



MapBiomass General “Handbook”

Algorithm Theoretical Basis Document (ATBD)

Collection 8

Version 1.0

August, 2023

Table of Contents

Executive Summary	3
1. Introduction	5
1.1. Scope and content of the document	5
1.2. Overview	5
1.3. Region of Interest	6
1.4. Key Science Applications	8
2. Overview and Background Information	9
2.1 Context and Key Information	9
2.1.1. MapBiomass Network	9
2.1.2. Remote Sensing Data	11
2.1.3 Google Earth Engine and MapBiomass Computer Applications	11
2.2 Historical Perspective: Existent Maps and Mapping Initiatives	12
2.2.1. International land cover and land use data	13
2.2.2. National land cover and land use data	13
2.2.3. Regional and biomes land cover and land use data	14
3. Methodological description	14
3.1. Landsat Mosaics	16
3.2. MapBiomass feature space	16
3.3. Classification	18
3.3.1. Legend	18
3.3.2. Sample collection	20
3.3.3. Classification	20
3.4. Post-classification	20
3.4.1. Gap fill	20
3.4.2. Spatial filter	21
3.4.3. Temporal filter	21
3.4.4. Frequency filter	21

3.4.5. Incident filter	21
3.4.6 Integration	22
3.4.7 Temporal and Spatial Filters on Integrated Maps	23
3.4.8. Transition Maps	23
3.4.9. Spatial Filter on Transition Maps	24
3.4.10. Statistics	23
3.5. Validation Strategies	24
3.5.1. Validation with reference maps	24
3.5.2. Validation with independent points	25
4. Map Collections and Analysis	27
5. Concluding Remarks and Perspectives	30
6. References	31

Executive Summary

MapBiomass 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 2022 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, MapBiomass has produced eight 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, 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 MapBiomass 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. MapBiomass Water classification was incorporated in the Water class. The Mining class was divided into industrial mining and artisanal (*garimpo*) mining, as well as categorizing the main substance being exploited.

The Collection 7 presented two additional 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, and 2021.

The current Collection 8 has the same 27 LCLU classes mapped in the previous collection and includes two additional classes: Floodable Forest and Palm Oil, totaling 29 classes, besides adding 2022 and other improvements in the land cover and land use classification.

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 8 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 MapBiomass Collection 8 and describe the datasets and analysis. All the MapBiomass maps and datasets are freely available on the project platform (<http://mapbiomas.org>), as well as all computational algorithms used in the MapBiomass 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 2022 of the MapBiomias Collection 8.

This document covers the classification methods of Collection 8, 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 MapBiomias Github (<https://github.com/mapbiomas-brazil>).

1.2. Overview

The MapBiomias 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 MapBiomias mapping evolved year by year and was divided into Collections (more about MapBiomias' 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)
- Collection 8: 1985 through 2022 (launched in August 2023)

MapBiomias 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 MapBiomias 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 datasets; iii) MapBiomias collaborative network of organizations and experts that share knowledge and mapping tools; and iv) visionary funding agencies that support the project (Souza Jr et al., 2020).

The products of the MapBiomias Collection 8 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 Area);
- 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 qualified according to the number of available observations, varying from 0 to 23 observations per year. The quality assessment of the Landsat mosaics is available on the MapBiomias website.
- LCLU data around buffers of infrastructure (transportation, energy, mining, agribusiness and telecommunication).
- Annual and accumulated deforestation maps (from 1987 to 2021).
- Secondary vegetation annual maps (from 1987 to 2022) and the age of secondary vegetation map.
- Irrigation maps (for center pivot irrigation systems, irrigated rice and other irrigation systems). There is also the pivot dynamics data, that provides information on the dynamics of irrigated crops on central pivots. This information includes: 1. number of times each pivot was cultivated (in a crop year); 2. beginning and end of each crop cycle, per pivot and per crop year; 3. duration of each crop cycle; 4. daily average rainfall of each crop cycle. This product is a beta version and for the 2016 to 2022 period.
- Pasture vigor condition annual maps (only from 2000 to 2022).
- Industrial and artisanal (*garimpo*) mining annual maps and its main substances.
- Temporal analysis (number of classes, stable areas and number of classes).

Besides these products, the MapBiomias network released MapBiomias Water Collection 2 and MapBiomias Fire Collection 2 with annual and monthly mapping of Brazil's water surface and fire scars from 1985 to 2022, respectively. Annual maps of topsoil (0 - 30 cm) organic carbon stocks from 1985 to 2021 were launched as part of the MapBiomias Soil Beta Collection along with the open soil research data repository SoilData.

The MapBiomias initiative also expanded in other regions and countries, such as the Chaco region with its Collection 4 (<https://chaco.mapbiomas.org/>), the Amazon region with its Collection 4 (<https://amazonia.mapbiomas.org/>), Collection 2 MapBiomias Atlantic Forest (<https://bosqueatlantico.mapbiomas.org/>), Collection 2 South American Pampa (<https://pampa.mapbiomas.org/>), and most recently in Indonesia (<https://mapbiomas.nusantara.earth/>), Peru (<https://peru.mapbiomas.org/>) and Bolivia (<https://bolivia.mapbiomas.org/>) with Collection 1. These new project areas follow the mapping protocol of MapBiomias Brazil with a adjustments considering

the peculiarities of their ecosystems and local teams. Detailed methodological information about these MapBiomias initiatives is found at the ATBD of these regions.

