Appendix - Urban Infrastructure

Cross-cutting theme

Collection 3.0

Responsible team

Cristina Cabrera
Cristina Aguilar
Katherine Terán
Stephany Cevallos

Developer

Fabricio Garcés

October 2025

Index

List of Figures

List of Tables

1. Introduction

The MapBiomas Ecuador Collection 3.0 incorporates urban infrastructure mapping as a binary, cross-cutting layer covering the entire territory of Ecuador for the period between 1985 and 2024. This land cover class consists of human settlement areas associated with large and small urban centers (towns) featuring built-environment infrastructure such as road and railway networks and associated lands, as well as other artificial areas such as hydrocarbon extraction sites, hydroelectric plants, military bases, airports, seaports, and non-conventional airstrips in rural zones. It also includes peripheral areas undergoing gradual urbanization for residential and/or industrial purposes.

Other global initiatives have developed urban land mapping using various remote sensing techniques and data sources (Schneider et al., 2010). Global products such as FROM-GLC (Gong et al., 2013) and GlobeLand30 (Chen et al., 2015) are available; however, most of this information covers a limited range of years, making it difficult to monitor urban dynamics over extended historical periods.

The Landsat data series has enhanced the availability of annual information for MapBiomas Ecuador at the regional level, serving as a starting point for this collection through automated cloud-based processing in Google Earth Engine. The methodological foundations were applied for the multitemporal mapping of Urban Infrastructure (ID:24) over the past 39 years.

2. Construction and definition of Landsat mosaics

The first step in generating land cover and land use maps is the construction of annual mosaics created from satellite images captured by the Landsat project. For this purpose, the MapBiomas Ecuador initiative employed two methods:

1. For the area within the RAISG boundary: the mosaics were processed based on a uniform grid derived from the International Map of the World at the Millionth Scale (1:250,000 scale), where each tile covers an area of 1°30' longitude by 1° latitude (see fig. 1).

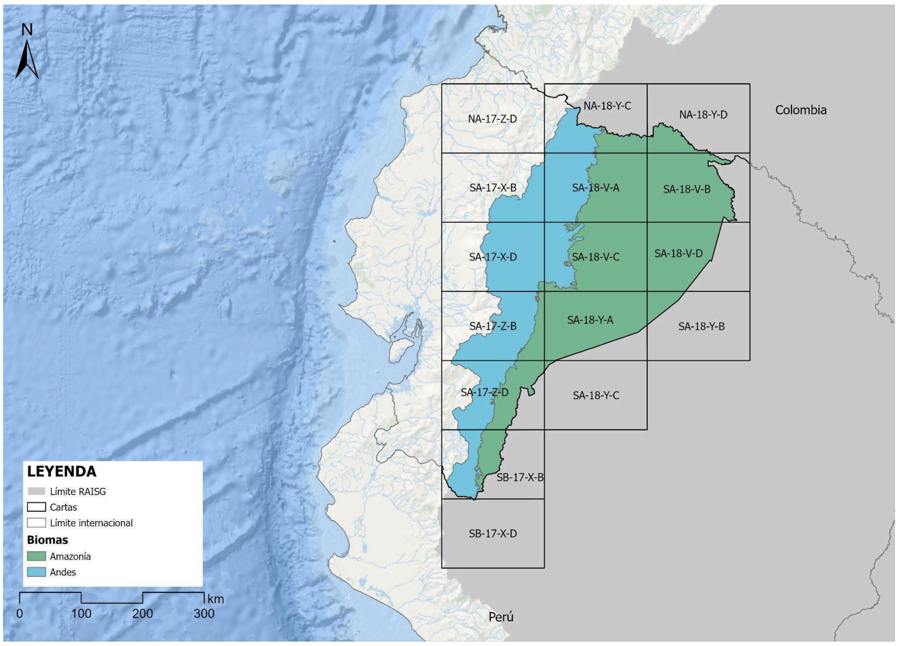


Figure 1: Distribution of the International Map of the World grid at a scale of one to one million, used for mosaic processing in the cross-cutting mining map. Source: EcoCiencia.

2. For the area outside the RAISG boundary: the reference unit for the mosaic construction process was the path and row grid. The entire Ecuadorian territory, both continental and insular, is covered by 17 path and row combinations.

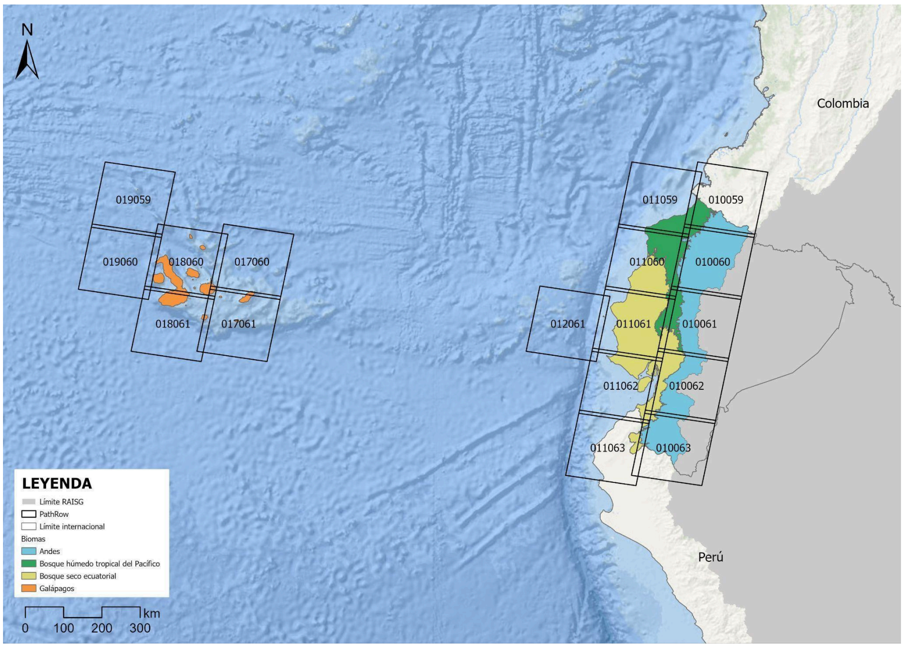


Figure 2: Distribution of Path/Row combinations used for mosaic processing in the cross-cutting mining map. Source: EcoCiencia.

