



Coastal Zone – Appendix

Collection 8

Version 1

General coordinator

Pedro Walfir Martins e Souza Filho

Team

Cesar Guerreiro Diniz

Luiz Cortinhas Ferreira Neto

Maria Luíze Silva Pinheiro

1 Overview

The Brazilian coastal zone presents diverse coastal environments that evolved during the Quaternary in response to changes in climate and sea-level changes, showing an interaction between different sediment supplies and a geologic heritage that dates back to the breakup of South America and Africa (Dominguez, 2009). Among this diversity of coastal features, four classes are mapped in the MapBiomias Collection 7; Mangroves, Beaches and Dunes, Aquaculture, and Hypersaline tidal flats.

Table 1, show the evolution of coastal features mapped in each collection as well as the changes in the methodological aspects of each collection.

Table 1 - Overview of the Cerrado MapBiomias Collections since their first version. In the method column, 'EDT' means 'Empirical Decision Tree,' and RF refers to 'Random Forest'.

Collection	Range	Method	Classes	Improvements
1.0	2008-2015	EDT	No Coastal-Specific Mappings	- First collection
2.0	2000-2016	EDT	Mangroves, Beaches & Dunes	- First two coastal classes
2.3	1985-2016	EDT	Same as Collection 2.0	--
3.0	1985-2017	RF	Mangroves, Beaches & Dunes	- Random Forest - Temporal stability is used to generate a large training dataset - Expanded to the entire Landsat Temporal Series - Better Quality Median Composites
3.1	1985-2017	RF	Same as Collection 3.0	--
4.0	1985-2018	RF and U-net	Mangroves, Beaches, and Dunes, Aquaculture	- Aquaculture/Salt-culture is added as a coastal feature - Improvements in temporal consistency through additional post-processing/ filters
4.1	1985-2018	RF and U-net	Same as Collection 4.0	--
5.0	1985-2019	RF and U-net	Mangroves, Beaches & Dunes, Aquaculture, Hypersaline Tidal Flats	- Hypersaline Tidal Flats are added as a coastal feature (also known as "Apicum")
6.0	1985-2020	RF and U-net	Same as Collection 5	- Snad Spots is now a feature that integrates Beach and Dune, coastal class
7.0	1985-2021	RF and U-net	Same as Collection 5	- A new version of the U-net classifier.
7.1	1985-2021	RF and U-net	Same as Collection 5	--
8.0	1985-2022	RF and U-net	Same as Collection 5	- Enhancements of the Deep-Learning Algorithms - Enhancements in temporal consistency through additional post-processing/ filters

In comparison to Collection 7, Collection 8 of the coastal zone classes presents small methodological changes, affecting only the Mangrove-related post-processing filters and a better-designed and trained version of the U-net derived “Hypersaline Tidal Flats”. In Collection 7, the “Apicum”/Salt Flat theme was also classified using the Deep Learning algorithm, the U-NET classifier. Now, supported by the improvement in the context domain, the “Hypersaline Tidal Flats” area oscillation, former “Apicum”, is much smaller. The whole classification process is described below in Figure 1.

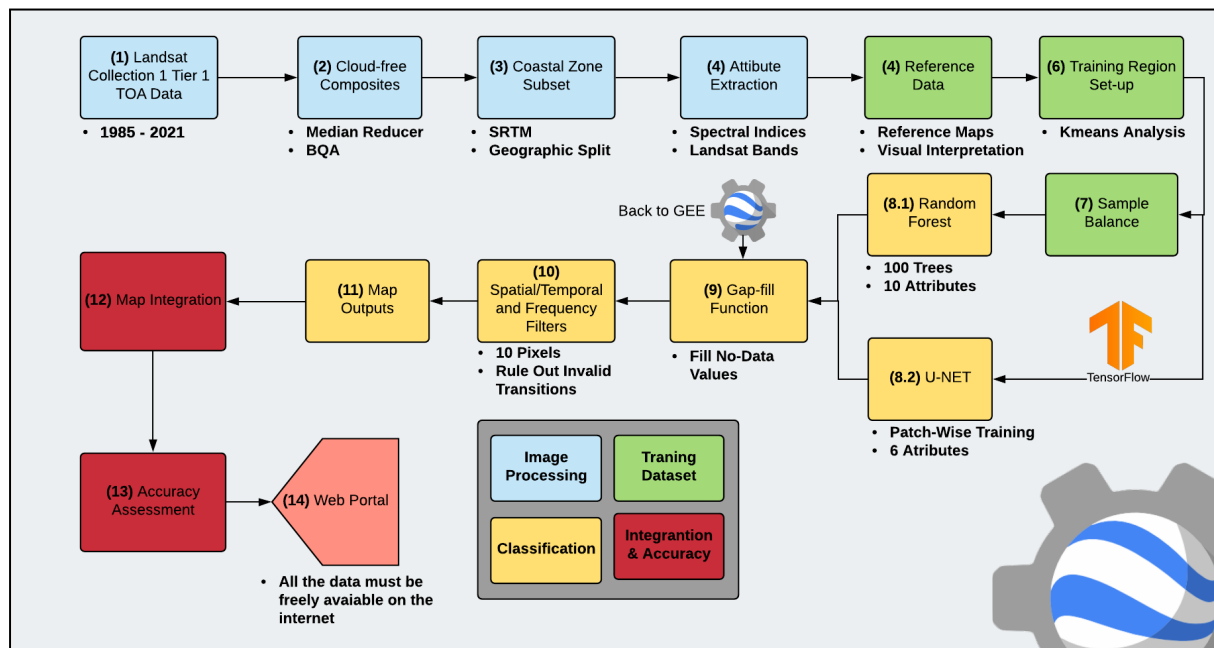


Figure 1 - All data processing occurs within the Google Earth Engine - GEE platform, except for the aquaculture/saline pattern and salt flat classification, which is dependent on the TensorFlow library. In green are steps related to sampling design. In yellow are steps related to classification. In red is the mapping accuracy evaluation stage.

