MAPPING NATURA 2000 HEATHLAND IN BELGIUM – AN EVALUATION OF ENSEMBLE CLASSIFIERS FOR SPACEBORNE ANGULAR CHRIS/PROBA IMAGERY

Jonathan C-W Chan[1], *Pieter Beckers*[2], *Frank Canters*[1], *Toon Spanhove*[3], *Jeroen Vanden Borre*[3] *and Desiré Paelinckx*[3]

[1] Cartography & GIS research group, Geography Department, Vrije Universiteit Brussel, Pleinlaan 2, Brussels, Belgium
[2] Division of Geography, Katholieke Universiteit Leuven, Celestijnenlaan 200E, B-3001 Leuven-Heverlee
[3] Research Institute for Nature and Forest (INBO), Kliniekstraat 25, 1070 Brussels

ABSTRACT

Natura 2000 is an ecological network of protected areas in the territory of the European Union (EU). With the introduction of the Habitats Directive in 1992, EU member states are obligated to report every six years the status of the Natura 2000 habitats so that better conservation policy can be formulated. This paper examines the use of angular hyperspectral CHRIS/Proba image for the mapping of heathland at a Belgian Natura 2000 site. We find that the use of angular images increases the overall classification rate as compared to using only the nadir image; with the incorporation of angular images the final mapping is also more homogenous with less salt and pepper effect. While the class accuracy of Calluna- and Erica-dominated heathlands are still low, class accuracy of Molinia-dominated heathland is generally more encouraging. Two tree-based ensemble classifiers, Random Forest (RF) and Adaboost, were compared with Support Vector Machines (SVM). When only the nadir image was used, SVM attained the highest accuracy. When angular images were included, all three classifiers obtained comparable accuracies though in general RF and Adaboost had faster training time. We also adopted an assessment approach which repeats the accuracy assessment in ten independent trials, instead of the common practice of having only one trial. Our results show that accuracy attainment can vary significantly among different trials and hence it is recommendable to have more than one trial in order that a more objective characterization of the classifiers is obtained. [1]

***Index Terms*—** Natura 2000, heathland, hyperspectral, CHRIS/Proba, angular images, ensemble classifiers, Random Forest, Adaboost, SVM

1. INTRODUCTION

With the implementation of the Habitats Directive in 1992, EU member states committed themselves to protect a range of highly threatened habitats within their territory. Monitoring and reporting on the status of the so-called 'Natura 2000 habitats' is an essential part of an effective conservation, and an important obligation under the Habitats Directive: every six years, member states have to report on the actual area, the range, the quality and the future prospects for each habitat type. In recent years, the conservation value of the semi-natural heaths has become much more appreciated. Even though heathland of western Europe are relatively species poor as compared to their counterparts in the southern hemisphere, these complex ecosystems have become habitats for a unique fauna and flora which adapted to the particular biotic and abiotic conditions [1]. Because of this typical biodiversity and ecological value of heathlands, they were included on the Annex I of the EU Habitats Directive and protected as Natura 2000 habitats [2][3].

Favourable	Non-favourable

Fig.1. Picture impressions of dry sand heaths with Calluna and Genista (2310) in favourable and non-favourable conditions.

Heath is a dwarf-shrub vegetation mainly found on poor, acidic soils. A typical heath ecosystem can be described as a mosaic of dwarf-shrubs, grass-dominated heath, bare soil and isolated shrubs or trees. European heath vegetations typically have a few dominant plant species. In the north of Belgium, heathland habitats are often Erica- or Calluna-dominated. In good conservation status, they show a complex structural variation consists of a mixture of dwarf scrub, open sand and patches of pioneer grasses and mosses, which is a prerequisite for many rare and specialized plant and animal species (Fig. 1, Left). In non-favourable conditions, encroachment with purple moor grass (Molinia

[1] The authors wish to thank the Belgian Science Policy Office for funding this research under the STEREO II HABISTAT project.

caerulea) leads to a monotonous vegetation with a heavily reduced ecological value (Fig. 1, Right).

