Pantanal - Appendix

Collection 11

Version 1

General coordinator

Eduardo Reis Rosa

Team

Marcos Reis Rosa

Mariana Dias

Julya Paes de Souza

August 2026

SUMMARY

1. ABSTRACT

This document provides the methodological description for Collection 11 of the MapBiomas project in the Pantanal biome, covering the period from 1985 to 2025. The classification methodology centers on the Random Forest algorithm using Landsat surface reflectance mosaics. These mosaics were generated 'on the fly' from monthly surface reflectance composites, utilizing deep learning methods to effectively recover information and reduce noise from clouds and shadows. Significant improvements in this collection include the implementation of monthly mapping for the addition class of the Flooded Savanna, and a refined feature space based on an analysis of variable importance across eight subregions. The model was trained using stable samples from previous collections and complementary samples collected through visual inspection to ensure high reliability. To maintain temporal consistency, several post-classification filters were applied, including gap filling, transition rules for forest and savanna formations, and a regeneration filter to prevent misclassification between pasture and natural grasslands. The final maps integrate cross-cutting themes such as agriculture and urban areas, achieving an overall accuracy of 85.9% at Level 1 for Collection 11.

How to cite: MapBiomas Project – Collection 11, Pantanal Appendix, Version 1, published August 2026. Available at: https://github.com/mapbiomas/brazil-pantanal

2. INTRODUCTION

The Pantanal biome is located within Brazilian territory between the states of Mato Grosso and Mato Grosso do Sul. It also extends into the territories of Bolivia and Paraguay. It covers an area of 150,000 km² and is surrounded by the Amazon, Cerrado, and Atlantic Forest biomes. Despite being the smallest Brazilian biome in area, it boasts the highest proportion of remaining native vegetation.

The Pantanal is renowned as one of the largest continuous wetlands in the world, characterized by a vast alluvial plain characterized by a seasonal flood pulse, with pronounced flood and dry periods. (Figure 1). This flood pulse is influenced by the interaction between global and regional climatic conditions and the hydro-sedimentological dynamics of the rivers in the Upper Paraguay Basin (Hamilton; Souza; Coutinho, 1998; Padovani, 2010).

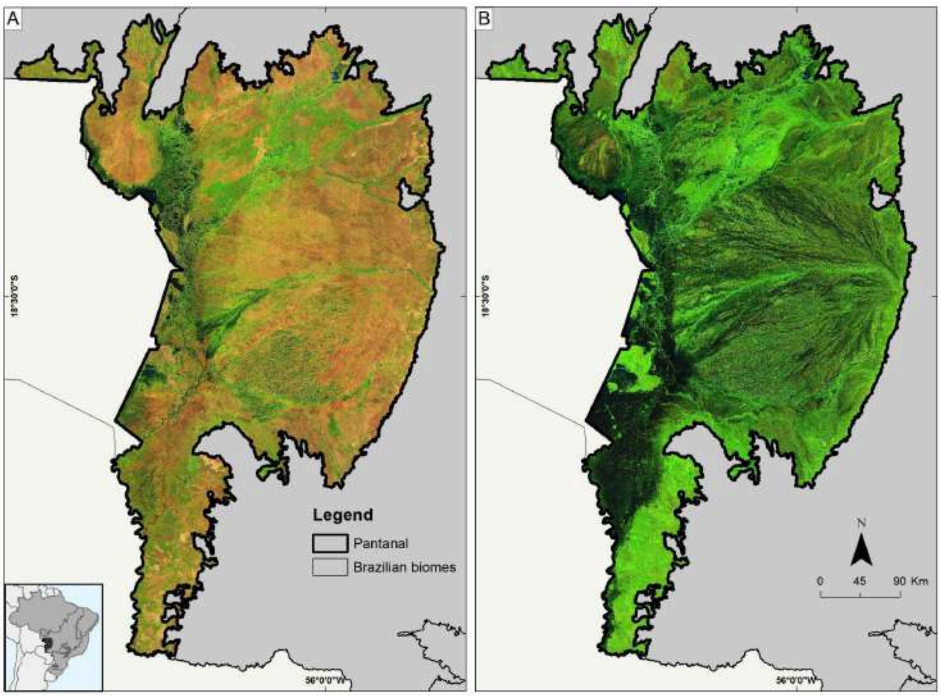


Figure 1: Location of the Pantanal biome showing RGB color composites of TM bands 5, 4 and 3 for Landsat mosaics from 1988. A) Dry-season Landsat mosaic; B) Wet-season Landsat mosaic.

Understanding the dynamics of flooding in the biome is crucial to comprehending changes in land use and land cover over time. As part of MapBiomas Collection 11, we conducted monthly mapping of water and wetland classes and made other improvements compared to previous collections, for example, the addition of the flooded savanna class.

This document describes the methodological flow used to map the land use and land cover in the Pantanal biome from 1985 to 2025. Also, all codes used are available in our public GitHub repository (https://github.com/mapbiomas/brazil-pantanal).

3. OVERVIEW OF CLASSIFICATION METHOD

The initial classification of the Pantanal biome within the MapBiomas project consisted of applying decision trees to generate annual maps of the predominant native vegetation (NV) types, distinguished as: Forest, Savanna, and Grassland as well as anthropic classes such as Pasture and Agriculture. The method used to generate these annual maps evolved, with significant improvements from the first MapBiomas Collection to the present. Collection 1.0 covered the period of 2008 to 2015 and was published in 2016. Collections 2.0 and 2.3 covered the period of 2000 to 2016 and were published in 2018.

The classification using Random Forest was implemented in Collection 2.3, and from this point onward, the empirical decision tree was used to generate stable samples, which were classified as the same NV type over the considered period (2000-2016). These stable samples were used to train the Random Forest models to classify the entire time series. Collections 3 and 3.1 expanded the period covered to 1985–2017.

Collections 4 and 5 used training samples based on the stable samples from the previous collection and reference maps. Collection 6.0 used stable samples from collection 5.0. Collections 7 and 7.1 (published in 2022) used stable samples from collection 6.0. Collection 8.0 used stable samples from Collection 7.1. Collection 9.0 used stable samples from Collection 8.0. Collection 10.0 used stable samples from Collection 9.0. Collection 11 covered the period from 1985 to 2025 and used stable samples from Collection 10.1, and also stable samples from collections 10.1, 9.0 and 8.0 (Table 1).

Table 1. The evolution of the Pantanal mapping collections in the MapBiomas Project, their periods, hierarchical level and number of classes, brief methodological description, and overall accuracy at levels 1 and 2.

