ARGENTINA

MapBiomas General "Handbook" Algorithm Theoretical Basis Document (ATBD)

Collection 1

Version 1

May, 2024

1. Introduction

1.1. Scope and content of the document

The objective of this document is to describe the theoretical basis, justification, and methods applied to produce annual maps of land use and land cover (LULC) of MapBiomas Argentina. This collection is the result of the integration of 5 working groups or "regions" and includes maps from 1998 to 2022. This document describes the classification methods of Collection 1, the approach to integrating the regions, background information, and a general description of the method applied. The specific procedures applied in each region are in the appendices. The classification algorithms are available on MapBiomas Argentina GitHub.

1.2. Overview

The MapBiomas project started in Argentina in August 2023, aiming to contribute to understanding LCLU dynamics in Argentina. The MapBiomas initiative combines people, algorithms, satellite information and large-scale processing in a methodology that has revolutionized the operational large-scale generation of LULC maps. MapBiomas provides an ideal environment to enhance and share skills and abilities by collaborators from different locations and perspectives but with the mutual core value of learning by doing. Thanks to Google Earth Engine and open source technology it was possible to access and process large-scale datasets of satellite imagery such as the one generated by the MapBiomas project.

MapBiomas collections aim to contribute to developing a fast, reliable, collaborative, and low-cost method to process large-scale datasets and generate historical time series of land use and land cover annual maps. All data, classification maps, codes, statistics, and further analyses are openly accessible through the MapBiomas Platform. This is possible thanks to i) the Google Earth Engine platform, which provides access to data, image processing, standard algorithms, and the cloud computing facilities; ii) the freely available Landsat time-series dataset; and iii) the MapBiomas collaborative network of organizations and experts that share knowledge and mapping tools. MapBiomas maps are an evolving product and subsequent collections will be available in the future.

Collection 1: from 1998 through 2022 (launched in May 2024)

The products of the MapBiomas Argentina Collection 1 are the following:

• Country and regions maps (Chaco, Pampas, Atlantic Forest, Puna, Monte and High Andes, and Patagonia).

• Pre-Processed feature mosaics generated from Landsat archive collections (Landsat 5, Landsat 7, and Landsat 8).

• Image processing infrastructure and algorithms (scripts in Google Earth Engine and source code).

•LCLU transition statistics and spatial analysis with political territories, watersheds, protected areas and ecoregions.

• Quality assessment of the Landsat mosaics. Each scene may have a proportion of clouds and other interference. Thus, each pixel in a given year was qualified according to the number of available observations, varying from 0 to 23 observations per year. The quality assessment of the Landsat mosaics is available on the MapBiomas website.

1.3. Region of Interest

MapBiomas was created to produce LCLU annual maps for the entire Argentinian territory. The work was divided into 5 "regions" or working groups: Chaco, Pampas, Atlantic Forest, Puna, Monte and High Andes, and Patagonia (Fig. 1); each one with its own team and methodological particularities. The working groups have different experiences related to the project. In the case of Chaco, they are in collection 4. Pampas and Atlantic Forests teams are in their third collections. Puna, Monte and High Andes and Patagonia are in their first collection.

Figure 1. Scope of the different regions or working groups

2. Overview and Background Information

2.1. MapBiomas Global Network

We are a network of non-governmental organizations (NGOs), universities, research institutions and tech startups that produce and promote the use of qualified information to stewardship the land use and land cover in South America and other tropical and subtropical regions. As a network, we are committed to generating data, methods, tools and information to qualify the understanding, analysis and decision-making on conservation and sustainable management of continental natural resources.

The core product of MapBiomas is the Land Use & Land Cover Annual Maps, the Transitions Maps (the class differences between a selected period) and Statistics of Land Use & Land Cover Change. There is also the Quality Map (based on the number of observations available related to the proportion of clouds and other interference).

In this initiative, we are dedicated to developing the most updated, completed, and detailed time series of annual maps covering 19998 onwards to support and advance sustainable management and conservation of natural resources.

