XIV WORLD FORESTRY CONGRESS, Durban, South Africa, 7-11 September 2015

Developments in forest monitoring under the Brazilian National Forest Inventory: multi-source and hybrid image classification approaches

Naíssa Batista da Luz[1], Yeda Maria Malheiros de Oliveira[2], Maria Augusta Doetzer Rosot[2], Marilice Cordeiro Garrastazú[2], Luziane Franciscon[2], Humberto Navarro de Mesquita Júnior[3], Joberto Veloso de Freitas[3], Clauber Rogerio da Costa[4]

[1] FAO/Embrapa Forestry; (naissa@gmail.com)

[2] Embrapa Forestry; (yeda.oliveira@; augusta.rosot@; marilice.garrastazu@; luziane.franciscon@ embrapa.br)

[3] Brazilian Forest Service; (joberto.freitas@, humberto.mesquita-junior@florestal.gov.br)

[4] Federal University of Paraná; (clauber@ufpr.br)

Abstract

Information on forest and tree resources as well as land use and land cover (LULC) maps are a growing demand which Brazilian National Forest Inventory (NFI-BR) is designed to meet through field and remote sensing surveys. Field data collection comprises biophysical variables for forest and environment condition assessment, as well as socioeconomic variables for characterization of how people living nearby forests use and perceive the forest resources. The landscape level, based on remote sensing survey and spatial analysis, focuses on variables such as forest fragmentation, changes in forest cover and land use, and the condition of forest along rivers and water bodies. Multi-temporal Landsat-8 (L-8) and RapidEye (RE) high resolution imagery and ancillary data are the sources of information for an intricate hybrid image classification approach. Object-oriented analysis coupled with pixel based multi-data classification is providing reliable information on forest, trees and LULC monitoring. Global forest cover data, Landsat-8 TOA reflectance as well as derived 32-day vegetation index composites along the year are being processed in a cloud computing environment, providing pixel-based 30m pre-classification results. These results and ancillary map information (i.e., urban areas, roads, rivers and water bodies) are included in an object-based approach based on RE 5m spatial resolution imagery to produce landscape sample units (LSU) LULC Maps. The described hybrid image classification technique takes advantage of multi-temporal Landsat-8 data, valuable ancillary information and high resolution RE data to produce good quality LULC maps for the landscape sample units of NFI-BR.

Keywords: object-based image analysis; cloud computing, RapidEye, landscape

Introduction, scope and main objectives

The Brazilian National Forest Inventory (NFI-BR) conducted by the Brazilian Forest Service (BFS) with support from Food and Agriculture Organization of the United Nations (FAO) through a GEF / FAO project, is established to allow forest resources continuous monitoring, providing information to subsidize public policy definitions for forests management and conservation.

Besides field data collected every 20 km all over the territory through Field Sample Units (FSU), NFI-BR includes also a Geospatial Component to provide information at landscape level through Landscape Sample Units (LSU) Land Use/Land Cover (LULC) mapping and spatial analysis. This component was designed to provide more accurate information on forest resources than previous official mapping initiatives at landscape level. Landscape is understood, under the scope of the NFI-BR, as a heterogeneous group of ecosystems embodied in different land use land cover types interacting with each other, including vegetation, soil, water, agriculture, pasture, urban areas, among others, composing a mosaic of LULC in which natural and anthropogenic components contribute to quality of existing forest resources. Such information can be obtained from LSU since they represent an intermediate data source between field data and existing LULC or forest cover maps.

Considering that more than 5500 LSU are being established over the Brazilian territory, the development of a consistent methodology for satellite imagery information extraction is essential, in a way that different levels of detail, aggregation and analysis can be generated to all Brazilian Biomes. With this goal, Embrapa Forestry, BFS and FAO initiated in 2013 the NFI Landscape Study as a pilot project, based on selected LSU satellite images located at Paraná State. Several digital image processing techniques are being evaluated to produce LULC maps and then employ spatial analysis to identify spatial patterns in different themes, such as forest fragmentation, conservation state, production and forest ecological sanity. Thus, this paper presents methodological procedures and preliminary results of a hybrid approach for satellite data LULC classification for those LSU in the pilot-study. The hybrid term is understood as the combination of automatic pixel-based classification and object-oriented image analysis.