1.3. Region of Interest

MapBiomias 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). A biome is a geographic region defined based on vegetation types associated with geomorphological and climatic conditions. 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 Area.

For the first MapBiomias collections, a 1:1,000,000 map of the biomes was produced 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. Since Collection 5, the new Brazilian biomes official map (1:250,000) of IBGE (2019) has been used.



Figure 1. Brazilian biomes were mapped in the MapBiomas project to generate the Collection 8 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 forest area of the biome 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

MapBiomass 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

The MapBiomass Award was created in 2019 to stimulate and expand technical and scientific applications that use any MapBiomass initiative and product data, including initiatives from other countries. The award is now in the fifth edition,

presenting six categories: General, Young People, Applications in Public Policy, Applications in Business and two new categories: Applications in Schools and Actions to Combat Deforestation (<https://mapbiomas.org/en/premio>).

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 (<https://www.oc.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 Area) 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 Goiás (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 MapBiomias. The Federal University of Technology - Paraná oversees the data repositories of the MapBiomias Network. Google provides the cloud computing infrastructure that allows data processing and analysis through Google Earth Engine and storage through Google Cloud Storage.

The funding for the MapBiomias 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 MapBiomias 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), presently composed by:

- Dr. Alexandre Camargo Coutinho (Embrapa)
- Dr. Edson Eygi Sano (IBAMA)
- Dr. Gerd Sparovek (University of São Paulo)
- Dra. Leila Maria Garcia Fonseca (INPE)
- Dra. Liana Oighenstein Anderson (CEMADEN)
- Dra. Marina Hirota (Federal University of Santa Catarina)

And also former members who contributed to the project's development on previous collections:

- 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 MapBiomass project, across Collections 1 to 8, 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, MapBiomass primarily utilized USGS Collection 1 Tier 1 from the top of the atmosphere reflectance (TOA). Starting from Collection 6, new Landsat mosaics were processed using surface reflectance (SR) data. However, Collections 7 and 8 used the USGS Landsat Collection 2 (Tier 1).

2.1.3. Google Earth Engine and MapBiomass Computer Applications

MapBiomass 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 MapBiomass 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, 7 and 8 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, some post-classification tasks, and statistical analysis were all performed in Earth Engine Python API.
- R scripts – Used to improve the processing of large datasets in the Google Earth Engine and to perform machine learning algorithms calibration, data analysis, and visualization.
- Toolkits – User’s Toolkits are collections of scripts with a friendly user interface in Google Earth Engine to download MapBiomass’ data by state, biome, municipality, or any other geometry. These toolkits were developed for the general public that is not familiar with programming languages used in Google Earth Engine. 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, R, 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 8, 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, soil organic carbon stocks, 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/en>).

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 coarse 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 reconstruct 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/en>),

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

2.2.1. Global 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 size. 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. Most recently, an automated approach for global land use land cover (LCLU) classification was published using deep learning on 10 m Sentinel-2 imagery (Brown et al. 2022). Other global products were produced using coarser spatial resolution (>500 m) (e.g. Friedl et al. 2010) but are not listed here because their resolution limits applications to assess MapBiomass products, which are produced at 30 m Landsat pixel size.

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 Brazilian territory. The project was conducted from 1975 to 1980 and was 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 Brazilian vegetation (Cardoso, 2009).

Only the Amazon and Atlantic Forest biomes were being monitored using other systems, 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 (Projeto de Conservação e Utilização 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 cartographic scale of 1:250.000. The accuracy assessment was based on digital imagery products at a 1:100.000 scale, 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 all biomes in the year 2008 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 area, dunes, rock outcrops, mining, and water on the 1:250.000 cartographic scale.

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

2.2.3. Regional and Biomes Land Cover and Land Use Data

There are also reference LCLU maps (i.e. maps emphasizing the location of geographic features) 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 LCLU maps for the biomes and cross-cutting themes are presented in Annex II.

3. Methodological Description

The Collection 8 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 feature space attributes from the Landsat bands to train one random forest classifier (feature space definition) for each year (Breiman, 2001). Then, yearly training samples were acquired in each biome and cross-cutting theme according to its information availability and statistical needs. The output of the random forest classifier is one LCLU map per year for the entire territory based on the training dataset of that year. The exceptions are

the classes Aquaculture, Mining, Irrigation, Rice, and Citrus, which had their areas of occurrence identified using the U-Net convolutional neural network classifier (CNN).

Spatial-temporal filters were applied over the classified data for noise removal and temporal stabilization. Subsequently, the filtered LCLU maps of each biome and cross-cutting themes were hierarchically merged (integrated) based on a set of prevalence rules. The prevalence rules were set using expert knowledge and are described in Table 4. Spatial and temporal filters were once again applied on the integrated maps to create the final Collection 8 product.

The accuracy assessment analysis was based on acquiring 85,000 independent samples per year from 1985 to 2022. 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.