2 Landsat image mosaics

The classification of the cross-cutting theme “**Coastal Zone**” used Landsat mosaics that differed from the mosaics used to classify the natural vegetation of the Brazilian biomes. The coastal mosaics were defined to preserve the maximum of the coastal zone land area while capturing the smallest possible cloud cover. These Landsat mosaics are the third generation of the methodology developed specifically for these cross-cutting themes.

2.1 Definition of the temporal period

Coastal areas are severely affected by atmospheric nebulosity, a condition that is intensified by its proximity to the oceans and its tropical location. On the other hand, the attempt to identify a time interval that covers only the driest season of the year, as an alternative to reduce cloud persistence ends up severely reducing the number of images available to cover the entire coastal region. Thus, the annual cloud-free composites are generated, ranging from the 1st of January to the 31st of December.

2.2 Mosaic Subsets

Since the Brazilian coastal zone (BCZ) is an extensive region, approximately 8,500 kilometers from Oiapoque to Chui (without counting reentrances), and affected by a variety of atmospheric systems, of lesser or greater influence of nebulosity, the BCZ is here divided into 7 different sectors, Figure 2.

Sector 1 - Amapa (AP), coastal region of Amapa. Sector 2 - Marajo Island (MAR), coastal region of Marajo Island. Sector 3 - Para / Maranhao (PAMA), a coastal sector of the states of Para and Maranhao. Piaui / Bahia (PIBA), a coastal sector of the states of Piaui to Bahia. Sector 5 - Espirito Santo / Sao Paulo (ESSP), a region that includes the states of Espirito and São Paulo. Sector 6 – Parana/Laguna (PRLA), a coastal region that goes from the state of Parana to the municipality of Laguna in Santa Catarina, and finally, Sector 7 (LARS), a region that ranges from Laguna to the state of Rio Grande do Sul.

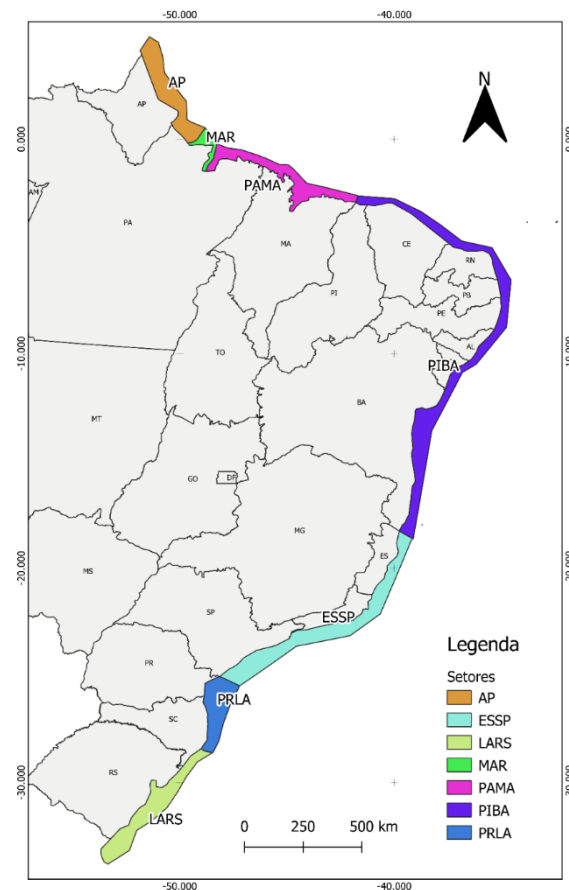


Figure 2 - The seven sectors delimited in different colors, along the Brazilian Coastal Zone (BCZ).

2.3 Image selection

The cloud/shadow removal script takes advantage of the Landsat Collection 1 Level-1 QA band and the GEE median reducer. In Collection 1 Tier 1 data, each pixel in the QA band

contains unsigned integer values representing certain surface, atmospheric, and sensor conditions that may affect the overall usefulness of a given pixel. When effectively used, QA values can improve the data integrity by indicating which pixels might be affected by instrument artifacts or subject to cloud contamination (USGS, 2017). In conjunction with that, GEE can be instructed to pick the median pixel values in a stack of images. By doing so, GEE rejects values that are too high (e.g., clouds) or too low (e.g., shadows) and picks the median pixel value in each band over time, Figure 3.



Figure 3 - Left, Collection 2 “cloud-free composite.” Right, Collection 8 “cloud-free composite.”

