basis based on classification performance.

The two thresholds were combined by selecting the maximum of the two values ($t_{combined} = \max(t_{urb}, t_{nonurb})$), ensuring that the final threshold simultaneously satisfied both the omission and commission constraints. In addition, an absolute minimum threshold of 75% was enforced to prevent excessively permissive threshold values in cells exhibiting atypical probability distributions. As a quality control measure, the recall of the urban reference samples and the commission error rate of the non-urban reference samples were calculated for each compositing period after applying the final threshold. These metrics provide a direct assessment of how effectively the selected threshold separates the two classes within the reference dataset for each grid cell and compositing period. The resulting probability thresholds used as urban-class membership thresholds are shown in Figure 7.

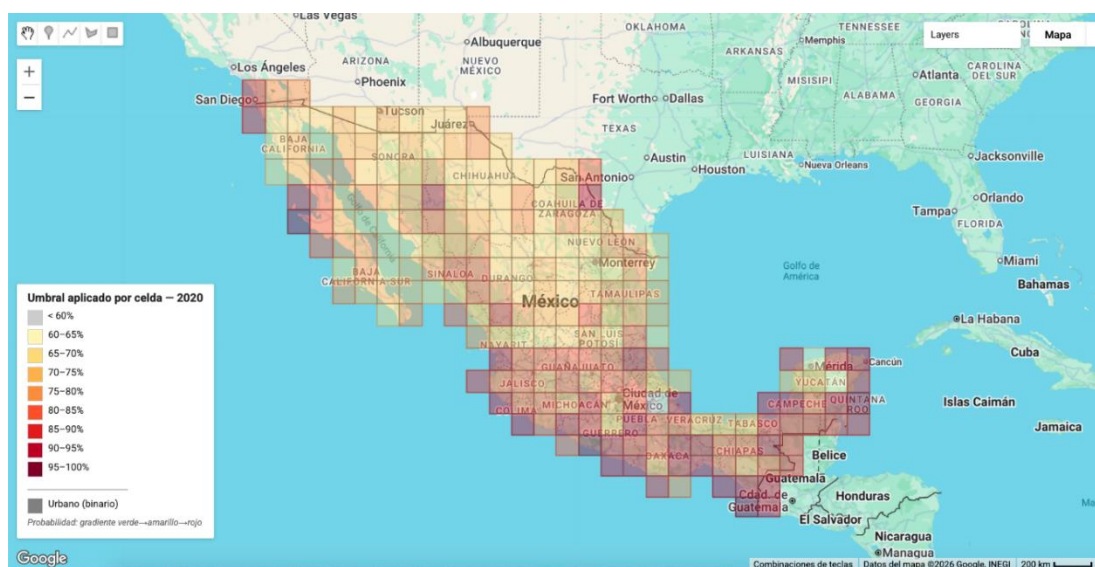


Figure 7. Probability threshold applied for urban mapping in 2020.

Figure 8 illustrates the urban map obtained after applying the selected probability threshold to grid cell 168, encompassing the Toluca Valley Metropolitan Area and the Mexico City Metropolitan Area. The classification was derived from the 2025 Landsat image mosaic.

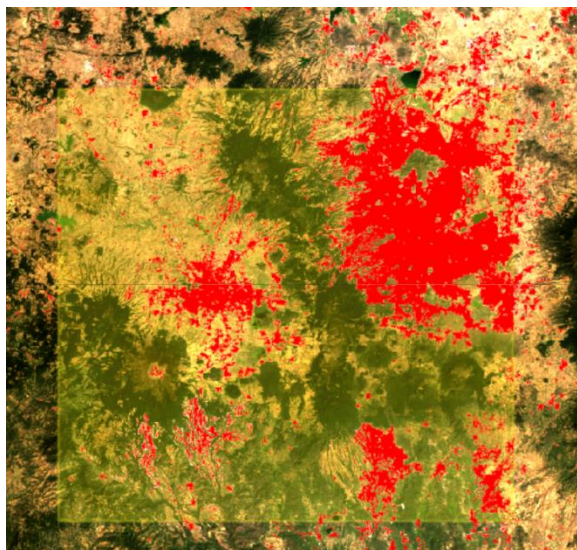


Figure 8. Initial urban footprint map for grid cell 168 in 2025.

