OBJECT-ORIENTED CLASSIFICATION OF FOREST SUCCESSIONAL STAGES

A. P. D. Corte[a]*, F. F. Kirchner[b], C. R. Sanquetta[a], B. Doubrawa[c], A. Machado[d]

[a]Federal University of Paraná, Curitiba, Brazil - (anacorte, sanquetta)@ufpr.br
[b]Instituto de Ecologia Terrestre, Curitiba, Brazil - flaviokirchner@gmail.com
[c]Federal University of Paraná, Curitiba, Brazil - betaaa_d@yahoo.com.br
[d]RagServ, Curitiba, Brazil – ayrton@ragserv.com.br

KEY WORDS: mapping, object-oriented classification, forests.

ABSTRACT

This study sought to develop a protocol for the classification of forests in the Atlantic Forest in Paraná, Brazil. SPOT5 images with spatial resolution of 5 meters were used. We tested various band combinations for the classification. The first principal component, and the red and green bands were used to create the RGB color composite image: PC1, R, G. Field data was used to create forest classes: native forests, homogeneous stands of trees, and non-forest. A multi-resolution segmentation to the first hierarchical level using the criterion of reference scale value of 30 was conducted. The color and form homogeneity criteria were set to 0.9 and 0.1, respectively. The classification was conducted in Definiens Professional, which separated scenes into two classes: forest and non-forest. Using the brightness information from band 3 and the criteria established for each class, another multi-resolution segmentation with a scale reference value 20 and the other targeting parameters equal was performed. Using the texture homogeneity criterion for all directions, it was possible to establish the pattern of homogeneous tree stands. By inverting the expression results for the remaining forest classes were obtained. To test the results generated by the classification a grid containing 50 sample points were selected as validation control points. The analysis showed that 90% accuracy in classification may be achieved through automated procedures. For the remaining 10%, confusion existed between the shaded areas and homogeneous forest stands. Meanwhile, confusion was low between the native forest class and homogenous forest stands.

1. INTRODUCTION

A forest is defined according to the *Global Forest Resources Assessments* – FRA document (2001) as land with a canopy cover (or its equivalent degree of coverage) of greater than 10% and with a surface area greater than 0.5 hectare. The general forest definition mentioned above considers typical natural forests and forest plantations.

The Ministry of Environment in 2007 released the results of its Brazilian Biome Land Cover Map, included in the Conservation and Sustainable Use of Diversity Project MMA/PROBIO (2007), which gave an overview of the country's remaining forests, and the Atlantic Forest biome in particular, where the present study is located.

According to IBGE (1992), the original vegetation of the study area is characterized as belonging to Mixed Atlantic Forest or Araucaria Forest. The floristic composition of this vegetation type is characterized by the association of Araucaria with different groups of species.

Mapping the remaining native vegetation in the State of Paraná and monitoring its evolution over time are essential steps for conservation planning and preservation at the landscape scale.

Add to this the fact that the existing vegetation maps conducted in the State, or the scope of their analysis, are outdated or were created at a scale whose poor detail does not permit accurate estimates native vegetation cover. They present limited value therefore in conservation planning or in obtaining reliable data.

Given the concern over strategic areas for Biodiversity, located in south central Paraná, an area of extreme importance that concentrates important fragments of native vegetation, there is need for focused work in this region to implement and monitor wildlife corridors, in order to maintain the biological fluxes of both flora and fauna.

In order for vegetation mapping in the south central Paraná to meets the needs mentioned above, it is vitally important that we be informed by detailed surveys including a representative sampling of points the existing vegetation classes.