The segmentation of the image allows derivation of attributes and features, favoring image interpretation and understanding (Blaschke et al., 2000). Semantic information necessary for image interpretation is mainly represented in meaningful image-objects and its mutual relations instead of individual pixels (Baatz e Schäpe 1999). Especially in high spatial resolution images it is very likely that neighbor pixels belong to the same land cover type even if spectral response is not the same, since in this kind of images a certain texture or heterogeneity is common (Blaschke et al., 2000; Blaschke e Strobl, 2001).

Although pixel-based classifications have presented several misclassification results, difficult to overcome, which leads to the research in object-based image analysis, some advantages of that first approach cannot be ignored. It provides knowledge on spectral (dis)similarities between land cover types and for some classes, improves class identification due to pixel purity instead of the mean image-object spectral response. The hybrid approach combines advantages of both methods, spectral knowledge of pixel-based classifications and improved semantic objects generation and classification due to the vast amount of features and attributes for object-based image analysis.

This paper presents advances in methodology development related to digital image processing workflow to be applied on the NFI-BR Geospatial Component LSU LULC mapping.

Methodology/approach

Study Area

As defined by the BFS, FSU were stablished on a regular 20x20km grid all over Brazil and the Landscape Sample Units (LSU) are located at 40x40km distance from each other, occupying an area of 10x10km (100 km^2) (Figure 1). The development of a methodology for Landscape Sample Units (LSU) analysis is based on a pilot project at Paraná State, comprehending 23 LSU selected from the 138 LSU defined for the whole State, according to the official NFI-BR grid (Figure 1). Selection of the 23 LSU for the pilot study was carefully done in a way that those selected would represent the variation of all phytogeographic units present in the State.

Data Source

RE and L-8 imagery were used as the basis for LULC mapping in the LSU areas. RE images were orthorectified with 3A level processing by the Brazilian Ministry of Environment through a web-based spatial data catalog. The 48 RE scenes necessary to cover the 23 LSU were acquired between April 2013 and May 2014. L-8 images were obtained from the USGS EarthExplorer site. The 28 scenes used were acquired during the same period, considering RE scenes acquisition dates for each LSU. Apart from RE and L-8 calibrated images, satellite imagery derived information and other data sources provide ancillary information for image classification, making the methodology a multi-source data approach.

Information obtained from the Google Earth Engine Data Catalog was derived from Global Forest Change (GFC) data, an initiative aimed at quantification of global forest loss and gain from 2000 to 2013 (Hansen *et al*., 2013). GFC dataset includes Tree Cover in the year 2000, Global Forest Cover Loss for the period 2000-2013 and Global Forest Cover Gain for the same period, amongst other layers of information. The data is publicly distributed by the authors with 30m spatial resolution and provide not only a general overview but also preliminary classification of forest and non-forest areas.