To date, the gross of the data needed for conservation status assessment are gathered through field surveys and the visual interpretation of aerial photos. Such an approach does, however, have some major drawbacks. First, these labour-intensive techniques are highly expensive. Second, field mapping is often slow, making it difficult to cover vast areas during the optimal season of inventory. Third, despite the existence of strict rules for field mapping, inter-observer errors are an issue [4]. Remote sensing techniques are often suggested as a proper alternative for this monitoring. In order to obtain adequate information on the diverse ecological and biological conditions at ecotope level, hyperspectral data with rich spectral information will be an effective choice [5].

The Compact High Resolution Imaging Spectrometer (CHRIS) is an imaging spectrometer carried on board a space platform called Proba (Project for On Board Autonomy). It acquires within the spectral region of 0.4-1 μm with up to 62 bands in mode 1 at a spatial resolution of 34 m. At mode 3, the sensor acquires images at a higher spatial resolution of 18m, but the spectral band number is reduced to 18. With the multi-angle scanning property of CHRIS, it is also possible to employ angular images for the mapping of heathlands. The results will provide interesting reference for future hyperspectral sensors with angular acquisition capacity such as the EnMap mission operated by the German Space Agency.

Many different classifiers have been proposed for remote sensing data. Support Vector Machines (SVM) is a very popular classifier used frequently on hyperspectral imagery [6][7]. Though not as widely used as classifiers like SVM, tree-based ensemble classifiers such as Random Forest (RF) [8] and Adaboost [9] are reportedly effective and accurate in classifying various remote sensing data [5][7][10-13]. The most notable advantages of these machine learning methods are that they are simple to use and robust to noise or missing values which can sometimes happen with remote sensing data. Limited need for parameter setting and fast training are also important attributes. It would be interesting therefore to investigate the performance of these ensemble classifiers for the mapping Natura 2000 habitats. In this study, we focused on RF and Adaboost and their performance will be compared to the benchmark classifier SVM.

2. METHODOLOGY

2.1. Classification scheme

A classification scheme with ten classes was adapted for this study: (1) Calluna, (2) Erica, (3) Molinia, (4) Grassland, (5) Coniferous forest, (6) Deciduous forest, (7) Bare sand and mosses, (8) water with vegetation, (9) water without vegetation, and (10) Cropland. Although these classes are inspired by the Habitats Directive, they do not completely correspond to the Natura 2000 habitat types. Class 1 is Calluna-dominated dry heathland, close to the Natura 2000 habitats encoded 2310 and 4030. Class 2, Erica-dominated heathland, belongs to wet heathland, Natura 2000 class 4010. Class 3 is Molinia-encroached heath, a degraded form of heathland with low ecological value which can originate from inland dunes (2330) as well as from dry (2310, 4030) and wet (4010) heathland. Class 6 is deciduous forest, potentially belonging to Natura 2000 class 9190. Classes 8 and 9 are oligotrophic water bodies; they could be either Natura 2000 classes 3110, 3130 or 3160. Coniferous forest and other land cover types such as urban and agricultural land do not correspond to any habitat listed in the Habitats Directive. After considering the geo-spatial distribution of each class, a total of 586 sampling points evenly distributed within the study area were selected.