2. MATERIALS AND METHODS

2.1 Study Area

The present study was conceived with the intention to update maps of remaining native forest in south central state of Paraná, Brazil, using images of high spatial resolution. The study areas included areas administered by the regional offices of the Environmental Institute of Paraná (IAP): the Guarapuava Office, the Pato Branco Office, the Irati Office, and the União de Victoria Office (Figure 1).
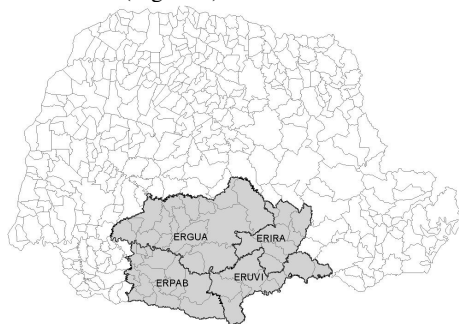

Figure 1 – Updated map project study area.

2.2. Methods

The satellite images used in this study were provided by the Paraná State Secretary of the Environment and Water Resources (SEMA), through the Office of Land, Cartography and Geosciences (ITCG). We used orthorectified SPOT 5 satellite imagery, with spatial resolution of 5 meters, and 4 bands. The scenes had temporal coverage over the years 2005 and 2007.

The aforementioned images were configured according to the following parameters: orthorectified to the UTM (Universal Transverse Mercator) Zone 22 South cartographic projection system, SAD 69 (South American Datum).

Sampling points were collected within the mapping area boundary in support of the development of the classification. The sampling points were collected during four field expeditions conducted between June and September 2010. Sampling points were collected from four different vegetation classes, namely: early stage forest succession (50 points), moderate and advanced stage forest succession (90 points), natural grasslands (50 points) and homogeneous tree stands or plantations for harvest (68 points).

The collection of sample points was performed by expeditions with routes predefined from satellite images and vegetation maps. A great majority of highways and secondary roads in the mapping area were travelled. For each sample point, the team visited the forest stand to record data using data collection sheets and to take photographs.

Sample points assignments to the various successional stages was made possible through the team members' field experience, and followed criteria established by CONAMA Resolution No. 2/94.

In order to improve the classification performance of images, tests on each were performed using several band compositions. Based on the tests, the best composition was selected for object-oriented classification. The selected band composition involved the first principal component of all four bands of the orthorectified scenes, to create the RGB color composite image: PC1, R, G. In general principal components perform well in vegetation studies, according to Jones & Vaughan (2010).

The classification was carried out using e-Cognition® software which is based on a new approach to automated image analysis.

e-Cognition® (Baatz et al., 2004) is commercial software that uses a system that interprets based on image segments and hierarchical networks. The aim is to employ the user's knowledge to categorize the hierarchical networks obtained from an automated multilevel segmentation of the image. Put more simply, the system can be thought of as a "knowledge-based system." The segmentation algorithm of the e-Cognition® software utilizes a region growing approach, where a similarity criterion based on the concept of internal heterogeneity of the regions is applied. Heterogeneity results from the difference between the value of some attribute of a region and the sum of the values of that attribute over the region. If the criterion is satisfied, any two given sub-regions will be merged. Coefficients of heterogeneity were calculated for color and form attributes (Baatz et al., 2004; Baatz et al., 2000).

In this way e-Cognition® provides a multi-resolution segmentation, a new patented procedure for the extraction of image objects.

Various combinations of input parameters in the software were also tested in order to corroborate its performance with the control classes taken from the field. It was thus possible to select the best combination to identify of the map classes.

Then after all the necessary edits were performed, edits were validated to identify possible topology errors. Topological structure is based essentially on three concepts: connectivity - arcs connect to one another at nodes; defining area - arcs that connect to surround an area to define a polygon; and adjacency - arcs have direction and left and right sides.

The verification process of a classification involves several steps. The first step was performed after an automatic classification of the images by the e-Cognition® software. The validation process was carried out scene-by-scene by the GIS team using ARCGIS software and the sampling points collected in the field as a reference. This was considerably time-consuming work, during which cases of confusion resulting classification of the e-Cognition® were corrected.

