Flood extent monitoring of the Amazon River floodplain using ScanSAR/ALOS data

A. S. Arnesen [a, *], T. S. F. Silva [a], L. L. Hess [b], E. M. L. M. Novo [a]

[a] Remote Sensing Division, National Institute for Space Research – INPE.
P.O. box 12227-010, São José dos Campos, SP, Brazil
(allansa, thiago, evlyn)@dsr.inpe.br
[b] Earth Research Institute, University of California, 93106 Santa Barbara, USA – lola@eri.ucsb.edu

KEY WORDS: flood pulse, wetlands, multi-temporal analysis, object-based image analysis, synthetic aperture radar, Kyoto & Carbon Initiative

ABSTRACT:

The lower Amazon River floodplain is subject to large seasonal variations in water level, which associated with the flat topography, result in significant variation in flood extent throughout the year. Synthetic Aperture Radar (SAR) data offers a good choice for mapping the flooded area in these wetlands, given its ability to provide timely and continuous information without being strongly affected by cloud cover and atmospheric conditions. As part of JAXA's Kyoto & Carbon Initiative, wide-swath, multi-temporal coverage of the Amazon basin has been obtained using the ScanSAR mode of ALOS PALSAR. One of the largest limitations of radar automated classification is the occurrence of speckle noise. Furthermore, the dynamic nature of the floodplain environment demands the use of advanced methods, capable of integrating multiple sources and scales of information. This study tests the applicability of object-based image analysis for monitoring flood extent changes as a function of river stage height, using ALOS ScanSAR images for the Curuai Lake floodplain. This study area is located at the lower Amazon River near the city of Óbidos (Pará State, Brazil). Seven ScanSAR scenes were acquired during the 2007 flood pulse. Water level records from two gauge stations (Curuai and Óbidos), field photographs collected during the rising water period of 2011 and optical images (Landsat-5/TM and MODIS/Terra and Aqua) were also used. A data mining algorithm allowed the identification of thresholds, later used to implement a hierarchical object-based classification algorithm to map the flooding status in the study area for all available dates. The accuracy of the classification was assessed for the first three hierarchical classification levels, as well as for flooding status. Levels 1 and 2 (one land cover map for the entire time series) had overall accuracies of 90% and 83%, respectively. Level 3 classifications (one map per date) were validated only for the lowest and highest water stages, with overall accuracies of 78% and 80%, respectively. Flooding status was mapped with 88% and 90% accuracies for the low and high water stages, respectively.

* Corresponding author.

1. INTRODUCTION

Flood extent monitoring in the Amazon River floodplain has great importance for the understanding and conserving this ecosystem. The floodplain fulfils several important roles, such as natural control of extreme events, habitat provision to flora and fauna species and as source of environmental services and resources to riverside communities (Mitsch & Gosselink, 2000). The Amazon wetlands also have an important role in the global carbon budget, both as a source of greenhouse gases (CO_2 and CH_4) and as a carbon sink (Melack et al., 2004; Richey et al., 2002).

Large seasonal variations in the spatial distribution of flooding in the Amazon floodplain are caused by the high amplitude of the annual flood wave. The periodicity and amplitude of this wave was introduced by Junk et al. (1989) for the Amazon system as the flood pulse concept.

Access, cost and logistics limit the use of in situ techniques for monitoring flood extent in the wetlands of large rivers. Remote sensing data, especially Synthetic Aperture Radar (SAR) data, represents an alternative due to its ability to provide timely and continuous information.

The wide swath of the ScanSAR mode from the PALSAR/ALOS sensor (L-band SAR) allows frequent image acquisition at several phases of the flooding season. This mode, however, has limitations such as speckle, coarse spatial resolution and a significant effect of the wide range of incidence angles on the recorded SAR backscatter (Ardila et al., 2010). Thus, object-based image analysis (OBIA) can be an alternative for an accurate monitoring of the flood extent, allowing the use of attributes other than per-pixel backscatter, and of complementary data such as optical images and topographic maps in the classification process.

Therefore, this paper reports an OBIA based methodology to monitor the flood extent in the Amazon River floodplain, using PALSAR ScanSAR data acquired as part of JAXA's Kyoto & Carbon Initiative during the 2007 flood pulse in the Curuai Lake floodplain.

