

# USING PATCH METRICS AS VALIDATION FOR CONTEXTUAL CLASSIFICATION OF HEATHLAND VEGETATION

Guy Thoonen<sup>a,\*</sup>, Koen Hufkens<sup>b</sup>, Jeroen Vanden Borre<sup>c</sup>, and Paul Scheunders<sup>a</sup>

<sup>a</sup> IBBT, Vision Lab, University of Antwerp, Universiteitsplein 1, B-2610 Antwerpen, Belgium - guy.thoonen@ua.ac.be

<sup>b</sup> Department of Geography & Environment, Boston University, 675 Commonwealth Avenue, Boston, MA, 02215

<sup>c</sup> Research Institute for Nature and Forest, Kliniekstraat 25, B-1070 Brussel, Belgium

**KEY WORDS:** Hyper spectral, Vegetation, Classification, Accuracy, Spatial, Contextual, Hierarchical, Quality

## ABSTRACT:

This article presents a method to assess the accuracy of vegetation maps for which contextual information has been included in the classification process. It is well known that land use classification may benefit from combining spatial and spectral information. Consequently, many classification techniques incorporating spatial information have been implemented. To compare various contextual classification techniques, the shape of vegetation patches, in the spatially enhanced maps, are statistically linked to their counterparts in the spectral classification result, on which these spatial enhancements are applied. To this end, measures for the change in shape of patches are introduced. The shape of any patch is characterized by the edges between the patch and its neighbors. Therefore, patch shape can be represented by an edge map in which each pixel gets the value of the number of classes that are different from the class label of the central pixel in a four-adjacency neighborhood. Rather than defining a single metric for the edge map difference, an error matrix is used to depict not only how many edges have changed with respect to the reference, but also by how much they have changed. The method is tested on contextual classification results of heathland vegetation.

## 1 INTRODUCTION

Elements such as habitat loss, climate change and invasive alien species are important causes of the current biodiversity crisis. Many Western European heathland sites are under pressure by this threat. The goal of the Habitat directive of the European Union is to protect rare or endangered habitats or species. One of the specific measures is the foundation of the Natura 2000 network (EEC, 1992), an ecological network of protected areas, spread over the whole continent.

Nature conservation for these areas is the responsibility of the European member states, each of which needs to take appropriate measures to bring and maintain each site in a good conservation status. Moreover, the member states are committed to report on the status of the Natura 2000 sites, habitats and species on a regular basis. This way, it is possible to keep track of the trends and take conservation measures whenever appropriate.

Such a commitment requires a thorough knowledge of the habitats of every Natura 2000 site, both in terms of the distribution of the habitat patches, as in terms of the quality of each patch. Traditional methods, such as field surveys and visual interpretation of aerial photographs, cannot fulfill these requirements by themselves. They are expensive, time consuming and are strongly subject to inter-surveyor errors.

Remote sensing can prove to be an important tool to aid in habitat monitoring. Unfortunately, few readily applicable procedures have been developed so far (Keramitsoglou et al., 2005, Kobler et al., 2006). One reason for this is that habitats are usually not homogeneous vegetation patches of a single or a few dominant species. Instead, they show a high variety in facies at different scale levels. At a large scale, the facies of the same habitat may differ between regions as a result of climatic or soil conditions. But also at a very fine scale, most habitats are in fact intricate mixtures of different land cover types.

To assess this gap, a framework has been developed to deal with the complete trajectory, from breaking down habitats into land cover types, to reconstructing the habitat types and their conservation status from land cover classification results. This framework has been developed for Western European heathlands and is part of a Belgian multidisciplinary project on habitat status monitoring, called Habistat (Haest et al., n.d.).

Within the Habistat project, a hierarchical classification scheme has been constructed, reflecting the structural dependencies that are present in the field. This scheme allows for incorporating contextual information in the classification process. To this end, classification is performed using a tree that resembles the hierarchy of the classification scheme, and modeling the spatial dependencies as Markov Random Fields (MRF) (Geman and Geman, 1984), one for each node of the tree.