Collection	Period	Levels /N. Classes	Method	Overall Accuracy
Beta & 1	8 years 2008-2015	1 / 7	Empirical Decision Tree	
2.0 & 2.3	16 years 2000-2016	3 / 13	Empirical Decision Tree & Random Forest (2.3)	
3.0 & 3.1	33 years 1985-2017	3 / 19	Random Forest	Level 1: 73.2% Level 3: 62.1% *
4.0 & 4.1	34 years 1985-2018	3 / 19	Random Forest	Level 1: 80.7% Level 3: 71.0% *
5.0	35 years 1985-2019	4 / 21	Random Forest	Level 1: 85.1% Level 3: 78.3% *
6.0	36 years 1985-2020	4 / 24	Random Forest	Level 1: 81.6% Level 2: 73.5%
7.0	37 years 1985-2021	4 / 27	Random Forest	Level 1: 85.9% Level 2: 79.5%
7.1	37 years 1985-2021	4 / 27	Random Forest	Level 1: 85.6% Level 2: 77.3%
8.0	38 years 1985-2022	4 / 29	Random Forest	Level 1: 84.0% Level 2: 77.6%
9.0	39 years 1985-2023	4 / 29	Random Forest	Level 1: 85.4% Level 2: 79.2%
10.1	40 years 1985-2024	4 / 30	Random Forest	Level 1: 87.5% Level 2: 81.3%
11.0	41 years 1985-2025	4 / 33	Random Forest	Level 1: 85.9% Level 2: 75.6%

* Due to changes in the hierarchy of the forest classes , level 2 of Collections 6 - 11 are being compared to level 3 of previous collections.

The production of the Collection 11.0 land use and land cover maps in the Pantanal biome followed steps similar to those used in previous Collections (4 - 10). The use of monthly data for mapping water and flooded areas follows the methodology used since Collection 9.0. However, some improvements were implemented, particularly in the mosaics, sample balancing, the feature space, the annual water and flooded area classification based on monthly data, post-classification filters, and the addition of a new class mapped: Flooded Savanna. The classification workflow is shown in Figure 2 and each step is described in the following sections.

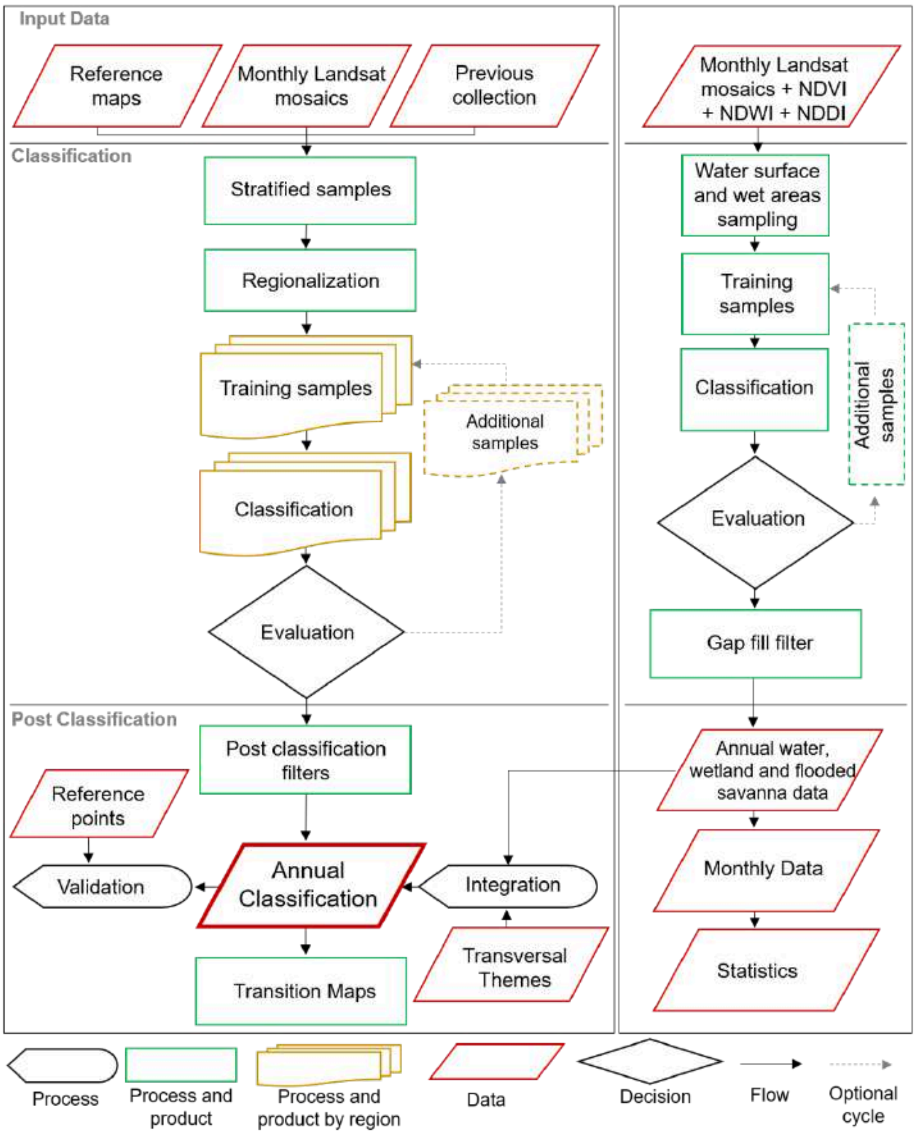


Figure 2: Workflow for generate Collection 11 land use and land cover maps for the Pantanal biome.

4. LANDSAT IMAGE MOSAICS

Until Collection 5.0, the classification was performed by using top-of-atmosphere (TOA) data from the Landsat 5 (TM), 7 (ETM+), and 8 (OLI) sensors. In Collection 6.0, we discontinued the use of TOA data and adopted surface reflectance (SR) data. From Collection 7.0 until Collection 9.0, we adopted the 'USGS Landsat Level 2, Collection 2, Tier 1' dataset available on the Google Earth Engine platform, maintaining the use of SR data.

The Landsat mosaics for Collections 10 and 11 were generated 'on the fly' from the Landsat series monthly mosaic asset, which is already atmospherically corrected and provides surface reflectance values (specifically, the 'LANDSAT/ COMPOSITES/ C02/ T1_L2_32DAY' asset). Due to these corrections and the use of deep learning methods to recover data information, the mosaics have fewer gaps than those used in previous collections. With reduced noise from cloud and shadow removal, it was possible to use all monthly images to generate the annual mosaic.