In each of the map sheets and path/row combinations, annual Landsat mosaics are constructed with the objective of generating cloud-free periodic images that enable classification and change detection. For further details, refer to Section 4 of the document: Algorithm Theoretical Basis Document (ATBD) of Ecuador – Collection 3.0.

3. Methodological summary

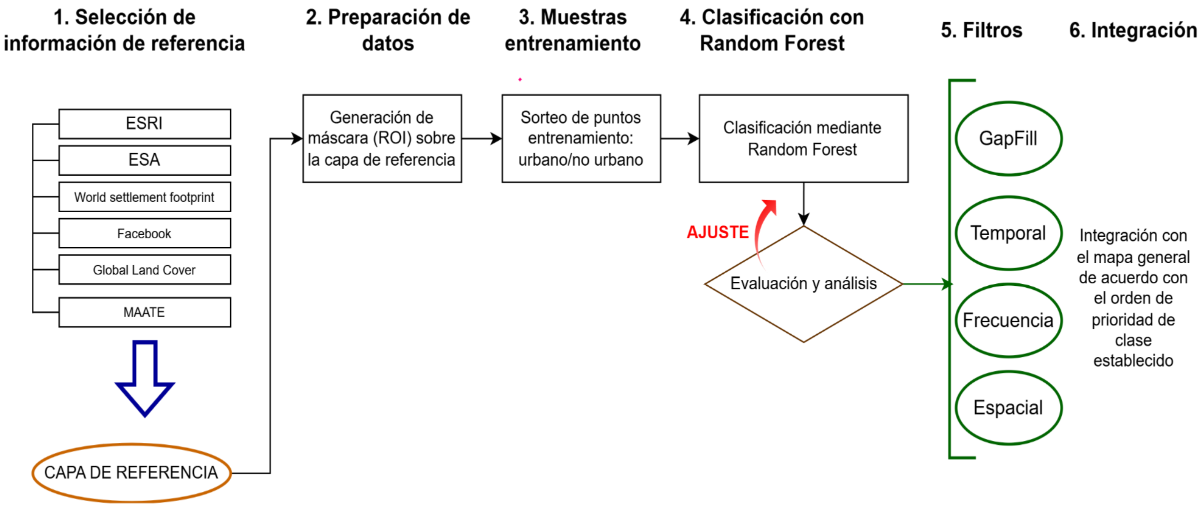


Figure 3. Methodological flow of Urban Infrastructure classification, Collection 3.0. Source: EcoCiencia

4. Classification variables

The next step consisted of calculating the spectral indices and fractions for each satellite observation (feature space), which include Landsat spectral bands, spectral indices, fractional and texture information derived from the previous indices and fractions. Static variables were also used—HAND, shademask2, slppost, altitude, slope, latitude, and longitude—as support for segmenting classes that are spectrally very similar but can be distinguished by their topographic conditions.

Table 1. Feature space: Variables for detecting Urban Infrastructure surfaces with the associated reducer.

Variable	Description	Reductor
blue	Landsat BLUE band	Median
green	Landsat GREEN band	Median
red	Landsat RED band	Median
nir	Landsat NIR band	Median
swir1		Median
swir2	Landsat SWIR2 band	Median
soil	Soil Fraction	Median
snow	Snow Fraction	Median
cloud	Cloud Fraction	Median
slope	Slope	Static topographic variable
ndvi	Normalized Difference Vegetation Index	Median
dwi_mcfeeters	Normalized Difference Water Index (Mcfeeters)	Median
ndbi	Normalized Difference Built-up	Median
nuaci	Normalized Urban Areas Composite Index	Median

Finally, a representative mosaic was created for each year, consisting of a total of 151 bands, based on the calculation of statistical reducers to generate the values of each pixel. These reducers correspond to:

Table 2. Statistical reducers used in the composition of Urban Infrastructure images, Collection 4.0.