However, due to the nature of the ground reference data, which is limited to isolated pixels in homogeneous areas, validating this technique with conventional methods, like resubstitution, hold-out or cross-validation, provides little information on transitions between vegetation patches, and their spatial extent.

To address this issue, a method has been developed to link the shape of vegetation patches, in any classification result that uses spatial information, statistically to their counterparts in the spectral classification results that these spatial methods stem from. The effects of using spatial information can be pictured by introducing measures for the change in shape of edges between patches with respect to the spectral classification result.

Using these measures together with conventional methods, the result from the hierarchical MRF classification is compared to the result of a simple smoothing technique. This comparison shows, as expected, that both methods provide smoothing, yet, at the same time, the MRF result preserves patch shapes much better than the simple smoothing result.

Section 2 describes the contextual classification framework and the use of Tree-Structured Markov Random Fields. Section 3

---

\*Corresponding author.

introduces the patch-based validation method. Section 4 shows experimental results on heathland data. Finally, section 5 presents the conclusions and elements for future research.

## 2 CONTEXTUAL CLASSIFICATION

### 2.1 Framework

Within the project framework, the Natura 2000 habitats have been translated into land cover types through a hierarchical classification scheme with 4 levels of detail. The first level comprises only 6 classes: heathland, grassland, forest, sand dunes, water bodies and arable fields. Level 2 and 3 mostly determine specific habitat types of the Natura 2000 programme, containing 11 and 17 classes respectively. Level 4 comprises 24 classes, focusing on vegetation structural elements that determine the conservation status of the habitat types.

The unique way in which the data have been subdivided offers the freedom of classifying on a per-level basis, but also allows us to classify hierarchically, one level after another. In addition, in search for obtaining more readily interpretable vegetation maps, including textural/contextual features in the classification process has been an important consideration.

It is well known that land use classification may benefit from combining spatial and spectral information. Consequently, many classification techniques incorporating spatial information have been implemented.

In general, there are several strategies for including spatial information in the classification process, among others:

- Performing classification after first segmenting an image (Akçay and Aksoy, 2008, Driesen et al., 2009).
- Including spatial information in the feature vectors together with the spectral information (Fauvel et al., 2008).
- Incorporating spatial information as prior information in a Bayesian classification framework and modeling the dependencies between neighboring pixels as Markov Random Fields (MRF) (Berthod et al., 1996).

In (Poggi et al., 2005) a recursive supervised segmentation algorithm based on a tree-structured Markov Random Field (TS-MRF) is proposed. A binary tree is constructed, exploiting the hierarchical structure exhibited in the image. The leaves of this tree correspond to the end classes. The TS-MRF model describes a K-ary label field by means of a sequence of binary MRFs, each one corresponding to a node in the tree. The hierarchical structure of the heathland data makes this technique especially interesting.

### 2.2 Tree-Structured Markov Random Fields

In the following, we assume  $y$  to be a data point and  $x$  to be a class label out of a total of  $K$  classes. The applied classification method is a supervised segmentation algorithm based on a Tree-Structured Markov Random Field Model (TS-MRF). A random label field  $X$  defined on a lattice  $\mathcal{S}$  is an MRF with respect to a neighborhood system  $\eta$ , if

$$\begin{aligned} p(x) &> 0, \forall x \in X, \text{ and} \\ p(x_s | x - x_s) &= p(x_s | x_{\eta(s)}) \end{aligned} \quad (1)$$

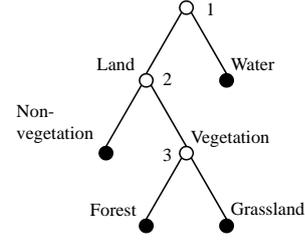


Figure 1: Example Tree.

for each site  $s$ , with  $\eta(s)$  the neighbors of  $s$ . The distribution of an MRF corresponds to a Gibbs form (Li, 2001)

$$p(x) = \frac{1}{Z} \exp[-U(x)] = \frac{1}{Z} \exp\left[-\sum_{c \in \mathcal{C}} V_c(x_c)\right] \quad (2)$$

where  $Z$  is a normalizing constant,  $U(x)$  is the energy,  $c$  is a clique, i.e., a set of neighboring sites, and the  $V_c(\cdot)$  are called potentials.

