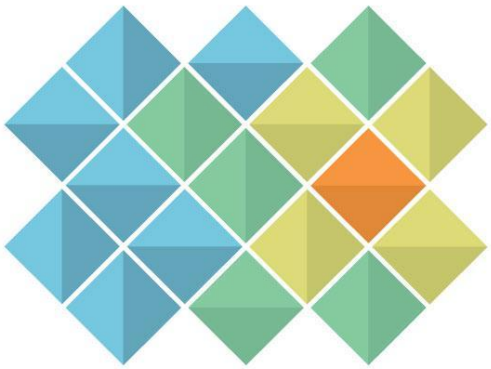


MAPBIOMAS

ARGENTINA

MapBiomas Argentina - Urban areas “Handbook” Algorithm Theoretical Basis Document (ATBD)

Collection 1

Version 1

Table of Contents

1. Introduction 2
2. MapBiomas Argentina - Urban areas 2
3. Methodological Description 3
3.1. Inputs preparation 4
3.1.1. Search Area 4
3.1.2. Mosaic and topographic maps 4
3.1.3. Samples 6
3.2. Processing 7
3.3.1 Classification with Random Forest 7
3.2.2. Spatial and temporal Filters 8
3.3. Validation 11
4. References 12

1. Introduction

This document is the explanations of all the process made to produce the first collection of urbanized areas for Argentina between 1985-2024, which are define as "*Urban areas with a significant density of roads, buildings, and a predominance of non-vegetated surfaces, including infrastructure and open spaces*". The report focusses on the methodological approach and decisions made by the team to produce this product. The methodology is based on the MapBiomas Brasil Urban areas Collection 9 experience, and most of the process was made using Google Earth Engine.

The report is divided into four parts. Being the first one the introduction to the product and the institutions that contribute to the generation of it. The second one is the Methodological description, where each step is explained and refer to the sources, codes and results generated for this collection. The third one is related to results and validation. While the last sections are References.

A relevant information is that due to the lack of information from Patagonia region for 1985 and between 1987 and 1997, the results obtained for the 1986 classification was repeated between this period as constants and updated in 1998. This must be considered now of using this information between this period for this region.

2. MapBiomas Argentina - Urban areas

MapBiomas Argentina - Urban areas is a new initiative in the project of MapBiomas Argentina. The working work started their activities during 2024 and was formalized in the beginning of 2025. Currently is a group of academics and NGO that focus the analysis in how the urbanized areas of Argentina has been changing in the past 40 years.

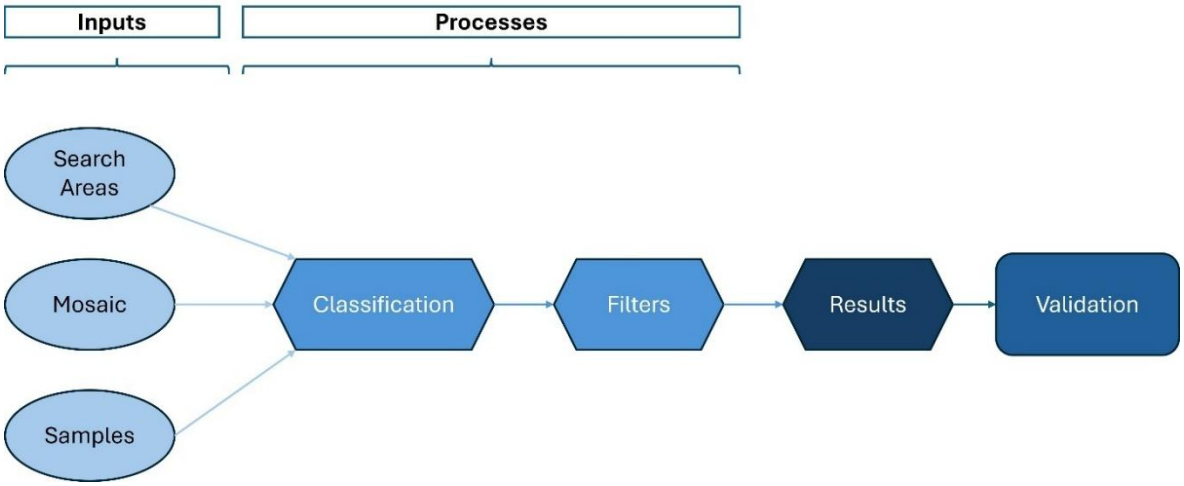
The group is coordinated by member of Universidad Nacional del Centro de la Provincia de Buenos Aires (Natasha Picone, Santiago Linares and Martina Álvarez) and Fundación Vida Silvestre Argentina (Ana Virginia Eljall Qüesta, Gonzalo Dieguez Gaviola and Luna Schteingart). There are also other three national universities that actively participate in the generation of this first collection: Universidad Nacional de Cuyo (Pablo Andrés Rizzo), Universidad Nacional del Nordeste (Osvaldo

Cardozo and Laura Fabiana Gomez) and Universidad Nacional de General Sarmiento (Andrea Pamela Flores).