The mosaic is generated using six original Landsat spectral bands: blue, red, green, NIR, SWIR1, and SWIR2. Afterwards, spectral indices are calculated and appended as bands to the mosaic. Auxiliary layers are also added as bands, including green-band entropy (texture) and a set of terrain and geomorphometric variables: MERIT DEM elevation, slope, and Geomorpho90m products (roughness, CTI, eastness and northness), of which elevation and roughness were retained in the final feature space of most regions. All of these bands constitute the candidate feature space used to classify the Pantanal biome. This candidate pool is the same set later screened by the feature-importance analysis described in Section 5.3 (roughly 160 candidate variables), from which the final per-region subsets of about 50 bands were drawn.

This process was conducted for each year (1985-2025), resulting in the annual Landsat mosaics that were subsequently submitted for classification. Most of the annual mosaics exhibit high quality and coverage rates; however, some years, such as 1985, 1986, and 1992, have lower image availability and thus require a more extensive post-classification procedure.

5. REGIONALIZATION

The Pantanal biome was divided into eight regions (Figure 3) based on dry and wet patterns, sub-basins and native vegetation distribution presented on different regionalization approaches (Silva & Abdon, 1998; Assine, 2015). The aim of this process was to reduce classification confusion and noise , improving the sample balance in more homogeneous regions.

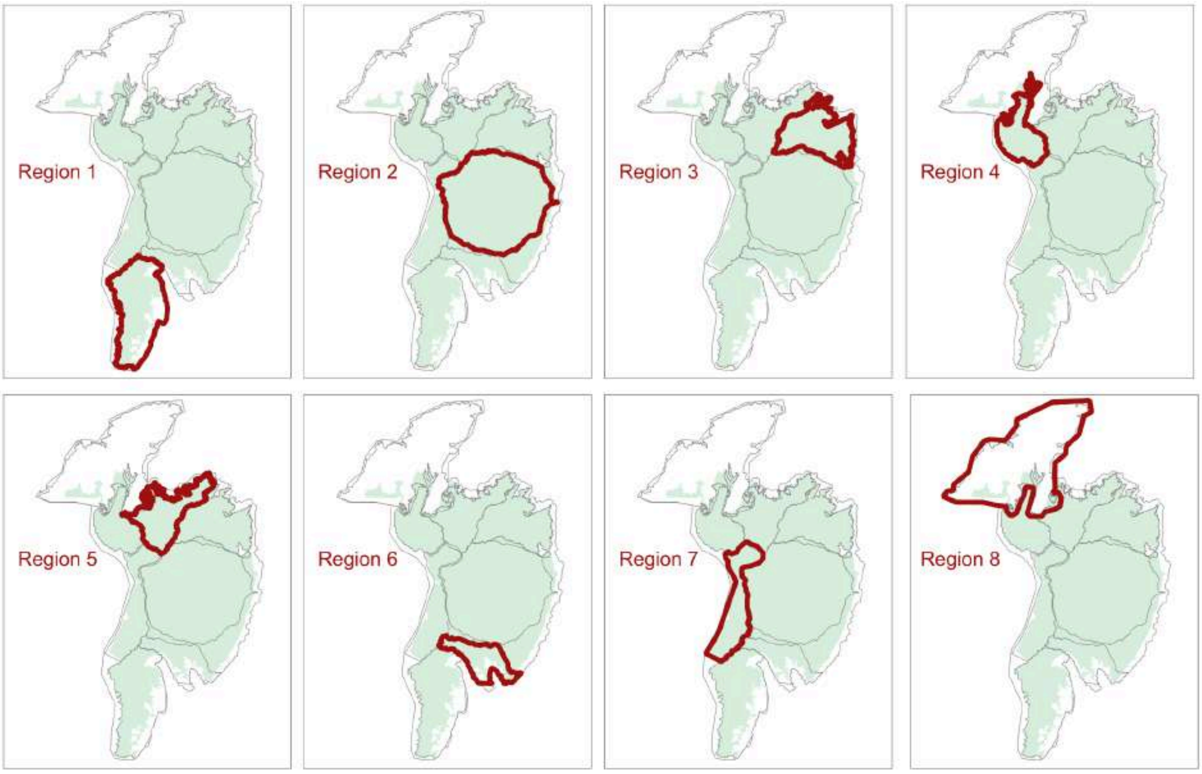
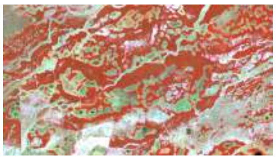
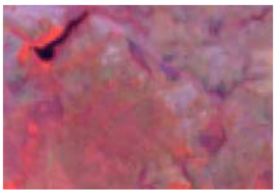


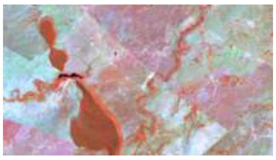
Figure 3: Regionalization used in the classification of the Collection 11 in the Pantanal biome.

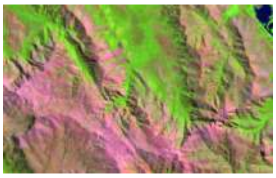
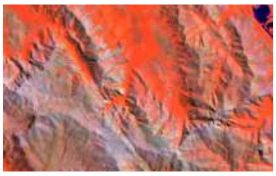
6. CLASSIFICATION

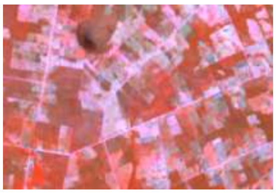
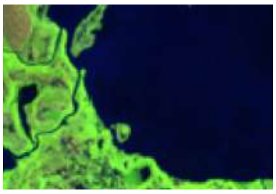
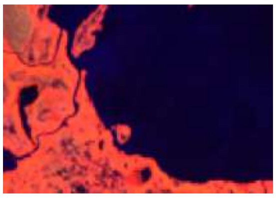
The digital classification of the Landsat mosaics for the Pantanal biome aimed to identify a subset of 10 classes from the complete legend of MapBiomas Collection 11 (Table 2), which were subsequently integrated with the cross-cutting themes in a further step (see section 7).

Table 2: Classes mapped by the Pantanal biome in the MapBiomas Collection 11.

1.1. Forest Formation			
Numeric ID: 3		Color #1f8d49	
Google Earth	Landsat (RGB 654)	Landsat (RGB 564)	Photo
			
Description: 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.			
1.2. Savanna Formation			
Numeric ID: 4		Color #7dc975	
Google Earth	Landsat (RGB 654)	Landsat (RGB 564)	Photo
			
Description: Small tree species, sparsely arranged in the shrub and herbaceous continuous vegetation. The herbaceous vegetation mixes with erect and decumbent shrubs.			
1.3. Flooded Savanna			
Numeric ID: 7		Color #228c70	
Google Earth	Landsat (RGB 654)	Landsat (RGB 564)	Photo
			
Description: Tree species, sparsely distributed and interspersed amidst continuous shrubby and herbaceous vegetation, established along watercourses and in the floodplain.			

