MapBiomas General "Handbook"

Algorithm Theoretical Basis Document (ATBD)

Collection 7

Version 1.0

August, 2022

Table of Contents

Executive Summary

MapBiomas initiative was formed in 2015 by universities, NGOs, and tech companies to develop a fast, reliable, collaborative, and low-cost method to produce an annual time series of Brazil's land cover and land use maps from 1985 to 2021 at 30 m resolution. The project is organized by biomes (such as the Amazon, Atlantic Forest, Caatinga, Cerrado, Pampa, and Pantanal) and cross-cutting themes (such as pasture, agriculture, forest plantation, coastal zone, mining, and urban area). A wide range of experts from remote sensing, geography, geology, ecology, environmental and forestry engineering, computer science, human science, journalists, and designers are involved in this project.

Since then, MapBiomas has produced seven sets of annual digital maps of land cover and land use (LCLU), named Collections. The Landsat satellite image classification methods and algorithms for each Collection evolved over the years. Collection 1, published in 2016, consisted of the first step of the mapping process, covering the period of 2008 to 2015 and having seven LCLU classes: forest, agriculture, pasture, forest plantation, mangrove, and water. Collection 2, released in 2017, by applying empirical decision tree classification, encompassed the period of 2000 through 2016 and included 13 LCLU classes with subclasses of forest, savanna, grassland, mangrove, beach, and urban infrastructure, and more. Collection 2.3 was based on a new approach of random forest machine learning to overcome empirical calibration of the input parameters for image classification. In 2018, Collection 3 was also based on the random forest algorithm but included a more robust sampling design for training the classifier and expanded the mapping period from 1985 through 2017. In 2019, Collection 4 was produced including 2018 in the time series and other new approaches and methods, such as 1) deep learning in the aquaculture mapping, 2) a per scene-based analysis for the Amazon biome, 3) the collection of 100 thousand samples for accuracy assessment and area estimation, and 4) reduction and better selection of feature space variables. In 2020, Collection 5 was produced by adding 2019 in the time series and other new improvements and methods; for example, 1) Wetlands were better mapped over the years in the Pantanal biome and also included in the Atlantic Forest biome; 2) The Amazon biome included the new classes Savanna and Grassland Formations in the classification; 3) In Agriculture, perennial and temporary crops were separated, adding the new classes soybean, sugar cane, other temporary crops, and perennial crops. In addition, a beta version of irrigated agriculture maps from 2000 to 2019 was released in the MapBiomas Collection 5.

Collection 6 presented new classes: Wooded Restinga (only in the Atlantic Forest), Rice, Coffee, Citrus, and Other Perennial Crops, totaling 25 mapped LCLU classes. The Wetland class expanded to the Cerrado and Amazon biomes. Until Collection 5, the classification was performed using Landsat top of atmosphere (TOA) data. In Collection 6, new Landsat mosaics were processed using surface reflectance (SR) data. Using all the images from each year, 90 spectral and temporal metrics were generated. Besides random forests, U-Net (convolutional neural network) was applied in Aquaculture, Mining, Irrigation, Rice and Citrus Classification. MapBiomas Water classification was incorporated in the Water class. The Mining class was divided into industrial mining and *garimpo*, as well as categorizing the primary substance being exploited.

The current Collection 7 presented two more new classes: Herbaceous Sandbank Vegetation (only in the Atlantic Forest, Pampa and Caatinga biomes) and Cotton (beta version). The Rocky Outcrop class was included in the Cerrado mapping. In Collection 7, new USGS Landsat surface reflectance images (Collection 2, Tier 1) were used in the classification. The accuracy assessment analysis based on the acquisition of 85,000 independent samples was updated and revised adding the years, 2019, 2020, 2021. Since Collection 6, spatial and temporal filters were not only applied in the mappings of biomes and cross-cutting themes but also at the end on integrated maps. The specific procedures applied in each biome and cross-cutting theme and its improvements in Collection 7 are in the appendices (https://mapbiomas.org/en/download-of-atbds).

This Algorithm Theoretical Basis Document (ATBD) aims to provide the methodological steps to produce MapBiomas Collection 7 and describe the datasets and analysis. All the MapBiomas maps and datasets are freely available on the project platform (http://mapbiomas.org), as well as all computational algorithms used in the MapBiomas classifications are available on Github (https://github.com/mapbiomas-brazil).

1. Introduction

1.1. Scope and content of the document

This document describes the theoretical basis, justification, and methods applied to produce annual maps of land cover and land use (LCLU) in Brazil from 1985 to 2021 of the MapBiomas Collection 7.

This document covers the classification methods of Collection 7, the image processing architecture, and the approach to integrating the biomes and cross-cutting theme maps. In addition, the document presents a historical context and background information, a general description of the satellite imagery datasets, feature inputs, and the accuracy assessment method applied.

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

1.2. Overview

The MapBiomas project was launched in July 2015, aiming to contribute to understanding LCLU dynamics in Brazil. The LCLU annual maps produced in this project were based on the Landsat archive available in the Google Earth Engine platform, encompassing the years from 1985 to the present. Since then, the MapBiomas mapping evolved year by year and was divided into Collections (more about MapBiomas' Collections comparisons in Table 5).

- Collection 1: 2008 through 2015 (launched in April 2016).
- Collection 2: 2000 through 2016 (launched in April 2017).
- Collection 2.3: a revised version of Collection 2 (launched in December 2017).
- Collection 3: 1985 through 2017 (launched in August 2018).
- Collection 4: 1985 through 2018 (launched in August 2019).
- Collection 5: 1985 through 2019 (launched in August 2020).
- Collection 6: 1985 through 2020 (launched in August 2021).
- Collection 7: 1985 through 2021 (launched in August 2022).

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 LCLU annual maps. All data, classification maps, codes, statistics, and further analyses are openly accessible through the MapBiomas Platform (https://plataforma.brasil.mapbiomas.org/). All these are possible thanks to i) Google Earth Engine platform, which provides access to data, image processing, standard algorithms, and the cloud computing facility; ii) freely available Landsat time-series dataset; iii) MapBiomas collaborative network of organizations and experts that share knowledge and mapping tools; and iv) to visionary funding agencies that support the project (Souza Jr et al., 2020).

The products of the MapBiomas Collection 7 are the following:

• Biome maps (Amazon, Atlantic Forest, Caatinga, Cerrado, Pampa, and Pantanal) and cross-cutting theme maps (Pasture, Agriculture, Forest Plantation, Coastal Zone, Mining, and Urban Areas);
• 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 other land tenure categorical maps.
• 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 classified 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.
• LCLU data around buffers of transportation and energy infrastructure.
• Annual deforestation maps (from 1987 to 2020).
• Secondary vegetation annual maps (from 1987 to 2019).
• Irrigation maps (for center pivot irrigation systems, irrigated rice and other irrigation systems).
• Pasture quality maps (only from 2000 to 2020).
• Industrial and *garimpo* mining by mineral type.
●Temporal analysis (stable areas and number of classes).

Besides these products, the MapBiomas network released MapBiomas Water Collection 2 and MapBiomas Fire Collection 1 with annual and monthly mapping of Brazil's water surface and fire scars from 1985 to 2021 and 2020, respectively.

The MapBiomas initiative also expanded in other regions and countries, such as the Chaco region with its Collection 2 (https://chaco.mapbiomas.org/), the Amazon region with its Collection 3 (https://amazonia.mapbiomas.org/), MapBiomas Atlantic Forest (https://bosqueatlantico.mapbiomas.org/), South American Pampa (https://pampa.mapbiomas.org/), and most recently in Indonesia (https://mapbiomas.nusantara.earth/) with Collection 1. These new project areas follow the mapping protocol of MapBiomas Brazil with a few adjustments considering the peculiarities of their ecosystems and local teams. Detailed methodological information about these MapBiomas initiatives is found at the ATBD of these regions.

1.3. Region of Interest

MapBiomas was created to produce LCLU annual maps for the entire Brazilian territory, thus covering all the six official biomes of the country: Amazon, Atlantic Forest, Caatinga, Cerrado, Pampa, and Pantanal (Figure 1). The biomes' division helps classify distinct LCLU classes and landscape patterns across the country (Table 1). The project was also divided into cross-cutting themes: Agriculture, Pasture, Forest Plantation, Coastal Zone, Mining, and Urban Areas. Although the Coastal Zone is not considered a biome officially, this region that covers dunes, beaches, hypersaline tidal flats (*Apicuns*), aquaculture, and mangroves along the Brazilian coast was treated as such.

Since Collection 5, the new Brazilian biomes official map (1:250,000) of IBGE (2019) has been used by the biome classifications and statistics. For the other past MapBiomas collections, it was produced a 1:1,000,000 map of the biomes based on the refinement of the official map of biomes in Brazil published by IBGE (1:5,000,000) considering the Brazilian boundaries map (1:250,000) and the physiognomies map (1:1,000,000), both from IBGE.

Figure 1. Brazilian biomes were mapped in the MapBiomas project to generate the Collection 7 products (source: IBGE, 2019).

Table 1. Land cover and land use characteristics of the Brazilian biomes.

Biome	Area (km^2) (Country %)	Land Cover	Predominant Land Use
Amazon	4,196,943 (49.29%)	Evergreen forest, with enclaves of savanna, natural grassland, and extensive wetlands and surface water, with almost 20% of the biome of the forest area cleared.	Cattle ranching, agriculture, mining, logging and non-timber forestry production.
Atlantic Forest	1,110,182 (13.04%)	Isolated forest fragments (Morellato & Haddad, 2000), mostly old secondary growth, surrounded by croplands, pasture, forest plantation, and urban area.	Agriculture, cattle ranching, urban, forest plantation, artificial water reservoir.
Caatinga	844,453 (9.92%)	Woody and deciduous forest, with at least 50% of the original converted (de Oliveira et al., 2012).	Agriculture, cattle ranching, smallholder livestock production, urbanization.
Cerrado	2,036,448 (23.92%)	Mosaic of savanna, grassland, and forest, 50% of the native vegetation cover has already been converted (PPCerrado/INPE).	Agriculture, cattle ranching.
Pampa	176,496 (2.07%)	Natural grassland, with scattered shrub and trees, rock outcrop formations (Roesch et al., 2009).	Agriculture (rice, soy, perennial crops), livestock production (in natural grasslands), forest plantation, and urbanization.
Pantanal	150,355 (1.76%)	Forest, savanna, grassland and wetland.	Agriculture and cattle ranching.

1.4. Key Science Applications

MapBiomas was originally designed to fill gaps in Brazil's greenhouse gas emission estimates of the LCLU change sector. However, other scientific applications can be derived from an annual time-series history of LCLU maps produced, including:

- Mapping and quantifying LCLU transitions.
- Quantification of gross and net forest cover loss and gain.
- Monitoring of regeneration and secondary growth forests.
- Monitoring of water resources and their interaction with LCLU classes.
- Monitoring agriculture and pasture expansion.
- Monitoring natural disasters.

- Expansion of infrastructure and urbanization.
- Identification of desertification process.
- Regional planning.
- Management of Protected Areas.
- Monitoring of land change modeling.
- Infectious disease risk modeling.
- Climate distribution modeling.

MapBiomas Award was created to stimulate and expand technical and scientific applications that use any MapBiomas initiative and product data, including initiatives from other countries. The award is now in the fourth edition, presenting four categories: General, Young People, Applications in Public Policy, and the new category of this edition, Applications in Business (https://mapbiomas.org/en/award).

2. Overview and Background Information

2.1. Context and Key Information

This section addresses complementary contextual and critical information relevant to understanding the MapBiomas products and methods to generate the collections.

2.1.1. MapBiomas Network

MapBiomas is a multi-institutional initiative of the Climate Observatory (a network of NGOs working on climate change in Brazil - http://www.observatoriodoclima.eco.br/en/). The co-creators of MapBiomas involve NGOs, universities, and technology companies (list of all organizations engaged in Annex I).

These organizations play specific or multiple roles and contribute to the project's overall development. Each biome and cross-cutting theme (Agriculture, Pasture, Forest Plantation, Coastal Zone, Mining, and Urban Areas) has a lead organization, as shown in the box below.

Biome coordination:

- **Amazon** – Institute of People and Environment of the Amazon (IMAZON).

• **Atlantic Forest** – SOS Atlantic Forest Foundation and ArcPlan.

• **Caatinga** – State University of Feira de Santana (UEFS), Northeast Plants Association (APNE) and Geodatin.

• **Cerrado** – Amazon Environmental Research Institute (IPAM).

• **Pampa** – Federal University of Rio Grande do Sul (UFRGS) and GeoKarten.

• **Pantanal** – SOS Pantanal Institute and ArcPlan.

Cross-cutting theme coordination:

• **Pasture** – Federal University of Goias (LAPIG/UFG).

• **Agriculture and Forest Plantation** – Agrosatelite.

• **Coastal Zone and Mining** – Solved and Vale Technological Institute (ITV).

• **Urban Area** – University of São Paulo (USP - QUAPÁ-FAU and YBY), Federal University of Bahia (UFBA) and Federal University of São Carlos (UFSCar - NEEPC).

The geospatial tech company Ecostage is responsible for the workspace/backend and dashboard/website/frontend of MapBiomas. Google provides the cloud computing infrastructure that allows data processing, analysis, and storage through Google Earth Engine.

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.).

Since both Climate Observatory and MapBiomas are not institutions, the initiative received generous institutional management to operational and financing tasks from partners, including Arapyaú Institute, Avina Foundation, and World Resources Institute (WRI).