3.6 Spatial Filtering

The binary classification maps generated in the previous step were processed with a spatial filtering procedure to fill gaps in the urban area classifications and remove isolated pixels (as in Couturier et al., 2011). Specifically, this filtering step applies a sequence of morphological operations to the binary classification map derived by thresholding the harmonized urban-class probability map.

First, a morphological closing operation (dilation followed by erosion using a circular structuring element with a radius of one pixel) was applied to connect nearby urban fragments and smooth the boundaries of urban patches. Next, a connectivity-based hole-filling procedure was performed by identifying contiguous groups of non-urban pixels smaller than 60 pixels and reclassifying them as urban. This operation removed small non-urban enclaves or holes enclosed within larger urban patches. Subsequently, a morphological opening operation (erosion followed by dilation) was applied to eliminate small protrusions and isolated boundary noise. Finally, a minimum mapping unit filter was applied, in which connected urban patches comprising five or fewer pixels were reclassified as non-urban, thereby removing isolated detections. The resulting map was then reclassified to assign the MapBiomass legend code for Urban Area (Class 24) exclusively to pixels that passed the complete morphological filtering procedure (Figure 9), while all remaining pixels were retained as non-urban (NoData).

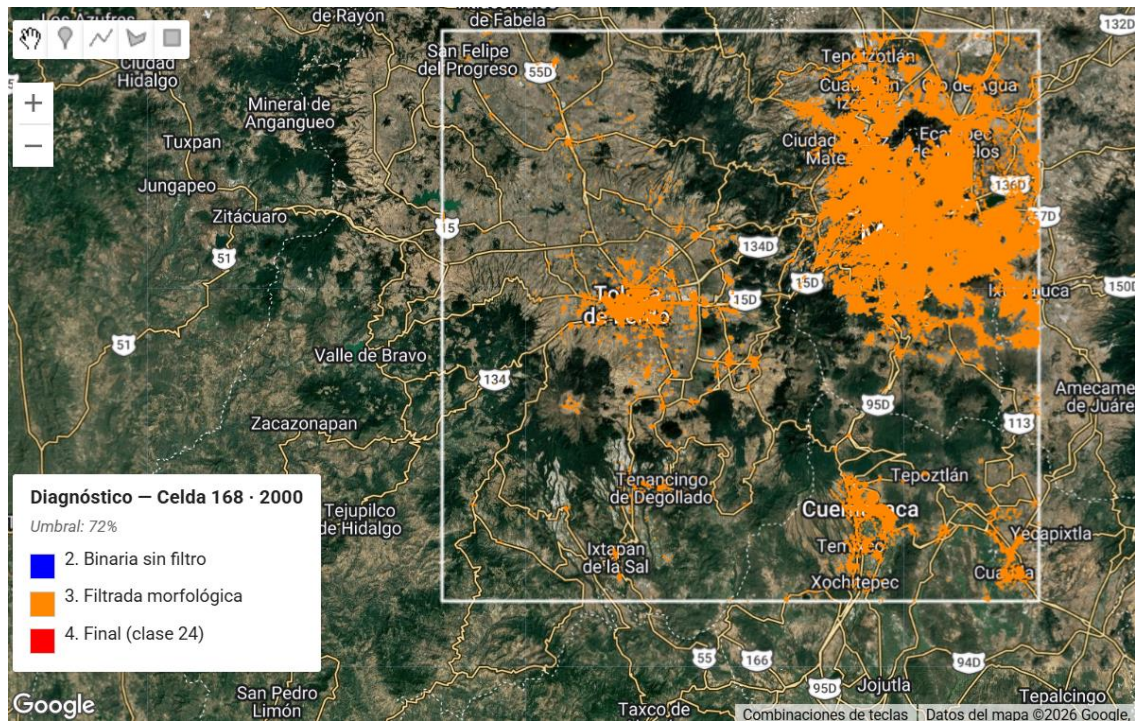


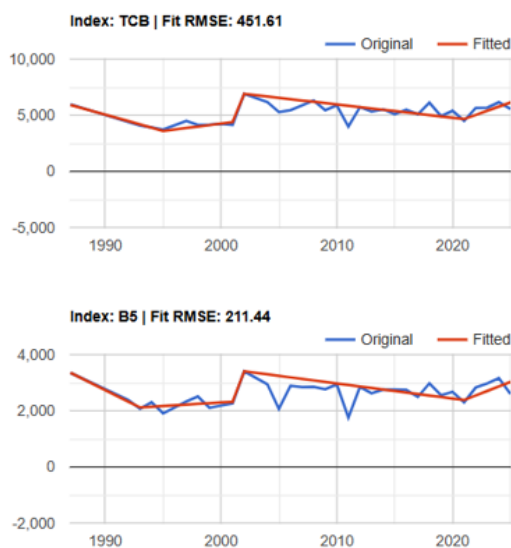
Figure 9. Application of the spatial connectivity filter to urban-classified pixels in grid cell 168.

3.7 Temporal Filtering

From a temporal perspective, a pixel may follow one of two trajectories: (a) it is initially classified as non-urban and subsequently transitions to the urban class, or (b) it is already classified as urban and remains persistently urban over time.

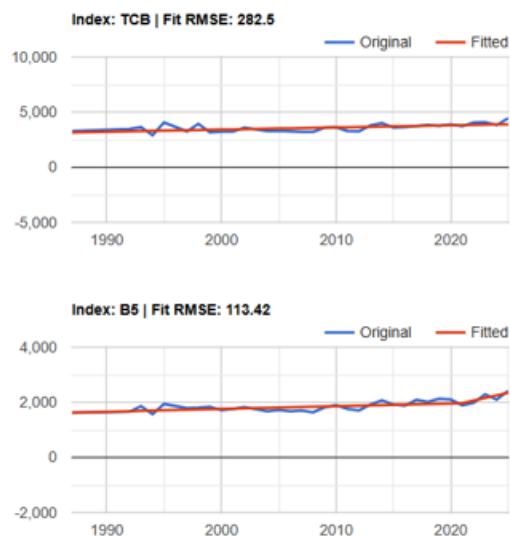
Figure 10 illustrates the temporal behavior of the Near-Infrared (NIR) spectral band and the Brightness component of the Tasseled Cap transformation, derived using the LandTrendr algorithm developed at Oregon State University.

