# TerraClass x MapBiomas: Comparative assessment of legend and mapping agreement analysis

**Conference Paper** · December 2017



**7 authors**, including:

Alana Kasahara Neves
National Institute for Space Research, Brazil


Thales Sehn Körting
National Institute for Space Research, Brazil


Gilberto Queiroz
National Institute for Space Research, Brazil


Lubia Vinhas
National Institute for Space Research, Brazil

# TerraClass x MapBiomas: Comparative assessment of legend and mapping agreement analysis

**Alana K. Neves , Thales S. Körting , Leila M. G. Fonseca , Gilberto R. de Queiroz , Lúbia Vinhas , Karine R. Ferreira , Maria Isabel S. Escada .**

National Institute for Space Research – INPE
Caixa Postal 515 – 12227-010 – São José dos Campos – SP, Brasil.

{alana.neves,thales.korting,leila.fonseca,gilberto.queiroz, lubia.vinhas,karine.ferreira}@inpe.br, isabel@dpi.inpe.br

***Abstract:*** *In this work, we evaluated the agreement between land cover maps generated by TerraClass and MapBiomas projects for Pará state and, more specifically: (1) describe the legends based on an international classification system, (2) analyze the differences among classes and (3) test* PostGIS Raster *from PostgreSQL database to work with classification products. The classifications were compared pixel by pixel and the evaluation was performed based on confusion matrices. The agreement between them was 84.40%. The different methodologies adopted by the two projects generate significant disagreements in class identification, so using both maps together as complementary is not recommended for land use and cover change analyzes.*

## 1. Introduction

The deforestation dynamics in the Legal Amazon has been monitored by remote sensing images since 1988 through PRODES project (Monitoring Program of Brazilian Amazon Forest by Satellite). Until 2015, an area of 76,990,300 hectares of Legal Amazon was deforested, which means 19.20% of the total forest initially available [INPE 2016]. To identify and quantify the drivers responsible for the deforestation, TerraClass project was created in 2008 to map land use and land cover in the Legal Amazon deforested areas [Almeida *et al.* 2016].

In 2015, MapBiomas (Brazilian Annual Land Use and Land Cover Mapping Project) was created by Greenhouse Gas Emissions Estimation System (SEEG) from the Climate Observatory's to map all Brazilian biomes annually (http://mapbiomas.org/). Its mapping methodology is fully automated and integrated with Google Earth Engine.

Maps from projects such as TerraClass and MapBiomas have been widely employed in land use and land cover modelling and climate change research. They can be used as support in the development of governmental projects and other initiatives [Mayax *et al.* 2006]; hence the need for assessments and products comparisons.

In this context, this work aims to evaluate the agreement between the classifications by TerraClass and MapBiomas, specifically (1) to describe the legends based on an international classification system, (2) to analyze the differences among classes and (3) to test PostGIS Raster from PostgreSQL database to work with classification products.

## 2. Methodology

The study area chosen for this work was Pará state, Brazil. With an area of approximately 1,248 million km , the entire state belongs to the Amazon biome. Over decades, this state as well as the Brazilian state of Mato Grosso have led the Amazon ranking of deforestation rate. TerraClass and MapBiomas classifications for Pará state were used, both for the year 2014, at 30m spatial resolution. Both classifications, which are available in raster format, were referenced in WGS84 and inserted in PostgreSQL database by "`raster2pgsql`" available in PostGIS extension (Table 1a).

**Table 1. SQL and R scripts used in the comparative assessment of mapping.**

***a)*** *Data insertion by* `raster2pgsql`*:*

```
raster2pgsql.exe -c -C -s 4326 -I -t 512x512 -b 1 -N 0 "raster/path.tif" public.patc | psql -U postgres -d TCxMap -h localhost -p 5432
```

***b)*** *Query which values are in the map:*

```
SELECT (pvc).* FROM (SELECT ST_ValueCount(patc.rast,1) AS pvc FROM patc) AS f ORDER BY (pvc).VALUE;
```

***c)*** *Map values reclassification:*

```
ALTER TABLE patc ADD COLUMN reclass raster;
UPDATE patc SET reclass=ST_Reclass(rast,1,'[3]:40,[6]:40,[8]:43,[10]:46,[21]:21,[26]:41,[27]:49,[28]:47', '32BF',0);
```