The project also has an independent Scientific Advisory Committee (SAC) composed by:

• Dr. Alexandre Camargo Coutinho (Embrapa)

- Dr. Edson Eygi Sano (IBAMA)
- Dr. Gilberto Camara Neto (INPE)
- Dr. Joberto Veloso de Freitas (Federal University of Amazonas)
- Dr. Matthew C. Hansen (Maryland University)
- Dr. Mercedes Bustamante (University of Brasília)
- Dr. Timothy Boucher (TNC)
- Dr. Robert Gilmore Pontius Jr (Clark University)

2.1.2. Remote Sensing Data

The imagery dataset used in the MapBiomas project, across Collections 1 to 7, 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.

In earlier collections, MapBiomas primarily utilized USGS Collection 1 Tier 1 from the top of the atmosphere reflectance (TOA). In Collections 6 and 7, new Landsat mosaics were processed using surface reflectance (SR) data. However, Collection 7 used the USGS Landsat Collection 2 (Tier 1).

2.1.3. 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."

The MapBiomas project has developed the following computer applications based on Google Earth Engine:

- Javascript scripts - these computer codes were written directly in the Google Earth Engine Code Editor and were used to prototype new image processing algorithms and test large-scale image processing to be implemented in the Workspace environment for Collections 1 and 2. Most image classification and post-classification of Collections 2.3, 3, 4, 5, 6, and 7 were written in Javascript.

- Python scripts – This code category was used to optimize the image processing of large datasets in Google Earth Engine. In addition, the map integration, post-classification tasks, and statistical analysis were all performed in Earth Engine Python API.

• Toolkits – User's Toolkits are collections of scripts in Google Earth Engine to download MapBiomas' maps and data by state, biome, municipality, or any other geometry. These toolkits were developed for the general public that is not used to programming to access the MapBiomas data differently. They are often reviewed and improved. All the toolkits are available at https://mapbiomas.org/en/tools. Instructions are also available on Github (https://github.com/mapbiomas-brazil/user-toolkit).

• Github repository - All Javascript, Python, toolkit, and dashboard codes are available at the public Github repository: (https://github.com/mapbiomas-brazil).

• Workspace - a web-based application to allow general users with no-programming experience to access imagery collections, process them, manage and store the results in databases, and map assets (*i.e.*, new collections). The biome maps of Collections 1 and 2 were produced using the Workspace application before the random forest approach classification. The biome teams of analysts worked simultaneously to set the image classification parameters, pre-process and evaluate the results and later submit tasks to large-scale image processing to generate the final products, which were the Landsat image mosaics and LCLU maps.

• Mapbiomas.org (Dashboard). The web platform of the MapBiomas initiative presents the Landsat image mosaics and their quality, land cover and land use annual maps of Collection 7, transitions analysis, statistics, and all the methodological information about the ATBD, tools, scripts, fact sheets, tutorial videos, and accuracy analysis. Besides the land cover and land use change data, the MapBiomas dashboard presents other products, such as deforestation, secondary vegetation, irrigation, infrastructure, pasture quality, fire scars, mining, and water surface mappings. All Landsat mosaics, maps, data, and methodological documents of the MapBiomas Collections are freely available to download, and information about the MapBiomas initiative at the MapBiomas website (http://mapbiomas.org/).

2.2. Historical Perspective: Existent Mapping Initiatives

The existing LCLU mapping efforts that covered all of Brazil before MapBiomas, were neither frequent nor updated (Annex II) and sometimes had lower resolution. MapBiomas and the available global and national land cover products can be used complementary, but there are potential advantages of MapBiomas maps. First, the MapBiomas maps reconstructed the entire Landsat annual time series (>35 years). The classification scheme is also more relevant for national applications because it follows the Brazilian vegetation classification legend (IBGE, 2012). In addition, MapBiomas has the potential to monitor primary forest changes (*i.e.,* deforestation), secondary forest regrowth, and land use classes (pasture, agriculture, forest plantation, mining, and urban area) along with this time series.

All products from MapBiomas and methods and codes to produce the maps are publicly available on the internet. This enables its reproduction in other contexts, just as it is currently taking place in all other Amazonian nations (Peru, Ecuador, Bolivia, Colombia, Venezuela, Guyanas, and Suriname - http://amazonia.mapbiomas.org/), the

Chaco region (Argentina, Bolivia, and Paraguay - http://chaco.mapbiomas.org/), the Pampa region (Brazil, Argentina, and Uruguay - https://pampa.mapbiomas.org/)

2.2.1. International land cover and land use data

Mapping initiatives at the global level complement national mapping efforts (Annex II). In collaboration with the University of Maryland, the USGS produced global land cover and tree cover layers. USGS also has a MODIS land cover map at a 500 m pixel scale. The GlobCover Portal is another initiative from the European Space Agency (ESA) which produced land cover maps with MERIS sensor at 300 m spatial resolution for two periods: December 2004 - June 2006 and January - December 2009. Later, the ESA Climate Change Initiative produced global land cover maps from 1992 to 2015 at 300 m spatial resolution and Global Sentinel 2 Land Cover for 2020 with 10 m resolution (https://viewer.esa-worldcover.org/worldcover/). Global Forest Watch (GFW) and Google Earth Engine provide the Global Forest Change (GFC) maps from 2000 to 2019 derived from the Landsat imagery at 30 m resolution produced by the University of Maryland (Global Land Cover Facility - GLCF). The National Geomatics Center of China (NGCC) had GlobeLand30 - a high-resolution (30 m) full coverage land cover map for the years 2000 and 2010 (Chen et al. 2014). Japan Aerospace Exploration Agency (JAXA) also produced a forest/non-forest map for 2007-2010 using a 25 m-resolution PALSAR mosaic. Other global products were made using lower spatial resolution (>500 m) (e.g. Friedl et al. 2010) but were not presented here because their resolutions limit applications to assess MapBiomas products, which are produced at 30 m Landsat pixel. Most recently, an automated approach for global land use land cover (LULC) classification was published using deep learning on 10 m Sentinel-2 imagery (Brown et al. 2022).

2.2.2. National land cover and land use data

The RadamBrasil project was the first national initiative to map the vegetation of the entire country of Brazil. This project was conducted from 1975 to 1980 based on airborne radar imagery, visual interpretation, and extensive and detailed fieldwork involving several organizations. The RadamBrasil project produced maps at a 1:250.000 scale, which is still a solid reference for scientific and technical studies about vegetation (Cardoso, 2009).

Only the Amazon and Atlantic Forest biomes were being monitored, so after the RadamBrasil project, the Minister of Environment launched the natural vegetation map of Brazil in 2004. This map was created as part of the Probio (Projecto de Conservação e Utilizaço Sustentável da Diversidade Biológica Brasileira) initiative. The Brazilian biome boundaries (IBGE, 2004a) were used as a reference for the national mapping initiative. The Probio project was based on Landsat imagery acquired in 2002, with a minimum mapping unit varying from 40 to 100 hectares and a mapping scale of 1:250.000. The accuracy assessment was based on digital imagery products at 1:100.000, with a minimum overall accuracy of 85%. The land cover classes followed the IBGE manual for vegetation mapping (IBGE, 2004b). The Probio project updated forest change mapping for the year 2008 for all biomes and the years 2009, 2010, and 2011 depending on the biome.

In the context of the National Inventories of GHG Emissions and Removals, the Ministry of Science and Technology commissioned the production of land cover and land use maps of Brazil for the years 1994, 2002, 2010 (also 2005 for the Amazon), and 2016. Those maps were produced based on segmentation and visual interpretation of Landsat Imagery and identifying natural vegetation (forest and non-forest), agriculture, pasture, silviculture, urban areas, dunes, rocky outcrops, mining, and water on the 1:250.000 visual scale.

More recently, IBGE has published a platform to monitor LCLU in Brazil, making available maps for 2000, 2010, 2012, 2014, 2016, and 2018 on a 1 km resolution and covering the classes of forests, savannas, agriculture, pasture, urban areas, and water and mosaics of those classes.

2.2.3. Regional and biomes land cover and land use data

There are also reference maps at the biome scale and through the cross-cutting themes. For example, the PRODES and the TerraClass maps are available for the Amazon biome and, more recently, in the Cerrado biome for some years. There are also maps available for subareas of the Pampa biome at the state level (e.g., Rio Grande do Sul, São Paulo, and Tocantins states). These reference land cover and land use maps for the biomes and cross-cutting themes are presented in Annex II.

3. Methodological description

The Collection 7 general methodological steps are presented in 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 biomes and cross-cutting themes (see the biome and cross-cutting theme Appendices for detailed information). The second step was to derive all attributes from the Landsat bands to guide the random forest classifier (feature space definition) (Breiman, 2001). Training samples were acquired in each biome and cross-cutting theme to adjust the training dataset according to its information availability and statistical needs. There is one LCLU map per year based on the adjusted training dataset. Based on the adjusted training dataset, except for the Aquaculture, Mining, Irrigation, Rice, and Citrus Classification, which had its identification based on the U-Net convolutional neural network classifier (CNN). For noise removal and temporal stabilization, spatial-temporal filters were applied over the classified data. Subsequently, the filtered LCLU maps of each biome and cross-cutting themes were hierarchically merged (integrated) based on a set of prevalence rules to create the final Collection 7 product. Spatial and temporal filters were once again applied at the end on integrated maps. Finally, the Accuracy assessment analysis was based on acquiring 85,000 independent samples per year from 1985 to 2021. The acquisition of the validation samples followed the good practices proposed by Olofsson et al. (2014), Stehman et al. (2014), and Stehman & Fody (2019). 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.

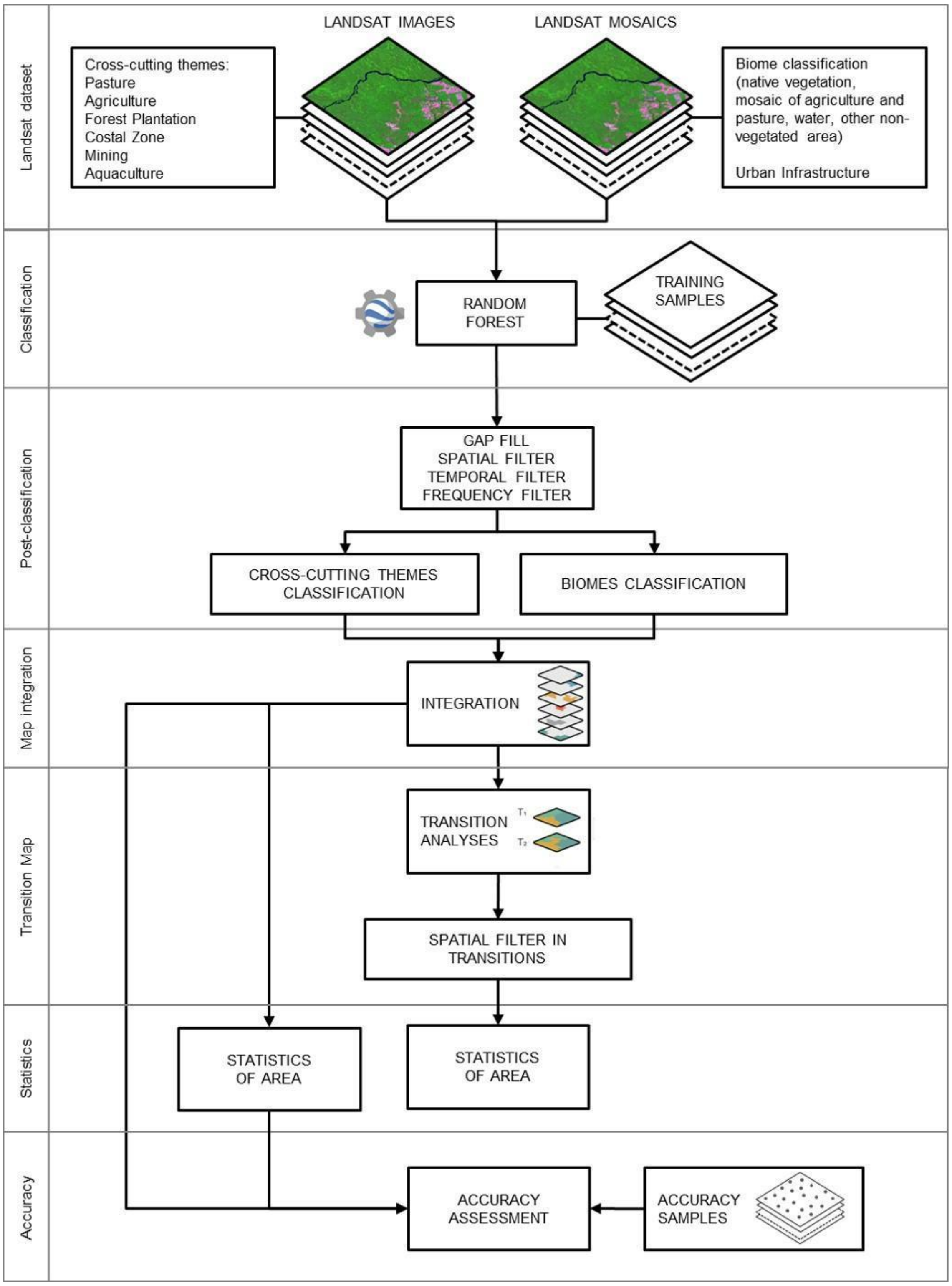


Figure 2. General methodological steps of Collection 7 to implement MapBiomas algorithms in the Google Earth Engine.