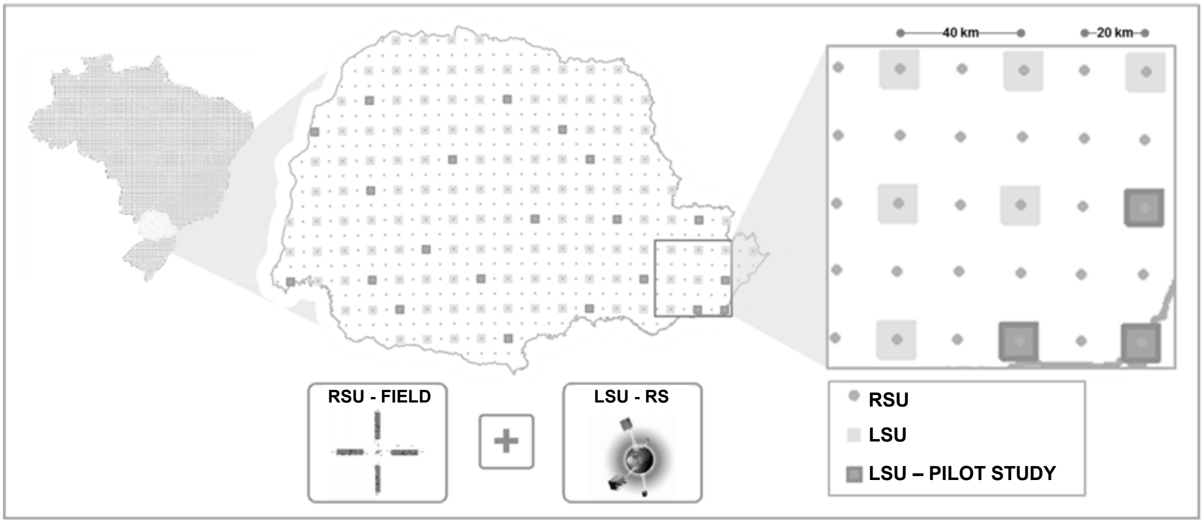


Figure 1. NFI-BR Field Sampling Units and Landscape Sampling Units

L-8 enhanced vegetation index (EVI) also derived from the Google Earth Engine Data Catalog in the form of 8-day composites were used to identify changes in land cover that could indicate agricultural and pasture land uses. Minimum, Maximum, Median and Mode EVI one year time series (April 2013 to April 2014) provide an indication of vegetation growth all year round, allowing detection of periodically covered x periodically exposed soil areas.

Ancillary data was also used to identify urban areas, in the form of a vector file, originated from a previous map.

Image Pre-processing and Pixel-based classification

After image downloading, radiometric calibration was performed, to convert digital numbers (DN) to Top-of-the-Atmosphere Reflectance (TOA Refl) using the IMPACT Tool developed by the Joint Research Centre of the European Commission (JRC). The software pre-processing tools allow automatic image decompression, layer-stacking into a single GeoTiff file, file renaming and DN to TOA conversion. Calibration parameters for Landsat images are retrieved from the image metadata. For RapidEye datasets, they come from metadata and RapidEye Product Specifications (Rapideye, 2013).

A temporal series of EVI images was analyzed using Google Earth Engine Javascript routines to produce maximum values for each pixel for the period from April 2013 to April 2014, aggregated into one image and then incorporated in the image classification process.

GFC data layers were processed using Google Earth Engine to produce an updated image of Tree Cover for the year 2013, an image containing areas where both tree cover Loss and Gain occurred during 2000-2013 period, and a Non-Forest mask for the year 2013.

RE and L-8 data were pixel-based classified using the IMPACT tool thus providing a preliminary classification to be included into OBIA. The tool uses the Phenology Based Synthesis algorithm (SIMONETTI *et al*., 2015) to automatically classify many images at one time using 16 unsupervised classes that can be further related to land use/land cover classes.

Image Segmentation and Spectral Indices

RapidEye images were segmented using a JRC developed segmentation tree that guarantees the minimum mapping unit (MMU) size of image-objects. An iterative process increases scale parameter while the MMU criteria of 95% of image-objects is not reached. After the scale parameter is adjusted, additional steps of image-objects fusion based on spectral thresholds help diminish over-segmentation of the image, reducing the number of polygons created.

After segmentation, several vegetation indexes were calculated for RE and L-8 images in eCognition software for image-objects generated during image segmentation to evaluate their potential for LULC classes discrimination (Table 1). The only index calculated on a L-8 pixel basis was Enhanced Vegetation Index (EVI) (Huete et al., 1997), available from Google Earth Engine Data Catalog in an aggregated form for every 8-days.

Table 1. Vegetation Indexes and Band Ratios used during experimentation with RapidEye and Landsat-8 imagery classification. All indexes were calculated for image-objects, instead of pixels, except for EVI, calculated for the temporal series of 8-days composite images for the period of one year (May 2013- April 2014).