LandTrendr Time Series Plots



a) Transition from non-urban to urban
Cuajimalpa (-99.2991,19.35526)

LandTrendr Time Series Plots



b) Persistence of the urban footprint
Mexico City Center(-99.1508,19.37967)

Figure 10. Possible temporal trajectories of a pixel over the analysis period. Generated using the Pixel Time Series Plotter function of the LandTrendr algorithm (Kennedy et al., 2018).

To refine the annual urban classification time series by removing year-to-year classification noise and enforcing temporal consistency in the urbanization trajectory of each pixel, a sequence of four consecutive temporal filters was applied, with each filter operating on the output generated by the previous one. The first two filters are restrictive, removing pixels that do not satisfy predefined temporal-consistency criteria, whereas the last two are additive, incorporating pixels that meet specific temporal-persistence conditions. Consequently, the workflow progressively transitions from refining to completing the urban classification time series (Figure 11).

Temporal Filter 1 first merges the spatially filtered classifications from the country's two processing regions (Northern and Southern Mexico) into a continuous national mosaic. It then applies a moving-window temporal consistency rule, whereby a pixel is retained as urban only if its urban classification is confirmed in at least 2 of the 3 nearest years (for the first and last two years of the time series) or in at least 3 of the 5 years within a centered moving window (for the intermediate years). This procedure removes isolated one-year urban detections that are likely attributable to classification noise.

Temporal Filter 2 further refines temporal consistency using broader, predominantly forward-looking windows. Across most of the time series, a pixel is retained as urban only if its classification is confirmed in at least 2 of the following 4 years. For the penultimate years, a mixed forward- and backward-looking window is applied, whereas for the final year, where no subsequent observations are available to confirm the classification, a more permissive retrospective rule is used, requiring confirmation in at least 1 of the preceding 3 years.

