Working Paper Series

Textural classification of land cover using support vector machines: an empirical comparison with parametric, non parametric and hybrid classifiers in the Bolivian Amazon

Jaime Paneque-Gálvez, Jean-François Mas, Gerard Moré, Jordi Cristóbal, Martí Orta-Martínez, Ana Catarina Luz, Maximiliem Guèze, Manuel Macía, Victoria Reyes-García

Tsimane' Amazonian Panel Study Working Paper # 69

Textural classification of land cover using support vector machines: an empirical comparison with parametric, non parametric and hybrid classifiers in the Bolivian Amazon

Jaime Paneque-Gálvez[a]*, Jean-François Mas[b], Gerard Moré[c], Jordi Cristóbal[d], Martí Orta-Martínez[a], Ana Catarina Luz[a], Maximiliem Guèze[a], Manuel Macía[e], Victoria Reyes-García[f]

Affiliations:

[a] Institut de Ciència i Tecnologia Ambientals (ICTA), Universitat Autònoma de Barcelona (UAB), 08193 Bellaterra, Barcelona, Spain

[b] Centro de Investigaciones en Geografía Ambiental (CIGA), Universidad Nacional Autónoma de México (UNAM), Antigua Carretera a Pátzcuaro No 8701, Col. Ex-Hacienda de San José de La Huerta, 58190, Morelia, Michoacán, Mexico

[c] Centre de Recerca Ecològica i Aplicacions Forestals (CREAF), Universitat Autònoma de Barcelona (UAB), 08193 Bellaterra, Barcelona, Spain

[d] Departament de Biologia Animal, Biologia Vegetal i Ecologia, Universitat Autònoma de Barcelona (UAB), 08193 Bellaterra, Barcelona, Spain

[e] Departamento de Biología, Unidad de Botánica, Universidad Autónoma de Madrid (UAM), Calle Darwin 2, 28049 Madrid, Spain

[f] ICREA and Institut de Ciència i Tecnologia Ambientals (ICTA), Universitat Autònoma de Barcelona (UAB), 08193 Bellaterra, Barcelona, Spain

* Corresponding author at: Institut de Ciència i Tecnologia Ambientals (ICTA), Universitat Autònoma de Barcelona (UAB), 08193 Bellaterra, Barcelona, Spain. Tel: (+34) 93 586 85 48; Fax: (+34) 93 586 80 08; E-mail: jpanequegalvez@gmail.com

Date: *28-11-2011*

ABSTRACT: Land cover classification is a key research field in remote sensing and land change science as thematic maps derived from remotely sensed data have become the basis for analyzing many socio-ecological issues. However, land cover classification remains a difficult task and it is especially challenging in heterogeneous tropical landscapes where nonetheless such maps are of great importance. The present study aims to establish an efficient classification approach to accurately map all broad land cover classes in a large, heterogeneous tropical area of Bolivia, as a basis for further studies (e.g., land cover-land use change). Specifically, we compare the performance of parametric (maximum likelihood), non-parametric (*k*-nearest neighbour and four different support vector machines - SVM), and hybrid classifiers, using both hard and soft (fuzzy) accuracy assessments. In addition, we test whether the inclusion of a textural index (homogeneity) in the classifications improves their performance. We classified Landsat imagery for two dates corresponding to dry and wet seasons and found that non-parametric, and particularly SVM classifiers, outperformed both parametric and hybrid classifiers. We also found that the use of the homogeneity index along with reflectance bands significantly increased the overall accuracy of all the classifications, but particularly of SVM algorithms. We observed that improvements in producer's and user's accuracies through the inclusion of the homogeneity index were different depending on land cover classes. Early-growth/degraded forests, pastures, grasslands and savanna were the classes most improved, especially with the SVM radial basis function and SVM sigmoid classifiers, though with both classifiers all land cover classes were mapped with producer's and user's accuracies of around 90%. Our approach seems very well suited to accurately map land cover in tropical regions, thus having the potential to contribute to conservation initiatives, climate change mitigation schemes such as REDD+, and rural development policies.

Keywords: remote sensing, thematic classification comparison, SVM, *k*-nearest neighbor, hybrid classification, homogeneity index, Bolivian Amazon

1. Introduction

Land cover (LC) classification is a key research field in remote sensing (Lu and Weng, 2007) and a fundamental component of land change science as LC maps derived from remotely sensed data are vital to analyze environmental change (Turner *et al.*, 2007). LC maps are critically needed in regions where few or no other up-to-date maps exist at local and landscape scales, which is a common situation in many tropical areas worldwide. In the Amazon basin, most LC classification efforts have been made in Brazil. Nevertheless, the rest of the Amazon basin provides environmental services well beyond their reaches too, playing a crucial role into global atmospheric circulation and regional precipitation (Malhi *et al.*, 2008), acting as a global carbon pool (Nobre and Borma, 2009) and a significant freshwater reservoir (Fearnside, 1997), hosting an incredible wealth of bio-cultural diversity (Maffi, 2005), and providing many commodities to global markets (Fearnside, 2005; Finer *et al.*, 2008). Therefore, producing high-quality LC maps also for these areas is paramount as a basis for undertaking further studies or designing management plans and policies.

LC classification remains a challenging task in highly heterogeneous tropical areas for several reasons. A major problem lies in the difficulty of acquiring cloud-free multispectral imagery (Asner, 2001), which may be partly overcome through the use of radar imagery (Saatchi *et al.*, 1997; Freitas *et al.*, 2008; Santos *et al.*, 2008). However, radar imagery interpretation is not straightforward in tropical areas (Almeida-Filho *et al.*, 2007) and there is often a lack of radar analysts in tropical countries. Another major

drawback is related to the limitations (in terms of cost, time, and feasibility) for carrying out fieldwork to collect sufficient information on LC classes, which hampers the training and validation stages of supervised or hybrid LC classification approaches. Other constraints that may severely impact on the potential to accurately map LC in tropical landscapes are the usual lack of aerial photography and videography, of previous LC maps, and of ancillary data (e.g., digital elevation models, land use maps, GIS layers) that may be used to support the classification process. Given all these limitations, in tropical regions it is essential to deploy an efficient LC classification scheme.

A fundamental issue is the selection of the classifier. In that respect, the use of machine learning algorithms for LC classification have gained momentum in recent years and some assessments of their relative performance compared to other classifiers have been conducted in the Amazon region (Lu *et al.*, 2004; Carreiras *et al.*, 2006a). Among machine learning classifiers, SVM have been shown to have some specific advantages that may render their use for LC mapping even more appropriate (Huang *et al.*, 2002; Foody and Mathur, 2004a; Pal and Mather, 2005; Dixon and Candade, 2008; Kavzoglu and Colkesen, 2009; Mountrakis *et al.*, 2011; Szuster *et al.*, 2011). For instance, since SVM classifiers seek to separate LC classes by finding a plane in the multidimensional feature space that maximizes their separation, rather than by characterizing such classes with statistics, they do not need a large training set. Foody and Mathur (2004b) found that SVM classifiers only require the training samples that are support vectors, which lie on part of the edge of the class distribution in feature space, as all the others do not provide useful information to them. Both authors demonstrated later on (Foody and Mathur, 2006) that when using SVM classifiers, rather than a conventional training approach it is advisable to follow an alternative approach that entails the use of small training sets composed of purposely selected mixed pixels containing the support vectors, as this may considerably reduce the time needed for the training stage (and therefore the classification cost) without compromising the classification accuracies obtained. Another major advantage of SVM algorithms is that they are independent of data dimensionality, which is a key feature when using many spectral bands or when ancillary data are included in the classification process, as for classifiers dependent on dimensionality (e.g., artificial neural networks) training sets must exponentially increase in size to maintain classifier performance (Dixon and Candade, 2008). All these SVM features and recent developments have made them very popular, being even portrayed as “an ultimate classifier that may possibly provide the best classification performance” (Chen and Ho, 2008).

Another fundamental issue to enhance LC classification is the adequate selection of input variables, which some authors suggest may have the same impact as the selection of the classifier (Heinl *et al.*, 2009). Nevertheless, it seems logical that the combination of an allegedly superior classifier such as SVM and appropriate ancillary data should improve results. In that vein, Watanachaturaporn *et al.* (2008) compared classification performance of SVM, decision tree and two neural network classifiers with and without ancillary data (NDVI and a DEM) and found that the SVM multisource classification outperformed all the rest. Another source of ancillary data is texture. The benefits of incorporating texture in LC classification have been pointed out in numerous studies making use of different classification algorithms and texture measures (Berberoglu *et al.*, 2000; Chica-Olmo and Abarca-Hernández, 2000; Haack and Bechdol, 2000; Maillard, 2003; Tsaneva *et al.*, 2010). The incorporation of texture measures for mapping forest age, forest types, detecting forest cover change, and characterizing canopy structure has also been widely examined (Woodcock *et al.*, 1994; Augusteijn *et al.*, 1995; Palubinskas *et al.*, 1995; Riou and Seyler, 1997; Franklin *et al.*, 2000; Franklin *et al.*, 2001; Chan *et al.*, 2003; Coburn and Roberts, 2004; Zhang

et al., 2004; Kayitakire *et al*., 2006; Malhi and Román-Cuesta, 2008; Ota *et al*., 2011). A significant advantage of using texture to enhance image classification in tropical regions (where other ancillary data sources may not exist) is that texture data can be extracted from the image itself. An approach which yields good results consists in extracting information from the image using the gray-level co-occurrence matrix (GLCM) method, and using such information as data bands in the classification process (Gong *et al*., 1992).

