MapBiomas General "Handbook"

Algorithm Theoretical Basis Document (ATBD)

Collection 3

Version 1.0

August, 2018

Table of Contents

Executive Summary

The MapBiomas initiative was formed in 2015 by universities, NGOs and companies to develop a fast, reliable, collaborative and low-cost method to produce an annual temporal series of land cover and land use maps of Brazil from 1985 to 2017. This mapping initiative is organized by biomes (Amazon, Atlantic Forest, Caatinga, Cerrado, Pampa and Pantanal) and cross-cutting themes (Pasture, Agriculture, Forest Plantation, Coastal Zone, Mining, Urban Infrastructure), and involves a wide range of specialists from from remote sensing, geography, geology, ecology, environmental and forestry engineering, computer science, human science, journalists, designers among others.

MapBiomas has produced three sets of digital annual maps of land use and land cover (LULC), named Collections. The satellite image classification methods and algorithms for each Collection evolved over time. Collection 1, which consisted on the first step of the mapping process, covered the period of 2008 to 2015 and focused on seven LULC classes: forest, agriculture, pasture, forest plantation, mangrove, and water. Collection 2 encompassesthe period of 2000 through 2016 and included 27 LULC classes with subclasses of forest, savanna, grasslands, mangroves, beaches, urban infrastructure and more. Collection 2.3 was launched based on random forest machine learning to overcome empirical calibration of the input parameters for image classification. Finally, Collection 3 also based on the random forest algorithm but included a more robust sampling designed for training the classifier, and expanded the mapping period for 1985 through 2017.

The objective of this This Algorithm Theoretical Basis Document (ATBD) this document is to provide the users of the MapBiomas data the understanding of the methodological steps and computational algorithms to produce Collection 3, and describe the datasets, statistics produced as well. All the MapBiomas maps and datasets are freely available at the project platform (http://mapbiomas.org).

1. Introduction

1.1. Scope and content of the document

The objective of this document is to describe the theoretical basis, justification and methods applied to produce annual maps of land cover and land use (LULC) in Brazil from 1985 to 2017 of the MapBiomas Collection 3.

This document covers the image classification methods of Collection 3, the image processing architecture, and the approach to integrate the biomes and cross-cutting theme maps. In addition, the document presents an historical context and background information, as well as a general description of the satellite imagery dataset, feature inputs, and of the accuracy assessment method applied. The algorithms and specific procedures applied in each biome and cross-cutting theme are present in the appendices.

1.2. Overview

The MapBiomas project was launched in July 2015, aiming to contribute with the understanding of LULC dynamics in Brazil. This LULC maps produced in this project were based on the Landsat Data Archive (LDA) available in the Google Earth Engine platform, encompassing the years from 1985 through the present days. The MapBiomas mapping efforts were divided in Collections for the following periods:

- 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 2.0 (launched in December 2017).
- Collection 3: 1985 through 2017 (launched in August 2018).

The Collection 2.3 marked the transition from empirical image classification approach to random forest machine learning classifier, from Collection 2 to Collection 3. Besides the annual classifications of digital maps, MapBiomas aims to contribute with the development of a fast, reliable, collaborative and low-cost method to process large-scale datasets to generate historical time-series of LULC annual maps. In addition, the project also produced a web-based platform (*i.e.,* MapBiomas Platform) to facilitate the implementation of the image processing method. Finally, all data, classification maps, software, statistics and further analyses are openly accessible through the MapBiomas Platform. 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) to organizations that are part of MapBiomas network that shared knowledge and mapping tools; and iii) to visionary funding agencies that support the project.

The products of the MapBiomas Collection 3 are the following:

- Biome maps (Amazon, Atlantic Forest, Caatinga, Cerrado, Pampa and Pantanal) and cross-cutting themes (Pasture, Agriculture, Forest Plantation, Coastal Zone, Mining, Urban Infrastructure);

• Pre-Processed feature mosaics generated from LDA collections (Landsat 5, Landsat 7 and Landsat 8).
• Image processing infrastructure and algorithms (scripts to run in Google Earth Engine, MapBiomas Workspace and source code).
• LULC transition statistics and spatial analysis with political, watershed, protected areas, and other categorical maps.

The MapBiomas project had also expanded to other regions and is now running in the Chaco region and in the Pan-Amazon countries. These new project areas also follow the mapping protocol of MapBiomas Brazil with a few adjustments to cope with peculiarities of these ecosystems. Detailed information about these MapBiomas initiatives can be found at the ATBD of these regions.

1.3. Region of Interest

MapBiomas was created to produce LULC annual maps for the entire Brazilian territory, thus covering all six official biomes of the country: Amazon, Atlantic Forest, Caatinga, Cerrado, Pampa and Pantanal (Figure 1). The division in biomes helps to classify distinct LULC classes and patterns across the country (Table 1). The project produced cross-cutting themes: agriculture, pasture, forest plantation, mining, urban infrastructure and water. Although the Coastal Zone is not considered a biome officially, this region that covers dunes, beaches and mangroves along the Brazilian coast was treated as such.

For the purpose of MapBiomas initiative the official Map of biomes from IBGE 1:5.000.000 was combined with the vegetation map of Brazil (1:250.000) to produce a 1:1.000.000 map of the limits of the biomes.

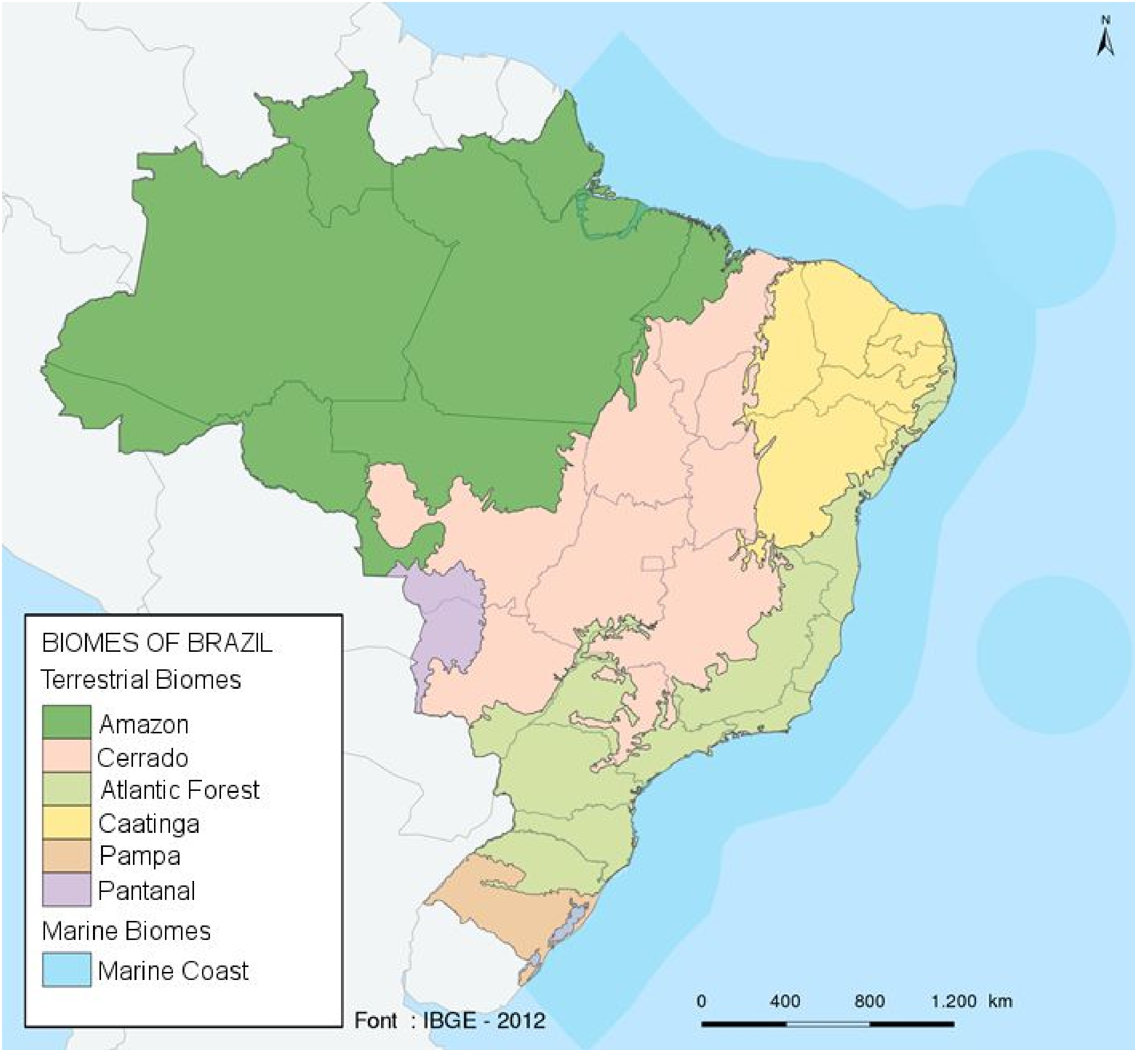

Figure 1. Brazilian biomes mapped in the MapBiomas project to generate the Collection 3 products (source: IBGE, 2012).

