

## Land cover update by supervised classification of segmented ASTER images

A. R. S. MARÇAL\*, J. S. BORGES, J. A. GOMES and J. F. PINTO DA COSTA  
Faculdade de Ciências, Universidade do Porto, Departamento de Matemática Aplicada,  
Rua do Campo Alegre, 687, 4169-007 Porto, Portugal

(Received 2 December 2003; in final form 1 June 2004)

The revision of the 1995 land cover dataset for the Vale do Sousa region, in the northwest of Portugal, was carried out by supervised classification of a multi-spectral image from the Advanced Spaceborne Thermal Emission and Reflectance Radiometer (ASTER) sensor. The nine reflective bands of ASTER were used, covering the spectral range from 0.52–2.43  $\mu\text{m}$ . The image was initially ortho-rectified and segmented into 51 186 objects, with an average object size of 135 pixels (about 3 ha). A total of 582 of these objects were identified for training nine land cover classes. The image was classified using an algorithm based on a fuzzy classifier, Support Vector Machines (SVM), K Nearest Neighbours (K-NN) and a Logistic Discrimination (LD) classifier. The results from the classification were evaluated using a set of 277 validation sites, independently gathered. The overall accuracy was 44.6% for the fuzzy classifier, 70.5% for the SVM, 60.9% for the K-NN and 72.2% for the LD classifier. The difficulty in discriminating between some of the forest land cover classes was examined by separability analysis and unsupervised classification with hierarchical clustering. The forest classes were found to overlap in the multi-spectral space defined by the nine ASTER bands used.

### 1. Introduction

Up-to-date knowledge of land use and land cover is an important tool for the various planning authorities with responsibilities for the management of territory at a regional level. This is particularly true for forest planners, as forests have a considerably higher temporal variability. This variability can be due to normal events, such as re-forestation and vegetation growth, or due to external causes, such as forest fires. The forest fires can be natural events or caused by human activities. Remote sensing data can be a valuable tool for an efficient update of land cover/land use maps. This is in fact one of the most common uses of multi-spectral satellite images. For this type of application, the spatial resolution of the sensor is an important parameter with direct implications on the scale of the derived maps. For scales in the range 1 : 25 000 to 1 : 50 000, a satellite image with a pixel size of about 15–30 m is adequate. There are numerous examples of land cover maps derived from multi-spectral satellite images from Landsat Thematic Mapper (TM) (Laba *et al.* 1997, Colby and Keating 1998) and Satellite pour l'Observation de la Terre (SPOT) High Resolution Visible (HRV) (Gao and Skillcorn 1998). Both these sensors have been providing high spatial resolution multi-spectral data for nearly 20 years. More

---

\*Corresponding author. Email: andre.marcal@fc.up.pt

recently, the Advanced Spaceborne Thermal Emission and Reflectance Radiometer (ASTER) sensor has become a valuable alternative, with an increased spatial and spectral resolution. The costs involved in using satellite images for land cover updates have been greatly reduced in the last few years, due to changes in data distribution policies and the developments in image processing equipment, both in hardware and software. However, the production of accurate and meaningful land cover maps is still a laborious task, mainly due to the need for identification of the actual land cover in selected locations in the field.

The land cover maps produced from the satellite imagery are usually imported into a Geographic Information System (GIS). One of the problems encountered in this process has to do with the pixel resolution of the satellite image, which is not a natural representation of the terrain. The most used approach is to classify each image pixel as an independent observation, regardless of its spatial context. A hardwoods forest parcel for example might have most of its pixels assigned to that class, but a few pine trees might result in a number of pixels in that parcel assigned to a different class. This is not a satisfactory result for the end user point of view. As a consequence, land cover maps produced from satellite image classification need to be extensively edited before they can be imported into a GIS. An alternative approach to the usual pixel-by-pixel classification is to perform image segmentation. The term segment refers to a cluster of adjacent pixels that represent a meaningful object on the terrain, from the user point of view (Geneletti and Gorte 2003). The multi-spectral image is first segmented into a set of objects that are then classified according to their average signal in each of the spectral bands.

The objective of this work was to test the ability of a number of supervised classification methods on a real situation, where the conditions are actually not very favourable. Another objective was to produce updated land cover maps for the Vale do Sousa region from ASTER data segmented into objects. An evaluation of the performance of each of the classification methods used was carried out using a set of validation sites. An evaluation of the similarities between the spectral signatures of the land cover classes was also performed, using hierarchical clustering techniques on unsupervised classification data.

