Mapping Brazilian Savanna Physiognomies using WorldView-2 Imagery and Geographic Object Based Image Analysis

Cesare Di Girolamo Neto [1], Leila Maria Garcia Fonseca [1], Thales Sehn Körting [1], Anderson Reis Soares [1]

1. National Institute for Space Research – INPE, Av. dos Astronautas, 1.758, 12227-010, São José dos Campos, São Paulo, Brazil. {cesare.neto; leila.fonseca; thales.korting; anderson.soares} @inpe.br

ABSTRACT. *Brazilian Savanna, or just "Cerrado", is considered one of the 25 hotspots for biodiversity conservation priority in the world. Cerrado occurs on the central part of Brazil and has three major natural formations: Grasslands, Savannas and Forests. However, the challenge on mapping Cerrado relies on the division of these major formations into specific physiognomies. Distinguishing each of these physiognomies is an important task to better evaluate smaller ecosystems, access carbon storage with greater precision and improve the exactitude of greenhouse gases emissions. Thus, the aim of this work is to evaluate the potential of very high spatial resolution imagery in order to improve the classification of 8 Cerrado physiognomies: Rocky Grasslands, Open Grasslands, Shrub Grasslands, Shrub Savanna, Typical Savanna, Dense Savanna, Flooded Plains with Palmtrees and Evergreen Forest. A WorldView-2 image was used for a protect area with over 30 thousand hectares of preserved Cerrado vegetation. Features such as surface reflectance, vegetation indices, tasseled cap transformation and spectral linear mixture models were used on the automatic classification. Random Forests algorithm was used with a 10-fold cross-validation. The Global Accuracy was of 67.7%. Values above 70% of User's Accuracy were obtained for classes such as Rocky Grasslands, Open Grasslands, Typical Savanna and Evergreen Forest. On the other hand, Flooded Plains with Palmtrees were omitted from the classification. Omission errors were also noticed for the classes of Shrub Savanna and Dense Savanna; they were sometimes misclassified as Typical Savanna which has a similar vegetation structure and tree cover percentage. The use of very high resolution images provided advantages on distinguishing Cerrado physiognomies on an automatic classification procedure. The detection of some classes was very precise and, despite the obtained misclassifications, it is an advance to distinguish some physiognomies that lower spatial resolution sensors are, hardly never, capable of distinguishing.*

KEYWORDS: *Data Mining; Random Forests; Image Classification; Very High Resolution Images; Cerrado; Carbon Storage; Remote Sensing.*

1. Introduction

Brazilian Savanna, or just Cerrado, occupies an area of approximately 2 million km^2 of the Brazilian territory. Cerrado biodiversity is one of the richest in the planet, containing more than 160 thousand species of plants, animals and fungi (Ferreira *et al.*, 2003). Considering Cerrado's full extent, it is also responsible for storing around 5.9 billion tons of carbon in the vegetation and 23.8 billion tons in the soil (MMA, 2014). On a global scale, Cerrado was considered one of the 25 hotspots for biodiversity conservation (Myers *et al.*, 2000).

In this sense, the loss of Cerrado natural vegetation reached almost 47% of its original area by 2015 (MMA, 2017). Only 8.7% of its area (about 175 thousand km^2) is in protected areas (MMA, 2015). In environmental terms, loss of biodiversity can directly or indirectly lead to problems such as soil erosion, water pollution, carbon cycle instability and probable microclimatic changes, as well as intensifying the biome fragmentation (Klink and Machado, 2005). Thus, promoting strategies to monitor and map areas of preserved Cerrado is mandatory.

However, mapping tropical heterogeneous biomes, such as Cerrado, must be done considering biological, climatic and topographic factors. Particularly for the Cerrado, it is important to consider the vegetation seasonality and typical burning events. These peculiarities generate different natural formations of the region, also called physiognomies. There are three major natural formations in the Cerrado: Grasslands, Savannas and Forests (Ribeiro and Walter, 2008). Forests represent areas with the predominance of tree species and continuous canopy formation. Savannas refer to areas with trees and shrubs spread over a stratum of grasses, without the formation of a continuous canopy. Grasslands correspond to areas with a predominance of herbaceous species and some shrubs.