***d)*** *Connect to database in R and sum reclassified maps:*

```
library(RPostgreSQL)
library(rpostgis)
drv <- dbDriver("PostgreSQL")
con <- dbConnect(drv,user = "postgres",password="",dbname = "TCxMap",host =
"localhost")
pa_tc_r<-pgGetRast(con,name=c("patc "),rast="reclass")
pa_mapbio_r<-pgGetRast(con,name=c("pamapbio "),rast="reclass")
sum_r<-pa_tc_r + pa_mapbio_r
```

***e)*** *Query and count map values:*

```
SELECT (dat).value, sum(dat.count) FROM (SELECT (pvc).* FROM (SELECT ST_ValueCount(sum_r.rast,1) AS pvc FROM sum_r) AS f ORDER BY (pvc).VALUE) as dat group by dat.value;
```

Each project has its own legend, so it was necessary to reclassify them (Table 2) to identify equivalent classes and also to group other ones. Some minority classes (*Agriculture or Pasture* from MapBiomas and *Mosaic of Uses*, *Mining* and *Deforestation 2014* from TerraClass) did not present equivalence between projects so their percentages were not evaluated in this work. It was necessary to use a SQL statement to find out which original values were presented in the classifications (Table 1b). After that, the function "ST_Reclass" (Table 1c) was used to reclassify the original values to the new values presented in Table 2, so that when the two classifications were added they would not present repeated values (i.e., Forest corresponds to the value 900 in the TerraClass map).

All adopted classes (Table 2) were described by the Land Cover Classification System – LCCS [Di Gregorio *et al.* 2016]. The use of LCCS aims to standardize class descriptions so that data produced in different ways can be used and compared, regardless of scale, level of detail and geographical location. This system uses a set of rules based on the physiognomy and stratification of biotic and abiotic elements [Di Gregorio *et al.* 2016].

**Table 2. Reclassification of both TerraClass and MapBiomas legends.**

| Adopted classes | TerraClass | | MapBiomas | |
|---|---|---|---|---|
| Forest | 900 | Forest | 40 | Dense forest<br>Open forest<br>Mangrove<br>Flooded forest<br>Degraded forest |
| Water bodies | 961 | Hydrograph | 41 | Water bodies |
| Planted forest | 1024 | Reforestation | 42 | Silviculture |
| Secondary vegetation | 1089 | Secondary vegetation | 43 | Secondary forest |
| Urban areas | 1156 | Urban area | 44 | Urban infrastructure |

| | | | | |
|---|---|---|---|---|
| Pasture | 1225 | Herbaceous pasture<br>Shrubby pasture<br>Regeneration with pasture<br>Pasture with exposed soil | 45 | Pasture |
| Non-forest natural vegetation - NFNV | 1296 | Non-forest | 46 | Non-forest natural formations<br>Non-forest natural wetlands<br>Other non-forest formations |
| Agriculture | 1369 | Annual crops | 47 | Annual crops<br>Mosaic of crops |
| Others | 1444 | Others | 48 | Beaches and dunes |
| Non-observed | 1521 | Non-observed areas | 49 | Non-observed |

To sum the two reclassified maps, PostgreSQL database was connected to R using the packages “RPostgreSQL” and “rpostgis” (Table 1d). R was also used for data visualization. By the function “ST_ValueCount” (Table 1e), the presented values in the resulting map were counted and then confusion matrices could be filled to analyze the agreements and disagreements between the classifications of the two land cover maps.

## 3. Results and Discussion

The description of reclassified classes in LCCS pattern is presented in Table 3. In a simplified way, it represents the classes in both legends, from TerraClass and MapBiomas. *Forest* (Table 3a), for example, has one of its stratums represented by water bodies so it can include *Flooded Forest* from MapBiomas.

*Non-Forest Natural Vegetation* (Table 3g) represents, in most cases, vegetation patches typical of another biome (such as Cerrado) remaining in Amazon. NFNV can represent rock surfaces too. In *Agriculture* pattern (Table 3h), there is only one stratum composed of graminae, forbs or bare soil. Each formation type in this pattern is conditioned by the presence of a temporal sequence depending on crop phenological cycles. The *Others* class (Table 3i) has only one stratum composed of loose and shifting sands. It can represent *Beaches and Dunes* from MapBiomas and *Others* from TerraClass, which stand for cover patterns such as river beaches and sandbars [Coutinho *et al.* 2013].

