Validation of TerraClass mapping for the Municipality of Paragominas state of Pará

Márcia Nazaré Rodrigues Barros¹, Alcione Ferreira Pinheiro², Vitor Mateus Carvalho Morais⁴, Luciana Barros Santos³, Andréa dos Santos Coelho¹, Luís Waldir Rodrigues Sadeck², Marcos Adami⁴, Alessandra Rodrigues Gomes⁴, Igor da Silva Narvaes⁴

¹Federal University of Para, Post-Graduate Program in Environmental Sciences, UFPA, PPGCA, Belém, Brazil. Email: [nnbarros, andrea_geoambiente]@hotmail.com  
²Foundation for Science, Technology and Space Applications, FUNCATE, São José dos Campos, Brazil. Email: [alcione.pinheiro, luís.sadeck]@funcate.org.br  
³Federal University of Para, Postgraduate Program in Amazon’s Natural Resource Management and Local Development, UFPA, PPGEDAM, Belém, Brazil. Email: lucyana_barros@hotmail.com  
⁴National Institute for Space Research – INPE – Amazon Regional Center – CRA, Brazil. Email: [marcos.adami, vitor.morais, alessandra.gomes, igor.narvaes]@inpe.br

Abstract - This work aims to evaluate the TerraClass mapping for the year 2014, in the municipality of Paragominas, State of Pará. The validation was made by comparing the mapping with the observations found in the field. Images of the Satélite Landsat-8, OLI sensor of the year 2014, path/row 222/062, 222/063, 223/062 and 223/063 were used to aid in the field. Using this data it was possible to analyze the main representative classes in the area, including agriculture, urban area, forest, clean pasture, dirty pasture, reforestation, regeneration with pasture and secondary vegetation. The secondary vegetation presented 2,198.16 km², clean pasture with 3,332.29 km², agriculture with 896.75 km² and the forest occupying 54.21% of the total area of Paragominas. The overall concordance index was 86%, corroborating the reliability of the mapping performed. The average error was 6% and the total value of discordance was of 14%. Concerning the secondary vegetation, pasture, agriculture, urban area and forest classes, they presented concordance higher to 50%, while regeneration with pasture and reforestation presented greater intensity of omission with 40.57% and 76.31% respectively. Inclusion errors were less than 40% for the secondary vegetation, pasture regeneration, clean pasture and dirty pasture classes. The field work was essential to validate and analyze the accuracy of the 2014 TerraClass Project for the studied region, which becomes important for the understanding of the dynamics of land use.

Keywords – image processing, land use, remote sensing, validation
(Macedo et al., 2013). State that this set of computational techniques has been strongly cooperating for a construction of knowledge about spatial and socioeconomic patterns. In addition, there are benefits from remote sensing application in studies based on the integration of field data with land use dynamics and vegetation cover (Almeida & Vieira, 2008).

Studies such as the one realized by Almeida et al. (2016) show that with the use of remote sensing, it is possible to understand the spatial patterns of land use and cover and it contributes to the studies of biodiversity, environmental modeling and climate change, essential for the creation and monitoring of land use policies. Forest formations are monitored by remote sensors in a time-scale that ranges from almost daily to annual (Diniz et al., 2015).

The conversion of natural landscapes to anthropic use or modification of management practices is known as land-use change (Foley et al., 2005). Which includes selective cutting, commercial planting of trees such as eucalyptus and paricá, conversion of forest to pastures, agricultural production areas, cut-crop agriculture and urbanization (Gardner et al., 2009). Which has expanded in the last decades, resulting in a fragmented landscape with different types of forest cover (Laurance et al., 2014).

The TerraClass Project, executed by the National Institute of Space Research (INPE), in a partnership between the Amazon Regional Center, Embrapa Oriental Amazon and Embrapa Livestock and Agriculture Informatics has as main objective qualifying the deforestation of the Legal Amazon, based on the deforestation areas mapped and published by PRODES Project and satellite images (Inpe, 2016), playing an important role in determining which land uses are replacing the Amazon forest (Almeida et al., 2016). In addition, this project is motivated by the concern of the scientific world and society in general with the threats to the greater biodiversity area of the planet, considering the role of Amazon forest in the context of global climate change (Inpe, 2016).