2.1. Wetland			
Numeric ID: 11		Color #519799	
Google Earth	Landsat (RGB 654)	Landsat (RGB 564)	Photo
			
Description: 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 within the landscape 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 (e.g. swamps). Areas with a water surface, but difficult to classify due to the amount of macrophytes, eutrophication or sediments, were also included in this category.			
2.2. Grassland Formation			
Numeric ID: 12		Color #d6bc74	
Google Earth	Landsat (RGB 654)	Landsat (RGB 564)	Photo
			
Description: 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 topographic gradients and pasture management (livestock). Patches of invasive exotic vegetation or forage use (planted pasture) may be present, forming mosaics with native vegetation.			
2.4. Rocky Outcrop			
Numeric ID: 29		Color #ad5100	
Google Earth	Landsat (RGB 654)	Landsat (RGB 564)	Photo

			
Description: Naturally exposed rocks without soil cover, often with the partial presence of rupicolous vegetation and steep slopes.			
3.1. Pasture			
Numeric ID: 15		Color #edde8e	
Google Earth	Landsat (RGB 654)	Landsat (RGB 564)	Photo
			
Description: 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. This class may occur on recently deforested areas, even without farming activities having started yet.			
3.2.1.5. Other Temporary Crops			
Numeric ID: 41		Color #f54ca9	
Google Earth	Landsat (RGB 654)	Landsat (RGB 564)	Photo
			
Description: Areas occupied by short- or medium-duration agricultural crops, generally with a vegetative cycle of less than one year.			
3.4. Mosaic of Uses			
Numeric ID: 21		Color #ffefc3	
Google Earth	Landsat (RGB 654)	Landsat (RGB 564)	Photo

			
Description: This class may include peri-urban areas, farmsteads, and transition areas from the pasture class to secondary vegetation in natural regeneration, or to degraded and/or abandoned pastures.			
5. Water			
Numeric ID: 33		Color #2532e4	
Google Earth	Landsat (RGB 654)	Landsat (RGB 564)	Photo
			
Description: Rivers, lakes, dams, reservoirs and other water bodies.			

The classes known as Forest, Savanna, Grassland, Pasture, Rocky Outcrop, Other Temporary Crops and Other non-vegetated areas were mapped based on annual Landsat mosaics. On the other hand, the water, flooded savanna and wetland classes represent a minimum frequency of two monthly Landsat mosaics classified as these wetland classes.

6.1. Stable Samples

The extraction of stable samples from the previous Collection 10.1 followed several steps, aiming to ensure their reliability for use as training areas. The areas that did not change class from 1985 to 2024 in Collection 10.1 were sampled using a class- and region-stratified scheme (ee.Image.stratifiedSample), with per-class quotas defined by region (e.g. 2,500 and 1,500 points for the dominant natural classes, 1,000 and 800 for secondary classes, and 500 for rarer strata) rather than a flat number of random points. In addition, a second stratum was drawn with the same stratified scheme from areas that did not change in a given year across Collections 10.1, 9 and 8, in order to obtain samples that are stable not over time but across collections built with different approaches. During the classification, sample balancing was performed by adjusting the number of stable samples for each class.

6.2. Complementary samples

The need for complementary samples was evaluated by visual inspection. Complementary sample were collection in the Google Earth Engine Code Editor with the drawing polygons tool. The same concept of stable samples was applied, checking the false-color composites of the Landsat mosaics for all the 41 years during the polygon drawing. Based on knowledge of each region, samples of forest, savanna, grassland, agriculture or pasture were added.

6.3. Feature Space

For this collection, a feature importance analysis was conducted for each region. A feature importance test was performed using a pool of 250 different bands, which included annual median, standard deviation, and quartile values from the spectral bands and a wide range of indices, in addition to texture and clustering bands. The test involved running the classifier with the full set of candidate bands for multiple years and for each region. Using a classifier composed of 200 decision trees, it was possible to obtain a ranking of the most relevant bands considered in the classification.

Subsequently, the 60 most relevant bands for each region were selected to compose the feature space subset used in the final classification of each region. Table 3 shows the list of the most important bands for the classification of the Pantanal biome.

Table 3. Feature space subset with the most frequent variables between regions considered in the classification of the Pantanal biome in Landsat image mosaics in the MapBiomas Collection 11 (1985-2025).

Type	Name	Formula	Statistics	Reference
Landsat band	Blue	Band 1 (L5 and L7) Band 2 (L8)	median, median_dry, median_wet	USGS
	Green	Band 2 (L5 and L7) Band 3 (L8)	median, median_dry, median_wet, median_texture	USGS

Type	Name	Formula	Statistics	Reference
	Red	Band 3 (L5 and L7) Band 4 (L8)	median, median_dry, median_wet	USGS
Landsat band	NIR	Band 4 (L5 and L7) Band 5 (L8)	median, median_dry, median_wet	USGS
	SWIR 1	Band 5 (L5 and L7) Band 6 (L8)	median, median_dry, median_wet	USGS
	SWIR 2	Band 7 (L5 and L7) Band 8 (L8)	median, median_dry, median_wet	USGS
	Advanced Vegetation Index	AVI = (NIR × (1.0 - Red) × (NIR - Red))$^{1/3}$	median, median_dry, median_wet	
Spectral Index	Automated Water Extraction Index	AWEI = Blue + 2.5 × Green- 15 × (NIR + SWIR1) - 0.25 × SWIR2	median, median_dry, median_wet	Feyisa et al., 2014
	Band Ratio for Built-up Area	BRBA = Red / SWIR2	median, median_dry, median_wet	Waqar et a., 2012
	Brightness Index	BI = ((Red^2 + Green^2) / 2)$^{1/2}$	median, median_dry, median_wet	Escadafal, 1989
	Enhanced Vegetation Index	EVI = 2.5 × (NIR - Red) / (NIR + 2.4 × Red + 1)	median_wet,	Parente et al., 2018
	Enhanced Vegetation Index 2	EVI 2 = 2.5 × (NIR - Red) / (NIR + 2.4 × Red + 1)	median, median_dry, median_wet, stdDev	Parente et al., 2018
	Global Environment Monitoring Index	GEMI = (2 * (NIR^2 - Red^2) + 1.5 * NIR + 0.5 * Red) / (NIR + Red + 0.5)	median, median_dry, median_wet	Pinty & Verstraete, 1992