3. Methodological Description

As mentioned before the based of this methodology is the one used by MapBiomas Brasil Urban areas (2024). For a better fit to the national context some adaptation has been made. Figure N° 1: shows the overview of the steps that's has been made in Argentina to obtain the urbanized areas from 1985 to 2024.

Figure N°1. Methodological overview



Source: MapBiomas Argentina - Urban areas (2025)

Most of the process has been made in Google Earth Engine, though some steps have been in QGIS 3.40. A variety of sources have been used to produce this first collection, each one of them has given a key piece of information. Table 1 shows the complete series of data sources used in this project, and in which part of the process had been used.

Table 1. Sources and data used

Source	Used Information	Methodological Step	Cover
INDEC	Census urban areas (2022)	Search areas	National
INDEC	Population per census urban areas (2022)	Search areas	National
RENABAP	Slums	Search areas	National
10 m Annual LULC	Class 7: "Built area" (2022)	Search area / Spatial filter	Global
Global Urban Settlement	GHS-S Urbanized area (1985/2005/2020)	Sample	Global
MapBiomas Argentina Collection 1	Stable classes for: agriculture, bushes, water and forest	Sample	National
LANDSAT 5, 7, 8 and 9 Collections	Mosaic on the fly per year (1985 – 2024)	Classification	Global
Google Open Buildings v3	Buildings polygons (2023)	Validation	Global

Source: MapBiomas Argentina - Urban areas (2025)

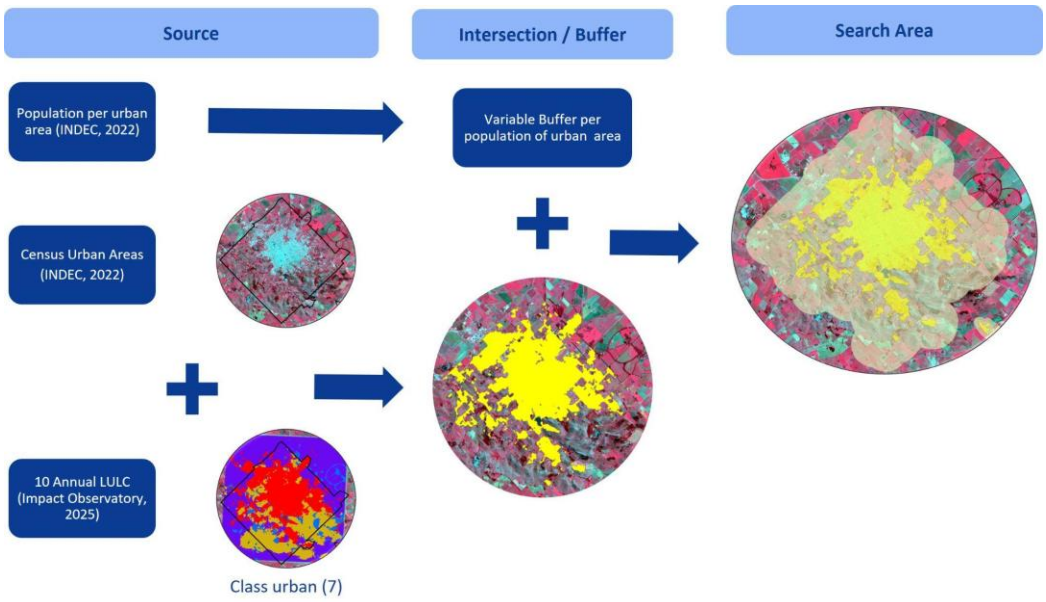
3.1. Inputs preparation

3.1.1. Search Area

Based on the preliminary analysis, two complementary sources were used to define the search areas of the algorithm to be applied for the Argentine territory. The search areas were constructed from an intersection between Census Urban Areas provided by INDEC (2022) and the Urban Class (7) of the Land Cover Land Use (LCLU) provided by Impact Observatory (Figure 2).

The population registered by the 2022 National Census of Population, Households, and Housing (INDEC, 2022) was annexed to these search areas. From this input, and in order to include transitional or expansion areas in cities, variable buffers were constructed based on the urban category (INDEC, 1998), defined as follows: Agglomerated rural population (less than 2000 inhabitants), 300 meters; Large Villages (2000 to 19999 people), 600 meters; Small Towns (20000 to 49999), 900 meters; Minor ATIS (50000 to 99999 people), 1200 meters; Intermediate ATIS (100000 to 399999), 1500 meters; Great ATIS (400000 to 999999 inhabitants), up to 1800 meters; and Large Cities (more than 1 million people), 2100 meters buffer (Figure 2).