Next a first validation of preliminary nature was performed on the classification, where a vegetation specialist alongside a GIS technician validated all classified scenes. The validation was carried out on an individual scene, using a 10,000 x 10,000 meter grid overlay, whose points of intersection defined sample regions for the purposed of assessing the quality of the image classification. Depending on the position and image cutlines, the number of validation samples varied from scene to scene, with a mean of seven samples. Some map scenes were verified in pairs, since they had also been classified in pairs.
Geographic coordinates were recorded for each validation point and after analysis of the classification, they given a score from the following scale:
• Score 0.0 $\Rightarrow$ Unsatisfactory – many misclassified polygons in most of the region represented the validation point;
• Score 0.5 $\Rightarrow$ Regular – a few polygons in the region represented by the validation point;
• Score 1.0 $\Rightarrow$ Satisfactory – the vast majority of polygons classified correctly with possibly a few misclassified polygons.

Also recorded were the principal classification errors at each validation point. With the results of this preliminary validation guidelines were established for the correction of the scenes that did not demonstrate a satisfactory score.

The next step was a second validation of the classification, also conducted by a vegetation specialist, this time using the Kappa index for all scenes. To this end, the validation was carried out individually by scene, using a 5000 x 5000 m grid overlay. Each polygon intercepted by a grid intersection was evaluated and list was generated of correctly and incorrectly classified points. An average of 25 to 33 polygons were evaluated in each full scene (because they were smaller than full scenes, fewer polygons were evaluated in boundary scenes).

3. RESULTS

The band combination yielding the best performance was one derived from principal component analysis of the orthorectified scene, and specifically the RGB color composition: PC1, R, G (see composition in Figure 1).

The parameter settings for the region growing segmentation algorithms should aim to obtain internally homogeneous regions and adjacent regions significantly different with respect to the characteristic in which homogeneity is being evaluated. These same conditions are valid for the results of the image segmentation algorithm in e-Cognition®. The results of first level segmentation are shown in Figure 2.
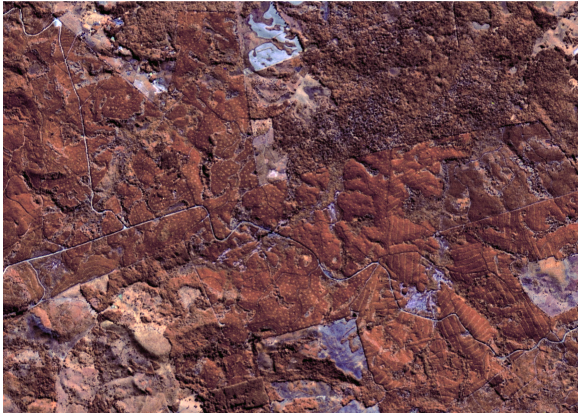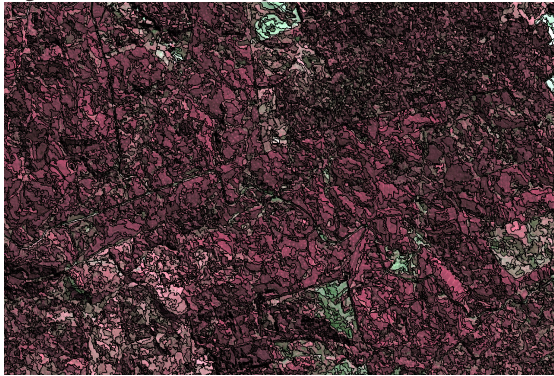
Figure 1 – RGB color composition: PC1,R,G.

Figurs 2 – First-level segmentation.

