

# An approach to classify polarimetric P-band SAR images for land use and land cover mapping in the Brazilian Amazonia

Luciana de Souza Soler  
Environmental Sciences Department  
Wageningen University - WUR  
Wageningen, The Netherlands  
lsoler@wur.nl

Sidnei João Siqueira Sant'Anna,  
Corina da Costa Freitas, Luciano Vieira Dutra,  
João Roberto dos Santos  
National Institute for Space Research - INPE  
São José dos Campos, Brazil  
{sidnei,corina,dutra}@dpi.inpe.br  
jroberto@ltid.inpe.br

**Abstract**— In this paper the potentiality of polarimetric P-band SAR data for Amazon tropical forest land cover mapping is assessed. The classifying approach is based on the Iterative Conditional Mode (ICM) algorithm, taking into account several specific distributions to SAR data. Distinct land cover classes are modeled considering different distributions. The results show that the P-band data is not capable to discriminate the nine classes initially used. However this capability improves significantly when classes having similar vegetation structure are grouped. The HV image is effective in differentiating primary and very old regeneration forest areas from other land cover classes, while VV image increases the classification of bare soil and crop/pasture areas. The results show the importance of polarimetric information for the classification of several land use classes.

**Keywords:** *polarimetry; P-band; ICM algorithm; land use; Amazon*

## I. INTRODUCTION

Different forest physiognomies of Brazilian Amazonia occupy about 76% of the 5 million km<sup>2</sup> area. Forest conversion into pasture or agriculture presented a rate of 18,793 km<sup>2</sup>/year in 2005 [1]. Recent studies have shown that although the old deforested areas present land use intensification, the recent ones present a higher tax of abandonment and forest regeneration. This fact can be an indication of land use intensification, but also the opening of new frontiers of deforestation [2]. Adequate mapping with the help of remote sensing data and techniques, are important to decision makers on environmental issues to establish appropriate inventory and control plans for the future. Furthermore, such datasets and maps are also used as a support for the estimation of carbon emission/re-absorption in global climate analysis resulting from large scale changes of land use/land cover [3].

Optical remote sensing used in official deforestation assessments in Brazil is systematically affected by clouds in some parts of the Brazilian Amazonia. Synthetic Aperture Radar (SAR) is able to detect targets response through clouds, thus it has been tested in land use/land cover change, in forest biophysical parameters estimation and in deforestation mapping [4]. Low frequency SAR's (L and P bands) are

especially applicable to tropical regions due to its penetration capability on forest covered areas. L-band data have been drawing most attention from the user community because its availability either on orbital or on airborne platforms. Due to restrictions on frequency allocation P-band is still only available on airborne platforms, but airborne data have presented interesting and significant results in forest areas [3]. However, there are still many issues in the P-band imagery usage not well known and understood, mainly its interaction with different vegetation types and complex statistical behavior when compared to optical data.

Some significant results using AIRSAR multifrequency P-band data for land cover classification by using correlation modeled by Beta distributions are presented in [5]. In this study, we further investigated these results considering several possible distributions from single to full polarimetric data. In addition, the use of P-band full polarimetric data acquired in a well-known and sampled area at the National Forest of Tapajós, where different stages of regeneration and other land use occur, provides a unique opportunity of analysis. Therefore, the main goal of the present work is to analyze the potentiality of full polarimetric P-band data in distinguishing different stages of secondary succession, as well as different types of land use in the Brazilian tropical environment.

## II. STUDY AREA AND DATA DESCRIPTION

The study area includes part of the National Forest of Tapajós (Pará State) and is located within the coordinates 54°49'36'' to 55°01'45''N and 02°56'38'' to 03°23'38''S. The radar image used in this work covers an area of approximately 27 km<sup>2</sup> and is located along of the BR-167 Cuiabá-Santarém highway. The area is highly covered by dense forest, but secondary succession areas at several stages, agriculture and cattle raising are also found. Eight classes of land use were defined, based on field campaigns from 1989, 1999 and 2000, together with a database of 10 Landsat/TM images acquired along the same period. The classes defined were primary forest (PF), very old regeneration (VO - more than 25 years-old), old regeneration (OR - between 12 and 25 years-old), intermediate regeneration (IR - between 6 and 12 years-old),

new regeneration (NR - newer than 6 years-old), crops/pasture (CP), bare soil (BS) and floodplain (FP). Floodplain was classified by a simple threshold procedure. The P-band SAR images were taken on September 2000 during a mission in National Tapajós Forest using an airborne radar system developed in AeroSensing Radarsysteme GmbH. They were set in slant range, with a pixel spacing of 1.50 m x 0.67 m, in one look to all polarimetric bands: HH, HV, VH and VV. These images were used to form a 3x3 multilook covariance matrix composed by averaging 2x5 (range x azimuth) pixels.