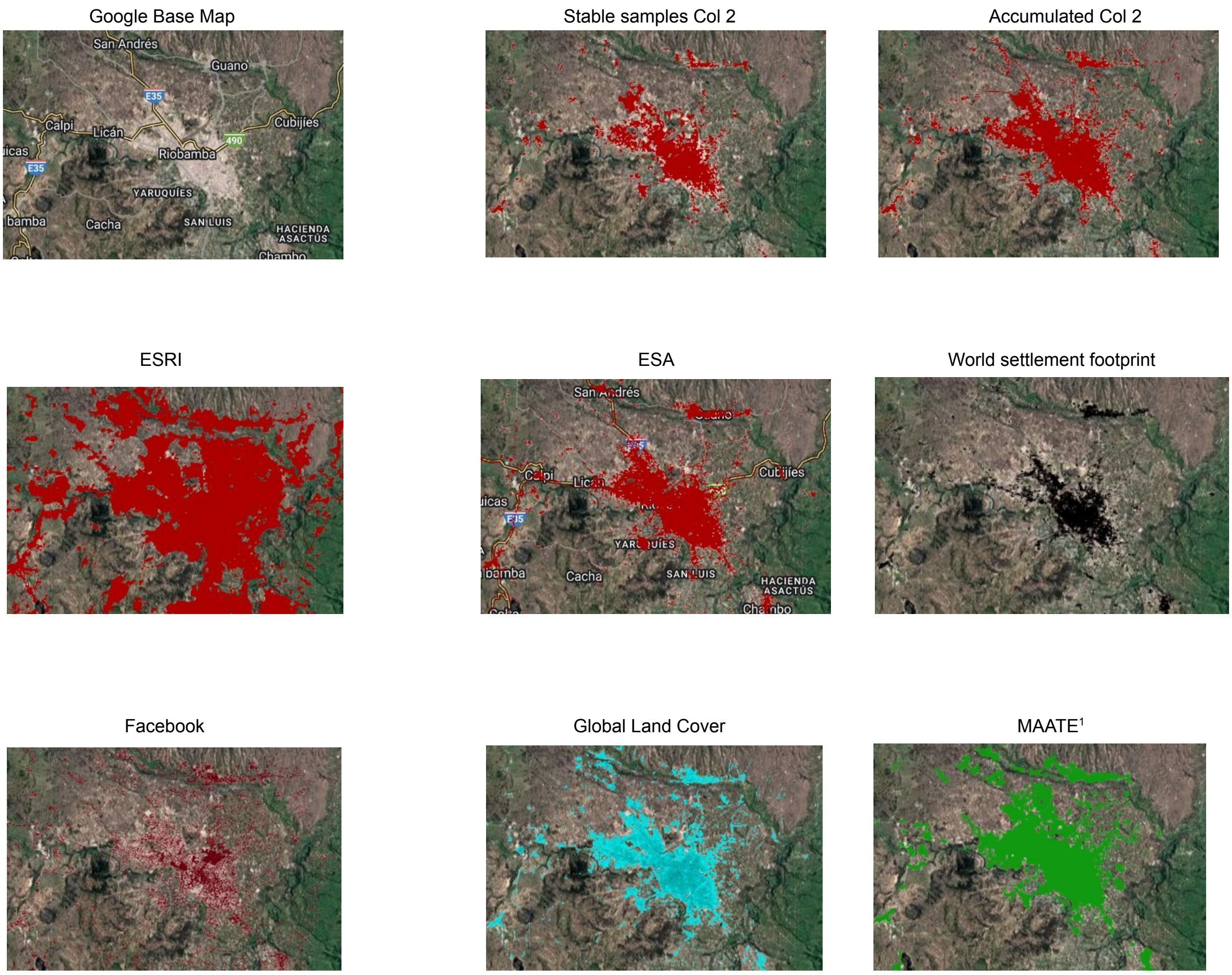
Reductor	Description
Median	Median of all available values in the annual mosaic for that location (pixel).
Median of the dry season	Calculation of the statistical median applied to the pixels of the 25th quartile (with the values) of NDVI (proxy for the dry season).
Median of the wet season	Calculation of the statistical median applied to the pixels of the 75th quartile (with the highest values) of NDVI (proxy for the rainy season).
Amplitude	Range of variation among all available pixels in the annual mosaic.
Standard deviation	Standard deviation of the values of all available pixels in the annual mosaic for a given location.
Minimum	Lowest value of all available pixels in the annual mosaic at a given location.
Maximum	Highest value of all available pixels in the annual mosaic at a given location.
Minimum of the dry period	Calculation of the lowest value of all available pixels from the images in the quartile with the lowest NDVI values (proxy for the dry season).
Minimum of the wet period	Calculation of the lowest value of all available pixels from the images in the quartile with the highest NDVI values (proxy for the rainy season).
Maximum of the dry period	Calculation of the highest value of all available pixels from the images in the quartile with the lowest NDVI values (proxy for the dry season).
Maximum of the wet period	Calculation of the highest value of all available pixels from the images in the quartile with the highest NDVI values (proxy for the rainy season).
QMO of the dry period	The highest value of the EVI2 band during the dry season. QMO of the rainy period. The highest value of the EVI2 band during the wet season.

5. Classification

5.1. Reference information

For Ecuador, the Random Forest algorithm used thirteen reference sources (see Table 3), compiled from projects containing geographic information with global and national datasets across different time periods. These also included accumulated and stable samples from previous MapBiomas Ecuador collections. The information was filtered based on the presence of urban infrastructure surfaces or urban conglomerates with attributes such as residential, commercial, service, industrial, and park areas, among others.

Table 3. Reference sources available for the MapBiomas Ecuador study area



[1] The reference used from MAATE (Ministry of Environment, Water and Ecological Transition) corresponds to the Land Cover and Land Use Map of Ecuador for the years 1990, 2000, 2008, 2014, 2016, 2018, 2020, and 2022. From this dataset, only the information related to "anthropic area" was extracted, which is disaggregated into populated areas (areas mainly occupied by housing and buildings intended for communities or public services) and infrastructure (civil works for transportation, communication, agro-industrial, and social purposes).

Table 4. Information used by Ecuador as reference in the urban infrastructure classification process.

Reference	MapBiomas	Global	Nacional
Ref1_stable_pixel_of_urban_ColAnt	●		
Ref2_accumulated_of_urban_ColAnt	●		
Ref3_ESRI_Built_Area		●	
Ref4_ESA_WorldCover_project_2020_Built_up		●	
Ref5_World Settlement Footprint 2015		●	
Ref6_World Settlement Footprint 2019		●	
Ref7_World Settlement Footprint Evolution 1985-2015		●	
Ref8_Mapa_densidad_población_alta_resolución_CIE SIN_Facebook		●	
Ref9_Global_land_cover_and_land_use_2019		●	
Ref10_MAATE_90 (1990)			●
Ref11_MAATE_00 (2000)			●
Ref12_MAATE_08 (2008)			●
Ref13_MAATE_14-22 (2014, 2016, 2018, 2020 y 2022)			●

Each reference was analyzed to determine which was the most suitable to apply in each classification region.