The interest in monitoring forest resources has been increasing by the past decades and so the demand for mapping plant cover at regional and global scales (Shimabukuuro, 2000). Thus, validating orbital remote sensing data through the verification of truth land points (Congalton e Green, 1999), has become difficult and sometimes impractical in some parts of the (Hess et al., 2002).

The Pará State has the largest deforested area of the BLA, with 262.088 km², equivalent to 33.72% of the Amazon’s total deforested area until 2016 (Inpe, 2017). This scenario is due to the installation of agromineral project in its territory, as the “Grande Carajás” Project for iron ore exploitation, the implementation of transport and energy infrastructures and the advancement of the agricultural frontier (Kohlhepp, 2002).

The accuracy assessment of the land use and cover maps has been well studied by researchers in remote sensing, however, the field measurements of large territorial mappings present obstacles because of the difficulty of access, operational and logistic costs, as well as the high planning and execution demanding time (Adami et al., 2012).

According to Almeida et al. (2016), the first TerraClass mapping had a global accuracy of 76%, when tested to the states of Pará and Mato Grosso, attending the accuracy’s need of the mapping and allowing statistic inferences (McRoberts, 2011).

In this sense, this paper has the objective of assessing the TerraClass mapping for the year of 2014, in a local scale, which will be applied to the municipality of Paragominas, in the state of Pará.

II. WORK METHODOLOGY

The study was realized in the municipality of Paragominas, situated at the margins of the Belém-Brasília Highway (BR-010), with connections with the PA-125 e PA-256 state highways. The population is estimated in 97,819 inhabitants, distributed in an area of 19,342.254 km² (IBGE, 2010). The municipality is part of the southeast region of Pará State and the Paragominas micro region. It presents limits with the municipalities of Ipixuna do Pará e Ulianópolis and its municipal seat is limited by the following geographic coordinates: 03°00’00” S e 47°21’30” W (Fig.1).

![Fig.1: Study Area Location Map](https://example.com/study_area_map.png)

The utilized images are from the Landsat-8 OLI (Operational Land Imager) satellite, which are available free of charge from the USGS (United States Geological Survey), with spatial resolution of 30 m. The images were from 2014, of the points 223 and 222 and orbits 62 and 63, dated from June, October and November.

The images were selected based on the lower presence of clouds and then submitted to geo referencing
and contrast enhancement treatment. At last, it was realized the procedures of equalization, to obtain better identification, and grouping of the land use and cover features (Fig. 2).

![Path/row of the 2014’s Landsat-8 images](image)

The municipality of Paragominas was chosen due to the fieldwork that was executed in 2014, through Sustainable Landscapes Project, which involved a land route and an overflight to register images of georeferenced points, generating pictures that based the mapping validation.

The validation of the land use and coverage map for the year 2014 was based on field data, collected with GPS navigation device for land and air. In total, 321 points were collected in the field, which were used to calculate global accuracy (Fig. 3).

![Fieldwork visited points in the municipality of Paragominas, PA](image)

The accuracy assessment is an important process in the remote sensing data analysis because it results in statistic inferences to the reliability of the generated data (McRoberts, 2011).

The overall accuracy is one of the simplest measures, calculated by the total sum of correctly classified pixels divided by the total number of pixels in the confusion matrix. In addition, individual category accuracy is calculated through the accuracy of the producer, indicating the likelihood of a reference pixel be correctly classified, and the user’s accuracy, which computes the probability of a mapped pixel be compatible with the field class (Congalton, 1991).

To calculate the Global Accuracy or Global Precision (GA), User Accuracy (UA) and Producer Accuracy (PA), the georeferenced pictures taken in the fieldwork were analyzed together with the TerraClass 2014’s mapping for the study area. For the calculation of accuracy, a contingency table was built (Pontius & SantaCruz, 2014) (Table 1).