### III. CLASSIFICATION PROCEDURE

The Iterative Conditional Mode (ICM) classification algorithm was used for the P-band imagery classification. The ICM classification method is a contextual procedure that classifies every pixel using both the observed value in the corresponding co-ordinate and the classification of the surrounding sites. A Markovian model known in the literature as Potts-Strauss is used for the classes. The classifications were performed in a system, called Polarimetric Classifier 7, developed at Image Processing Division (DPI-INPE) in IDL language. This system is available on the website <http://www.dpi.inpe.br/cursos/ser410/laboratorio.php>. It allows classifying individual or a combination of amplitude/intensity images, and full polarimetric images by modeling the covariance matrix for areas with no texture, fine texture and coarse texture [6]. These models lead, respectively, to the following distributions for the covariance matrix: Wishart [7], Multivariate  $K$  [8] and Multivariate  $G^0$  [9]. The ICM algorithm implementation in this system uses the Maximum Likelihood (ML) classification as the initial classification. More details about the system implementation can be found in the above mentioned website.

The system allows the user to choose among several distributions the one that is best adjusted for the data for each class. Based on training samples, the system performs a  $\chi^2$  goodness of fit test for each class indicating which distribution should be used. Concerning the full polarimetric classification, the multivariate distributions used to fit the covariance matrix are the Wishart, Multivariate  $K$  and Multivariate  $G^0$ , as mentioned before. For the individual amplitude images, the Square Root of Gamma,  $K$ -Amplitude,  $G_A^0$ -Amplitude, Log-Normal, and Normal distributions are available. For the intensities images the multivariate possible distributions are Gamma,  $K_I$  and  $G_I^0$ .

Both classifications (ML and ICM) are supervised and require the specification of training sets for parameter estimation. These sets were informed through regions of interest, previously defined with fieldwork information and Landsat/TM data analysis. The equivalent number of looks ( $n$ ) is also an input parameter, and it was estimated only once for the whole image by taking samples from homogeneous areas in each amplitude image (HH, VV and HV). A  $\chi^2$  goodness of fit test for the Square Root of Gamma distribution is performed and the individuals  $n$  were estimated from the sample moments for those areas that had a proper fit. The final estimate of  $n$  is given by the mean of these individuals  $n$ .

The 3x3 multi-look covariance matrix was built by averaging 2x5 (range x azimuth) pixels, making the pixel approximately square. Three elements of this matrix are intensity data and the remaining are complex data. We also built the three multi-look amplitude images (HH, VV and HV), and a biomass index image (BMI). The amplitude images were formed by taking the square root of the multi-look intensities images. The biomass index (BMI) was defined as the average of the HH and VV amplitude images. BMI for P-band is an indicator of the relative amount of woody compared to leafy biomass [10].

Nine ML/ICM classifications were carried out with the eight classes defined in study area section. The classifications were performed for the following data: individual amplitude images HH, VV, and HV; pair of intensities HH-VV, HH-HV and HV-VV; set of three intensities HH-HV-VV; the covariance matrix (full polarimetric classification) and the BMI image. Based on the analysis of these nine classifications, the land use classes were merged considering the misclassification among them indicated on test samples. After defining two sets of five and four classes, we executed the same classifications listed before with these new sets. It is important to note that the test samples had approximately the same size for all classes. We computed the Kappa coefficient of agreement and their corresponding variances for all classifications. These values were computed based on confusion matrices using test samples. The final step included the analysis of all the classification results with the help of two-sided statistical tests of equality of Kappa for all pair of classifications.

### IV. RESULTS AND DISCUSSION

The Kappa coefficients of agreement and their respective variance for all classifications are shown in Table I. The best classification results are highlighted in this table. As a reference, Fig. 1a shows a piece of a geo-referenced Landsat-7/TM image from August, 2000. Fig. 1b shows the color composition of intensities HH, HV and VV, showing the training and test samples for all defined classes. The best classification using eight classes was obtained with the BMI image. This classification is depicted in Fig. 1c. From the confusion matrices of classified images, not shown here, it was observed significant confusion among primary forest, old and very old regeneration, between new and intermediate regeneration and also between crops/pasture and bare soil. The results with eight classes were relatively poor, indicating that P-band might not be suitable for differentiating among many regeneration classes.