Other subproducts/modules of Land Use & Land Cover time series maps that at the moment are produced only in Brazil are: i) Deforestation (annual data, frequency and accumulated area); ii) Secondary vegetation areas (age and annual data); iii) Infrastructure (with buffer area from each type of infrastructure); iv) Pasture information (ranging by level of degradation); v) Mining (artisanal/garimpo and industrial types), vi) irrigation; vii) temporal analysis.

The geospatial tech company Ecostage is responsible for the workspace/backend and dashboard/website/frontend of MapBiomas. The Federal University of Technology - Paraná oversees the data repositories of the MapBiomas Network. Google provides the cloud computing infrastructure that allows data processing and analysis through Google Earth Engine and storage through Google Cloud Storage. The funding for the MapBiomas initiative's implementation and operationalization comes from the Arapyaú Institute, Children's Investment Fund Foundation (CIFF), Climate and Land Use Alliance (CLUA), Good Energies Foundation, Gordon & Betty Moore Foundation, Humanize Institute, Institute for Climate and Society (iCS), Norway's International Climate and Forest Initiative (NICFI), Global Wildlife Conservation (GWC), OAK Foundation, Quadrature Climate Foundation (QCF), and Walmart Foundation (U.S.).

2.2. MapBiomas Argentina Network

2.2.1 Origin

The project began in our country in 2017, alongside Bolivia and Paraguay, as Argentina initiated mapping of the Gran Chaco Americano, and its inaugural collection was published in 2020. In 2019, in collaboration with Paraguay, Brazil, and Uruguay, mapping expanded to the regions of the Atlantic Forest and the Pampas, with the first collections published in 2021. With over half of the country mapped, a national-level initiative was consolidated in August 2023, aiming to understand transformation processes across the entire territory using remote sensing data and capitalizing on the experience of local groups in Argentina.

MapBiomas Argentina represents a qualitative leap in automating processes, generating and disseminating spatially explicit information about Argentine territory openly, freely, and at no cost. Its goal is to produce annual collections of land cover and land use maps for the entire country based on the local adaptation of the methodology developed in Brazil in 2015, which is periodically updated. Additionally,

the platform provides users with codes, technical reports, and protocols used in generating information, ensuring process transparency. The platform offers statistics and graphics to facilitate access for different audiences.

2.2.3 Institutions

MapBiomas Argentina is a collaborative network of NGOs, Universities, and Research Institutes that is constantly growing. The network is a group of specialists in remote sensing, ecology, Geographic Information Systems (GIS), and programming from different ecoregions of the country, focused on producing annual collections of land cover and land use maps of Argentina since 1998.

Co-creators:

- Fundación Vida Silvestre Argentina (FVS)

- Facultad de Agronomía - UBA (FAUBA)

- Instituto Nacional de Tecnología Agropecuaria (INTA)

- Instituto de Biología Subtropical (IBS)

- Facultad de Ciencias Forestales - UNaM (FCF)

- Grupo de Estudios Ambientales - Instituto de Matemática Aplicada de San Luis (GEA - IMASL)

- Facultad de Ciencias Agrarias - UNCuyo

- Instituto Argentino de Investigaciones de Zonas Áridas (IADIZA)

- Natura Argentina

- Instituto de Investigaciones en Biodiversidad y Medioambiente (INBIOMA)

- Instituto de Investigaciones en Recursos Naturales, Agroecología y Desarrollo Rural (IRNAD) (UNRN-CONICET)

- Dirección de Bosques Rio Negro

- Centro Regional de Geomática (CEREGEO - UADER)

- Centro de Investigación Científica y de Transferencia Tecnológica a la Producción (CICVTTP-CONICET)

- Centro de Investigaciones de Biociencias y Tecnologías Ambientales para el Desarrollo (CIBTA-UADER)

- GeoSimple

- Instituto de Investigaciones Forestales y Agropecuarias Bariloche (IFAB-CONICET)

- Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)

- Instituto de Investigaciones Fisiológicas y Ecológicas Vinculadas a la Agricultura (IFEVA-CONICET)