If, for sake of simplicity, the binary Potts model is used and only two-site cliques are considered, then the clique potentials are a function of the clique site, given by

$$V_c(x_c) = \begin{cases} \beta^{(x_c, x_s)}, & \text{if } x_c \neq x_s \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

with  $\beta^{(k,l)} > 0$  the edge-penalty parameters for all pairs of classes  $k$  and  $l$ .

The TS-MRF model describes a  $K$ -ary label field by means of a sequence of binary MRFs, each one corresponding to a node  $t$  in the tree  $T$ . As a result, rather than using a single set of parameters which are estimated based on the whole image, multiple sets of parameters become possible, each set adapted to the separation between two classes. Fig. 1 shows an example tree with 4 end classes.

In general, given a binary tree structure with  $K$  terminal nodes, and consequently  $K$  classes, the number of internal nodes, and therefore the number of parameters, when using the Potts model, will be  $K - 1$ , rather than  $(1/2)(K - 1)K$ .

Assuming the  $K$  classes follow a Gaussian distribution for which estimates of the means and the covariance matrices are known, the algorithm iteratively optimizes the realization  $x^t$  and the value of  $\beta^t$  for each node  $t$  of the tree, each time starting from a purely spectral segmentation. In general, the algorithm looks as follows:

1. initialization: node  $t = 1$ , initialize  $\mathcal{S}^t$ , the set of locations classified as the class corresponding to node  $t$ , as  $\mathcal{S}$ , the original image
2. estimation:  $\forall s \in \mathcal{S}^t$ , compute  $p(y_s | x_s^t = l(t))$  and  $p(y_s | x_s^t = r(t))$ , with  $l(t)$  and  $r(t)$  the left and right child of  $t$ , using Maximum Likelihood Estimation (MLE).
3. optimization: find the estimates of  $x^t$  and  $\beta^t$  as

$$(x^t, \beta^t) = \arg \max_{(x^t, \beta^t)} p(x^t | \beta^t) p(y | x^t) \quad (4)$$

by iteratively fixing  $\beta^t$  while computing  $x^t$  and vice versa until convergence, starting from  $\beta^t = 0$ .

4. segmentation:  $S^{l(t)}$  is the set of locations classified as  $l(t)$ ,  $S^{r(t)}$  the set of locations classified as  $r(t)$ , if  $t < K - 1$ , then  $t = t + 1$  and goto step 2, else exit.

Although, for simplicity, we described the TS-MRF technique using the binary Potts model, it is possible to generalize this to the K-ary Potts model. Moreover, it is possible to build a TS-MRF by integrating several reference MRF models, not necessarily of the same type nor just binary. Specifically, in this work, the applied tree is not binary and is no longer determined during classification. Rather, it is built from the structural dependencies which follow naturally from the hierarchy, that is already present in the ground reference data (Thoonen et al., n.d.).

### 3 PATCH METRICS VALIDATION

Traditional ways of assessing the accuracy of classification results involve using the confusion matrix (Congalton and Green, 1999). A confusion matrix is a square array of numbers set out in rows and columns that expresses the number of sample units assigned to a particular category in a classified map relative to the number of sample units assigned to a particular category in a reference map. The accuracy obtained in this way is called the thematic or labeling accuracy.

Moreover, the confusion matrix is a very effective representation of map accuracy because the accuracies of each map category are described individually. A popular statistical measure for the agreement of two classification results is the Kappa coefficient (Bishop et al., 1977).

However, collecting the reference samples is tedious and expensive. As a result, more often than not, the reference samples are sparse and do not model the shapes of structures in the scene. In particular, when using classification methods that include contextual information, changes in shapes (i.e. changes in transitions between classes) can be very subtle with respect to a pure spectral classification result.