Figure 2. Search area generation



Source: MapBiomas Argentina - Urban areas (2025)

3.1.2. Mosaic and topographic maps

The mosaics are generated in Google Earth Engine using the Landsat Collection from satellite 5, 7, 8 and 9. Each annual mosaic is created on the fly and contains a total of 39 bands, which area a combination of Indexes, Fractions and Bands. The complete list is presented in Table 2.

Table 2: List of Bands, Indexes and Fraction that are created to generate the annual Mosaic.

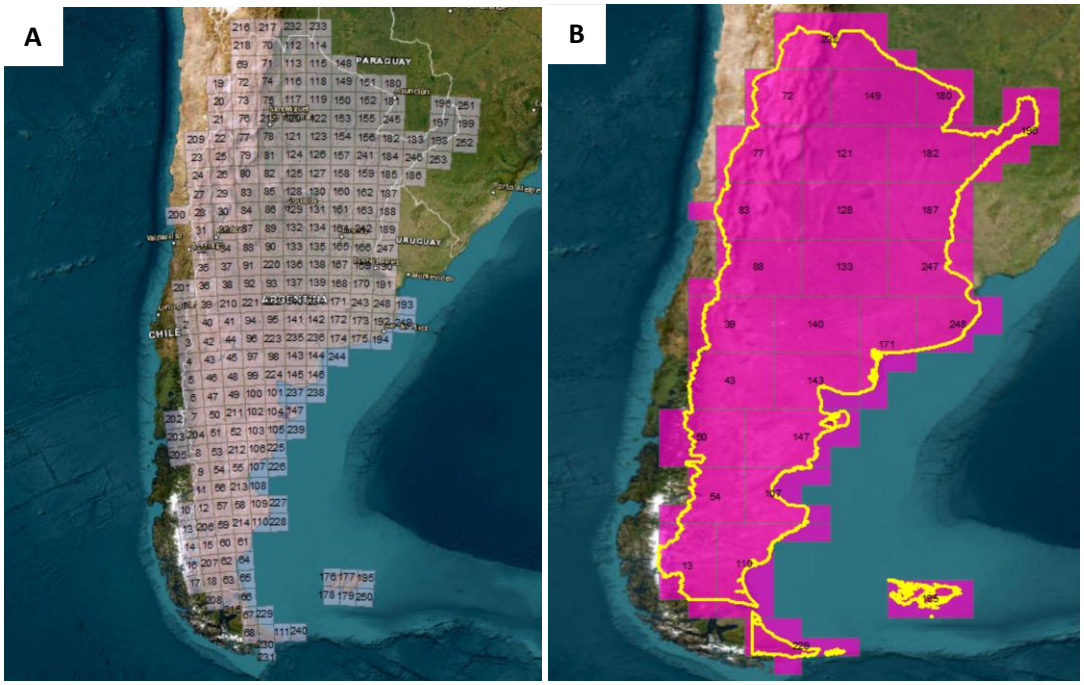
N°	BAND / INDEX / FRACTION	DESCRIPTION	REDUCTION METHOD	SCRIPT ACRONIC
1	BLUE	Landsat band	Median	BLUE_median
2	GREEN	Landsat band	Median	GREEN_median
3	RED	Landsat band	Median	RED_median
4	NIR	Landsat band	Median	NIR_median