**Table 3. Classes patterns described in LCCS.**

| | |
|---|---|
| **a)** | Forest:<br>Horizontal pattern 1:<br>Stratum 1 (mandatory): trees –natural or semi-natural vegetation, leaf phenology = evergreen and leaf type = broadleaved;<br>Stratum 2 (optional): shrubs – natural or semi-natural vegetation;<br>Stratum 3 (optional): graminae –natural or semi-natural vegetation;<br>Stratum 4 (optional): water bodies. |
| **b)** | Water bodies:<br>Horizontal pattern 1:<br>Stratum 1 (mandatory): water bodies – position = above surface. |
| **c)** | Planted forest:<br>Horizontal pattern 1:<br>Stratum 1 (mandatory): trees – cultivated and managed vegetation, planted forest;<br>Stratum 2 (optional): bare soil;<br>Stratum 3 (optional): herbaceous growth forms. |
| **d)** | Secondary vegetation:<br>Horizontal pattern 1:<br>Stratum 1 (mandatory): woody growth forms – natural or semi-natural vegetation, height up to 3m;<br>Stratum 2 (optional): herbaceous growth forms – natural or semi-natural vegetation. |
| **e)** | Urban areas:<br>Horizontal pattern 1:<br>Stratum 1 (mandatory): buildings;<br>Stratum 2 (optional): woody growth forms;<br>Stratum 3 (optional): herbaceous growth forms.<br>Horizontal pattern 2:<br>Stratum 1 (mandatory): roads. |
| **f)** | Pasture: |

| | |
|---|---|
| | Horizontal pattern 1:<br>Stratum 1 (mandatory): graminae – cultivated and managed vegetation;<br>Stratum 2 (optional): shrubs – natural or semi-natural vegetation;<br>Stratum 3 (optional): trees – cover between 0 and 4%. |
| **g)** | Non-forest natural vegetation:<br>Horizontal pattern 1:<br>Stratum 1 (mandatory): trees– cover between 20 and 70%, height up to 5m and leaf phenology = deciduous;<br>Stratum 2 (optional): herbaceous growth forms. |
| **h)** | Agriculture:<br>Horizontal pattern 1:<br>Stratum 1 (mandatory): graminae, forbs or bare soil – sequential temporal relationship, cultivated and managed vegetation, orchard and other plantations. |
| **i)** | Others:<br>Horizontal pattern 1:<br>Stratum 1 (mandatory): loose and shifting sands. |

After describing the classes, the agreement among them for the year 2014 was analyzed. The overall classification agreement for Pará state was 84.40%. In the confusion matrices (Tables 4 and 5), the agreements and disagreements among classes are presented in more detail. The main diagonal of Table 4 represents the agreement of TerraClass if MapBiomas is considered as reference, while the main diagonal of Table 5 represents the agreement of MapBiomas if TerraClass is considered as reference.

*Forest* had a high agreement (98.23 and 85.72%, Tables 4 and 5, respectively) and a small percentage of MapBiomas *Forest* was classified as *Secondary Vegetation* (5.06%), *Pasture* (4.77%) and NFNV (3.76%) in TerraClass. *Planted Forest* had 0% of agreement. In MapBiomas, a few pixels represent this class and most of them (50%) are classified as *Forest* in TerraClass. Despite its large area, the exclusion of *Forest* slightly decreased the overall agreement from 84.40% to 84.38%. This occurred because this class is the source of confusion for other classes. For example, 80.77% of MapBiomas *Secondary Vegetation* was classified as *Forest* by TerraClass (Table 4).