3.1. Landsat Mosaics

All biomes except Amazon generated Landsat cloud-free composites based on specific periods of time to optimize the spectral contrast and help within the discrimination of LCLU classes. The cloud/shadow removal script takes advantage of the quality assessment (QA) band and the GEE median reducer. QA values can improve data integrity by indicating which pixels might be affected by artifacts or subject to cloud contamination (USGS, 2017). In conjunction, GEE can be instructed to pick the median pixel value in a stack of images. By doing so, the engine rejects values that are too bright (e.g., clouds) or too dark (e.g., shadows) and picks the median pixel value in each band over time.

In Collection 7, new USGS Landsat surface reflectance images (Collection 2, Tier 1) were used in the classification. For each chart, a specific temporal mosaic of Landsat images was built based on the following selection criteria/parameters: 1. The selected Landsat data must allow an annual analysis, and 2. The period for Landsat scene selection (t0 and t1 in day/month/year) must provide enough spectral contrast to distinguish LCLU classes better.

The cross-cutting themes (Pasture, Agriculture, Forest Plantation, Urban Area, Coastal Zone, and Mining) processed Landsat mosaics per scene basis (more details available in the biome and cross-cutting theme Appendices). To improve image classification and reduce noises, a new tool was developed to evaluate the images individually, excluding images, to improve the mosaic quality. The Amazon biome classified each Landsat image using Random Forests and reclassified the results to create the annual LCLU map.

3.2. MapBiomas feature space

The total available bands of the MapBiomas feature space are composed of 90 input variables, including the original Landsat bands and fractional and textural information derived from these bands (Table 2). Table 2 presents the formula or the description to obtain these feature variables, as well as highlights in green all the bands, indices, and fractions available in the feature space. In addition, statistical reducers were used to generate temporal features such as:

- Median: Median of the pixel values within the defined stack of images.
- Median_dry: median of the quartile of the lowest pixel NDVI values.
- Median_wet: median of the quartile of the highest pixel NDVI values.
- Amplitude: amplitude of variation of the index considering all the year's images.
- stdDev: stdDev of the pixel values within the defined stack of images.
- Min: the lower annual value of the pixels of each band.

Table 2. List, description, and reference of bands, fractions and indices available in the feature space.

			Reducer					
	band or index name	**formula**	**median**	**median _dry**	**median _wet**	**amplitude**	**stdDev**	**min**
bands	blue	B1 (L5 e L7); B2 (L8)	■	■	■		■	■
	green	B2 (L5 e L7); B3 (L8)	■	■	■		■	■
	red	B3 (L5 e L7); B4 (L8)	■	■	■		■	■
	nir	B4 (L5 e L7); B5 (L8)	■	■	■		■	■
	swir1	B5 (L5 e L7); B6 (L8)	■	■	■		■	■
	swir2	B7 (L5); B8 (L7); B7 (L8)	■	■	■		■	■
	temp	B6 (L5 e L7); B10 (L8)	■				■	■
index	ndvi	(nir - red)/(nir + red)	■	■	■	■	■	
	evi2	(2.5 * (nir - red)/(nir + 2.4 * red + 1)	■	■	■	■	■	
	cai	(swir2 / swir1)	■	■	■		■	
	ndwi	(nir - swir1)/(nir + swir1)	■	■	■	■	■	
	gcvi	(nir / green - 1)	■	■	■		■	
	hall_cover	(-red*0.017 - nir*0.007 - swir2*0.079 + 5.22)	■				■	
	pri	(blue - green)/(blue + green)	■	■	■		■	
	savi	(1 + L) * (nir - red)/(nir + red + 0,5)	■	■	■		■	
	textG	('median_green') .entropy(ee.Kernel .square({radius: 5}))	■					
fraction	gv	fractional abundance of green vegetation within the pixel	■			■	■	
	npv	fractional abundance of non-photosynthetic vegetation within the pixel	■			■	■	
	soil	fractional abundance of soil within the pixel	■				■	
	cloud	fractional abundance of cloud within the pixel	■			■	■	

	shade	100 - (gv + npv + soil + cloud)						
MEM index	gvs	gv / (gv + npv + soil + cloud)						
	ndfi	(gvs - (npv + soil))/(gvs + (npv + soil))						
	sefi	(gv+npv_s - soil)/(gv+npv_s + soil)						
	wefi	((gv+npv) - (soil+shade)) /((gv+npv) + (soil+shade))						
	fns	((gv+shade) - soil) / ((gv+shade) + soil)						
slope		ALOS DSM: Global 30m						

Each biome executed a feature selection mechanism to choose the most appropriate subset of variables to run the random forest algorithm later. Each biome and cross-cutting theme selected their feature variables, and more details are available in the Appendices.

3.3. Classification

3.3.1. Legend

The MapBiomas classification scheme is a hierarchical system with a combination of LULC classes (Table 3). At Level 1, there are six classes: Forest, Non-Forest Formation, Farming, Non-Vegetated Area, Water, and Not-Observed. Level 2 has 16 classes that also have a combination of LULC classes. The Forest Level 1 class is divided into four subclasses: Forest Formation, Savanna Formation, Mangrove, and Wooded Sandbank Vegetation. Then, Non-Forest Natural Formation is divided into Wetland, Grassland Formation, Salt Flat, Rocky Outcrop, Other Non-Forest Formations, and the new class Herbaceous Sandbank Vegetation; Farming into Pasture, Agriculture, Forest Plantation, and Mosaic of Uses; the Non-Vegetated Area into Beach, Dune and Sand Spot, Urban Area, Mining and Other Non-Vegetated Areas; Water into River, Lake and Ocean, and Aquaculture. Agriculture classes go down to Levels 3 and 4. Agriculture was subdivided into Temporary Crops and Perennial Crops. Temporary crop class is divided into Soybean, Sugar cane, Rice, Cotton and Other Temporary Crops. Meanwhile, the Perennial Crop is detailed in the classes Coffee, Citrus, and Other Perennial Crops.

Annex III presents the cross-reference of the MapBiomas LCLU classes with classes from other classification systems (*i.e.*, FAO, IBGE, and National GHG Emissions Inventory). Annex IV presents the classification scheme of the previous collections of MapBiomas.

Table 3. Classes of land cover and land use of MapBiomas Collection 7 in Brazil.

ID	COLLECTION 7 CLASSES	NATURAL/ ANTHROPIC	LAND COVER/ LAND USER	BIOMES/ THEMES
1	**1. Forest**	NATURAL	COVER	-
3	1.1. Forest Formation	NATURAL	COVER	BIOMES
4	1.2. Savanna Formation	NATURAL	COVER	BIOMES
5	1.3. Mangrove	NATURAL	COVER	THEMES
49	1.4. Wooded Sandbank Vegetation	NATURAL	COVER	BIOMES
10	**2. Non Forest Natural Formation**	NATURAL	COVER	-
11	2.1. Wetland	NATURAL	COVER	BIOMES
12	2.2. Grassland Formation	NATURAL	COVER	BIOMES
32	2.3. Salt Flat	NATURAL	COVER	THEMES
29	2.4. Rocky Outcrop	NATURAL	COVER	BIOMES
50	2.5. Herbaceous Sandbank Vegetation	NATURAL	COVER	BIOMES
13	2.6. Other non Forest Formations	NATURAL	COVER	BIOMES
14	**3. Farming**	ANTHROPIC	USE	-
15	3.1. Pasture	ANTHROPIC	USE	THEMES
18	3.2. Agriculture	ANTHROPIC	USE	THEMES
19	3.2.1. Temporary Crop	ANTHROPIC	USE	THEMES
39	3.2.1.1. Soybean	ANTHROPIC	USE	THEMES
20	3.2.1.2. Sugar cane	ANTHROPIC	USE	THEMES
40	3.2.1.3. Rice	ANTHROPIC	USE	THEMES
62	3.2.1.4. cotton	ANTHROPIC	USE	THEMES
41	3.2.1.5. Other Temporary Crops	ANTHROPIC	USE	THEMES
36	3.2.2. Perennial Crop	ANTHROPIC	USE	THEMES
46	3.2.2.1. Coffee	ANTHROPIC	USE	THEMES
47	3.2.2.2. Citrus	ANTHROPIC	USE	THEMES
48	3.2.2.3. Other Perennial Crops	ANTHROPIC	USE	THEMES
9	3.3. Forest Plantation	ANTHROPIC	USE	THEMES
21	3.3. Mosaic of uses	ANTHROPIC	USE	BIOMES
22	**4. Non Vegetated Area**	NATURAL/ ANTHROPIC	COVER/USE	-
23	4.1. Beach, Dune and Sand Spot	NATURAL	COVER	THEMES
24	4.2. Urban Area	ANTHROPIC	USE	THEMES
30	4.3. Mining	ANTHROPIC	USE	THEMES
25	4.4. Other Non Vegetated Areas	NATURAL/	COVER/USE	BIOMES

		ANTHROPIC		
26	**5. Water**	NATURAL/ ANTHROPIC	COVER/USE	-
33	5.1. River, Lake and Ocean	NATURAL	COVER	BIOMES
31	5.2. Aquaculture	ANTHROPIC	USE	THEMES
27	6. Non Observed	NONE	NONE	NONE

3.3.2. Sample collection

Samples for the training and calibration of the random forest classifier were extracted from classes that did not change their values across all years of Collection 6 (stable classes). When necessary, additional samples were collected. The Amazon biome and cross-cutting themes used different sampling designs (see more details in the Appendices).

3.3.3. Classification

Random forest demands the definition of a few classification parameters, such as the number of trees, a list of variables, and training samples. The biomes and cross-cutting appendices detail these parameters, variables, and the number of training samples. Besides random forest, deep learning algorithms were used, such as U-Net (convolutional neural network) in Aquaculture, Mining, Irrigation, Rice, and Citrus classification.

3.4. Post-classification

Due to the pixel-based classification method and the long temporal series, post-classification filters were applied in the biome and cross-cutting maps. The first post-classification action involves a gap fill filter and the application of temporal filters followed by a spatial filter. The application of these filters removes classification noises. These post-classification procedures were implemented in the Google Earth Engine platform and are described in more detail below.

3.4.1. Gap fill

The Gap fill is a temporal filter used to fill possible no-data values. In a long time series of severely cloud-affected regions, it is expected that no-data values may populate some of the resultant median composite pixels. No-data values (“gaps”) are theoretically not allowed in this filter and replaced by the temporally nearest valid classification. The last valid observed class replaces the no-data value in this procedure. If no “previous” class is available, then the no-data value is replaced by the first valid class in the time series. Therefore, gaps should only exist if a given pixel has been permanently classified as no-data throughout the entire temporal domain.

3.4.2. Spatial filter

A spatial filter that removes pixels with isolated values was built based on the "connectedPixelCount" function to avoid unwanted modifications to the edges of the pixel groups (blobs). This function in GEE connects components (neighbors) that share the same pixel value. Thus, only pixels that do not share connections to a predefined number of identical neighbors are considered isolated. This filter needs at least five connected pixels to reach the minimum connection value. Consequently, the minimum mapping unit is directly affected by the spatial filter applied, and it was defined as 5 pixels (~0.5 ha).

3.4.3. Temporal filter

The temporal filter uses sequential classifications in a three-to-five-years unidirectional moving window to identify temporally non-permitted transitions. Based on generic rules (GR), the temporal filter inspects the central position for three to five consecutive years. If the extremities of the successive years are identical, but the central position is not, then the central pixels are reclassified to match their temporal neighbor class. A single central position shall exist for the three years based temporal filter. Two and three central positions are considered for the four and five years filters.

Another generic temporal rule is applied to the extremity of consecutive years. In this case, a three successive years window is used, and if the classifications of the first and last years are different from their neighbors, these values are replaced by the classification of its matching neighbors.

3.4.4. Frequency filter

This filter considers the occurrence frequency throughout the entire time series. Thus, all-natural class occurrences with less than a given percentage of temporal persistence (e.g. 3 years or fewer out of 33) are filtered out. This mechanism reduces the temporal oscillation associated with a given natural class, decreasing the number of false positives and preserving consolidated trajectories. Each biome customized applications of frequency filters; see more details in their respective appendices.

3.4.5. Incident filter

An incident filter was applied to remove pixels that had changed too many times in the 36 years of spam. All pixels that changed more than eight times and are connected to less than 6 pixels were replaced by the MODE value of that given pixel position in the stack of years. This avoids changes in the classes' border and helps stabilize noise pixel trajectories originally. Each biome and cross-cutting themes customized applications of incident filters; see more details in its respective appendices.

3.4.6. Integration

The maps of each biome with the maps of cross-cutting themes were integrated through the hierarchical overlap of each mapped class (Table 4), according to specific prevalence rules. Biomes prevalence rules details are described in the Appendices. The integration process was made on a pixel-by-pixel basis. However, there were specific integration rules as follows:

Table 4. Collection 7 general prevalence rules for integrating biomes and cross-cutting themes maps.