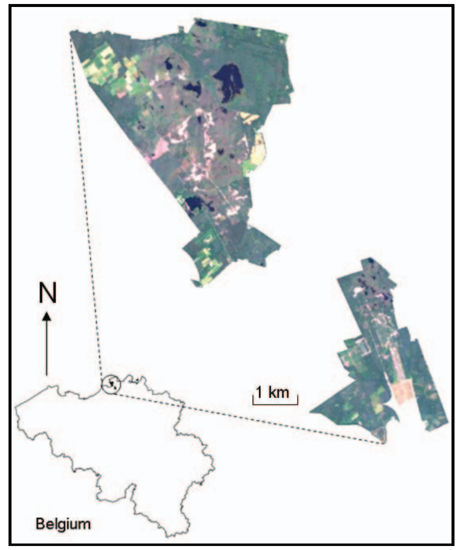

Fig. 2. Location of Kalmthoutse Heide in northern Belgium

2.2. Accuracy assessment

Accuracy assessment of a classification is usually carried out through cross-tabulating the predicted classes and the reference classes, thus building a confusion matrix. Initially, a full set of reference data will be divided into a training set and a test set. For a supervised classification, the training pixels are used to build (train) a classifier and then this classifier is used to classify the blind test pixels. The procedure of separating the training and the testing pixels is normally done in a random stratified manner where for each class the number in training and in testing is more or less equal. However, bias could happen by chance through this random separation process leading to extreme results. To further investigate this issue, we decided to have more elaborated experiments by repeating the procedure in ten different trials. The results would shed light on the sensitivity of each classifier to this random procedure and provide a better characterization of each classifier.

To understand if angular measurements are useful for mapping heath habitats, the experiments were further divided into using only the nadir images, and with off-nadir images incorporated. Accuracies for both cases were compared. In addition to overall accuracy and kappa values, mean class accuracy was also calculated to show the strength of a classifier over the separation of all classes.

2.3. Study area

Our study area is Kalmthoutse Heide, a Natura 2000 site in northern Belgium (51°22'N, 4°27'E) (Fig. 1). It consists mainly of dry and wet heathland, inland dunes, water bodies and vast forests. The whole area measures 3750 ha (9375 acres) and is cut through by the Belgian-Dutch border. A CHRIS image of excellent quality was acquired at Mode 3 (18 bands at 18m resolution) on 1st July 2008. The image was atmospherically corrected and de-noised using the BEAM toolbox freely available through the ESA website. Since the off-nadir images at ±55° have serious geometrical distortion, only the nadir and the ±36° images were used.

3. RESULTS AND DISCUSSION

3.1. Parameter tuning

For RF, there are two parameters: the number of trees and the number of random variables at each split. The best settings were found using the Out-of-bag (OOB) estimate [5]. The AdaBoost algorithm used is AdaBoost.M1, which is a generalization of AdaBoost for more than two classes. The algorithm has been implemented in the R software environment [14], which uses classification trees as base learners. The type of trees is similar to the classification trees described in [15]. More details of these trees can be found in the R package of 'rpart' [16]. Two parameters are important for this algorithm: the number of trees (iterations) and how deep the tree is going to grow. The evaluation was done using a 5-fold cross-validation. In terms of the number of trees, initial experiments have shown that a number between 30 -120 is adequate. The depth is given as an integer, with 0 being the root node. In [14], it is suggested the number to be equal to the number of classes.

SVMs separate two classes by maximizing the margin between the classes' closest points. Often, data points are projected into a higher-dimensional space via kernel techniques to enable a linear separability. Tuning of SVM involves the trials of different kernel functions such as linear, polynomial, sigmoid and radial basis functions (RBF) Comparatively, RBF gave better results and was much faster with two main parameters: gamma and cost.

3.2. Accuracy performance

The overall accuracy averaged over ten independent runs for RF, Adaboost and SVM are 57.1%, 57.5% and 61.8%, respectively. SVM outperformed the two ensemble classifiers by around 4.5%. The averaged mean class accuracy for RF, Adaboost and SVM are 51.4%, 51.2% and 54.6%, respectively. The difference between the two ensemble classifiers and SVM is a little more than 3%.

	RF		**Adaboost**		**SVM**	
	Only Nadir	3 angles	Only Nadir	3 angles	Only Nadir	3 angles
Calluna	38.7	43.5	41.1	44.9	44.1	45.1
Erica	41.9	48.1	44.2	47.3	37.3	37.7
Molinia	69.7	74.9	70.4	73.6	80.7	79.0

Table 1. Accuracy (%) of healthland classes using angular images.