Index	Formulae	Sensor	Reference
EVI	2.5*(NIR−R)/(NIR+6*R−7.5*B+1)	L-8	Huete *et al*. (1997)
Max B;G;R	MaxB+MaxG+MaxR	RE	Simonetti et al., 2014
Max G;R;NIR;Ed	MaxG+MaxR+MaxEd+MaxNIR	RE	
Max G;R;NIR	MaxG+MaxR+MaxEd	RE	
Max Ed;NIR	MaxEd + MaxNIR	RE	
Min B;G;R	MinB+MinG+MinR	RE	
Min G;R;NIR;Ed	MinG+MinR+MinEd+MinNIR	RE	
Min G;R;NIR	MinG+MinR+MinEd	RE	
NDSI	((MeanB-MeanSWIR1)/ (MeanG+MeanSWIR1))	L-8	Hall et al. 1995
NDVI Ed;R	(MeanEd-MeanR)/(MeanEd+MeanR)	RE;L-8	Modified from Rouse *et al.* (1974)
NDVI SWIR1;NIR	(MeanSWIR1-MeanNIR)/(MeanSWIR1+MeanNIR)	L-8	
NDVI SWIR1;R	(MeanSWIR1-MeanR)/(MeanSWIR1+MeanR)	L-8	
NDVI SWIR2;NIR	(MeanSWIR2-MeanR)/(MeanSWIR2+MeanR)	L-8	
NDVI SWIR2;R	(MeanSWIR2-MeanR)/(MeanSWIR2+MeanR)	L-8	
NDVI Ed;R	((MeanEd-MeanR)/ (MeanEd+MeanR))	RE	
NDVI NIR;R	(MeanNIR-MeanR)/ (MeanNIR+MeanR)	RE	Rouse *et al.* (1974)
NDVI NIR;Ed	(MeanNIR-MeanED)/ (MeanNIR+MeanED)	RE	Gitelson *et al.* (1996)
NDWI Ed;G	(MeanG-MeanED)/ (MeanG+MeanED)*10	RE	
NDWI NIR;G	((MeanG-MeanNIR)/ (MeanG+MeanNIR))*10	RE	McFeeters in 1996
NDWI G;SWIR 1	(MeanG-MeanSWIR1)/ (MeanG+MeanSWIR1)	L-8	Modified from McFeeters (1996)
NDWI G;SWIR 2	(MeanG-MeanSWIR2)/ (MeanG+MeanSWIR2)	L-8	
SAVI	((1+L)*MeanNIR−MeanR)/ ((MeanNIR+MeanR)+L)	RE	Huete (1988)
SAVI Ed;R	((1+L)*MeanEd−MeanR)/ ((MeanEd+MeanR)+L)	RE	Modified from Huete (1988)
SR Ed;R	MeanED/ MeanR	RE	
SR NIR; R	MeanNIR/ MeanED	RE;L-8	Jordan (1969)
SR SWIR 1/G	MeanSWIR1/ MeanG	L-8	Modified from Jordan (1969)
SR SWIR 1/NIR	MeanSWIR1/ MeanNIR	L-8	
SR SWIR 1/R	MeanSWIR1/ MeanR	L-8	
SR SWIR 2/G	MeanSWIR2/ MeanG	L-8	
SR SWIR 2/NIR	MeanSWIR2/ MeanNIR	L-8	
SR SWIR 2/R	MeanSWIR2/ MeanR	L-8	
SR SWIR 2/SWIR1	MeanSWIR2/ MeanSWIR1	L-8	

B = Blue Band; G = Green Band; R = Red Band; Ed= Red Edge Band; NIR = Near Infra-Red Band; Max = Maximum Object Pixel Value; Min = Minimum Object Pixel Value; SWIR = Short-Wavelength Infrared; EVI= Enhanced Vegetation Index; SR = simple ratio, NDVI = Normalized Difference Vegetation Index, NDSI = Normalized Difference Snow Index; NDWI = Normalized Difference Water Index; SAVI = Soil Adjusted Vegetation Index, L = Soil Correction Factor.

Object-Oriented Image Analysis

The segmented RE image-objects were classified using object-oriented image analysis. The hierarchical classification scheme uses multi-source ancillary data to separate image-objects into LULC classes. Several experiments have been carried out to determine the best image-object derived feature to separate and classify polygons into LULC classes.

Due to the hierarchical classification approach, that allows sequential image-objects division into LULC classes, it is possible to construct a dichotomy-like classification system that separates polygons in pairs of classes. Using samples collected by the interpreter, the Feature Space Optimization Tool calculates separability distance for selected image-object descriptors, resulting in the best separation distance optimization. The tool was developed as an aid for nearest neighbor classifier but can be very helpful in identifying those descriptors that mostly contribute for image-objects classification.