Amazon Basin: The Amazon Basin was divided into five regions. The references used in each region are detailed below.

Table 5. Information used by Ecuador as reference in the urban infrastructure classification process in the Amazon Basin.

URBAN INFRASTRUCTURE REGIONS OF THE AMAZON BASIN

	40101	40102	40201	40202	40601
Ref1_stable_pixel_of_urban_ColAnt		●			
Ref2_accumulated_of_urban_ColAnt	●		●	●	●
Ref4_ESA_WorldCover_project_2020_B uilt_up	●	●	●	●	●
Ref6_World Settlement Footprint 2019	●	●	●	●	●
Ref7_World Settlement Footprint		●	●		●

Pacific Basin: The Pacific Basin was divided into 13 regions. Unlike the Amazon Basin, this basin is characterized by areas that, either due to natural conditions or anthropogenic activities, exhibit low vegetation cover density. This created confusion in differentiating between infrastructure and other land cover and land use classes. Therefore, an approach based on selecting references by period was implemented to mitigate possible overestimations in data interpretation. The references used by region and period are detailed below.

Table 6. Information used by Ecuador as reference in the urban infrastructure classification process in the Pacific Basin regions.

URBAN INFRASTRUCTURE REGIONS OF THE PACIFIC BASIN

	40907	40905	40908	40901	40902	40903	40904	40906	40603	40604	40605	40606	40607
Ref7_World Settlemer Footprint Evolution 1985-2015	●	●	●	●	●	●	●	●	●	●	●	●	●
Ref10_MAATE_90	●	●	●	●	●	●	●	●	●	●	●	●	●
Ref11_MAATE_00	●	●	●	●	●	●	●	●	●	●	●	●	●
Ref12_MAATE_08	●	●	●	●	●	●	●	●	●	●	●	●	●
Ref13_MAATE_14-22	●	●	●	●	●	●	●	●	●	●	●	●	●

Table 7. Information used by Ecuador as reference in the urban infrastructure classification process for the periods defined for each region of the Pacific Basin.

	Ref7_World Settlement Footprint Evolution 1985-2015	Ref10_MAATE_90	Ref11_MAATE_00	Ref12_MAATE_08	Ref13_MAATE_14-22
1985 - 1990	●	●			
1991 - 2000	●		●		
2001 - 2008	●			●	
2009 - 2023	●				●
2024	●				●

2.1. Training samples (urban/non-urban)

The training samples were randomly selected within the references (see Tables 6 and 7). The number of points to be sampled for each class was defined accordingly. Two types of points were sampled: Urban (ID 24), which strictly contains pixels belonging to urban infrastructure areas, and Non-Urban (ID 27), which includes all areas that do not correspond to urban infrastructure—i.e., all other land cover classes (see Fig. 4).

Depending on the region, the number of references to be used was determined. Out of the 13 available references, it was established how many would be employed. A point is considered valid as a training sample only when all selected references agree in classifying it as urban or non-urban.

The training samples are used to generate the classification of pixels into urban and non-urban areas. Areas outside the region of interest (reference boundaries) are not classified within this cross-cutting theme.

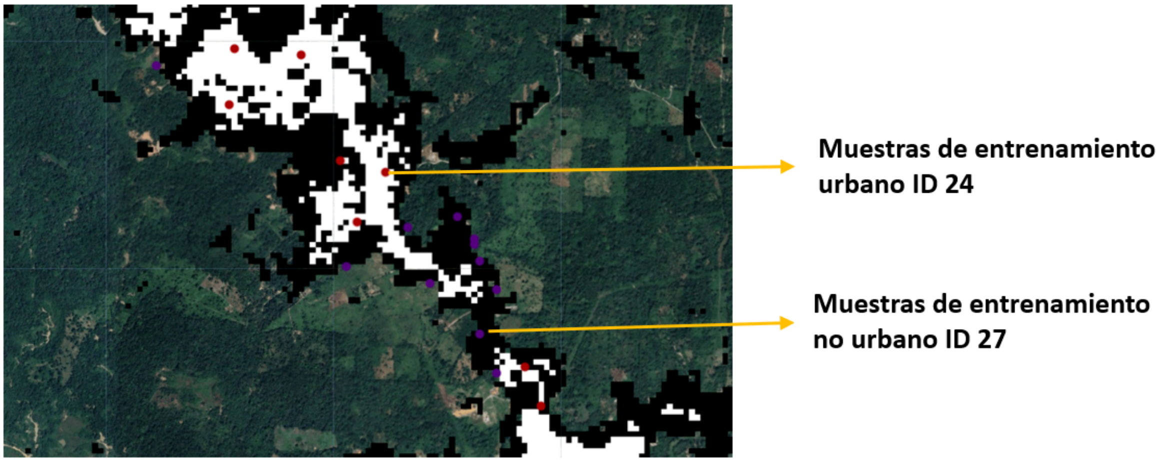


Figure 4. Urban and Non-Urban training samples. Source: EcoCiencia

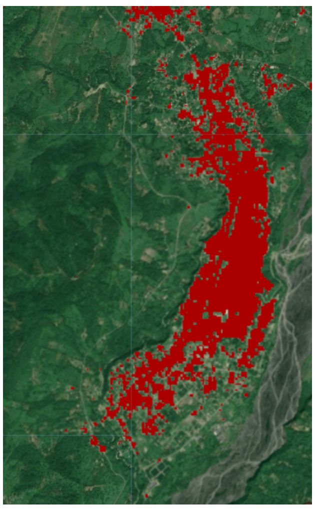
2.2. Classification using the Random Forest algorithm

The objective of the classification was to produce a binary layer distinguishing between urban infrastructure and non-urban areas, considering only two possible values for each pixel (1: Presence of Urban Infrastructure, 2: Absence of Urban Infrastructure). The classification, based on training samples over the Landsat mosaics, was performed using the Random Forest (RF) classification algorithm.