Temporal Filter 3 adopts a different strategy. Rather than removing pixels, it fills single-year temporal gaps by assuming that if a pixel was classified as urban in both the preceding and following years, it was likely urban during the intervening year, even if the classifier failed to detect it. For the final year of the time series (2025), where no subsequent observation is

available for confirmation, temporal continuity is assumed if the pixel was classified as urban in the immediately preceding year.

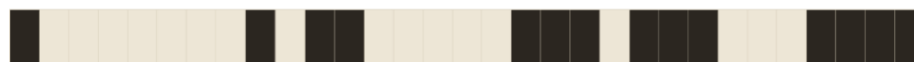
Finally, Temporal Filter 4 applies a cumulative consolidation based on the principle of “once urban, always urban.” For each year, the urban classification is updated using the historical maximum across all preceding years, producing a strictly non-decreasing time series and thereby eliminating spurious reversals in which pixels transition from urban to non-urban. To prevent an erroneous urban classification in the initial year (1985) from propagating throughout the entire time series, the baseline year is first validated before the cumulative rule is applied.

Sequential temporal filtering of one pixel's trajectory

Urban / non-urban series, Mexico, 1995–2025. Four simultaneous problems, resolved one filter at a time.

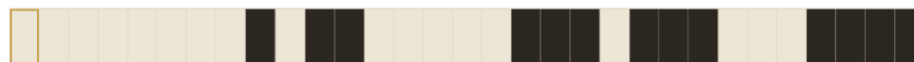
Raw classification

Isolated noise, a noise cluster, a one-year gap, and a multi-year reversion, all present at once.



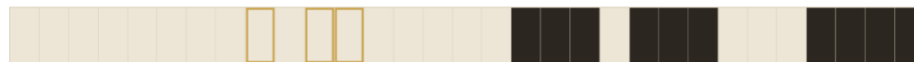
1 Filter 1

Removes the isolated single-year detection.



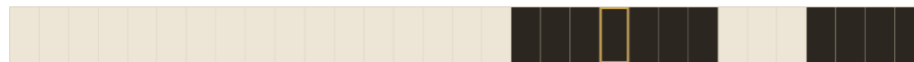
2 Filter 2

Wider window removes the remaining noise cluster.



3 Filter 3

Fills the single-year gap.



4 Filter 4

“Once urban, always urban” rule resolves the reversion.

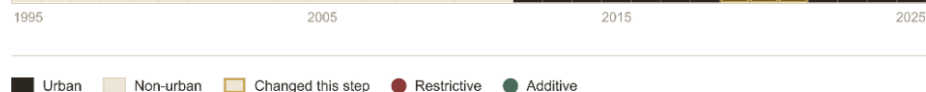


Figure 11. Illustration of the temporal filtering process applied to the trajectory of an individual pixel.

The net result of this sequence is an urban classification layer that is both spatially and temporally consistent (Figure 12). The last two filters, being additive—and, in the case of the fourth filter, cumulative and irreversible—tend to produce a slight overestimation of the urban extent. This behavior is intentional and reflects a methodological decision to minimize omission errors as much as possible. Nevertheless, the resulting maps underwent a final visual

quality control assessment by the mapping team before being accepted as the definitive products.