2.4 Final quality

The mosaic quality is related to Landsat’s cloud-free availability during the image selection period. However, from 1985 to March 1998, only the Landsat 5 satellite remained operational. In this period, for the BCZ, the average number of images per year was ~500. In the last decade, between 2008 and 2018, this figure tripled to ~1500 images per year, as shown in Figure 4.

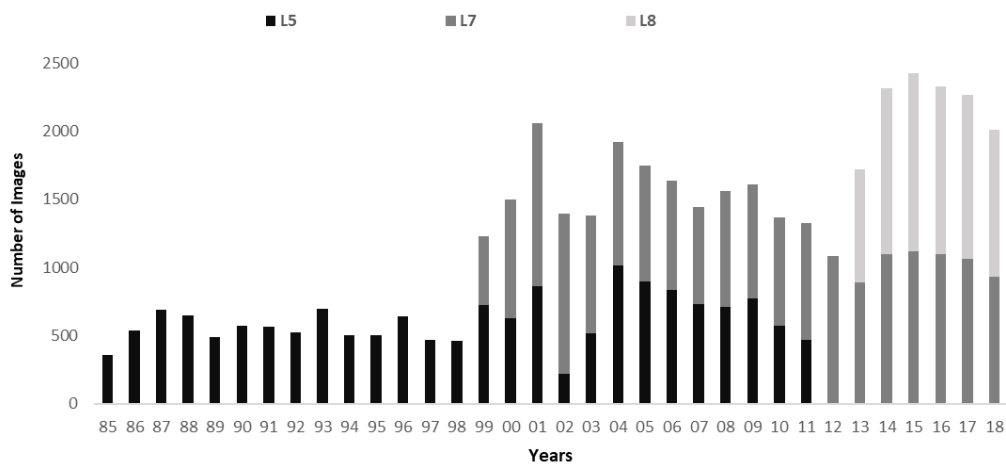


Figure 4 – Landsat image availability from 1985 to 2018. The bars show the distribution of Landsat images along the time series. L5 stands for Landsat 5, L7 refers to Landsat 7, and L8 stands for Landsat 8.

3 Classification

The automatic classification of the Landsat mosaics was mainly performed on the Google Earth Engine platform, based on the Random Forest classifier (Breiman, 2001). The Salt-flat and Aquaculture classes were Deep-Learning dependent and were classified outside the GEE.

3.1 Classification scheme

Each interested class was classified separately. In account of this, four independent classification processes were performed: 1) Mangrove; 2) Beaches and dunes; 3) Apicum; and 4) Aquaculture. The classification process was carried out considering only two possible classes for each pixel, the interest class (Mangrove, Beaches, Dunes, and Sand Spot, Salt flat, and Aquaculture) or the non-interest class (all different from the interest class).

We have selected training points based on reference maps and the MapBiomias Collection 6 for the supervised classification of the Landsat mosaics. The details of the parameters used in the Random Forest classifier, the reference maps used for each interest class, and the feature space produced for each classification are presented in the following sections.

3.2 Reference Data

For each class, a dataset of reference data was used to guide the generation of training samples. Table 2 shows the references used for each one of the coastal zone classes.

Table 2 - Reference datasets to guide training samples of coastal zone classes in Collection 7.

Class	References
Mangrove	MapBiomias Collection 6, Giri et al., 2011, ICMBio Mangrove Atlas (ICMBio, 2018), Global Mangrove Watch (Bunting <i>et al.</i> , 2018; Thomas <i>et al.</i> , 2018), plus visual inspection.
Aqua/Salt-Culture	MapBiomias Collection 6, Atlas Dos Remanescentes Florestais da Mata Atlântica (SOS Mata Atlântica, 2020), Barbier and Cox, 2003; Guimarães et al., 2010; Prates, Gonçalves and Rosa, 2010, Queiroz et al., 2013; Tenório et al., 2015; Thomas et al., 2017, plus visual inspection
Apicum/Salt flat	MapBiomias Collection 6, Atlas Dos Remanescentes Florestais da Mata Atlântica (SOS Mata Atlântica, 2020), Prates, Gonçalves and Rosa, 2010, plus visual inspection.

Beaches, Dunes and Sand Spots	MapBiomass Collection 6, Atlas Dos Remanescentes Florestais da Mata Atlântica (SOS Mata Atlântica, 2020), Prates, Gonçalves and Rosa, 2010, plus visual inspection.
-------------------------------	---

3.3 Coastal Zone Feature Space

Tables 3 and 4 show all spectral indices and bands used for the BCZ classification.

Table 3 – Spectral Indices used for coastal zone classification.

Index	Expression	Reducer	Reference
EVI2	$2.5 * ((NIR - RED) / (NIR + 2.4 * RED + 1))$	Median and Standard Deviation	Liu and Huete, 1995
NDVI	$(NIR - RED) / (NIR + RED)$	Median and Standard Deviation	Tucker, 1979
MNDWI	$(GREEN - SWIR1) / (GREEN + SWIR1)$	Median and Standard Deviation	Xu, 2006
NDSI	$(SWIR1 - NIR) / (SWIR1 + NIR)$	Median and Standard Deviation	Rogers and Kearney, 2004
MMDI	Modular Mangrove Detection Index	Median and Standard Deviation	Diniz et al., 2019

Table 4 - Table of bands used to classify coastal zone classes.