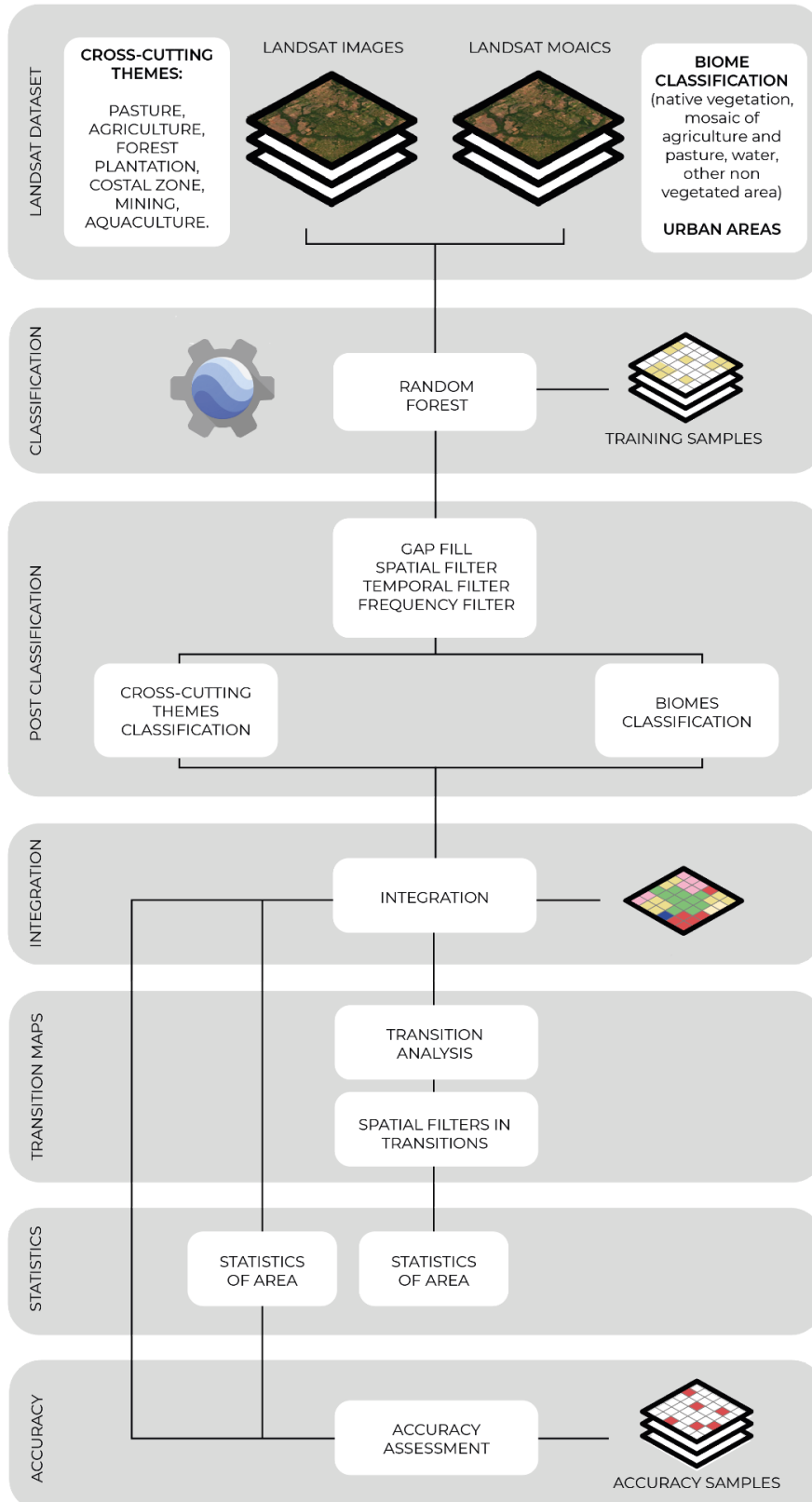


Figure 2. General methodological steps of Collection 8 to implement MapBiomas algorithms in the Google Earth Engine.

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. The 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 and 8, 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 enable 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 reduce noise and improve the mosaic quality, a tool was developed to evaluate the images individually, excluding uninformative images (excess cloud cover). The Amazon biome classified each Landsat image using Random Forests and reclassified the results to create the annual LCLU map.

3.2. MapBiomass feature space

The feature space for LCLU classification is composed of 90 input variables per year, including the original Landsat bands and fractional and textural information derived from those bands (Table 2). Table 2 presents the formula or the description to obtain these 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.

	band or index name	formula	Reducer					
			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	<code>('median_green').entropy(ee.Kernel.square({radius: 5}))</code>						
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 and cross-cutting theme executed a feature selection algorithm to choose the most appropriate subset of variables to train the respective random forest classifier. More details are available in the Appendices.

3.3. Classification

3.3.1. Legend

The MapBiomass classification scheme is a hierarchical system comprising four categorical levels (Table 3). At Level 1, there are six classes: 1) Forest, 2) Non Forest Natural Formation, 3) Farming, 4) Non-Vegetated Area, 5) Water, and 6) Not Observed. Level 2 has 16 classes across the six classes of the first categorical level. Agriculture (3.2) is the only class with further subdivisions down to the fourth categorical level, comprising nine LCLU classes.