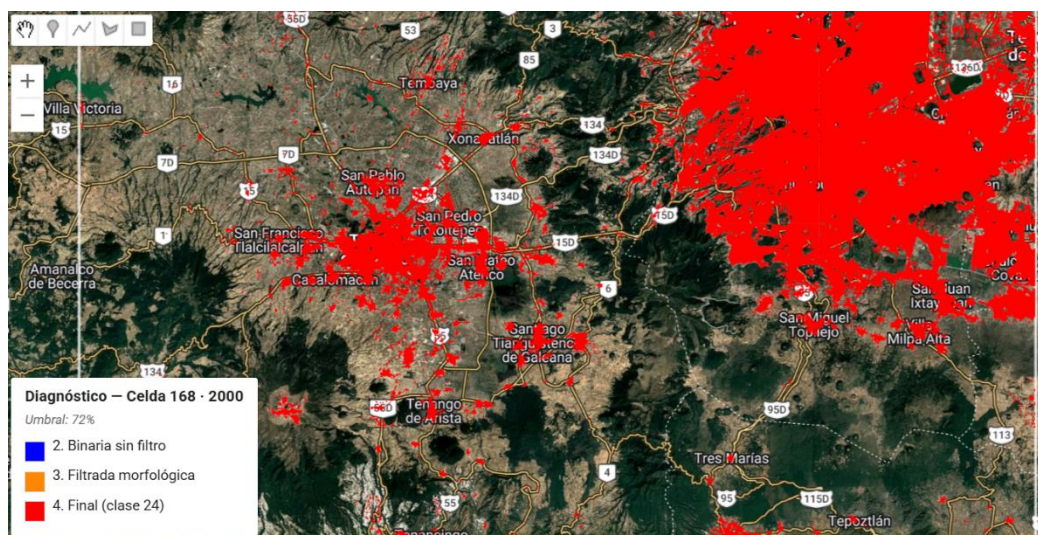


Figure 12. Final urban land-use map for grid cell 168 in 2025.

3.8 Spatial Mask Application

Finally, to reduce commission errors in areas located far from established urban settlements (e.g., highly reflective surfaces such as bare soil, coastlines, and fallow agricultural fields), a spatial mask was applied using the polygons of rural and urban localities provided by INEGI (2025).

As a result, the Urban Area class (Class 24) defined in the official MapBiomias Mexico legend was retained only within these polygons, whereas pixels outside their boundaries were reclassified as NoData (NA) to preserve the land-cover classification assigned by the main MapBiomias land-use and land-cover mapping workflow.

4. Quality Control

4.1 Collaborative assessment

To define the search radius and probability threshold, and to assess classification quality, a visual inspection of the final cartographic product was conducted for each grid cell and each processed year.

This visual assessment considered the mapping to be satisfactory when the urban class accurately preserved the shape, size, and spatial connectivity of urban areas (Figure 13).

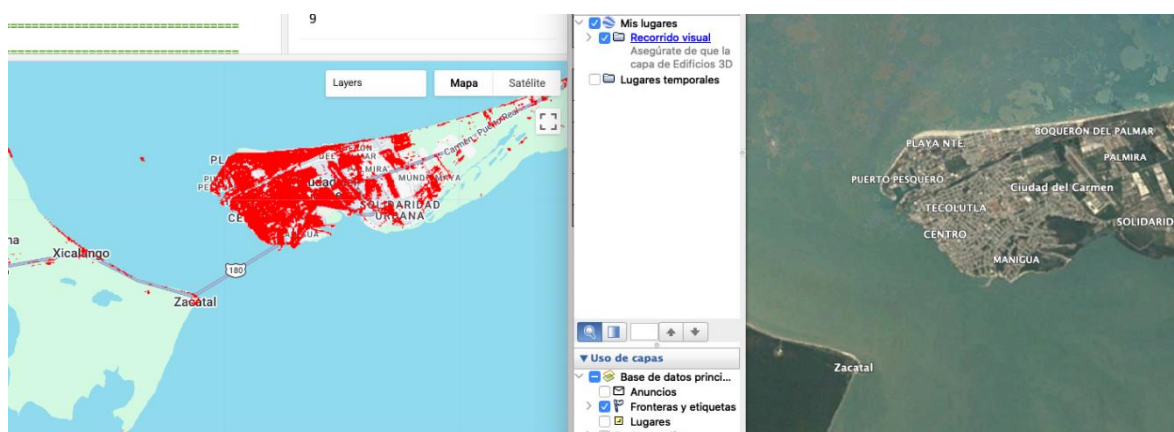


Figure 13. Example of accurate delineation of the shape and spatial extent of urban areas.

By contrast, when the urban extent was overestimated—primarily due to commission errors associated with agricultural land, gullies, coastlines, mines, gravel pits, uncultivated land, and other highly reflective surfaces—the classification parameters were adjusted to minimize this overestimation (Figure 14).