Variable	Description	Reducer
GREEN	Landsat Green band median value	Median and Standard Deviation
RED	Landsat Red band median value	Median and Standard Deviation
NIR	Landsat NIR band median value	Median and Standard Deviation
SWIR1	Landsat SWIR1 band median value	Median and Standard Deviation
SWIR2	Landsat SWIR2 band median value	Median and Standard Deviation

3.4 Classification algorithm, training samples, and parameters

When lacking reference maps that match the classes and/or year to be classified, reference maps of the closest possible timeframe to the median composites were used. When no reference map was available, then the classification results of the previous year were used for subsequent training. This was done for each one of the years without an external reference training guide. Table 5 and 6, shows the Random Forest and U-net parameters used to classify each one of the years.

Table 5 - Random Forest parameters used to classify each one of the years. Mangroves and Beaches, Dunes and Sand spots.

Parameter	Value
Number of trees	100
Number of points	100000
Number of Variables	20 (Coastal Zone)
Classes	2 (binary classification)

Table 6 - U-Net parameters used to classify each one of the years. Aquaculture and Salt flat classes.

Parameter	Value
Classifier	U-Net
Tile-Size	256 x 256 px
Optimizer	SGD
Learning Rate	0.1
Momentum	0.9
Decay	1e-4
Samples	10000 (geometries)
Attributes	MNDWI, NDVI, NDSI
Classes	2 (binary classification)

3.4.1 Mangroves

As in any supervised method, the Random Forest classifier needs to rely on a training dataset. For mangrove cover recognition, the training data was obtained from MapBiomass Collection 6, Giri *et al.*, 2011, Atlas dos Manguezais do Brasil (ICMBio, 2018), Global Mangrove Watch (Bunting *et al.*, 2018; Thomas *et al.*, 2018) and visual inspection, Figure 5. The consolidated results of the mangrove distributions are available in Diniz *et al.*, 2019.

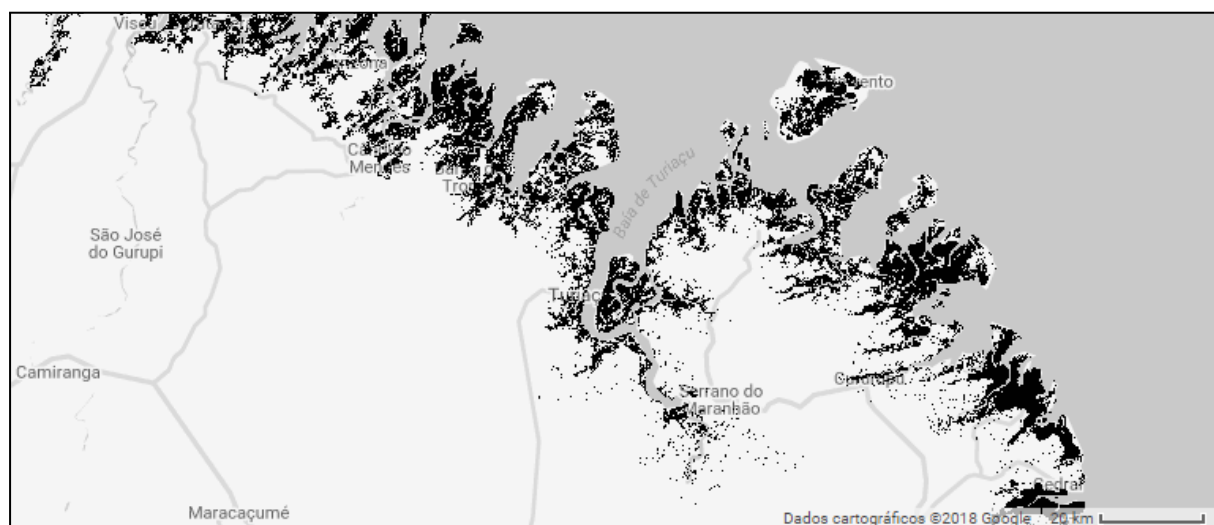


Figure 5 - Global Mangrove Cover data was used as a mangrove mapping reference, from 1999 to 2002.

3.4.2 Hypersaline Tidal Flat

Generally, the less frequently flooded area of a mangrove swamp, in the transition to topographically elevated lands, is usually devoid of arboreal vegetation. In Brazil, this area is called “*Apicum*”, or *Hypersaline Tidal Flat*. In the international scientific literature, this transition zone is usually called salt flat or hypersaline tidal flat. As shown in Table 1, three different reference maps were here used, the “Atlas dos Remanescentes Florestais da Mata Atlântica” (SOS Mata Atlântica, 2020) from 2019/2020, covering the Mata Atlantica coastal region and the “Carta de Sensibilidade Ambiental ao Oleo -Para-Maranhão-Barreirinhas” referent to 2017 and covering most of the Brazilian north coastal region and the data from the MapBiomias Collection 6, Figure 6.



Figure 6 – Apicum reference maps, the “Atlas Dos Remanescentes Florestais da Mata Atlântica” from 2019/2020, covering the Forest Atlantic coastal region and the “Carta de Sensibilidade Ambiental ao Oleo -Para-Maranhão-Barreirinhas 2017”, covering most of the Brazilian north coast region

3.4.3 Beaches and Dunes

Mapped without distinction between one another, here the “Beaches and Dunes” class refers to sandy strands, bright white in color, where there is no predominance of vegetation of any kind. As shown in Table 1, the training data for this land cover was obtained from MapBiomias Collection 6 and available reference, Figure 7.