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

Biome	Area (km^2) (Country %)	Land Cover	Land Use
Amazon	4,196,943 (49.29%)	Evergreen forest, with enclaves of savanna, natural grassland, and extensive wetland and surface water.	Cattle ranching, agriculture, mining, logging and non-timber forestry production.
Atlantic Forest	1,110,182 (13.04%)	Small isolated forest fragments covering 7% of the biome (Morellato & Haddad, 2000), mostly old secondary growth, surrounded by croplands, pasture, forest plantation, and urban infrastructure.	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%)	Savanna, grassland and wetland	Agriculture, cattle ranching and ecotourism.

1.4. Key Science Applications

MapBiomas was originally designed to fill gaps in greenhouse gas emission estimates of the LULC change sector in Brazil. However, other scientific applications can be derived with an annual time-series history of land use and land cover maps produced, including:

- Mapping and quantifying land cover and land use transitions.
- Quantification of gross and net forest losses and gains.
- Monitoring of regeneration and secondary growth forests.
- Monitoring of water resources and their interaction with land cover classes.
- Monitoring agriculture and pasture expansion.
- Expansion of infrastructure and urbanization.
- Identification of desertification process.
- Regional planning.
- Management of Protected Areas.
- Monitoring of Forest Concessions.
- Develop land use indicators for risk of infectious diseases.

2. Overview and Background Information

2.1 Context and Key Information

This section addresses complementary contextual and key information relevant to the understanding of the MapBiomas products and methods to generate the map Collections.

2.1.1. MapBiomas Network

MapBiomas is a multi-institutional initiative of the Greenhouse Gas Emissions Estimation System (SEEG - http://seeg.eco.br/en/) promoted by the Climate Observatory (a network of 40+ NGOs working on climate change in Brazil - http://www.observatoriodoclima.eco.br/en/). The co-creators of the MapBiomas involve NGOs, universities and technology companies (list of all organizations involved in the Annex I).

Organizations play specific or multiple roles as well as contributes to the overall development of the project. Each biome and cross-cutting theme (agriculture, pasture, forest plantation, coastal zone, mining and urban infrastructure) has a lead organization, as showed in the box below.

Biome coordination:

• **Amazon** – Institute of Man and Environment of the Amazon (IMAZON).

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

• **Caatinga** – State University of Feira de Santana (UEFS) and Plantas do Nordeste Association (APNE).

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

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

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

Cross-cutting themes coordination:

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

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

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

• **Urban Infrastructure** – Terras.

Two geospatial tech companies, Terras and Ecostage, are responsible for the workspace/backend and dashboard/website/frontend of the MapBiomas, respectively. Google provides the cloud computing infrastructure that allows data processing, analysis and storage in through Earth Engine.

In summary, a team of specialists on remote sensing and biomes/cross-cutting themes works to develop the methods and algorithms for satellite image processing through the Code Editor interface of Google Earth Engine. When the algorithms are defined, MapBiomas programmers translated them to run in the Workspace web application that allow non-programmer analyst team of specialists to set the different parameters for making LULC maps.

Funding to implement and operationalize the MapBiomas Initiative comes from the Norway's International Climate and Forest Initiative (NICFI), Gordon & Betty Moore Foundation, Good Energies Foundation, Arapyaú Institute, Climate and Land Use Alliance (CLUA) and Good Energies Foundation.

Since both Climate Observatory and MapBiomas are not institutions, the initiative receives the generous institutional manage to operational and financing tasks from partners which include Avina Foundation, World Resources Institute (WRI), The Nature Conservancy (TNC) and Instituto Democracia e Sustentabilidade (IDS).

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

• Dr. Alexandre Camargo Coutinho (Embrapa)
• Dr. Edson Eygi Sano (IBAMA)
• Dr. Gilberto Camara Neto (INPE)
• Dr. Joberto Veloso de Freitas (Brazilian Forest Service)
• Dr. Matthew C. Hansen (Maryland University)
• Dr. Mercedes Bustamante (University of Brasília)
• Dr. Timothy Boucher (TNC)

2.1.2. Remote Sensing Data

The imagery dataset used in the MapBiomas project, across Collections 1 to 3, was obtained by the Landsat sensors Thematic Mapper ™, Enhanced Thematic Mapper Plus (ETM+) and the Operational Land Imager and Thermal Infrared Sensor (OLI-TIRS), on board of Landsat 5, Landsat 7 and Landsat 8, respectively. The MapBiomas has used Collection 1 Tier 1 from USGS and top of the atmosphere reflectance (TOA), which underwent through radiometric calibration and orthorectification correction based on ground control points and digital elevation model to account for pixel co-registration and correction of displacement errors. The Landsat imagery collections with 30 pixel resolution were accessible via Google Earth Engine, and source by NASA and USGS.

2.1.3 Google Earth Engine and MapBiomas Computer Applications

MapBiomas image processing chain 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 Earth Engine is defined by Google as: "a platform for petabyte-scale scientific analysis and visualization of geospatial datasets, both for public benefit and for business and government users."

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

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

• Python scripts – This category of code was used to optimize image processing of large datasets in Google Earth Engine. In addition, the image processing of Collection 3, and statistical analysis of the classification results were implemented in Earth Engine Platform with the Python API.

• WebCollect: a front-end web application to allow image analysts to collect and interpret sample points (*i.e.,* at the pixel level) by visual interpretation of Landsat color composite images. The main application of the WebCollect is to derive reference LULC classes for accuracy assessment.

• Workspace - a web-based application to allow general user with no-programming experience to access imagery collections, process them, and manage store the results in databases and map assets (*i.e.*, new collections) (Figure 2). The biome maps of Collections 1 and 2 were produced using the Workspace application. The Workspace environment allows to manage each image individually, define and store image classification parameters on a per map sheet basis (Figure 2). The biome teams of analysts can work 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 are Landsat image mosaics, LULC maps, transition analysis and statistics. All these products are publicly available on the web platform named MapBiomas Dashboard. More details on how the Workspace was used to parameterize the image processing and the classification, and control the processing workflow by biome teams are presented in the specific ATBDs of Collections 1 and 2 of the biomes.

• Mapbiomas.org (Dashboard). The web-platform of the MapBiomas initiative presents the Landsat image mosaics and its quality, land cover and land use annual maps of the Collection 3, transitions analysis, statistics, and all the methodological information about the ATBDs, tools, scripts, and accuracy analysis. All the maps and Landsat mosaics of the MapBiomas Collections are publicly available to download at the MapBiomas website.

Figure 2. MapBiomas Workspace interface. Brazil is divided in map charts of 1 degree and 1.5 degree to establish classification parameters and store them in a database.

2.2 Historical Perspective: Existent Maps and Mapping Initiatives

The existing mapping efforts that cover all Brazil, before MapBiomas, were neither frequent and nor updated (Annex II) and sometimes have lower resolution. MapBiomas and the available global and national land cover products can be used complementary but there are potential advantages of MapBiomas maps. First, the MapBiomas maps will reconstruct the entire Landsat time-series (>35 years) on an annual basis. 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 and forest degradation), secondary forest regrowth, and land use classes (pasture, agriculture, forest plantation and urban infrastructure) along this time series. Even though the MapBiomas project focus at a national scale in Brazil, there is a potential to replicate it in other countries. Currently, the Chaco region and the Pan-Amazon countries are already adopting the MapBiomas methods.

2.2.1. International land cover data

Mapping initiatives at the global level complement national mapping efforts (Annex II). The USGS in collaboration with the University of Maryland produced global land and tree cover layers circa 2010. USGS also produces a MODIS land cover map at

500m pixel scale. The GlobCover Portal is another initiative from the European Space Agency (ESA) which produced land cover maps with MERIS sensor at 300m spatial resolution for two periods: December 2004 - June 2006 and January - December 2009. Global Forest Watch (GFW) and Google Earth Engine provide the Global Forest Change (GFC) maps from 2000 to 2014 derived from the Landsat imagery at 30 m resolution produced by University of Maryland (Global Land Cover Facility - GLCF). The National Geomatics Center of China (NGCC) produced GlobeLand30 - a high-resolution (30 m) full coverage land cover maps for years 2000 and 2010. Finally, Japan Aerospace Exploration Agency (JAXA) also produced a forest/non-forest map for 2007-2010 using a 25m-resolution PALSAR mosaic. There are other global products that were produced using lower spatial resolution (>500m) but are not presented here because their resolutions limits applications to assess MapBiomas products, which are produced at 30m Landsat pixel.