The Random Forest machine learning algorithm (Breiman, 2001) was implemented on the Google Earth Engine platform. Training samples were generated and classified according to the behavior of the selected indices, supported by regional reference maps such as OpenStreetMap and national and global reference data.

The classification was carried out over the 13 classification regions for the cross-cutting theme of Collection 3.0, focusing on areas with the presence of urban infrastructure according to known reference sources and high-resolution imagery.

Table 8. Urban Infrastructure Detection Collection 3.0 (A) - Landsat 2024 Mosaic (RGB 654), (B) - Urban Surface Classification 2024, (C) - High-Resolution Google Earth Image. (Quito - Ecuador).

		
A	B	C

6. Post-classification

Due to the pixel-based nature of the classification method and the extended temporal series, a post-classification process was implemented. This included the application of temporal and morphological filters (gap-fill, spatial, and frequency filters) on the binary discrimination of Urban Infrastructure. The objective was to reduce temporal inconsistencies, minimize classification noise below the minimum mapping unit, and fill data gaps.

6.1.Gap-fill filter

The sequence of filters for urban infrastructure coverage, as a cross-cutting theme, begins with the filling of data gaps — the Gap-fill filter. Considering the spatial extent covered by MapBiomas Ecuador and the atmospheric and climatic conditions affecting urban areas, annual mosaics often contain pixels with missing or unobserved satellite data ("gaps"). This filter reduces these residual gaps by replacing them with the temporally closest available value.

When a "future" pixel lacks data, the Gap-fill filter assigns it the value from the nearest available year. The filter reviews the entire time series, first filling gaps in a "backward-to-forward" sweep, where missing pixels are filled using data from the closest preceding years. If gaps remain, they are then filled using data from the closest subsequent year.

For each pixel whose value has been completed using this filter, the change is recorded in a metadata file that tracks the pixel's history. Multiple years can be used to fill gaps. Therefore, data gaps persist only when a pixel has been permanently classified as missing throughout the entire temporal series.

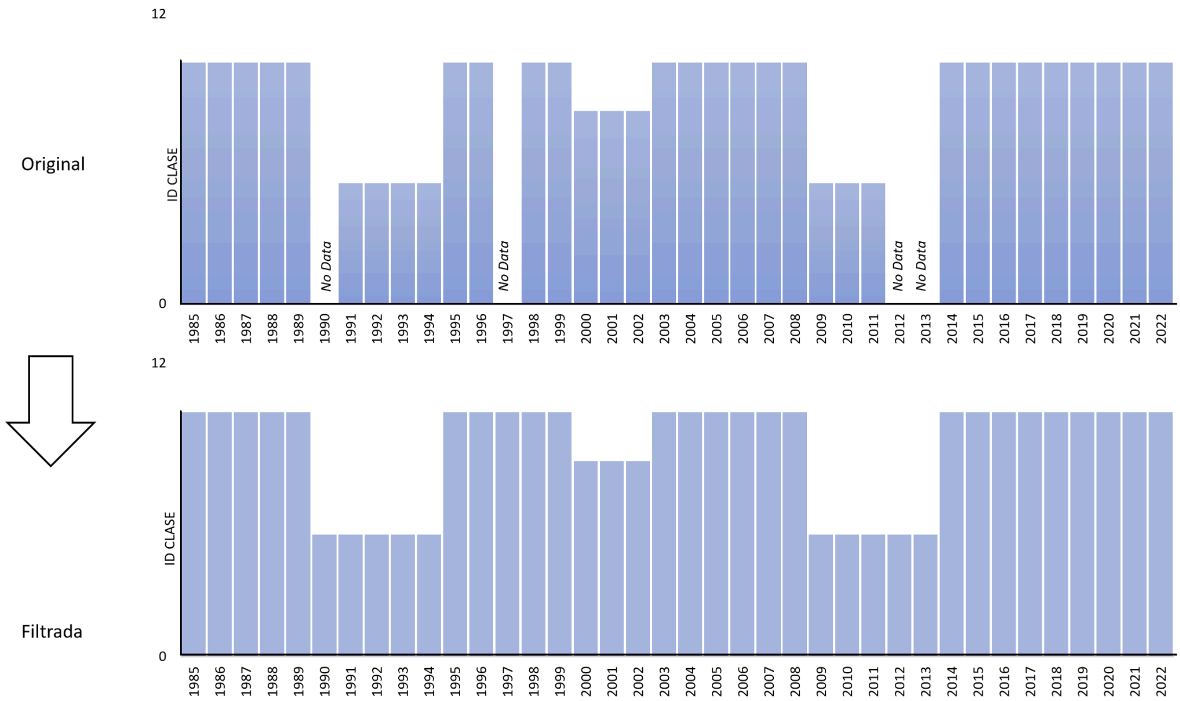


Figure 5: Schematic representation of the Gap-fill filter functionality. Source: MapBiomas

6.2. Temporal filter

A temporal filter was applied using moving windows of 3, 4, and 5 years across the entire time series. This filter inspects the value of each classified pixel in relation to its values in temporally consecutive classifications. It employs a unidirectional moving window that considers sequences of 3 to 5 years to identify non-permitted temporal transitions. The temporal filter is applied to every pixel for all years in the collection. Depending on which year the rule modifies, three types of rules are used:

General Rules (GR). Applied to pixels in intermediate-year positions within 3- to 5-year sequences. This rule is only applied in cases of temporal inconsistency — for example, when consecutive years have identical values except for the pixel in the central position. In such cases, the filter modifies the value of the central pixel to make it consistent with its preceding and succeeding years. In 3-year sequences, there is only one central position (the intermediate year), while in 4- or 5-year sequences, there are two or three possible central positions. This rule modifies classifications from 1986 to 2024.