3. Methodological Description

3.1. Remote Sensing Data

The imagery dataset used in the MapBiomas project was obtained by the Landsat sensors Thematic Mapper (TM), Enhanced Thematic Mapper Plus (ETM+), and the Operational Land Imager and Thermal Infrared Sensor (OLI-TIRS), onboard of Landsat 5, Landsat 7, and Landsat 8, respectively. The Landsat imagery collections with 30 meters pixel resolution were accessible via Google Earth Engine and produced by NASA and USGS. Specifically, the surface reflectance images from the 2 Tier 1 collection were used. This dataset contains atmospherically corrected surface reflectance and land surface temperature. These images contain 5 visible and near-infrared (VNIR) bands and 2 short-wave infrared (SWIR) bands processed to orthorectified surface reflectance, and one thermal infrared (TIR) band processed to orthorectified surface temperature. They also contain a QA band.

3.2 Google Earth Engine and MapBiomas Computer Applications

MapBiomas image processing is based on Google technology, which includes image processing in cloud computing infrastructure, programming with Javascript and Python via Google Earth Engine, and data storage using Google Cloud Storage. Google defines Google Earth Engine as a platform for petabyte-scale scientific analysis and visualization of geospatial datasets, both for public benefit and for business and government users.

3.3 Overview of methodological process

The common methodological procedures applied in the different working groups or "regions" included several steps (Figure 2). The first step was to generate annual Landsat mosaics comprising specific temporal windows to optimize the spectral contrast and better discriminate the LCLU classes across the regions. The second step was to derive all feature space attributes from the Landsat bands to train one random forest classifier (feature space definition) for each year (Breiman, 2001). Then, yearly training samples were acquired in each region according to its

information availability and statistical needs. The output of the random forest classifier is one LCLU map per year for the entire territory based on the training dataset of that year. Additionally, to improve the classification results, complementary samples were generated, defining georeferenced points of different classes by visual interpretation of historical satellite images (high and very high resolution images) and time series of vegetation indices. Based on the adjusted training data set, a supervised classification using the random forest algorithm was run. Spatial-temporal filters were applied over the classified data for noise removal and temporal stabilization. For more details of each step, see the ATDB from each region. Subsequently, the filtered LCLU maps of each region were hierarchically merged (integrated) based on a set of prevalence rules. The integration rules were set using expert knowledge and are described in Table 1. The transitions and statistics of each class were derived from the annual LCLU maps. The statistical analysis covered different spatial categories, such as biome, state, municipality, watershed, and protected areas.
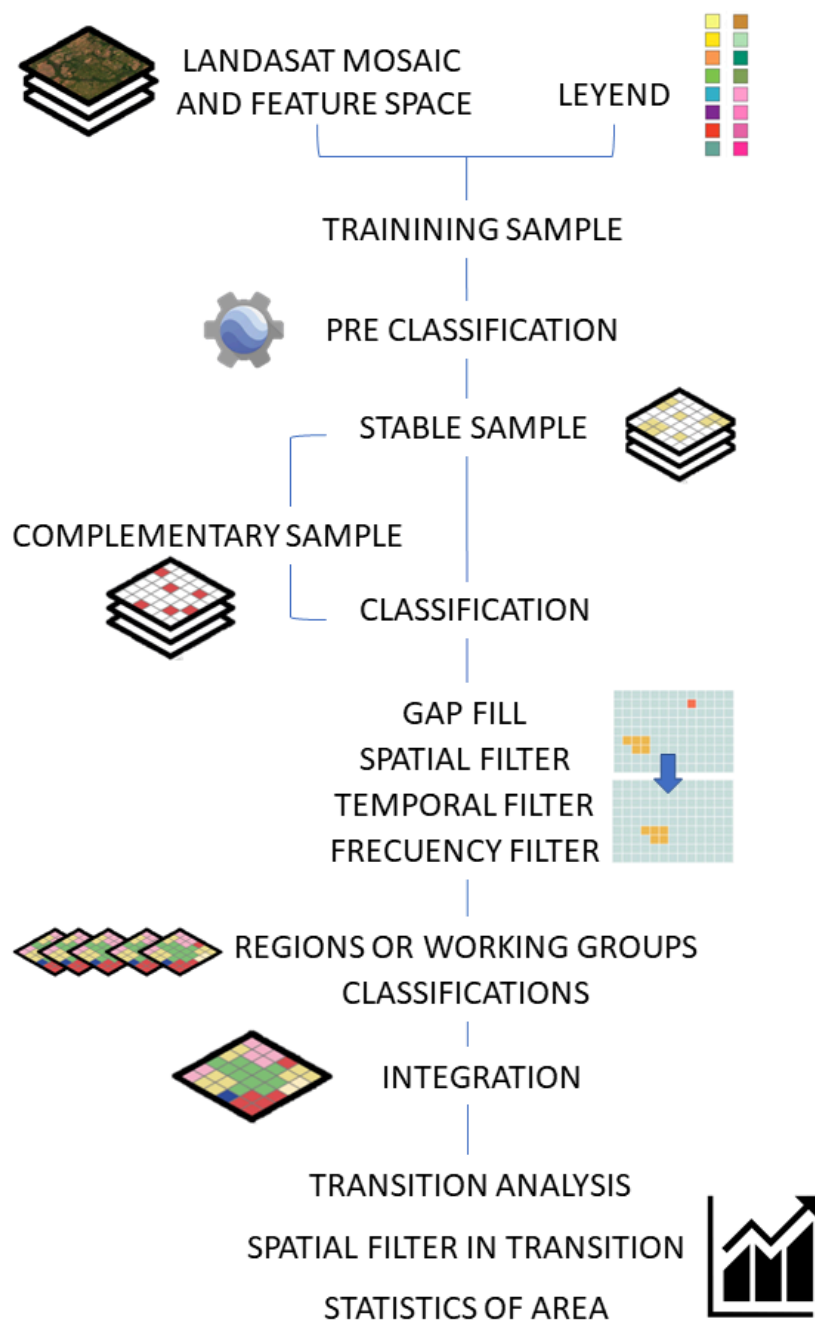
Figure 2. Overview of the classification process for MapBiomas Argentina Collection 1.