In the Bolivian Amazon there have been few attempts to provide LC classifications; the most comprehensive study was carried out by Killeen *et al.* (2007), which served as the basis for a historical land use change analysis of all Bolivian territory below 3000 m of altitude (Killeen *et al*., 2008). However, owing to the spatial extent of their analysis, Killeen and colleagues used very broad LC classes and thereby their results may be of limited usefulness for studies that need to analyze LC or LC change at local or landscape scales. To fill that gap, the main goal of this study was to establish an appropriate classification approach to accurately map all LC classes considered in our study area in the Bolivian Amazon. Our specific aims were: 1) to test if the use of SVM classifiers improved LC classification with regard to conventional parametric, non-parametric, and hybrid classifiers; and 2) to assess whether the inclusion of a textural index (homogeneity, which has seldom been used for LC classification), could improve classification results.

2. Study Area, field surveys and map legend definition

2.1. Study area

The study area is located in the Department of Beni, Bolivia (Fig. 1). We selected this large area because its landscapes are highly heterogeneous as a transition across three biogeographic areas: 1) montane tropical forests covering the foothills of the Andes to the West, 2) lowland tropical forests to the South and Centre of the study area, and 3) wet savannas to the North and East (Navarro and Maldonado, 2002). Montane tropical forests are possibly the most plant diverse area of Bolivia (Ibisch and Mérida, 2004) and in our study area are found over 400m. Lowland forests are located below 400m and conform a rolling landscape. These forests contain some deciduous species owing to a marked seasonality (dry and wet seasons) and are not as species-rich as montane tropical forests or typical Amazonian rainforests, though they are very similar in species composition and structure to the latter ones (Killeen *et al*., 1993). In the case of wet savannas, vegetation is controlled by small variations in ground elevation and relief, which in turn are shaped by river dynamics and periodic flooding. They consist of swampy areas and lagoons with aquatic vegetation in the lowest areas; scrublands, semi-natural grasslands and pastures in areas less prone to be flooded; and patches of forests on mounds that do not get seasonally flooded and are seemingly the result of past civilizations (Mann, 2008; Lombardo and Prümers, 2010). The vegetation formations of the study area are also shaped by the land use type and intensity of its different inhabitants, who range from Andean indigenous peoples in montane forests, to local peasants, cattle ranchers, and different native and colonist indigenous peoples in lowland forests and savanna areas.

2.2. Field surveys

Two field surveys were undertaken to cover the study area with the purpose of training and testing data collection. The first focused on forested areas (old-growth, early-growth, and degraded forests), water, bare soil, and infrastructure/urban categories, and was carried out in June-August 2009 (dry season). The second one took place in

April-May 2010 (end of the wet season) and was conducted on the large savanna areas that are present across the study area, which mix with patches of pastures, semi-natural grasslands and scrublands. Planning the acquisition of ground-truth data was done upon preliminary analyses of the most recent Landsat-5 TM scenes (April 2009). Ground data were acquired with handheld GPS units, with typical mean positioning errors of 2-4 m in open areas and 4-6 m in forested areas. Information on land cover-land use plus a set of ecological and geomorphological features were recorded along with each GPS reading. Additionally, to assist in the processes of geometric correction and geometric accuracy assessment, GPS points were collected at road crossings and other human-made features on the ground, together with GPS tracks along the major roads and rivers in the study area.

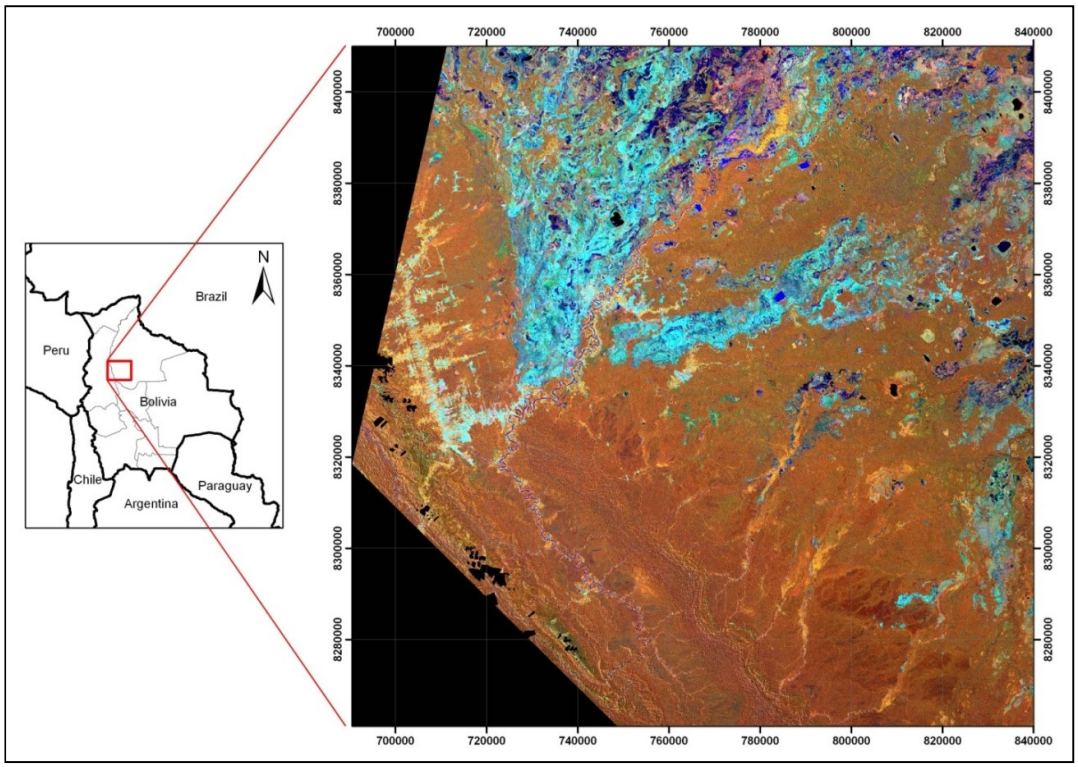

Figure 1. RGB (4-5-3) composite from a Landsat-5 TM mosaic (17/04/2009) showing the study area. Projected coordinate system: UTM19S (WGS84).

2.3. Map legend definition

The definition of broad LC classes was carried out prior to the field surveys based on previous knowledge of the area and initial remotely sensed data exploration. This exploration consisted in 1) carrying out several unsupervised ISODATA classifications on the most recent Landsat imagery we had, and 2) checking the classification obtained by Killeen *et al.* (2007) for our study area. Nevertheless, the definition of LC classes was modified according to our field observations and thorough examination of the spectral signatures extracted for our field data. Eight broad LC categories were finally considered (Table 1). Agriculture was not defined as a specific category because nearly all the agricultural plots in this area consisted in subsistence slash and burn agriculture. Such plots, that account for a small percentage of the study area, are small (<0.5 ha), have usually great abundance of non-photosynthetic biomass (dead trees and logs) lying on the ground, include banana and other tree crops, and are mixed with

pioneer tree species (e.g., *Cecropia spp.*). In addition, they do not usually have a clear spatial distribution pattern and are highly dynamic, rapidly evolving into young regenerating forests. As a consequence, these agricultural plots exhibit in Landsat imagery a very similar spectral response to older regenerating forests (Chan *et al.*, 2001; Vieira *et al.*, 2003), which makes it very hard to accurately identify this LC class. Therefore, it was deemed preferable to include agriculture within the early-growth forest class. Likewise, the infrastructure/urban class was included within the bare soil class owing to the lack of paved roads and streets in the area. Moreover, the spectral response of these linear features was relatively similar to that of roofs in urban settlements and, to a lesser extent, to that of sand banks along the main rivers. The spectral similarity of agriculture to early-growth forests and of urban areas to bare soil was tested by the Jeffries-Matusita transformed divergence index, which confirmed that in both instances they could be grouped under a single category.

Table 1
Definition of LC classes included in the study.

LC Class	Definition
Early-growth / degraded forest (EGDF)	Forested areas with varying degrees of disturbance due to human activities (e.g., typically slash and burn agriculture or logging) or natural dynamics (e.g., flooding regimes). Typically composed of regenerating trees, dead trees and logs, crops such as rice, manioc and bananas, sometimes with scattered old big trees. The canopy is rather open, structurally simple, and the average tree height is 3-10 m.
Old-growth forest (OGF)	Forested areas with low levels of disturbance that consist of mature trees forming a dense and structurally complex canopy with few gaps and a typical height range of 15-40 m.
Water (W)	Water bodies such as creeks, rivers, shallow lakes, and deep lakes.
Bare soil / urban (BSU)	Sand banks along rivers, urban areas including towns, unpaved streets and roads.
Pasture (P)	Areas typically used for cattle ranching, both in deforested and savanna areas. In deforested areas, pasture species are frequently sown, while in savanna areas pastures are frequently not sown. In both instances it is common to have varying amounts of bare soil.
Savanna (S)	Low relief savanna areas that are seasonally inundated and may form swamps or marshes.
Semi-natural grassland (G)	Grassland patches that occur mostly across the savanna areas, with very little or total absence of woody species.
Scrubland (SC)	Open canopy areas dominated by bushes or short trees, commonly present across the savanna areas, growing on dry ground of low quality; sometimes in the fringe or vicinity of forested areas.