Type	Name	Formula	Statistics	Reference
Spectral Index	Green Chlorophyll Vegetation Index	GCVI = (NIR / Green - 1)	amplitude, median, median_dry, median_wet	Burke et al., 2017
Spectral Index	Modified Normalized Difference Water Index	MNDWI = (Green - SWIR1) / (Green + SWIR1)	median, median_dry, median_wet,	Xu, 2006
	Normalized Difference Drought Index	NDDI = (NDVI - NDWI) / (NDVI + NDWI)	median_wet	Gu et al., 2007
	Normalized Difference Vegetation Index	NDVI = (NIR - Red) / (NIR + Red)	amplitude, median, median_dry, median_wet	Rouse et al., 1974
	Normalized Difference Water Index	NDWI = (NIR - SWIR1) / (NIR + SWIR1)	amplitude, median, median_dry, median_wet	Gao et al., 1996
	Optimized Soil Adjusted Vegetation Index	OSAVI = ((NIR - Red) / (NIR + Red + 0.16)) × (1 + 0.16)	median, median_wet	Rondeaux et al., 1996
	Ratio Vegetation Index	RVI = Red/NIR	median, median_wet	Jordan, 1969
	Redness Index	RI = (Red - Green) / (Red + Green)	median, median_wet,	Escadafal & Huete, 1991
	Soil-Adjusted Vegetation Index	SAVI = 1.5 × (NIR - Red) / (NIR + Red + 0.5)	median_wet, amplitude, median, median_dry, median_wet	Huete, 1988
Surface Index	Hall's Forest Cover	HFC = −0.017 × RED − 0.007 × NIR − 0.079 × SWIR2 + 5.22	median, median_dry, median_wet	Hall et al., 2006
	Hall's Forest Height	HFH = −0.039 × RED − 0.011 × NIR − 0.026 × SWIR1 + 4.13	median, median_dry, median_wet	

Type	Name	Formula	Statistics	Reference
Terrain	Slope	ALOS DSM: Global 30 m	identity	Tadono et al., 2014
Coords	Latitude and Longitude		Latitude, Longitude	

6.4. Classification algorithm

A digital classification was performed year by year using a variation of the Random Forest algorithm (Breiman, 2001), available in Google Earth Engine. The annual classifier (ee.Classifier.smileRandomForest) was configured with 200 trees and one variable per split (variablesPerSplit = 1), with a fixed random seed for reproducibility. Random Forest was trained with stable and complementary samples, using the corresponding feature space subset. Each subregion had a specific sample balance and classification model . After the classification processing, the subregions were merged to form the Pantanal biome territory.

6.5. Rocky Outcrop Classification

The mapping of the rocky outcrop was conducted in a parallel process, starting with a specific regionalization of the Serra do Amolar area. This was followed by clustering of the region and a supervised classification using the Random Forest algorithm with manually collected samples. The feature space for this binary classification was composed of the SAVI, EVI, GCVI and BSI indexes.

Considering that the rocky outcrop in this region is a stable area over the time series, a frequency filter was applied to stabilize the raw mapping, which had shown a tendency to vary depending on wet and dry years. This class can be found on the map with the code 29.

6.6. Water and Flooded Monthly and Annual Classification

We performed a monthly classification to map water, savanna and wetland classes, aiming to map the maximum floodplain area each year. To guarantee the monthly classification of the maximum available images, we used monthly mosaics of each satellite and also mosaics composed of data from different satellites, if they were available for the same period (Table 4). Note that the monthly water/wetland classification uses the individual

Landsat Collection 2 Tier 1 Level 2 scenes per sensor, following the sensor-by-period scheme in Table 4, whereas the annual mosaics used for the non-wet classification are generated on the fly from the unified LANDSAT/COMPOSITES/C02/T1_L2_32DAY asset, which selects the best available pixel across sensors and is therefore not governed by Table 4.

Table 4: Satellites for each period

Satellite	Years
Landsat 5	1985-1999; 2003
Landsat 5 + 7	2000-2002; 2004-2011
Landsat 7	2012
Landsat 8	2013-2021
Landsat 8 + 9	2022-2025

A sampling toolkit was developed to collect water and wetland points on the monthly Landsat mosaics (Figure 4).

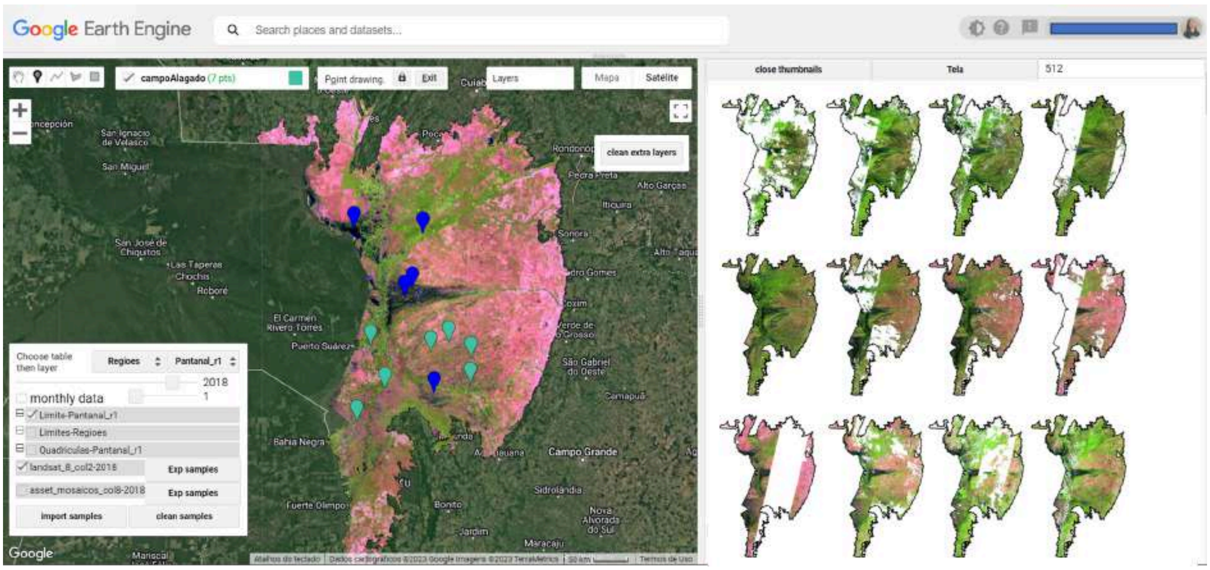


Figure 4: Monthly sampling toolkit.

For classes considered 'non-wet', we randomly sampled points in pixels that had a stable classification over the 40 years of the Collection 10.0. The Random Forest classifier was trained using the sample points and all Landsat bands plus Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index (NDWI) and Normalized Difference Drought Index (NDDI).