First a multi-resolution segmentation to the first hierarchical level using the criterion of reference scale value of 30 was conducted. The color homogeneity criterion was set to 0.9 and form value was set to 0.1. Next, an object classification hierarchy was built using the Definiens Professional software, which separated the scenes into 2 classes: forest and non-forest. To this end, we used the brightness information from band 3, with the criteria established for each class, resulting in Figure 4. Next, we performed a new multi-resolution segmentation with reference scale value set at 20 while all other parameters were left unchanged. Using the texture homogeneity criterion for all directions, it was possible to establish the pattern of homogeneous tree stands. To separate forest classes, the separation scheme depicted in Figure 3 was created.
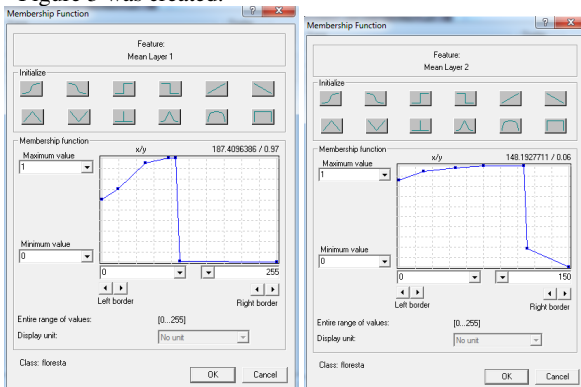

Figure 3 – Forest classification scheme.

By inverting the expression we obtained the result for the other forest classes. This first stage of classification permitted

differentiation of forest (gray) and non-forest (yellow) classes in the map.
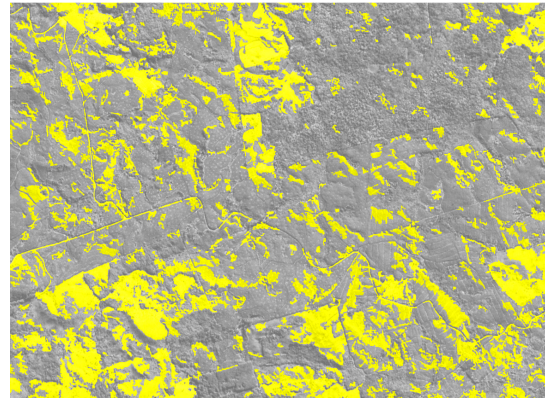

Figure 4 – First hierarchical classification level.

In the next phase a second level classification of native and planted forests was conducted. To accomplish this the criteria settings were used for the planted forest class as in Figure 5.

To obtain the native forest class, we used the inversion criterion on the planted forest class. In this way forested areas (gray in Figure 4) that do not fit the criteria of Figure 6 (planted forest) were classified as native forest.

The resulting classified image using this second step is presented in Figure 5.

The map showed that there are 385,213.67 hectares of native forests in early stages of succession (9.60% of the area covered by this study), 1,084,199.02 hectares of native forests in moderate and advanced stages of succession (27.01% of area) and 131,681.30 hectares of planted forest (see Table 1).

To evaluate the performance of the classification created for the study area, a sampling of control points was carried out. The Kappa index was calculated by vegetation class for each scene. Thus, scenes with an accuracy index lower than 85% in a certain class were submitted to a further correction.
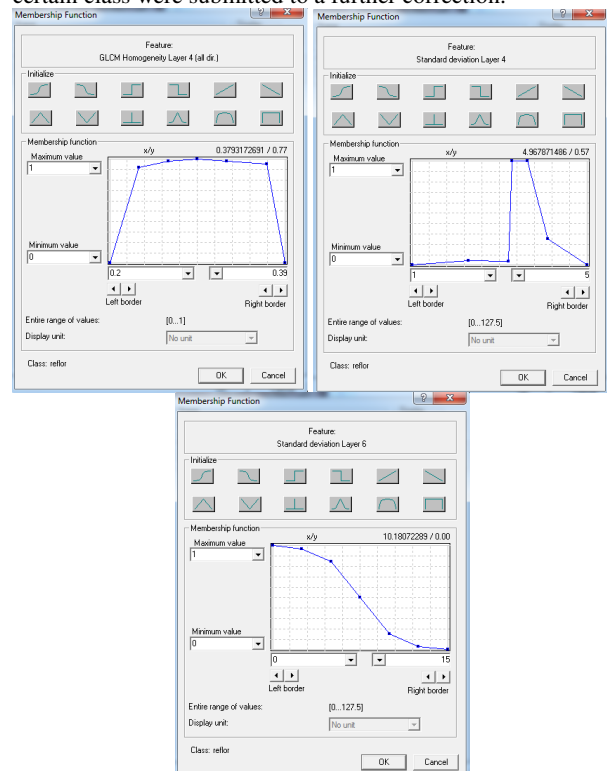
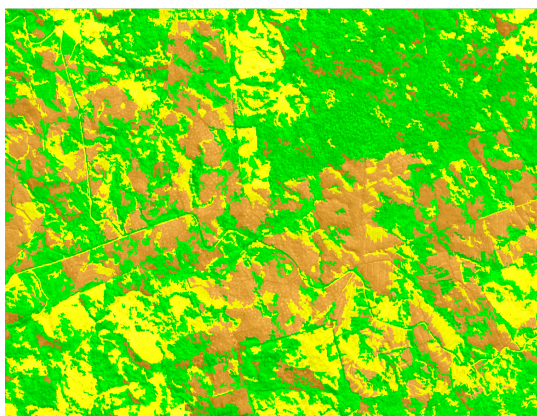Figure 5 – Classification criteria for planted forests.