5	SWIR1	Landsat band	Median	SWIR1_median
6	SWIR2	Landsat band	Median	SWIR2_median
7	NDVI	Normalized Difference Vegetation Index	Median	NDVI_median
8	EVI1	Enhanced Vegetation Index 1	Median	EVI_median
9	EVI1	Enhanced Vegetation Index percentiles	10p, 90p	EVI_p10, EVI_p90
10	EVI1	Enhanced Vegetation Index percentiles difference	Difference	EVI_dif9010
11	EVI2	Enhanced Vegetation Index 2	Median	EVI2_median
12	EVI2	Enhanced Vegetation Index percentiles	10p, 90p	EVI2_p10, EVI2_p90
13	EVI2	Enhanced Vegetation Index percentiles difference	Difference	EVI2_dif9010
14	MNDWI	Modified Normalized Difference Water Index	Median	MNDWI_median
15	NDWI	Normalized Difference Water Index	Median	NDWI_median
16	NDBI	Normalized Difference Built-Up Index	Median	NDBI_median
17	NBR	Normalized Burn Ratio	Median	NBR_median
18	NDRI	Normalized Difference Road Index	Median	NDRI_median
19	BAI	Bare Soil Area Index	Median	BAI_median
20	UI	Urban Index	Median	UI_median
21	NDUI	Normalized Difference Urban Index	Median	NDUI_median
22	BSI	Bare-Soil Index	Median	BSI_median
23	BU	Built-up Index	Median	BU_median
24	GV	Green Vegetation Fraction	Median	GV_median
25	NPV	Non-Photosynthetic Vegetation Fraction	Median	NPV_median
26	SOIL	Soil Fraction	Median	SOIL_median
27	CLOUD	Cloud Fraction	Median	CLOUD_median
28	SHADE	Shade Fraction	Median	SHADE_median
29	GVS	Green Vegetation + Soil Fraction	Median	GVS_median
30	NDFI	Normalized Difference Fraction	Median	NDFI_median
31	SUBS	Index Substrate Fraction	Median	SUBS_median
32	VEG	Vegetation Fraction	Median	VEG_median
33	DARK	Dark Fraction	Median	DARK_median
34	EBBI	Enhanced Built-Up and Bareness Index	Median	EBBI_median
35	EBBI	Enhanced Built-Up and Bareness Index percentiles	25p, 90p	EBBI_p25, EBBI_p90
36	EBBI	Enhanced Built-Up and Bareness Index percentiles difference	Difference	EBBI_dif7525
37	EBBIsNeg	Positive part of the Enhanced Built-Up and Bareness Index	Median	EBBIsNeg_median
38	EBBIsNeg	Positive part of the Enhanced Built-Up and Bareness Index percentiles	25p, 75p	EBBIsNeg_p25, EBBIsNeg_p75
39	EBBIsNeg	Positive part of the Enhanced Built-Up and Bareness Index percentiles difference	Median	EBBIsNeg_dif7525

Source: MapBiomas Brasil Urban areas (2024)

The Instituto Geográfíco Nacional divide Argentina into 253 topographic maps at 1:250000 scale (Figure 3a). Based on this topographic grid, the homogeneity of the territories that they cover and the particularities of the cities, 28 macro topographic maps were generate (Figure 3b). This final macro maps were used to run the classification algorithm.

Figure 3. Topographic maps and Macro-maps



Source: MapBiomas Argentina - Urban areas (2025)

3.1.3. Samples

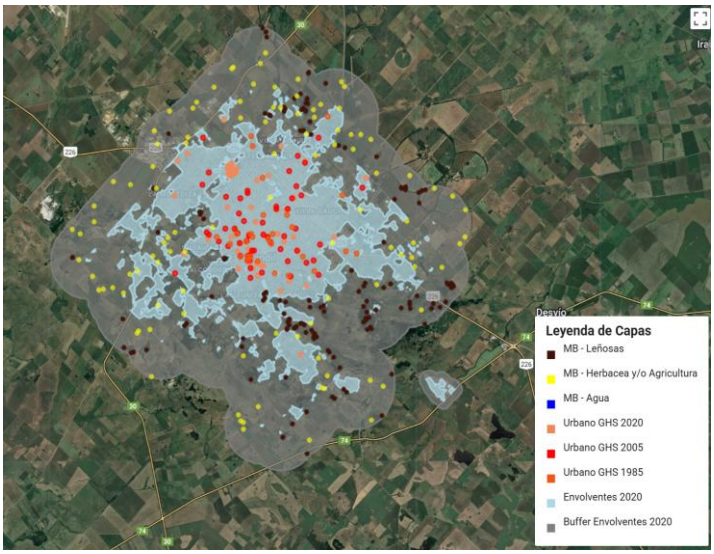
To obtain samples of urbanized areas, the Built-up Surface (GHS-S) product was used (European Commission, 2024). This product, developed by the Joint Research Centre (JRC) of the European Commission, represents the physical extension of the built environment (built surfaces). This product offers us comparable multitemporal data for different times (1975, 1980, 1985, 1990, 1995, 2000, 2005, 2010, 2015 and 2020), with a spatial resolution of 100 meters. The product quantifies the proportion of built-up area in each pixel, allowing the analysis of urban sprawl, densification and urbanization patterns on a global and regional scale.

For this first version of MapBiomas Argentina - Urban areas, data for the years 1985, 2005, 2020 (Figure 4) were selected, using only those pixels that had at least 30 % of their constructed area (>= 3000 m2). On this total number of pixels, samples were randomly extracted, using a stratified methodology to get representation of the different region of MapBiomas Argentina and categories of urban areas.

On the other hand, for the samples of the uses and covers that surround the cities, stable samples (those that have not changed throughout all the years of classification) from collection 1 of MapBiomas Argentina were used, these represent: water, herbaceous, cultivated areas and forests. From these samples only those located between the perimeter of the search area and the corresponding buffer area according to urban category were selected (Figure 4).