Mapping these three major natural formations in the Cerrado is not a difficult task anymore, once Brazilian Ministry of Environment promotes several projects to perform this task. Some examples are the projects called PROBIO (Project for the Conservation and Sustainable Use of Brazilian Biological Diversity - Sano *et al.*, 2008) and TerraClass Cerrado (Mapping Land Use and Cover for Cerrado - MMA, 2015). Both projects mapped the entire Cerrado biome with visual interpretation of medium resolution images (30m Landsat-like data). On the other hand, they required from two to three years and an average of 30 geospatial analysts to complete the results of the visual interpretation. This execution time can be reduced by performing automatic classification of remote sensing data. Authors such as Paneque-Galvez *et al.* (2013), Muller *et al.* (2015) and Silva and Sano (2016) used this technique to map Grasslands, Savannas and Forest on Cerrado. The mentioned works obtained accuracies of over 80% for all of these classes.

The real challenge when classifying Cerrado physiognomies relies on the use of a more detailed classification legend. Ribeiro and Walter (2008) proposed a legend with greater detail of Cerrado vegetation, dividing the three major natural formations into 13 physiognomies (Rocky Grasslands, Open Grasslands, Shrub Grasslands, Shrub Savanna, Typical Savanna, Dense Savanna, Rocky Savanna, Flooded Plains with Palmtrees, Savanna Parkland, Forested Savannah, Evergreen Forest, Dry Forest and Semideciduous Forest). Discriminating the vegetation into a greater number of physiognomic types implies into detailing the biodiversity of smaller ecosystems and also allows a better estimative of carbon storage and possible greenhouse gases emissions. This is especially important for Cerrado when considering the typical burning events on the biome.

This detailed legend was used by Ferreira *et al.* (2007), Oliveira *et al.* (2007), Costa *et al.* (2014) and Schwieder *et al.* (2016) in order to automatically classify Cerrado physiognomies using 30m spatial resolution images. Oliveira *et al.* (2007) and Costa *et al.* (2014) worked with physiognomies such as Rocky Grasslands, Open Grasslands and Shrub Grasslands. The Global Accuracy was of 65.70% and 67.20%, respectively. Costa *et al.* (2014) pointed out a great difficulty to distinguish the patterns of Rocky Grasslands and Open/Shrub Grasslands. Ferreira *et al.* (2007) and Schwieder *et al.* (2016) worked with the Spectra Linear Mixture Model (Shimabukuro and Ponzoni, 2017) and the Tasseled Cap Transformation (Crist and Cicone, 1984), respectively. Both authors pointed misclassification errors between similar physiognomies. These errors were related to incorrect labeling of some classes, which happened due to the difficulty of identifying then on the medium resolution images and also to the presence of smooth transition areas between two or more physiognomies.

In order to reduce these problems, the use of high (4 to 10m) spatial resolution images was evaluated by Teixeira *et al.* (2015) and Girolamo Neto *et al.* (2017). They worked with 5m spatial resolution images in order to classify Cerrado physiognomies. Using spectral data and the techniques mentioned on the works of Ferreira *et al.* (2007) and Schwieder *et al.* (2016), they were able identify different forest classes and also reduced errors between the transition of grasslands and savanna (especially between Shrub Grasslands and Shrub Savanna). Pinheiro and Durigan (2009) worked with very high spatial (1 to 4m) resolution images to identify Cerrado physiognomies. No automatic classification was performed, but the authors reported that vegetation structural parameters were distinguished on the visual classification. Therefore, these images could reduce even more the classification errors mentioned for the medium resolution images.

Therefore, it is evident that high and very high resolution images have potential to improve Cerrado physiognomies classification. Hence, the aim of this work is to evaluate the potential of WorldView-2 (2m spatial resolution) images in order to classify Cerrado physiognomies according to a detailed legend developed by Ribeiro and Walter (2008).

2. Methodology

Figure 1 presents the methodology flowchart proposed to automatically classify Cerrado physiognomies according to Ribeiro and Walter (2008) legend.

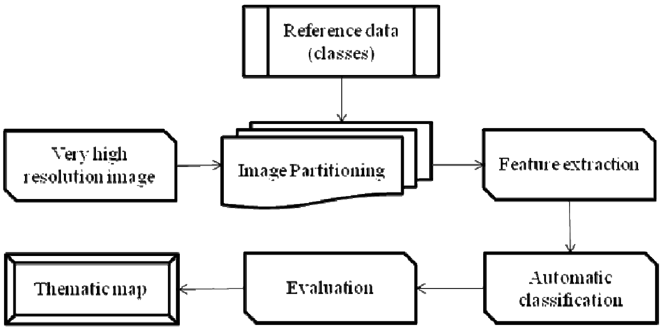

FIGURE *1. Methodology flowchart for automatic classification of Cerrado physiognomies.*

2.1. Study Site and WorldView-2 Image

The study area is located within the Brasília National Park on Distrito Federal State, Brazil, which has approximately 30,000 ha of preserved natural Cerrado vegetation. Figure 2 shows the major part of the park, in which a red line highlights the study area. For this work, a Multispectral WorldView-2 image was used (tile ID 103001003373A600). This image was acquired in 07/22/2014 at 12:48:56 pm in a level 2A product (Digital Globe, 2018). WorldView-2 multispectral band characteristics are on Table 1. The acquisition was made through an academic cooperation contract between Digital Globe Foundation and National Institute for Space Research (INPE, 2016).