Annex III presents the cross-reference of the MapBiomass 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 MapBiomass.

Table 3. Classes of land cover and land use of MapBiomass Collection 8 in Brazil.

ID	COLLECTION 8 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

6	1.4. Floodable Forest	NATURAL	COVER	
49	1.5. 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
35	3.2.2.3. Palm Oil	ANTHROPIC	USE	THEMES
48	3.2.2.4. 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/ ANTHROPIC	COVER/USE	BIOMES
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. Not Observed	NONE	NONE	NONE
----	-----------------	------	------	------

3.3.2. Training samples

Samples for training the yearly random forest classifiers were obtained using random sampling from areas with LCLU classes that did not change across all years of Collection 7 (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 parameters, such as the number of trees, a list of variables, and training samples. The appendices of the biomes and cross-cutting themes 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, Citrus, and Palm Oil classification.

3.4. Post-Classification

Due to the pixel-based classification method and the long time series, post-classification filters were applied to the biome and cross-cutting maps. The first post-classification action involved a gap filling algorithm and the application of temporal filters followed by a spatial filter. The application of those 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. Time-wise gap fill

The gap fill is a temporal filter used to fill no-data values. In a long time series of severely cloud-affected regions, it is expected that no-data values will populate some of the resultant median composite pixels. No-data values (“gaps”) are not allowed in this filter and replaced by the temporally nearest valid past or future classification. First, the filter operates backwards in time, replacing the no-data value with the last valid class of the time series. If no “previous” valid class is available, then the filter operates forward in time, and the no-data value is replaced with the first valid class found in the time series. Therefore, entire gaps in the final product will 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 isolated pixels 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. A pixel that does not share connections to at least four identical neighbors is considered isolated and its value is replaced with the mode value of the surrounding. This also means that this spatial filter affects the minimum mapping unit, defined as five pixels (~0.5 ha).

3.4.3. Temporal Filter

The temporal filter consists of backward moving windows of three and four (all classes) and five (forest) years that identify temporally unrealistic transitions between LCLU classes. Based on heuristic rules, the temporal filter inspects the central positions (one, two, and three years for the three, four, and five years windows, respectively) of the time window. If the two years at the extremities of the time window are identical, but the central positions are not, then the central pixels are reclassified to match the class of the extremities.

Another heuristic 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 frequency with which a LCLU class occurs throughout the entire time series. Thus, all natural class occurrences with less than a given percentage of temporal persistence (e.g. three or fewer years out of 38) 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 and cross-cutting theme customized the frequency filters accordingly; 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 along the 38 years period. All pixels that changed more than eight times and are connected to less than six pixels were replaced by the MODE value of that given pixel position in the stack of years. This removes unrealistic changes in the borders of mapped classes and eliminates noisy pixel trajectories. Each biome and cross-cutting theme customized the application of incident filters accordingly; see more details in its respective appendices.

3.4.6. Integration

The maps of each biome and the maps of cross-cutting themes were integrated on a pixel-by-pixel basis through the hierarchical overlap of each mapped class, following prevalence rules defined by experts (Table 4). Details about biome-wise prevalence rules are described in the Appendices.

Table 4. Collection 8 general prevalence rules for integrating biomes and cross-cutting themes maps.

ID	COLLECTION 8 CLASSES	PREVALENCE ID
1	1. Forest	-
3	1.1. Forest Formation	20
4	1.2. Savanna Formation	21
5	1.3. Mangrove	3
6	1.4. Floodable Forest	24
49	1.5. Wooded Sandbank Vegetation	22
10	2. Non Forest Natural Formation	-
11	2.1. Wetland	23
12	2.2. Grassland Formation	26
32	2.3. Salt Flat	5
29	2.4. Rocky Outcrop	17
50	2.5. Herbaceous Sandbank Vegetation	25
13	2.6. Other Non Forest Natural Formations	27
14	3. Farming	-
15	3.1. Pasture	28
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	12
36	3.2.2. Perennial Crop	-
46	3.2.2.1. Coffee	13

47	3.2.2.2. Citrus	14
35	3.2.2.3. Palm Oil	15
48	3.2.2.4. Other Perennial Crops	16
9	3.3. Forest Plantation	7
21	3.4. Mosaic of Uses	29
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	18
26	5. Water	-
33	5.1. River, Lake and Ocean	19
31	5.2. Aquaculture	4
27	6. Not Observed	-

3.4.7. Temporal and Spatial Filters on Integrated Maps

A temporal filter was applied on isolated classes over the time series, that is, classes with less than three occurrences in the whole 38 year period. After the application of the temporal filter on integrated maps, a spatial filter similar to the one described in 3.4.2 was applied on 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 any two maps were computed for the following periods: (A) any consecutive years (e.g. 2001-2002); (B) five-year periods (e.g. 2000-2005); (C) Forest Code period (2008-2022); (D) Forest Code approval (2012-2022); (E) National GHG Inventory (1994-2002; 2002-2010; 2010-2016); (F) all the years (1985-2022). The class transitions represent LCLU changes such as: Transitions from classes of agriculture or non-vegetated areas to forest cover or natural non-forest areas; Transitions that add water surface; Transitions that reduce water surface; Transitions with gain in forest plantation areas; Transitions from forest cover or natural non-forest areas to agriculture or non-vegetated areas; and Areas without transition or transitions that involve not observed areas or transitions between classes within level 1 of legend.