### Table 1: Example of a contingency table containing the functions for calculating the GA, UA e PA.

| Reference Classes | Class1 | Class2 | Classc | Total |
|-------------------|-------|-------|-------|-------|
| Mapping Classes   | P1    | P2    | Pc    | P+    |
| Class1            | P11   | P12   | P1c   | P1+   |
| Class2            | P21   | P22   | P2c   | P2+   |
| Classc            | Pc1   | Pc2   | Pcc   | Pc+   |

The values allocated on the main diagonal of Table 1 represent the correctly classified elements, whereas the elements outside the main diagonal represent errors of omission and inclusion (Stehman and Foody 2009).

III. RESULTS AND DISCUSSIONS

The technique of visual interpretation of the spectral targets of a given satellite image is part of data generated through Remote Sensing. In this sense, the acquisition of points of terrestrial truth has become an important tool for validating the data generated by interpreters through orbital images (Espírito-Santo, 2005).

The TerraClass Project maps the following thematic classes: Annual agriculture, Mosaic occupations, Clean pasture, Dirty grass, Regeneration with pasture, Pasture with exposed soil, Secondary vegetation, Reforestation, Urban area, Others, Mining and Area not observed (Adami et al., 2015). With the exception of the annual Agriculture class, all other classes are mapped through photointerpretation. The annual Agriculture class is mapped from an automatic method based on the spectral-temporal behavior of the Normalized Difference Vegetation Index (NDVI), obtained by the MODIS sensor (Roerink et al., 2000).

The validation for the municipality of Paragominas was carried out with the following thematic classes of mapping of the TerraClass Project, year of 2014: agriculture, urban area, forest, clean pasture, dirty pasture, reforestation, pasture regeneration and secondary vegetation, representative in the municipality.

Based on the TerraClass 2014 mapping, the forest occupies 54.21% of the municipality, with 10,483.37 km². The secondary vegetation presented 2,198.16 km², the clean pasture with 3,332.39 km², the annual agriculture with 896.75 km² (Fig. 4).
Fig. 4: Use and land cover map year 2014 of the municipality of Paragominas, PA.

In all, these three classes correspond to approximately 85% of the area of the municipality. Another interesting factor to be observed is that the predominance of forest is inversely associated with the distance of the roads that cut the municipality. From the satellite image and photographs, it was possible to represent some validation points that can also be used for training in photointerpretation.

In point 1 it was possible to observe and validate in the field the class of urban area, with characteristics of streets, squares, houses positioned in a very closely and with regular spatial distribution (Fig. 5).

Fig. 5: Class of Urban area.

In section 2 the validation was verified by the observation in the field of the annual agriculture class represented with soybean planting, in the magenta color, smooth texture and regular polygons characterize the feature (Fig. 6).

Fig. 6: Class of Annual agriculture.

It was observed in section 3 the class clean pasture validated in the field according to the representation of well managed pasture area, with low infestation of herbaceous and shrubby weeds (Fig. 7).

Fig. 7: Pasture clean class
For point 4, the regenerativo pasture class was confirmed with the field truth by the presence of invasive species at various levels of development (Fig. 8).

Fig. 8: Regeneration pasture class.

For point 5 the validation of the Reforestation class, represented by areas with regular polygons, division of plots and homogeneous texture (Fig. 9).

Fig. 9: Reforestation class.

At point 6 the field truth was verified by the Forest class represented by the forest management reserve area, in the satellite image with rough texture and dark green color characteristics (Fig. 10).

Fig. 10: Forest class.

Point 7 represents the dirty pasture class with presence of many herbaceous and some shrub species (Fig. 11).

Fig. 11: Shrubby pasture class.

For the representation of point 8 the secondary vegetation class was validated proving the aspect of natural regeneration of native shrub and tree vegetation (Fig. 12).
Fig. 12: Secondary vegetation class.

Fig. 12: Secondary vegetation class.

The classification accuracy is expressed using the error matrix also called the confusion or contingency matrix, derived from the classification and reference data set as the starting point of the precision measurements (Story and Congalton, 1986), which shows the cross tabulation of the mapped land use and cover versus the actual one assessed in the field (Foody, 2002). The contingency matrix lists land use and cover reference values in the columns and the classified data in the rows, being the main diagonal of the matrix the correctly classified data (Banko, 1998).

Based on the contingency matrix, global accuracies, omission and inclusion errors per class and indices of disagreement (displacement, quantity and change) were calculated (Pontius and Santa Cruz, 2014). Studies performed in Cerrado areas show high reliability with a general agreement index of 80.2% (Maurano and Adami, 2017).