ID	COLLECTION 7 CLASSES	PREVALENCE ID
1	**1. Forest**	-
3	1.1. Forest Formation	18
4	1.2. Savanna Formation	19
5	1.3. Mangrove	3
49	1.4. Wooded Sandbank Vegetation	20
10	**2. Non Forest Natural Formation**	-
11	2.1. Wetland	21
12	2.2. Grassland Formation	22
32	2.3. Hypersaline Tidal Flat	5
29	2.4. Rocky Outcrop	15
50	Herbaceous Sandbank Vegetation	23
13	2.3. Other Non Forest Natural Formations	23
14	**3. Farming**	-
15	3.1. Pasture	24
18	3.2. Agriculture	-
19	3.2.1. Temporary Crop	-
39	3.2.1.1. Soybean	9
20	3.2.1.2. Sugar Cane	8
40	3.2.1.3. Rice	10
62	3.2.1.4. Cotton	11
41	3.2.1.5 Other Temporary Crops	11
36	3.2.2. Perennial Crop	-
46	3.2.2.1. Coffee	12

47	3.2.2.2. Citrus	13
48	3.2.2.3. Other Perennial Crops	14
9	3.3. Forest Plantation	7
21	3.4. Mosaic of Uses	25
22	**4. Non Vegetated Area**	-
23	4.1. Beach, Dune and Sand Spot	2
24	4.2. Urban Area	6
30	4.3. Mining	1
25	4.4. Other Non-Vegetated Areas	16
26	**5. Water**	-
33	5.1. River, Lake and Ocean	17
31	5.2. Aquaculture	4
27	6. Non-Observed	-

3.4.7. Temporal and Spatial Filters on Integrated Maps

A temporal filter was applied in isolated classes over the time series with less than three occurrences in the whole period. After the temporal filter on integrated maps, a spatial filter similar to the one described in 3.4.2 was applied in the integrated maps to remove isolated classes with less than half a hectare as well as noise resulting from eventual Landsat data misregistration.

3.4.8. Transition Maps

The pixel-by-pixel class differences between the maps follow the periods: (A) any consecutive years (e.g. 2001-2002); (B) five-year periods (e.g. 2000-2005); (C) Forest Code period (2008-2021); (D) Forest Code approval (2012-2021); (E) National GHG Inventory (1994-2002; 2002-2010; 2010-2016); (F) all the years (1985-2021). The class transitions represent land use 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 forestry areas, Transitions from forest cover or natural non-forest areas to agriculture or non-vegetated areas, and Areas without transition or transitions that involve unobserved areas or transitions between classes within level 1 of legend.

3.4.9. Spatial Filter on Transition Maps

A spatial filter similar to the one described in 3.4.2 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) stream of up to five pixels with two or one neighbor pixel in the same transition class.

3.4.10. Temporal analysis

In the Collection 7, new temporal analysis besides the transition maps were presented in the MapBiomas platform: stable areas and number of classes.

The stable area tool shows areas that stayed in the same land cover and land use class throughout all the years in the selected temporal extent. This version considers a temporal resolution of one year and allows the user to choose the following temporal extents: (a) The whole temporal extent mapped by MapBiomas, from 1985 to 2021; (b) five years e.g. 1985-1990, 1990-1995, 1995-2000; (c) ten years e.g. 1985-1995, 1995-2005, 2005-2015; (d) half of the temporal extent mapped by MapBiomas, from 1985 to 2000 and from 2000 to 2021.

The number of classes tool shows the number of land cover and land use classes a pixel was classified as during the selected temporal extent. This version considers a temporal resolution of one year and allows the user to choose the following temporal extents: (a) The whole temporal extent mapped by MapBiomas, from 1985 to 2021; (b) five years e.g. 1985-1990, 1990-1995, 1995-2000; (c) ten years e.g. 1985-1995, 1995-2005, 2005-2015; (d) half of the temporal extent mapped by MapBiomas, from 1985 to 2000 and from 2000 to 2021. Results can also be obtained for an individual class and for various legend levels.

3.4.11. 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 (including indigenous lands and conservation units) were included in the zonal statistics. A toolkit in the Google Earth Engine is available to insert user-defined polygons and download the LCLU maps (https://drive.google.com/open?id=1xyGPmsKt14PI1X-bY6pVlAZ6oWtWiDX5).

3.5. Validation Strategies

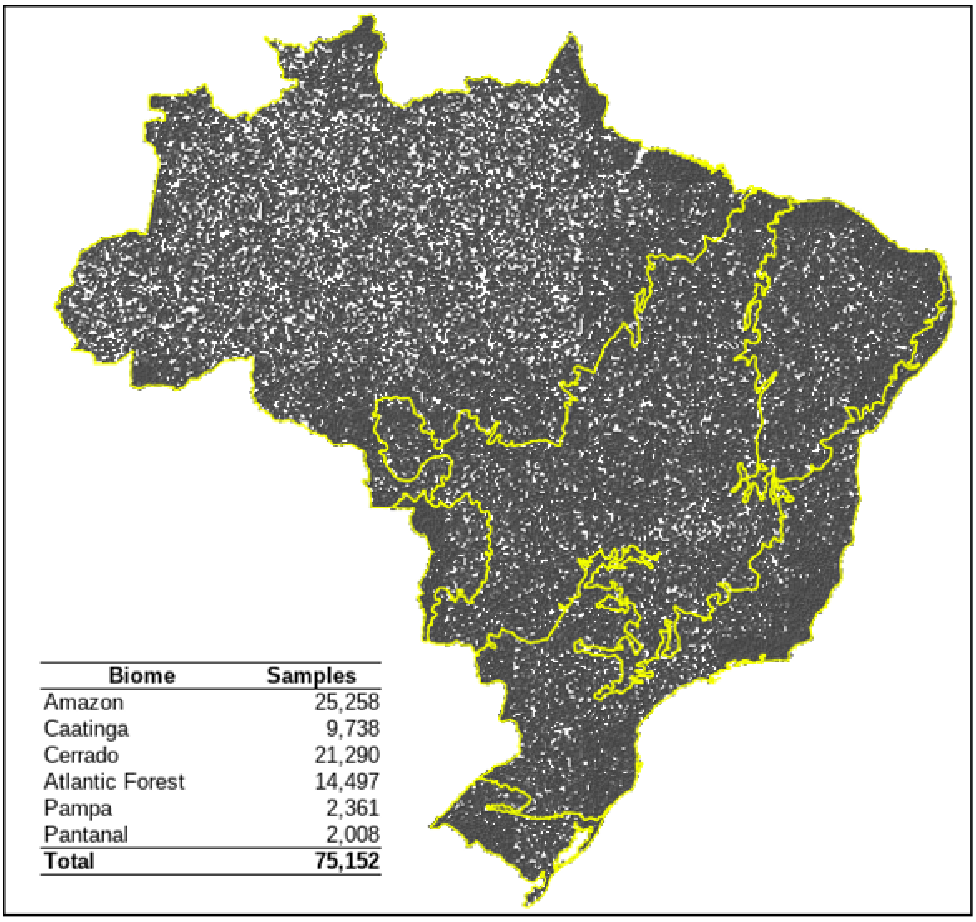
The validation strategy was based on two approaches: (i) comparative analyses with reference maps of specific biomes/regions and years, and (ii) accuracy analyses based on statistical techniques using independent sample points with visual interpretation over entire Brazil and for the entire time series.

3.5.1. Validation with reference maps

Each biome and cross-cutting theme conducted the spatial agreement analyses with reference maps according to their availability (more details are available in the Appendices and on the reference maps webpage https://mapbiomas.org/en/mapas-de-referencia?cama_set_language=en).

3.5.2. Validation with independent points

The examination of almost 85,000 samples from 1985 to 2018 was led by visual interpretation of Landsat data, MODIS-NDVI times series, and high-resolution imagery from Google Earth (when available), for each of the MapBiomas classes. However, 10,000 samples were used as training samples for the Amazon biome. Thus, the error assessment analysis was conducted over ~75,000 samples per year (Figure 3). In Collection 7, the 85,000 independent samples was updated and revised adding the recent years, 2019, 2020, 2021.



Biome	Samples
Amazon	25,258
Caatinga	9,738
Cerrado	21,290
Atlantic Forest	14,497
Pampa	2,361
Pantanal	2,008
Total	**75,152**

Figure 3. Independent random samples were used in the Brazilian biomes for the error assessment analysis of MapBiomas Collections.

The sampling design considered a minimum analysis unit group of four IBGE charts and six slope classes according to SRTM data (Shuttle Radar Topography Mission) (Figure 4 and 5). The accuracy analysis was based on Stehman et al. 2014 and

Stehman & Fody, 2019, using the population error matrix and the global, user, and producer accuracies.

Three independent interpreters inspected each sample; in case of confusion, a senior interpreter decided the final class of the pixel. This evaluation was based on the web platform Temporal Visual Inspection (TVI), developed by LAPIG/UFG. The TVI platform allowed the assessment of all the classes mapped by MapBiomas since Collection 3.1 (https://mapbiomas.org/accuracy-statistics?cama_set_language=en).

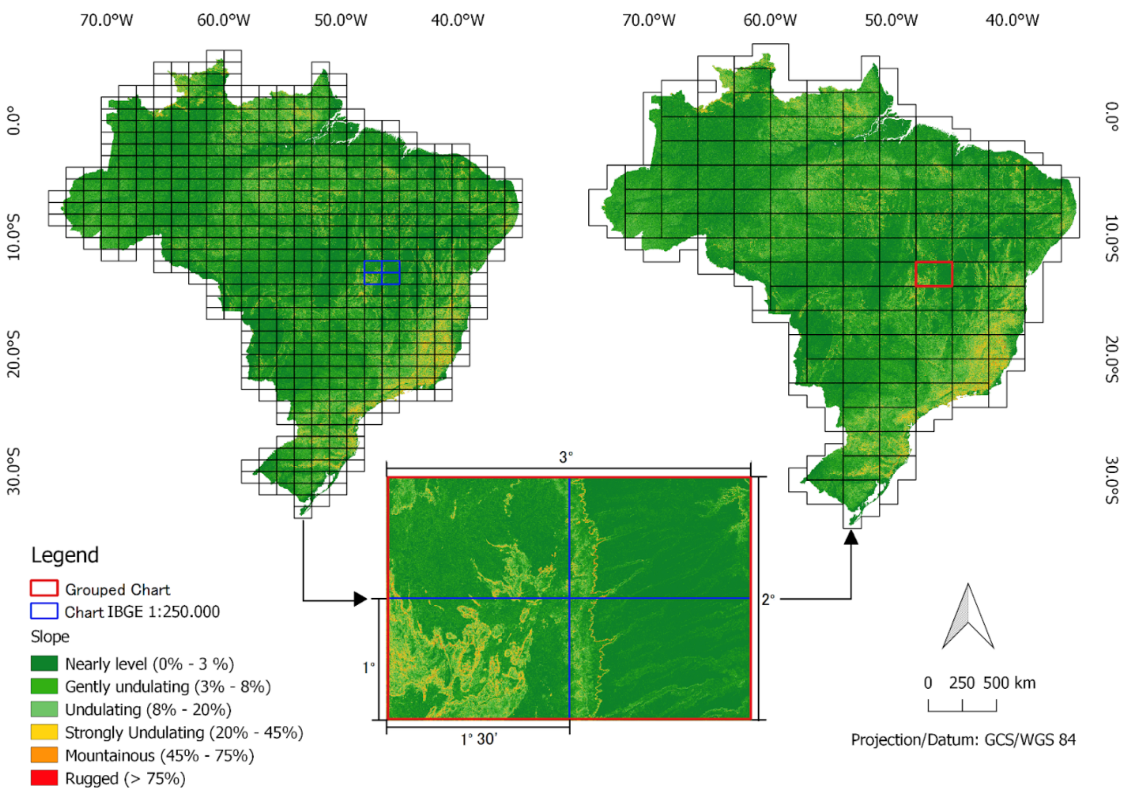


Figure 4. Slope categories used in sampling design.

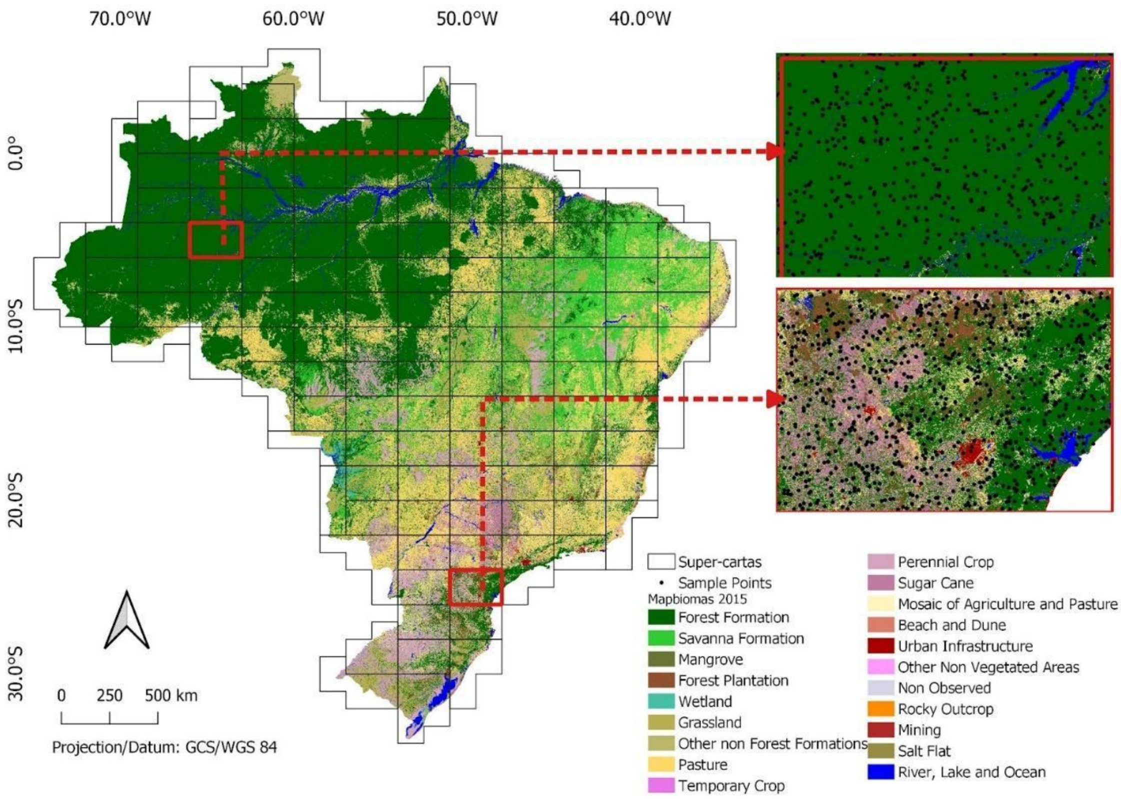


Figure 5. Super-chart size used as subset in the sampling design.