Then, a supervised classification was carried out using the Random Forest (RF) algorithm (Breiman, 2001) to classify each monthly mosaic into three classes: water, wet and 'non-wet' areas.

The RF classification result was used to mask the 'non-wet' monthly mosaic areas, and then a water-wetland separation threshold was defined for each satellite based on the NDDI value and applied on the 'wet' RF masked area. In this Collection, the NDDI-based separation threshold was set at 1300 for Landsat 8/9, 1200 for Landsat 7 and mixed-sensor mosaics, and 1100 for Landsat 5.

Some monthly Landsat mosaics had no data pixels because of a lack of images or the cloud mask filter action. To overcome this challenge, we developed a 'gap fill' filter on the classified data. The filter considered the local precipitation pattern (Figure 5) to fill each no data pixel with the classification of the closest and driest month available. This was considered a conservative strategy to not overestimate the mapping of the Pantanal flooded areas in periods without images, since the wettest months are also those with the highest levels of rainfall and, therefore, the cloudiest.

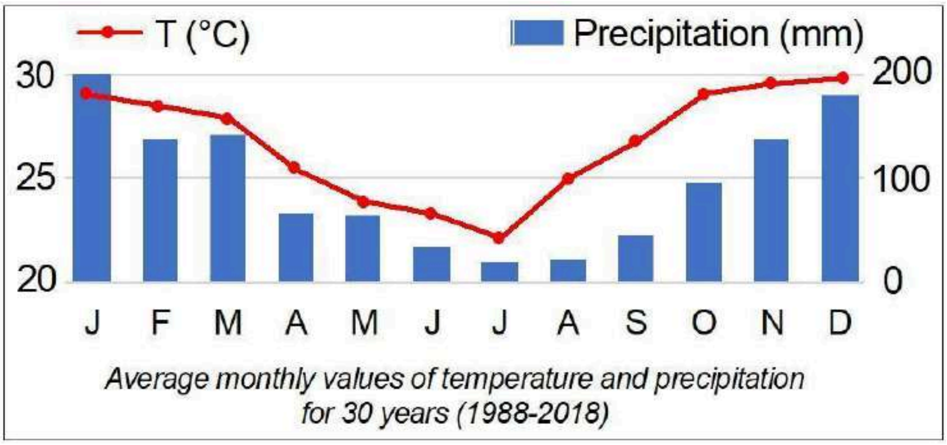


Figure 5: Climograph from 1988 to 2018 with CHIRPS (Climate Hazards Group InfraRed Precipitation with Station) precipitation data and MODIS (Moderate-Resolution Imaging Spectroradiometer) temperature data. (Funk; Peterson; Landsfield, 2015).

With the available monthly water and wetland classifications for each year since 1985, it was possible to generate frequency data. To integrate these into the annual maps for MapBiomas Collection 11.0, we included all pixels classified as either water surface or wet area in at least two months of a given year. In cases where a pixel was classified as both within the same year, the water class was given precedence over the wet class, wet classes crossing savanna classification was turned into flooded savanna, wet classification crossing grassland classification was turned into wetland classification.

The Flooded Savanna class (ID 7) is specifically identified when a pixel classified as Savanna Formation (ID 4) in the annual mosaic intersects a minimum frequency of two monthly classifications of water or wetland within the same year. This rule addresses the historical challenge of identifying flooding beneath woody canopies using only annual spectral composites.

Furthermore, we incorporated the annual data classified by the MapBiomas Water methodology to supplement the mapped water surface area for each year. This was done because the MapBiomas Water data tends to provide an information gain for both smaller and more stable (permanent) water bodies.

7. POST CLASSIFICATION

7.1. Masks

Some natural grassland areas in the western portion of the Pantanal can be easily confused and classified as exotic pastures by RF because of overgrazing or the soil exposure during a dry period. To avoid this misclassification, a mask of 'non-pasture' was manually drawn based on historical reference maps (Ci et al., 2009), PRODES data (Assis et al., 2019) and high-resolution images, and then applied to exclude pixels misclassified as pasture. For the same purpose, a second 'non-pasture' mask was applied considering flooded pixels (classified monthly as water or wetland) in more than 75% of the entire time series.

Also, in order to improve the classification for the most recent years, a mask of 'deforested areas' was built from the MapBiomas Alerta validated polygons. Subsequently , all native vegetation pixels intersecting the mask were reclassified to the farming class in the deforestation event year and in the following years. It is worth noting that these deforestation polygons are available from 2019. The integration of MapBiomas Alerta validated polygons ensures that land use transitions in recent years (since 2019) are anchored in verified deforestation events, thereby enhancing the overall reliability and consistency of the time series.

7.2. Water, Wetland and Flooded Savanna data integration

Annual water and wetland data were added to the map of 'non-wet' classes only in areas classified as grasslands or savanna, considering that the methodology for the latter is not adapted to identify flooded forest. This is also a strategy to avoid false positives from humid areas in the shade of the relief, urban areas or roads.

7.3. Filters

Some temporal and spatial filters were applied to the annual maps to correct classification errors or invalid class transitions using long-term information from the temporal series.

7.3.1 Gap Fill filter

In this filter, no-data values ("gaps") are theoretically not allowed and are replaced by the temporally nearest valid classification. In this procedure, if no subsequent valid position is available, then the no-data value is replaced by its previous valid class. Therefore, gaps should only exist if a given pixel has been permanently classified as no-data throughout the entire temporal domain.

7.3.2. Transitions Forest Formation x Savanna Formation

Forest (3) and Savanna (4) formations are stabilized with a set of rules based on the frequency of presence and on the trajectory of each pixel , rather than on a single fixed percentage. The main rules are: where savanna presence is greater than or equal to forest presence, forest is reclassified as savanna, and vice-versa; forest pixels with an absence-alternation-absence trajectory (trajectory type 6) are reclassified as savanna; savanna pixels with trajectory 6 and fewer than six years of presence are reclassified as grassland (12); and forest pixels with more than three changes and a presence-alternation-presence trajectory (type 5) are reclassified as savanna. These rules produce a more stable classification of the natural classes and remove noise, including in the first and last years of the series.

7.3.3. Moving window temporal filter

The temporal filter (p02_temporalFilter) uses neighbouring years to replace pixels with invalid transitions and is applied over the full 1985–2025 series. Before the moving windows, a remap step forces water (33) to prevail and collapses a set of transitional codes back to grassland (12) or to the non-vegetated class (25). The first step uses a 3-year moving window that corrects a central year when the previous and next years share the same class; it is applied sequentially to classes 19, 4, 12, 21 and 3. The second step uses a 4-year moving window (processed in odd and even years) applied sequentially to classes 19, 21, 4, 12, 3 and 25. The first and last years of the series are preserved unaltered by both windows.