Results

As a methodology development project, the main result of this study is the definition of necessary steps and data to be used for image classification. The workflow to correctly classify image-objects in LULC classes is shown in Figure 2. The bases of the classification scheme are RapidEye images, used to generate image-objects through image segmentation. Other input data are represented by Landsat-8 imagery, Global Forest Change derived information layers such as tree cover, forest gain and forest loss during the period 2000 to 2013 and L-8 EVI 8-days composite analyzed for an one-year period (May 2013 to April 2014), used in the form of a single image containing maximum EVI for the period.

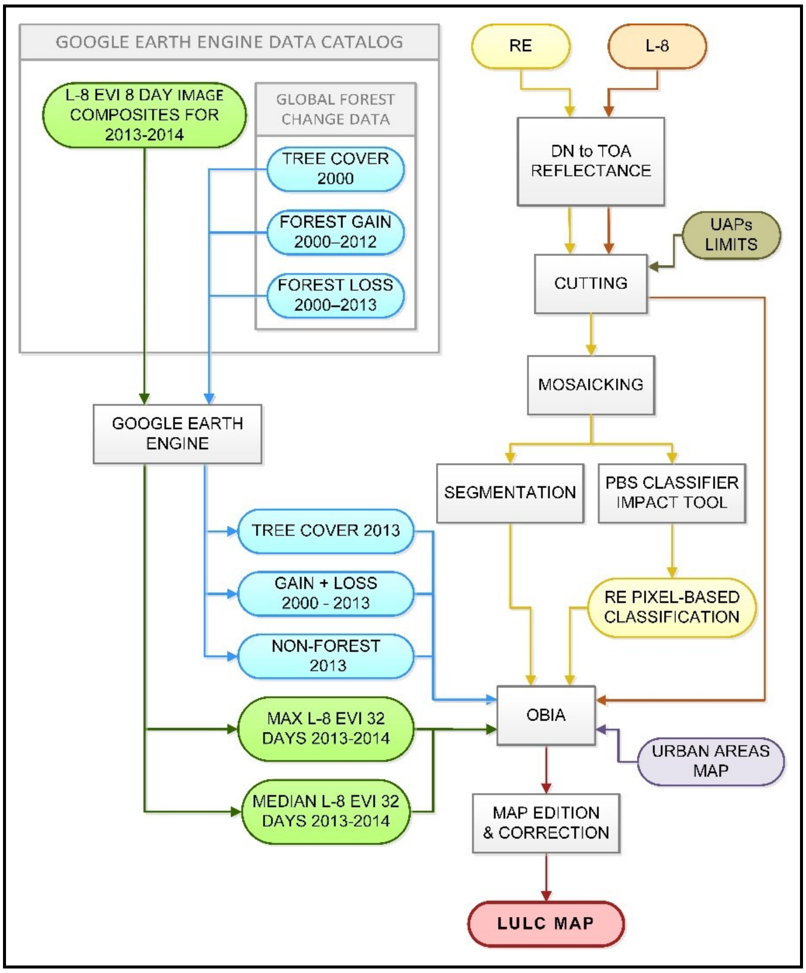


Figure 2. Workflow illustrating the defined methodology for image classification of Landscape Sample Units land use/land cover mapping.

The classification scheme adopted in object-based image analysis is illustrated in Figure 4 and classification steps used in OBIA are illustrated in a summarized form in Figure 4. The RapidEye segmentation generated image-objects that are firstly separated into tree-cover, non-tree cover and water bodies using the GFC derived layers and Normalized Difference Water Index (NDWI).

The tree cover class is further separated into planted forests and native forests using GFC derived forest gain, a combination of forest loss and gain, RE Normalized Difference Water Index calculated using Near Infra-red and Red bands and also the L-8 Normalized Difference Vegetation calculated using SWIR 2 and Red bands. Forest loss layer is used to re-classify image-objects of originally classified tree cover into exposed soil.