3.4.9. 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) streams of up to five pixels with two or one neighbor pixel in the same transition class.

3.4.10. Temporal Analysis

Since Collection 7, new temporal analyses besides the transition maps are present in the MapBiomass platform: stable areas, number of classes and number of changes. The stable area tool shows areas that stayed in the same LCLU class throughout the years in the temporal extent selected by the user. The number of classes tool shows the number of LCLU classes a pixel was classified as during the temporal extent selected by the user, while the number of changes shows the number of changes between LCLU classes a pixel went through during the temporal extent selected. Both tools consider a temporal resolution of one year and allows the user to choose the following temporal extents: (a) The whole temporal extent mapped by MapBiomass, from 1985 to 2022; (b) five years e.g. 1985-1990, 1990-1995, 1995-2000; (c) ten years e.g. 1985-1995, 1995-2005, 2005-2015; and (d) about half of the temporal extent mapped by MapBiomass, from 1985 to 2000 and from 2000 to 2022. 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). A toolkit in the Google Earth Engine is available to upload user-defined polygons and download the LCLU maps (<https://code.earthengine.google.com/?scriptPath=users%2Fmapbiomas%2Fuser-toolkit%3Amapbiomas-user-toolkit-lulc.js>).

3.5. Validation Strategies

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

3.5.1. Validation with Reference Maps

Each biome and cross-cutting theme conducted the spatial agreement analysis with reference maps where available. 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

Some 85,000 spatially independent samples, replicated for each of the 38 years between 1985 and 2022, were labeled according to MapBiomas LULC classes by experts after the visual interpretation of Landsat data, MODIS-NDVI times series, and high-resolution imagery from Google Earth (when available). Out of these , 10,000 samples were used as training samples for the Amazon biome. Thus, the error assessment analysis was conducted using ~75,000 samples per year (Figure 3). In Collection 7, the label of the 85,000 independent samples was updated and revised adding the recent years, 2019, 2020, 2021, with the year 2022 added in the Collection 8.

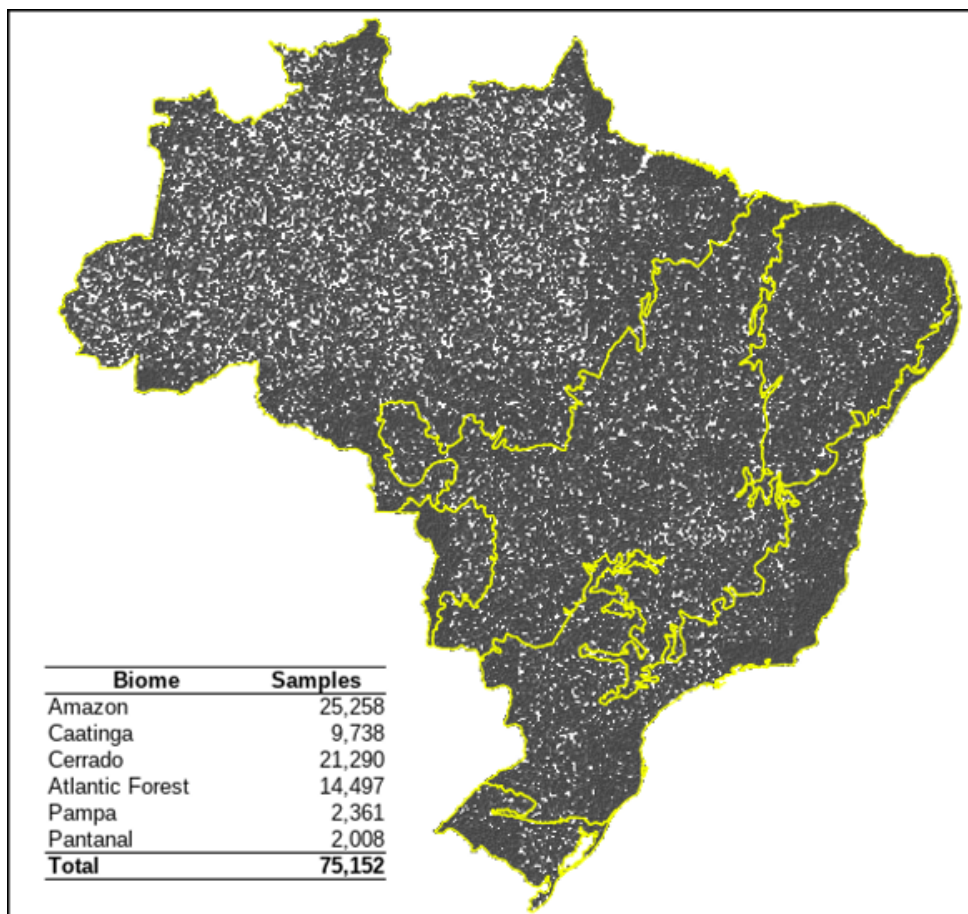


Figure 3. Independent random samples were used in the Brazilian biomes for the error assessment analysis of MapBiomas Collections.