2.2.2. National land cover data

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

In 2004, the Minister of Environment launched the natural vegetation map of Brazil developed in the context of Probio (*Projeto de Conservação e Utilização Sustentável da Diversidade Biológica Brasileira*) providing updated information about land cover in Brazil, considering that only the Amazon and Mata Atlantic biomes were being monitored after RadamBrasil project. The Brazilian biome boundaries (IBGE, 2004a) were used as reference for national mapping initiative. The Probio project was based on Landsat imagery acquired in 2002, with minimum mapping unit varying from 40 to 100 hectares, and mapping scale of 1:250.000. Accuracy assessment was based on digital imagery products at 1:100.000, with a minimal overall accuracy of 85%. The land cover classes followed IBGE manual for vegetation mapping (IBGE, 2004b). The Probio project updated forest change mapping for the year 2008 for all biomes and for 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 and 2010 (also 2005 for the Amazon). Those maps were produced by FUNCATE based on segmentation and visual interpretation of Landsat Imagery and identifying natural vegetation (forest and no-forest), agriculture, pasture, silviculture, urban areas and water.

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

2.2.3. Regional and biomes land cover data

There are also reference maps at the biome scale and though 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 state). These reference land cover and land use maps for the biomes and cross-cutting themes are presented in the Annex II.

3. Methodological description

The methodological steps of Collection 3 are presented in the Figure 3 and detailed above. The first step was to generate annual Landsat image mosaics based on specific periods of time to optimize the spectral contrast and discriminate the LULC classes across the biomes (see the biome Appendices for detailed information). The second step was to establish the spectral feature inputs derived from the Landsat bands to run the random forest classification. The classification was based on training samples acquired from temporal stable LULC classes generated in Collection 2.3, reference maps and manual sampling. Once the training samples were defined, a random forest classifier was run. Following that, spatial and temporal filters were applied to remove classification noise. The LULC maps of each biome and cross-cutting themes were integrated based on prevalence rules to generate the final map Collection 3. Accuracy assessment analysis was conducted in the map Collection following the good practices proposed by Olofsson et al. (2014). LULC change transition analysis (with spatial filter application) statistics, and spatial analysis with political, watershed, protected areas, and other categorical maps, were generated to understand LULC within these boundaries.

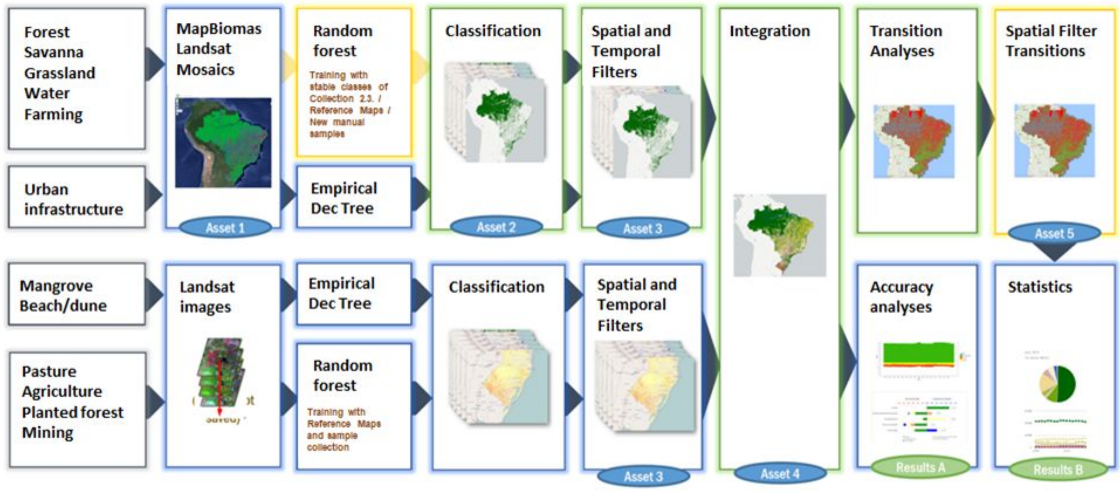

Figure 3. Methodological steps of Collection 3 to implement MapBiomas algorithms in the Google Earth Engine.

3.1. Mapping unit/charts

The mapping unit adopted in the MapBiomas project was defined based on the subdivision of the International Chart of the World to the Millionth on the 1:250,000 scale. Each rectangle of this subdivision covers an area of 1°30' of longitude by 1° of latitude, totalizing 558 charts (or sheets) for the Brazilian territory (Figure 4). Charts intercepting more than one biome were processed separately, with parameters adjusted for the specificities of the portions of each biome and were subsequently concatenated in the post-classification step.

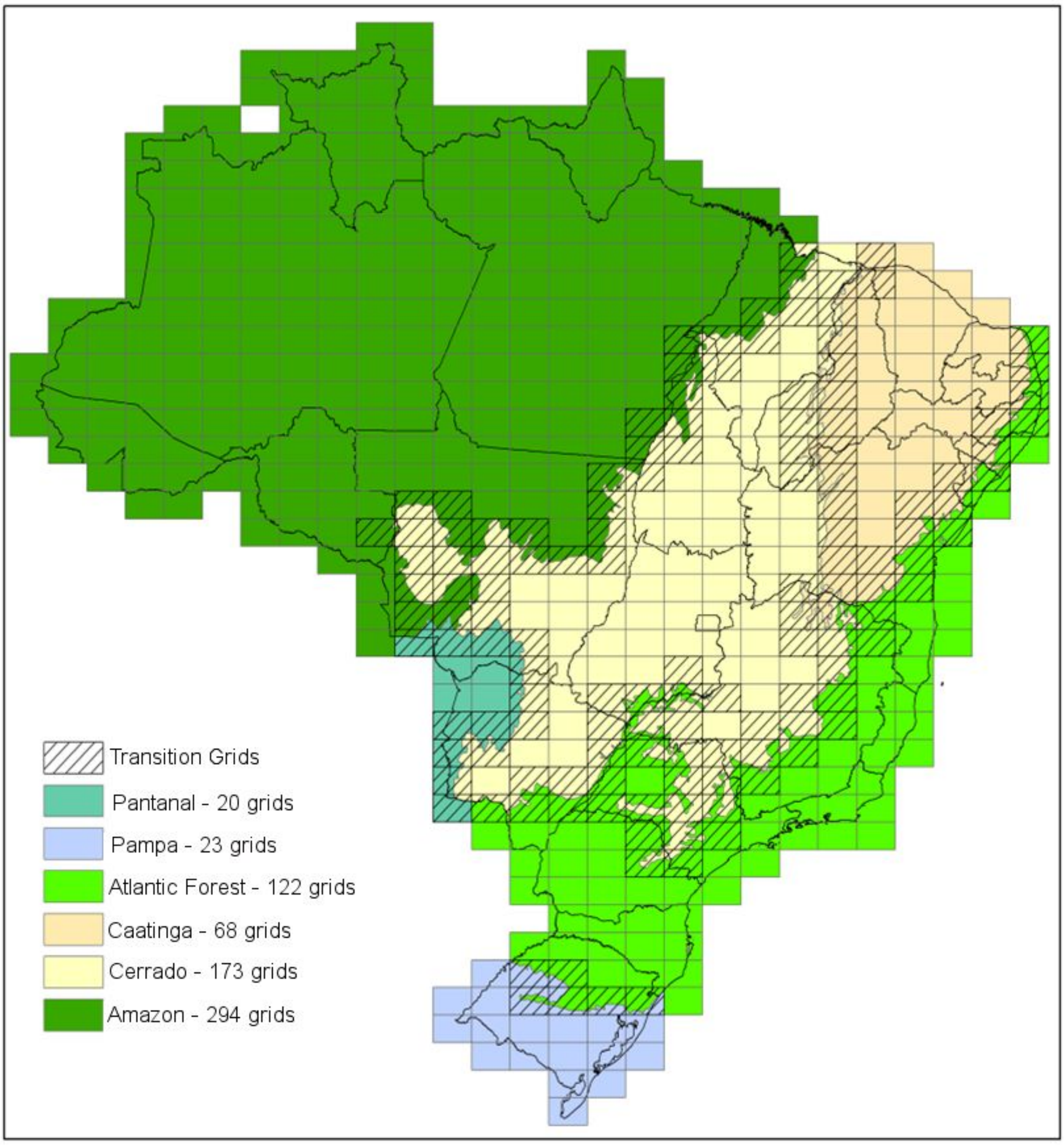

Figure 4: Distribution and number of the charts along the Brazilian biomes used for processing Landsat images within the MapBiomas Workspace environment.