2. METHODS

2.1 Study area and input data

The study area comprises the Curuai Lake floodplain, along the Lower Amazon River (Figure 1). Water level variation in the lake shows seasonal amplitude of ~ 6 meters. This amplitude, associated with its flat topography, results in significant variation in the flooding extent throughout the year.

Image data was acquired by the PALSAR sensor, aboard the ALOS satellite. Seven scenes of the 409 orbit, ScanSAR mode were acquired at 1.5 processing level, in amplitude (digital numbers), with a pixel spacing of 100 meters. Each image date represents a different water level condition.

Optical data was used to minimize classification confusion between land cover types. Landsat-5/TM and MODIS surface reflectance (Terra and Aqua) images acquired at dates close to ScanSAR acquisition and at similar water levels (Figure 2) were used in the OBIA process. Also, the SRTM (Shuttle Radar Topographic Mission) elevation model was used to exclude upland areas from the analysis.

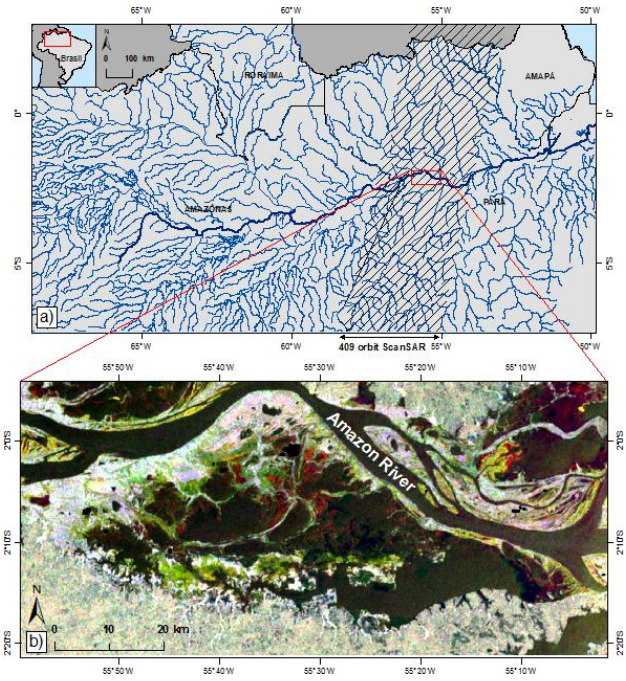

Figure 1. Curuai floodplain. a) Hatched area represents the 409 ScanSAR/ALOS orbit; b) Filtered ScanSAR composite image (R:30/11/2006 - G:15/01/2007- B:02/06/2007).
© JAXA/METI 2007.

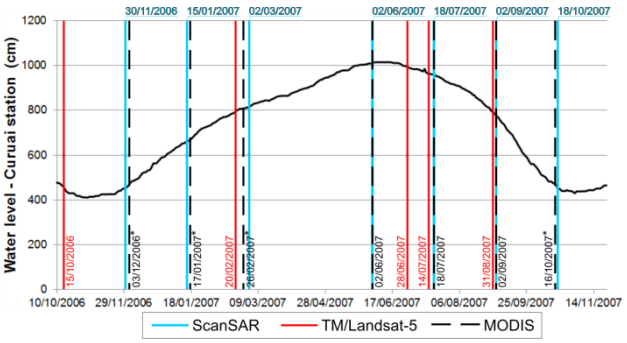

Figure 2. Image acquisition dates along the 2007 flood pulse.
* MODIS 8 days surface reflectance data.

The methodology of this study was structured in five main steps (Figure 3).

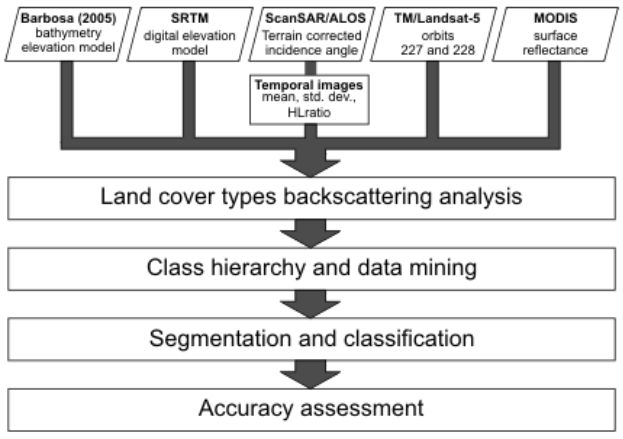

Figure 3. Steps of the OBIA analysis.