3. Materials and methods

3.1. Satellite data and pre-processing

LC classifications were carried out on Landsat satellite mosaics of two scenes (path 233, rows 70 and 71) from two dates (25/08/2001 and 17/04/2009, corresponding to the dry and the end of the wet periods respectively) so as to account for differences in phenology, illumination, and reflectance, and hence strengthen the classification comparison among different algorithms. We chose Landsat data because Landsat is arguably the world's most commonly used satellite to undertake ecological studies in

tropical landscapes, including LC classifications (Cohen and Goward, 2004), which renders our results more comparable to other studies than if we had used a less conventional sensor. In addition, we used a digital elevation model (ASTER GDEM v.1).

The two 25/08/2001 Landsat-7 ETM+ scenes were acquired through the USGS and their geometric accuracy was assessed through the ground control points and GPS tracks we had collected in the field. Geometric accuracy was high (misalignments around 0.5 pixels) in both cases, which was deemed appropriate for the objectives of the study. The two 17/04/2009 Landsat-5 TM scenes acquired from INPE required geometric and topographic corrections, which were carried out with MiraMon software (Pons, 2000) using the procedure developed by Palà and Pons (1995). The Landsat-7 ETM+ scenes were used as reference images and the geometric errors obtained for both Landsat-5 TM images after the corrections were consistent with those of the reference scenes (less than a pixel). The nearest neighbor algorithm was used in the corrections so as to preserve the original radiometry of the images. Subsequently, each pair of images corresponding to the same date was mosaicked., and radiometric corrections were performed using MiraMon, which implements a radiometric correction model that includes atmospheric and illumination corrections. The atmospheric correction is done in terms of additive and multiplicative factors that account for solar irradiance, atmospheric effects and sensor calibration, operating on an image-wide basis. The illumination correction is carried out on a pixel by pixel basis using a digital elevation model, and it minimizes the effects of differential illumination conditions due to sun locations and relief, as well as atmospheric conditions (Pons and Solé-Sugrañes, 1994). Upon radiometric correction completion, the two mosaics were cropped to the extent of the area of interest and a cloud and cloud-shadow mask was applied to all of them, being thus ready for classification analysis.

3.2. Training data

Training data were retrieved for each mosaic starting with the most recent one (2009). This process was based on field data and careful examination of spectral signatures. In most cases training data consisted of small polygons, though there were few instances in which single pixels were chosen in narrow areas (e.g., roads, sand banks and rivers). Much care was taken to scatter training areas across the image to ensure they were representative of the entire mosaic, and to retrieve as many training areas for each class to satisfy previously suggested criteria to establish an appropriate minimum sample size (Mather, 2004; Congalton and Green, 2009; Foody, 2009). To enhance the comparability of results between the classifications of both dates we tried to use the same training areas as much as possible. The Jeffries-Matusita transformed divergence index was used to assess the separability of training data for both dates. We confirmed that separability was high for some LC classes (i.e., water, bare soil/urban, old-growth forest), and lower for the others.

3.3. Classification algorithms

We used a parametric classifier (maximum likelihood – ML), non-parametric classifiers (*k*-nearest neighbors – KNN, and four different support vector machines – SVM: lineal, polynomial, radial basis function and sigmoid), and a hybrid classifier (contained in MiraMon software – MMHC). We do not explain here how the ML, KNN and SVM algorithms work since detailed descriptions abound in remote sensing and pattern recognition textbooks (Richards and Jia, 2006; Tso and Mather, 2009). However, we provide a very brief explanation of MMHC as it is not a conventional classification method, (for further details see Serra *et al.* (2003)). MMHC classification approach

involves the use of an unsupervised ISODATA algorithm based on the methodology proposed by Duda and Hart (1973) to retrieve spectral classes, and a subsequent supervised classification performed on the ISODATA results using training areas to obtain thematic classes. MMHC has been successfully used to classify Mediterranean environments (Serra *et al.*, 2003; Serra *et al.*, 2005) and has been used in tropical dry areas of Nicaragua to classify vegetation (García-Millán and Moré, 2008). To our knowledge this is the first time MMHC has been used to classify tropical forests and savannas.

3.4. Textural data

We explored the use of textural homogeneity in the LC classification process to assess whether its usage could improve classification results. Textural homogeneity indicates the amount of local similarities within the chosen area (Huvenne *et al.*, 2002). It is higher for regions with a uniform reflectance and lower for regions that are spectrally heterogeneous and thus have varying reflectance. Textural homogeneity can be calculated by the homogeneity index (HI) as follows:

$$HI = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \frac{f(i,j)}{1+|i-j|}$$

where f (i, j) is the brightness value of the pixel located at i^{th} row and j^{th} column in the operation window, and N is the pixel number of the operation window (Zhang, 2001). We calculated HI using moving windows of 3x3 and 7x7 pixels based on the GLCM, which is often employed to extract textural indices from remote sensing images (Haralick *et al.*, 1973; Gong *et al.*, 1992; Zhang, 1999). Eventually we used the 6 HI bands calculated with a 7x7-pixel window from Landsat reflectance bands 1-5 and 7.

3.5. Classification post-processing and accuracy assessment

We applied a 3x3-pixel majority filter to all the classifications to eliminate the salt and pepper effect prior to their accuracy assessment. Reference data retrieval for accuracy assessment was based on a stratified random sample selection, with sample units taken at a minimum distance of 2.0 Km to avoid the potential effects of spatial autocorrelation (Congalton, 1988) and ground-truthed by expert-knowledge from the images themselves. Sample units lying on the fringe of two or more LC classes were not discarded so as not to affect the randomness principle of accuracy assessment. For practical reasons we used points rather than clusters of 3x3 pixels (as suggested by Congalton and Green (2009) as sample units for accuracy assessment of Landsat-derived classifications). However, since we performed a 3x3-pixel majority filter on every classification prior to their accuracy assessment, each reference point represents not the thematic class of a single pixel, but the most common class classified within the 3x3-pixel window centered on the reference point. Regarding the minimum reference sample set size required for accuracy assessment we used the rule of thumb proposed by Congalton and Green (2009), whereby 75-100 testing sample units per thematic class should suffice for large areas and less than 12 thematic classes.

We carried out both hard and soft (also known as fuzzy) classification accuracy assessments. The soft classification assessment may enable a better evaluation of the behavior of a classifier, particularly regarding points that are challenging because they lie on transition or mixed zones (Woodcock and Gopal, 2000), being thus well suited to compare classifiers. For the soft assessment we considered two possible LC classes for each reference point: a primary class, which coincided with that used in the hard

classification assessment and that was supposed to represent the ground *truth*, and a secondary class, which was specific to the soft assessment and was considered to be right too because it represented a *good* or *acceptable answer* given the location of the reference point. However, whenever a reference point was located in a homogeneous area and its LC class was deemed clear, we assigned the same LC class to both primary and secondary classes. For both types of assessments and for each classification obtained we generated a confusion matrix (CM – also known as error matrix), which is the most standard method for remote sensing classification accuracy assessment (Congalton and Green, 2009). Through the construction of CM we retrieved for each assessment the classification overall accuracy (OA) as a global measure of classification accuracy, and the producer's and user's accuracies (PA and UA, respectively) as specific accuracy measures to each of the eight thematic classes considered in this study. We did not retrieve the kappa coefficient as some authors reported this measure of global map accuracy is problematic (Stehman, 1997; Foody, 2004) since it does not have a probabilistic interpretation (unlike OA, PA or UA). Moreover, the kappa coefficient has been shown not to be an appropriate map accuracy measure for comparing the accuracy of thematic maps, particularly when (as in this study) the reference data used have always been the same (Foody, 2004).

3.6. Classification overall accuracy comparison

We used the McNemar test to assess the statistical significance of the difference in OA between each pair of classifications because we had used identical reference data to generate the CM and thus obtain the proportion of correctly allocated cases (Agresti, 1996; Foody, 2004). This test is based on a 2x2 matrix and analyzes the level of agreement with respect to correct and incorrect allocations between two classifications based on the following formula:

$$Z = \frac{f_{12} - f_{21}}{\sqrt{f_{12} + f_{21}}}$$

where f_{ij} indicates the frequency of allocations lying in element *i,j* of the 2x2 matrix. This test compares the frequencies of cases correctly allocated in one classification but misclassified in the other. Two classifications are considered to be significantly different at the 95% level of confidence if $Z > |1.96|$ (de Leeuw *et al.*, 2006; Foody, 2006). In this study we carried out McNemar tests to evaluate the statistical significance of differences in classification OA observed 1) among the seven classifications without HI (i.e., 21 tests), 2) between each of the seven classifiers with and without HI (7 tests), and 3) among the seven classifications with HI (21 tests).

4. Results

4.1. Improvements in overall LC classification results using SVM

We find that, for both dates (2001 and 2009) and for both types of accuracy assessments (hard and soft), all four SVM classifications attain the highest OA and only KNN for 2001 imagery is comparable to them (Table 2). On the contrary, MMHC attains the lowest OA for both dates and assessments though these are similar to ML and KNN for 2009 imagery. ML results are significantly worse than those of SVM for both dates and assessments whereas KNN results appear somehow contradictory as they are as good as SVM for 2001 but as bad as MMHC and ML for 2009, irrespective of the type of accuracy assessment. We used the McNemar test to evaluate differences only in OA of hard accuracy assessments as the differences in OA of soft

assessments were very similar (see Table 2). McNemar results are shown in Table 3 and a few things are worth noting. First, the statistical significance of the differences between any SVM and MMHC is always maximum regardless of the date, whereas that between any SVM and ML ranges from significant to extremely significant for 2001 and is always extremely significant for 2009. Second, KNN shows no statistically significant differences with the worst classifiers (ML and MMHC) for 2009 imagery (although its OA is slightly higher than theirs) and extremely significant differences with all SVM classifiers. However, for 2001 imagery KNN shows no statistically significant differences with the best SVM classifiers and a slightly significant difference with SVM linear, attaining in fact the highest OA of all the classifiers. Third, the relative performance of the different SVM algorithms is very similar. There are no statistically significant differences among them for 2009 and minor differences for 2001 imagery, being perhaps in the latter case the SVM polynomial of 6th grade the best one though this classifier is no significantly better than the SVM RBF and SVM sigmoid.