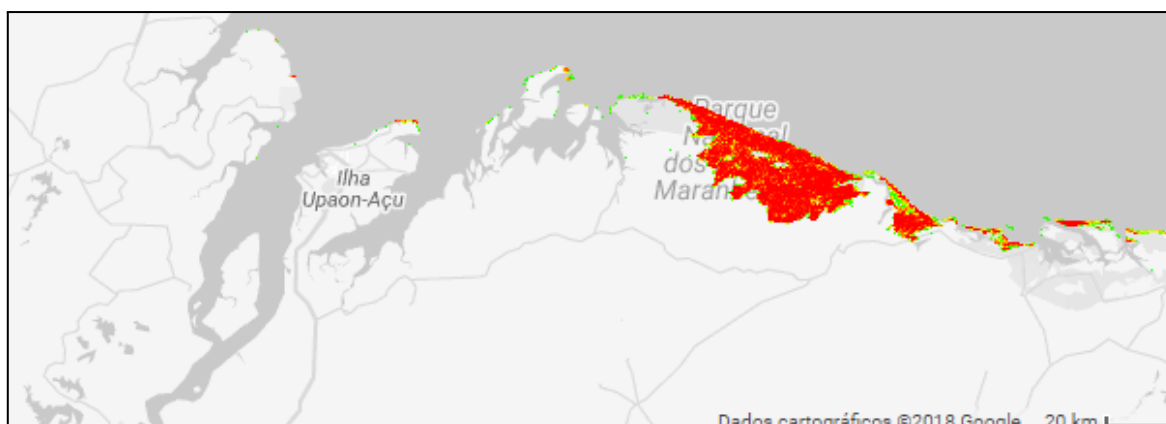


Figure 7 - The training data for this land cover was obtained from MapBiomas Collection 6 and from available reference, as shown in Table1”.

3.4.4 Aquaculture/Salt Culture

Compared to previous Collections, Collection 8 aquaculture mapping consolidated the use of the Deep-Learning model in replacement of the traditional Random Forest Algorithm. Its results are now published (Diniz *et al.*, 2021). In this scenario, traditional machine learning algorithms use spectral-temporal data to classify targets according to similarities of their spectral-temporal patterns (Breiman, 2001). Although temporal and spectral properties might not be enough to discriminate “super-similar” targets. Targets that behave similarly in both spectral and temporal domains. That is the case for most surface water targets, such as aquaculture ponds, rivers, lakes, and open waters, Figure 8.

Water is water, and unless it presents a high concentration of external compounds (minerals, suspended sediments, algae, etc.), not much can be done to differentiate between numerous surface water targets spectrally. On the other hand, the temporal domain may not present much valid discriminatory data either. In Brazil, aquaculture is a traditional and coastal-related economic activity. Thus, in 35 years of data, a diverse set of aquaculture frequencies may exist (Barbier and Cox, 2003; Guimarães *et al.*, 2010; Queiroz *et al.*, 2013; Tenório *et al.*, 2015; Thomas *et al.*, 2017). As a result, the temporal domain renounces to distinguish between well-consolidated aquaculture, main river channels, and open waters once all these features present high temporal persistence throughout the entire time series.