2.2 Backscattering analysis of land cover types

Backscattering analysis was performed to identify possible overlaps between class distributions that could result in confusion during the classification process. Eight land cover types were sampled in the ScanSAR images, according to their backscattering signals: Rough and Smooth Open Water (ROW and SOW); Flooded and Non-flooded Forest (FF and NFF); Dry/Smooth and Wet/Rough soil (DS and WS); and Emergent and Floating Macrophytes (EM and FM). Landsat-5/TM images were used as reference for sample collection of these classes (> 500 pixels). A field campaign during the rising water period in 2011 also helped class identification.

The backscattering characteristics were examined in terms of amplitude, and later converted to normalized backscattering coefficients (σ^0) to allow for comparison with previous SAR studies. Backscattering samples were analysed using descriptive statistics and boxplots showing class overlap and temporal variation.

2.3 Class hierarchy and data mining

The classification of the land cover types of the floodplain was necessary as a preliminary step for the flood mapping, since the radar response to the flooding status varies between classes. The eight classes analysed were organized into four hierarchical levels, as a function of their backscattering and temporal variation characteristics.

Similarly to Silva et al. (2010) study, the first classification level identified three classes, defined by their annual flooding pattern: Upland (non-floodable areas), Floodplain (variable flooding) and Permanent Open Water (open water surface during the lowest water level in 2007). At level 2, the Floodplain class was split into two main classes, assumed as spatially constant along the year: Forest and Non-forest. The third level segregated Forest areas into two classes, according to flooding status (Flooded -'FF' - and Non-Flooded forest – 'NFF'), and also segregated Non-Forest areas into two intermediary groups according to radiometric similarity: i) 'Bright' (including Wet/Rough soil – 'WS', Emergent macrophyte - 'EM' - and Rough open water – 'ROW'), and ii) 'Dark' (with Dry/Smooth soil – 'DS', Floating macrophyte - 'FM' and Smooth open water – 'SOW'). Discrimination of each individual class within these radiometric groups was based on optical data, given their very similar backscattering. Finally, Level 3 classes were merged into a fourth level, according to their flooded condition ('Flooding Status', Figure 4).

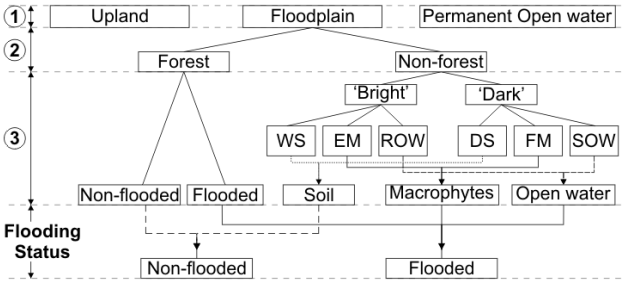

Figure 4. Hierarchical classification scheme for classifying land cover types and flooding status at the Curuai Lake floodplain, lower, Amazon River, Brazil.

The WEKA data mining tool (Witten & Frank, 2005) was used to identify backscattering patterns in the defined classes, as well as extract thresholds that could be later used as classification rules. The WEKA J4.8 algorithm was used to define the classification rules based on the input SAR data class attributes provided by the samples. When class overlaps were not resolved using SAR data (water and soil classes, for instance), MODIS infrared reflectance was used to further separate the classes. The optical classification thresholds were extracted through visual feature inspection in eCognition 8.

2.4 Segmentation and classification

Image segmentation was performed using the multi-resolution algorithm implemented in eCognition 8, which allows the integration of images of multiple sources and resolutions into the object-generation process (Definiens, 2008). Temporal images were generated of the same backscattering attributes selected for the data mining approach (Table 1). They were used at the first three classification levels and had specific objectives. Beyond the temporal images two other data were used at Level 1: SRTM because of its importance on segregating Upland than Floodplain; and infrared band of Landsat-5/TM (B5) to segregate the Permanent Open Water.