A particular case was the generation of samples for Islas Malvinas. A unique set of samples for urban and non-urban areas was used for the hold series (1985-2024), this came from Stable samples from MapBiomas Argentina Collection 1 and manual selection of sample for urban category.

Figure 4. Example of urban and no-urban samples



Source: MapBiomas Argentina - Urban areas (2025)

3.2. Processing

3.3.1 Classification with Random Forest

The process of classifying urbanized areas in Google Earth Engine (GEE) is based on the use of machine learning algorithms, in particular random decision trees (Random Forest). This approach makes it possible to predict, for each pixel of the satellite image, the probability of belonging to the class we call "urbanized area".

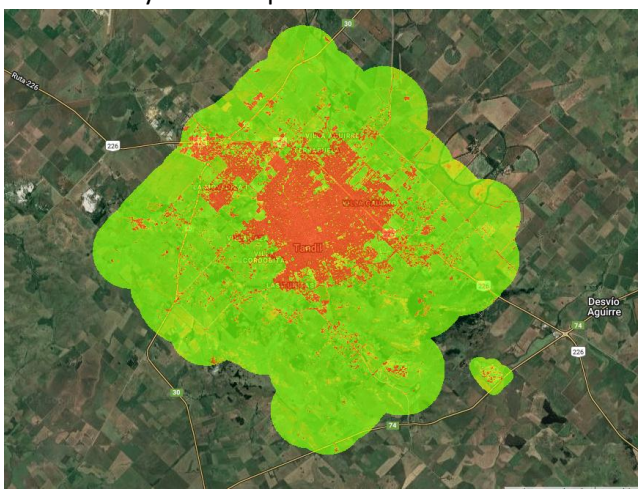
The algorithm is founded on the one developed by MapBiomas Brasil Urban areas, which is based on the implementation of the native Google Earth Engine (GEE) Random Forest algorithm, a set of multiple decision trees built from different random partitions of the data. In the case of MapBiomas Brasil Urban areas – Collection 9 (MapBiomas Brasil Urban areas, 2024), 100 decision trees were used on each annual mosaic of the Brazilian territory, although this number can be adjusted depending on the region or biome.

Each branch is built on a random sample with bootstrap of the training set and considers a random subset of variables in each split, which increases the diversity of the branch set and reduces the risk of overfitting. The final probability assigned to a pixel corresponds to the percentage of branches that classify it as an “urbanized area”. *For example, if 45 out of 50 branches do this, the pixel receives a 90% chance.*

To train the model, representative georeferenced samples of urbanized areas throughout the Argentine territory, obtained from GHS-S for the years 1985, 2005 and 2020, were used. These samples allow us to capture the spectral, temporal and contextual variability of urban areas in different ecoregions and urban scales. In addition, negative samples (non-urban) were incorporated as part of the discrimination process, essential for the calibration of the algorithm and to strengthen the spectral separation between classes (urban vs. non-urban), particularly relevant in areas of spectral mixture such as urban edges, industrial zones with vegetation and non-residential artificial surfaces.

The Random Forest machine learning algorithm (Breiman, 2001) evaluates for each pixel the reflectance values for thirty-nine bands, indexes and fractions that are the inputs (Table 1) and from there its spectral signature is constructed by the analysis of spectral mixtures. The resulting product is a continuous probability map (0–100%) that represents, for each pixel and year, the certainty estimated by the model that this pixel corresponds to an urbanized area (Figure 5). This map undergoes a process of harmonization, using a temporal window of ±2 year, were the final value for each year is the average value for a period of 5 years. This allows to prevent abrupt transitions of probabilities caused by external factors to the dynamics of changes in land use.

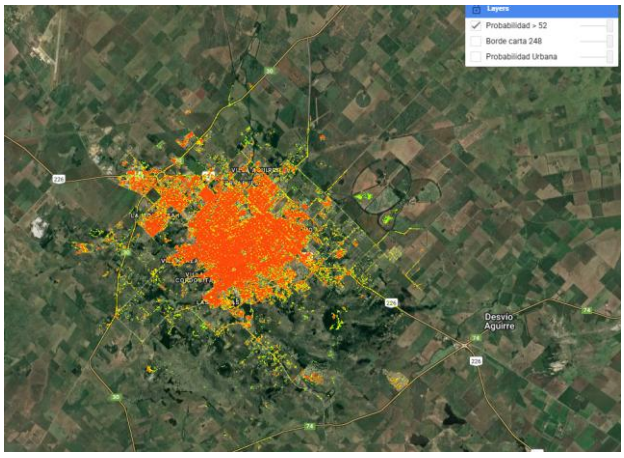
Figure 5. Probability for each pixel to be classified as urbanized area



Source: MapBiomas Argentina - Urban areas (2025)

To extract the layer of urbanized areas for the MapBiomas Argentina - Urban areas collection, we transformed the continuous urbanization probability map (0–100 %) into a binary classification (urban/non-urban). To determine the appropriate value three methods of threshold are calculated: ROC curve with Youden's index, Accuracy, and F1 Score (Figure 6). Per topographic grid one value is selected to extract the urban area.