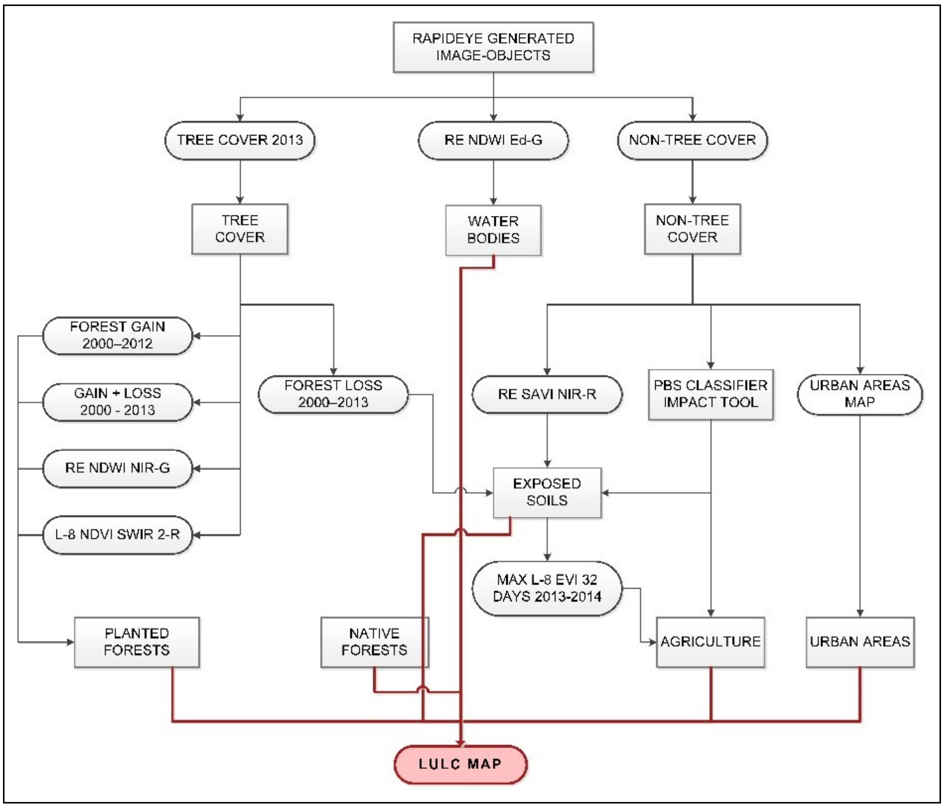


Figure 3. Classification scheme adopted in object-based image analysis of Landscape Sample Units land use/land cover mapping.

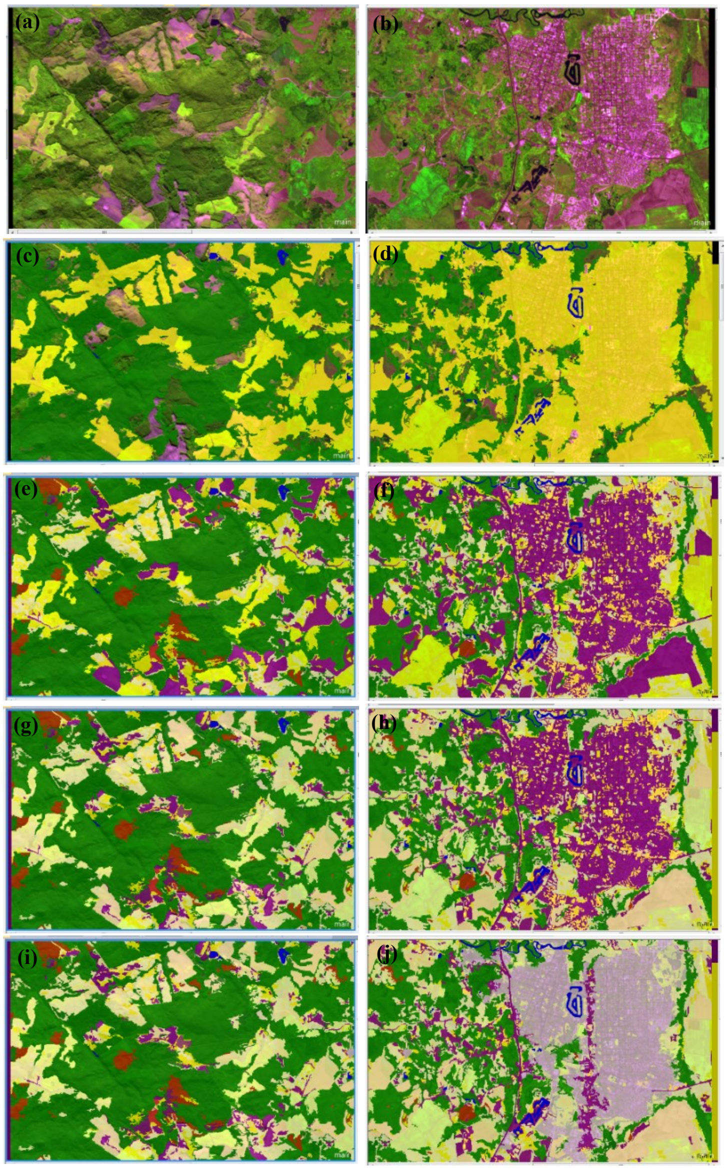
The non-forest cover initial class is further separated into exposed soil, urban areas and agriculture. Exposed soils are identified, in this stage, using Soil Advanced Vegetation Index calculated with the Near Infra-red and Red bands and also PBS pre-classified RE images. Urban areas are identified from the map and corrected using distance of exposed soil to the mapped urban area polygons. Agricultural areas are classified using PBS pre-classified RE images, at first. A second step identifies temporary exposed soils that are actually agricultural fields using the Maximum values of Enhanced Vegetation Index for one year period derived from Google Earth Engine Data Catalog.