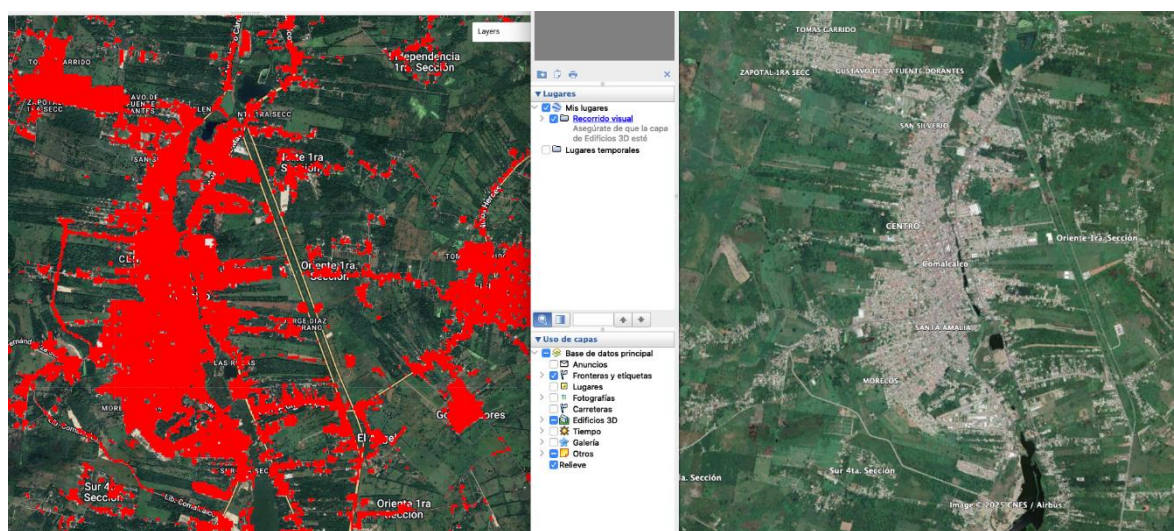


Figure 14. Example of urban extent overestimation.

The opposite case corresponds to omission errors in the urban classification. These were addressed with considerably greater caution because, unlike overestimation errors, they cannot be corrected through post-classification filtering (Figure 15). Instead, underestimation was mitigated by adjusting the classification probability thresholds and increasing the search radius used to collect training samples.



Figure 15. Underestimation of the urban extent creates gaps in the urban classification.

4.2 Limitations of the final product

The quality of the input image mosaic used for the Random Forest classification is a major determinant of the overall performance of the mapping workflow. Regardless of the number of training samples, spectral variables, and processing steps incorporated, a poor-quality input mosaic inevitably yields a lower-quality cartographic product, even after applying filtering procedures and parameter adjustments. This limitation was particularly evident during the 1985–1995 period, when the quality of the available image mosaics did not support satisfactory urban classifications for this first collection. Consequently, 1995 was selected as the baseline year because it provided the most reliable classification available. The resulting map was then propagated to the preceding years only in areas classified as NoData (ND), where insufficient information prevented a reliable classification. As a consequence, urban land-cover dynamics between 1985 and 1994 could not be reliably assessed.

Another limitation of the final product is that not all erroneously classified urban pixels associated with protected agricultural structures could be removed, even after applying the filtering procedures. Consequently, these features may still be misclassified as Urban Area (Class 24), particularly when located near established urban centers.

5. Lessons Learned

The first UTC collection demonstrated that Mexico's environmental diversity requires differentiated classification strategies. Initial experiments showed that a single combination of input variables was insufficient to adequately represent urban areas across the entire country. Consequently, adopting an operational regionalization scheme, together with variables specifically designed for arid environments, proved essential to improving model performance.

Likewise, using urban-class probability surfaces as an intermediate product proved more flexible than direct binary classification, as it enabled probability thresholds to be adjusted, classification results to be harmonized, and classification errors to be reduced prior to generating the final maps. The first collection also confirmed that temporal consistency is an essential component of representing urban expansion. Accordingly, applying temporal filters substantially improved the coherence of the historical time series.

Finally, the experience demonstrated that combining automated processing and collaborative expert assessment is an effective strategy for producing national-scale urban maps and supporting the continuous improvement of future collections.