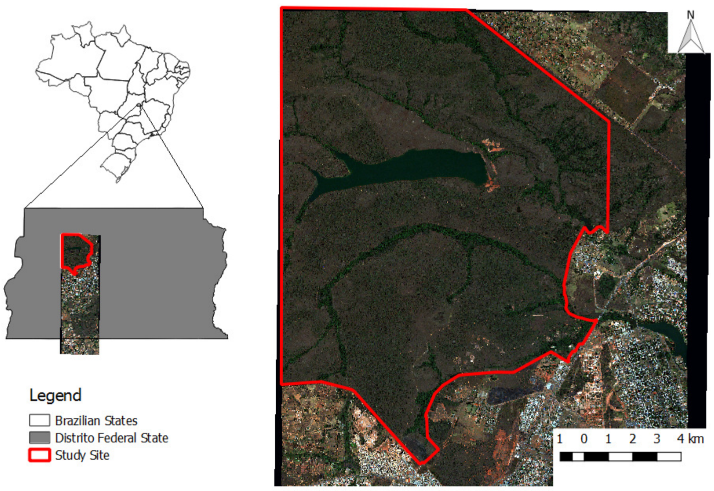

FIGURE 2. *Study site highlighted on with a red line on the WorldView-2 image (true color RGB composition).*

TABLE *1. WorldView-2 multispectral band characteristics.*

Band number	Band name	Wavelength (ηm)
1	Coastal	400-450
2	Blue	450-510
3	Green	510-580
4	Yellow	585-625
5	Red	630-690
6	Red edge	705-745
7	NIR1	770-895
8	NIR2	860-1040

As a reference sample, a total of 1062 points were randomly selected. These samples were validated by a field work realized on 07/2017 and also visual interpretation of pan-sharpened WorldView-2 image (with 0.5m of spatial resolution). The detailed system proposed by Ribeiro and Walter (2008) was used to classify 8 Cerrado physiognomies described in Table 2. Other classes such as Water Bodies, Marsh, Reforestation, Bare Soil, Roads and Constructed Areas were removed from the dataset.

TABLE 2. *Description of Cerrado physiognomies (Ribeiro and Walter, 2008).*

Physiognomy name	Vegetation description	Tree cover (%)	Tree height (m)
Rocky Grassland (RG)	Rocks and Grasses	0	-
Open Grassland (OG)	Grasses	0	-
Shrub Grassland (SG)	Grasses and Shrubs	0-5	-
Shrub Savanna (SS)	Shrubs and a few trees	5-20	2-3
Typical Savanna (TS)	Trees and a few Shrubs	20-50	3-6
Dense Savanna (DS)	Dense trees and a few Shrubs	50-70	5-8
Flooded Plains with Palmtrees (FP)	Grasses and Palmtrees	0-80	8-15
Evergreen Forest (EF)	Trees	70-95	15-30

2.2. Image partitioning

In order to generate objects for the automatic classification, the image was partitioned into square objects of 15 x 15 pixels size. The use of square objects instead of polygons extracted from segmentation algorithms based on similarity allowed us to evaluate features that capture the natural heterogeneity in the scene. This object size generated similar areas to Landsat-like pixel, allowing a better comparison of the results with other studies realized on medium resolution images. The field work and the visual interpretation of the WorldView-2 image (including the pan band with 0.5m of spatial resolution) allowed us to estimate tree cover percentage according to the legend proposed by Ribeiro and Walter (2008). Figure 3 illustrates a part of the image, with the square objects generated and also the excluded classes mentioned on section 2.1.