Discussion

A hybrid approach that combines advanced algorithm for pixel-based image classification and object based-image analysis seems to be a promising methodology to produce Landscape Sample Units land use/land cover Maps for the Brazilian National Forest Inventory. It combines the advantages of each classification method, reducing spectral confusion in pixel-based classification, allowing the use of ancillary data and profiting from the object-based classification benefits. The Phenology Based Synthesis algorithm developed by JRC team (Simonetti *et al*., 2015) available on IMPACT tool provides a preliminary classification and increases image-objects separability classification at a later stage.

The great number of data obtained from different sources not only improves classification but also makes it a more stable scheme to be applied to other areas. This is especially important since the presented methodology is supposed to be used for the whole country, under very different ecological conditions (as the Brazilian biomes) and types of human land use (agricultural practices and forms of land occupation).

Data layers used as ancillary data were very helpful in labeling confused land use classes. Discrimination of land use is not always a straightforward task since, differently from land cover classes, land use is not necessarily well characterized in a single date satellite scene. The adoption of Enhanced Vegetation Index for the period of one year provided identification of areas temporarily exposed but annually cultivated, classified as agriculture. The EVI 8-day composite images are publicly available and can be easily obtained using the Google Earth Engine tool since it allows production of data for areas as big as the planet at once in a cloud based processing environment. Other very important information for classification, obtained from GEE, is Global Land Cover derived information. The initial separation of areas with trees or no trees was considered very stable, since information of forest cover, forest gain and forest loss for the period of 2000 to 2013 is available.



RE segmented image (image-objects lines are not shown for better visualization) representing different areas of the LSU. In (a) a rural area containing planted forests and in (b) an urban area.

Tree-cover (green), non-tree cover (yellow) and water bodies (blue) classification in (c) and (d). Non-classified polygons refer to loss or gain areas subtracted from tree cover for the year 2000 in Google Earth Engine.

The next step of image classification, illustrating discrimination of planted forests (red), exposed soils (purple) and agricultural areas (light yellow), in (e) and (f). Planted forests were identified using spectral indices and GLF data derived layers.

Classification of agriculture from non-tree cover and exposed soils based on Maximum Enhanced Vegetation Index calculated for L-8 8-day composites for the period of 1 year (May 2013-April 2014) in (g) and (h).

Classification of urban areas based on a previous map updated using distance of previously classified exposed soils to the map urban area polygon in (i) and (j).

Figure 4. Image classification steps illustration. RE segmented image is shown on left (without lines for better visualization) and classified image at right.