Table 2
Hard versus soft overall accuracy classification results obtained for 2009 and 2001 imagery using only reflectance bands. OA=Overall accuracy; ML=Maximum Likelihood; SVM=Support Vector Machine; KNN=*k*-nearest neighbors; MMHC=Hybrid classification.

Hard vs. Soft Assessment	**17/04/2009**		**25/08/2001**	
Classifier	Hard OA	Soft OA	Hard OA	Soft OA
ML	**71.50**	**80.50**	**70.25**	**70.88**
SVM Linear	**79.25**	**86.75**	**73.25**	**76.63**
SVM Polynomial (6th grade)	**79.25**	**86.63**	**74.50**	**74.88**
SVM Radial Basis Function	**79.38**	**87.00**	**73.88**	**74.25**
SVM Sigmoid	**80.13**	**87.75**	**75.25**	**75.63**
KNN	**72.63**	**80.50**	**75.75**	**76.75**
MMHC	**70.75**	**79.75**	**64.13**	**67.88**

Table 3
McNemar tests showing the statistical significance of the differences in overall accuracy from a hard assessment among classifiers, without using HI and for both 2009 and 2001 imagery. Codes are as follows (with 95% confidence interval): 0 – No significant ($p>0.05$), 1 – Hardly significant ($p\sim0.5$), 2 – Significant ($0.5<p\leq0.01$), 3 – Very significant ($p=0.001$), 4 – Extremely significant ($p=0.0001$). Positive values indicate better performance of the row classifier whereas negative values indicate better performance of the column classifier. Left values refer to 2009 and right values to 2001 classifications

17/04/2009 | 25/08/2001	**SVM Lineal**	**SVM Polynomial**	**SVM RBF**	**SVM Sigmoid**	**KNN**	**MMHC**
ML	**-4 | -2**	**-4 | -3**	**-4 | -2**	**-4 | -3**	**0 | -4**	**0 | 3**
SVM Lineal		**0 | -2**	**0 | 0**	**0 | -1**	**4 | -1**	**4 | 4**
SVM Polynomial			**0 | 0**	**0 | 0**	**4 | 0**	**4 | 4**
SVM RBF				**0 | 0**	**4 | 0**	**4 | 4**
SVM Sigmoid					**4 | 0**	**4 | 4**
KNN						**0 | 4**

4.2. Improvements in overall LC classification results using the homogeneity index

The use of HI greatly improves the results obtained by any classifier regardless of the date and the type of accuracy assessment (Table 4). Even though the differences in OA attained by each classifier with and without HI are striking, they too have been evaluated with the McNemar test albeit only for hard assessments as the differences in OA for soft assessments follow the same pattern. We find that the differences in OA

are extremely significant for all the classifiers, with the sole exception of the MMHC classification for 2009 imagery, which is not statistically significant. Yet, even in that case, the OA increased 1.25% with the inclusion of HI. Looking carefully at the magnitude of the improvements achieved with the inclusion of HI in the classifications (Table 4), we find that there is a gradient from small to moderate improvements in MMHC (1.25% & 7.38%), moderate to large improvements in ML (7.73% & 9.13%) and KNN (7.12% & 11.00%), and large to very large improvements in the four SVM classifiers (ranging from around 11% to 13% for 2009 imagery, to around 16% and up to 22% for 2001 imagery). We applied McNemar tests to evaluate whether there was statistical significance regarding the differences obtained in map (hard) accuracy by all the classifiers when incorporating HI in the classification process (Table 5). The most striking finding here is that with HI all four SVM algorithms outperform even further all the other algorithms. For instance, it is remarkable that without HI and for 2001 imagery, KNN shows no significant difference with any SVM and actually performs a bit better than SVM linear (see Table 3), whereas with HI the statistical significance of the superiority of SVM algorithms over KNN ranges from significant to extremely significant (Table 5). Similarly, the superiority of all SVM over ML classifiers increases for 2001 imagery with the inclusion of HI, as evidenced by the increase in the statistical significance of their differences (see Table 3 vs. Table 5). Therefore, all four SVM classifiers optimize the use of HI compared to KNN, ML and MMHC. Looking at the differences in performance among the four SVM algorithms it is worth mentioning that, with HI, SVM sigmoid and particularly SVM RBF obtained the best results for both dates, thus maximizing the usefulness of HI (without HI all SVM performed equally well).

Table 4
Hard classification results obtained for the images of 2009 and 2001 using both reflectance and HI bands. IOA=Improvement in Overall Accuracy owing to the inclusion of HI in the classification.

Hard vs. Soft Assessment	17/04/2009				25/08/2001			
Classifier	Hard OA	Hard IOA	Soft OA	Soft IOA	Hard OA	Hard IOA	Soft OA	Soft IOA
ML	**78.88**	**7.38**	**85.25**	**4.75**	**79.38**	**9.13**	**81.50**	**10.62**
SVM Linear	**90.50**	**11.25**	**93.88**	**7.13**	**90.38**	**17.13**	**90.50**	**16.87**
SVM Polynomial	**89.13**	**9.88**	**92.25**	**5.62**	**89.88**	**15.38**	**90.13**	**15.25**
SVM RBF	**92.63**	**13.25**	**95.63**	**8.63**	**96.13**	**22.25**	**96.25**	**22.00**
SVM Sigmoid	**92.75**	**12.62**	**95.50**	**7.75**	**90.63**	**15.38**	**92.00**	**16.37**
KNN	**79.75**	**7.12**	**86.25**	**5.75**	**86.75**	**11.00**	**88.63**	**11.88**
MMHC	**72.00**	**1.25**	**81.38**	**1.63**	**71.25**	**7.12**	**75.38**	**7.50**

Table 5
McNemar tests showing the statistical significance of the differences in overall accuracy from a hard assessment among classifiers, using HI and for both 2009 and 2001 imagery. Notation as in Table 3.

17/04/2009 | 25/08/2001	SVM Lineal + HI	SVM Polynomial + HI	SVM RBF + HI	SVM Sigmoid + HI	KNN + HI	MMHC + HI
ML + HI	**-4 | -4**	**-4 | -4**	**-4 | -4**	**-4 | -4**	**0 | -4**	**4 | 4**
SVM Lineal + HI		**0 | 0**	**-4 | -4**	**-3 | 0**	**4 | 3**	**4 | 4**
SVM Polynomial + HI			**-4 | -4**	**-4 | 0**	**4 | 2**	**4 | 4**
SVM RBF + HI				**0 | 4**	**4 | 4**	**4 | 4**
SVM Sigmoid + HI					**4 | 3**	**4 | 4**
KNN + HI						**4 | 4**

4.3 Improvements in classification results by LC classes using the homogeneity index

To assess what LC classes benefit more with the inclusion of HI in terms of an increase in their producer's (PA) and user's accuracy (UA), we present here the confusion matrices (CM) of ML, KNN and SVM RBF classifications. We do not show CM of SVM linear, SVM polynomial and SVM sigmoid as overall SVM RBF appears to be the best classification of all when HI is included. Neither do we show CM of MMHC because its improvement with HI for 2009 is not significant, only moderate for 2001, and the OA attained with this classifier for either date is not good enough compared to the rest of classifiers tested here. We focus on hard accuracy assessments and show CM only for 2009 classifications because we ground-truthed the area in 2009-2010 but not in 2001. Nevertheless, the results presented are coherent with those obtained for 2001 imagery and with the fuzzy assessments of both dates unless otherwise stated.

Table 6 shows the CM of ML without and with HI. Regarding PA we observe very large improvements in early-growth/degraded forest (20%) and pasture (17%), and a moderate improvement in grassland (7%) when HI is included. Both savanna and scrubland remain with the same PA and old-growth forest is slightly worse (4%) but still has a 90% PA. Regarding UA we see moderate improvements in grassland (7.14%) and scrubland (6.39%), and larger ones in old-growth forest (11.32%) and savanna (15.13%). Both early-growth/degraded forest and pasture remain with the same UA. The results from the CM from 2001 imagery are similar. The main differences in relation to PA are higher increases in pasture and grassland (48% and 23%, respectively) seemingly at the expense of savanna and scrubland that decrease 8% and 7% respectively, whereas for UA the main differences relate to greater improvements in savanna and scrubland (28.47% and 19.53%), and a significant decrease in early-growth/degraded forest (14.74%).

Table 7 shows CM of KNN. Remarkably, these CM are very similar both in values and trends to those of ML. The main differences with what has been shown for ML are that, for KNN, savanna's PA is improved in one date and that early-growth/degraded forest's UA is not affected in either date.