The global accuracy for each level of LCLU classes in the Collection 7 legend was calculated for each year, class, and biome (more details can be explored in the MapBiomas web platform (https://mapbiomas.org/en/accuracy-analysis). In Level 1 classes, the LCLU mapping product in the Collection 7 presented global accuracy and 7.5% allocation disagreement with 1.7% area disagreement. At Level 2, the global accuracy was 87., with 9.3% allocation disagreement and 3.3% area disagreement. Finally, at Level 3, the global accuracy was 87, with 9.2% allocation disagreement and 3.4% area disagreement. The global accuracy was stable over the mapped period, varying across biomes from 81,8% to 96,6% in Level 1.

4. Map Collections and Analysis

The MapBiomas Collections produced so far are listed and summarized below (Table 5):

• Collection 1 - comprised the period of 2008 to 2016 and was based on empirical decision trees for the biomes and Coastal Zone, random forest classification for the Pasture and Agriculture themes. Before launching Collection 1, a Beta Collection was produced to test the methodology used in Collection 1.

• Collection 2 - comprised the period of 2000 to 2016 and was based on empirical decision trees for the biomes and Coastal Zone, random forest classification for the Pasture and Agriculture themes.

• Collection 2.3 - comprised the period of 2000 to 2016 and was based on random forest decision trees for all biomes and the Coastal Zone, Pasture, and Agriculture themes.

• Collection 3 - comprised the period of 1985 to 2017 and was based on random forest decision trees for all biomes and the Coastal Zone, Urban Area, Mining, Pasture, and Agriculture themes. Collection 3.1 was published in Remote Sensing (Souza Jr. et al. 2020).

• Collection 4 - comprised the period of 1985 to 2018 and was based on random forest decision trees for all biomes and the Coastal Zone, Urban Area, Mining, Pasture and Agriculture themes, except the Aquaculture that had its classification based on the U-Net convolutional neural network classifier.

• Collection 5 - comprised the period of 1985 to 2019 and was based on random forest decision trees for all biomes and the Coastal Zone, Urban Area, Mining, Pasture and Agriculture themes, except the Aquaculture and Irrigated Agriculture (central pivot) that the classification was based on the U-Net convolutional neural network classifier.

• Collection 6 - comprised the period of 1985 to 2020 and was based on random forest decision trees for all biomes and the Coastal Zone, Urban Area, Pasture and Agriculture themes, except the Aquaculture, Mining, Irrigation, Rice and Citrus that applied U-Net convolutional neural networks in the classification.

• Collection 7 - comprised the period of 1985 to 2021 and was based on random forest classification for all biomes and the Coastal Zone, Urban Area, Pasture, and Agriculture themes, except the Aquaculture, Mining, Irrigation, Rice and Citrus that applied U-Net convolutional neural networks in the classification.

Table 5. MapBiomas' Collection evolution and its period, number of levels and land cover and land use classes, methods and global accuracy.

Collection	Time interval	Classes	Method	Global Accuracy
Beta & 1	8 years 2008-2015	1 levels / 7 classes	Empirical Decision Tree + Random Forest	n.a.

			(Farming)	
2.0 2.3	16 years 2000-2016	3 levels / 13 classes	Empirical Decision Tree + Random Forest (Farming) All Random Forest (C2.3)	[C2.3] L1 - Gb **79%** Ar 7% All: 14% L2 - Gb **80%** Ar 11% All: 10% L3 - Gb **74%** Ar 12% All: 14%
3.0 & 3.1	33 years 1985-2017	3 levels / 19 classes	Random Forest	L1 - Gb **89%** Ar 7% All: 4% L2 - Gb **84%** Ar 7% All: 9% L3 - n.a.
4.0 & 4.1	34 years 1985-2018	3 levels / 19 classes	Random Forest + U-Net (Aquaculture)	L1 - Gb **90%** Ar 9% All: 1% L2 - Gb **89%** Ar 9% All: 2% L3 - Gb **86%** Ar 11% All: 3%
5.0	35 years 1985-2019	4 levels / 21 classes	Random Forest + U-Net (Aquaculture & irrigation)	L1 - Gb **91%** Ar 7% All: 2% L2 - Gb **90%** Ar 7% All: 3% L3 - Gb **88%** Ar 9% All: 4%
6.0	36 years 1985-2020	4 levels / 25 classes	Random Forest + U-Net (Aquaculture, irrigation, Mining, Rice and Citrus)	L1 - Gb **91%** Ar 7% All: 2% L2 - Gb **87%** Ar 9% All: 3% L3 - Gb **87%** Ar 9% All: 3%
7.0	37 years 1985-2021	4 levels / 27 classes	Random Forest + U-Net (Aquaculture, irrigation, Mining, Rice and Citrus)	L1 - Gb 91% Ar 7% All: 2% L2 - Gb 87% Ar 9% All: 3% L3 - Gb 87% Ar 9% All: 3%

Collection 7 resulted not only in a long time series, adding the year 2021, but more spatially and temporally consistent annual LCLU maps of Brazil. Two new classes were mapped: Herbaceous Sandbank Vegetation (only in the Atlantic Forest, Pampa and Caatinga biomes), totaling 27 mapped LCLU classes. At the same time, the Rocky Outcrop class expanded to the Cerrado biome. Significant improvements were done in Collection 7 by improving the random forest classification, such as smoothing transitions in area variations of each class along the time series. U-Net (convolutional neural network) was expanded not only in the Aquaculture and Irrigation mapping but as well as in the Rice and Citrus classification and Mining mapping detailing. However, challenges remain, and more improvements will be done in the following updated MapBiomas collection. All the programming codes for running the MapBiomas classifications are publicly available and accessible through mapbiomas.org.

6. Concluding Remarks and Perspectives

The proposal algorithms for pre-processing and classifying Landsat imagery hold promise for revolutionizing the production of LCLU maps on a large scale. Thanks to Google Earth Engine and open source technology, it is possible to access and process large-scale satellite imagery datasets such as the one generated by the MapBiomas project. The replication of this type of project is viable for other areas of the planet. The MapBiomas initiative has expanded to other regions such as Amazon, Chaco, Pampa, Atlantic Forest, and Indonesia. In addition, the MapBiomas team will keep improving the following collections in subsequent years (2022 and so long). The open-access MapBiomas LCLU dataset allowed several scientific publications in Brazil and abroad. Policymakers and stakeholders also use the dataset for public policies and decision-makers in the country.

7. References

Breiman, L. Random forests. Machine learning, v. 45, n. 1, p. 5-32, 2001.

Brown, C.F., Brumby, S.P., Guzder-Williams, B. et al. Dynamic World, Near real-time global 10 m land use land cover mapping. Sci Data 9, 251 (2022). https://doi.org/10.1038/s41597-022-01307-4.

Cardoso, M. I. 2009. Projeto Radam: uma saga na Amazônia.

Chen, J.; Chen, J.; Liao, A.; Cao, X.; Chen, L.; Chen, X.; He, C.; Han, G.; Peng, S.; Lu, M.; et al. Global land cover mapping at 30 m resolution: A POK-based operational approach. ISPRS J. Photogramm. Remote Sens. 2014, 103, 7–27.

Friedl, M.A.; Sulla-Menashe, D.; Tan, B.; Schneider, A.; Ramankutty, N.; Sibley, A.; Huang, X. MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. Remote Sens. Environ. 2010, 114, 168–182.FAO, 2012. Manual for integrated field data collection. Rome: FAO. 175p.

Hasenack, H.; Cordeiro, J.L.P; Weber, E.J. (Org.). Uso e cobertura vegetal do Estado do Rio Grande do Sul – situação em 2002. Porto Alegre: UFRGS IB Centro de Ecologia, 2015. 1a ed. ISBN 978-85-63843-15-9. Disponível em: http://www.ecologia.ufrgs.br/labgeo

Hoffmann, G.S.; Weber, E.J.; Hasenack, H. (Org.). Uso e cobertura vegetal do Estado do Rio Grande do Sul – situação em 2015. Porto Alegre: UFRGS IB Centro de Ecologia, 2018. 1a ed. ISBN 978-85-63843-22-7. Disponível em: http://www.ecologia.ufrgs.br/labgeo

IBGE. 2004a. Mapa de biomas do Brasil (escala 1:5.000.000), Rio de Janeiro: IBGE. Mapa e nota técnica.

IBGE, 2004b. Mapa de biomas do Brasil (escala 1:5.000.000), Rio de Janeiro: IBGE.

IBGE, 2010. Uso da terra no Estado do Rio Grande do Sul: relatório técnico. Rio de Janeiro: IBGE. 151 p.

IBGE, 1999. Manual técnico de uso da terra. Rio de Janeiro: IBGE. 58p.

IBGE, 2012. Manual técnico da vegetação brasileira. 2ed. Rio de Janeiro: IBGE. 275p.

de Oliveira G, Araújo MB, Rangel TF, Alagador D, Diniz-Filho JAF. Conserving the Brazilian semiarid (Caatinga) biome under climate change. Biodivers Conserv. 2012; 21: 2913–2926. doi:10.1007/s10531-012-0346-7.

Morellato LPC, Haddad CFB. Introduction: The Brazilian Atlantic Forest. Biotropica. 2000; 32: 786–792. doi:10.1111/j.1744-7429.2000.tb00618.x.

Olofsson P, Foody GM, Herold M, Stehman SV, Woodcock CE, Wulder MA. Good practices for estimating area and assessing accuracy of land change. Remote Sensing of Environment, 2014. 148, pp.42-57.

Roesch LFW, Vieira FCB, Pereira VA, Schünemann AL, Teixeira IF, Senna AJT, et al. The Brazilian Pampa: A fragile biome. Diversity. 2009. pp. 182–198. doi:10.3390/d1020182.

Souza, C.M., Jr.; Z. Shimbo, J.; Rosa, M.R.; Parente, L.L.; A. Alencar, A.; Rudorff, B.F.T.; Hasenack, H.; Matsumoto, M.; G. Ferreira, L.; Souza-Filho, P.W.M.; de Oliveira, S.W.; Rocha, W.F.; Fonseca, A.V.; Marques, C.B.; Diniz, C.G.; Costa, D.; Monteiro, D.; Rosa, E.R.; Vélez-Martin, E.; Weber, E.J.; Lenti, F.E.B.; Paternost, F.F.; Pareyn, F.G.C.; Siqueira, J.V.; Viera, J.L.; Neto, L.C.F.; Saraiva, M.M.; Sales, M.H.; Salgado, M.P.G.; Vasconcelos, R.; Galano, S.; Mesquita, V.V.; Azevedo, T. Reconstructing Three Decades of Land Use and Land Cover Changes in Brazilian Biomes with Landsat Archive and Earth Engine. Remote Sens. 2020, 12, 2735. https://doi.org/10.3390/rs12172735.

Stehman, Stephen V. Sampling designs for accuracy assessment of land cover. International Journal of Remote Sensing, 2019, 30 pp. 5243-5272. doi:10.1080/01431160903131000

Stehman, S. V. Estimating area and map accuracy for stratified random sampling when the strata are different from the map classes. International journal of remote sensing, 2014. 34 pp. 4923-4939. doi:10.1080/01431161.2014.930207

USGS Landsat. USGS Landsat Collection 1 Level 1 Product Definition; USGS Landsat: Sioux Falls, SD, USA, 2017; Volume 26.

Weber, E.J.; Hoffmann, G.S.; Oliveira, C.V.; Hasenack, H. (Org.). Uso e cobertura vegetal do Estado do Rio Grande do Sul – situação em 2009. Porto Alegre: UFRGS IB Centro de Ecologia, 2016. 1a ed. ISBN 978-85-63843-20-3. Disponível em: http://www.ecologia.ufrgs.br/labgeo.

APPENDICES

Appendix 1 - Amazon biome

Appendix 2 - Atlantic Forest biome

Appendix 3 - Caatinga biome

Appendix 4 - Cerrado biome

Appendix 5 - Pampa biome

Appendix 6 - Pantanal biome

Appendix 7 - Agriculture and Forest Plantation

Appendix 8 - Pasture

Appendix 9 - Coastal Zone

Appendix 10 - Mining

Appendix 11 - Urban Area

Appendix 12 - Irrigation

Appendix 13 - Infrastructure layers (Transportation, Energy and Mining)

Appendix 14 - Deforestation and Secondary Vegetation

ANNEXES

Annex I: MapBiomas Network

MapBiomas is an initiative of the Climate Observatory, involving a collaborative network of NGOs, universities and technology companies organized by biomes and cross-cutting themes.