Figure 6 – Second hierarchical classification level.

Table 1 – Area results for the study area.

Summary Statistics	Area (ha)	
	Native forest	Planted forests
Values for the study area	1,469,412.69	131,681.30
Percent of total area (study area)	36.61	3.28

The validation process was completed with the 3rd validation of the classification, where all scenes considered unsatisfactory in the previous validation were again checked. This final step involved analysis only of polygons belonging to the class that did not achieve 85% accuracy in the previous validation. Thus, in each scene 12 polygons were randomly selected for validation of the classification.

4. CONCLUSIONS

This work showed that 9.60% of forest areas are in an initial stage of succession in the study area, or a total of 385,213.67 hectares while 27.01% of forest areas are in a moderate and advanced stage of succession, or a total of 1,084,199.02 hectares. Planted forest comprised 3.28% of the study area (131,681.30 hectares).

The Guarapuava regional office had the greatest concentration of native forests in absolute terms, with 157,118.68 hectares in early stage and 434,017.95 hectares in moderate and advanced stages of succession.

Meanwhile, the Pato Branco regional office had the smallest concentration of native forests in absolute terms with 60,020.95 hectares of forests in the initial stage and 186,512.67 hectares of forests in moderate and advanced stages of succession.

The map validation process gave rise to an average accuracy greater than 90% in the classification.

5. REFERENCES

Baatz, M.; Benz, U.; Dehghani, S.; Heynen, M.; Höltje, A.; Hofmann, P.; Lingenfelder, I.; Mimler, M.; Sohlbach, M.; Weber, M.; Willhauck, **G. e-Cognition Professional User Guide 4**. München, Germany: Definiens Imaging GmbH, 2004. 486 p.

Campbell, James B. **Introduction to remote sensing**. New York: Guilford Press, 3 Edição, 2002.

Conama - Conselho Nacional do Meio Ambiente. Resolução nº 2/94.

FAO - Food and Agriculture Organization of the United Nations. Global Forest Assessment 2000. Main report. FAO Forestry Paper 140. Rome. Italy. 2001.

Hamlyn G. Jones; Robin A. Vaughan. Remote Sensing of Vegetation: Principles, Techniques, and Applications. New York: Oxford, 2010.

IBGE. **Manual Técnico da Vegetação Brasileira**. Rio de Janeiro: Fundação Instituto Brasileiro de Geografia e Estatística – IBGE. 92 p. 1992.

MMA – MINISTÉRIO DO MEIO AMBIENTE. Mapeamento de Cobertura Vegetal dos Biomas Brasileiros. inserido no Projeto de Conservação e Utilização Sustentável da Diversidade – PROBIO. 2007.