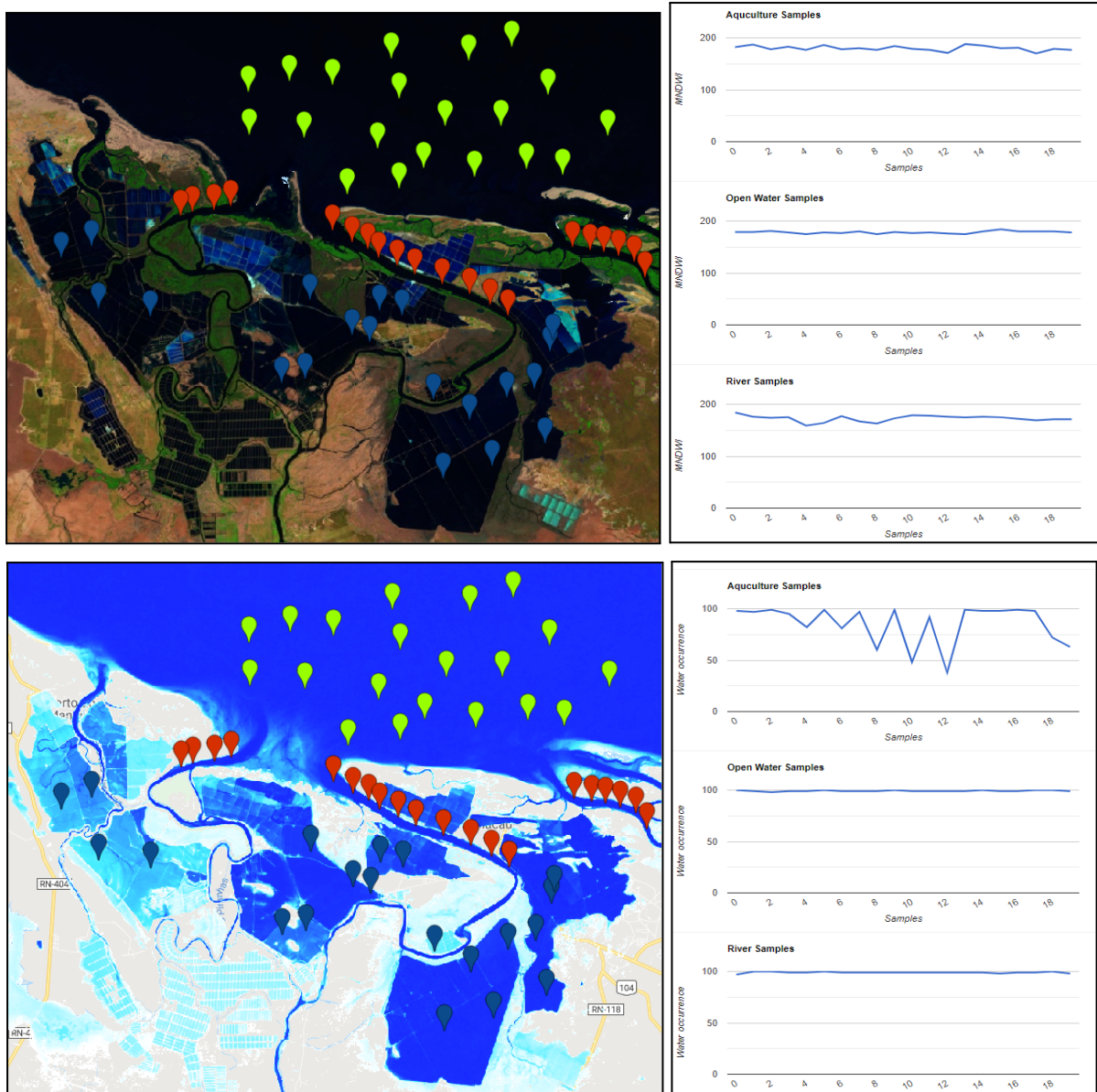


Figure 8 – Spectral and temporal patterns of the aquaculture, rivers, and open waters classes. In the top-left corner, the median cloud-free composite from Macau-RN, northeast of Brazil. The markers in dark-blue, green, and red represent aquaculture, open water, and river samples. In the top-right, NMDWI values for each one of the samples. In the bottom-left, JRC occurrence data. In the bottom-right is the occurrence frequency of each one of the samples.

In cases like this, the “context domain” may be key to distinguishing between rivers, aquaculture, and open waters pixels. In the context analysis scenario, the U-Net: Convolutional Networks (Abadi *et al.*, 2015) have the advantage of predicting the class label of each pixel by providing as input a local region (patches or chips) around that pixel. Such a characteristic of working with “patches” or “chips” is what gives the U-Net the ability to access the “context domain” of the image instead of using isolated pixels. The U-Net initial training was guided by Collection 6 and available reference data.

4 Post-classification

Due to the pixel-based nature of the classification method and the very long temporal series, a chain of post-classification filters was applied. The post-classification process includes the application of a gap-fill, a temporal, a spatial, and a frequency filter.

4.1 Gap-Fill filter

The chain starts by filling in possible no-data values. In a long time-series of severely cloud-affected regions, such as tropical coastal zones, it is expected that no-data values may populate some of the resultant median composite pixels. 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 “future” valid position is available, the no-data value is replaced by its previous valid class. Up to three prior years can be used to fill in persistent no-data positions. Therefore, gaps should only exist if a given pixel has been permanently classified as no-data throughout the entire temporal domain. A mask of years was built to keep track of pixel temporal origins, as shown in Figure 9.

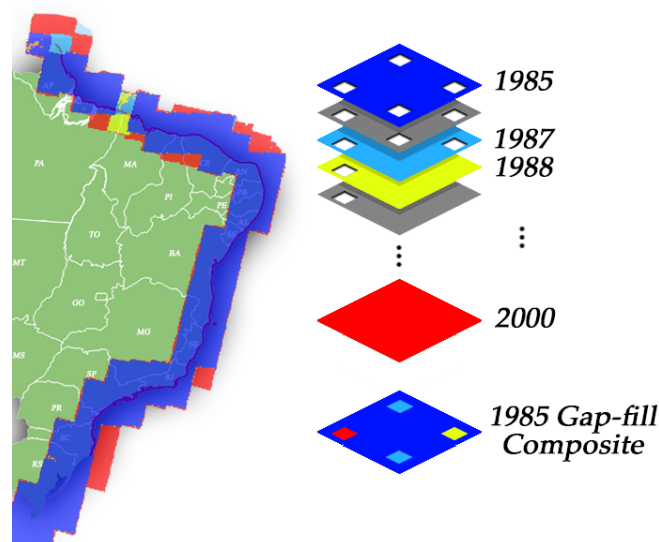


Figure 9 – Gap-filling mechanism. The following valid classification replaces existing no-data values. If no “future” valid position is available, then the no-data value is replaced by its previous valid classification based on up to a maximum of three (3) prior years. A mask of years was built to keep track of pixel temporal origins.