References

- As-syakur, A. R., Adnyana, I. W. S., Arthana, I. W., & Nuarsa, I. W. (2012). Enhanced Built-Up and Bareness Index (EBBI) for mapping built-up and bare land in an urban area. *Remote Sensing*, 4(10), 2957-2970. <https://doi.org/10.3390/rs4102957>
- Belgiu, M. & Dragut, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing* 114, 24-31. <https://doi.org/10.1016/j.isprsjprs.2016.01.011>.
- Bhatti, S. S., & Tripathi, N. K. (2014). Built-up area extraction using Landsat 8 OLI imagery. *GIScience & Remote Sensing*, 51(4), 445-467.
- Bouhennache, R., Bouden, T., Taleb-Ahmed, A., & Chaddad, A. (2019). A new spectral index for the extraction of built-up land features from Landsat 8 satellite imagery. *Geocarto International*, 34(14), 1531-1551. <https://doi.org/10.1080/10106049.2018.1497094>
- Couturier, S., Huerta Luna, R., Osorno Covarrubias, J., Flores Rangel, A., Kuffer, M., Sliuzas, R., Camacho Cervantes, F.M., & López Martínez, R. (2020). "Monitoring the land consumption rate of urban growth from the Urban Footprint of Mexico, an online national cartographic platform". *Terra Digitalis: international journal of interactive maps and geovisualization*, 4 (1), 1-8. <https://doi.org/10.22201/igg.25940694e.2025.1.119>
- Couturier, S., Ricárdez, M., Osorno Covarrubias, J., & López-Martínez, R. (2011). Morpho-spatial extraction of urban nuclei in diffusely urbanized metropolitan areas. *Landscape and urban planning*, 101(4), 338-348. <https://doi.org/10.1016/j.landurbplan.2011.02.039>
- Estoque, R. C., & Murayama, Y. (2015). Classification and change detection of built-up lands from Landsat-7 ETM+ and Landsat-8 OLI/TIRS imageries. *Remote Sensing*, 7(3), 2630-2653. <https://doi.org/10.3390/rs70302630>
- Gao, B. C. (1996). NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, 58(3), 257-266. [https://doi.org/10.1016/S0034-4257\(96\)00067-3](https://doi.org/10.1016/S0034-4257(96)00067-3)
- Haralick, R. M., Shanmugam, K., & Dinstein, I. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3(6), 610-621. <https://doi.org/10.1109/TSMC.1973.4309314>
- Harrak, A., Bouchta, A., El Mahradi, B., El Ouadih, A., Benaddou, D., Ait Lhaj, L., & El Youssi, M. (2025). Assessment of spectral indices for urban built-up area extraction in semi-arid environments using Landsat 8 imagery. *Urban Science*, 9(3), 104. <https://doi.org/10.3390/urbansci9030104>
- Huete, A. R. (1988). A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, 25(3), 295-309. [https://doi.org/10.1016/0034-4257\(88\)90106-X](https://doi.org/10.1016/0034-4257(88)90106-X)
- 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(1-2), 195-213. [https://doi.org/10.1016/S0034-4257\(02\)00096-2](https://doi.org/10.1016/S0034-4257(02)00096-2)
- INEGI (2025). *Marco Geoestadístico Nacional*, 2025. <https://www.inegi.org.mx/app/biblioteca/ficha.html?upc=794551163061>
- INEGI. (2023). Conjunto de datos vectoriales de la Malla para la publicación de información estadística y geográfica. Malla Gesoestadística. <https://www.inegi.org.mx/infraestructura/malla/#descargas>