The purpose of having ten independent trials is to investigate the possible sensitivity of a classifier to the random process of separating the ground truth samples into the training and test sets. It is observed that the gap between the lowest and the highest accuracy attainments is substantial. For RF, the min-max accuracy range is 51.7%-62.9%; it is 53.9%-61.1% for Adaboost and 58.4%-65.5% for SVM. Comparatively, the gap with SVM is the smallest, but still around 7%. These results imply that, for accuracy assessments using a similar approach, it will be more informative to repeat the procedure more than once. An assessment based on only one trial may lead to biased conclusions as the accuracy could be very poor or very good simply by chance.

Including angular images adds 2x18=36 bands to the input. Using the same training and test sampling points as in the experiment of nadir image, ten independent trials were run. After off-nadir images were incorporated, comparable accuracies were obtained from all three classifiers. Notably, accuracy of SVM did not increase with the additional angular images. The reason for that is not thoroughly investigated. However, it could be related to feature selection which is often done with sophisticated algorithms to reduce computing time and improve accuracy [6]. On the average, all classifiers achieved accuracy at 61%. The average 'mean class accuracy' over ten trials was around 54% for all three classifiers. The accuracies obtained from the classifiers are very similar at each trial with marginal differences of only 1-2%. The difference between the minimum and maximum accuracies among the ten trials is 7% for RF, 9% for Adaboost, and almost 10% for SVM. Again, it shows that variation among different trials can be quite large.

3.3. Accuracy and mapping of heathland types

Table 1 shows the accuracy of three target heathland classes. On the average over ten trials, SVM has the highest accuracy for Calluna and Molinia as compared with the two ensemble classifiers. However, SVM's accuracy with Erica is 10% lower. It is also observed that using angular input has brought significant increase in classification accuracy of these heathland classes. While the accuracy of Calluna and Erica is still quite low, the accuracy of Molinia is approaching the 80% level showing the potential of mapping this important ecotope. Fig. 3 is a comparison of

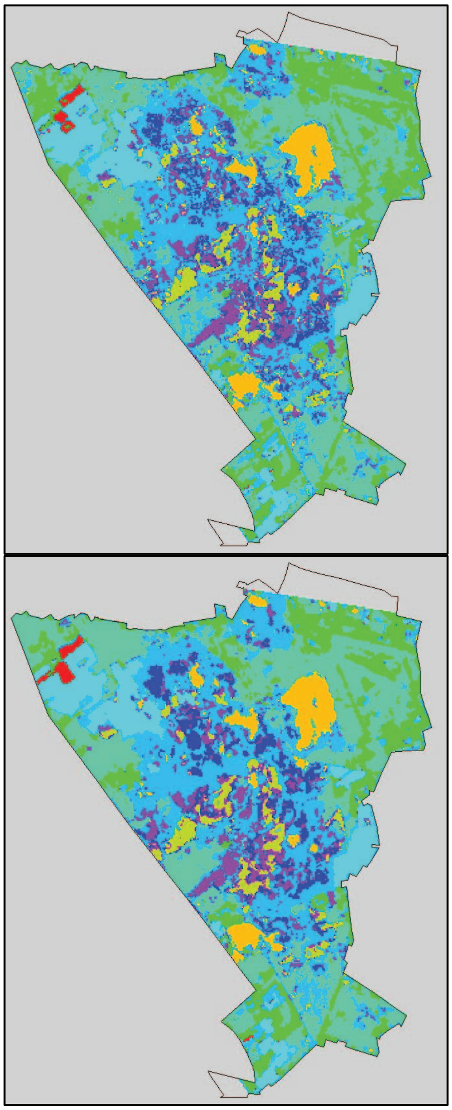

Fig. 3.Mapping with angular CHRIS (Below) shows better formation of patches and with less salt and pepper effect as mapping using only the nadir image (Above).

the classification maps generated by RF with the top representing the result using only the nadir image and the bottom using all angular images. The map produced with angular images shows clearly a better formation of patches with less pepper and salt effect.