7.3.4. Regeneration filter

A regeneration filter prevents Forest or Savanna formations converted to anthropic use from being reclassified as Natural Grassland, reducing confusion between pasture and grassland. In practice the filter is triggered when a pixel is classified as pasture (21) in two consecutive years, after which the pasture class is propagated (locked) for the current year and for up to seven subsequent years. This filter was used to ensure that any area converted to pasture stays as pasture for, at least, seven consecutive years to avoid some confusion in the classification.

Additionally, a forest stabilization filter was implemented in the time series from 1985 to 2025. Using the Stable Forest data, confusion between savanna and wetland areas that overestimated the Forest class, especially in the early years, was corrected.

Still, to avoid false regeneration in agriculture areas, we used the agricultural data from the previous year and updated the classification year by year from 1985 onward. This filter considers the field knowledge in the Pantanal (in the southern part of subregion 6).

7.3.5. Trajectory filter

To avoid false transitions between savanna formation and grassland in short time intervals, a filter was applied that considered the trajectory of the pixel according to Pontius (2022). Considering the trajectory of absence-alternation-absence, the number of changes that the class was involved in and also the pixel mode along the time series, the filter stabilizes these two classes.

7.3.6. Spatial filter

A spatial filter was also applied to avoid unwanted modifications to the edges of the pixel groups, using the "connectedPixelCount" function. Native to the GEE platform, this function locates connected pixels (neighbors) with the same value. Thus, only pixels that do not share connections to a predefined number of identical neighbors are considered isolated. In this filter, connectivity is evaluated using eight-connected neighborhoods and at least six connected pixels are needed to reach the minimum connection value; isolated groups below this threshold are replaced by the local mode. Consequently, the minimum mapping unit is directly affected by the spatial filter applied, and it was defined as 6 Landsat pixels (~0.5 ha). This is the final step before integration with cross-cutting themes.

8. INTEGRATION WITH CROSS-CUTTING THEMES

The final map generated by the Pantanal biome team was integrated with maps from some cross-cutting themes, which represent rare classes or classes that demand a specific mapping strategy. These external themes comprise urban areas (24), mining (30), agriculture (39, 20, 41), forest plantation (9) and exotic pasture (15). They were superimposed on the Pantanal data (Figure 7), resulting in final maps of the biome containing 14 classes (Table 5).

Table 5: Classes observed in the Pantanal biome.

Legend class of Collection 11	Numeric ID	Color
1.1. Forest Formation	3	
1.2. Savanna Formation	4	
1.3. Flooded Savanna	7	
2.1. Wetland	11	
2.2. Grassland Formation	12	
2.4. Rocky Outcrop	29	
3.1. Pasture	15	

Legend class of Collection 11	Numeric ID	Color
3.2.1.1. Soybean	39	
3.2.1.2. Sugar cane	20	
3.2.1.5. Other Temporary Crops	41	
3.3. Forest Plantation	9	
3.4. Mosaic of Uses	21	
4.2. Urban area	24	
4.3. Mining	30	
4.4. Other non vegetated area	25	
5. Water	33	

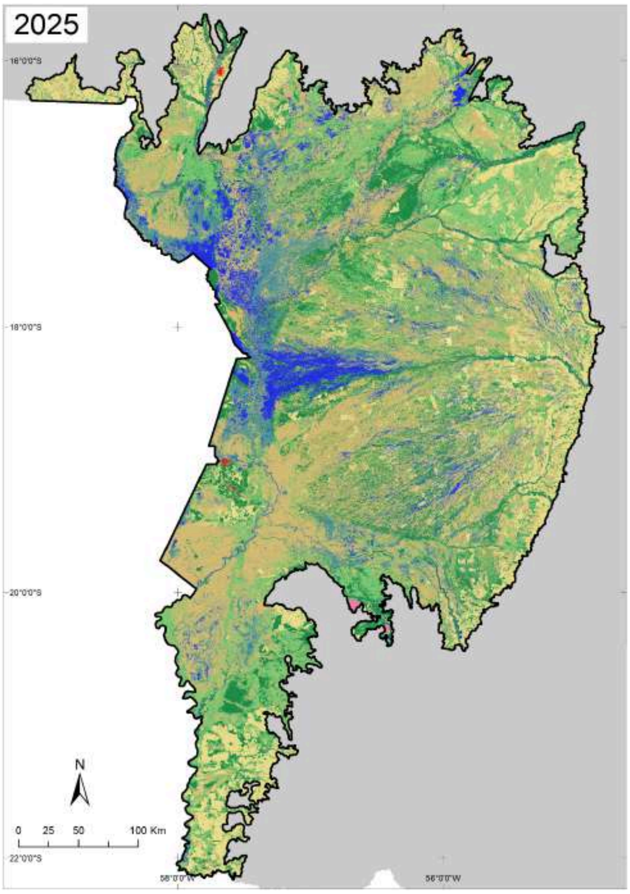


Figure 7: Final land use and land cover map of the Pantanal for the year 2025.

8. VALIDATION

In Collection 11, the overall accuracy reached 85.9% and 75.6% in levels 1 and 2, respectively. Wetland and water classes were added to the map based on the maximum area in each year, but this criterion is different from the validation methodology. To address this difference, the validation points of grassland or wetland that overlapped on the classified map as water were considered correct.

This nominal variation in accuracy compared to previous versions primarily reflects the increased thematic complexity and the adoption of more rigorous monthly monitoring of hydrological dynamics. The implementation of monthly mapping to capture annual flood peaks provides a far more realistic representation of the biome's seasonal pulse, even though it creates inherent discrepancies with standard validation protocols that may not account for the maximum annual extent. Furthermore, the application of stricter temporal stabilization filters and the integration of validated deforestation data since 2019 prioritize the long-term consistency of the 41-year series over annual spectral snapshots.Considering the Global Accuracy for all collections in legend level 1 (Figure 8), the maps in Collection 11 embody over four decades of expertise in mapping land use and land cover within the Pantanal. This is evident from the collection's accuracy and enhanced mapping methodologies. Time-series analysis reveals that Collection 11 is more stable and consistent.