Table 8 shows CM of SVM RBF. With respect to PA everything improves except old-growth forest, which decreases to 93%. PA gains are most remarkable for early-growth/degraded forest (31%) and pasture (20%), but notable for savanna (10%), grassland (13%), and scrubland (11%). UA improvements are very large for old-growth forest (21.57%), savanna (25.33%), and scrubland (18.07%), and moderate for pasture (8.34%), grassland (6.06%), and early-growth/degraded forest (4.10%). The results from 2001 classifications are very similar to the results from 2009 regarding the improvement of classes and the final PA and UA values attained. Again, neither PA nor UA get worse with HI and, in fact, improvements with HI are even higher for 2001: pasture's and grassland's PA increase by 62% and 41% respectively whereas savanna's, grassland's, and scrubland's UA enhance by 36.32%, 32.40% and 34.54% respectively. In sum, SVM RBF maximizes the use of HI as it boosts both PA and UA. Furthermore, it does so for all the LC classes (not just for some as ML and KNN) and up to very high levels of PA and UA (always higher than 90% but scrubland's UA for 2009, unlike ML and KNN that did accurately map only old-growth forest and old-growth forest and pasture, respectively). Lastly, it is to be highlighted that even though the other 3 SVM classifiers did not perform as well as SVM RBF, they show very similar PA and UA values (usually higher than 85-90%) and trends of improvement by LC class with the inclusion of HI.

Table 6
Hard assessment confusion matrices for the maximum likelihood classifications of Landsat data (17/04/2009). Left values refer to the classification without HI and right values to the classification with HI.

Classification Data	Reference Data																			
	EGDF		OGF		W		BSU		P		S		G		SC		Total		UA	
EGDF	**57**	**77**	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**3**	**58**	**81**	98,3	95,1
OGF	**15**	**3**	**94**	**90**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**110**	**93**	85,4	96,8
W	**0**	**0**	**0**	**0**	**57**	**66**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**57**	**66**	100,0	100,0
BSU	**0**	**0**	**0**	**1**	**11**	**14**	**88**	**96**	**6**	**8**	**0**	**1**	**9**	**5**	**0**	**0**	**114**	**125**	77,2	76,8
P	**0**	**0**	**0**	**0**	**0**	**0**	**5**	**4**	**67**	**84**	**7**	**9**	**0**	**0**	**1**	**1**	**80**	**98**	83,7	85,7
S	**0**	**0**	**1**	**1**	**30**	**20**	**0**	**0**	**13**	**0**	**65**	**66**	**0**	**1**	**9**	**6**	**118**	**94**	55,1	70,2
G	**4**	**1**	**1**	**0**	**1**	**0**	**1**	**0**	**1**	**0**	**0**	**2**	**60**	**67**	**5**	**5**	**73**	**75**	82,2	89,3
SC	**24**	**19**	**4**	**7**	**0**	**0**	**6**	**0**	**13**	**8**	**28**	**22**	**31**	**27**	**84**	**85**	**190**	**168**	44,2	50,6
Unclassified	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	0,0	0,0
Total	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**800**	**800**	78,3	83,1
PA	57,0	77,0	94,0	90,0	57,0	66,0	88,0	96,0	67,0	84,0	65,0	66,0	60,0	67,0	84,0	85,0	71,5	78,9		

Table 7
Hard assessment confusion matrices for *k*-nearest neighbor classifications of Landsat data (17/04/2009). Left values refer to the classification without HI and right values to the classification with HI.

Classification Data	Reference Data																			
	EGDF		OGF		W		BSU		P		S		G		SC		Total		UA	
EGDF	**45**	**63**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**1**	**1**	**46**	**65**	97,8	96,9
OGF	**27**	**14**	**94**	**93**	**5**	**10**	**0**	**0**	**0**	**0**	**0**	**2**	**0**	**0**	**1**	**1**	**127**	**120**	74,0	77,5
W	**0**	**0**	**0**	**0**	**78**	**81**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**78**	**82**	100,0	98,8
BSU	**0**	**0**	**0**	**0**	**0**	**1**	**85**	**87**	**3**	**2**	**0**	**0**	**4**	**1**	**0**	**0**	**92**	**91**	92,4	95,6
P	**0**	**0**	**0**	**0**	**0**	**1**	**9**	**10**	**68**	**87**	**5**	**6**	**0**	**0**	**0**	**0**	**82**	**104**	82,9	83,6
S	**0**	**0**	**0**	**0**	**10**	**1**	**0**	**0**	**11**	**3**	**50**	**62**	**0**	**0**	**5**	**4**	**76**	**70**	65,8	88,6
G	**2**	**1**	**1**	**1**	**5**	**3**	**2**	**1**	**2**	**0**	**0**	**1**	**73**	**78**	**5**	**6**	**90**	**91**	81,1	85,7
SC	**26**	**22**	**5**	**6**	**2**	**3**	**4**	**2**	**16**	**8**	**45**	**29**	**23**	**20**	**88**	**87**	**209**	**177**	42,1	49,1
Unclassified	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	0,0	0,0
Total	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**800**	**800**	79,5	84,5
PA	45,0	63,0	94,0	93,0	78,0	81,0	85,0	87,0	68,0	87,0	50,0	62,0	73,0	78,0	88,0	87,0	72,6	79,7		

Table 8
Hard assessment confusion matrices for SVM radial basis function classifications of Landsat data (17/04/2009). Left values refer to the classification without HI and right values to the classification with HI.

Classification Data	Reference Data																			
	EGDF		OGF		W		BSU		P		S		G		SC		Total		UA	
EGDF	**60**	**91**	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**2**	**0**	**1**	**2**	**64**	**93**	93,7	97,8
OGF	**28**	**6**	**95**	**93**	**2**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**8**	**1**	**133**	**100**	71,4	93,0
W	**0**	**0**	**0**	**0**	**79**	**98**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**1**	**79**	**100**	100,0	98,0
BSU	**0**	**0**	**0**	**0**	**0**	**0**	**88**	**92**	**2**	**1**	**0**	**0**	**3**	**0**	**0**	**0**	**93**	**93**	94,6	98,9
P	**0**	**0**	**0**	**0**	**0**	**0**	**7**	**5**	**71**	**91**	**5**	**2**	**0**	**0**	**1**	**0**	**84**	**98**	84,5	92,9
S	**0**	**0**	**1**	**1**	**15**	**0**	**0**	**0**	**14**	**2**	**84**	**94**	**0**	**0**	**10**	**4**	**124**	**101**	67,7	93,1
G	**1**	**0**	**1**	**0**	**4**	**1**	**3**	**3**	**2**	**1**	**0**	**0**	**80**	**93**	**2**	**3**	**93**	**101**	86,0	92,1
SC	**11**	**3**	**3**	**6**	**0**	**1**	**1**	**0**	**11**	**5**	**11**	**3**	**15**	**7**	**78**	**89**	**130**	**114**	60,0	78,1
Unclassified	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	0,0	0,
Total	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**800**	**800**	82,3	93,0
PA	60,0	91,0	95,0	93,0	79,0	98,0	88,0	92,0	71,0	91,0	84,0	94,0	80,0	93,0	78,0	89,0	79,4	92,6		

5. Discussion

5.1. Comparative performance among classifiers

In this study we have found that, when mapping the LC of heterogeneous tropical landscapes of Bolivia, SVM classifiers outperform conventional parametric (ML), non-parametric (KNN), and hybrid (MMHC) classifiers. This finding is consistent with research that has shown the superiority of other non-parametric machine learning algorithms for LC mapping in the Amazon basin (Lu *et al.*, 2004; Carreiras *et al.*, 2006a). However, there are very few examples of the use of SVM for mapping forests and other LC classes in the Amazon (e.g., Wijaya and Gloaguen (2007)). There are several major advantages that have been highlighted regarding the use of SVM (and other non-parametric classifiers) for LC mapping, some of which may make them achieve better results than parametric or hybrid classifiers. A substantial advantage relates to the fact that there is no need to assume any particular data distribution, which facilitates the use of ancillary data in the classification process (Lu and Weng, 2007). This is possibly one major reason why in our study SVM have significantly outperformed the rest of classifiers, as we have verified that some training data do not follow a normal distribution. This may also explain the poor performance of ML even for the case of soft assessments, as this algorithm requires training data be normally distributed. Another reason for the superiority of SVM may be related to the training sets we have used and the occurrence of mixed pixels within them. As Foody and Mathur (2006) demonstrated, SVM use mixed pixels to get the support vectors they need for classifying data, while the rest of classifiers cannot deal properly with mixed pixels as they derive LC class statistics from training samples to characterize such classes.

We do not have a clear explanation with regard to the contrasting performance of KNN for 2001 and 2009 imagery irrespective of the type of accuracy assessment. We tested the separability of all pairs of LC classes for both dates through the Jeffries-Matusita transformed divergence index and found out that, unlike for 2001 imagery, some of them were hardly separable for 2009 imagery, which may suggest that under challenging conditions KNN cannot perform too well. Finally, the performance of MMHC is very poor for both dates and assessments compared to all SVM classifiers despite the many trials we carried out using different parameterization in both unsupervised and supervised stages. Although MMHC has been successfully used in Mediterranean environments (Serra *et al.*, 2003; Serra *et al.*, 2005), we believe it may not be appropriate for classifying tropical areas as they are typically too complex spectrally and, therefore, a very large training set may be needed to derive a supervised classification from the many spectral classes obtained with the ISODATA classification. This finding is consistent with that of García-Millán and Moré (2008) for LC classification of a tropical dry area in Nicaragua, as they obtained 0% of PA for several LC classes and of UA for one class.