The validation process is a necessary step in a research that aims to analyze in a statistical way the accuracy of the classified images. Thus, table 2 shows the results of the contingency matrix of the classification for the municipality of Paragominas.

| TerraClass mapping class | Agriculture | Urban area | Forest | Clean pasture | Shrubby pasture | Reforestation | Regeneration with pasture | Secondary vegetation | Total |
|--------------------------|-------------|------------|--------|---------------|-----------------|---------------|---------------------------|----------------------|-------|
| Agriculture              | 50          | 1          | 3      |               |                 |               |                           |                      | 54    |
| Urban area               | 7           |            |        |               |                 |               |                           |                      | 7     |
| Forest                   | 110         | 3          | 1      | 3             |                 |               |                           |                      | 117   |
| Clean pasture            | 9           | 49         | 2      | 12            | 3               |               |                           |                      | 75    |
| Shrubby pasture          | 1           | 7          | 1      | 2             | 1               |               |                           |                      | 12    |
| Reforestation            | 7           |            |        |               |                 |               |                           |                      | 7     |
| Regeneration with pasture|             |            |        |               |                 | 7             | 6                         | 13                   |       |
| Secondary vegetation     | 4           | 3          | 4      |               |                 |               |                           |                      | 25    |
| Total                    | 59          | 7          | 114    | 57            | 9               | 27            | 13                        | 35                   | 321   |

The general agreement index between mapping and the validation was 86%, demonstrating high reliability of the mapping performed by TerraClass, based on the result of the contingency matrix.

According to Pontius & Santa Cruz (2014), the mean values of agreement and disagreement are divided into quantity, displacement, and change. The quantity errors refer to the absolute difference between omission and inclusion errors, presenting an indicative of the mean error of the mapped area. The displacement refers to the unpaired allocation difference between the analyzed classes in the contingency matrix and, at last, the change...
refers to the error caused by paired exchanges between classes in the contingency matrix (Foody, 2009).

Thus, the mean error presented for the area was 6%, displacement of 7% and the change of 1%. When these disagreements are summed, a total error of 14% is obtained, which is considered low due to the complexity of the mapped classes and size of the study area (Fig. 13).

**Fig. 13: Global Accuracy Results for the TerraClass 2014 in the municipality of Paragominas**

When analyzing the results by thematic class, it was observed that the classes of secondary vegetation, clean pasture, agriculture, urban area and forest presented agreement superior to 50%. The classes that presented the greatest intensity of omission were grass regeneration and reforestation, with 40.57% and 76.31%, respectively. Regarding the inclusion, secondary vegetation, regeneration with pasture, clean pasture and shrubby pasture presented errors in the order of 24.39%, 27.43%, 29.87% and 35.44% respectively (Fig. 14).

**Fig. 14: Global Accuracy Results of the TerraClass mapping for the municipality of Paragominas per class**

Possibly, the errors in the regeneration class with pasture are associated with the very characteristic of this class, which corresponds to a transition process from pasture to secondary vegetation or vice versa. In what concerns to the confusion between reforestation and pasture classes, it may be associated with the beginning of reforestation planting, when there is great soil and grassland exposure, making it possible to confuse with pasture classes.

IV. CONCLUSION

The presentation of the contingency matrix is fundamental, since it is possible to visualize the confusions that occurred between classes. The global accuracy may be sufficient, depending on the purpose of the map. However, if there are specific classes of interest or more than others, individual classes’ accuracy may be relevant.

The global accuracy of the TerraClass mapping for the municipality of Paragominas was 86%. Therefore, future research on this project should consider improvements for the mapping of regeneration with pasture and reforestation classes.

The validation, as part of the mapping process, estimates the data accuracy of the land cover use dynamics, giving support to the information consolidation.

ACKNOWLEDGEMENTS

We thank the staff at the Global Environment Facility Sustainable Landscapes Project for the logistical support. This research was carried out within the scope of the Environmental Monitoring by Satellite Project in the Amazon Biome, TerraClass with a partnership between the National Institute for Space Research - INPE / Amazon Regional Center - CRA, Embrapa Amazônia Oriental (CPATU) both located in Belém - PA and Embrapa Information Technology (CNPTIA) located in Campinas-SP.