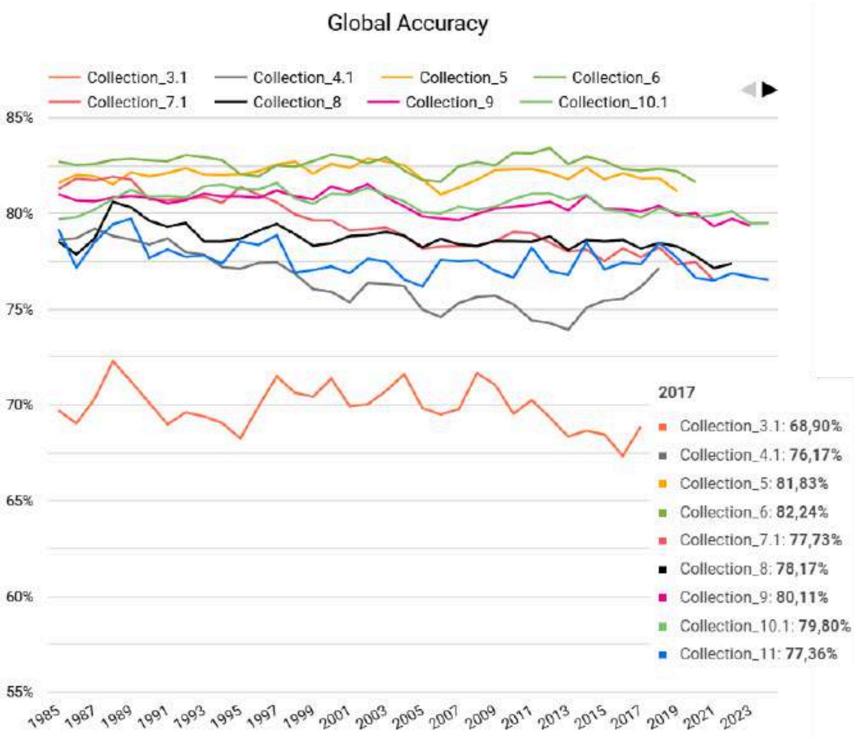


Figure 8. Global accuracy for all collections in Pantanal over the classification period and 2017.

9. REFERENCES

ASSINE, Mario Luis et al. The Quaternary alluvial systems tract of the Pantanal Basin, Brazil. Brazilian Journal of Geology, v. 45, n. 3, p. 475-489, 2015.

FG ASSIS, Luiz Fernando et al. TerraBrasilis: A spatial data analytics infrastructure for large-scale thematic mapping. ISPRS International Journal of Geo-Information, v. 8, n. 11, p. 513, 2019.

BREIMAN, Leo. Random forests. Machine learning, v. 45, n. 1, p. 5-32, 2001.

BURKE, Marshall; LOBELL, David B. Satellite-based assessment of yield variation and its determinants in smallholder African systems. Proceedings of the National Academy of Sciences, v. 114, n. 9, p. 2189-2194, 2017.

CI – CONSERVAÇÃO INTERNACIONAL; ECOA – ECOLOGIA E AÇÃO; FUNDACIÓN AVINA; INSTITUTO SOS PANTANAL; WWF-BRASIL. Monitoramento das alterações da cobertura vegetal e uso do solo na Bacia do Alto Paraguai – Porção Brasileira – Período de Análise: 2002 a 2008. Brasília, 2009. 58 p. ISBN 978-85-86440-25-0.

DA SILVA, João dos Santos Vila; DE MOURA ABDON, Myrian. Delimitação do Pantanal brasileiro e suas sub-regiões. Pesquisa agropecuária brasileira, p. 1703-1711, 1998.

ESCADAFAL, Richard. Remote sensing of arid soil surface color with Landsat thematic mapper. Advances in Space Research, v. 9, n. 1, p. 159-163, 1989.

ESCADAFAL, Richard; HUETE, A. Etude des propriétés spectrales des sols arides appliquée à l'amélioration des indices de végétation obtenus par télédétection. Comptes rendus de l'Académie des sciences. Série 2, Mécanique, Physique, Chimie, Sciences de l'univers, Sciences de la Terre, v. 312, n. 11, p. 1385-1391, 1991.

FEYISA, Gudina L. et al. Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery. Remote sensing of environment, v. 140, p. 23-35, 2014.

FUNK, Chris et al. The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. Scientific data, v. 2, n. 1, p. 150066, 2015.

GAO, Bo-Cai. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. Remote sensing of environment, v. 58, n. 3, p. 257-266, 1996.

GU, Yingxin et al. A five‐year analysis of MODIS NDVI and NDWI for grassland drought assessment over the central Great Plains of the United States. Geophysical research letters, v. 34, n. 6, 2007.

HALL, R. J. et al. Modeling forest stand structure attributes using Landsat ETM+ data: Application to mapping of aboveground biomass and stand volume. Forest ecology and management, v. 225, n. 1-3, p. 378-390, 2006.

HAMILTON, Stephen K.; DE SOUZA, Osni Corrêa; COUTINHO, Marcos Eduardo. Dynamics of floodplain inundation in the alluvial fan of the Taquari River (Pantanal, Brazil). Internationale Vereinigung für theoretische und angewandte Limnologie: Verhandlungen, v. 26, n. 3, p. 916-922, 1998.

HUETE, Alfredo R. A soil-adjusted vegetation index (SAVI). Remote sensing of environment, v. 25, n. 3, p. 295-309, 1988.

JORDAN, Carl F. Derivation of leaf‐area index from quality of light on the forest floor. Ecology, v. 50, n. 4, p. 663-666, 1969.

PADOVANI, Carlos Roberto. Dinâmica espaço-temporal das inundações do Pantanal. 2010. Tese de Doutorado. Universidade de São Paulo.

PARENTE, Leandro; FERREIRA, Laerte. Assessing the spatial and occupation dynamics of the Brazilian pasturelands based on the automated classification of MODIS images from 2000 to 2016. Remote Sensing, v. 10, n. 4, p. 606, 2018.

PINTY, B.; VERSTRAETE, M. M. GEMI: a non-linear index to monitor global vegetation from satellites. Vegetatio, v. 101, n. 1, p. 15-20, 1992.

PONTIUS JR, Robert Gilmore. Metrics that make a difference. Switzerland. ISBN-10: 3030707644, 2022.

RONDEAUX, Geneviève; STEVEN, Michael; BARET, Frédéric. Optimization of soil-adjusted vegetation indices. Remote sensing of environment, v. 55, n. 2, p. 95-107, 1996.

ROUSE, W. Monitoring vegetation system in the great plain with ERTS. In: 3rd ERTS symposium, NASA Washington DC, 1974. 1974. p. 309-317.

TADONO, Takeo et al. Precise global DEM generation by ALOS PRISM. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, v. 2, p. 71-76, 2014.

WAQAR, Mirza Muhammad et al. Development of new indices for extraction of built-up area & bare soil from landsat data. Open Access Sci. Rep, v. 1, n. 1, p. 4, 2012.

XU, Hanqiu. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. International journal of remote sensing, v. 27, n. 14, p. 3025-3033, 2006.