3.2. Annual Landsat Image Mosaicing

All biomes generated annual Landsat image mosaics based on specific periods of time that optimize the spectral contrast to discriminate the LULC classes target for mapping. A cloud masking procedure (described below) and a median reduction algorithm were applied to generated the best ground observed pixel of the temporal period and produce the image mosaic for each year. For each chart, specific temporal mosaic of Landsat images was built, based on the following selection criteria/parameters:

1. Mapping year covered in the map Collection (*i.e.,* 1985 through 2017 for Collection 3);

2. Period of the year for selecting images (t0 and t1 in day/month/year) to build the annual mosaic that provides the maximum spectral contrast between forest and non-forest;
3. Maximum cloud coverage acceptable in the Landsat image that will be used for processing;
4. Sensor type (for the years 2000-2002 Landsat 5 [L5], 2003-2011 Landsat [L7], and 2013-2016 Landsat 8 [L8]). These parameters are obtained from the metadata file of the Landsat image collections already available on the Google Earth Engine platform.

The Coastal Zone and Mining theme generated annual mosaics using images available in the entire year (*i.e.*, January 1st to December 31st) for the entire region as a mapping unit, and the other themes (Pasture, Agriculture and Forest Plantation) processed the Landsat imagery in a per scene basis (more details available in the Appendices).

Cloud mask: The cloud and cloud shadow pixels were removed by a cloud mask applying the BQA band (Quality Assessment Band), single cloud score and Temporal Dark Outlier Mask (TDOM) algorithm in each single Landsat image during the process to build the mosaics.

Quality assessment of the Landsat mosaics: The mosaic is a composition of pixels from various Landsat scenes over a period of time. These scenes have a percentage of clouds and other atmospheric interferences. Thus each Landsat scene in each year was classified into three categories according to cloud percentage and verified interference: 1 – Bad: > 30%; 2 – Regular: 10 a 30%; 3 – Good: < 10%. The quality assessment of the Landsat mosaics are available at MapBiomas website.

3.3. MapBiomas feature space

The feature space of MapBiomas Collection 3 has 104 input variables, including the original Landsat bands, fractional and textural information derived from these bands. These feature variables were used in the random forest classification of LULC classes (Table 2). Table 4 presents the formula or the description to obtain these feature variables. In addition, statistics were obtained to capture the variability of the pixel ground observation for the period used to generate the annual image mosaics, including:

- Median - Median of the pixel values of the best mapping period defined by each biome.
- 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 images of each year.
- stdDev = standard deviation of all pixel values of all images of each year.
- Min = lower annual value of the pixels of each band.

The name of each band in the GEE asset is defined by the statistic applied plus the band or index name, as for example, median_blue which means the median of the blue band. Table 2 shows in green all the bands, indices and fractions available in the data feature asset.

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

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

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

Each biome selected empirically a set of features o to optimize the spectral separability of the LULC (Table 3). Each cross-cutting theme used their own feature variables and more details are available in the Appendices (Pasture, Agriculture, Coastal Zone, Mining, Urban Infrastructure).

Table 3. Bands, indices and fractions of the feature space applied by each biome in the MapBiomas Collection 3.

Index	Name	Amazon	Atlantic Forest	Caatinga	Cerrado	Pampa	Pantanal
1	amp_evi2						
2	amp_gv						
3	amp_ndfi						
4	amp_ndvi						
5	amp_ndwi						
6	amp_npv						
7	amp_sefi						
8	amp_soil						
9	amp_wefi						
10	median_blue						
11	median_blue_dry						
12	median_blue_wet						
13	median_cai						
14	median_cai_dry						

15	median_cai_wet						
16	median_cloud						
17	median_evi2						
18	median_evi2_dry						
19	median_evi2_wet						
20	median_fns						
21	median_fns_dry						
22	median_fns_wet						
23	median_gcvi						
24	median_gcvi_dry						
25	median_gcvi_wet						
26	median_green						
27	median_green_dry						
28	median_green_wet						
29	median_gv						
30	median_gvs						
31	median_gvs_dry						
32	median_gvs_wet						
33	median_hallcover						
34	median_ndfi						
35	median_ndfi_dry						
36	median_ndfi_wet						
37	median_ndvi						
38	median_ndvi_dry						
39	median_ndvi_wet						
40	median_ndwi						
41	median_ndwi_dry						
42	median_ndwi_wet						
43	median_nir						
44	median_nir_dry						
45	median_nir_wet						
46	median_npv						
47	median_pri						
48	median_pri_dry						
49	median_pri_wet						
50	median_red						
51	median_red_dry						
52	median_red_wet						
53	median_savi						
54	median_savi_dry						
55	median_savi_wet						
56	median_sefi						
57	median_sefi_dry						
58	median_sefi_wet						

59	median_shade						
60	median_soil						
61	median_swir1						
62	median_swir1_dry						
63	median_swir1_wet						
64	median_swir2						
65	median_swir2_dry						
66	median_swir2_wet						
67	median_temp						
68	median_wefi						
69	median_wefi_dry						
70	median_wefi_wet						
71	min_blue						
72	min_green						
73	min_nir						
74	min_red						
75	min_swir1						
76	min_swir2						
77	min_temp						
78	stdDev_blue						
79	stdDev_cai						
80	stdDev_cloud						
81	stdDev_evi2						
82	stdDev_fns						
83	stdDev_gcvi						
84	stdDev_green						
85	stdDev_gv						
86	stdDev_gvs						
87	stdDev_hallcover						
88	stdDev_ndfi						
89	stdDev_ndvi						
90	stdDev_ndwi						
91	stdDev_nir						
92	stdDev_npv						
93	stdDev_pri						
94	stdDev_red						
95	stdDev_savi						
96	stdDev_sefi						
97	stdDev_shade						
98	stdDev_soil						
99	stdDev_swir1						
100	stdDev_swir2						
101	stdDev_temp						
102	stdDev_wefi						

3.4. Classification

3.4.1. Legend

MapBiomas aims to classify the most comprehensive LULC classes at the Landsat pixel level. The classification scheme for the Collection 3 is described in Table 4. The Annex III presents the cross reference of the MapBiomas LULC classes with classes from other classification schemes (*i.e.*, FAO, IBGE and National GHG Emissions Inventory). The Annex IV presents the classification scheme of the previous collections of MapBiomas.

Table 4. Classes of land cover and land use of MapBiomas Collection 3.

ID	COLLECTION 3 CLASSES	NATURAL/ ANTHROPIC	LAND COVER/ LAND USER	BIOMES/ THEMES
1	**1. Forest**	NATURAL/ ANTHROPIC	COVER/USE	-
2	1.1. Natural Forest	NATURAL	COVER	-
3	1.1.1. Forest Formation	NATURAL	COVER	BIOMES
4	1.1.2. Savanna Formation	NATURAL	COVER	BIOMES
5	1.1.3. Mangrove	NATURAL	COVER	THEMES
9	1.2. Forest Plantation	ANTHROPIC	USE	THEMES
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
13	2.3. Other non Forest Natural Formation	NATURAL	COVER	BIOMES
14	**3. Farming**	ANTHROPIC	USE	-
15	3.1. Pasture	ANTHROPIC	USE	THEMES
18	3.2. Agriculture	ANTHROPIC	USE	THEMES
21	3.3. Mosaic of Agriculture and Pasture	ANTHROPIC	USE	BIOMES
22	**4. Non vegetated area**	NATURAL/ ANTHROPIC	COVER/USE	-
23	4.1. Beach and Dune	NATURAL	COVER	THEMES
24	4.2. Urban Infrastructure	ANTHROPIC	USE	THEMES
29	4.3. Rocky outcrop	NATURAL	COVER	BIOMES

30	4.4. Mining	ANTHROPIC	USE	THEMES
25	4.5. Other Non Vegetated Area	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. Non Observed	NONE	NONE	NONE

3.4.2. Sample collection

Samples for the training and calibration of the random forest classifier were obtained as followed. First, for the Amazon biome, samples were acquired from a set of reference maps including GlobeLand30, Prodes, TerraClass and Global Forest Cover. The LULC of these maps were matched with the MapBiomas classes. A set of 5,000 points were obtained for each map sheet to extract the pixel values of the feature variables matching the year of the reference map with the year of the map sheet mosaic. These samples were used in the random forest classifier using sample data for each map sheet as an input. For map sheet with low representation of sample classes, the samples of the adjacent map sheets were used in the random forest classifier.