3.4 Integration

In order to integrate the regions, the legends of the regional maps up to level 2 were used. For the overlapping areas, it was decided which region overlapped which one in the following order: Pampas, Puna, Monte and High Andes, Patagonia, Chaco,

Bosque Atlántico. In some areas, exceptions were made by prioritizing a class over a region. In addition, the regional legends were readjusted, merging conceptually similar classes (Table 1).

Table 1:. Collection 1 general rules for integrating regions maps.

Rule	
Remap - Pampa	[0,3,4,45,6,66,11,12,43,42,44,67,18,19,57,58,15,65,48, 9,21,22,24,61,25,26,33,34] to [0,3,4,45,6,4,11,12,12,12,63,67,18,18,18,18,15,36,36,9, 21,22,22,22,22,33,33,34])
Remap- Atlantic Forest	[0,3,4,45,6,66,11,12,43,42,44,67,18,19,57,58,15,65,48, 9,21,22,24,61,25,26,33,34] to [0,3,4,45,6,4,11,12,12,12,63,67,18,18,18,18,15,36,36,9, 21,22,22,22,22,33,33,34]
Remap- Chaco	[0,3,4,45,6,66,11,12,43,42,44,67,18,19,57,58,15,65,48, 9,21,22,24,61,25,26,33,34] to [0,3,4,45,6,4,11,12,12,12,63,67,18,18,18,18,15,36,36,9, 21,22,22,22,22,33,33,34]
Remap- Puna, Monte and High Andes	[0,3,4,45,6,66,11,12,43,42,44,67,18,19,57,58,15,65,48, 9,21,22,24,61,25,26,33,34] to [0,3,4,45,6,4,11,63,12,12,12,67,18,18,18,18,15,36,36,9, 21,22,22,22,22,33,33,34]
Remap- Patagonia	[0,3,4,45,6,66,11,12,43,42,44,18,19,57,58,15,65,48,9,2 1,22,24,61,25,26,33,34,67,63]to [0,3,4,45,6,4,11,12,12,12,12,18,18,18,18,15,36,36,9,21, 22,22,22,22,33,33,34, 4,63]
General Rules:	Pampas pastures, above Chaco classification
Specific rule for an area	Pampas grassland, above Chaco classification Pampas Open woodlands, above Chaco classification Pampas Agriculture, above Chaco classification Chaco flooded grassland, above Atlantic Forest classification Open grasslands from chaco to steppe