4.2 Temporal filter

After gap filling, a temporal filter was executed. The temporal filter uses sequential classifications in a 3-year unidirectional moving window to identify temporally non-permitted transitions. Based on a single generic rule (GR), the temporal filter inspects the central position of three consecutive years (“ternary”). If the extremities of the ternary are identical, but the center position is not, then the central pixel is reclassified to match its temporal neighbor class, as shown in Table 6.

Table 6 - The temporal filter inspects the central position for three consecutive years, and in cases of identical extremities, the center position is reclassified to match its neighbor. T1, T2, and T3 stand for positions one (1), two (2), and three (3), respectively. GR means “generic rule”, while Mg and N-Mg represent mangrove and non-mangrove pixels.

Rule	Input (Year)			Output		
	T1	T2	T3	T1	T2	T3
GR	Mg	N-Mg	Mg	Mg	Mg	Mg
GR	N-Mg	Mg	N-Mg	N-Mg	N-Mg	N-Mg

4.3 Spatial filter

Next, a spatial filter was applied. To avoid unwanted modifications to the edges of the pixel groups (blobs), a spatial filter was built based on the "connectedPixelCount" function. Native to the GEE platform, this function locates connected components (neighbors) that share the same pixel value. Thus, only pixels that do not share connections to a predefined number of identical neighbors are considered isolated, as shown in Figure 10. In this filter, at least ten connected pixels are needed to reach the minimum connection value. Consequently, the minimum mapping unit is directly affected by the spatial filter applied, and it was defined as 10 pixels (~1 ha).

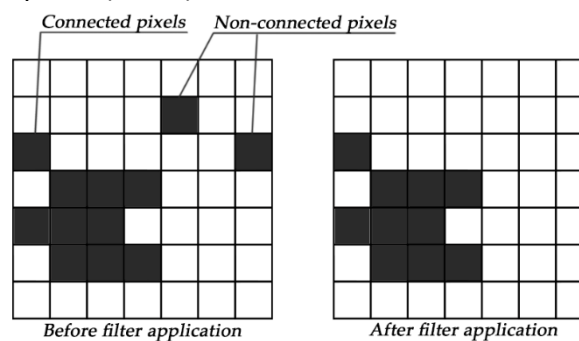


Figure 10 – The spatial filter removes pixels that do not share neighbors of identical value. The minimum connection value was 10 pixels.

4.4 Frequency filter

The last step of the filter chain is the frequency filter, as shown in Figure 11. This filter considers the occurrence frequency of a given class throughout the entire time series. Thus, all class occurrences with less than 10% temporal persistence (3 years or fewer out of 37) are filtered out and incorporated into the non-class binary. This mechanism contributes to reducing the temporal oscillation of the classification signal, decreasing the number of false positives, and preserving consolidated class pixels.

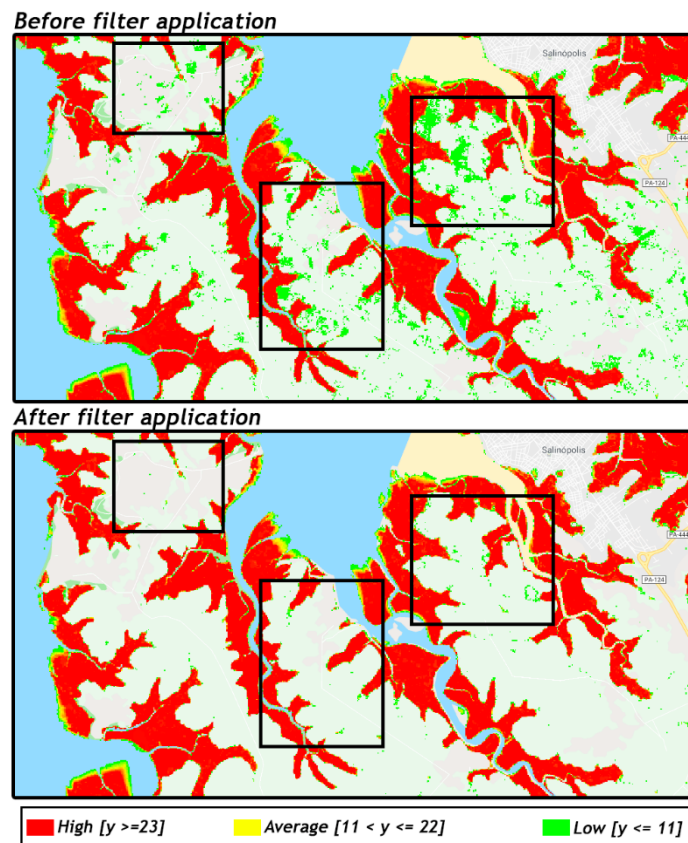


Figure 11 – Red, yellow and green represent mangrove pixels with high (23 or more years, $y \geq 23$), average (between 11 and 22 years, $11 < y \leq 22$), and low (ten years or less, $y < 11$) occurrence frequencies, respectively. The top image shows mangrove pixels before applying the frequency filter. The bottom image shows mangrove pixels after applying the frequency filter. The black boxes are centered on areas significantly affected by the filter. All mangrove occurrences with less than 10% temporal persistence (3 years in 33 possible years) were filtered out.

4.5 Integration with biomes themes

After applying the filter chain, the cross-cutting themes and the biomes data are integrated. This integration is guided by specific hierarchical prevalence rules (Table 6). As the output of this step, a final land cover/land use map of Brazil for each year.