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 and labeled each independent sample; in case of disagreement between interpreters, a senior interpreter decided the final LCLU class of the sample. 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 MapBiomass since Collection 3.1 (https://mapbiomas.org/accuracy-statistics?cama_set_language=en).

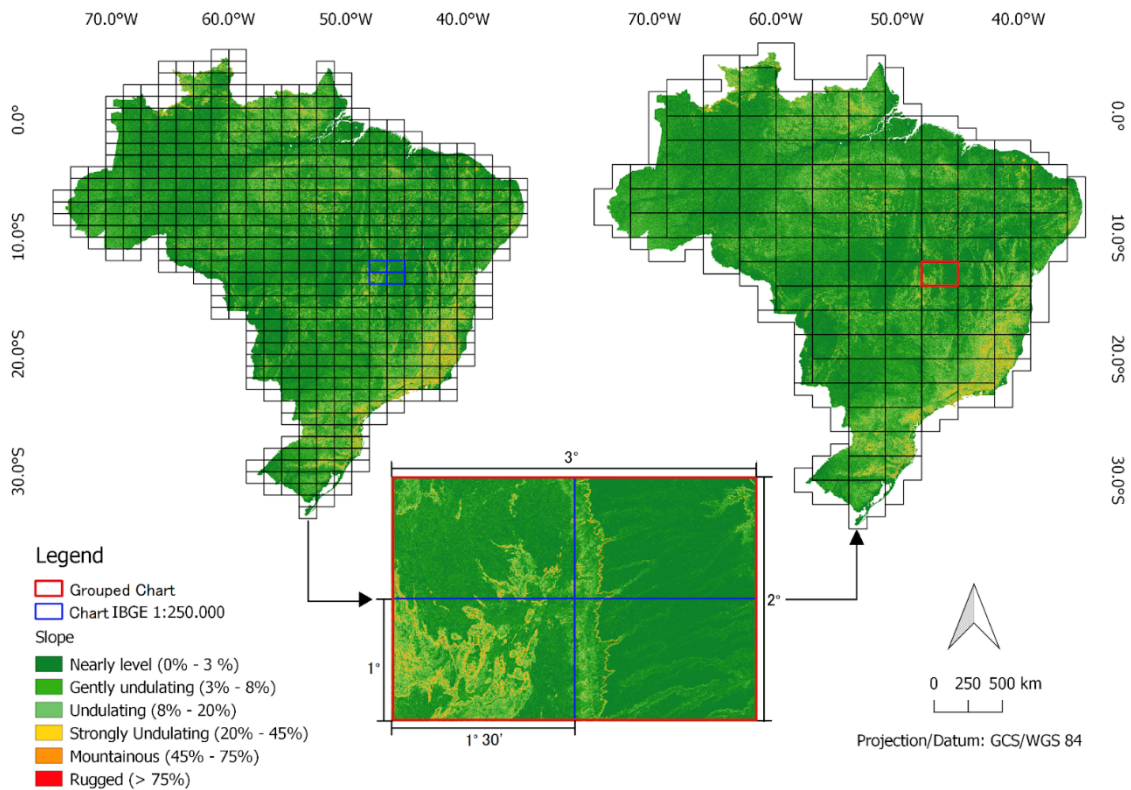


Figure 4. Slope categories used in sampling design for validation with independent points.