In order to compare thematic maps resulting from spatial classification techniques, another reference and another methodology is needed in addition to traditional references and methods.

#### 3.1 Reference

In (Baraldi et al., 2005), the references are multiple unsupervised cluster maps. In (Persello and Bruzzone, 2010), reference objects are defined by photo interpretation, using the high resolution property of the investigated images.

An alternative strategy is to take the pure spectral classification result that the contextual methods are closest related to as a reference. ‘Closest related’ means that the contextual methods use the same spectral classifier in all cases.

When using this type of reference, the assumption is made that the desired shapes are present in the spectral classification result, but at the same time, they are clouded with noise.

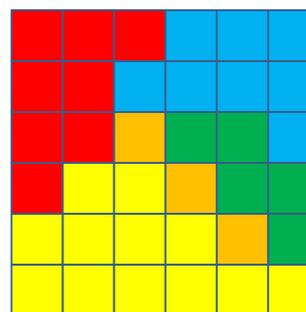
With respect to the pure spectral classification result, the desired properties of the contextual classification techniques are:

- Removing the noise, i.e. increasing the thematic accuracy, as measured by traditional methods.
- Conserving the shape information, as measured by an alternative quality assessment method.

#### 3.2 Methodology

As thematic accuracy is covered by traditional methodology, the focus is on how to assess shapes in an appropriate fashion. A way is to consider that a thematic map consists out of patches, collections of connected pixels with the same class label. The shape of any of these patches is characterized by the edges between the patch and its neighbors.

As a consequence, the shapes of the patches can be represented by an edge map, for instance a 4-neighbor edge map, in which each pixel gets the value of the number of classes that are different from the class label of the central pixel in a four-adjacency neighborhood, as shown in Fig. 2.



(a) Patches

0	0	1	1	0	0
0	1	2	1	1	0
0	2	4	2	1	1
1	1	1	2	1	1
1	0	0	1	2	2
0	0	0	0	1	1

(b) 4-neighbor edge map

Figure 2: A number of patches and the corresponding 4-neighbor edge map

It follows that the 4-neighbor edge maps of both the classification result and the reference can be compared to indicate how well the shape information has been conserved.

In (Baraldi et al., 2005) the 4-neighbor edge map, alternatively defined as a map in which each pixel gets the value of the number of four-adjacency pixels that do not have the same label as the central pixel, was used to determine the spatial fidelity, by calculating the mean and standard deviation of the absolute difference between the edge map of the classification result and the unsupervised reference cluster maps.

However, the alternative definition presented here provides valuable information on how much each edge changes: Small changes in patch shapes lead to smaller changes in absolute values of the edge map. For instance, when an isolated single-pixel patch in a homogeneous region disappears, its edge map value changes from 1 to 0.

Additionally, a suitable way of representing this information is by using a confusion matrix, allowing to extract more information than possible from a single metric. The categories in this confusion matrix are no longer vegetation classes, but one of the five possible values the 4-neighbor edge map can take. Not only

does the confusion matrix indicate which edges remain the same and which don't, but it also indicates how drastic the change is.

#### 4 EXPERIMENTS AND RESULTS

For the study site of Kalmthoutse Heide, the airborne hyperspectral data were obtained in June 2007 with an AHS sensor with a ground resolution of approximately 2.5m. The range of 450nm-2550nm is covered by 63 spectral bands. Fig. 3 shows a part of the Kalmthoutse Heide image data.

During summer 2007, ground reference data were collected in homogeneous plots of 10 meters diameter. The vegetation data of the sampled plots, approximately 1200 in total, were analyzed and plots were grouped in the 4-level hierarchical classification scheme.

For the following analysis, only the plots in the nature reserve areas of Kalmthoutse Heide were used. Furthermore, not all plots were categorized up to level 4. These limitations leave us with a total of 678 plots for the 18 remaining level 4 categories. Following a hold-out strategy, 50% of these plots were per class randomly selected for training, while the remaining 50% were used to assess the thematic accuracy.