Figure 6. Thresholds application to extract urbanized area.



Method	Threshold macro map 248
F1 score	52
Accuracy	52
ROC	52

Source: MapBiomas Argentina - Urban areas (2025)

3.2.2. Spatial and temporal Filters

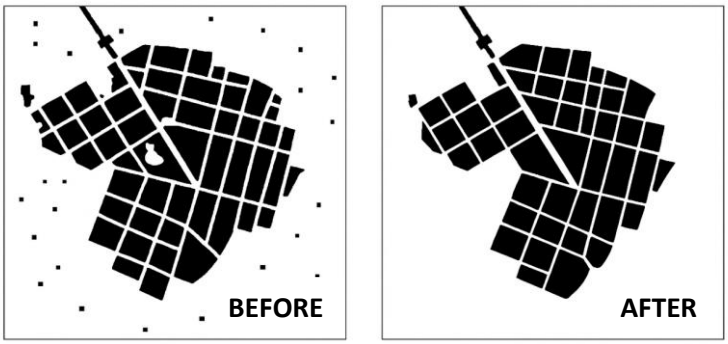
Over the previous layer (extraction of urbanized area per year, base don the threshold per macro map) six filter are applied. Two of them are spatial filter, applied in the beginning and end of

this process. In the middle of this two, four temporal filter are applied. The result is a harmonized series of urbanized area for Argentina.

3.2.2.1. Spatial Filters

Two spatial filters are used. The first of them (SF1) is based on a morphological operation, using a moving window of 3 x 3 pixels (which represents an area of 90 x 90 meters - Figure 7). This filter has a double purpose. First, to erased pixels that have been erroneously classified as urbanized in undeveloped areas, such as scattered structures or isolated pixels in rural areas. Second, using the same function, it allows filling gaps in urbanized areas that were not detected by the classification algorithm or constitute voids in the urban area.

Figure 7. Spatial Filter: morphological error correction.



Source: MapBiomas Argentina - Urban areas (2025)

As an example, Figure 8 shows the application of the SP1. The first screenshot shows the result of the extraction using thresholds of the urbanized area (a) and the second one shows the results of spatial filter, erasing errors outside of the urban continuity or filling gaps inside (b).

Figure 8. Example of SP1 applied to 2024 in Tandil city

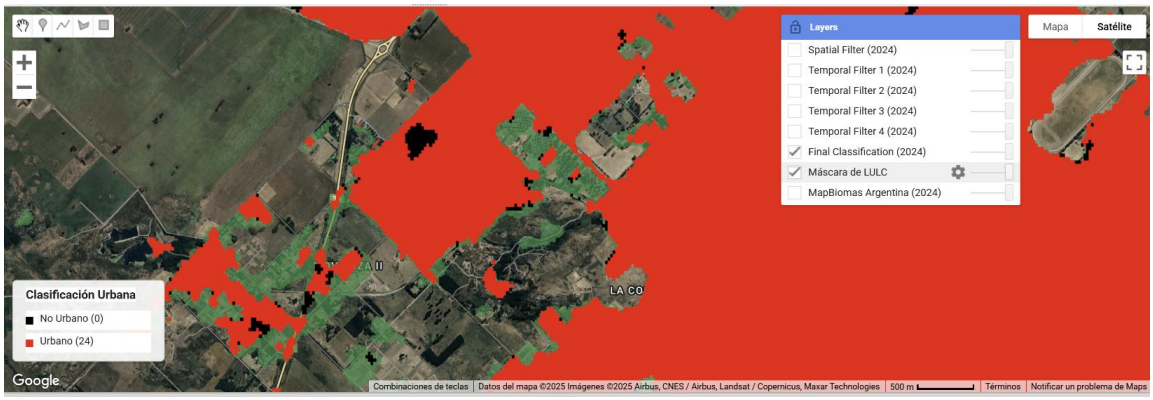


Source: MapBiomas Argentina - Urban areas (2025)

The second of the spatial filters (SF2) is applied after the four temporal filters and consists of an urban area mask made for the year 2024, based on Sentinel images of higher spatial resolution. This mask is obtained from the Copernicus LULC product. All pixels outside this mask will be excluded

from the urbanized area category. Figure 9 shows an example of how the mask take out areas that has been classified as urbanized area by error. (Figure 9).