Biomes Coordination:

- Amazon – Institute of People and Environment of the Amazon (IMAZON)
- Caatinga – State University of Feira de Santana (UEFS), Northeast Plants Association (APNE), and Geodatin
- Cerrado – Amazon Environmental Research Institute (IPAM)
- Atlantic Forest – Foundation SOS Atlantic Forest and ArcPlan
- Pampa – Federal University of Rio Grande do Sul (UFRGS) and GeoKarten
- Pantanal – Institute SOS Pantanal and ArcPlan

Cross-cutting Themes Coordination:

- Pasture – Federal University of Goias (LAPIG/UFG)
- Agriculture – Agrosatelite
- Coastal Zone and Mining – Vale Technological Institute (ITV) and Solved
- Urban Area – University of São Paulo (USP - QUAPÁ-FAU and YBY), Federal University of Bahia (UFBA) and Federal University of São Carlos (UFSCar - NEEPC)

Technology Partners:

- Google
- EcoStage
- Terras App

Financing:

- 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)
- Walmart Foundation
- Sequoia Foundation
- Skoll Foundation

Institutional Partners:

- Arapyaú Institute
- WRI Brasil
- AVINA Foundation

General Coordination: Tasso Azevedo (SEEG/OC)

Technical Coordination: Marcos Rosa (ArcPlan)

Scientific Coordination: Julia Shimbo (IPAM)

The project counts on an Independent Committee of Scientific Advice composed by renowned specialists:

- Alexandre Camargo Coutinho (Embrapa)
- Edson Eygi Sano (IBAMA)
- Gilberto Camara Neto (INPE)
- Joberto Veloso de Freitas (Brazilian Forestry Service)
- Matthew C. Hansen (Maryland University)
- Mercedes Bustamante (University of Brasília)
- Timothy Boucher (TNC)
- Robert Gilmore Pontius Jr (Clark University)

Technical Partners:

- Institute of Agricultural and Forest Management and Certification - Imaflora (IMAFLORA)
- Energy and Environment Institute (IEMA)
- Socioambiental Institute (ISA)
- Institute for Democracy and Sustainability (IDS)
- The Nature Conservancy (TNC)
- Life Center Institute (ICV)
- WWF Brasil
- Brasil I.O

Annex II: Mapping initiatives at global scale, in Brazil, biomes and cross-cutting themes, and respectively references/sources in Collection 7.

MAP	SOURCE	DESCRIPTION	DOWNLOAD
Mapa de Geomorfologia do Brasil - 2021	IBGE	Base de dados em meio digital (1: 250000), disponibilizadas para todo o Brasil, com consistência geométrica, estruturada e padronizada para uso em Sistemas de Informação Geográfica sobre o tema da Geomorfologia.	https://geoftp.ibge.gov.br/informacoes_ambientais/geomorfologia/vetores/escala_250_mil/versao_2021/
Agriculture Irrigated by Center Pivots in Brazil	ANA / Embrapa	Mapping of the area and number of central pivot irrigation equipment in Brazil between 1985 and 2017. Study carried out through a partnership between the National Water Agency - ANA and Embrapa Milho e Sorgo.	https://metadados.ana.gov.br/geonetwork/srv/pt/main.home?uuid=e2d38e3f-5e62-41ad-87ab-990490841073
Amazon Mining Map	Social and Environmental Institute - ISA	Mining data in the Brazilian Amazon compiled by Instituto Socioambiental - ISA	https://www.amazoniasocioambiental.org/es/mapas/#descargas
Annual agriculture maps for the Amazon	Agrosatélite	Agrosatellite "Maps that include cotton and corn crops and soybeans, cultivated on a large scale, for the crop years: 2000/2001, 2006/2007 and 2013/2014, 2016/2017, 2019/2020"	https://www.agrosatelite.com.br/static/img/cases/pdf/en/Amazon_Report_2018.pdf
Annual agriculture maps for the Cerrado (project: Geospatial analyzes of the annual crops dynamic in the Brazilian Cerrado biome: 2000 to 2014) and update	Agrosatélite	Maps that include cotton, corn crops and soybeans, cultivated on a large scale, for the crop years: 2000/2001, 2006/2007, 2013/2014. 2018/2019	https://www.agrosatelite.com.br/static/img/cases/pdf/pt-br/Analise_geoespacial_da_dinamica_das_culturas_anuais_no_bioma_Cerrado_2000a2014.pdf
Atlas of the Atlantic Forest	SOS Atlantic Forest/INPE	Mapping of forest formations and associated ecosystems, reference year 2018/2019	http://mapas.sosma.org.br/dados/
Atlas of the Atlantic Forest	SOS Atlantic Forest/INPE	Detailed 1ha mapping for the State of São Paulo	
Atlas of the Atlantic Forest	SOS Atlantic Forest/INPE	Detailed 1ha mapping for the State of Rio de Janeiro	
Atlas of the Atlantic Forest	SOS Atlantic Forest/INPE	Detailed 1ha mapping for the State of Paraná	
Atlas of the Atlantic Forest	SOS Atlantic Forest/INPE	Detailed 1ha mapping for the State of Santa Catarina	

Atlas of the Mangroves of Brazil	MMA / ICMBio	Partnership between the Chico Mendes Institute for Biodiversity Conservation and the Project "Effective Conservation and Sustainable Use of Mangrove Ecosystems in Brazil", implemented by the United Nations Development Program - Brazil (UNDP), with the support of the Global Fund for the Environment Environment (GEF).	https://www.icmbio.gov.br/portal/images/stories/comunicacao/downloads/Mangues.zip
Biomes Boundary Map 1:250,000	IBGE	Together with the methodological report that brings new limits between the six Brazilian biomes, Amazon, Atlantic Forest, Caatinga, Cerrado, Pantanal and Pampa, compatible with the 1:250 000 scale.	https://www.ibge.gov.br/geociencias/informacoes-ambientais/estudos-ambientais/15842-biomas.html?=&t=acesso-ao-produto
Brazil Mineral Resources Map	Geological Survey of Brazil (CPRM) - GeoSGB	The spatial representation of Brazil's mineral resources is maintained by the Geological Service of Brazil (CPRM) through its geoportal called GeoSGB.	https://geoportal.cprm.gov.br/geosgb/
Brazilian Mining Map	Brazil-Germany Chamber of Commerce and Industry / GeoAnsata Projects	Portraits of the mining industry and intermediation of interested parties in the area, experts and producers in the sector with mining companies.	https://www.google.com/maps/d/viewer?mid=19ps2n5FI62X-ib2V2teFhaqcUCbS2BZJ&ll=-14.64391762573763%2C-58.49807411843837&z=4
Canasat	INPE/Agrosatelite	Sugarcane map for south-central Brazil	
CAR Thematic Digital Vector Base - State of Tocantins	Tocantins State Secretariat for the Environment and Water Resources	Prepared from satellite images (Plêiades Year 2015), and constitutes a vectorial, digital base, compatible with a 1:25,000 scale, consisting of features of the hydrographic grid, municipal boundaries, road system grid, transmission lines and special areas of the State of Tocantins, cut and articulated according to sheets 1:25,000 of the National Cartographic System – SCN.	https://www.to.gov.br/semarh/base-vetorial-digital-tematica-do-car/1knojozyng4n
Census Sectors 2020	IBGE	The 2020 Census Sectors was updated for the Demographic Census collection scheduled for 2021. Since it is an Intermediate Mesh it does not contemplate statistical information on population or domicile, but contains the updated situation of the Brazilian Political and Administrative Division - DPA, in effect on 06/30/2020, contained in the Municipal Mesh, also produced annually by the Coordination of Territorial Structures of the Directorate of Geosciences of IBGE.	https://www.ibge.gov.br/geociencias/organizacao-do-territorio/malhas-territoriais/26565-malhas-de-setores-censitarios-divisoes-intramunicipais.html?=&t=o-que-e

Citrus map in SP	Agrosatélite	Translated with www.DeepL.com/Translator (free version)	
Deforestation Alert Map - Mining Class - DETER Project	INPE	Mining Class of the Near Real Time Alert System (DETER), maintained by the National Institute for Space Research (INPE)	http://terrabrasilis.dpi.inpe.br/
Forest Mapping of the State of Sergipe	SEMARH-SE	Survey of Forest Coverage in the State of Sergipe	
Global Distribution of Mangroves USGS	USGS	This dataset shows the global distribution of mangrove forests, derived from earth observation satellite imagery	http://sedac.ciesin.columbia.edu/data/set/lulc-global-mangrove-forests-distribution-2000/data-download
Global Forest Change 2000–2015	University of Maryland	Results from time-series analysis of Landsat images in characterizing global forest extent and change from 2000 through 2015. For additional information about these results, please see the associated journal article (Hansen et al., Science 2013).	https://earthenginepartners.appspot.com/science-2013-global-forest/download_v1.7.html
Global Human Settlement Layer (GHS-BUILT e GHS_BUILT_S2)	European Commission Joint Research Centre (JRC)	A layer of multitemporal information on the presence of built-up area derived from Landsat image collections (GLS1975, GLS1990, GLS2000 and Landsat 8 ad-hoc 2013/2014 collection).	https://ghsl.jrc.ec.europa.eu/download.php
Global Land Analysis Discovery - Commodity Crop Mapping and Monitoring in South America	University of Maryland	Annual Soy Coverage Maps in South America, Soy Coverage 2000 onwards to enable tracking of indirect effects of soy expansion on deforestation and inform multi-stakeholder negotiations on soy sustainability, Corn Coverage Maps in South America South America, updated semi-annually.	https://glad.umd.edu/projects/commodity-crop-mapping-and-monitoring-south-america
Homogeneous Urban Land Use and Occupation Units (UHCT) for the State of São Paulo	Geological Institute (IG) / Department of the Environment of the State of São Paulo	Sectoring or dividing the territory into areas with similar characteristics regarding certain physical aspects, shape and intrinsic texture of the occupation.	http://s.ambiente.sp.gov.br/cpla/UHCT_112015_v2.zip
Index of Roads and Structures (IRS)	Justiniano et. al	Presents a new easily reproducible methodology for urban mapping. The methodology allows for the combined processing of OpenStreeMap and Remote Sensing data, where a metric called the road and structure index (LVI) is proposed for mapping urban areas. The IVE is used with NDVI and MNDWI to map the urban surface with high accuracy, with a reference year of 2020.	https://www.sciencedirect.com/science/article/pii/S0303243422001179?via%3Dihub
III National Inventory of Greenhouse	MCTIC	Map of agricultural areas (annual, semi-perennial and	

Gas Emissions (LULUCF sector)		perennial)	
Land Cover Rondônia 1984 -2010	University of California Santa Barbara (UCSB)	Collaborative research between UCSB, Salisbury University (Maryland) and North Carolina State University to produce and manage an archive of annual Landsat images without gaps (1984-2009) covering most of the state of Rondônia	https://sites.google.com/site/ucsbviperlab/project-definition#TOC-Land-use-and-land-cover-change-in-Rondonia-Brazil
Land Use and Coverage Mapping of the State of Paraná	State Secretariat for Planning and Structuring Projects	Mapping carried out through orbital images of high spatial resolution (2 meter) satellites for the period 2011 to 2016 - WorldView2 and Pleiades 1A and 1B). Supervised automatic classification (GEOBIA).	ftp://200.189.114.112/Mapeamento_Uso_e_Cobertura_da_Terra/
Land Use Map of Brazil	IBGE	Land use map based on Modis (250mt) with comparisons between the years 2000, 2010 and 2012	http://geoftp.ibge.gov.br/informacoes_ambientais/cobertura_e_uso_da_terra/uso_250mil/vetores/
Land use mapping for the Cerrado and Atlantic Forest	FBDS	Land use mapping for the Cerrado and Atlantic Forest. Based on RapidEye high resolution images with 5m resolution.	http://geo.fbds.org.br/
Map of mangroves in northeastern Brazil	Pereira, E.A., Souza-Filho, P.W.M., et al.	Map of mangrove areas from Ponta de Tubarão-MA to the south of the State of Bahia from the classification of Landsat and ALOS PALSAR images from 2008	
Map of mangroves in the northern region of Brazil	Nascimento Jr, W.R; Souza-Filho, P.W.M., et al.	Map of mangrove areas from Oiapoque-AP to Ponta de Tubarão-MA generated from the classification of Landsat and ALOS PALSAR images from 2008	
Mapping of Irrigated Rice in Brazil	National Water Agency (ANA) / National Supply Company (Conab)	Mapping of Irrigated Rice in Brazil	https://metadados.snirh.gov.br/geonetwork/srv/por/catalog.search#/metadata/1ac9b37f-0745-44f9-a60b-6a2bd366bbe1
Mapping of Native Vegetal Coverage and Land Use 1/25,000 in the State of Espírito Santo	State Institute for the Environment and Water Resources (IEMA)	Mapping carried out through orthophotos with 25cm and photointerpretation and manual vectorization of limits between use and coverage classes with a minimum area of 0.5ha.	https://geobases.es.gov.br/downloads
Mapping of the Forest Inventory of the State of São Paulo	Infrastructure and Environment Secretariat of the State of São Paulo - Forestry Institute	Mapping carried out using orbital satellite images of high spatial resolution (0.5 meter), for the period 2017 to 2019, belonging to the collection of the Infrastructure and Environment Department.	
Mapping of the Upper Paraguay Basin	SOS Pantanal/WWF Brasil	Monitoring the use and vegetation cover of the Upper Paraguay Basin, which includes the Pantanal and its headwaters. Data available for 2002, 2008, 2010, 2012, 2014	https://www.sospantanal.org.br/atlas/