5.2. Taking advantage of textural information: the case of the homogeneity index

Our results demonstrate that HI is a very useful textural index for LC mapping. The improvements in OA of both hard and soft accuracy assessments for 2001 and 2009 classifications are extremely significant irrespective of the classifier employed with the sole exception of MMHC for 2009 imagery. This finding is consistent with other studies that have sought to evaluate whether improvements in OA can be achieved when using spectral-textural classification approaches rather than approaches based solely on spectral features (Gong *et al.*, 1992; Franklin *et al.*, 2000; Haack and Bechdol, 2000; Zhang *et al.*, 2004; Ota *et al.*, 2011), including studies in tropical forest areas (Riou and Seyler, 1997; Chan *et al.*, 2001; Chan *et al.*, 2003). Nonetheless, few studies have explored the potential usefulness of HI for textural

analysis of remotely sensed images. For instance, HI has been shown to describe nearly all the textural information contained in sonar imagery along with entropy (Huvenne *et al.*, 2002; Blondel and Gómez Sichi, 2009). In the context of LC mapping from satellite or airborne imagery, Chan and colleagues (2001, 2003) used HI together with other seven textural indices extracted from the GLCM and four indices derived from the grey level difference vector for forest classification with Landsat TM imagery in Peru and the Congo Republic respectively, but apparently did not find HI useful as they did not include it in their final classifications (though this fact was not discussed). Chehade *et al.* (2009) explored the use of HI together with other three conventional textural indices (energy, contrast and entropy) and NDVI, for classifying vegetation types from an aerial color infrared image with SVM algorithms, obtaining good results. We have not found, however, any LC classification study that exploits HI on its own together with spectral bands. In fact, studies comparing the performance of different textural indices for image classification have not regarded HI as one of the most useful ones (Haralick *et al.*, 1973; Gong *et al.*, 1992; Baraldi and Parmiggiani, 1995).

Nevertheless, our study shows that HI may be a very powerful index for improving LC classification accuracies, at least in heterogeneous tropical landscapes, which we believe may be due to two main reasons. First, our results suggest that this index by itself is able to characterize textural variability to a great extent and therefore enhance discriminability of spectral information alone. In this respect, it would be interesting to assess whether, as suggested for sonar imagery (Huvenne *et al.*, 2002; Blondel and Gómez Sichi, 2009), the use of entropy in addition to HI may yield even better results for classifying multispectral imagery. Second, we have observed that the inclusion of HI has often enabled the classifier to correctly allocate ambiguous reference points, i.e., points located in the transition between different covers (mixed pixels) or corresponding to transition covers not considered specifically in the classification scheme. For instance, for 2009 imagery, when HI was used the SVM radial basis function algorithm correctly classified 31 reference points more for the early-growth/degraded forest category. We verified that 28 out of those 31 points lied in between two or more LC classes (normally either old-growth forest or scrubland), and only 3 points were located in homogeneous areas. Similarly, 21 reference points more were correctly classified by using HI, from which 17 were ambiguous. This trend is followed by other LC classes and explains the improvements obtained using soft accuracy assessments. Therefore, the inclusion of HI seems to alleviate the classification problem posed by spectrally mixed pixels. Additionally, we have found two facts in conflict with previous research. First, contrary to what was suggested by Augusteijn *et al.* (1995), small window sizes (7x7-pixel in our study) seem to accurately characterize GLCM textural information as related to LC classes. We believe this responds to the relatively small size of patches of some LC classes such as early-growth/degraded forest, pasture and grassland, which therefore makes texture to change over small areas in many instances. Second, contrary to the findings of Ota *et al.* (2011), classification accuracy improves with the inclusion of textural information at 30-m spatial resolution, including PA and UA of forest types.

Regarding the improvements achieved by different classifiers through the use of HI alongside spectral data, it is clear that SVM classifiers take further advantage of HI. One possible explanation for this improvement is that SVM classifiers are independent of data dimensionality (Dixon and Candade, 2008) unlike the rest of the classifiers tested in this study. Since we have not increased the size of the training set after including the 6 HI bands in the classification, the improvements in performance of SVM classifiers may have been more significant. In addition, we have verified that some training data extracted from the 6 HI data bands calculated for each Landsat mosaic do not follow a normal distribution, which may explain why non-parametric

algorithms deal with the inclusion of HI in the classification in a better way. This fact is exacerbated when looking at HI values for different LC classes, which may possibly explain why certain classes show large improvements in PA and/or UA while others are not seemingly affected or may even decrease their accuracies. This is further discussed in the next subsection.

5.3. Accurate mapping of all LC classes using SVM and the homogeneity index

We have found that the combination of HI and spectral bands is particularly useful when using SVM classifiers (particularly SVM RBF and SVM sigmoid) as all the LC classes considered could be accurately mapped both in terms of their PA and UA. This is important as many studies need to attain reasonably high accuracies for all LC classes considered, for instance studies aiming to monitor LC change trajectories (Brink and Eva, 2009; Schulz *et al.*, 2010) or to model future LC changes (Guerrero *et al.*, 2008). Similarly, in studies that seek to quantify a specific LC class or just a few, attaining high accuracies for such classes is of critical importance. This is common for many studies focusing on mapping forests to assess deforestation, forest degradation and/or forest regrowth (Lambin, 1999; Lucas *et al.*, 2000; Lu *et al.*, 2003; Porter-Bolland *et al.*, 2007; Díaz-Gallegos *et al.*, 2010), which sometimes barely go beyond the forest/non-forest legend (Messina *et al.*, 2006; Walsh *et al.*, 2008; Marsik *et al.*, 2011), as well as for studies concerned with mapping some other typical LC or land use classes such as agriculture, pasture, grassland, savanna areas (Carreiras *et al.*, 2006b; Baldi and Paruelo, 2008; Brannstrom and Filippi, 2008).

In our study, we needed to develop an approach that enabled us to obtain high accuracies for all LC classes aside from water and urban/bare soil as a first step to assess LC change and landscape dynamics. Though we found that HI significantly increased the OA regardless of the classifier used, we observed that for non-SVM classifiers (i.e., KNN, ML, MMHC), improvements in PA and UA were highly dependent on LC class as some classes were largely improved (early-growth/degraded forest, pasture, and grassland for PA, and savanna for UA), while others did not appear to improve or even worsened (scrubland for PA). However, SVM classifiers improved all LC classes both in terms of their PA and UA, with the only exception of old-growth forest (its PA slightly worsened though it always retained an extremely high value). We believe the main reasons underpinning this finding are very similar to what has been explained in the previous subsection, i.e., the optimal use of HI by SVM classifiers compared to the rest (as our training set was not enlarged and non-SVM classifiers would exponentially need more training samples to keep up their performance as data dimensionality increases), and the no normality found in the training samples for some HI bands, which is dependent on LC class and may explain why improvements are not evenly spread among LC classes.

6. Conclusions

LC mapping efforts in tropical regions are key to conservation initiatives, climate change mitigation strategies, and rural development policies. In this study we set out to establish a classification approach to accurately map LC in a large, heterogeneous tropical area of Bolivia. We verified that SVM classifiers outperformed other parametric (ML), non-parametric (KNN) and hybrid (MMHC) classifiers. Nevertheless, the overall accuracies attained based solely on Landsat spectral bands were not satisfactory even for SVM classifiers, and some important LC classes (e.g., early-growth/degraded forest, pasture) were mapped with low producer's and/or user's accuracies. Therefore, the use of some source of ancillary data such as texture was deemed necessary to enhance classification results. We observed that the inclusion of textural

homogeneity (as calculated with the homogeneity index) significantly improved the overall accuracy of the classifications regardless of the algorithm used. This is an important finding as the use of textural homogeneity has been neglected so far in LC mapping efforts. The best results through the inclusion of homogeneity were achieved by SVM classifiers (particularly SVM radial basis function and SVM sigmoid), which further outperformed the rest of classifiers compared in the study, thus optimizing the use of the homogeneity index.

The use of both spectral and textural homogeneity information for LC classification with SVM algorithms enabled us to map the two forested categories (early-growth/degraded forest & old-growth forest) with producer's and user's accuracies greater than 90% for both imagery dates and types of accuracy assessment (hard and soft) used, which makes our approach very suitable to map and monitor tropical forest cover change as needed for ecological assessments and REDD+ schemes. Similarly, the rest of LC classes included in this study were mapped to producer's and user's accuracies of around 90%, rendering our approach very interesting too for land change analysis, biodiversity studies, and natural resource assessments in areas other than forests. Finally, our approach presents the advantage of being easy to implement (as both the calculation of the homogeneity index and the presence of SVM classifiers are readily available in common remote sensing software), and cost-effective (as SVM classifiers may use smaller training sets without compromising classification accuracy).

Acknowledgements

This work was funded through a FBBVA research grant (BIOCON_06_106-07) to the project *Conservación del Bosque Amazónico y Territorios Indígenas: del Conflicto a la Colaboración. Estudio de Caso en la Amazonía Boliviana*. The authors are grateful to Milenka Aguilar and Evaristo Tayo for help with fieldwork in the savanna areas and to TAPS for their assistance with logistics while conducting fieldwork in the study area.

References

Agresti, A., 1996. An Introduction to Categorical Data Analysis. Wiley, New York, N.Y.

Almeida-Filho, R., Rosenqvist, A., Shimabukuro, Y.E., Silva-Gomez, R., 2007. Detecting deforestation with multitemporal L-band SAR imagery: a case study in western Brazilian Amazônia. International Journal of Remote Sensing 28, 1383-1390.

Asner, G.P., 2001. Cloud cover in Landsat observations of the Brazilian Amazon. International Journal of Remote Sensing 22, 3855-3862.