4. CONCLUSIONS

The use of angular hyperspectral CHRIS images shows significant improvements in mapping of Natura 2000 heathland habitats. Other than increasing accuracy, it also enhances mapping result with better patch formation and less pepper and salt effect. Though accuracy with Erica and Calluna is still low, classification of Molinia is reaching 80%. The results generated from ten trials of accuracy assessment show that more than one trial is necessary to provide an objective evaluation of the classifiers.

5. REFERENCES

[1] Y.L. Dupont, and B. Overgaard Nielsen, "Species composition, feeding specificity and larval trophic level of flower visiting insects in fragmented versus continuous heathlands in Denmark," *Biol. Conserv.* 131: 475–485, 2006.

[2] N.R. Webb, "The traditional management of European heathlands," *J. of App. Ecol.* 35, 987–990, 1998.

[3] A. Ssymank, U. Hauke, C. Rückriem, and E. Schröder, "Das europäische Schutzgebietssystem NATURA 2000 – BfNHandbuch zur Umsetzung der Fauna-Flora-Habitat-Richtlinie (92/43/WG) und der Vogelschutz-Richtlinie (79/409/EWG)", *Schriftenr. Landschaftspfl. und Natursch* 53, 1–560, 1998.

[4] J.P. Stevens, T.H. Blackstock, E.A. Howe and D.P. Stevens, "Repeatability of Phase 1 habitat survey," *J. of Env. Mgt.* 73: 53-59, 2004.

[5] J.C.-W. Chan and D. Paelinckx, "Evaluation of Random Forest and Adaboost tree-based ensemble classification and spectral band selection for ecotope mapping using airborne hyperspectral imagery," *Remote Sens. of Env.* 112, 2999-3011, 2008.

[6] D. Vyas, N.S.R. Krishnayya, K.R. Manjunath, S.S. Ray and S. Panigrahy, "Evaluation of classifiers for processing Hyperion (EO-1) data of tropical vegetation," *Int. J. of Appl. Earth Obs. & Geoin.* 13, 228-235, 2011.

[7] B. Waske, J.A. Benediktsson, K. Árnason and J.R. Sveinsson, "Mapping of hyperspectral AVIRIS data using machine learning algorithms," *Can. J. of Remote Sens.* 35, 106-116, 2009.

[8] L. Breiman, "Random Forest," *Machine Learning* 45, 5-32, 2001.

[9] Y. Freund and R.E. Schapire, "Experiments with a new boosting algorithm," Proc. of the 13th Int. Conf. on Machine Learning (pp. 148-156). Bari, Italy. 1996

[10] J. Ham, Y.C. Chen, M.M. Crawford and J. Ghosh, "Investigation of the random forest framework for classification of hyperspectral data," *IEEE Trans. Geos. & Remote Sens.* 43, 492-501, 2005.

[11] M. Pal and P.M. Mather, "An assessment of the effectiveness of decision tree methods for land cover classification," *Remote Sens. of Env.* 86, 554-565, 2003.

[12] J.C.-W. Chan, C. Huang and R.S. DeFries, "Enhanced algorithm performance for land cover classification from remotely sensed data using bagging and boosting," *IEEE Trans. of Geos. & Remote Sens.* 39, 693-695, 2001.

[13] R.S. DeFries and J.C.-W. Chan, "Multiple criteria for evaluating machine learning algorithms for land cover classification from satellite data," *Remote Sens. of Env.* 74, 503-515, 2000.

[14] E.A. Cortes, M.G. Martinez and N.G. Rubio, *Package 'adabag': applies AdaBoost.M1 and Bagging, CRAN project*, 2007. (http://cran.r-project.org/web/packages/adabag/index.html)

[15] L. Breiman, J.H. Friedman, R.A. Olshen and C.J. Stone, *Classification and regression trees*. Monterey, Calif., U.S.A.: Wadsworth, Inc. 1984.

[16] T.M. Therneau and B. Atkinson. *Package 'rpart': recursive partitioning and regression trees, CRAN project*. 2010. (http://cran.r-project.org/web/packages/rpart/index.html)