REFERENCES

[1] Abreu, K. M. P & Coutinho, L. M. Sensoriamento remoto aplicado ao estudo da vegetação com ênfase em índice de vegetação e métricas da paisagem. Vértices, v. 16, n. 1, p. 173-198. 2014.

[2] Adami, M.; Gomes, A. R.; Coutinho, A. C. Dinâmica do uso e cobertura da terra no estado do Pará entre os anos de 2008 a 2012. Anais XVII Simpósio Brasileiro de Sensoriamento Remoto – SBSR. 2015.

[3] Adami, M.; Mello, M. P.; Mguiar, D. A.; Rudorff, B. F. T.; Souza, A. F. . A Web Platform Development to Perform Thematic Accuracy Assessment of Sugarcane Mapping in South-Central Brazil. Remote Sensing, v. 4, p. 3201-3214, 2012.

[4] Almeida, A. S.; Vieira, I. C. G. Dinâmica da Cobertura vegetal e uso da terra no município de São Francisco do Pará (Pará, Brasil) com o uso da técnica de sensoriamento remoto. Museu Paraense Emílio Goeldi. Ciências Naturais, Belém, v.3, n°1, p. 81-92, 2008.

[5] Almeida, C. A.; Coutinho, A. C.; Esquerdo, J. 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 w the
Brazilian Legal Amazon in 2008 using Landsat-5/TM and MODIS data. Acta Amazonica, v. 46, n. 3, p. 291-302, Sept. 2016.

[6] Banko G. A review of Assessing the accuracy of classifications of Remotely Sensed Data and of Methods Including Remote Sensing Data in Forest Inventory. Interim Report. IIASA. A-2361.Laxenburg-Austria. November 1998.

[7] Becker, B. K. Geopolítica da Amazônia. Estudos Avançados. v.19, n. 53. 2005.

[8] Congalton, R. G. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment, 37, 35-46. 1991.

[9] Congalton, R. G.; Green, K. Assessing the accuracy of remotely sensed data: principles and practices. New York: Lewis Publishers, 1999. 137 p.

[10] Diniz, C. G.; Souza, A. A. A.; Santos, D. C. S.; Dias, M. C.; Luz, N. C.; Moraes, D. R. V.; Maia, J. S.; Gomes, A. R.; Narvaez, I. S.; Valeriano, D. M.; Maurano, L. E. P.; Adami, M. DETER-B: The New Amazon Near Real-Time Deforestation Detection System. IEEE journal of selected topics in applied earth observations in applied earth observations and remote sensing, vol. 8, no. 7, july 2015.

[11] Espírito-Santo, F. D. B.; Shimabukuro, Y. E. Validation of tropical forest area mapping using aerial videography images and data from field work survey. Revista Árvore. V. 29. N. 2 Viçosa. Mar/Apr. 2005.

[12] Fearnside, P. M. Desmatamento na Amazônia: dinâmica, impactos e controle. Acta Amazonica. v. 36(3). P. 395-400.2006.

[13] Ferreira, J., Pardini, R., Metzger, JP., Fonseca, C., Pompeu, P., Sparovek, G., Louzada, J. Towards environmentally sustainable agriculture in Brazil: challenges and opportunities for applied ecological research. Journal of Applied Ecology, v. 49 (3), p. 535541. 2012.

[14] Foley, J.A., DeFries, R., Asner, G.P., Barford, C., Bonan, G.B., Carpenter, S.R., Chapin, F.S., Coe, M.T., Daily, G.C., Gibbs, H.K., Helkowski, J.H., Holloway, T., Howard, E.A., Kucharik, C.J., Monfreda, C., Patz, J.A., Prentice, I.C., Ramankutty, N., & Snyder, P.K., 2005. Global consequences of land use. Science, 309, 570–574.

[15] Foody, G. M. Status of land cover classification accuracy assessment. Remote Sensing of Environment, 80, p. 185–201. 2002.