Figure 9. Example of Spatial Filter 2 for 2024 Tandil city



Source: MapBiomas Argentina - Urban areas (2025)

3.2.2.2. Temporal Filters

The temporal filter come after spatial filter 1 (SP1) and consist of four filter applied in a sequential form (Table 3). The temporal filters applied in MapBiomas Urbano have a crucial role in the purification of inconsistencies in the annual series of land use classification. These four sequential filters allow to correct spurious transitions, smooth out unjustified fluctuations, and reinforce the temporal consistency of the data. Its main objective is to reduce the effect of year-on-year classification errors—such as abrupt alternations between urban and non-urban classes—that do not reflect the real processes of territorial transformation. By doing so, the fidelity of urban mapping is improved, allowing more robust analyses on expansion, consolidation and vulnerability in the Argentine urban fabric.

Table 3. Description of the temporal filter's rules

TF	SCOPE	TYPE	YEARS	KERNEL						CONDITIONALS
				i-2	i-1	i	i+1	i+2	i+3	
1	It acts on the pixels that were classified as 'UA' in the SF results and sets the mask for the filters up to TF2.	FYR	1985 y 1986			x	x	x		If the pixel under analysis is classified as 'UA' within two or more years of the kernel, then the 'UA' is validated to the next step.
		GR	1987 a 2022	x	x	x	x	x		If the pixel under analysis is classified as 'UA' within three or more years of the interval, then the 'UA' is validated to the next step.
		LYR	2023 a 2024	x	x	x				If the pixel under analysis is classified as 'UA' within two or more years of the interval, then the 'UA' is validated to the next step.
2	It acts on pixels that have been validated as 'UA' in the TF1 results.	GR	1985 a 2021			x	x	x	x	If the pixel under analysis is classified as 'UA' within two or more years of the range, then the 'UA' is validated to the next step.
		LYR	2022 y 2023	x	x	x	x			If the pixel is classified as 'UA' within two or more years of the range, then the 'UA' is validated to the next step.
		LYR	2024	x	x	x				If the pixel is classified as 'UA' within one or more years of the range, then the 'UA' is validated to TF3.

3	Extends the filter mask and acts on pixels not classified as urban in TF2.	FYR	1985		x		The results obtained for TF2 are assumed.
		GR	1986 a 2023	x	x	x	If the pixel is classified as 'UA' for i-1 and i+1, then 'UA' is assigned.
		LYR	2024	x	x		If the pixel is rated 'UA' for i-1, then 'UA' is assigned.
4	Area Consolidation Filter.	FYR	1985		x	x	If a pixel under analysis is rated 'UA' for i and not rated for i+1, then it becomes unranked.
		FYR	1986		x		The results obtained for TF3 are assumed.
		GR	1986 a 2024		x	x	If a pixel under analysis is classified as 'UA' for i, then for i+1 it will also be.

FYR= first year rule: LYR= last year rule; UA = Urbanized area; GR = general year rule

Source: MapBiomas Brasil (2024).

3.3. Validation

The validation of the supervised classifications is carried out by means of a cross-confusion matrix, which summarizes the successes and errors with respect to the reference data (Chuvieco, 2016, pp. 399-404). A confusion matrix was constructed and includes two layers, the one from the Open Buildings v3, which detects buildings by artificial intelligence, versus the classification of urbanized areas obtained from the methodology proposed in MapBiomas Argentina - Urban areas. The Open Buildings layer is considered this time as the field truth, given the precision with which it identifies the built surface, especially reliable because it is based on very high-resolution images. To implement this layer, it was necessary to transform its vector format to the raster format, using the same projection system, extension and pixel size as the classification of urbanized area object of validation.

Table 4 presents the evaluation of successes by Argentine province, both in terms of the number of buildings included in the urbanized areas of MapBiomas Argentina - Urban areas, as well as with the surface area in square meters (m2). 97.4% of the buildings from Open Buildings in Argentina are located within the urbanized area class obtained from the MapBiomas classification, this accuracy is greater if the m2 are considered, since 98.01% of the built area is detected, being able to affirm that there are a high spatial concordance and a good precision in the classification. Only 3.4% of the buildings and 2.6% of the built-up area are located outside the urbanized areas, which mostly correspond to rural buildings, detection errors or blurred boundaries in the definition of the urban area.

It is also observed that there is a variation in confidence levels across the country. Neuquén and La Rioja provinces are the only ones that present confidence levels that barely exceed 80%, while Catamarca and San Luis have values very close to 90%. The remaining provinces have values above 95%, with greater accuracy in those provinces belonging to the Pampas region than in the extra-Pampas ones.