The second approach to sample data to train and calibrate the random forest classifier used the classification results of Collection 2.3. Samples were extracted from classes that did not change their value across all years of the this collection (*i.e.*, 2000 through 2016). Additional samples were collected in map sheets that did not have enough class areas that did not change over time. A polygon was drawn and random points were generated inside the polygons to obtain these additional samples. This approach was applied to all other biomes.

3.4.3. Classification

We use the random forest classifier available in Earth Engine to classify the LULC of Table 4. The minimum number of trees in the random forest classifier was 50 with 100 iterations. The number of features selected was the default value for this parameter (*i.e.*, mtree which is given by the square root of the number of feature as define by each biome).

3.5. Post-classification

The final classification result for each map sheet resulted in three final products: classification, classification with application of the spatial and temporal filters and the classification resulting from the integration with the cross-cutting themes.

The first post-classification action involves the application of spatial and temporal filters to the maps generated in the LULC classification step. The application of these filter is necessary to remove classification noise and disallowable LULC class transitions. The temporal filter was also used to fill information gap due to cloud. These pos-classification procedures were implemented the Google Earth Engine platform and are described in more detailed below.

3.5.1 Spatial filter

The spatial filter segments and indexes the classes of each collection into contiguous regions, which are subsequently identified and reclassified based on the following criteria: areas less than or equal to half a hectare (*i.e.*, approximately 5 pixels) are reclassified based on the majority of the neighboring classes. For instance, a patch belonging to a given class of up to 5 pixels is first identified along with its neighboring pixels; this patch is then reclassified as the predominant class value of the neighboring pixels. This process is applied to all segments of the classes selected for filtering. Spatial filters applied in each biome and cross-cutting theme are described in their respective Appendices.

3.5.2 Temporal filter

The temporal filter seeks to identify and correct class transitions that expected along a series of consecutive years (*i.e.*, 3 to 5), as well as to fill in pixels with no data caused by cloud cover. For example, a pixel classified as non-Forest in a given year ti (where i = 2008, 2009, ..., 2015), and Forest in year ti1 and ti+1, is reclassified as Forest for the year ti. Several transition rules were defined and applied to be used in the temporal filter for each biome to deal with specific phenological and land use transitions. Temporal filters applied to each biome and cross-cutting theme are described in their respective Appendices.

3.5.3 Integration

The integration of the maps of each biome with the maps of cross-cutting themes was accomplished through hierarchical overlap of each mapped class (Table 5), according to specific rules of prevalence defined empirically (*e.g.* in the Cerrado biome, Pasture class overlayed natural vegetation cover; in the Pantanal biome, the class 21 became 15; in the Pampa biome, the left over of class 21 was converted to 15; in the Cerrado and Amazon transitions, the Natural Non-Forest Formation areas in the Amazon were replaced by Cerrado classification). Biomes prevalence rules details are described in the Appendices. The integration process was made on a pixel by pixel basis.

Table 5. Collection 3 general prevalence rules for integrating biomes and crosscutting themes maps.

ID	COLLECTION 3 CLASSES	PREVALENCE ID
1	**1. Forest**	-
2	1.1. Natural Forest	-
3	1.1.1. Forest Formation	12
4	1.1.2. Savanna Formation	13
5	1.1.3. Mangrove	3
9	1.2. Forest Plantation	8
10	**2. Non Forest Natural Formation**	-
11	2.1. Wetland	15
12	2.2. Grassland Formation	16
32	2.3. Salt flat	7
13	2.3. Other non Forest Natural Formation	17
14	**3. Farming**	-
15	3.1. Pasture	18
18	3.2. Agriculture	9,10,11
21	3.3. Mosaic of Agriculture and Pasture	21
22	**4. Non vegetated area**	20
23	4.1. Beach and Dune	2
24	4.2. Urban Infrastructure	6
29	4.3. Rocky outcrop	14
30	4.4. Mining	1
25	4.5. Other Non Vegetated Area	19
26	**5. Water**	-
33	5.1. River, Lake and Ocean	5
31	5.2. Aquaculture	4
27	6. Non Observed	NONE

3.5.4 Spatial Filter on Integrated Maps

A spatial filter similar to the one described in 3.5.1 was applied in the integrated maps to remove isolated classes with less than half hectares as well as noise resulting from eventual Landsat data misregistration.

3.5.5. Transition Maps

The pixel to pixel class differences between the maps follow the periods: (A) any consecutive years (eg 2001-2002); (B) five-year periods 2000-2005 / 2005-2010 / 2010-2015; (C) 2008-2015 and (D) 2000-2016. The class transitions represent land use changes available in maps and Sankey diagram.

3.5.6. Spatial Filter on Transition Maps

A spatial filter similar to the one described in 3.5.1 was applied in the transition maps. The target is to eliminate single pixels or streams of pixels in the border of different classes derived from the creation of transition maps. The general rules for this filter were: (i) pixels with only one neighbor pixel in the same transition class; (ii) stream of up to five pixels with two or one neighbor pixel in the same transition class.

3.5.7. Statistics

Zonal statistics of the mapped classes were calculated for different spatial units, such as the biomes themselves, states and municipalities. In the Collection 3, river basins, protected areas (including indigenous lands and conservation units), rural settlements were included in the zonal statistics, and in a newer future will be included the user-defined polygons.

3.6. Validation Strategies

The validation strategy was based in two approaches: (i) comparative analyses with reference maps existed for specific regions and years, and (ii) accuracy analyses based on statistical techniques to define sample points based on the extent and on the number of classes of each biome.

3.6.1. Validation with reference maps

The spatial agreement analyses with reference maps were conducted by each biome and cross-cutting themes, according with their availability (more details available in the Appendices).

3.6.2. Validation with independent points

A preliminary analysis was performed based on ~30,000 pixel samples to each one of the years based on visual interpretation of Landsat data. Each sample was inspected by three independent interpreter, in case of confusion a senior interpreter decides the final class of the pixel. This evaluation was based on the web platform Temporal Visual Inspection (TVI - tvi.lapig.iesa.ufg.br:5000), developed by LAPIG/UFG. These samples were ideally collected for pasture mapping. The TVI platform allowed the evaluation of 12 classes, thus MapBiomas Collection 3 classes were simplified to feet TVI classes. Therefore, the TVI and MapBiomas classes were simplified to seven classes as follow: Natural Forest (Forest Formation, Savanna Formation and Mangroves), Grassland, Agriculture, Farming (Pasture and Mosaic of Agriculture or Pasture classes), Urban Infrastructure, and Water. Once adjusted, a confusion matrix was created. The global accuracy was calculated to each year and biome. This preliminary accuracy results showed that on country level the overall accuracy of Collection 3 is higher than Collection 2 (Figure 5). Not only for the country level but also to each biome there was an improvement in the global accuracy (Figure 6).

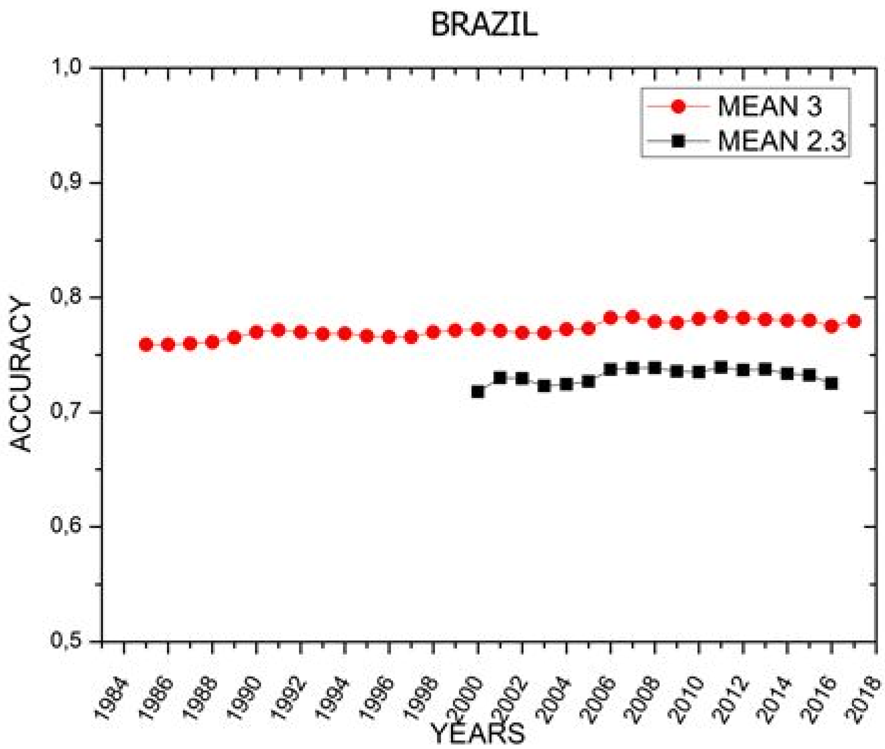

Figure 5. Global Annual Accuracy comparing Collections 2.3 and 3 in Brazil.