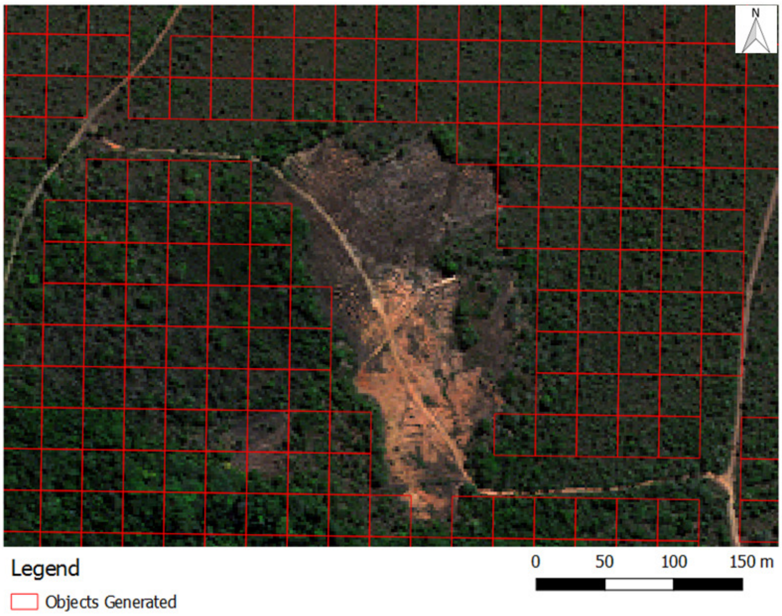

FIGURE *3. Results of the image partitioning and the removal of the classes of bare soil and roads.*

2.3. Feature extraction

From the square objects, the features for classification were extracted. The original WorldView-2 image was processed to Surface Reflectance using the Flaash algorithm on ENVI software (Exelis, 2013). The mean values for each object was extracted for the bands 2 to 8 from Table 1.

Considering the works of Ferreira *et al.* (2007) and Schwieder *et al.* (2016) we also extracted the components of the Spectral Linear Mixture Model and the Tasseled Cap Transformation. For the Spectral Linear Mixture Model, 10 pure pixels of the endmembers of soil, shadow and vegetation were selected. From these pixels each fraction image was generated and the mean values were extracted. A similar procedure was performed for the Tasseled Cap transformation. The coefficients for the transformation were obtained according to Yarbrough *et al.*

(2014). The images generated were for the greenness, brightness and wetness components.

Besides these features, we also considered a few vegetation indices for this approach. Liesenberg *et al.* (2007) stated that vegetation indexes can distinguish some Cerrado physiognomies and that, depending on the season, some indices might be better than others. In this sense, 5 vegetation indices were extracted from the objects. These indices were used on several studies mentioned before. A summary of all features used on the automatic classification is presented on Table 3.

TABLE *3. Features used on the automatic classification.*

Feature name	Description
Mean_Band_2	Surface reflectance from blue band
Mean_Band_3	Surface reflectance from green band
Mean_Band_4	Surface reflectance from yellow band
Mean_Band_5	Surface reflectance from red band
Mean_Band_6	Surface reflectance from red-edge band
Mean_Band_7	Surface reflectance from NIR1 band
Mean_Band_8	Surface reflectance from NIR2
SLMM_Soil	Soil component from the Spectral Linear Mixture Model
SLMM_Shadow	Shadow component from the Spectral Linear Mixture Model
SLMM_Vegetation	Vegetation component from the Spectral Linear Mixture Model
TC_Greenness	Greenness component from the Tasseled Cap Transformation
TC_Brightness	Brightness component from the Tasseled Cap Transformation
TC_Wetness	Wetness component from the Tasseled Cap Transformation
NDVI	Normalized Difference Vegetation Index
EVI	Enhanced Vegetation Index
EVI2	Enhanced Vegetation Index-2
SAVI	Soil Adjusted Vegetation Index
MSAVI2	Modified Soil Adjusted Vegetation Index-2
Class	Physiognomy name

2.4. Classification

The automatic classification was performed using the Random Forests algorithm (Breiman *et al.*, 2001), implemented in the WEKA software (Hall *et al.*, 2009). The number of trees on each forest was set to 100, following Breiman *et al.* (2001) recommendation. In order to get more reliable results, 10 different Random Forests were generated and the evaluation metrics were averaged. For the validation process, we used metrics such as Global Accuracy, Sensitivity (Producer's Accuracy), Precision (User's Accuracy), False Positive Rate (Commission Error) and False Negative Rate (Omission Error).

TABLE *4. Performance metrics for evaluating the classification.*

$$\text{Global Accuracy} = (TP+TN) / n \qquad (1)$$

$$\text{Producer's Accuracy} = TP / (TP+FN) \qquad (2)$$

$$\text{User's Accuracy} = TP / (TP+FP) \qquad (3)$$

$$\text{Commission Error} = FP / (FP+TN) \qquad (4)$$

$$\text{Omission Error} = FN / (FN+TP) \qquad (5)$$

in which: TP = True Positive, FP = False Positive, TN = True Negative, FN = False Negative and n = number of samples.