Image / attribute	Description	Classification Hierarchy
TAB	Temporal Average Backscattering	Levels 1, 2 and 3
TSD	Temporal Standard Deviation backscattering	Levels 1, 2 and 3
HWL	Highest Water Level backscattering	Levels 1 and 3
LWL	Lowest Water Level backscattering	Levels 1 and 3
HLR	Highest/Lowest water level backscattering Ratio	Level 1
SRTM	Digital Elevation Model	Level 1
B5/TM	Landsat-5/TM Infrared band of the low water stage (15/10/2006)	Level 1

Table 1. Images/attributes used at the segmentation process of the classification levels.

A hierarchical, object-based classification was applied at eCognition 8 not only the ScanSAR data but also optical classification thresholds to deal with the high variability of radar backscatter values in the Amazon Floodplain.

2.5 Accuracy assessment

Accuracy assessment was performed for all hierarchical levels using different approaches. For Levels 1 and 2, a set of 200 points (50 per L2 class) was randomly generated. For the third classification level sample size was increased to 500 points, which were similarly interpreted, and with the final assignment of two water stages, corresponding to low (30/11/2006, 452 cm) and high (18/07/2007, 959 cm) water level. Validation samples were visually interpreted based on field and high resolution optical data, in addition to the SAR and TM images.

Since no field data was available to validate the flooding status of forest areas during the studied period, a validation sample was created by simulating flood extent using the merged SRTM/bathymetry DEM produced by Barbosa (2005). An algorithm based on elevation thresholds was applied to generate an approximation of the flood extent based on the available Curuai water stage height data.

3. RESULTS

3.1 Land cover types backscatter

Important temporal aspects were observed in the ScanSAR data. Although the main confusion was observed between soil (Figure 5) and open water classes, as expected for L-band SAR (Hess et al., 2003), soil cover types had a wider temporal variation throughout the year (Figures 6 and 7). This occurs in response to changes in moisture content and the transition from bare to herbaceous vegetation cover (Balenzano et al., 2011), while water surface roughness is mostly affected by wind, resulting in radiometric variations of shorter duration that averaged to similar backscattering values over time.

The results also supported the grouping of land cover types during the classification, since the Non-forest classes (macrophytes, soil and open water) presented bi-modal backscattering distributions. Therefore, Non-forest classes could be divided into two main classes according to backscatter values: 'Bright' (-10 dB to -17.5 dB, EM, WS and ROW) and 'Dark' (-19.5 to -24 dB DS, FM and SOW).

Despite the EM class present the mean value inside the 'Bright' group (i.e. -13.4 dB), it presented the highest radiometric variability, with values distributed along the three groups defined. This behaviour had to be considered in the classification rules definition.

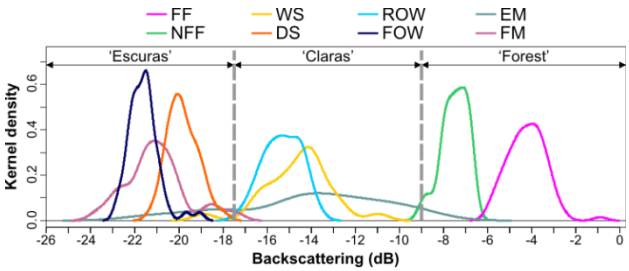

Figure 5. Backscattering kernel density of the eight classes sampled.

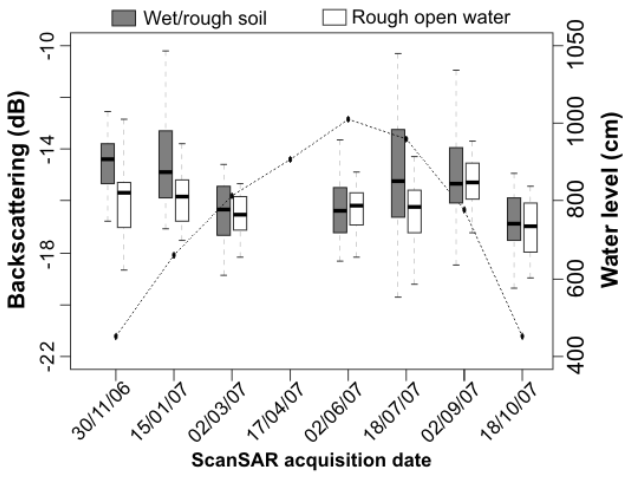

Figure 6. Temporal variation of the classes Wet/rough soil (WS) and Rough open water (ROW).

