GROUND TRUTH METHOD ASSESSMENT FOR SVM-BASED LANDSCAPE CLASSIFICATION

Robin Pouteau, Benoit Stoll, Sébastien Chabrier

GROUND TRUTH METHOD ASSESSMENT FOR SVM-BASED LANDSCAPE CLASSIFICATION

R. Pouteau, B. Stoll, S. Chabrier

South Pacific Geosciences (GePaSud) Laboratory - French Polynesia University (UPF)
BP 6570, 98702 FAA'A - TAHITI - French Polynesia

ABSTRACT

Researches on land cover classification have a complete lack of ground truth methodology description. We propose a method to track ecotones as privileged training areas for SVM-based natural vegetation classification. This guidance method combines (i) the construction of a principal component analysis (PCA) on spectral bands and gray level co-occurence matrix texture attributes calculated on very high resolution images and (ii) the use of the Sobel's edge detection algorithm on this PCA. The experiment is successfully applied with an overall accuracy of 82 %. Using SVM, a minimum number of mixed pixels is necessary but they can help significantly in locating an appropriate hyperplane. Moreover, the presented results show that the training stage could be more influential on classifier accuracy than classifiers themselves.

Index Terms— Ground truth, support vector machines (SVM), maximum likelihood, vegetation, classification.

1. INTRODUCTION

Understanding spatial organization of natural and human structures in the context of global changes is a modern challenge. Detailed land cover thematic mapping is a simplified view for complex objects and a key tool for decision makers. Remotely sensed data becomes a powerful instrument to monitor landscapes, the integration level of management decisions.

Over the past four decades, classification of remote sensing imagery is developed, initially from signal processing methods (e.g. maximum likelihood estimation (MLC)). Then, the development of remote sensing data with increasing spectral and spatial resolution and the enhanced computer processing capability have lead to the development of many new classifying techniques to map more precisely land covers. Numerous comparative studies come to the consensus that support vector machines (SVM) [1] are presently one of the most efficient classifiers [2].

SVM are a semi-supervised method and need thus adapted training sets to be optimally functional. Nevertheless, in order to compare classifiers objectively, training sets have to be specified and adapted. Moreover, the dependence of the classifier to the training sets suggests that they are a key point to outperform the present classification accuracies.

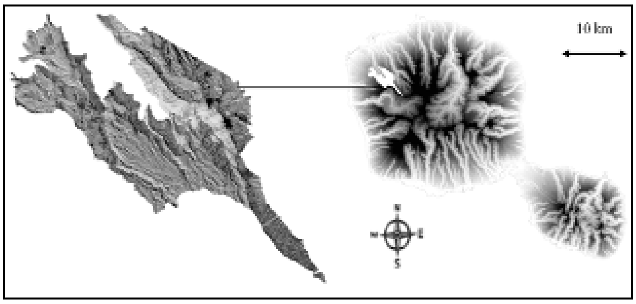

Figure 1: The study site is mont Marau (left) located in the northwest side of Tahiti (right).

This essay has two major objectives (i) while the nature of an ideal training set is not clear, to explore this key stage (ii) to suggest and apply a generic ground truth method to train efficiently our classifications

2. MATERIAL AND METHODS

The study site is located in mont Marau, at the northwest of Tahiti, French Polynesia (Fig.1). Mont Marau is a threatened area of exceptional ecological value [3]. In mountainous areas such as Pacific volcanic islands, access is limited and resources are difficult to evaluate in situ. That's why remotely sensed data represents a helpful tool for eco-environmental monitoring.

A four channels (R, G, B, NIR) multispectral Quickbird scene from 2003 is used for the analyses. This very high resolution image allows computing efficient analysis on texture [4] to help species discrimination.

By nature, accuracy of a supervised classification depends highly on the quantity and quality of the data used in the learning and assessment steps. The chosen classifier accuracy may thus be impacted by the used training set.

In a large majority of studies, pure pixels are used for the SVM training stage. Nevertheless, Lesparre and Gorte [5] denote that mixed pixels can be successfully used for the

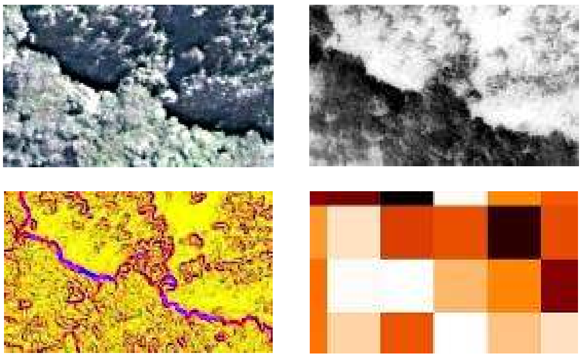

Figure 2: RGB composition of a subset of the original Quickbird image (up left), PCA (up right), edge detection (bottom left) and pixel purity detection (bottom right) calculated on the PCA. Images represent an ecotone between a *Pinus caribaea* (coniferous) forest and a *Falcataria moluccana* (decidiuous) one.