The Producer's Accuracy refers to the map accuracy from the perspective of the map producer. This is how often the real features on the ground are correctly shown on the classified map or the probability that a certain land cover of an area on the ground is classified as such. The Producer's Accuracy is complementary to the Omission Error (exclusion). The User's Accuracy is the accuracy from the point of view of a map user, not the map maker. The User's Accuracy essentially shows how often the class on the map will actually be present on the ground. The User's Accuracy is related to Commission Error (inclusion) (Humboldt, 2016).

3. Results and discussion

In order to present the results, Table 5 shows the averaged performance metrics mentioned on section 2.4. for each class and Table 6 represents the confusion matrix for the fold that generated the model with best performance.

The Global Accuracy of the classifier was of 67.7 ± 0.5%. When compared to works that used a less detailed classification legend, it can be considered low, once these studies have at least 80% of Global Accuracy. When considering studies that used Ribeiro and Walter (2008) detailed classification legend, the Global Accuracy ranged between 63% from 83%. It is also important to evaluate how many physiognomies were present on each case, for example Ferreira *et al.* (2008) reached an accuracy of 83% classifying only 4 physiognomies. When the number of classes improved, the Global Accuracy tended to fall. Working with 5 classes, Girolamo Neto *et al.* (2017) reached 74.3%, while working with 6 and 7 classes, Oliveira *et al.* (2007) and Schwieder *et al.* (2016) obtained 65.7% and 63.0%, respectively. It is also important to state that neither of the mentioned studies tried to classify Rocky

Grasslands and Flooded Plains with Palmtrees. Costa *et al.* (2014) worked with 4 physiognomies, including Rocky Grasslands and the Global Accuracy was of 67.2%. Considering that this work has 8 classes, including Rocky Grasslands and Flooded Plains with Palmtrees, the Global Accuracy of 67.7% can be considered above average.

TABLE 5. *Classification results*

Physiognomy	PA (%)	UA (%)	CE (%)	OE (%)
Rocky Grassland (RG)	50.8 ± 2.6	75.3 ± 0.9	0.2 ± 0.0	49,2 ± 2.6
Open Grassland (OG)	56.5 ± 0.9	70.0 ± 1.4	2.8 ± 0.2	43.5 ± 0.9
Shrub Grassland (SG)	64.5 ± 0.8	61.1 ± 0.5	10.2 ± 0.2	35.5 ± 0.8
Shrub Savanna (SS)	38.8 ± 1.2	46.2 ± 1.1	9.8 ± 0.4	61.2 ± 1.2
Typical Savanna (TS)	86.1 ± 0.7	73.2 ± 0.7	17.1 ± 0.5	13.9 ± 0.7
Dense Savanna (DS)	42.7 ± 1.7	57.3 ± 2.0	2.0 ± 0.1	57.3 ± 1.7
Flooded Plains with Palmtrees (FP)	0.0 ± 0.0	0.0 ± 0.0	0.1 ± 0.0	100.0 ± 0.0
Evergreen Forest (EF)	98.9 ± 0.0	95.6 ± 0.4	0.4 ± 0.0	1.1 ± 0.0

in which: PA = Producer's Accuracy, UA = User's Accuracy, CE = Commission Errors, OE = Omission Errors.

The best classification results were obtained to the Evergreen Forest class. Both User's and Producer's Accuracy are considered high and the errors were almost zero. This is the only class that is considered a Forest among all other classes, as mentioned on Table 2, it is the only one that can have a continuous canopy with almost 100% tree cover. Due to this propriety, some features such as vegetations indices, the Greenness and Vegetation components have higher values on this class when compared to others. This difference was captured by the classifiers and the high accuracy values were obtained. Authors such as Oliveira *et al.* (2007) and Ferreira *et al.* (2007) also obtained high accuracy values for this class. Classification errors of Evergreen Forest with Dry Forest were spotted by Teixeira *et al.* (2015) and Sano *et al.* (2008) noticed that some areas of reforestation were incorrectly classified as Evergreen Forest. These errors do not apply to this study, once these classes were not evaluated by the classifiers.

On the other hand, the worst classification results belong to the class of Flooded Plains with Palmtrees. The omission error of 100% means that the model was not able to recognize this pattern. When analyzing the confusion matrix presented on Table 6, it is possible to evaluate that this class was misclassified as Evergreen Forest, Dense Savanna and Typical Savanna. These errors can be explained on a first moment due to the small number of samples collected from this class. The lack of samples hardens the patter identification by the classifiers, leading misclassification errors. On the field work these physiognomy was noticed only a few times on the entire study area. It is also important to evaluate the vegetation structure of this class. As presented on Table 2, the tree canopy variation from 0% to

80% makes this class very heterogenic. On spots that have a very high tree cover percentage, the misclassification with Evergreen Forest is inevitable. This repeats for the classes of Dense Savanna and Typical Savanna, as the tree cover percentage diminishes.