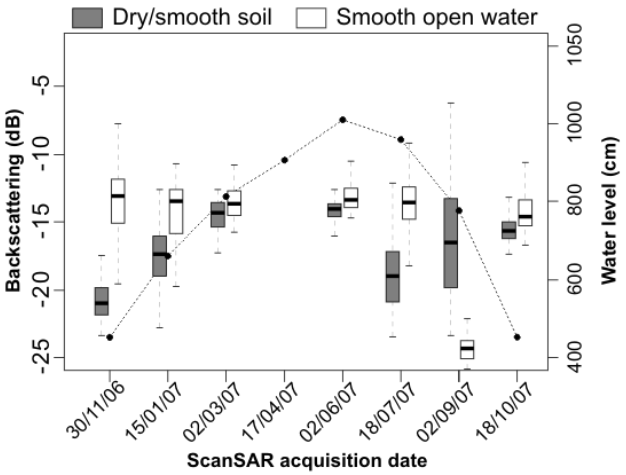

Figure 7. Temporal variation of the classes Dry/smooth soil (DS) and Smooth open water (SOW).

3.2 Object-based classification

Images used for segmentation had different importance on the object generation results. The TAB image allowed the identification of areas with similar backscattering along the entire flood pulse, while the TSD revealed those with larger radiometric variation along the year, which have higher spatial-temporal dynamic. The LWL, HLW and HLR images were important in the identification of classes' objects that present very specific backscattering behaviour according to the water level.

The attributes explored from the classes allowed the definition of the classification rules that segregated the land cover types using SAR, optical and topographic data (Figure 8, 9 and 10).

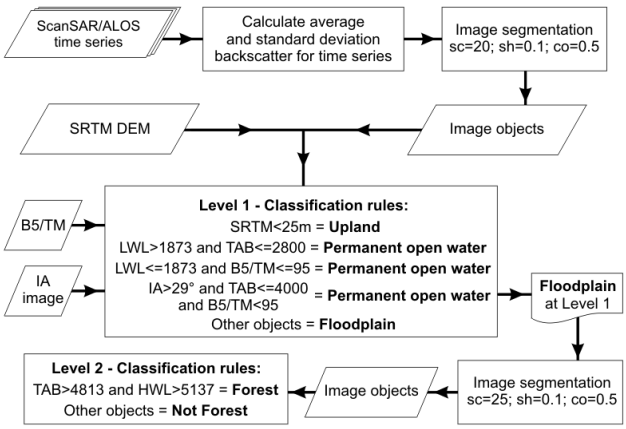

Figure 8. Object-based classification rules for levels 1 and 2. sc=scale; sh=shape; co=compactness; LWL=Lowest Water Level backscattering; TAB=Temporal Average Backscattering; B5/TM= Landsat-5/TM band 5 acquired at low water stage; IA=Incidence angle.

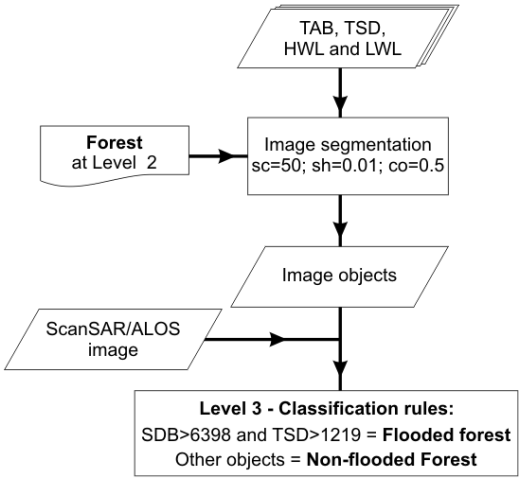

Figure 9. Object-based classification algorithm for Forest class at level 3. sc=scale; sh=shape; co=compactness; HWL=Highest Water Level backscattering; LWL=Low Water Level backscattering; TAB= Temporal Average Backscatter; TSD= Temporal Standard Deviation backscattering; SDB= Single Date Backscattering.