Augusteijn, M.F., Clemens, L.E., Shaw, K.A., 1995. Performance evaluation of texture measures for ground cover identification in satellite images by means of a neural network classifier. IEEE Transactions on Geoscience and Remote Sensing 33, 616-626.

Baldi, G., Paruelo, J., 2008. Land-Use and Land Cover Dynamics in South American Temperate Grasslands. Ecology and Society 13, art6.

Baraldi, A., Parmiggiani, F., 1995. An investigation of the textural characteristics associated with gray level cooccurrence matrix statistical parameters. IEEE Transactions on Geoscience and Remote Sensing 33, 293-304.

Berberoglu, S., Lloyd, C.D., Atkinson, P.M., Curran, P.J., 2000. The integration of spectral and textural information using neural networks for land cover mapping in the Mediterranean. Computers and Geosciences 26, 385-396.

Blondel, P., Gómez Sichi, O., 2009. Textural analyses of multibeam sonar imagery from Stanton Banks, Northern Ireland continental shelf. Applied Acoustics 70, 1288-1297.

Brannstrom, C., Filippi, A.M., 2008. Remote classification of Cerrado (Savanna) and agricultural land covers in northeastern Brazil. Geocarto International 23, 109-134.

Brink, A.B., Eva, H.D., 2009. Monitoring 25 years of land cover change dynamics in Africa: A sample based remote sensing approach. Applied Geography 29, 501-512.

Carreiras, J.M.B., Pereira, J.M.C., Shimabukuro, Y.E., 2006a. Land-cover mapping in the Brazilian Amazon using SPOT-4 vegetation data and machine learning classification methods. Photogrammetric Engineering and Remote Sensing 72, 897-910.

Carreiras, J.M.B., Pereira, J.M.C., Campagnolo, M.L., Shimabukuro, Y.E., 2006b. Assessing the extent of agriculture/pasture and secondary succession forest in the Brazilian Legal Amazon using SPOT VEGETATION data. Remote Sensing of Environment 101, 283-298.

Coburn, C.A., Roberts, A.C.B., 2004. A multiscale texture analysis procedure for improved forest stand classification. International Journal of Remote Sensing 25, 4287-4308.

Cohen, W.B., Goward, S.N., 2004. Landsat's Role in Ecological Applications of Remote Sensing. BioScience 54, 535-545.

Congalton, R.G., 1988. Using spatial autocorrelation analysis to explore the errors in maps generated from remotely sensed data. Photogrammetric Engineering and Remote Sensing 54, 587-592.

Congalton, R.G., Green, K., 2009. Assessing the Accuracy of Remotely Sensed Data—Principles and Practices. CRC Press, Taylor & Francis Group, Boca Ratón.

Chan, J.C.W., Defries, R.S., Townshend, J.R.G., 2001. Improved Recognition of Spectrally Mixed Land Cover Classes Using Spatial Textures and Voting Classifications. In: Skarbek,

W. (Ed.), Computer Analysis of Images and Patterns. Springer Berlin / Heidelberg, 9th International Conference, CAIP 2001 Warsaw, Poland, September 5–7, 2001, pp. 217-227.

Chan, J.C.W., Laporte, N., Defries, R.S., 2003. Texture classification of logged forests in tropical Africa using machine-learning algorithms. International Journal of Remote Sensing 24, 1401-1407.

Chehade, N.H., Boureau, J.G., Vidal, C., Zerubia, J., 2009. Multi-class SVM for forestry classification. 2009 16th IEEE International Conference on Image Processing, pp. 1653-1656.

Chen, C.H., Ho, P.-G.P., 2008. Statistical pattern recognition in remote sensing. Pattern Recognition 41, 2731-2741.

Chica-Olmo, M., Abarca-Hernández, F., 2000. Computing geostatistical image texture for remotely sensed data classification. Computers and Geosciences 26, 373-383.

de Leeuw, J., Jia, H., Yang, L., Liu, X., Schmidt, K., Skidmore, A.K., 2006. Comparing accuracy assessments to infer superiority of image classification methods. International Journal of Remote Sensing 27, 223-232.

Díaz-Gallegos, J.R., Mas, J., Velázquez, A., 2010. Trends of tropical deforestation in Southeast Mexico. Singapore Journal of Tropical Geography 31, 180-196.

Dixon, B., Candade, N., 2008. Multispectral landuse classification using neural networks and support vector machines: one or the other, or both? International Journal of Remote Sensing 29, 1185-1206.

Duda, O., Hart, P.E., 1973. Pattern Classification and Scene Analysis. Wiley, New York.

Fearnside, P.M., 1997. Environmental services as a strategy for sustainable development in rural Amazonia. Ecological Economics 20, 53-70.

Fearnside, P.M., 2005. Deforestation in Brazilian Amazonia: History, rates and consequences. Conservation Biology 19, 680-688.

Finer, M., Jenkins, C.N., Pimm, S.L., Keane, B., Ross, C., 2008. Oil and Gas Projects in the Western Amazon: Threats to Wilderness, Biodiversity, and Indigenous Peoples. PLoS ONE 3, e2932.

Foody, G.M., 2004. Thematic map comparison: evaluating the statistical significance of differences in classification accuracy. Photogrammetric Engineering and Remote Sensing 70, 627-633.

Foody, G.M., 2006. The evaluation and comparison of thematic maps derived from remote sensing. 7th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, Lisbon, Portugal, pp. 18-31.

Foody, G.M., 2009. Sample size determination for image classification accuracy assessment and comparison. International Journal of Remote Sensing 30, 5273-5291.

Foody, G.M., Mathur, A., 2004a. A relative evaluation of multiclass image classification by support vector machines. IEEE Transactions on Geoscience and Remote Sensing 42, 1335-1343.

Foody, G.M., Mathur, A., 2004b. Toward intelligent training of supervised image classifications: directing training data acquisition for SVM classification. Remote Sensing of Environment 93, 107-117.

Foody, G.M., Mathur, A., 2006. The use of small training sets containing mixed pixels for accurate hard image classification: Training on mixed spectral responses for classification by a SVM. Remote Sensing of Environment 103, 179-189.

Franklin, S.E., Hall, R.J., Moskal, L.M., Maudie, A.J., Lavigne, M.B., 2000. Incorporating texture into classification of forest species composition from airborne multispectral images. International Journal of Remote Sensing 21, 61-79.

Franklin, S.E., Wulder, M.A., Gerylo, G.R., 2001. Texture analysis of IKONOS panchromatic data for Douglas-fir forest age class separability in British Columbia. International Journal of Remote Sensing 22, 2627-2632.

Freitas, C., Soler, L., Sant'Anna, S.J.S., Dutra, L.V., dos Santos, J.R., Mura, J.C., Correia, A.H., 2008. Land Use and Land Cover Mapping in the Brazilian Amazon Using Polarimetric Airborne P-Band SAR Data. IEEE Transactions on Geoscience and Remote Sensing 46, 2956-2970.

García-Millán, V., Moré, G., 2008. Generación de cartografía de vegetación a partir de imágenes de satélite en el paisaje protegido de Miraflor-Moropotente, Nicaragua. Universitat Autònoma de Barcelona, Bellaterra (Barcelona), pp. 1-20.

Gong, P., Marceau, D.J., Howarth, P.J., 1992. A comparison of spatial feature extraction algorithms for land-use classification with SPOT HRV data. Remote Sensing of Environment 40, 137-151.

Guerrero, G., Masera, O., Mas, J.F., 2008. Land use / Land cover change dynamics in the Mexican highlands: current situation and long term scenarios. In: Paegelow, M., Camacho-Olmedo, M.T. (Eds.), Modelling Environmental Dynamics. Advances in Geomatic Solutions. Springer Berlin Heidelberg, pp. 57-76.

Haack, B., Bechdol, M., 2000. Integrating multisensor data and RADAR texture measures for land cover mapping. Computers and Geosciences 26, 411-421.

Haralick, R.M., Shanmugam, K., Dinstein, I., 1973. Textural features for image classification. IEEE Transactions on Systems, Man and Cybernetics 3, 610-621.

Heinl, M., Walde, J., Tappeiner, G., Tappeiner, U., 2009. Classifiers vs. input variables—The drivers in image classification for land cover mapping. International Journal of Applied Earth Observation and Geoinformation 11, 423-430.

Huang, C., Davis, L.S., Townshend, J.R.G., 2002. An assessment of support vector machines for land cover classification. International Journal of Remote Sensing 23, 725-749.

Huvenne, V.A.I., Blondel, P., Henriet, J.P., 2002. Textural analyses of sidescan sonar imagery from two mound provinces in the Porcupine Seabight. Marine Geology 189, 323-341.

Ibisch, P.L., Mérida, G., 2004. Biodiversity: the richness of Bolivia. State of knowledge and conservation. Editorial Fan, Santa Cruz de la Sierra, Bolivia.

Kavzoglu, T., Colkesen, I., 2009. A kernel functions analysis for support vector machines for land cover classification. International Journal of Applied Earth Observation and Geoinformation 11, 352-359.

Kayitakire, F., Hamel, C., Defourny, P., 2006. Retrieving forest structure variables based on image texture analysis and IKONOS-2 imagery. Remote Sensing of Environment 102, 390-401.

Killeen, T.J., Calderon, V., Soria, L., Quezada, B., Steininger, M.K., Harper, G., Solórzano, L.A., Tucker, C.J., 2007. Thirty years of land-cover change in Bolivia. Ambio 36, 600-606.