First-Year Rules (FYR). Applied exclusively to the first year of the time series. This rule modifies the classification values for the year 1985.

Last-Year Rules (LYR). Applied to the last year of the classification. This rule modifies the classification values for the year 2024.

In this way, temporal filters reduce data gaps and temporal inconsistencies, as well as unrealistic or non-permitted classification changes.

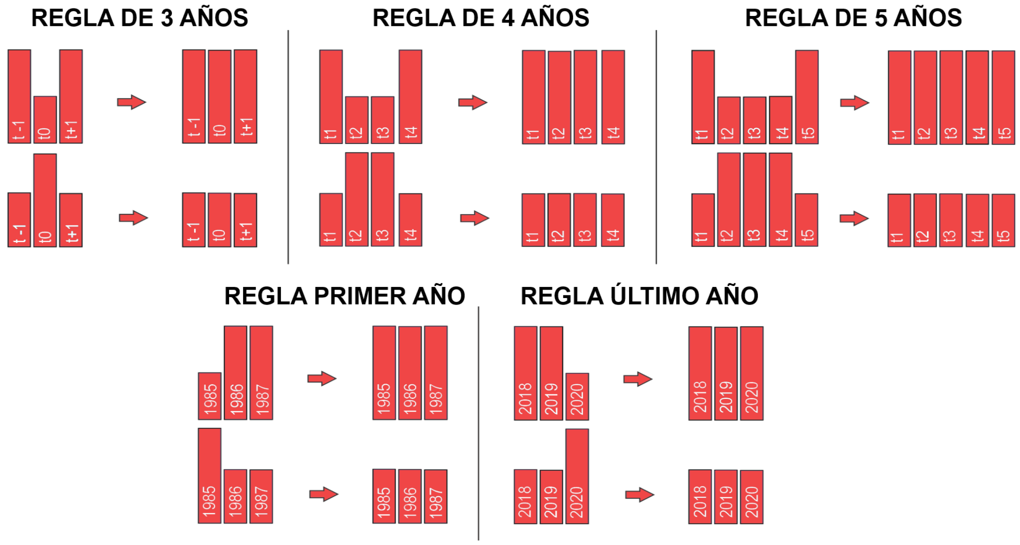


Figure 6: Effect of the temporal filter. Source: MapBiomas

6.3. Adjusted frequency filter

This filter considers the occurrence of the class throughout the time series to normalize its historical trajectory. Given that urban infrastructure typically exhibits an incremental and expanding trend over time, two filters (first-year and last-year) were applied to regulate the consistent increase of urban pixels and to prevent anomalous year-to-year fluctuations.

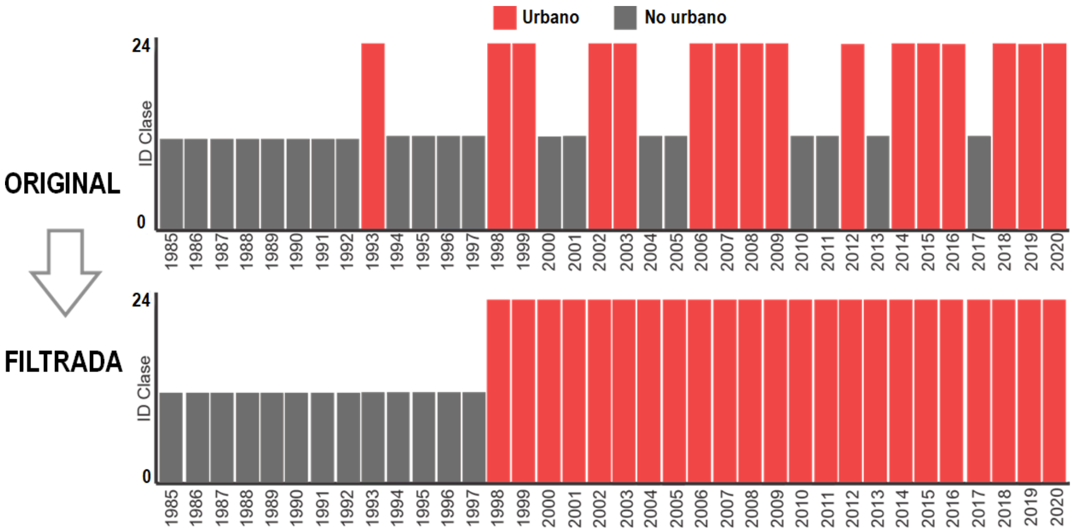


Figure 7: Schematic representation of the frequency filter functionality. Source: MapBiomas.

6.4. Spatial filter

The final filter applied in the post-classification sequence was the spatial filter, based on the "connectedPixelCount" function from Google Earth Engine. This function identifies connected (neighboring) pixels that share the same value by using a moving window. Only pixels that do not share a connection with a predefined number of identical neighboring pixels are considered isolated pixels.