TABLE 6. *Confusion matrix.*

		Classified as							
		RG	OG	SG	SS	TS	DS	FP	EF
Reference	RG	6	3	2	1	0	0	0	0
	OG	2	64	30	9	7	0	0	0
	SG	0	21	138	42	10	0	0	0
	SS	0	3	51	74	60	0	0	0
	TS	0	1	3	32	323	14	0	0
	DS	0	0	0	0	32	29	1	1
	FP	0	0	0	0	3	4	0	3
	EF	0	0	0	0	0	1	0	92

Classes: Rocky Grassland (RG); Open Grassland (OG); Open Grassland (OG); Shrub Grassland (SG); Shrub Savanna (SS); Typical Savanna (TS); Dense Savanna (DS); Flooded Plains with Palmtrees (FP) and Evergreen Forest (EF)

Another class that obtained very solid results on the classification was the Typical Savanna. Its Producer's Accuracy was of 86.1% and the User's Accuracy a bit lower, 73.2%. These values were higher than those obtained by Oliveira *et al.* (2007), Schwieder *et al.* (2016) and Girolamo Neto *et al.* (2017) and equivalent with Ferreira *et al.* (2007). The reduction of User's Accuracy occurred due to the increase of commission errors (17.1%), which were the misclassification of some samples of Dense Savanna and Shrub Savanna as Typical Savanna. These errors are also related to the similarity of each physiognomy in terms of tree cover %. Besides Pinheiro and Durigan (2009) statement, that high resolution images can distinguish physiognomy succession, the transition between them may be very smooth and, sometimes, very difficult to detect even on field (Ribeiro and Walter, 2008). Despite good accuracy values for typical savanna, Dense Savanna got worst classification results when compared to Schwieder *et al.* (2016). They obtained values of 60% and 64% for User's and Producer's Accuracy, respectively. Other studies mentioned before did not classified Dense Savanna.

The omission errors for Shrub Savanna made this class achieve poor results as well. Together with the Shrub Grasslands, theses physiognomies represent the transition between the major formation of grasslands and savannas. This transition is pointed out as a source of error on almost all works involving classification of Cerrado physiognomies. This work is no exception. This class is misclassified as Typical Savanna and also Shrub Grasslands. Ferreira *et al.* (2007) pointed out the components of the Spectral Linear Mixture Model were able to reduce these errors, but when analyzing our data all these 3 physiognomies (specially Shrub Savanna and Shrub Grasslands) presented very close values for these components.

Nevertheless, when we evaluate User's Accuracy for the classes of Shrub Grasslands and Open Grasslands, they were superior to other works, such as Schwieder *et al.* (2016) and Oliveira *et al.* (2007). Shrub grasslands suffered from the inclusion errors of the Shrub Savanna and presented a User's Accuracy of 61% and Open Grasslands of 70%. Girolamo Neto *et al.* (2017) got better results for Open grasslands and worst results for Shrub Grasslands. The great results for Open Grasslands were usually related to texture features that were used. The presence of some shrubs and even a few small trees on the Shrub Savanna and the complete absence for the Open Grasslands represented a change on the entropy feature for these classes, and thus, the better classification results.

Finally, the physiognomy of Rocky Grasslands obtained a User's Accuracy of 75.3% and Commission error of 0.2%. Costa *et al.* (2014) had a lot of difficulties trying to classify different grasslands physiognomies and only got values higher than 75% when the authors used open, shrub and Rocky Grasslands on a single class. The authors did not present the confusion matrix on the work; therefore we cannot compare different metrics for each physiognomy. However, the discrimination of Rocky Grasslands can be considered an advance using very high resolution images when compared to other studies with medium resolution data. The fully classified image is presented on Figure 4. It is important to remember that the gaps that are not classified are classes such as Water Bodies, Marsh, Reforestation, Bare Soil and Constructed Areas that were removed from the dataset.