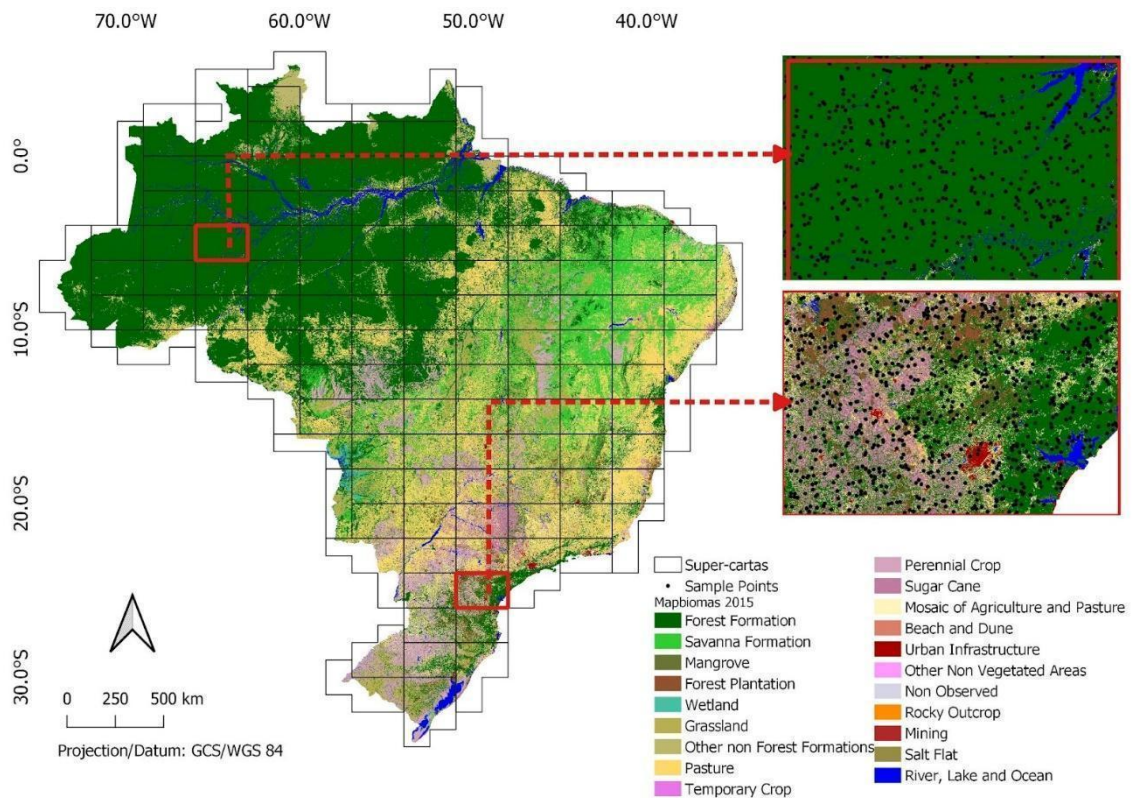


Figure 5. Super-chart size used as a subset in the sampling design for validation with independent points.

The global accuracy for each level of LCLU classes in the Collection 8 legend was calculated for each year, class, and biome (more details can be explored in the MapBiomas web platform (brasil.mapbiomas.org/en/analise-de-accuracia/)). In Level 1 classes, the LCLU mapping product in the Collection 8 presented 90.0% global accuracy and 9.0 % allocation disagreement with 1.0% area disagreement. At Level 2, the global accuracy was 85.8%, with 9.4% allocation disagreement and 4.8% area disagreement. Finally, at Level 3, the global accuracy was 85.8%, with 9.3% allocation disagreement and 4.9% area disagreement. The global accuracy was stable over the mapped period, varying across biomes from 82.3% to 96.8% in Level 1.

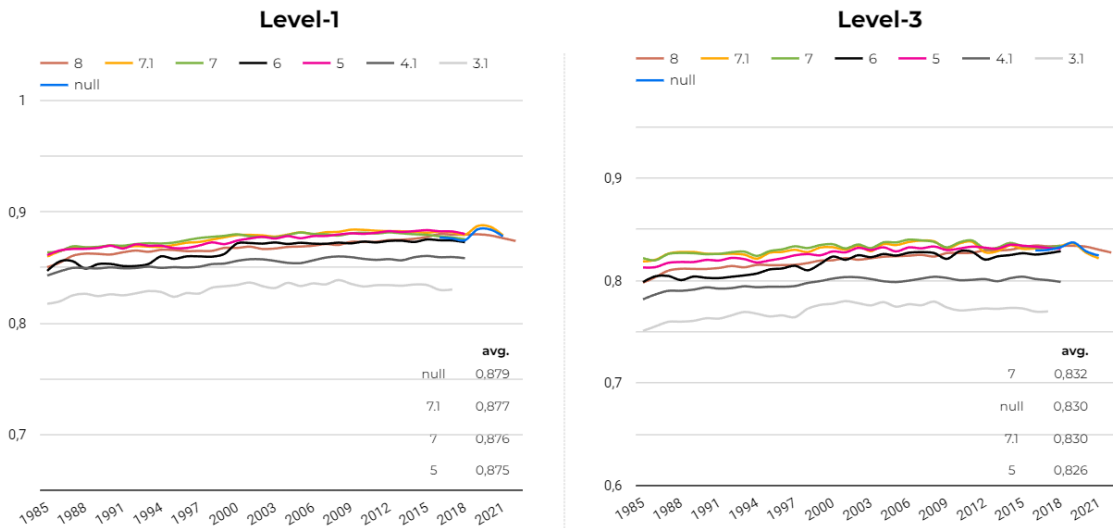


Figure X:

4. Map Collections and Analysis

The MapBiomass 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. 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.
- 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.
- Collection 8 - comprised the period of 1985 to 2022 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, Citrus, and Palm Oil that applied U-Net convolutional neural networks in the classification.

Table 5. MapBiomias' 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 level / 7 classes	Empirical Decision Tree + Random Forest (Farming)	n.a.
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%
8.0	38 years 1985-2022	4 levels / 29 classes	Random Forest + U-Net (Aquaculture, Irrigation, Mining, Rice, Citrus, and Palm Oil)	L1 - Gb 90% Ar 9% All: 1% L2 - Gb 86% Ar 9% All: 5% L3 - Gb 86% Ar 9% All: 5%

Collection 8 resulted not only in a long time series, adding the year 2022, but more spatially and temporally consistent annual LCLU maps of Brazil. Two new classes were mapped: Floodable Forest and Palm Oil, totaling 29 mapped LCLU classes. At the same time, the Wetlands class went through major improvements. Considerable improvements were done in Collection 8 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 to mapping Palm Oil besides Aquaculture, Irrigation, Rice, Citrus, and Mining. However, challenges remain, and more improvements will be made in the following updates of the MapBiomias collection. All the programming codes for running the MapBiomias classifications are publicly available and accessible through <https://mapbiomas.org/en>. All documents and data are available and referenced with a unique digital identifier DOI through the MapBiomias Network Data Repository (<https://data.mapbiomas.org/>).