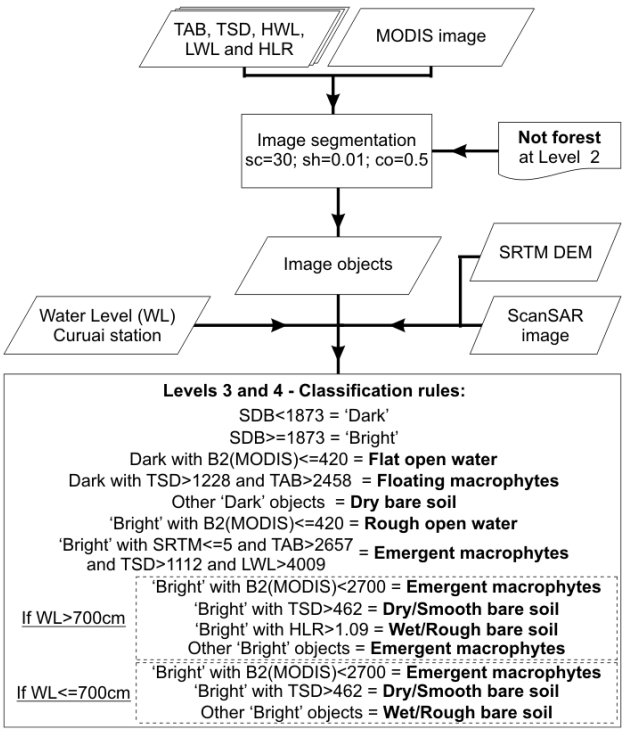

Figure 10. Object-based classification algorithm for Forest class at level 3. sc=scale; sh=shape; co=compactness; HWL=High water stage backscatter; LWL=Low water stage backscatter; TAB= Temporal Average Backscatter; TSD= Temporal Standard Deviation backscatter; SDB= Single Date Backscatter. B5/TM= Landsat-5/TM band 5 acquired at low water stage; IA=Incidence angle; and WL=water level at Curuai station.

Level 1 classification had the best overall accuracy (91%) and kappa index value (0.86). The overall accuracy and kappa index of Level 2 were reduced to 83% and 0.77, respectively. At this level the highest classification errors were observed for Forest and Non-Forest. Level 3 low water stage had an overall accuracy of 77%, while the high water stage Level 3 overall accuracy was slightly lower (76%). The decrease in the accuracy throughout the classification hierarchy is explained both by error propagation and by the increasing similarity between classes at the higher classification levels (Walker et al., 2010).

The final merging of classes into a binary flooding status map increased the accuracy of Level 3 classification, as a large portion of the confusion occurred among classes with the same status. The overall accuracies of the flood maps for low and high water stages were, respectively, 88% and 90% (Figure 13).

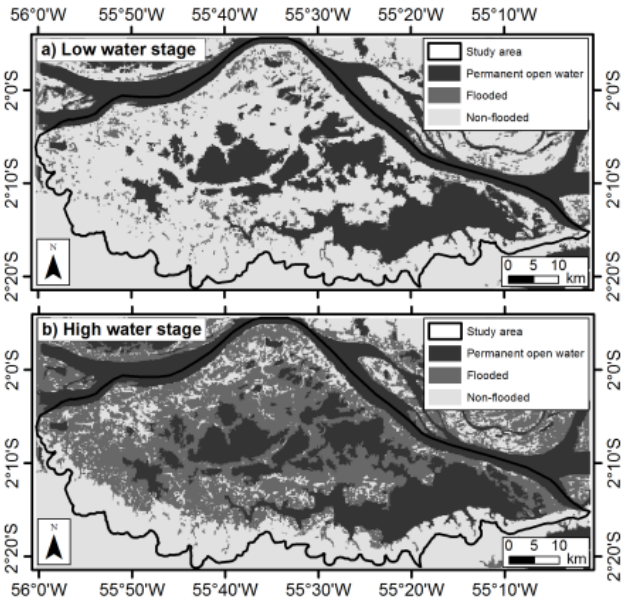

Figure 11. Mapped flooded area for the Curuai Lake floodplain, Lower Amazon River, for two dates: 30/11/2006 representing the low water stage (a) and 18/07/2007 representing the high water stage (b). The black line represents the polygon considered for area calculations.