3.5 Map Legend

The classification for the MapBiomas Argentina initiative included 15 land use and land cover (LULC) classes (Table 1): Closed woodland (3), Open woodland (4), Sparse woodland (45), Flooded woodland (6) Non-wooded floodable vegetation (11), Grassland (12), Steppe (63), Pasture (15), Agriculture (18), Forest Plantation (9), Shub plantation (36), Mosaic of agriculture and pasture (21), Non vegetated area (22), River, lake or ocean (33) Ice and snow on surface (34) and Not observed (27).A full description of the legend is described in the document Legend Description.

Table 2. Land cover and land use classes for MapBiomas Argentina - Collection 1.

COLLECTION 1 - CLASSES	Pixel id	Hexadecimal code	New color
1. Natural wooded vegetation	1	#1f8d49	
1.1. Closed woodland	3	#1f8d49	
1.2. Open woodland	4	#7dc975	
1.3. Sparse woodland	45	#807a40	
1.4. Flooded woodland	6	#026975	
2. Non-wooded natural formation	10	#d6bc74	
2.1. Non-wooded floodable vegetation	11	#519799	
2.2. Grassland	12	#d6bc74	
2.3. Steppe	63	#ebf8b5	
3. Agricultural and livestock area	14	#ffefc3	
3.1. Pasture	15	#edde8e	
3.2. Agriculture	18	#E974ED	
3.3. Forest plantation	9	#7a5900	
3.4. Shrub plantation	36	#d082de	
3.5. Mosaic of agriculture and pasture	21	#ffefc3	
4. Non-vegetated area	22	#d4271e	
5. Water body	26	#2532e4	
5.1. River, lake or ocean	33	#2532e4	
5.2. Ice and snow on surface	34	#93dfe6	
6. Not observed	27	#ffffff	

3.6 Transition Maps

The pixel-by-pixel class differences between any two maps were computed for the following periods: (A) any consecutive years (e.g. 2001-2002); (B) five-year periods (e.g. 2000-2005); (C) all the years (1998-2022). The class transitions represent LCLU changes such as: Transitions from classes of agriculture or non-vegetated areas to forest cover or natural non-forest areas; Transitions that add water surface; Transitions that reduce water surface; Transitions with gain in forest plantation areas; Transitions from forest cover or natural non-forest areas to agriculture or non-vegetated areas; and Areas without transition or transitions that involve not observed areas or transitions between classes within level 1 of legend.

3.7. Spatial Filter on Transition Maps

A spatial filter similar to the one described in 3.6 section was applied in the transition maps. The target is to eliminate single pixels or streams of pixels in the border of different classes derived from the creation of transition maps. The general rules for this filter were: (i) pixels with only one neighbor pixel in the same transition class; (ii) streams of up to five pixels with two or one neighbor pixel in the same transition class.

3.8. Statistics

Zonal statistics of the mapped classes were calculated for different spatial units, such as the biomes, states, and municipalities, as well as watersheds and protected areas. A toolkit in the Google Earth Engine is available to upload user-defined polygons and download the LCLU maps.

4. References

- Breiman, L. 2001. Random forests. Machine learning, v. 45, n. 1, p. 5-32.
- USGS Landsat (2017). USGS Landsat Collection 1 Level 1 Product Definition; USGS Landsat: Sioux Falls,SD,USA, 2017; Volume 26.