- Jiang, Z., Huete, A. R., Didan, K., & Miura, T. (2008). Development of a two-band enhanced vegetation index without a blue band. *Remote Sensing of Environment*, 112(10), 3833-3845. <https://doi.org/10.1016/j.rse.2008.06.006>
- Kawamura, M., Jayamanna, S., & Tsujiko, Y. (1996). Relation between social and environmental conditions in Colombo Sri Lanka and the urban index estimated by satellite remote sensing data. *Proceedings of the 17th Asian Conference on Remote Sensing*, Sri Lanka.
- Kennedy, R. E., Yang, Z., Gorelick, N., Braaten, J., Cavalcante, L., Cohen, W. B., & Healey, S. (2018). Implementation of the LandTrendr Algorithm on Google Earth Engine. *Remote Sensing*, 10(5), 691.
- Key, C. H., & Benson, N. C. (2006). Landscape assessment: Ground measure of severity, the Composite Burn Index, and remote sensing of severity, the Normalized Burn Ratio. In *FIREMON: Fire Effects Monitoring and Inventory System*. USDA Forest Service.
- Lin, Y., & Wang, Z. (2011). A new index for mapping built-up and impervious surfaces from remotely sensed imagery. *Proceedings of the International Conference on Remote Sensing and GIS*, Shanghai, China.
- Martin, M. J., Blesius, L., & Liu, X. (2026). Classifying Desert Urban Landscapes with Multi-Spectral Analysis Using Landsat 8-9 Imagery. *Remote Sensing*, 18(8), 1241. <https://doi.org/10.3390/rs18081241>
- McFeeters, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, 17(7), 1425-1432. <https://doi.org/10.1080/01431169608948714>
- Nguyen, K. A., Liou, Y. A., Tran, H. P., Hoang, P. P., & Nguyen, T. H. (2021). Soil and built-up area discrimination using modified bare soil indices from Landsat imagery. *Remote Sensing Applications: Society and Environment*, 21, 100459.
- Pal, M. (2005). Random forest classifier for remote sensing classification. *International Journal of Remote Sensing*, 26(1), 217-222. <https://doi.org/10.1080/01431160412331269698>
- Pesaresi, M., Politis, P. (2023): GHS-BUILT-S R2023A. Red de superficie construida de GHS, derivada de la composición de Sentinel-2 y Landsat, multitemporal (1975-2030). European Commission, Joint Research Centre (JRC). PID: <http://data.europa.eu/89h/9f06f36f-4b11-47ec-abb0-4f8b7b1d72ea> doi:10.2905/9F06F36F-4B11-47EC-ABB0-4F8B7B1D72EA
- Pinho, C. M. D., Rennó, C. D., Kux, H. J. H., & Novack, T. (2011). Urban areas classification from remote sensing images using texture and spectral indices. *Revista Brasileira de Cartografia*, 63(3), 421-431.
- Rasul, A., Balzter, H., Ibrahim, G. R. F., Hameed, H. M., Wheeler, J., Adamu, B., Ibrahim, S., & Najmaddin, P. M. (2018). Applying Built-Up and Bare-Soil Indices from Landsat 8 to Cities in Dry Climates. *Land*, 7(3), 81. <https://doi.org/10.3390/land7030081>
- Rikimaru, A., Roy, P. S., & Miyatake, S. (1996). Tropical forest cover density mapping. *Tropical Ecology*, 37(1), 39-47.
- Rodriguez-Galiano, F., B. Ghimire, J. Rogan, M. Chica-Olmo, J.P. Rigol-Sanchez, (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing* 67: 93-104. <https://doi.org/10.1016/j.isprsjprs.2011.11.002>.

Rouse, J. W., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. Proceedings of the Third Earth Resources Technology Satellite-1 Symposium, NASA SP-351, 3010-3017. <https://ntrs.nasa.gov/citations/19740022614>

Santos, A. R., Marcato Junior, J., Silva, T. S. F., et al. (2022). Evaluation of spectral indices for urban mapping using Landsat imagery. Remote Sensing, 14(3), 704. <https://doi.org/10.3390/rs14030704>

Souza, C. M., Roberts, D. A., & Cochrane, M. A. (2005). Combining spectral and spatial information to map canopy damage from selective logging and forest fires. Remote Sensing of Environment, 98(2-3), 329-343. <https://doi.org/10.1016/j.rse.2005.07.013>

Waqar, M. M., Mirza, J. F., Mumtaz, R., & Hussain, E. (2012). Development of new indices for extraction of built-up area and bare soil from Landsat data. Egyptian Journal of Remote Sensing and Space Sciences, 15(1), 59-68.

Xu, H. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. International Journal of Remote Sensing, 27(14), 3025-3033. <https://doi.org/10.1080/01431160600589179>

Xu, H. (2008). A new index for delineating built-up land features in satellite imagery. International Journal of Remote Sensing, 29(14), 4269-4276. <https://doi.org/10.1080/01431160802039957>

Zha, Y., Gao, J., & Ni, S. (2003). Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. International Journal of Remote Sensing, 24(3), 583-594. <https://doi.org/10.1080/01431160304987>