4. CONCLUSION

The use of OBIA allowed the integration of optical and SAR data, and the exploration of the spatial and temporal variation of PALSAR ScanSAR backscattering observed for the Amazon floodplain, and supported the monitoring of flooding extent during 2007, while reducing the effect of image speckle and incidence angle variability on the SAR data classification.

The major restriction of the algorithm is the requirement of optical data to improve classification accuracy for open water and soil areas. A viable alternative would be the use of SAR data acquired at shorter wavelengths, which more sensitive to small roughness variations in smooth targets (Aubert et al., 2011).

A remarkable advantage of the present approach is the short processing time of the object-based algorithm, yielding accurate flood extent maps without the need for laborious and time-consuming manual editing.

5. ACKNOWLEDGEMENTS

The authors wish to acknowledge the Brazilian National Council for Scientific and Technological Development (CNPq) for the fellowship supporting Allan Arnesen through his graduate program (Process 130519/2010-3), and the São Paulo Research Foundation (FAPESP) for the financial support during the field campaign of April 2011 (grant 2008/07537-1) and Dr. Silva's post-doctoral grant (grant 2010/11269-2). The authors also thank Dr. Claudio Barbosa for providing the bathymetry elevation model. This work has been undertaken within the framework of JAXA's Kyoto & Carbon Initiative, with ALOS PALSAR data provided by JAXA EORC.

6. REFERENCE

Ardila, J. P., Tolpekin, V., Bijker, W., 2010. Angular Backscatter Variation in L-Band ALOS ScanSAR Images of Tropical Forest Areas. *IEEE Geoscience and Remote Sensing Letters*, 7(4), pp.821-825.

Balenzano, A., Mattia, F., Satalino, G., Davidson, M.W.J., 2011. Dense Temporal Series of C- and L-band SAR Data for Soil Moisture Retrieval Over Agricultural Crops. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 4(2), pp.439-450.

Definiens. 2008. Definiens developer 7: user guide. The Imaging Intelligence Company, Munich. 506p.

Junk, W.J., Bayley, P.B., Sparks, R. E., 1989. The flood pulse concept in the river-floodplain systems. *Canadian Special Publication in Fisheries and Aquatic Sciences*, 106, pp.110-127.

Melack, J.M., Hess, L.L., Gastil, M., Forsberg, B.R., Hamilton, S.K., Lima, I.B.T., Novo, E.M.L.M., 2004. Regionalization of methane emissions in the Amazon Basin with microwave remote sensing. *Global Change Biology*, 10(5), pp.530-544.

Mitsch, W. J., Gosselink, J. G., 2000. The value of wetlands: importance of scale and landscape setting. *Ecological Economics*, 35(200), pp. 25-33.

Richey, J.E., Melack, J.M., Aufdenkampe, A.K., Ballester, V.M., Hess, L.L., 2002. Outgassing from Amazonian rivers and wetlands as a large tropical source of atmospheric CO2. *Nature*, 416(6881), pp.617-20.

Silva, T.S.F., Costa, M.P.F., Melack, J.M., 2010. Spatial and temporal variability of macrophyte cover and productivity in the eastern Amazon floodplain: A remote sensing approach. *Remote Sensing of Environment*, 114(9), pp.1998-2010.

Walker, W.S., Stickler, C.M., Kellndorfer, J.M., Kirsch, K.M., Nepstad, D.C., 2010. Large-Area Classification and Mapping of Forest and Land Cover in the Brazilian Amazon : A Comparative Analysis of ALOS / PALSAR and Landsat Data Sources. *IEEE Journal of selected topics in applied Earth observations and Remote Sensing*, 3(4), pp.594-604.

Witten, I.H., Frank, E., 2005. *Data mining:* practical machine learning tools and techniques. Elsevier, San Francisco. 524p.