training of a maximum likelihood classifier and consider that this facililitates the estimation of the spectra of the pure classes using mixed pixels, provided that the mixture proportions are known. In the same way, Foody and Mathur [6] show that the use of small training sets containing mixed pixels as boundaries between agricultural fields improves SVM classification. They consider that unlike the conventional classifiers, the aim of SVM training is not to describe accurately the classes but to provide information that will help fitting the classification decision boundaries - the hyperplanes - to separate them. Such boundaries are privileged training areas thanks to the aggregation of information on pixels [7].

3.1. Training set size

Training set size has to maximize accuracy without increasing needlessly ground sampling and computational times. For example, [8] state that the classification overall accuracy (OA) achieved by SVM is affected by the size of the training data set, as noted in the case of other classifiers. This behaviour could be related to the capability of the training pixels to adequately represent the characteristics of their respective classes. As the number of training pixel increases, SVM find pixels that better define inter-class discriminating surfaces. We formulate the hypothesis that the classification accuracy and the number of training pixels are not linearly correlated because of their redundancy.

3.2. Pure pixels tracking

See Fig.2. Training sets are usually made out of pure pixels characterizing homogenous areas. Pure pixels selection is simply carried out calculating a purity index. The calculation is processed on a principal component analysis (PCA) which is computed on the 4 spectral bands and 6 gray level co-

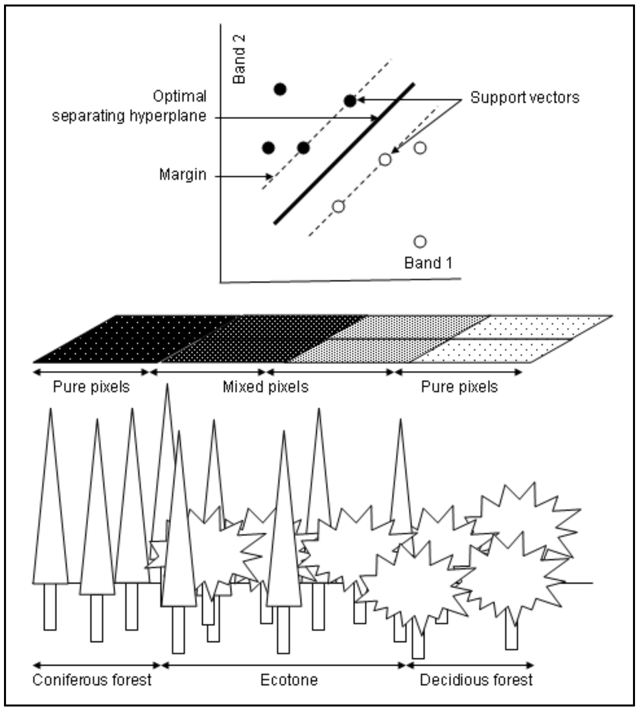

Figure 3: A simplified view of the location of pure and mixed class spectral responses in a feature space. Mixed pixels from ecotones are more capable of being support vectors than pure ones.

occurence matrix (GLCM) texture attributes calculated on a 9x9 pixels window for each band. The intrinsic purity coefficient of pixel aggregates is calculated in aggregation surface of 450 m².

3.3. Mixed pixels tracking

See Fig.2. Because satellite data are spatially correlated, ecotones (the transition area between two adjacent but different plant communities) are privileged areas of mixed spectral responses (Fig.3). Ecotones are often hard to distinguish in imagery when vegetal communities are a complex mosaic of different taxa. Moreover it is difficult to quantify their magnitude. Consequently, ecotones as landscape breakings are located using the Sobel's edge detection algorithm [9] since it proves to have a good accuracy/confusion trade-off [10].