		and 2016	
Mapping of the Vegetal Coverage of the Atlantic Forest of Minas Gerais	State Forest Institute (IEF)	Mapping carried out a sweep of 30,673,854.99 hectares, which included the legal limit of the biome established by Federal Law 11,428/2006, plus a five-kilometer buffer, considering the transition areas for the other biomes.	https://geoserver.meioambiente.mg.gov.br/ows?service=WFS&version=1.0.0&request=GetFeature&typeName=WebGis:0301_mg_cobertura_mata_atlantica_2019_lote_2_pol&outputFormat=SHAPE-ZIP
OpenStreetMap	OpenStreetMap Foundation	OpenStreetMap is an initiative to create and provide free geographic data, such as street maps, to anyone.	https://www.openstreetmap.org/
Planted forests map	GFW/WRI	Map of planted forests for Brazil	http://data.globalforestwatch.org/datasets/baae47df61ed4a73a6f54f00cb4207e0_5
Prodes	INPE	Satellite monitoring of clear-cut deforestation in the Legal Amazon and produces, since 1988, the annual deforestation rates in the region, which are used by the Brazilian government to establish public policies.	http://www.dpi.inpe.br/prodesdigital/dadosn/
Substandard Agglomerates 2019	IBGE	This preliminary version incorporates updates through December 2019 and is comprised of 13 152 Subnormal Agglomerations. These agglomerations are located in 734 Municipalities, in all States and in the Federal District, and total 5 127 747 household	https://www.ibge.gov.br/geociencias/organizacao-do-territorio/tipologias-do-territorio/15788-aglomerados-subnormais.html?=&t=saiba-mais-edicao
Synthesis Map of Pastures of Brazil - v8	LAPIG/UFG	Mapping of pasture areas, from TerraClass Amazon data compilation; Funcate; PROBIO; Canasat and TNC.	http://maps.lapig.iesa.ufg.br/?layers=pa_br_areas_pastagens_250_2016_lapig
Terra Class Amazônia	INPE	Mapping the dynamics of use and coverage of the Brazilian Legal Amazon.	http://www.inpe.br/cra/projetos_pesquisas/dados_terraclass.php
Terra Class Cerrado	MMA, IBAMA, EMBRAPA, INPE, UFG e UFU	Cerrado Land Use and Coverage Mapping	http://www.dpi.inpe.br/tccerrado/download.php
Urban Areas in Brazil in 2015	EMBRAPA	Identification, mapping and quantification of urban areas in Brazil that seek to generate a thematic set of geo-referenced information	http://geoinfo.cnpm.embrapa.br/layers/geonode%3Aareas_urbanas_br_15

Use and vegetation cover in the State of Rio Grande do Sul – situation in 2002.	Hasenack, H.; Cordeiro, J.L.P; Weber, E.J. (Org.). Porto Alegre: UFRGS IB Centro de Ecologia, 2015. 1a ed. ISBN 978-85-63843-15-9.	Vegetation Cover Map of Rio Grande do Sul - 2002 base year, obtained by visual interpretation of Landsat images. Level of detail compatible with 1:250,000 scale	https://www.ufrgs.br/labgeo/index.php/uso-e-cobertura-vegetal-do-rio-grande-do-sul-situacao-em-2002/
Use and vegetation cover in the State of Rio Grande do Sul – situation in 2009.	Weber, E.J.; Hofmann, G.S.; Oliveira, C.V.; Hasenack, H. (Org.). Porto Alegre: UFRGS IB Centro de Ecologia, 2016. 1a ed. ISBN 978-85-63843-20-3.	Vegetation Cover Map of Rio Grande do Sul - base year 2009, obtained by visual interpretation of Landsat images. Level of detail compatible with 1:250,000 scale	https://www.ufrgs.br/labgeo/index.php/uso-e-cobertura-vegetal-do-rio-grande-do-sul-situacao-em-2009/
Use and vegetation cover in the State of Rio Grande do Sul – situation in 2015.	Hofmann, G.S.; Weber, E.J.; Hasenack, H. (Org.). Porto Alegre: UFRGS IB Centro de Ecologia, 2018. 1a ed. ISBN 978-85-63843-22-7.	Vegetation cover map of Rio Grande do Sul - base year 2015, obtained by visual interpretation of Landsat images. Level of detail compatible with 1:250,000 scale	https://www.ufrgs.br/labgeo/index.php/downloads/dados-geoespaciais/uso-e-cobertura-vegetal-do-rio-grande-do-sul-situacao-em-2015/
Vegetal Coverage Maps of Brazilian Biomes - ProBio	UEFS/APNE/EMBRAPA-Solos/UFCE/UFRN/UFRPE/UFPB/CRA/SEMARH-MMA	Survey of vegetable coverage and land use of the Caatinga Bioma	http://mapas.mma.gov.br/geodados/brasil/vegetacao/vegetacao2002/mosaicos_vegetacao/caatinga.zip
Mapping of the Evolution of Vegetation Cover.	Bahia Forestry Forum	The objective of the independent monitoring of the vegetation cover and soil use of the Discovery Coast, Extremo Sul, and South Coast Parcel (areas where Suzano and Veracel operate), is to map the evolution of the vegetation cover.	https://forumflorestalbahia-worldresources.hub.arcgis.com/
Cofee Map	CONAB	The mapping of agricultural crops is carried out by means of remote sensing. It aims to contribute to the estimation of area and productivity, offering precise information about the geographic distribution in each state. In area estimation, the mapping result helps in the analysis of the declared information, as a verifiable data in the field. In the productivity estimation, the knowledge of the location of the cultivation areas enables the monitoring of the productive areas through agrometeorological parameters, offering indicatives on the forecast of the yield of the crops. The mappings are available for download in shapefile format.	https://www.conab.gov.br/info-agro/safras/mapeamentos-agricolas

VIIRS Stray Light Corrected Nighttime Day/Night Band Composites Version 1	NASA		https://eogdata.mines.edu/download_dnb_composites.html
WorldPop Project Population Data: Estimated Residential Population	University of Southampton	Translated with www.DeepL.com/Translator (free version)	https://www.worldpop.org/project/categories?id=18
Illegal mineral activity map of RAISG	RAISG	Data of illegal mining of the Amazon Socio-environmental Georeference Information Network.	http://www.amazoniasocioambiental.org
Amazon Mining Watch Network Mining Map	AMW	Amazon Mining Watch Mining Activity Data	https://amazonminingwatch.org/

Annex III: Cross-reference of MapBiomas land use/land cover classes in the Collection 7 with FAO, IBGE and National GHG Emissions Inventory classes.

Level 1	Level 2	Level 3	Level 4	Biome	Brief description	IBGE (1999; 2012) Classification	FAO (2012) Classification	National Inventory of GHG Emissions (2015) Classification
Forest	Natural Forest Formation	Forest Formation		Amazon	Dense Ombrophilous Forest, Evergreen Seasonal Forest, Open Ombrophilous Forest, Semi-deciduous Seasonal Forest, Deciduous Seasonal Forest, Wooded Savannah, Areas that had fire or logging, Forest resulting from natural successional processes, after total or partial primary vegetation suppression by anthropogenic actions or natural causes, which may have remaining trees of primary vegetation. Alluvial Open Ombrophilous Forest established along the watercourses, occurring in periodically or permanently flooded plains and terraces, where in the Amazon represent the physiognomies of igapó and lowland forests, respectively. Bamboo florest (Acre).	Da, Db, Ds, Dm, Ha, Hb, Hs, Ld, La, Aa, Ab, As, Am, Fa, Fb, Fs, Fm, Ca, Cb, Cs, Cm, Vsp	FDP, FEP, FSP, FEM, FDM, FSM	FMN, FM, FSec
				Caatinga	Vegetation types with predominance of continuous canopy - Wooded Steppe Savanna, Semi-deciduous and Deciduous Seasonal Forest.	Td, Cs, Cm, Fm, Fs, Pa, As, Fb, Pf, Pm, Fa, Cb, Ds, Am, Ab, Sd	FEP, FSP	FMN, FM
				Cerrado	Vegetation types with predominance of tree species, with continuous canopy formation (Riparian Forest, Gallery Forest, Dry Forest and Forested Savanna) (Ribeiro & Walter, 2008), as well as Semi-deciduous Seasonal Forests.	Aa, Ab, As, Cb, Cm, Cs, Da, Dm, Ds, F, Ml, Mm, P, Sd, Td	FEP, FDP, FSP	FMN, FM

			Atlantic Forest	Dense, Open and Mixed Ombrophilous Forest, Semi-deciduous and Deciduous Seasonal Forest, and Pioneer Formation.	D, A, M, F, C, Pma	FEP, FSP	FMN, FM
			Pampa	Vegetation with predominance of tree species and continuous canopy. It includes forest typologies: Ombrophilous, Semi-deciduous and Deciduous and part of the pioneer formations.	Da, Db, Ds, Dm, Ma, Ms, Mm, Ml, Fa, Fb, Fs, Fm, Ca, Cb, Cs, Cm, P, Pa, Pm	FEP, FDP, FSP	FMN, FM, FSec, CS
			Pantanal	Tall trees and shrubs in the lower stratum: Deciduous and Semi-deciduous Seasonal Forest, Wooded Savanna, Wooded Steppe Savanna, and Fluvial and/or Lacustre Influenced Pioneer Formations.	Ca, Cb, Cs, Fa, Fb, Fs, SN, Sd, Td, Pa	FEP, FSP	FMN, FM
		Savanna Formation	Amazon	Open plant formation with a more or less developed shrub and/or arboreal layer, herbaceous layer always present.	Sa, Ta	WS	FMN, FM
			Caatinga	Vegetation types with predominance of semi-continuous canopy species - Wooded Steppe Savanna and Wooded Savanna.	Ta, Sa,	FDP	FMN, FM
			Cerrado	Savanna formations with defined tree and shrub-herbaceous stratum (Cerrado Stricto Sensu: Dense, Typical, Sparse and Rupestrian Savanna).	Sa, Ta	WS	FMN, FM
			Atlantic Forest	Steppe, Forested and Wooded Savannas.	Sd, Td, Sa, Ta	FDP, FSP, WS	FMN, FM
			Pantanal	Small tree species, sparsely arranged in the shrub and herbaceous continuous vegetation. The herbaceous vegetation mixes with erect and decumbent shrubs.	Sa, Sp, Sg, Td, Ta, Tp	FDP, FSP, WS	FMN, FM

		Mangrove		Dense and Evergreen Forest formations, often flooded by tide and associated with the mangrove coastal ecosystem.	Pf	FEP, FEM	FMN, FM
		Wooded Restinga	Atlantic Forest	Forest formations on sandy soils in the coastal region.	Pma	FEP, FEM	FMN, FM
Non-Forest Natural Formation	Wetland		Amazônia	Lowland or grassland vegetation that suffers fluvial and/or lacustrine influence.	Pa	OM	GNM, GM, GSec
			Cerrado	Vegetation with a predominance of herbaceous strata subject to seasonal flooding (e.g. Campo Umido) or under fluvial/lacustrine influence (e.g. Brejo). In some regions, the herbaceous matrix is associated with arboreal species of savanna formation (e.g. Parque de Cerrado) or palm trees (Vereda, Palmeiral).	Pa, Sp	OM	GNM, GM, GSec
			Atlantic Forest	Wetlands with fluvial influence.	Pa	OM	GNM, GM, GSec
			Pampa	Wetland areas, regionally called banhados or marshes (saline influence). Vegetation typically hygrophilous, with aquatic plants emerging, submerged or floating. They occupy plains and depressions of the terrain with waterlogged soil and also the shallow edges of ponds or water reservoirs.	P, Pa, Pm	OM	A, Res
			Pantanal	Herbaceous vegetation with a predominance of grasses subject to permanent or temporary flooding (at least once a year) according to the natural flood pulses. The woody element can be present on the country matrix forming a mosaic with shrub or tree plants (e.g.: cambarazal, paratudal and carandazal). Swampy areas generally occur on the banks of temporary or permanent lagoons occupied by emergent, submerged or floating aquatic plants (eg swamps and barns). Areas with a water surface, but difficult to classify due to the amount of macrophytes, eutrophication or sediments, were also included in this category.	Tg, Sp, Pa, Tp	OM	GNM, GM, GSec

	Grassland Formation	Amazon	Savanna, Park Savana (Marajó), Steppe-Savana (Roraima), Grassland Savanna, Campinarana, for regions outside the Amazon/Cerrado Ecotone. And for regions within the Amazon/Cerrado Ecotone, predominance of herbaceous strata.	Sa, Sp, Sg, Ta, Tp, Tg	WG, OG, WS	GNM, GM, GSec
		Caatinga	Vegetation type with predominance of herbaceous species (Park Steppe Savanna, Herbaceous-Woody Steppe Savanna, Park Savanna, Herbaceous-Woody Savanna) + (Flooded areas with an interconnected pond net, located along waterways and in lowlands areas that accumulate water, vegetation predominantly herbaceous to shrub).	Tp, Sg, Rm, Sp, Tg, Rl	WG, OG, WS	GNM, GM, GSec
		Cerrado	Grassland formations with a predominance of herbaceous strata (dirty, clean and rupestrian fields) and some areas of savanna formations such as the rupestrian cerrado.	Sg, Tp, Tg	WG, OG	GNM, GM, GSec
		Atlantic Forest	Savannas, Park and Grassland Steppe Savannas, Steppe and Shrub and Herbaceous Pioneers.	Sp, Sg, Tp, Tg, E, Pa	WS,OG	GNM, GM, GSec
		Pampa	Vegetation with a predominance of grassy strata, with the presence of herbaceous and sub-shrub dicots. The botanical composition is influenced by edaphic and topographic gradients and by pasture management (livestock). They occur in deep to shallow soils, including rocky (rupestrian fields) and sandy (sandy or psamophilic fields). Occupy well-drained soil (mesic fields) to soils with higher moisture content (wet fields - with a strong presence of sedges). In most cases, it corresponds to native vegetation, but patches of exotic invasive vegetation or exotic forage use (planted pasture) may be present.	E, Ea, Ep, Eg, T, Ta, Tp, P, Pa, Pm	WG, OG	GNM, GM, GSec
		Pantanal	Vegetation with a predominance of grassy stratum, with the presence of isolated and stunted woody shrubs. The botanical composition is influenced by the edaphic and topographical gradients and pasture management (livestock). Patches of invasive exotic vegetation or forage use (planted pasture) may be present forming mosaics with native vegetation.	Sg, Sp, Ta, Tg	WG, OG	GNM, GM, GSec