## **2. Study area and datasets**

### **2.1 *The study area***

The Vale do Sousa region (River Sousa valley) is located in northwest Portugal, and is crossed by the rivers Sousa and Douro in the south (figure 1). The region has an area of 764 km<sup>2</sup> (about 300 square miles) and ground altitudes ranging from 100 to 800 m above sea level. About two thirds of the region is covered by forests or uncultivated soil. The uncultivated areas are mainly due to the numerous forest fires. The main forest types are eucalyptus and areas with mixed eucalyptus and pine (CNIG 1990). There is also agricultural and many small urban areas. The land use is generally small private fields (5 ha or less) with the forest areas in much larger portions of the land. The existing land cover dataset was produced by photointerpretation in 1995. On this land cover dataset only areas above 0.5 ha were classified. This dataset was out-of-date and not very reliable. The production of an updated land cover dataset is helpful for forest planning, allowing, for example, for

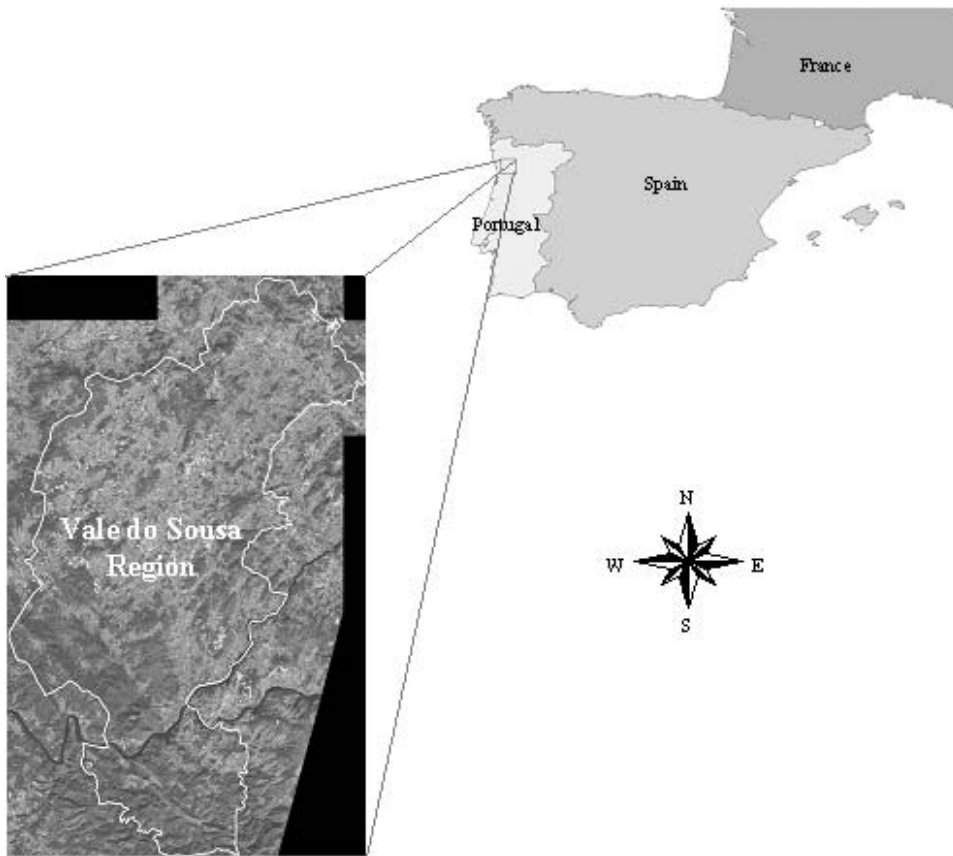


Figure 1. Location of the Vale do Sousa region in the northwest of Portugal. The detached area is 30.9 km by 50.1 km, and shows an RGB colour composite (bands 2/3/1) of the ASTER mosaic file used; 24 October 2001, 11:42.

the detection of burned areas and the subsequent undesirable spread of species like eucalyptus.

## 2.2 ASTER data selection and acquisition

The ASTER sensor provides multi-spectral images of the Earth with each image granule covering roughly 60 km by 60 km on the ground. The ASTER images have 14 spectral bands, three in the visible and near-infrared (VNIR instrument, 0.52–0.86  $\mu\text{m}$ ), six in the short-wave infrared (SWIR instrument, 1.60–2.43  $\mu\text{m}$ ) and five in the thermal infrared (TIR instrument, 8.12–11.65  $\mu\text{m}$ ) (Yamaguchi *et al.* 1998). The spatial resolution is 15 m (VNIR), 30 m (SWIR) and 90 m (TIR). Two ASTER image granules were collected, covering the north (Path 205, Row 90, ID 408566) and south (Path 205, Row 91, ID 408567) of the interest area. These two image granules were acquired consecutively, at 11:42, 24 October 2001, with a view angle of 7°. Both images are nearly 100% cloud free and combined provide complete coverage of the whole Vale do Sousa region. Only the VNIR and SWIR image bands were used, as the TIR data has a much lower spatial resolution and should not provide useful information about vegetation cover.