Figure 3: Part of the Kalmthoutse Heide image data

To make a fair comparison, all classification results have been obtained using the same spectral classifier, a Support Vector Machine (SVM) classifier with a Radial Basis Function kernel. The techniques under comparison are

- a classification using the TS-MRF technique with the extended tree that resembles the classification scheme and a  $3 \times 3$  window (or second-order neighborhood).
- a pixel-based classification followed by a simple smoothing done by a majority analysis (MA), in which a pixel gets the class of the majority of its direct neighbors. Window sizes of  $3 \times 3$ ,  $5 \times 5$  and  $7 \times 7$ , respectively, are used.

Fig. 4 shows a part of the TS-MRF classification result. The shades of purple represent heathland classes, yellow colors are sand dunes, dark green are forest types while light green are grasslands and blue represents water.

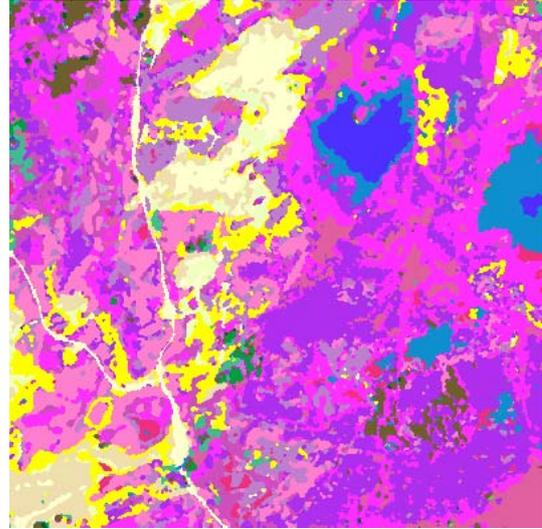


Figure 4: TS-MRF classification result

#### 4.1 Thematic Accuracy

Table 1 shows the overall accuracy and the Kappa coefficient for the various techniques. All three of the simple majority analysis results show a higher thematic accuracy than the TS-MRF result, that also shows an increase with respect to the spectral classification result. However, visually, the result from the majority analysis using a  $7 \times 7$  window does not look promising, as shown in Fig. 5. Indeed, the thematic accuracy does not provide a lot of information on the spatial fidelity.

Classification	Overall Accuracy	Kappa coefficient
Spectral	85.88%	0.8457
TS-MRF	86.77%	0.8553
MA $3 \times 3$	89.34%	0.8835
MA $5 \times 5$	89.60%	0.8863
MA $7 \times 7$	89.01%	0.8798

Table 1: Thematic Accuracy

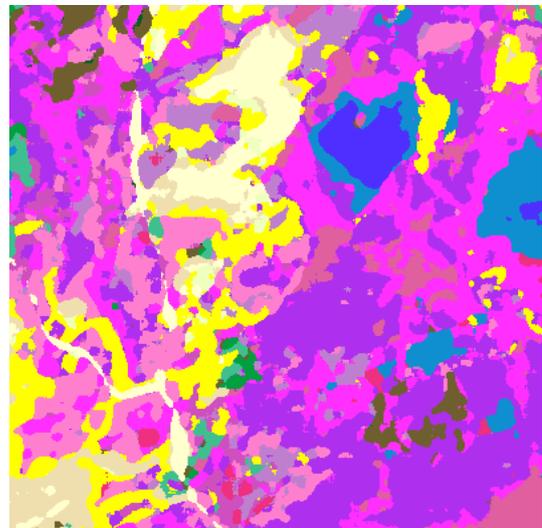


Figure 5: Majority analysis with a  $7 \times 7$  window

## 4.2 Spatial Fidelity

To measure the spatial fidelity, confusion matrices are constructed for the 4-neighbor edge maps, as defined in section 3.2, using the spectral classification result as the reference map.