	Salt flat		"Apicuns" or Salt flats are formations often without tree vegetation, associated to a higher, hypersaline and less flooded area in the mangrove, generally in the transition between this area and the continent.	Pf, Pfh	OM, OX	O
	Rocky Outcrop	Caatinga	Naturally exposed rocks without soil cover, often with the partial presence of rupicolous vegetation and high slope.	Ar	OX	ArM, ArNM
		Cerrado	Naturally exposed rocks without soil cover, often with the partial presence of rupicolous vegetation and high slope.	Ar	OX	ArM, ArNM
		Atlantic Forest	Naturally exposed rocks without soil cover, often with the partial presence of rupicolous vegetation and high slope.	Ar	OX	ArM, ArNM
		Pampa	Naturally exposed rocks without soil cover, often with the partial presence of rupicolous vegetation and high slope.	Ar	OX	ArM, ArNM
	Herbaceous Sandbank Vegetation	Atlantic Forest	Vegetação herbácea com influência fluviomarinha.	Pmb, Pmh	WG, OG	GNM, GM
		Caatinga	Vegetação herbácea com influência fluviomarinha.	Pmb, Pmh	WG, OG	GNM, GM
		Pampa	Herbaceous vegetation that is established on sandy soils or on dunes in the coastal zone.	Pmb, Pmh	WG, OG	GNM, GM
	Other Non-Forest Formations	Atlantic Forest	Marshes (with fluvio-marine influence).	Pfh, Pmb, Pmh	WG, OG	GNM, GM, GSec
Farming	Pasture		Pasture area, predominantly planted, linked livestock production activities. Areas of natural pasture are predominantly classified as grassland formation that may or may not be grazed.	AP, PE, PS	OP, OG	Ap

	Agriculture	Temporary Crop	Soybean	Areas cultivated with soybean.	AMc (s)	OCA	AC
			Sugar cane	Cultivated areas with sugar cane.	AMc (c)	OCA	AC
			Rice	Areas cultivated with rice, exclusively under irrigation, in the states of Rio Grande do Sul, Tocantins, Santa Catarina and Coast of Paraná. This map is the same one presented in the irrigation module in the "Irrigated Rice" class.	AMc	OCA	AC
			Cotton (BETA version)	Areas cultivated with cotton cultivation.	AMc (s)	OCA	AC
			Other Temporary Crops	Areas occupied with short or medium-term agricultural crops, generally with a vegetative cycle of less than one year, which after harvesting need to be planted again to produce.	AMc	OCA	AC
		Perennial Crop	Coffee	Areas cultivated with coffee plantation.	AMp (c)	OCP	PER
			Citrus	Areas cultivated with citrus cultivation.	AMp	OCP	PER
			Other Perennial Crops	Areas occupied with agricultural crops with a long vegetative cycle (more than one year), which allow successive harvests, without the need for new planting. In this version, the map covers mostly cashew areas on the northeast coast and oil palm in the northeast region of Pará, but without distinction between them.	AMp	OCP	PER
	Forest Plantation			Tree species planted for commercial purposes (e.g. pinus, eucalyptus, araucaria).	R	FPB, FPC, FPM	Ref
	Mosaic of Uses		Caatinga	Farming areas where it was not possible to distinguish between pasture and agriculture.	AP, PE, PS, ATp, ATc, ATpc	OCA, OCM, OP, OG	AC, PER, Ap, APD
			Cerrado	Areas of agricultural use where it was not possible to distinguish between pasture and agriculture.	AP, PE, PS, ATp, ATc, ATpc	OCA, OCM, OP, OG	AC, PER
			Atlantic Forest	Farming areas where it was not possible to distinguish between pasture and agriculture.	AP, PE, PS, ATp, ATc, ATpc	OCA, OCM, OP, OG	AC, PER

		Pampa	Areas of agricultural use, where it was not possible to distinguish between pasture and agriculture. It may include cropland, winter or summer pasture and horticulture. Includes rest areas between agricultural crops (fallow).	AP, AS, AT, AM, PE, PS, Ag, Ap, Ac, Acc, Acp, AA	OCA, OCM, OP, OG, OF	AC, PER, Ap, APD
		Urban Areas	Areas of urban vegetation, including cultivated vegetation and natural forest and non-forest vegetation.		OB	S
Non Vegetated Area	Beach, Dune and Sandy		Sandy areas, with bright white color, where there is no vegetation predominance of any kind.	Dn	OX	DnM,DnNM
	Urban Area		Urban areas with predominance of non-vegetated surfaces, including roads, highways and constructions.		OB	S
	Mining		Areas where clear signs of extensive mineral extractions are present, shows clear exposure of the soil by the action of heavy machinery. Only regions surrounding the AhkBrasilien (AHK) and the CPRM digital reference data were considered.	MCA	OQ	Min
	Other Non Vegetated Areas	Amazon	Non-permeable surface areas (infrastructure, urban expansion or mining) not mapped into their classes	AU, MCA	OB, OQ	S, Min
		Caatinga	Non-permeable surface areas (infrastructure, urban expansion or mining) not mapped into their classes	AU,MCA	OB, OQ	S, Min
		Cerrado	Non-permeable surface areas (infrastructure, urban expansion or mining) not mapped into their classes	AU, MCA	OB, OQ	S, Min
		Atlantic Forest	Non-permeable surface areas (infrastructure, urban expansion or mining) not mapped into their classes and regions of exposed soil in natural or crop areas	AU, MCA	OB, OQ	S, Min
		Pampa	Mixed class that includes natural and anthropic areas. Natural areas include exposed sandy surfaces including mainly river and sandy beaches. Anthropic	AU, MCA, Dn, Iu	OB, OQ, OX	S, SE, DnM, DnNM, Min

			areas include non-permeable surfaces (roads and infrastructure for rural developments).			
		Pantanal	Exposed soil areas (mainly sandy soil) not classified as Grassland Formation or Pasture.	PE, Sg	OX	Ap, GNM, GSec
Water	River, Lake and Ocean		Rivers, lakes, dams, reservoir and other water bodies		IRP, IRS, IL, ID	A, Res
	Aquaculture		Artificial lakes, where aquaculture and/or salt production activities predominate			
Not Observed			Areas blocked by clouds or atmospheric noise, or with absence of ground observation masked out from analysis.			NO

References: Instituto Brasileiro de Geografia e Estatística - IBGE. Manual técnico de uso da terra, IBGE: Rio de Janeiro, Brazil, 1999, 58p.; Instituto Brasileiro de Geografia e Estatistica - IBGE. Manual técnico da vegetação brasileira, 2nd ed., IBGE: Rio de Janeiro, Brazil, 2012. pp.157-160; Food and Agriculture Organization of the United Nations - FAO. Manual for integrated field data collection. FAO: Rome, Italy, 2012, 175p.; Ministério da Ciência, Tecnologia e Inovações. Secretaria de Pesquisa e Formação Científica. Quarta Comunicação Nacional do Brasil à Convenção-Quadro das Nações Unidas sobre Mudança do Clima, Brasília, 2020, 620p.

Annex IV: Classes of land cover and land use of Collections 1, 2, 2.3, 3, 4, 5 and 6 of MapBiomas.

Collection 1

- Forest
- Forest in Costal Zone
- Planted Forest
- Agriculture
- Pasture
- Water
- Other
- Non-Observed

7 Classes

Collection 2

- 1. Forest
  - 1.1. Natural Forest
    - 1.1.1. Natural Forest Formation
    - 1.1.2. Savanna Formation
    - 1.1.3. Mangrove
  - 1.2. Forest Plantations
- 2. Non-Forest Natural Formations
  - 2.1. Non-forest Natural Wetlands
  - 2.2. Grasslands
- 3. Farming
  - 3.1. Pasture
  - 3.2. Agriculture
  - 3.3 Agriculture or Pasture
- 4. Non-Vegetated areas
  - 4.1. Beach and dune
  - 4.3. Other non-vegetated areas
  - 4.2. Urban Infrastructure
- 5. Water
- 6. Non-Observed

13 Classes

Collection 3

- 1. Forest
  - 1.1. Natural Forest
    - 1.1.1. Forest Formation
    - 1.1.2. Savanna Formation
    - 1.1.3. Mangrove
  - 1.2. Forest Plantation
- 2. Non Forest Natural Formation
  - 2.1. Wetland
  - 2.2. Grassland
  - 2.3. Salt flat
  - 2.3. Other non forest natural formation
- 3. Farming
  - 3.1. Pasture
  - 3.2. Agriculture
    - 3.2.1. Annual and Perennial Crop
    - 3.2.2. Semi-perennial Crop
  - 3.3. Mosaic of Agriculture and Pasture
- 4. Non vegetated area
  - 4.1. Beach and Dune
  - 4.2. Urban Infrastructure
  - 4.3. Rocky outcrop
  - 4.4. Mining
  - 4.5. Other non vegetated area
- 5. Water
  - 5.1. River, Lake and Ocean
  - 5.2. Aquaculture
- 6. Non Observed

19 Classes

Collection 4

- 1. Forest
  - 1.1. Natural Forest
    - 1.1.1. Forest Formation
    - 1.1.2. Savanna Formation
    - 1.1.3. Mangrove
  - 1.2. Forest Plantation
- 2. Non Forest Natural Formation
  - 2.1. Wetland
  - 2.2. Grassland
  - 2.3. Salt flat
  - 2.4. Rocky outcrop
  - 2.5. Other non forest natural formation
- 3. Farming
  - 3.1. Pasture
  - 3.2. Agriculture
    - 3.2.1. Annual and Perennial Crop
    - 3.2.2. Semi-perennial Crop
  - 3.3. Mosaic of Agriculture and Pasture
- 4. Non vegetated area
  - 4.1. Beach and Dune
  - 4.2. Urban Infrastructure
  - 4.3. Mining
  - 4.4. Other non vegetated area
- 5. Water
  - 5.1. River, Lake and Ocean
  - 5.2. Aquaculture
- 6. Non Observed

19 Classes

Collection 5

- 1. Forest
  - 1.1. Natural Forest
    - 1.1.1. Forest Formation
    - 1.1.2. Savanna Formation
    - 1.1.3. Mangrove
  - 1.2. Forest Plantation
- 2. Non Forest Natural Formation
  - 2.1. Flooded Grassland and Swamped Area
  - 2.2. Grassland
  - 2.3. Salt flat
  - 2.4. Rocky outcrop
  - 2.5. Other non forest natural formation
- 3. Farming
  - 3.1. Pasture
  - 3.2. Agriculture
    - 3.2.1. Annual Crop
      - 3.2.1.1. Soy bean
      - 3.2.1.2. Sugar Cane
      - 3.2.1.3. Other annual crops
    - 3.2.2. Perennial Crop
  - 3.3. Mosaic of Agriculture and Pasture
- 4. Non vegetated area
  - 4.1. Beach and Dune
  - 4.2. Urban Infrastructure
  - 4.3. Mining
  - 4.4. Other non vegetated area
- 5. Water
  - 5.1. River, Lake and Ocean
  - 5.2. Aquaculture
- 6. Non Observed

21 Classes

Collection 6

- 1. Forest
  - 1.1. Forest Formation
  - 1.2. Savanna Formation
  - 1.3. Mangrove
  - 1.5. Wooded Restinga
- 2. Non Forest Natural Formation
  - 2.1. Flooded Grassland and Swamped Area
  - 2.2. Grassland
  - 2.3. Salt Flat
  - 2.4. Rocky Outcrop
  - 2.6. Other non Forest Formations
- 3. Farming
  - 3.1. Pasture
  - 3.2. Agriculture
    - 3.2.1. Temporary Crop
      - 3.2.1.1. Soybean
      - 3.2.1.2. Sugar cane
      - 3.2.1.3. Rice
      - 3.2.1.4. Other Temporary Crops
    - 3.2.2. Perennial Crop
      - 3.2.1.1. Coffee
      - 3.2.1.2. Citrus
      - 3.2.1.3. Other Perennial Crops
  - 3.3. Forest Plantation
  - 3.4. Mosaic of Agriculture and Pasture
- 4. Non vegetated area
  - 4.1. Beach, Dune and Sand Spot
  - 4.2. Urban Area
  - 4.3. Mining
  - 4.4. Other non Vegetated Areas
- 5. Water
  - 5.1. River, Lake and Ocean
  - 5.2. Aquaculture

25 Classes

MAPBIOMAS