[16] Foody, GM. Classification accuracy comparison: Hypothesis tests and the use of confidence intervals in evaluations of difference, equivalence and non-inferiority. Remote Sensing of Environment, 113, 1658-1663, 2009.

[17] Gardner, T.A., Barlow, J. Chazdon, R.L., Ewers, R., Harvey, C.A., Peres, C.A., & Sodhi, N. 2009. Prospects for tropical forest biodiversity in a human-modified world. Ecology Letters, 12, 561–582.

[18] Hess, L. L. et al. Geocoded digital videography for validation of land cover mapping in the Amazon basin. International Journal of Remote Sensing, v. 23, n. 7, p. 1527-1556, 2002.

[19] INPE. Instituto Nacional de Pesquisas Espaciais. Monitoramento da Floresta Amazônica Brasileira por Satélite. 2017. Disponível em http://www.obt.inpe.br/prodes/dashboard/prodes-rates.html. Acesso em 22/01/2018.

[20] INPE. Instituto Nacional de Pesquisas Espaciais. Projeto TerraClass 2014: Mapeamento do uso e cobertura da terra na Amazônia Legal Brasileira. Brasília. 2016. Disponível em <http://www.inpe.br/terraclasse/ofertas/index.php>. Acesso em nov. 2016.

[21] Instituto Brasileiro de Geografia e Estatística. Base de Informações do censo demográfico 2010. Brasil: IBGE. 2016. Disponível em: http://download.ibge.gov.br/downloads_estatistica.html. Acesso em nov. 2016.

[22] Kohlhepp, G. Conflitos de interesse no ordenamento territorial da Amazônia brasileira. Estudos Avançados. v. 16(45), p.37-61. 2002.

[23] Laurance, W.F., Sayer, J., & Cassman, K.G. 2014. Agricultural expansion and its impacts on tropical nature. Trends in Ecology & Evolution, 29, 107–116.

[24] Macedo et al. Configuração espacial do desflorestamento em fronteira agrícola na Amazônia: um estudo de caso na região de São Félix do Xingu, estado do Pará. Revista Nera. Ano 22 – janeiro/Junho de 2013 – ISSN: 1806-6755.

[25] Maurano, L. E. & Adami, M. Ferramenta Web-Gis para avaliar exatidão de mapeamento de uso e cobertura da terra no Cerrado brasileiro. In: Simpósio Brasileiro de Sensoriamento Remoto, 18. (SBSR), 2017, Santos. Acessível em: <http://www.inpe.br/cra/projetos_pesquisas/terraclasss2014.php>. Acesso em nov. 2016.

[26] Michele, E.G. Desmatamento na Amazônia brasileira: um estudo de caso na região de São Félix do Xingu, estado do Pará. Revista Nera – Ano 16, N.º. 22 – janeiro/Junho de 2013 – ISSN: 1806-6755.

[27] Maurano, L. E. & Adami, M. Ferramenta Web-Gís para avaliar exatidão de mapeamento de uso e cobertura da terra no Cerrado brasileiro. In: Simpósio Brasileiro de Sensoriamento Remoto, 18. (SBSR), 2017, Santos. Acessível em: <http://www.inpe.br/cra/projetos_pesquisas/terraclasss2014.php>. Acesso em nov. 2016.

[28] Roerink, G. J., Menenti, M., & Verhofen, W. Reconstructing cloud free NDVI composites using Fourier analysis of time series. International Journal of Remote Sensing, v. 21, n. 9, p.1911-1917, 2000.
[29] Shimabukuro, Y. E.; Rudorff, B. F. T. Fraction images derived from NOAA AVHRR data for global studies. International Journal of Remote Sensing, v. 21, n. 17, p. 3191-3194, 2000.

[30] Stehman, S.V., Foody, G.M. Accuracy assessment. In: T.A. Warner, M.D. Nellis, G.M. Foody (Eds.), The Sage Handbook of Remote Sensing, p. 297–309. London: SAGE, 2009.

[31] Story, M.; Congalton, R. G. Accuracy assessment: A user's perspective. Photogrammetric Engineering and Remote Sensing, v.52, p.397–399. 1986.

[32] USGS. United States Geological Survey. Landsat 8 (LDCM). Disponível em: http://landsat.usgs.gov/. Acesso em março/2013.