Killeen, T.J., E., E.G., Beck, 1993. Guía de árboles de Bolivia. Herbario Nacional de Bolivia. Herbario Nacional de Bolivia, La Paz, and Missouri Botanical Garden, St. Louis.

Killeen, T.J., Guerra, A., Calzada, M., Correa, L., Calderon, V., Soria, L., Quezada, B., Steininger, M.K., 2008. Total historical land-use change in eastern Bolivia: Who, where, when, and how much? Ecology and Society 13, art36.

Lambin, E.F., 1999. Monitoring forest degradation in tropical regions by remote sensing: some methodological issues. Global Ecology and Biogeography 8, 191-198.

Lombardo, U., Prümers, H., 2010. Pre-Columbian human occupation patterns in the eastern plains of the Llanos de Moxos, Bolivian Amazonia. Journal of Archaeological Science 37, 1875-1885.

Lu, D., Mausel, P., Brondízio, E., Moran, E., 2003. Classification of successional forest stages in the Brazilian Amazon basin. Forest Ecology and Management 181, 301-312.

Lu, D., Weng, Q., 2007. A survey of image classification methods and techniques for improving classification performance. International Journal of Remote Sensing 28, 823-870.

Lu, D.S., Mausel, P., Batistella, M., Moran, E., 2004. Comparison of land-cover classification methods in the Brazilian Amazon Basin. Photogrammetric Engineering and Remote Sensing 70, 723-731.

Lucas, R.M., Honz, aacute, k, M., Curran, P.J., Foody, G.M., Milne, R., Brown, T., Amaral, S., 2000. Mapping the regional extent of tropical forest regeneration stages in the Brazilian Legal Amazon using NOAA AVHRR data. International Journal of Remote Sensing 21, 2855-2881.

Maffi, L., 2005. Linguistic, cultural and biological diversity. Annual Review of Anthropology 29, 599-617.

Maillard, P., 2003. Comparing Texture Analysis Methods through Classification. Photogrammetric Engineering and Remote Sensing 69, 357-368

Malhi, Y., Roberts, J.T., Betts, R.A., Killeen, T.J., Li, W., Nobre, C.A., 2008. Climate change, deforestation, and the fate of the Amazon. Science 319, 169-172.

Malhi, Y., Román-Cuesta, R.M., 2008. Analysis of lacunarity and scales of spatial homogeneity in IKONOS images of Amazonian tropical forest canopies. Remote Sensing of Environment 112, 2074-2087.

Mann, C.C., 2008. Ancient earthmovers of the Amazon. Science 321, 1148-1152.

Marsik, M., Stevens, F.R., Southworth, J., 2011. Amazon deforestation: Rates and patterns of land cover change and fragmentation in Pando, northern Bolivia, 1986 to 2005. Progress in Physical Geography 35, 353-374.

Mather, P.M., 2004. Computer Processing of Remotely-Sensed Images: An Introduction. John Wiley & Sons Ltd, England.

Messina, J., Walsh, S., Mena, C., Delamater, P., 2006. Land tenure and deforestation patterns in the Ecuadorian Amazon: Conflicts in land conservation in frontier settings. Applied Geography 26, 113-128.

Mountrakis, G., Im, J., Ogole, C., 2011. Support vector machines in remote sensing: A review. ISPRS Journal of Photogrammetry and Remote Sensing 66, 247-259.

Navarro, G., Maldonado, M., 2002. Geografía ecológica de Bolivia: Vegetación y ambientes acuáticos. Centro de Ecología Simón I. Patiño. Cochabamba, Bolivia.

Nobre, C.A., Borma, L.S., 2009. 'Tipping points' for the Amazon forest. Current Opinion in Environmental Sustainability 1, 28-36.

Ota, T., Mizoue, N., Yoshida, S., 2011. Influence of using texture information in remote sensed data on the accuracy of forest type classification at different levels of spatial resolution. Journal of Forest Research 16, 432-437.

Pal, M., Mather, P.M., 2005. Support vector machines for classification in remote sensing. International Journal of Remote Sensing 26, 1007-1011.

Palà, V., Pons, X., 1995. Incorporation of Relief in Polynomial Based Geometric Corrections. Photogrammetric Engineering and Remote Sensing 61, 935-944.

Palubinskas, G., Lucas, R.M., Foody, G.M., Curran, P.J., 1995. An evaluation of fuzzy and texture-based classification approaches for mapping regenerating tropical forest classes from Landsat-TM data. International Journal of Remote Sensing 16, 747 - 759.

Pons, X., 2000. MiraMon. Geographic Information System and Remote Sensing software. Centre de Recerca Ecològica i Aplicacions Forestals, CREAF. Bellaterra (Barcelona).

Pons, X., Solé-Sugrañes, L., 1994. A simple radiometric correction model to improve automatic mapping of vegetation from multispectral satellite data. Remote Sensing of Environment 48, 191-204.

Porter-Bolland, L., Ellis, E.A., Gholz, H.L., 2007. Land use dynamics and landscape history in La Montaña, Campeche, Mexico. Landscape and Urban Planning 82, 198-207.

Richards, J.A., Jia, X., 2006. Remote Sensing Digital Image Analysis. Springer-Verlag Berlin Heidelberg, Germany.

Riou, R., Seyler, F., 1997. Texture analysis of tropical rain forest infrared satellite images. Photogrammetric Engineering and Remote Sensing 63, 515-521.

Saatchi, S.S., Soares, J.V., Alves, D.S., 1997. Mapping deforestation and land use in amazon rainforest by using SIR-C imagery. Remote Sensing of Environment 59, 191-202.

Santos, J.R., Mura, J.C., Paradella, W.R., Dutra, L.V., Goncalves, F.G., 2008. Mapping recent deforestation in the Brazilian Amazon using simulated L-band MAPSAR images. International Journal of Remote Sensing 29, 4879-4884.

Schulz, J.J., Cayuela, Echeverria, C., Salas, J., Rey Benayas, J.M., 2010. Monitoring land cover change of the dryland forest landscape of Central Chile (1975–2008). Applied Geography 30, 436-447.

Serra, P., Pons, X., Saurí, D., 2003. Post-classification change detection with data from different sensors: Some accuracy considerations. International Journal of Remote Sensing 24, 3311-3340.

Serra, P., Moré, G., Pons, X., 2005. Application of a hybrid classifier to discriminate Mediterranean crops and forests. Different problems and solutions. XXII International Cartographic Conference (ICC2005), A Coruña, Spain, pp. 1-8.

Stehman, S.V., 1997. Selecting and interpreting measures of thematic classification accuracy. Remote Sensing of Environment 62, 77-89.

Szuster, B.W., Chen, Q., Borger, M., 2011. A comparison of classification techniques to support land cover and land use analysis in tropical coastal zones. Applied Geography 31, 525-532.

Tsaneva, M.G., Krezhova, D.D., Yanev, T.K., 2010. Development and testing of a statistical texture model for land cover classification of the Black Sea region with MODIS imagery. Advances in Space Research 46, 872-878.

Tso, B., Mather, P.M., 2009. Classification methods for remotely sensed data. CRC Press, Taylor & Francis Group, Boca Raton.

Turner, B.L., II, Lambin, E.F., Reenberg, A., 2007. The emergence of land change science for global environmental change and sustainability. Proceedings of the National Academy of Sciences of the United States of America 104, 20666-20671.

Vieira, I.C.G., de Almeida, A.S., Davidson, E.A., Stone, T.A., Reis de Carvalho, C.J., Guerrero, J.B., 2003. Classifying successional forests using Landsat spectral properties and ecological characteristics in eastern Amazônia. Remote Sensing of Environment 87, 470-481.

Walsh, S.J., Shao, Y., Mena, C.F., McCleary, A.L., 2008. Integration of Hyperion Satellite Data and a Household Social Survey to Characterize the Causes and Consequences of Reforestation Patterns in the Northern Ecuadorian Amazon. Photogrammetric Engineering and Remote Sensing 74, 725-736

Watanachaturaporn, P., Arora, M.K., Varshney, P.K., 2008. Multisource classification using support vector machines: an empirical comparison with decision tree and neural network classifiers. Photogrammetric Engineering and Remote Sensing 74, 239-246.

Wijaya, A., Gloaguen, R., 2007. Comparison of multisource data support vector Machine classification for mapping of forest cover. Geoscience and Remote Sensing Symposium, 2007. IGARSS 2007. IEEE International, pp. 1275-1278.

Woodcock, C.E., Collins, J.B., Gopal, S., Jakabhazy, V.D., Li, X., Macomber, S., Ryherd, S., Harward, J.V., Levitan, J., Wu, Y., Warbington, R., 1994. Mapping forest vegetation using Landsat TM imagery and a canopy reflectance model. Remote Sensing of Environment 50, 240-254.

Woodcock, C.E., Gopal, S., 2000. Fuzzy set theory and thematic maps: accuracy assessment and area estimation. International Journal of Geographical Information Science 14, 153-172.

Zhang, C., Franklin, S.E., Wulder, M.A., 2004. Geostatistical and texture analysis of airborne-acquired images used in forest classification. International Journal of Remote Sensing 25, 859-865.

Zhang, Y., 1999. Optimisation of building detection in satellite images by combining multispectral classification and texture filtering. ISPRS Journal of Photogrammetry and Remote Sensing 54, 50-60.

Zhang, Y., 2001. Texture-integrated classification of urban treed areas in high-resolution color-infrared imagery. Photogrammetric Engineering and Remote Sensing 67, 1359-1365.