In the case of MapBiomas Amazonia, the minimum mapping unit was defined as 0.5 ha (approximately 5 pixels). Therefore, at least five connected pixels were required to meet the minimum connection criterion.

This spatial filter smooths local variations by removing isolated or edge pixels smaller than 0.5 ha, thereby increasing the spatial consistency of the class.

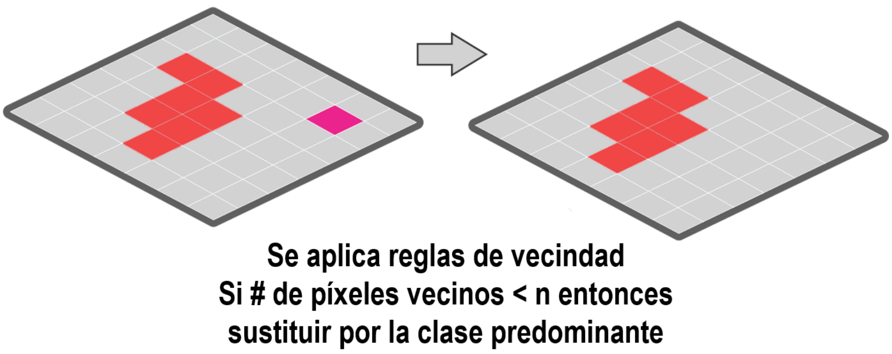


Figure 8: Schematic representation of the spatial filter functionality. Source: MapBiomas

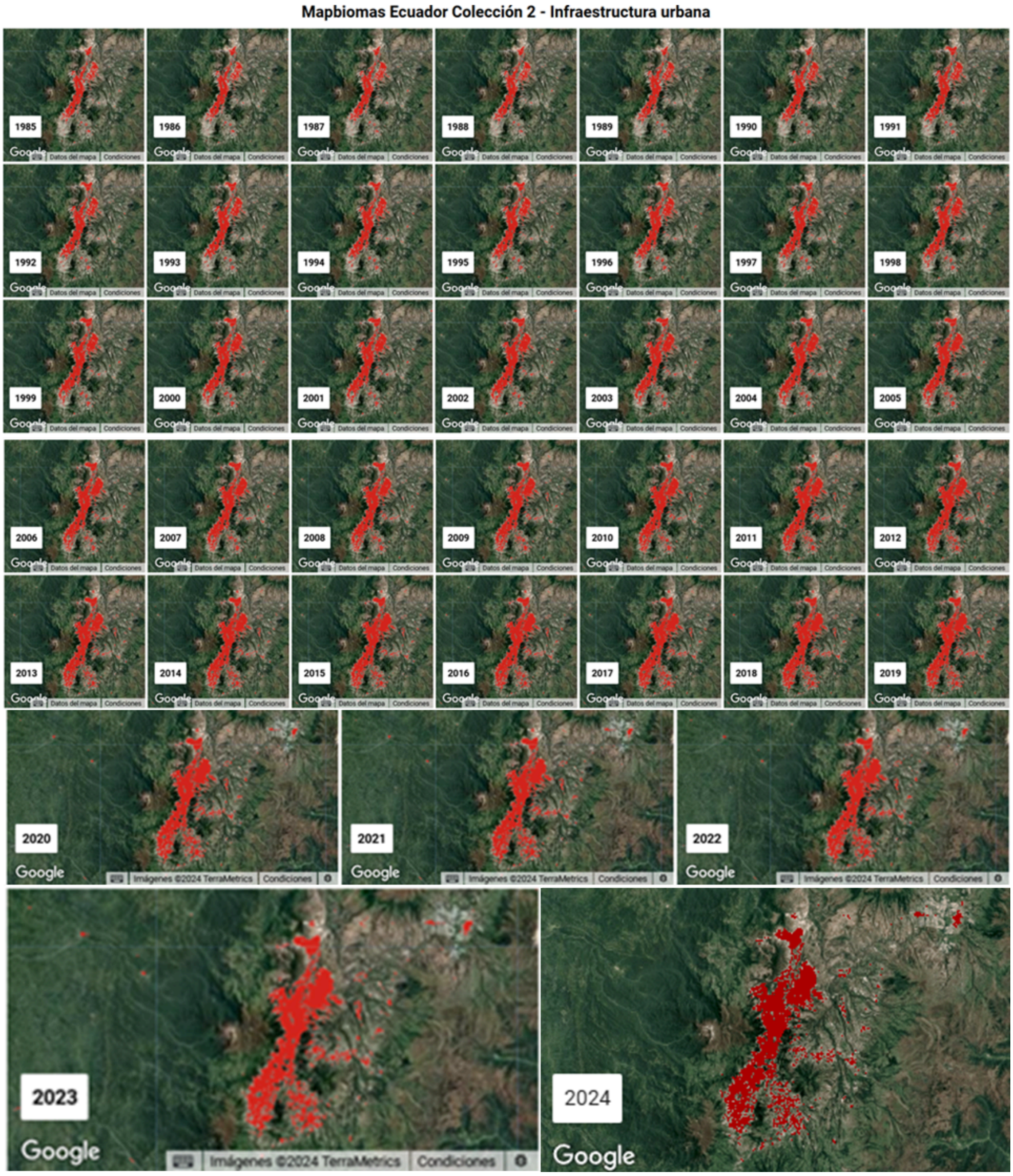


Figure 9. Multitemporal classification of Urban Infrastructure Collection 2.0 – 1985 to 2022 (case study: Quito, Ecuador).