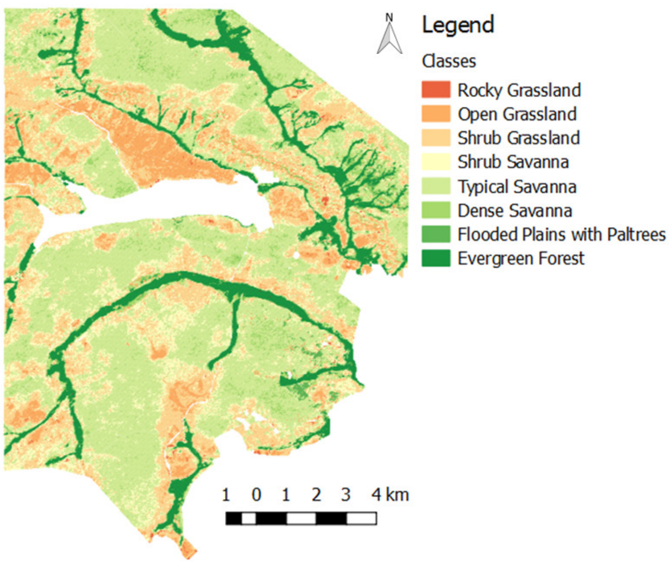

FIGURE 4. *Thematic map generated after the classification.*

4. Conclusion

In this study we evaluated a very high resolution image (2m) in order to classify 8 different Cerrado physiognomies in Brazil. The Global Accuracy of the classifier was of 67.7 ± 0.5%. Some physiognomies such as Rocky Grasslands, Open Grasslands, Typical Savanna and Evergreen Forest got good User's Accuracy values (over 70%). Most classification errors on were related to the transition of grasslands to savannas, represented by the classes of Shrub grasslands and Shrub savanna. They have a very smooth transition and were very difficult to be distinguished from each other. Problems were also reported for physiognomies with greater tree cover percentage, such as Dense Savanna and Flooded Plains with Palmtrees. Dense Savanna was sometimes misclassified as Typical Savanna and the Flooded Plains with Palmtrees were not detected at all for the classifiers. Besides being a very heterogeneous class, Flooded Plains with Palmtrees had just a few patches on the study area, and more samples could improve the classifiers results. As future works, we recommend the use of texture in order to distinguish classes with different tree cover percentage and also the application of more powerful learning techniques, such as deep learning algorithms for example.

Acknowledgements:

We want to thank the Digital Globe foundation for supplying the image for this study and cooperating with postgraduate students at INPE. We also want to thanks the staff of Brasilia National Park, which received us very well and helped several times during the field work.

References

Breiman, L. Random forests. (2001) *Machine Learning Journal*, v.45, p.5-32.

Costa, W. S.; Fonseca, L. M. G.; Körting, T. S. (2014) Mapping Grasslands Formations and Cultivated Pastures in the Brazilian Cerrado Using Data Mining. *6º GEOProcessing - International Conference on Advanced Geographic Information Systems, Applications, and Services*, Barcelona, Spain.

Crist, E. P.; Cicone, R. C. (1984). A physically-based transformation of Thematic Mapper data---The TM Tasseled Cap. *IEEE Transactions on Geoscience and Remote sensing*, v.3, p.256-263.

Ferreira, L. G.; Yoshioka, H.; Huete, A.; Sano, E. E. (2003) Seasonal landscape and spectral vegetation index dynamics in the Brazilian Cerrado: An analysis within the Large-Scale Biosphere–Atmosphere Experiment in Amazônia (LBA). *Remote Sensing of Environment*, v.87, n.4, p.534-550.

Ferreira, M. E.; Ferreira, L. G.; Sano, E. E.; Shimabukuro, Y. E. (2007) Spectral linear mixture modeling approaches for land cover mapping of tropical savanna areas in Brazil. *International Journal of Remote Sensing*, v.2, n.28, p.413-429.

Girolamo Neto, C. D.; Fonseca, L. M. G.; Körting, T. S. (2017). Assessment of texture features for Brazilian savanna classification: a case study in Brasilia National Park.

Brazilian Journal of Cartography, v.5, n.69, p.891-901.

Hall, M. A.; Frank, E.; Holmes, G.; Pfahringer, B.; Reutemann, P.; Witten, I. H. (2009) The WEKA Data Mining Software: An Update. *SIGKDD Explorations*. v.11, n.1, p.10-18.

Humboldt State University. (2006). *Introduction to remote sensing - Accuracy Metrics,* http://gsp.humboldt.edu/olm_2016/Courses/GSP_216_Online/lesson6-2/metrics.html

INPE – National Institute for Space Research (2016). *Academic Cooperation Agreement*. available at: http://www.inpe.br/institucional/sobre_inpe/relacoes_internacionais/arquivos/theDigitalGlobeFoundation.pdf

Klink, C.; Machado, R. (2005) Conservation of the Brazilian Cerrado. *Conservation Biology*, v.19, n.3, p.707-713.