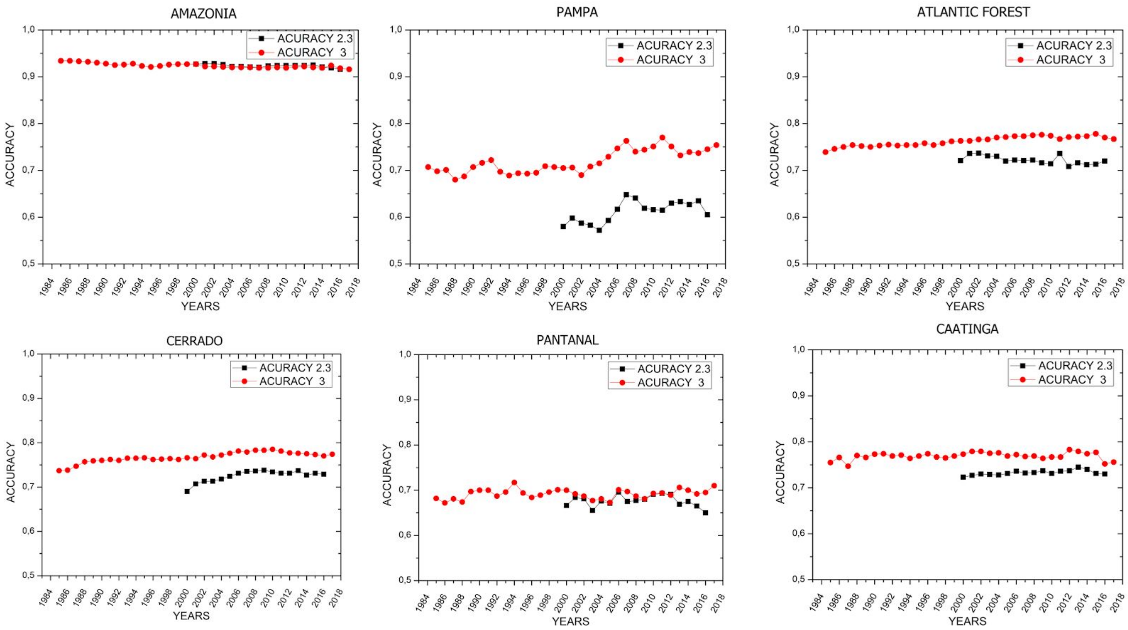

Figure 6. Biomes annual global accuracy comparison of Collections 2.3 and 3.

As mentioned the accuracy assessment described here is preliminar, the complete accuracy assessment including biomes and cross-cutting themes is under evaluation.

4. Map Collections and Analysis

The MapBiomas Collections produced so far are listed and summarized below:

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

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

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

• Collection 3 - comprised the period of 1985 to 2017 and was based on random forest decision trees for all biomes and the Coastal Zone, Urban Infrastructure, Mining, Pasture and Agriculture themes.

5. Practical Considerations

The Collection 3 resulted not only in a longer time series but more spatially and temporally consistent annual LULC maps of Brazil. Significant improvements were done in the Collection 3 by applying the random forest classification, such as smoothing of transitions between charts and biomes, as well as the variations in the areas of each class mapped along the time series. However, challenges still remain and more improvements will be done in the next MapBiomas collection. On the other hand, the programming codes for running the MapBiomas algorithms are publicly available and accessible through mapbiomas.org.

6. Concluding Remarks and Perspectives

The proposal algorithms for pre-processing and classifying Landsat imagery hold promise for revolutionizing the production of LULC maps at a large scale. Thanks to Google Earth Engine and open source technology it is possible to access and process large scale datasets of satellite imagery such as the one generated by MapBiomas project. The replication of this type of project is viable for other areas of the planet. The next step of this project is to expand mapping and monitoring of the Pan-Amazon, Chaco and other tropical forest regions. In addition, the project team will improve the Collection 3 with subsequent years (2018 and so long). Future developments include using the entire spectral-temporal information of Landsat data in a per pixel basis and integration with other sensors such as Sentinel-2 and AWiFs-Resourcesat. The data produced are important not only for estimating greenhouse gas emissions, but also for subsidizing several public policies.

7. References

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

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

IBGE, 2012. Manual técnico da vegetação brasileira. 2ed. Rio de Janeiro: IBGE. p.157-160.

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.

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.

APPENDIX

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 and Mining

Appendix 10 - Urban Infrastructure - in progress

ANNEX

Annex I: MapBiomas Network

MapBiomas is an initiative of the Greenhouse Gas Emissions Estimation System (SEEG) from the Climate Observatory's and is produced by a collaborative network of co-creators made up of NGOs, universities and technology companies organized by biomes and cross-cutting themes.

Biomes Coordination:

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

Cross-cutting Themes Coordination:

- Pasture – Federal University of Goias (LAPIG/UFG)
- Agriculture – Agrosatelite
- Coastal Zone and Mining – Vale Technological Institute (ITV) / Solved
- Urban infrastructure – Terras

Technology Partners:

- Google
- EcoStage
- Terras App

Financing:

- Norway's International Climate and Forest Initiative (NICFI)
- Gordon & Betty Moore Foundation
- Arapyaú Institute
- Climate and Land Use Alliance (CLUA)
- Good Energies Foundation

Institutional Partners:

- WRI Brasil
- Institute for Democracy and Sustainability (IDS)
- AVINA Foundation
- The Nature Conservancy (TNC)

- Climate, Forest and Agriculture Coalition

Technical and Scientific Coordination: Carlos Souza (IMAZON)

General Coordination: Tasso Azevedo (SEEG/OC)

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

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

Annex II: Mapping initiatives at global scale, in Brazil, biomes and cross-cutting themes, period/years mapped and respectively references/sources (in portuguese).

Mapa	Referência	Descrição	Escala	Ano do Mapa
Mapa de Uso do Solo do Brasil	IBGE, 2015	Mapa de uso do solo baseado em Modis (250mt) com comparaçÕes entre os anos 2000, 2010 e 2012	1:250 mil	2000, 2010, 2012
Terra Class Cerrado	MMA, IBAMA, EMBRAPA, INPE, UFG e UFU, 2015	Mapeamento do Uso e Cobertura da Terra do Cerrado	1:250 mil	2013
Uso e cobertura vegetal do Estado do Rio Grande do Sul – situação em 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, 2015	Mapa de Cobertura vegetal do Rio Grande do Sul - ano base 2002, obtido por interpretação visual de imagens Landsat. Nível de detalhe compatível com escala 1:250.000	1:250 mil	2002
Uso e cobertura vegetal do Estado do Rio Grande do Sul – situação em 2009.	Weber, E.J.; Hoffmann, G.S.; Oliveira, C.V.; Hasenack, H. (Org.). Porto Alegre: UFRGS IB Centro de Ecologia, 2016. 1a ed. ISBN 978-85-63843-20-3, 2016	Mapa de Cobertura vegetal do Rio Grande do Sul - ano base 2009, obtido por interpretação visual de imagens Landsat. Nível de detalhe compatível com escala 1:250.000	1:250 mil	2009
Uso e cobertura vegetal do Estado do Rio Grande do Sul	Hoffmann, G.S.; Weber, E.J.; Hasenack, H. (Org.). Porto Alegre: UFRGS IB Centro de Ecologia, 2018. 1a ed. ISBN	Mapa de Cobertura vegetal do Rio Grande do Sul - ano base 2015, obtido por interpretação visual de imagens Landsat. Nível de detalhe compatível com escala 1:250.000	1:250 mil	2015