Table 2 shows the percentage confusion matrix for the TS-MRF result and Table 3, 4 and 5 show the same matrix for the Majority Analysis results with window sizes of  $3 \times 3$ ,  $5 \times 5$  and  $7 \times 7$  respectively.

	0	1	2	3	4	
0	95.9%	68.6%	41.9%	23.5%	13.2%	64.0%
1	4.0%	29.7%	43.4%	41.3%	33.5%	28.5%
2	0.1%	1.6%	13.9%	27.3%	32.7%	6.6%
3	0.01%	0.04%	0.8%	7.5%	16.7%	0.9%
4	0%	0.003%	0.02%	0.3%	4.0%	0.07%
	100%	100%	100%	100%	100%	100%

Table 2: Edge percentage confusion matrix of the TS-MRF result

	0	1	2	3	4	
0	53.7%	50.8%	50.5%	50.0%	49.5%	51.2%
1	36.7%	38.9%	38.9%	38.8%	38.8%	38.4%
2	8.6%	9.2%	9.5%	10.0%	10.4%	9.2%
3	1.0%	1.1%	1.1%	1.2%	1.2%	1.1%
4	0.05%	0.05%	0.06%	0.06%	0.05%	0.05%
	100%	100%	100%	100%	100%	100%

Table 3: Edge percentage confusion matrix of the MA  $3 \times 3$  result

	0	1	2	3	4	
0	68.7%	67.3%	67.0%	66.8%	66.4%	67.5%
1	27.2%	28.5%	28.6%	28.7%	29.0%	28.2%
2	3.8%	3.9%	4.1%	4.3%	4.4%	4.0%
3	0.3%	0.3%	0.3%	0.3%	0.3%	0.3%
4	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%
	100%	100%	100%	100%	100%	100%

Table 4: Edge percentage confusion matrix of the MA  $5 \times 5$  result

	0	1	2	3	4	
0	76.5%	75.8%	75.4%	75.3%	75.1%	75.8%
1	21.3%	22.0%	22.2%	22.2%	22.5%	21.9%
2	2.1%	2.2%	2.3%	2.3%	2.3%	2.2%
3	0.1%	0.1%	0.1%	0.1%	0.2%	0.1%
4	0.002%	0.003%	0.003%	0.003%	0%	0.003%
	100%	100%	100%	100%	100%	100%

Table 5: Edge percentage confusion matrix of the MA  $7 \times 7$  result

Edge value '0' indicates areas with no edges, the homogeneous areas. Most of these areas are kept unchanged in the TS-MRF result. Furthermore, the results for edge value '1' are more or less comparable across all the techniques. This is to be expected, as it comprises the isolated-pixel areas, that are easily filtered away by all the techniques in the experiment.

For the more complex edges, it is clear that, not only does the TS-MRF result have the highest values on the diagonal, moreover, the values that have changed are closer to the diagonal than in the majority analysis case. Therefore, the changes in edge complexity are less drastic. The rows for the majority analysis results are more or less constant, no matter what window size was selected.

## 5 CONCLUSIONS AND FUTURE WORK

A qualitative method was introduced to assess the accuracy of thematic maps that were constructed incorporating spatial information in the classification process. The method is built upon the selection of an alternative reference, the spectral classification result that the spatial enhancements are related to, and an alternative

methodology, focusing on measures for the change in shapes of edges between the patches.

The method was tested on real data of heathland vegetation. On this data, a recursive supervised segmentation algorithm, based on a Tree-structured Markov Random Field (TS-MRF) was applied as one of the contextual methods.

The quality assessment confirms what can be expected from intuition: the TS-MRF technique provides smoothing, but, at the same time, conserves edges better than methods like, for instance, majority filtering.

The method presented here is not intended for stand-alone use. It only provides information about the spatial fidelity of a target map. Therefore, it should be used together with conventional methods that measure thematic accuracy. Naturally, it is possible to use the presented method together with another type of reference map.

For future work, investigations should be done on how to summarize the information drawn from the edge map confusion matrices (for instance a weighted Kappa coefficient). More alternative contextual methods should be compared using the technique to explore the potential.