Liesenberg, V.; Ponzoni, F. J.; Galvão, L. S. (2007) Análise da dinâmica sazonal e separabilidade espectral de algumas fitofisionomias do Cerrado com índices de vegetação dos sensores Modis Terra e Aqua. *Revista Árvore*, v.31, n.2, p.295-305, 2007.

MMA - Ministério do Meio Ambiente (Brazilian Ministry of Environment) (2014) *Biomas do Brasil*, available at: www.mma.gov.br/biomas/cerrado, 2014.

MMA - Ministério do Meio Ambiente (Brazilian Ministry of Environment) (2015) *Mapeamento do uso e cobertura da terra do cerrado – Projeto TerraClass Cerrado 2013*, available at: http://www.mma.gov.br/publicacoes/biomas/category/62-cerrado, 2015.

MMA - Ministério do Meio Ambiente (Brazilian Ministry of Environment) (2017) *Cerrado deforestation rates*, available at: http://www.dpi.inpe.br/fipcerrado/dashboard/cerrado-rates.html, 2017.

Müller, H; Rufin, P.; Griffiths, P.; Siqueira, A. J. B.; Hostert, P. (2015) Mining dense Landsat time series for separating cropland and pasture in a heterogeneous Brazilian savanna landscape. *17º Brazilian Symposium of Remote Sensing*, 2015, João Pessoa, Brazil.

Myers, N.; Mittermeier, R. A.; Mittermeier, C. G.; Fonseca, G. A. B.; Kent, J. (2000) Biodiversity hotspots for conservation priorities. *Nature***,** v.403, p.853-858.

Oliveira, M. E. A.; Martins, F. R.; Castro, A. A. J. F.; Ssantos, J. R. (2007) Classes de cobertura vegetal do Parque Nacional de Sete Cidades (transição campo-floresta) utilizando imagens TM/Landsat, NE do Brasil. *13º Brazilian Symposium of Remote Sensing*, 2007, Florianópolis, Brazil.

Paneque-Gálvez, J.; Mas, J. F.; Moré, G.; Cristobál, J.; Orta-Martinez, M.; Luz, A. C.; Guèze, M.; Macía, M. J.; Reyes-Garcia, V. (2013) Enhanced land use/cover classification of heterogeneous tropical landscapes using support vector machines and textural homogeneity. *International Journal of Applied Earth Observation and Geoinformation,* v.23, n.1, p.372-383.

Pinheiro, E. D. S.; Durigan, G. I. S. E. L. D. A. (2009) Dinâmica espaço-temporal (1962-2006) das fitofisionomias em unidade de conservação do Cerrado no sudeste do Brasil. *Revista Brasileira de Botânica*, v.32, n.3, p.441-454.

Ribeiro, J. F.; Walter, B. M. T. (2008). As principais fitofisionomias do Bioma Cerrado. In: Sano, S. M.; Almeida, S. P.; Ribeiro, J. F. *Cerrado: ecologia e flora*, Embrapa, Brasília.

Sano, E. E.; Rosa, R.; Brito, J. L. S.; Ferreira, L. G. (2007). Mapeamento de cobertura vegetal do bioma Cerrado. *Embrapa Cerrados: Brasília*.

Schwieder, M.; Leitão, P. J.; Bustamante, M. M. C., Ferreira, L. G.; Rabe, A.; Hostert, P.

(2016). Mapping Brazilian savanna vegetation gradients with Landsat time series. *International Journal of Applied Earth Observation and Geoinformation*, v.*52*, p.361-370.

Shimabukuro, Y. E.; Ponzoni, F. J. (2017). *Mistura Espectral - Modelo Linear e Aplicações*, Oficina de textos, São Paulo.

Silva, L. R.; Sano, E. E. (2016). Análise das imagens do satélite RapidEye para discriminação da cobertura vegetal do bioma Cerrado. *Brazilian Journal of Cartography*, v.7, n.69, p.1269-1283.

Teixeira, L. R.; Nunes, G. M., Finger, Z.; Siqueira, A. J. B. (2015). Potencialidades da Classificação Orientada a Objetos em Imagens SPOT5 no Mapeamento de Fitofisionomias do Cerrado. *Revista ESPACIOS,* v.36, n.20.

Yarbrough, L. D.; Navulur, K.; Ravi, R. (2014). Presentation of the Kauth–Thomas transform for WorldView-2 reflectance data. *Remote Sensing Letters*, v.5, n.2, p131-138.