Niche-specific classes or uses; likewise, coastal-related features such as Mangroves, Beaches, Dunes, and Aquaculture, as well as anthropic transition widely distributed throughout Brazil's territory, tend to occupy the top positions of the Prevalence rank, as seen below in Table 6.

Table 6 - Prevalence rules for combining the output of digital classification with the cross-cutting themes in Collection 8.

Class	Pixel Value	Prevalence
Mining	30	1
Beach, Dune, and Sand Spot	23	2
Mangrove	5	3
Aquaculture/Salt-Culture	31	4
Hypersaline Tidal Flat	32	5
Urban Infrastructure	24	6
Sugar Cane	20	7
Soybean	39	8
Rice	40	9
Other Temporary Crop	41	10
Perennial Crop	36	11
Coffee	46	12
Citrus	47	13
Other Perennial Crop	48	14
Temporary Crop	19	15
Forest Plantation	9	16
Rocky Outcrop	29	17
Other Non-Vegetated Areas	25	18
River, Lake, and Ocean	33	19
Forest Formation	3	20
Savanna Formation	4	21
Wetland	11	22
Grassland Formation	12	23
Pasture	15	24
Mosaic of Uses	21	25

References

- ABADI, M. *et al.* TensorFlow: Large-scale machine learning on heterogeneous systems. **Methods in Enzymology**, 2015.
- BARBIER, E. B.; COX, M. Does Economic Development Lead to Mangrove Loss? A Cross-Country Analysis. **Contemporary Economic Policy**, v. 21, n. 4, p. 418–432, 1 out. 2003.
- BREIMAN, L. Random Forests. **Machine Learning**, v. 45, n. 1, p. 5–32, 2001.
- BUNTING, P. *et al.* **The Global Mangrove Watch—A New 2010 Global Baseline of Mangrove Extent Remote Sensing**, 2018.
- DINIZ, C. *et al.* Brazilian Mangrove Status: Three Decades of Satellite Data Analysis. . 2019.
- _____. **A Large-Scale Deep-Learning Approach for Multi-Temporal Aqua and Salt-Culture Mapping Remote Sensing**, 2021.
- DOMINGUEZ, J. M. L. The Coastal Zone of Brazil. *In: Geology and Geomorphology of Holocene Coastal Barriers of Brazil*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009. p. 17–51.
- GIRI, C. *et al.* Status and distribution of mangrove forests of the world using earth observation satellite data. **Global Ecology and Biogeography**, v. 20, n. 1, p. 154–159, 2011.
- GUIMARÃES, A. S. *et al.* Impact of aquaculture on mangrove areas in the northern Pernambuco Coast (Brazil) using remote sensing and geographic information system. **Aquaculture Research**, v. 41, n. 6, p. 828–838, 13 maio 2010.
- ICMBIO. **Atlas dos manguezais do Brasil**. 1. ed. Brasília, Brazil: ICMBio, 2018.
- LIU, H. Q.; HUETE, A. Feedback based modification of the NDVI to minimize canopy background and atmospheric noise. **IEEE Transactions on Geoscience and Remote Sensing**, 1995.
- PRATES, A. P. L.; GONÇALVES, M. A.; ROSA, M. R. Panorama da conservação dos ecossistemas costeiros e marinhos no Brasil. **Brasília: MMA/SBF/GBA**, 2010.
- QUEIROZ, L. *et al.* **Shrimp aquaculture in the federal state of Ceará, 1970–2012: Trends after mangrove forest privatization in Brazil**. [s.l.: s.n.]. v. 73
- ROGERS, A. S.; KEARNEY, M. S. Reducing signature variability in unmixed coastal marsh Thematic Mapper scenes using spectral indices. **International Journal of Remote Sensing**, 2004.
- SOS MATA ATLÂNTICA. Atlas dos remanescentes florestais da Mata Atlântica, período 2019-2020. **São Paulo, Brasil. Fundação SOS Mata Atlântica. Instituto Nacional das Pesquisas Espaciais**, 2020.
- SOSMA; INPE. ATLAS DOS REMANESCENTES FLORESTAIS DA MATA ATLÂNTICA PERÍODO 2016-2017. **Fundação SOS Mata Atlântica e Instituto de Pesquisas Espaciais**, 2018.
- TENÓRIO, G. S. *et al.* Mangrove shrimp farm mapping and productivity on the Brazilian Amazon coast: Environmental and economic reasons for coastal conservation. **Ocean & Coastal Management**, v. 104, p. 65–77, 2015.
- THOMAS, N. *et al.* Distribution and drivers of global mangrove forest change, 1996–2010. **PLOS ONE**, v. 12, n. 6, p. e0179302, 8 jun. 2017.
- _____. **Mapping Mangrove Extent and Change: A Globally Applicable Approach Remote Sensing**, 2018.
- TUCKER, C. J. Red and photographic infrared linear combinations for monitoring vegetation. **Remote Sensing of Environment**, v. 8, n. 2, p. 127–150, 1979.
- USGS. **LANDSAT COLLECTION 1 LEVEL 1 PRODUCT DEFINITION**. [s.l.] Earth Resources Observation and Science (EROS) Center, 2017.
- XU, H. 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, 20 jul. 2006.