3.4. Ground truth sampling method

Transects are drawn at 100 m in parallel to a road. When an area is detected i.e. when the transect crosses a very pure area or a marked ecotone, one homogeneous region of interest (ROI) is sampled for the pure pixels sampling method or two ROI, in each side of the detected ecotone and in the transect direction, for the mixed pixels sampling method. ROI surface is 450 m² because, according to phyto-

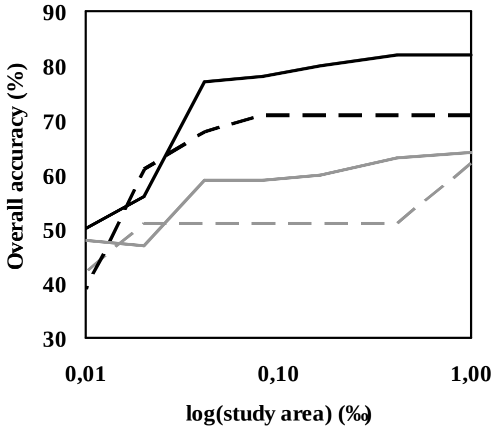

Figure 4: OA as a function of the training set area ; 0.01‰=250 pixels and 1‰=25 000 pixels. MLC with pure pixels is represented in grey dashed line, MLC with mixed pixels in black dashed line, SVM with pure pixels in grey continuous line and SVM with mixed pixels in black continuous line ; n=7 points for each line.

sociologists, the minimal area (designated as the smallest area which can contain an adequate representation of a species association) for tree and shrub communities is more than 400 m² [11].
Validation set is composed by 25,000 pixels - i.e. ~ 1 ‰ of the mont Marau area - equally distributed.

4. RESULTS

Fig.4 shows that the training pixels area - or the number of pixels - is not linearly correlated with the OA, reaching a plateau. Mixed pixels constitute the best training set for both classifiers. If classification schemes give similar results for the smallest training sets i.e. ~ 500 pixels, significant differences appear when this minimal size is outreached. This remark doesn't agree with [6], considering that the use of mixed pixels allows the use of smaller training sets in the set size range they studied. The most interesting point is the synergy between SVM and the training method based on mixed pixels areas.

Comparing classifiers in Table 1, OA and Kappa coefficients are sensibly improved by the use of the SVM face to the familiar Gaussian maximum likelihood classifier for both sampling methods. The same observation can be found in [12] and [13].

On the other hand, difference between sampling methods for the same classifier are significantly bigger than the inter-classifiers one which clearly shows the importance of training sets comparing classifiers. The optimal OA of 82% was obtained combining SVM and mixed pixels-based training set. Ecotones are thus very informative training areas providing aggregated spectral response of two different classes. Such a training set is particularly capable

Table 1: OA and Kappa coefficients of the different classification schemes

	MLC		SVM	
	Pure pixels	Mixed pixels	Pure pixels	Mixed pixels
OA (%)	62	71	64	82
Kappa	0.49	0.62	0.53	0.77

to bring adjancent classes support vectors closer, optimizing the fitting of the separating hyperplane. Our results corroborate observations of [5] and [6] and validate the proposed method.

5. DISCUSSION

Foody and Mathur mentioned that judgemental sampling is often viewed unfavourably and avoided. The proposed method has the advantage to be objective. Another benefit of the proposed technique is that the chosen training mixed pixels are the most separable ones because they are selected in function of their change rate (algorithm of Sobel).

This study proves that the use of mixed pixels is efficient in complex systems such as tropical landscapes, where intrusive vegetation cover boundaries represent a large area misclassified with conventional sampling methods. Sampling at the ecotones level consists to "show" to the SVM the most complex spectral situations to guarantee the effectiveness of their classification, hypothesizing that trained in difficult situations, the algorithm will classify easily the simplest cases. Moreover, training on pure pixels doesn't allow knowing precisely where limits between two classes are localized. In contrary, sampling at the ecotones level allows to compel the definition of each class in situ.

ACKNOWLEDGEMENT

The authors are grateful to the Government of French Polynesia and its Urbanism Department for the provision of the Quickbird data.

REFERENCES AND NOTES

[1] V. Vapnik and A. Chervonenkis, *Statistical learning theory*. New York: Wiley, 1998.
[2] B. Waske and J. A. Benediktsson, "Fusion of support vector machines for classification of multisensor data, " *IEEE Trans. Geosci. Remote Sens.*, vol.45, pp. 3858-3866, 2007.
[3] J. Y. Meyer, "La biodiversité du mont Marau en péril," Technical Report, Délégation à la recherche, Papeete, French Polynesia, 2004.