– situação em 2015.	978-85-63843-22-7, 2018			
Atlas da Mata Atlântica	SOS Mata Atlântica/INPE, 2016	Mapeamento das formações florestais e ecossistemas associados, ano de referência 2016	1:50.000	2016
Mapeamento da Bacia do Alto Paraguai	SOS Pantanal/WWF Brasil, 2015	Monitoramento do uso e cobertura vegetal da Bacia do Alto Paraguai, que inclui o Pantanal e suas cabeceiras. Dados disponíveis para 2002, 2008, 2010, 2012 e 2014	1:250.000 adaptado para 1:50.000	2014
Mapa dos manguezais da região norte do Brasil	Nascimento Jr, W.R; Souza-Filho, P.W.M., et al., 2016	Mapa das áreas de manguezal do Oiapoque-AP até a Ponta de Tubarão-MA gerado a partir da classificação de imagens Landsat e ALOS PALSAR do ano de 2008	1:100.000	2008
Mapa dos manguezais da região nordeste do Brasil	Pereira, E.A., Souza-Filho, P.W.M., et al., 2016	Mapa das áreas de manguezal da Ponta de Tubarão-MA até o sul do Estado da Bahia a partir da classificação de imagens Landsat e ALOS PALSAR do ano de 2008	1:100.000	2008
Mapas da agricultura anual para a amazônia	Agrosatélite, 2015	Mapas que incluem as culturas de algodão, milho e soja, cultivados em larga escala, para os anos/safra: 2000/2001, 2006/2007 e 2013/2014, 2016/2017	1:100.000	2000/2001, 2006/2007, 2013/2014 e 2016/2017
Mapas da agricultura anual para o cerrado (projeto: Geospatial analyses of the annual crops dynamic in the Brazilian Cerrado biome: 2000 to 2014)	Agrosatélite, 2015	Mapas que incluem as culturas de algodão, milho e soja, cultivados em larga escala, para os anos/safra: 2000/2001, 2006/2007 e 2013/2014.	1:100.000	2000/2001, 2006/2007 e 2013/2014
Mapa de florestas plantadas	GFW/WRI, 2015	Mapa de florestas plantadas para o Brasil	1:100.000	2014
III Inventário Nacional de Emissões de Gases de Efeito Estufa	MCTIC, 2015	Mapa das áreas de agricultura (anual, semi-perene e perene)	1:250.000	2010

(setor LULUCF)				
Canasat	INPE/Agrosatelite, 2003 a 2015	Mapa da cana-de-açúcar para o centro-sul do Brasil	1:100.000	2003 a 2015
Mapas de Cobertura Vegetal dos Biomas Brasileiros - ProBio	UEFS/APNE/EMBRAPA-Solos/UFCE/UFRN/UFRPE/UFPB/CRA/SEMARH-MMA, 2006	LEVANTAMENTO DA COBERTURA VEGETAL E DO USO DO SOLO DO BIOMA CAATINGA	1:100.000	2002-2004
Mapeamento Florestal do Estado de Sergipe	SEMARH-SE, 2010	Levantamento da cobertura Florestal do Estado de Sergipe	1:250.000	2010
Mapas da Agricultura do Cerrado	Agrosatélite, 2015	Map of the cultivated area of 1st crop season of soybean, corn, and cotton in the Cerrado biome	1:100.000	2000/2001, 2006/2007 e 2013/2014
Mapeamento de uso da terra para o Cerrado e Mata Atlântica	FBDS, 2015	Mapeamento de uso da terra para o Cerrado e Mata Atlântica. Baseado em imagens de alta resolução RapidEye com resolução de 5m.	1:25.000	2013
Atlas da Mata Atlântica	SOS Mata Atlântica/INPE, 2015	Mapeamento detalhado com 1ha para o Estado de São Paulo	1:50.000	2014
Atlas da Mata Atlântica	SOS Mata Atlântica/INPE, 2015	Mapeamento detalhado com 1ha para o Estado do Rio de Janeiro	1:50.000	2014
Atlas da Mata Atlântica	SOS Mata Atlântica/INPE, 2015	Mapeamento detalhado com 1ha para o Estado do Paraná	1:50.000	2014
Atlas da Mata Atlântica	SOS Mata Atlântica/INPE, 2015	Mapeamento detalhado com 1ha para o Estado de Santa Catarina	1:50.000	2014
Global Distribution of Mangroves USGS	USGS, 2011	This dataset shows the global distribution of mangrove forests, derived from earth observation satellite imagery	1:50.000	1997-2000

Mapa Síntese de Pastagens do Brasil - v8	LAPIG/UFG, 2016	Mapeamento de áreas de pastagem, a partir de compilação de dados TerraClass Amazon; Funcate; PROBIO; Canasat e TNC. Os mapeamentos que o compõem são: Bioma Pantanal - Mapeamento da Bacia do Alto Paraguai para 2014; Bioma Caatinga (Mapeamento Lapig - Versão 2 2014-2016) - Esse mapeamento classificou as áreas de pasto limpo, pasto sujo e área degradada (solo exposto); Bioma Mata Atlântica (Mapeamento Lapig 2014-2016) - Esse mapeamento classificou as áreas de pasto limpo, pasto sujo e área degradada (solo exposto); Bioma Pampa - Mapeamento realizado pelo IBGE para o estado do Rio Grande do Sul (2012); Bioma Amazônia - Mapeamento TerraClass Amazônia 2014; Bioma Cerrado - Mapeamento TerraClass Cerrado 2013.	1:250.000	2016
Prodes	INPE, 2016	O projeto PRODES realiza o monitoramento por satélites do desmatamento por corte raso na Amazônia Legal e produz, desde 1988, as taxas anuais de desmatamento na região, que são usadas pelo governo brasileiro para o estabelecimento de políticas públicas. As taxas anuais são estimadas a partir dos incrementos de desmatamento identificados em cada imagem de satélite que cobre a Amazônia Legal. A primeira apresentação dos dados é realizada para dezembro de cada ano, na forma de estimativa. Os dados consolidados são apresentados no primeiro semestre do ano seguinte.	1:250.000	2016

Terra Class Amazônia	INPE, 2016	Como principal resultado deste mapeamento é possível entender a dinâmica de uso e cobertura da Amazônia Legal Brasileira. Para isto já foram mapeados cinco anos de uso e cobertura (2004, 2008, 2010, 2012 e 2014) permitirão uma análise evolutiva de uma década que se inicia no ano da implantação do Plano de Prevenção e Controle do Desmatamento na Amazônia Legal (PPCDAm). Com estes resultados é possível fazer uma avaliação da dinâmica do uso e ocupação das áreas desflorestadas nestes 10 anos da implementação do PPCDAm.	1:250.000	2014
Global Forest Change 2000–2015	University of Maryland, 2015	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).	1:250.000	2015
Land Cover Rondônia 1984 -2010	University of California Santa Barbara (UCSB), 2012	This project is part of a collaborative research effort between UCSB, Salisbury University (Maryland) and North Carolina State University. Under the direction of principal investigators Jill Caviglia-Harris, Daniel Harris (Salisbury), Erin Sills (NCSU) and Dar Roberts (UCSB), we have been active in the production and management of a gapless annual Landsat imagery archive (1984-2009) covering the majority of the state of Rondonia	1:250.000	2010

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

Classe Nivel 1	Classe Nivel 2	Classe Nivel 3	Biomas	Descrição
Floresta	Floresta Natural	Formação Florestal	Amazônia	Floresta Ombrófila Densa, Floresta Estacional Sempre-Verde, Floresta Ombrófila Aberta, Floresta Estacional Semidecidual, Floresta Estacional Decidual, Savana Arborizada, Áreas que sofreram ação do fogo ou exploração madeireira, Floresta resultante de processos naturais de sucessão, após supressão total ou parcial de vegetação primária por ações antrópicas ou causas naturais, podendo ocorrer árvores remanescentes de vegetação primária. Floresta Ombrófila Aberta Aluvial estabelecida ao longo dos cursos de água, ocupa as planícies e terraços periodicamente ou permanentemente inundados, que na Amazônia constituem fisionomias de matas-de-várzea ou matas-de-igapó, respectivamente.
			Caatinga	Tipos de vegetação com predomínio de dossel contínuo - Savana-Estépica Florestada, Floresta Estacional Semi-Decidual e Decidual.
			Cerrado	Tipos de vegetação com predomínio de espécies arbóreas, com formação de dossel contínuo (Mata Ciliar, Mata de Galeria, Mata Seca e Cerradão) (Ribeiro & Walter, 2008), além de florestas estacionais semideciduais.
			Mata Atlântica	Floresta Ombrófila Densa, Aberta e Mista e Floresta Estacional Semi-Decidual, Floresta Estacional Decidual e Formação Pioneira Arbórea.
			Pampa	Vegetação com predomínio de espécies arbóreas, com dossel contínuo. Inclui as tipologias florestais: ombrófila, decidual e semidecidual e parte das formações pioneiras.
			Pantanal	Árvores altas e arbustos no estrato inferior: Floresta Estacional Decidual e Semidecidual, Savana Florestada, Savana-Estépica Florestada e Formações Pioneiras com influência fluvial e/ou lacustre.
		Formação Savânica	Amazônia	n.a.
			Caatinga	Tipos de vegetação com predomínio de espécies de dossel semi-contínuo - Savana-Estépica Arborizada, Savana Arborizada.
			Cerrado	Formações savânicas com estratos arbóreo e arbustivo-herbáceos definidos [Cerrado Sentido Restrito (Cerrado denso, Cerrado típico, Cerrado ralo e Cerrado rupestre) e Parque de Cerrado].
			Mata Atlântica	Savanas, Savanas-Estépicas Florestadas e Arborizadas.
			Pampa	n.a.