When the number of classes is reduced to five classes (by merging primary forest with very old regeneration classes, new regeneration with intermediate regeneration, and bare soil with crops/pasture) the classification accuracy increases. Based on the Kappa coefficient of agreement, the best classification results were achieved for Amplitude HV and Bivariate Intensities HH-HV images, which can be considered statistically equal at the 95% confidence level. Therefore, it leads to the conclusion that HH may not bring any new information in addition to that already given by HV image. The Amplitude HV classification is shown in Fig. 1d.

TABLE I. KAPPA COEFFICIENT OF AGREEMENT AND THEIR VARIANCE VALUES FOR ALL OF THE CLASSIFICATIONS PERFORMED. THE BOLD VALUES INDICATE THE HIGHER VALUES FOR EACH SET OF CLASSES.

Classifications	Number of classes					
	Eight Classes		Five Classes		Four Classes	
	$\hat{\kappa}$	$\sigma_{\hat{\kappa}}^2 \cdot (10^{-5})$	$\hat{\kappa}$	$\sigma_{\hat{\kappa}}^2 \cdot (10^{-5})$	$\hat{\kappa}$	$\sigma_{\hat{\kappa}}^2 \cdot (10^{-5})$
Amplitude HH	0.3356	1.27	0.5893	2.62	0.6626	3.48
Amplitude VV	0.4230	1.37	0.6319	2.42	0.7208	3.03
Amplitude HV	0.3292	1.18	<b>0.6675</b>	2.32	0.7390	2.89
Bivariate Intensities HH-VV	0.3566	1.46	0.6207	2.43	0.7278	2.92
Bivariate Intensities HH-HV	0.3699	1.48	<b>0.6657</b>	2.55	0.7528	2.98
Bivariate Intensities HV-VV	0.4438	1.52	0.6201	2.46	<b>0.7681</b>	2.62
Multivariate Intensities HH-HV-VV	0.4148	1.38	0.6539	2.33	<b>0.7599</b>	2.68
Full polarimetric BMI	0.3754	1.43	0.4286	2.76	0.7111	3.12
BMI	<b>0.4959</b>	1.33	0.6421	2.31	0.7345	2.76

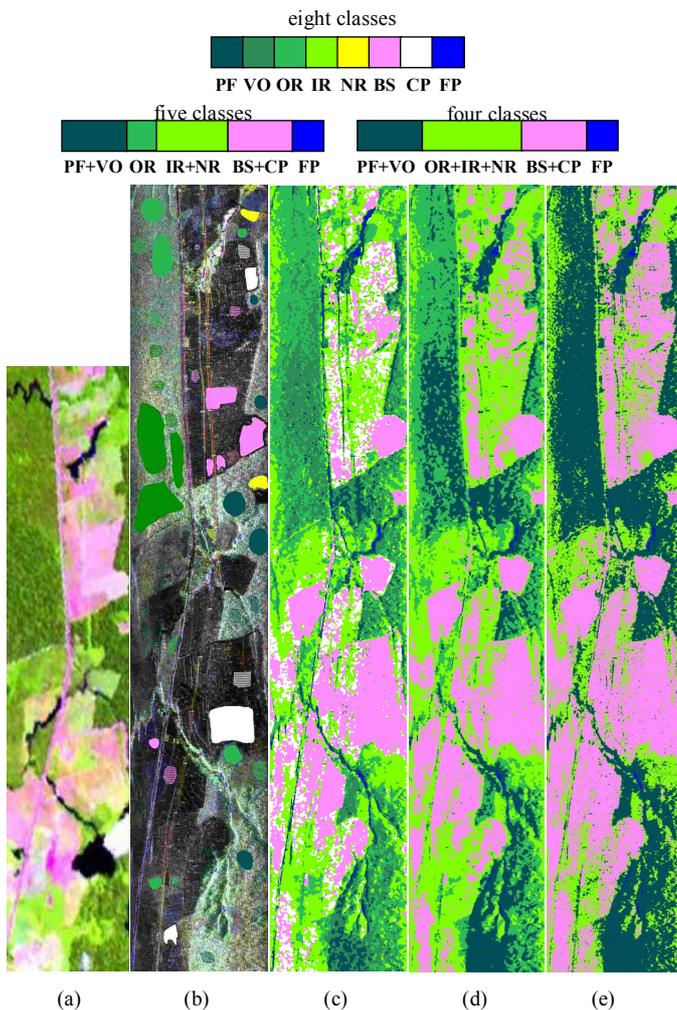


Figure 1. (a) Color composition 5R4G3B of Landsat-7/TM August, 2000, (b) Color composition of P-band intensities R-HH, G-HV, B-VV with the training (solid polygons) and test (hachured polygons) samples, (c) BMI classification with eight classes, (d) HV classification with five classes, (e) bivariate intensities HV-VV classification with four classes.