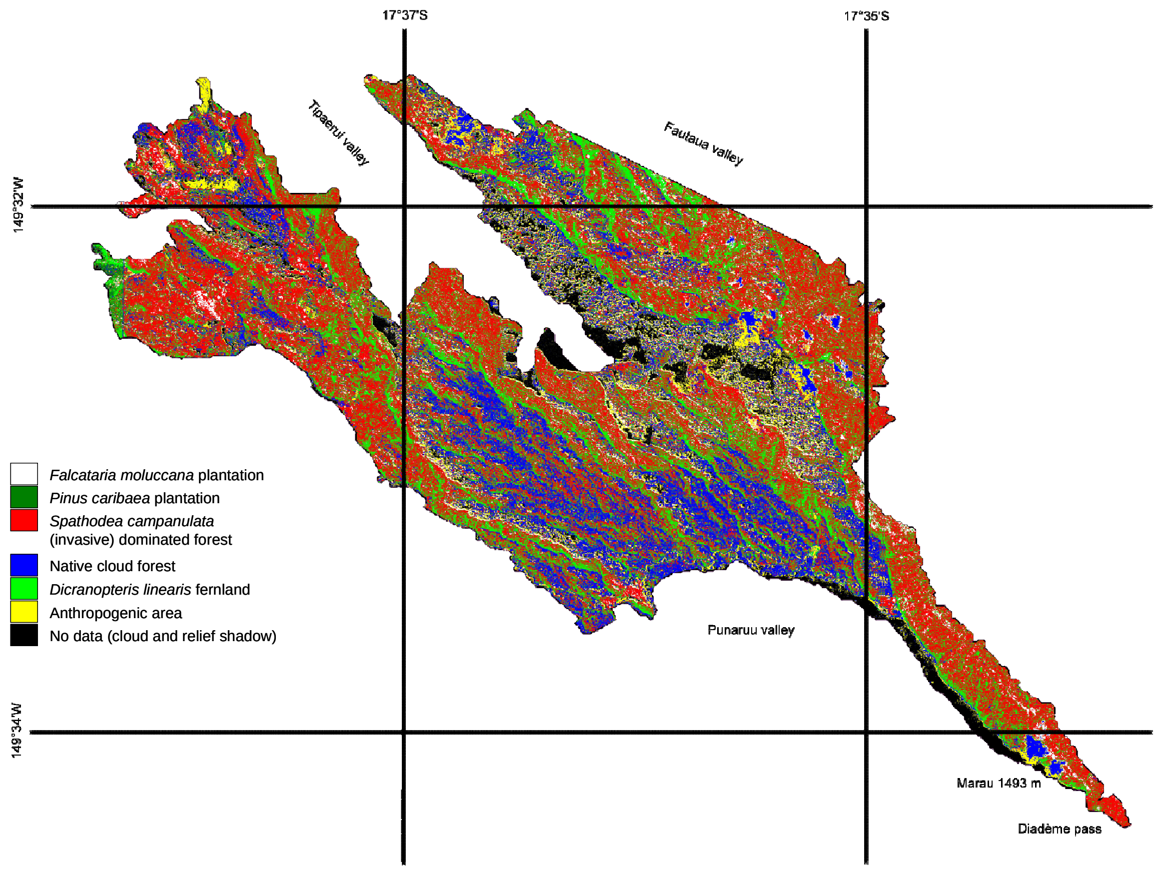

Figure 5: Result of the SVM classification trained with mixed pixels.

[4] R. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification", *IEEE Trans. Systems Man Cybernet*, , vol. 3, pp. 610-621, 1973.

[5] J. Lesparre, and B. G. H. Gorte, "Using mixed pixels for the training of a maximum likelihood classification," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 34, Part XXX, 2006.

[6] G.M. Foody, and A. Mathur, "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*, vol. 103, pp. 179-189, 2006.

[7] W. Turner, S. Spector, N. Gardiner, M. Fladeland, E. Streling, and M. Steininger, "Remote sensing for biodiversity science and conservation, " *TRENDS in Ecology and Evolution*, vol. 18, pp. 306-314, 2003.

[8] M. Pal, and P.M. Mather, "An assessment of effectiveness of decision tree methods for land cover classification," *Remote Sensing of Environment*, vol. 86, pp. 554-565, 2003.

[9] I.E. Sobel, "Camera models and machine perception", PhD thesis, Standford Univ., 1970.

[10] M. Jeon, M. Alexander, W. Pedrycz, and N. Pizzi, "Unsupervised hierarchical image segmentation with level set and additive operator splitting," *Pattern Recognition Letters*, vol. 26, pp. 1461-1469, 2005.

[11] F. Gillet, *La phytosociologie synusiale intégrée. Guide méthodologique*, Documents du Laboratoire d'Ecologie végétale, Institut de Botanique, Univ. de Neuchâtel., 2000.

[12] S. Fukuda, and H. Hirosawa, "Support Vector Machine Classification of Land Cover: Application to Polarimetric SAR Data," *IEEEInternational Geosci. Remote Sens. Symposium*, vol. 1, pp. 187-189, 2001.

[13] Jin Shuying, Li Deren, and Wang Jingwen, "A comparison of support vector machine with maximum likelihood classification algorithms on texture features," *IEEE Trans. Geosci. Remote Sens.*, vol. 5, pp. 3717-3720, 2005.