Another great advantage of this approach is the possibility of using different spatial and spectral resolution images. Using RapidEye spatial resolution as the basis for image segmentation preserved at most object-feature details for the 5m resolution. Spectral resolution of Landsat-8, in turn, provides very important information from spectral regions for objects classification, especially discrimination between forest plantations and native forests.

Conclusions

This methodology will be further enhanced to be applied to the State of Paraná, South Region of Brazil and, at a later stage, to the other Brazilian states. The development of the basic methodology is considered ready to be applied and adjusted to the other pilot study LSUs in the State of Paraná in the near future. The adoption of the so-called multi-source hybrid-approach has shown to be robust enough to produce LSU land use/land cover maps that can be used in the Brazilian National Forest Inventory Landscape Analysis.

Although it may seem a complex approach several steps are automatized. RE and L-8 calibration and pixel-based classification using PBS algorithm are available from IMPACT Toolbox, thus allowing a great number of images to be processed sequentially. The temporal analysis of Maximum EVI for one year and GFC data (tree cover, forest loss and gain) can be obtained for the entire State in a few minutes, or for the entire country under half hour. Spectral indices can be automatically generated in eCognition once a project is built and applied to other LSU image mosaics. The process trees for OBIA are being constructed using as stable as possible information for image classification, such as spectral indices (instead of image-to-image thresholds), previous maps available for the entire country, as well as time series of enhanced vegetation index and global data on forest cover.

Apart from application of this methodology for LSUs mapping, other aspects will be addressed under the Landscape Study of the NFI-BR. Trees outside forests estimation, forest fragmentation and connectivity and riparian zones evaluation will be carried out based on the LSU maps.

Acknowledgements

We would like to offer our great appreciation to our colleagues René Beuchle and Dario Simonetti from the Institute of Environment and Sustainability of the Joint Research Centre and Rastislav Rasi from the National Forest Centre – Forest Research Institute of Slovakia (NFC/FRI Slovakia) who provided their support and valuable advice when necessary, including provision of the Impact Tool software used for pixel-based classification.

We acknowledge partial funding for shared facilities used in this research from the GEF/FAO project under award number GCP/BRA/079/GFF and also the European Union - Brazil Sector Dialogues Support Facility, jointly coordinated by the Ministry of Planning, Budget and Management (MPOG) and the European Union Delegation in Brazil (Delbra).

References

RapidEye, 2013. http://www.blackbridge.com/rapideye/upload/RE_Product_Specifications_ENG.pdf

Simonetti, D., Simonetti, E., Szantoi, Z., Lupi, A., & Eva, H. D. 2015. First Results from the Phenology-Based Synthesis Classifier Using Landsat 8 Imagery. **Ieee Geoscience and Remote Sensing Letters**, Volume:PP, Issue: 99, 2015. In press…

McFeeters, S.K. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. **Int. J. Remote Sens**. 1996, 17, 1425–1432.

Hall, D.K., Foster, J.L., Chien, J.Y.L. & Riggs, G.A. 1995: Determination of actual snow-covered area using Landsat TM and digital elevation model data in Glacier National Park, Montana. - **Polar Record** 31(177): 191–198.

Gitelson, A. A.; Merzyak, M. N.; Lichtenthaler, H., 1996. Detection of red edge position and chlorophyll content by reflectance measurements near 700 nm. **Journal of Plant Physiology**.

Rouse, J. W., Haas, R. H., Schell, J. A., Deering, D. W., and Har- lan, J. C.: **Monitoring the vernal advancements and retrogradation (greenwave effect) of nature vegetation**, NASA/GSFC Final Report, Greenbelt, 1974.

Huete, A. R. et al. A Comparison of Vegetation Indices over a Global Set of TM Images for EOS-MODIS. **Remote Sens. Environ**, n.59, p.440-451, 1997.

Jordan, C.F., 1969. Derivation of leaf area index from quality of light on the floor. **Ecology** 50, 663–666.

Simonetti, E.; Simonetti, D.; Preatoni, D. **Phenology-based land cover classification using Landsat 8 time series**. Joint Res. Centre, Ispra, Italy, EUR 26841 EN, 2014.