Table 4. Validation results per province. Open Buildings vs MapBiomas Argentina Urban-areas

PROVINCE	OPEN BUILDINGS									
	Inside urbanized area (buildings)	Inside urbanized area (km2)	Inside urbanized area (Buildings %)	Inside urbanized area (m2 - %)	Outside urbanized area (buildings)	Outside urbanized area (km2)	Outside urbanized area (Buildings %)	Outside urbanized area (m2 - %)	Total Buildings - Open Buildings	Total km2 Open Buildings
CIUDAD AUTÓNOMA DE BUENOS AIRES	2390727	210.5	99,9	99,9	1460	0.1	0,1	0,1	2392187	210.6
BUENOS AIRES	9193453	814.3	98,9	99,5	102253	4.1	1,1	0,5	9295706	818.4
CATAMARCA	270630	24.1	90,8	92,4	27362	2.0	9,2	7,6	297992	26.1
CHACO	833409	68.3	98,4	99,0	13890	0.7	1,6	1,0	847299	69.0
CHUBUT	316320	29.4	94,8	96,2	17201	1.2	5,2	3,8	333521	30.6
CÓRDOBA	2550324	244.4	96,1	97,2	104101	7.1	3,9	2,8	2654425	251.5
CORRIENTES	404293	38.7	97,5	97,5	10164	1.0	2,5	2,5	414457	39.7
ENTRE RÍOS	902754	78.6	97,6	98,0	22634	1.6	2,4	2,0	925388	80.1
FORMOSA	297575	24.3	97,3	98,4	8222	0.4	2,7	1,6	305797	24.7
JUJUY	377474	32.6	96,1	96,8	15413	1.1	3,9	3,2	392887	33.7
LA PAMPA	239262	24.2	97,7	98,5	5663	0.4	2,3	1,5	244925	24.6
LA RIOJA	218721	20.3	84,9	88,1	38888	2.7	15,1	11,9	257609	23.0
MENDOZA	1073028	109.0	97,3	97,8	30147	2.4	2,7	2,2	1103175	111.4
MISIONES	501638	48.0	97,3	97,9	14183	1.0	2,7	2,1	515821	49.0
NEUQUÉN	223507	19.1	80,0	82,8	55973	4.0	20,0	17,2	279480	23.1
RÍO NEGRO	639093	58.4	95,5	96,4	29882	2.2	4,5	3,6	668975	60.6
SALTA	692722	62.0	95,5	96,1	32812	2.5	4,5	3,9	725534	64.4
SAN JUAN	372023	38.3	94,2	95,8	22741	1.7	5,8	4,2	394764	40.0
SAN LUIS	291887	27.2	90,8	92,8	29542	2.1	9,2	7,2	321429	29.3
SANTA CRUZ	175128	15.9	95,6	97,0	8028	0.5	4,4	3,0	183156	16.4
SANTA FE	2231329	209.1	98,8	99,1	28212	1.9	1,2	0,9	2259541	211.0
SANTIAGO DEL ESTERO	531135	40.8	96,9	97,4	17273	1.1	3,1	2,6	548408	41.9
TIERRA DEL FUEGO	42718	4.5	98,1	98,8	820	0.1	1,9	1,2	43538	4.6
TUCUMÁN	959497	79.9	95,8	96,3	42097	3.1	4,2	3,7	1001594	83.0
ARGENTINA TOTAL	**25728647**	**2321.9**	**97,4**	**98,1**	**678961**	**44.8**	**2,6**	**1,9**	**26407608**	**2366.7**

Source: MapBiomas Argentina - Urban areas (2025)

4. References

- Chuvieco, E. (2016). Teledetección Ambiental. La observación de la Tierra desde el Espacio. 3rd ed. Ariel, Barcelona.
- Impact Observatory (2025). *10 m Annual Land Use Land Cover (LULC) map* [Dataset]. Copernicus / Sentinel Hub Collections. Retrieved September 11, 2025, from https://collections.sentinel-hub.com/impact-observatory-lulc-map/
- European Commission (2024). *Global Human Settlement Layer (GHSL) – Download Portal*. Copernicus Emergency Management Service. Recuperado de https://human-settlement.emergency.copernicus.eu/download.php

- INDEC (1998). CENSO 91. El concepto de localidad: definición, estudios de casos y fundamentos teóricos- metodológicos. Serie D N° 4, Buenos Aires.
- MapBiomas Brasil (2024). Urban Areas - Algorithm's Theoretical Basis Document (ATBD) (Collection 9, Version 1), Brasil. https://brasil.mapbiomas.org/.
- MapBiomas Argentina (2024). MapBiomas Argentina - Colección 1 de los Mapas Anuales de Cobertura y Uso del Suelo en Argentina, accedido en 13 de junio de 2025 a través del enlace: [https://argentina.mapbiomas.org/mapas-de-la-coleccion/].