**Table 4. TerraClass agreement, considering MapBiomas as reference.**

| | | TerraClass 2014 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Forest | Water bodies | Planted forest | Secondary vegetation | Urban areas | Pasture | NFNV | Agriculture | Others |
| **MapBiomas 2014** | Forest | **98.23** | 9.50 | 73.55 | 80.77 | 27.67 | 30.43 | 53.91 | 18.95 | 53.72 |
| | Water bodies | 0.44 | **89.03** | 0.16 | 0.48 | 2.55 | 0.26 | 6.11 | 0.41 | 16.80 |
| | Planted forest | 0.00 | 0.00 | **0.00** | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | Secondary vegetation | 0.50 | 0.21 | 4.44 | **5.70** | 0.77 | 1.81 | 0.28 | 4.60 | 1.09 |
| | Urban area | 0.00 | 0.02 | 0.00 | 0.01 | **31.45** | 0.04 | 0.08 | 0.20 | 0.03 |
| | Pasture | 0.61 | 0.68 | 21.27 | 12.19 | 31.39 | **66.00** | 8.45 | 67.41 | 20.11 |
| | NFNV | 0.21 | 0.43 | 0.40 | 0.82 | 4.91 | 1.26 | **28.36** | 0.19 | 7.60 |
| | Agriculture | 0.01 | 0.12 | 0.17 | 0.03 | 1.24 | 0.21 | 2.82 | **8.22** | 0.62 |
| | Others | 0.00 | 0.00 | 0.00 | 0.00 | 0.02 | 0.00 | 0.00 | 0.00 | **0.04** |

TerraClass mapping is executed in PRODES deforestation polygons and, in this project, deforested areas do not go back to being *Forest* even after many years of regeneration. So these areas become *Secondary Vegetation* by TerraClass. This restriction does not exist in MapBiomas and, therefore, there was a high disagreement in the classification of *Secondary Vegetation*. 80.77% of the TerraClass *Secondary Vegetation* was considered as *Forest* by MapBiomas (Table 4). In addition, 38.76 and 25.08% of MapBiomas *Secondary Vegetation* were classified as *Forest* and *Pasture* by TerraClass, respectively (Table 5).

It is known that the use of time series assists in the identification of agricultural patterns due to the seasonal profiles of these targets. In the MapBiomas methodology [IMAZON 2017], time series are not used for the classification of *Agriculture*, while TerraClass uses MODIS images time series for its identification [Almeida *et al.* 2016]. Thus, 67.41% of the TerraClass *Agriculture* was classified as *Pasture* by MapBiomas (Table 4) and 12.18 and 73.36% of MapBiomas *Agriculture* were classified as *Pasture* and NFNV by TerraClass, respectively (Table 5).

**Table 5. MapBiomas agreement, considering TerraClass as reference.**

| | MapBiomas 2014 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| TerraClass 2014 | Forest | Water bodies | Planted forest | Secondary vegetation | Urban areas | Pasture | NFNV | Agriculture | Others |
| Forest | **85.72** | 7.98 | 50.00 | 38.76 | 0.87 | 4.26 | 7.51 | 1.82 | 28.19 |
| Water bodies | 0.42 | **81.18** | 0.00 | 0.83 | 0.97 | 0.24 | 0.78 | 1.92 | 14.06 |
| Planted forest | 0.11 | 0.00 | **0.00** | 0.56 | 0.00 | 0.24 | 0.02 | 0.09 | 0.00 |
| Secondary vegetation | 5.06 | 0.62 | 25.00 | **31.63** | 0.99 | 6.08 | 2.10 | 0.76 | 0.85 |
| Urban area | 0.05 | 0.10 | 0.00 | 0.13 | **80.26** | 0.47 | 0.37 | 0.86 | 11.61 |
| Pasture | 4.77 | 0.83 | 25.00 | 25.08 | 8.72 | **82.31** | 8.05 | 12.18 | 0.32 |
| NFNV | 3.76 | 8.81 | 0.00 | 1.74 | 7.35 | 4.69 | **80.74** | 73.36 | 26.55 |
| Annual crop | 0.05 | 0.02 | 0.00 | 1.15 | 0.78 | 1.52 | 0.02 | **8.70** | 0.00 |
| Others | 0.07 | 0.45 | 0.00 | 0.13 | 0.05 | 0.21 | 0.40 | 0.30 | **18.42** |

In Figure 2, there are crops of both project classifications where some existing disagreements can be seen. TerraClass mapping generates consolidated polygons because most of its methodology is visual. MapBiomas, on the other hand, has a fully automatic and per pixel classification and does not consider the context each pixel is inserted. Thus, in polygons classified by TerraClass, MapBiomas identified, for example, pixels of other classes, such as *Agriculture or Pasture* in areas of NFNV or *Forest* in *Urban Areas*. Methodological differences like that generate disagreements as it can be seen in NFNV class (53.91% of TerraClass NFNV was classified as *Forest* by MapBiomas, Table 4).