			Pantanal	Espécies arbóreas de pequeno porte, distribuídas de forma esparsa e dispostas em meio à vegetação contínua de porte arbustivo e herbáceo. A vegetação herbácea se mistura com arbustos eretos e decumbentes.
		Mangue		Formações florestais, densas, sempre-verdes, frequentemente inundadas pela maré e associadas ao ecossistema costeiro de Manguezal
	Floresta Plantada			Espécies arbóreas plantadas para fins comerciais (ex. eucalipto, pinus, araucária)
Formação Natural não Florestal	Área Úmida Natural Não Florestal		Amazônia	n.a.
			Caatinga	n.a.
			Cerrado	n.a.
			Mata Atlântica	n.a.
			Pampa	Banhados (influência fluvial ou lagunar) e marismas (influência marinha).
			Pantanal	Áreas úmidas de planície, sujeitas a inundações periódicas ou permanentes, localizadas ao longo dos cursos de água e em áreas de depressões que acumulam água. Vegetação herbáceo / arbustivo e/ou de porte arbóreo e formações pioneiras (ex: brejos, cambarazal, paratudal, carandazal).
	Formação Campestre (Campo)		Amazônia	n.a.
			Caatinga	Tipos de vegetação com predomínio de espécies herbáceas (Savana-Estépica Parque, Savana-Estépica Gramíneo-Lenhosa, Savana Parque, Savana Gramíneo-Lenhosa) + (Áreas inundáveis com uma rede de lagoas interligadas, localizadas ao longo dos cursos de água e em áreas de depressões que acumulam água, vegetação predominantemente herbácea a arbustiva).
			Cerrado	Formações campestres com predominância de estrato herbáceo (campo sujo, campo limpo e campo rupestre).
			Mata Atlântica	Savanas e Savanas-Estépicas Parque e Gramíneo-Lenhosa, Estepe e Pioneiras Arbustivas e Herbáceas.
			Pampa	Formações campestres com predominância de estrato herbáceo ou de estrato herbáceo-lenhoso. Inclui também manchas com estrato arbustivo-herbáceo desenvolvido. Em especial nos biomas Pampa e Pantanal, uma parte da área classificada como Formação Campestre inclui também áreas pastejadas.

		Pantanal	Vegetação graminóide / herbácea que possibilitam o uso da Pastagem natural ou plantada para o gado. Em especial nos biomas Pampa e Pantanal, uma parte da área classificada como Formação Campestre inclui também áreas pastejadas.
	Apicum		Apicuns ou Salgados são formações quase sempre desprovidas de vegetação arbórea, associadas a uma zona mais alta, hipersalina e menos inundada do manguezal, em geral na transição entre este e a terra firme.
	Outra Formação Natural não Florestal	Amazônia	Savana, Savana Parque (Marajó), Savana-Estépica (Roraima), Savana Gramíneo-Lenhosa, Campinarana
		Caatinga	n.a.
		Cerrado	n.a.
		Mata Atlântica	n.a.
		Pampa	n.a.
		Pantanal	n.a.
Agropecuária	Pastagem		Áreas de pastagens, naturais ou plantadas, vinculadas a atividade agropecuária. Em especial nos biomas Pampa e Pantanal, uma parte da área classificada como Formação Campestre inclui também áreas pastejadas.
	Agricultura	Cultura Anual e Perene	Áreas predominantemente ocupadas com cultivos anuais e, em algumas regiões (principalmente para a região Nordeste) com a presença de cultivos perenes.
		Cultura Semi-Perene	Áreas cultivadas com a cultura da cana-de-açúcar
		Mosaico de Cultivos (Caatinga)	Áreas ocupadas com diferentes tipos de agricultura (diferentes ciclos e culturas agrícolas) e com cultivos intermitentes ou esporádicos.
	Mosaico de agricultura e pastagem		Áreas de uso agropecuário onde não foi possível distinguir entre pastagem e agricultura.
Área Não Vegetada	Praia e Duna		Cordões arenosos, de coloração branco brilhante, onde não há o predomínio de vegetação de nenhum tipo.
	Infraestrutura Urbana		Áreas urbanizadas com predomínio de superfícies não vegetadas, incluindo estradas, vias e construções.
	Afloramento rochoso	Amazônia	n.a.

		Caatinga	Rochas naturalmente expostas na superfície terrestre sem cobertura de solo, muitas vezes com presença parcial de vegetação rupícola e alta declividade
		Cerrado	Rochas naturalmente expostas na superfície terrestre sem cobertura de solo, muitas vezes com presença parcial de vegetação rupícola
		Mata Atlântica	Rochas naturalmente expostas na superfície terrestre sem cobertura de solo, muitas vezes com presença parcial de vegetação rupícola e alta declividade
		Pampa	Rochas naturalmente expostas na superfície terrestre sem cobertura de solo, muitas vezes com presença parcial de vegetação rupícola
		Pantanal	n.a.
	Mineração		Áreas referentes a extração mineral de grande porte, havendo clara exposição do solo por ação de maquinário pesado. Somente são consideradas áreas pertencentes a malha digital do DNPM (SIGMINE).
	Outra Área não Vegetada	Amazônia	Áreas urbanizadas com predomínio de superfícies não vegetadas, incluindo estradas, vias e construções.
		Caatinga	Áreas de superfícies não permeáveis (infra-estrutura, expansão urbana ou mineração) não mapeadas em suas classes
		Cerrado	n.a.
		Mata Atlântica	Áreas de superfícies não permeáveis (infra-estrutura, expansão urbana ou mineração) não mapeadas em suas classes
		Pampa	n.a.
		Pantanal	Áreas de solo exposto (principalmente solo arenoso) não classificadas na classe de Formação Campestre ou Pastagem.
Corpos D'água	Rio, Lago e Oceano		Rios, lagos, represas, reservatórios e outros corpos d'água
	Aquicultura		Área referente a lagos artificiais, onde predominam atividades aquícolas e/ou de salicultura
Não Observado			

Annex IV: Classes of land cover and land use of of Collections 1.0, 2.0 and 2.3 of MapBiomas

Collection 1 Classes
Forest
Forest in Coastal Zone
Forest Plantation
Agriculture
Pasture
Water
Other
Non-Observed

Collection 2 Classes
1. Forest
1.1. Natural Forest Formations
1.1.1. Dense Forest
1.1.2. Open Forest
1.1.3. Mangrove
1.1.4. Flooded Forest
1.1.5. Degraded Forest
1.1.6. Secondary Forest
1.2. Silviculture
2. Non-Forest Natural Formations
2.1. Non-forest Natural Wetlands
2.2. Grasslands
2.3. Other non-forest natural formations
3. Farming
3.1. Pasture
3.1.1. Pastude in natural grasslands
3.1.2. Other pasture
3.2. Agriculture
3.2.1. Annual crops
3.2.2. Semi-Perennial crops

3.2.3. Mosaic of crops
3.3 Agriculture or Pasture
4. Non-Vegetated areas
4.1. Dunes and Beach
4.3. Other non-vegetated areas
4.2. Urban Infrastructure
5. Water
6. Non-Observed

Collection 2.3 Classes
1. Forest
1.1. Natural Forest
1.1.1. Natural Forest Formation
1.1.2. Savanna Formation
1.1.3. Mangrove
1.2. Forest Plantations
2. Non-Forest Natural Formations
2.1. Non-forest Natural Wetlands
2.2. Grasslands
3. Farming
3.1. Pasture
3.2. Agriculture
3.3 Agriculture or Pasture
4. Non-Vegetated areas
4.1. Beach and dune
4.3. Other non-vegetated areas
4.2. Urban Infrastructure
5. Water
6. Non-Observed