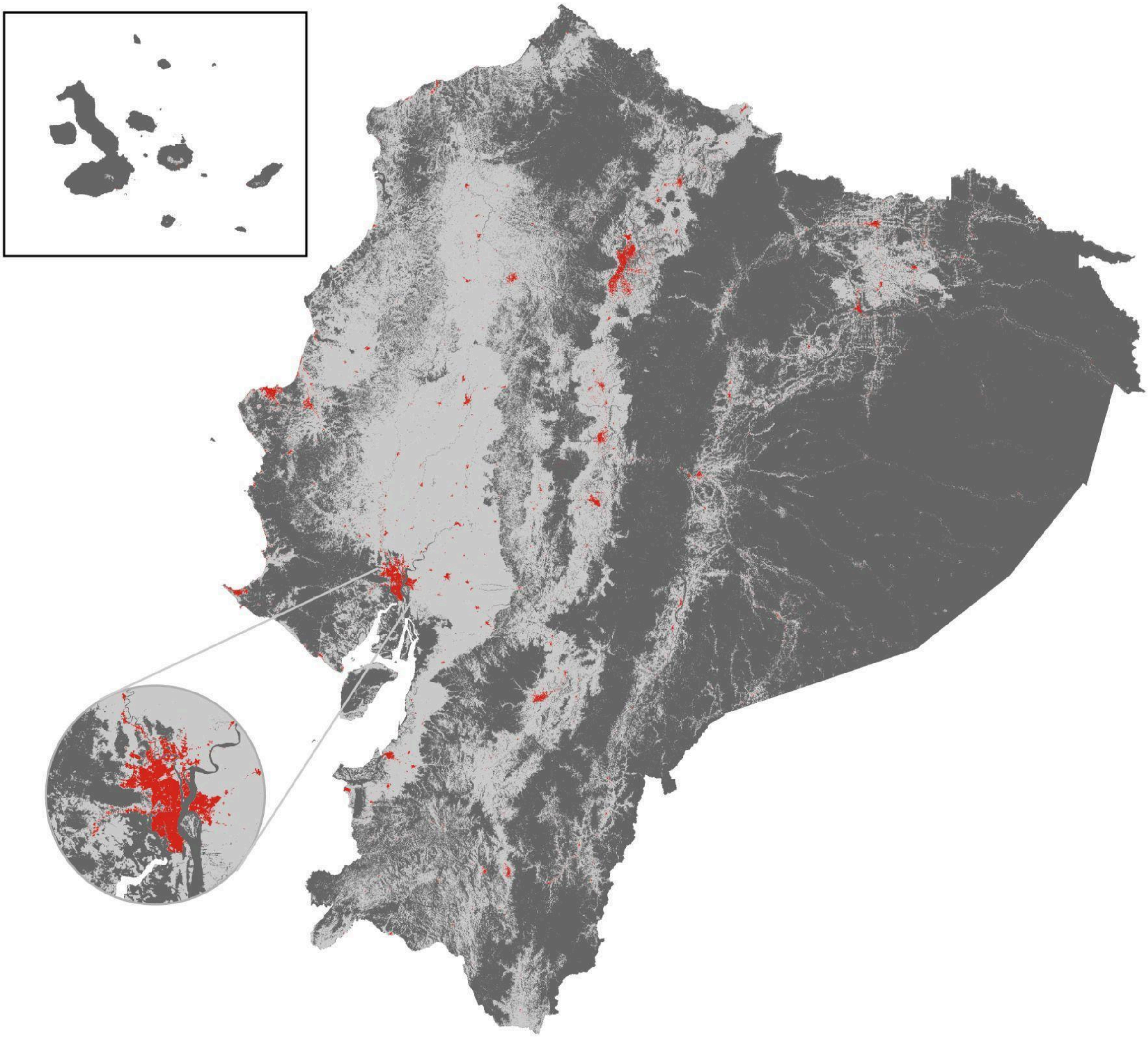


Figure 10: Urban infrastructure map 2023. Source: EcoCiencia

7. References

Breiman, L. (2001). Random Forest. Statistics Department University of California Berkeley. 1-30.

Chen, J., Chen J., Liao, A., Cao, X., Chen, L., Chen, X., He, C., Han, G., Peng, S. fi Lu, M. (2015). Global land cover mapping at 30 m resolution: a POK-based operational approach. ISPRS J. Photogramm. Remote Sens. 103, 7-27.

Gong, P., Wang, J., Yu, L., Zhao, Y., Liang, L., Niu, Z., Huang, X., Fu, H. fi Liu, S. (2013). Finer resolution observation and monitoring of global land cover: first mapping results with Landsat ™ and ETM+ data. Int. J. Remote Sens. 34, 2607-2654.

Olofsson, P.; Foody, G.M.; Herold, M.; Stehman, S.V.; Woodcock, C.E.; Wulder, M.A. (2014) Good Practices for Estimating Area and Assessing Accuracy of Land Change. Remote Sens. Environ. 148, 42–57, doi:10.1016/j.rse.2014.02.015.

Schneider, A., Friedl, M.A., Potere, D. (2010). Mapping global urban areas using MODIS 500-m data: new methods and datasets based on ‘urban ecoregions’. Remote Sens. Environ. 114, 1773-1746.

Souza, C., Monteiro, D., Shimbo, J. fi Azevedo, T. (2019). Apéndice Infraestructura Urbana. Versión 1. Recuperado el 2 de Julio de 2021:https://mapbiomas-br-site.s3.amazonaws.com/Infraestrutura_Urbana_Appendix_-_ATBD_Cole%C3%A7%C3%A3o_4_dez_2019.pdf

UCLG (2018). Amazonian Cities. Peer Learning about the sustainable use of terrestrial ecosystems #23. United Cities and Local Governments. Bolivia. 36 p.