5. Concluding Remarks and Perspectives

The algorithms developed 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 MapBiomias project. The replication of this type of project is viable for other areas of the planet. The MapBiomias initiative has expanded to other regions such as Amazon, Chaco, Pampa, Atlantic Forest, and Indonesia. In addition, the MapBiomias team will keep

improving the following collections in subsequent years. 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.

6. 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: MapBiomias Network

MapBiomias 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 Goiás (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)
- Montpellier Foundation
- Mulago Foundation
- 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
- MapBiomass Support Institute (IAMap)
- 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:

- Dr. Alexandre Camargo Coutinho (Embrapa)
- Dr. Edson Eygi Sano (IBAMA)
- Dr. Gerd Sparovek (University of São Paulo)
- Dra. Leila Maria Garcia Fonseca (INPE)
- Dra. Liana Oighenstein Anderson (CEMADEN)
- Dra. Marina Hirota (Federal University of Santa Catarina)

Former members:

- 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)

Technical Partners:

- Institute of Agricultural and Forest Management and Certification - Imaflora (IMAFLOA)
- 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 8.

MAP	SOURCE	DESCRIPTION	DOWNLOAD
Geomorphology Map of Brazil - 2021	IBGE	Database in digital media (1:250000), available throughout Brazil, with geometric consistency, structured and standardized for use in Geographic Information Systems on the subject of Geomorphology.	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/geonet/work/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 (GEF).	https://www.icmbio.gov.br/portal/imagens/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=19ps2n5F162X-ib2V2teFhagcUCbS2BZl&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-vectorial-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	https://www.ibge.gov.br/geociencias/organizacao-do-territorio/malhas-territoriais/26565-malhas-de-setores-censitarios-diviso-es-intramunicipais.html?="

		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.	&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	Geological Institute (IG) /	Sectoring or dividing the territory into areas with similar	http://s.ambiente.sp.gov.br/cpla/UHC

Occupation Units (UHCT) for the State of São Paulo	Department of the Environment of the State of São Paulo	characteristics regarding certain physical aspects, shape and intrinsic texture of the occupation.	T_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 OpenStreetMap 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 Gas Emissions (LULUCF sector)	MCTIC	Map of agricultural areas (annual, semi-perennial and 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 5 m 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	Nascimento Jr, W.R;	Map of mangrove areas from Oiapoque-AP to Ponta de	

region of Brazil	Souza-Filho, P.W.M., et al.	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 25 cm 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 and 2016	https://www.sospantanal.org.br/atlas/
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=WobGis: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 comprises 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	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/

	de Ecologia, 2016. 1a ed. ISBN 978-85-63843-20-3.		
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/
Coffee 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 verifiable data in the field. In productivity estimation, the knowledge of the location of the cultivation areas enables the monitoring of the productive areas through agrometeorological parameters, offering indicators 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:	University of Southampton	Translated with www.DeepL.com/Translator (free version)	https://www.worldpop.org/project/ca

Estimated Residential Population			tegories?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 MapBiomass land use/land cover classes in the Collection 8 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	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. Bamboo forest (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 Savannah, 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 Savannah) (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	FDP, FSP, 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
	Floodable Forest (beta)	Amazon	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.	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
	Wooded Sandbank Vegetation	Atlantic Forest	Forest formations on sandy soils in the coastal region.	Pma	FEP, FEM	FMN, FM
		Pampa	Forest formations on sandy soils in the coastal region.	Pma	FEP, FEM	FMN, FM
		Caatinga	Forest formations on sandy soils in the coastal region.	Pma	FEP, FEM	FMN, FM
Non-Forest Natural Formation	Wetland	Amazon	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	Tp, Sg, Rm, Sp, Tg, Rl	WG, OG, WS	GNM, GM, GSec

			areas that accumulate water, vegetation predominantly herbaceous to shrub).			
		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	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

	Hypersaline Tidal Flat	"Apicuns" or hypersaline tidal 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	Amazon	Naturally exposed rocks without soil cover, often with the partial presence of rupicolous vegetation and high slope.	Ar	OX	ArM, ArNM
		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	Herbaceous vegetation that is established on sandy soils or on dunes in the coastal zone.	Pmb, Pmh	WG, OG	GNM, GM
		Caatinga	Herbaceous vegetation that is established on sandy soils or on dunes in the coastal zone.	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	Amazon	Pasture area, predominantly planted, linked to livestock production activities. Areas of natural pasture are predominantly classified as grassland or wetland, 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)	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
			Palm Oil (beta)	Areas cultivated with palm oil plantation.	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 Sand Spot			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 areas include non-permeable surfaces (roads and infrastructure for rural developments).	AU, MCA, Dn, lu	OB, OQ, OX	S, SE, DnM, DnNM, Min
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
<p>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 Estatística - 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.</p>					

Annex IV: Classes of land cover and land use of Collections 1, 2, 2.3, 3, 4, 5, 6 and 7 of MapBiomias.