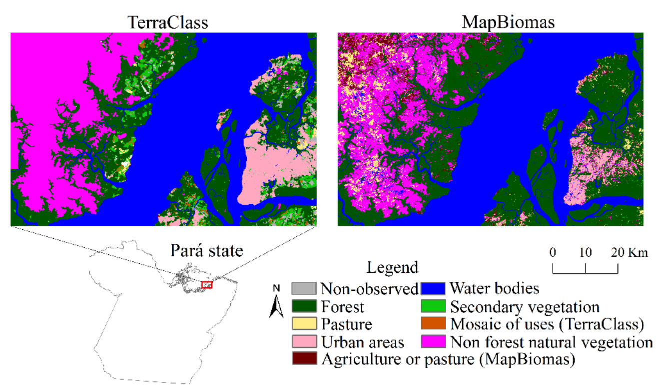


**Figure 1. Classification crops to see, in detail, some existing disagreements.**

The spatialization of agreement and disagreement areas in Pará state between both projects can be seen in Figure 2. Large consolidated areas of disagreement occurred in Marajó Island and close to the Amazon River channel, most of which represent TerraClass NFNV that was mapped into other classes by MapBiomas. In the northwest of the state, a great concentration of small polygons occurred and the disagreement between the two projects was very visible.

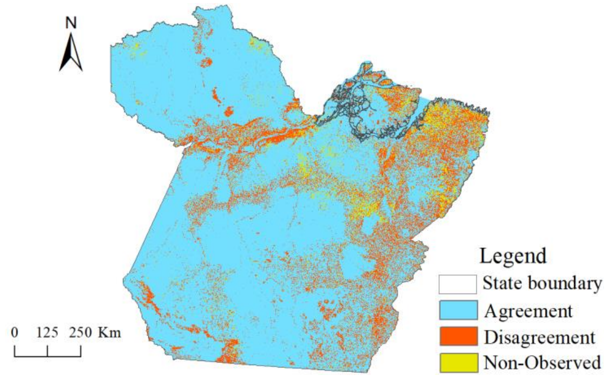


**Figure 2. Spatialization of the agreement analysis of classifications.**

## 4. Conclusions

Although TerraClass methodology has several visual stages and produces data every two years, there is a greater consistence in the identification of its classes. MapBiomas data still have some inconsistencies such as the existence of few pixels of other classes in already consolidated areas, but it has a fully automated data generation. The approach to verify the agreements between the classifications in the databases was efficient and not very time consuming.

Despite the high overall agreement (84.40%) between TerraClass and MapBiomas classifications, the methodological differences of these projects result in significant disagreements in the mapping results. For this reason, using the two maps as complementary ones without a proper adaptation of legends is not recommended for an analysis of land use and land cover change.

## References

ALMEIDA, C. A.; COUTINHO, A. C.; ESQUERDO, J. C. D. M.; ADAMI, M.; VENTURIERI, A.; DINIZ, C. G.; DESSAY, N.; DURIEUX, L.; GOMES, A. R. High spatial resolution land use and land cover mapping of the Brazilian Legal Amazon in 2008 using Landsat-5/TM and MODIS data. **Acta Amazonica**, v. 46. n. 3, p. 291-302, 2016.

COUTINHO, A. C.; ALMEIDA, C.; VENTURIERI, A.; ESQUERDO, J. C. D. M.; SILVA, M. Projeto TerraClass: Uso e cobertura da terra nas áreas desflorestadas na Amazônia Legal. Brasília, DF: Embrapa; Belém: INPE, 2013.

DI GREGORIO, A.; HENRY, M.; DONEGAN, E.; FINEGOLD, Y.; LATHAM, J.; JONCKHEERE, I.; CUMANI, R. Land Cover Classification System: Classification Concepts. Software version 3. 2016.

IMAZON - Institute of Man and Environment of the Amazon. Algorithm Theoretical Basis Document & Results – Amazon Biome. **MapBiomas**. 2017. Available at: <http://mapbiomas.org/pages/methodology>.

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS (INPE). **Projeto de Monitoramento do Desflorestamento na Amazônia Legal – PRODES**. São José dos Campos, 2016. Disponível em: <http://www.obt.inpe.br/prodes/index.php>.

MAYAUX, P.; EVA, H.; GALLEGO, J.; STRAHLER, A. H. Validation of the global land cover 2000 map. **IEEE Transactions on Geoscience and Remote Sensing**, v. 44, n. 7, p. 1728-1739, 2006.