Reducing the number of classes to four, by further merging old, intermediate and new regenerations, the best classification result with individual band was HV. It can be noted that there was a statistically significant improvement in the classification accuracy when the VV image is used in conjunction with the HV image. Based on the confusion matrices, it seems that HV is more efficient than VV to differentiate primary and very old regeneration forest areas from other land cover classes (because HV return is more related to volume scattering) and VV improves the classification for bare soil and crop/pasture areas. We obtained the best results with Bivariate Intensities HV-VV and Multivariate Intensities HH-HV-VV, where their Kappa coefficients are statistically equal at the 95% confidence level, showing again the lack of additional information of HH image. The classification of the Bivariate Intensities HV-VV is presented in Fig. 1e.

## V. CONCLUSIONS

Using the ML/ICM algorithm for land use classification, the polarimetric P-band data were not able to distinguish several land cover classes. However, it was able to differentiate primary forest/very old regeneration from other types of regeneration. This is probably justified by their different vegetation structures, which is sensible to P band frequency. This fact is clearly perceptible through HV classification. In general, the HV image presented better classification result than HH or VV ones, when individual images are considered. It was also observed that VV improves the classification for bare soil and crop/pasture areas. The use of polarimetric data is essential to increase the number of land use classes, which can be uniquely identified.

## REFERENCES

- [1] INPE. "Monitoramento da Floresta Amazônica Brasileira por Satélite. São José dos Campos, Brazil: Instituto Nacional de Pesquisas Espaciais, 2006. <http://www.obt.inpe.br/prodes/>, October, 2006.
- [2] D.S. Alves, M.I.S. Escada, J.L.G. Pereira, C.D. Linhares, "Land use intensification and abandonment in Rondonia, Brazilian Amazonia," *Int. J. Remote Sens.*, v.24(4), pp. 899-903, February 2003.
- [3] T. Neeff, L.V. Dutra, J.R. Santos, C.C. Freitas, L. S. Araújo, "Tropical forest biomass measurement by interferometric height modelling and P-band radar backscatter," *Forest Science*, v. 51, pp. 585-594, 2005.
- [4] J.R. Santos, C.C. Freitas, L.S. Araujo, L.D. Dutra, J.C.Mura, F.F.Gama, L.S.Soler, S.J.S. Sant'Anna, "Airborne P-band SAR applied to the above ground biomass studies in the Brazilian tropical rainforest," *Remote Sens. Environ.* v.87, pp. 482-493, Special Issue of LBA, 2003.
- [5] D.H. Hoekman, M.J. Quiñones, "Biophysical forest type characterization in the Colombian amazon by airborne polarimetric SAR," *IEEE Trans. Geosci. Remote Sens.*, v.40(6), pp. 1288-1300, 2002.
- [6] A.C. Frery, H.J. Müller, C.C.F. Yanasse, S.J.S. Sant'Anna, "A model for extremely heterogeneous clutter," *IEEE Trans. Geosci. Remote Sens.*, v.35(3), pp. 648-659, 1997.
- [7] M.S. Srivastava, "On the complex Wishart distribution," *Annals of Mathematical Statistics*, v.36(1), pp. 313-315, 1963.
- [8] J.S. Lee, K.W. Hoppel, S.A. Mango, A.R. Miller, "Intensity and phase statistics of multi-look polarimetric and interferometric SAR imagery," *IEEE Trans. Geosci. Remote Sens.*, v.32, pp. 1017-1028, 1994.
- [9] C.C. Freitas, A.C. Frery, A.H. Correia, "The polarimetric G distributions for SAR data analysis," *Environmetrics*, v. 16, pp. 13-31, 2005.
- [10] K.O.Pope, J.M. Rey-Benayas, J.F. Paris, "Radar remote sensing of forest and wetland ecosystems in the Central American Tropics," *Remote Sens. Environ.*, v.48, pp. 205-219, 1994.