## ACKNOWLEDGEMENTS

This work is supported by the Belgian Science Policy Office under the framework of the STEREO programme project HABI-STAT (contract SR/00/103). The partners involved in the project are: Flemish Institute for Technological Research (VITO), Vrije Universiteit Brussel (VUB), Research Institute for Nature and Forest (INBO), Alterra and University of Antwerp (Vision lab).

## REFERENCES

- Akçay, H. and Aksoy, S., 2008. Automatic detection of geospatial objects using multiple hierarchical segmentations. *IEEE Transactions on Geoscience and Remote Sensing* 46(7), pp. 2097–2111.
- Baraldi, A., Bruzzone, L. and Blonda, P., 2005. Quality assessment of classification and cluster maps without ground truth knowledge. *IEEE Transactions on Geoscience and Remote Sensing* 43(4), pp. 857–873.
- Berthod, M., Kato, Z., Yu, S. and Zerubia, J., 1996. Bayesian image classification using markov random fields. *Image and Vision Computing* 14(4), pp. 285–295.
- Bishop, Y., Fienberg, S. and Holland, P., 1977. *Discrete Multivariate Analysis: Theory and Practice*. MIT Press, Cambridge, MA.
- Congalton, R. and Green, K., 1999. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*. CRC/Lewis Press, Boca Raton, FL.
- Driesen, J., Thoonen, G. and Scheunders, P., 2009. Spatial hyperspectral image classification by prior segmentation. In: *Geoscience and Remote Sensing Symposium, 2009 IEEE International, IGARSS 2009*, Vol. 3, pp. 709–712.
- EEC, 1992. Council directive 92/43/EEC of 21 May 1992 on the conservation of natural habitats and of wild fauna and flora. *Official Journal, L* 206 pp. 7–50.
- Fauvel, M., Benediktsson, J., Chanussot, J. and Sveinsson, J., 2008. Spectral and spatial classification of hyperspectral data using svms and morphological profiles. *IEEE Transactions on Geoscience and Remote Sensing* 46(11), pp. 3804–3814.

Geman, S. and Geman, D., 1984. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Trans. Pattern Anal. Machine Intell.* 6(6), pp. 721–741.

Haest, B., Delalieux, S., Chan, J.-W., Hufkens, K., Kooistra, L., Ma, J., Schmidt, A., Thoonen, G., Vanden Borre, J., Canters, F., Kempeneers, P., Knaeps, E., Múcher, S., Paelinckx, D. and Scheunders, P., n.d. A classification framework for habitat status reporting with remote sensing methods. In preparation.

Keramitsoglou, I., Kontoes, C., Sifakis, N., Mitchley, J. and Xofis, P., 2005. Kernel based re-classification of earth observation data for fine scale habitat mapping. *Journal for Nature Conservation* 13(2-3), pp. 91–99.

Kobler, A., Deroski, S. and Keramitsoglou, I., 2006. Habitat mapping using machine learning-extended kernel-based reclassification of an ikonos satellite image. *Ecological Modelling* 191(1), pp. 83–95.

Li, S., 2001. *Markov Random Field Modeling in Image Analysis*. second edn, New York: Springer-Verlag.

Persello, C. and Bruzzone, L., 2010. A novel protocol for accuracy assessment in classification of very high resolution images. *IEEE Transactions on Geoscience and Remote Sensing* 48(3), pp. 1232–1244.

Poggi, G., Scarpa, G. and Zerubia, J. B., 2005. Supervised segmentation of remote sensing images based on a tree-structured mrf model. *IEEE Transactions on Geoscience and Remote Sensing* 43(8), pp. 1901–1911.

Thoonen, G., Hufkens, K., Vanden Borre, J., De Backer, S. and Scheunders, P., n.d. Mapping the conservation status of heathland vegetation using hyperspectral remote sensing in a hierarchical classification framework. Submitted to *IEEE Transaction on Geoscience and Remote Sensing*.