### 2.3 Geometric correction of the ASTER data

A grid of reference points with latitude and longitude is provided on the header files of ASTER data Level 1B. The accuracy of an image transformation based on these reference points was evaluated. It was verified that an image rectification based on these points alone is reasonably accurate for flat areas, except for a small offset that can be easily fixed. However, in mountainous areas, such as the Vale do Sousa region, the incorporation of an elevation model is required for the establishment of an accurate geometric correction function—ortho-rectification. The ortho-rectification of satellite images is a similar process to that applied to aerial photographs to produce ortho-photographs. The OrthoEngine Satellite Edition software from PCI-Geomatics was used (PCI Geomatics 2000). Ground Control Points (GCPs) were identified on each original ASTER image using 1 : 25 000 raster maps (IGEOE 1998). A total of 11 GCPs were identified on the image ‘north’ and 10 GCPs on the image ‘south’. The mean residual errors for a 2nd order model were 0.54/0.53 (‘north’) and 0.37/0.29 (‘south’), in pixel units. A Digital Elevation Model (DEM) of the area with a 25-m grid was used. Each image was ortho-rectified into a 15-m pixel geo-reference image file. Vector data with the roads network was also available, derived from the 1 : 25 000 raster maps. Overlaying the road vector data tested the accuracy of the ortho-rectified images. The differences between the roads on the vector and image data were below or about the same size as the pixel.

The two ortho-rectified ASTER images were combined into a single multichannel image file of 2060 by 3340 15-m pixels (30.9 km by 50.1 km), covering the whole of the Vale do Sousa region. Data from the nine bands of the VNIR and SWIR instruments were used, stored in this file with a 15 m × 15 m pixel. The ‘mosaic’ ASTER image is totally cloud free, although there are some non-observed areas outside the Vale do Sousa region. About three quarters of the interest area was provided by the ‘north’ image. Figure 1 shows an RGB colour composite of the ‘mosaic’ image.

### 2.4 Image segmentation into objects

The segmentation of an image allows for groups of pixels to be considered as a single unit, or an object, within the image. The image segmentation was performed using eCognition software (Baatz *et al.* 2001). This multi-resolution segmentation is a bottom-up two region-merging technique starting with one-pixel objects. In numerous subsequent steps, smaller image objects are merged into bigger ones. Throughout this pairwise clustering process, the underlying optimization procedure minimizes the weighted heterogeneity  $nh$  of resulting image objects, where  $n$  is the size of a segment and  $h$  an arbitrary definition of heterogeneity. In each step, that pair of adjacent image objects is merged which stands for the smallest growth of the defined heterogeneity. If the smallest growth exceeds the threshold defined by the scale parameter, the process stops. Doing so, multi-resolution segmentation is a local optimization procedure.

There are three criteria for heterogeneity, and they can be applied in a mixed form. The first criterion has to do with spectral or colour heterogeneity, the second and the third ones have to do with spatial heterogeneity. It is useful to mix these criteria in order to avoid a fractal shaping of objects. The scale parameter is a measure of the maximum change in heterogeneity that may occur when merging two image objects. Internally, this value is squared and serves as the threshold that

terminates the segmentation algorithm. When the merger of a pair of image objects is examined, a fusion value between those objects is calculated and compared to the squared scale parameter. Throughout the segmentation procedure, the whole image is segmented and image objects are generated based upon several adjustable criteria of homogeneity or heterogeneity in colour and shape. Adjusting the so-called scale parameter indirectly influences the average object size: a larger value leads to bigger objects and vice versa. Additionally the influence of shape as well as the image's channels on the object's homogeneity can be adjusted. During the segmentation process all generated image objects are linked to each other automatically.

A set of segmentation parameters was established after visual inspection of a number of segmented images with different parameters, and prior knowledge of the study area. The ASTER bands 1, 2 and 3 were given a weighting factor of 1 and the remaining bands were given a weighting factor of 0.1. These values were chosen to account for the different spatial resolution of the VNIR and SWIR images. The eCognition software was used to produce segmented images with five different levels, hierarchically structured. Figure 2 shows a section of the segmented image at levels 1 (top), 3 (middle) and 5 (bottom) overlaid on a RGB colour composite of the ASTER VNIR bands (R:3/2, G:1/2, B:1; 3/2 means the average of bands 3 and 2). The whole image was segmented into a total of 51 186 objects in level 1, with an average size of 135 pixels, that is about 30 000 m<sup>2</sup> or 3 ha. The level 2 segmentation produced 14 133 objects, with an average size of 487 pixels (11 ha). The segmentation levels 3, 4 and 5 produced 4857, 2269 and 651 objects, with an average object size of 1417 pixels (32 ha), 3032 pixels (68 ha), 10 569 pixels (238 ha). Taking into account the size of the forest areas and small parcelling of land in this part of the country, segmentation levels 1 and 2 should be the most appropriate.

## 2.5 Land cover classes

The land cover classes were chosen according to the Portuguese national forest inventory (Tomé *et al.* 2002). A total of six main classes are used: AG, Agriculture; FL, Forest; IC, Uncultivated; IP, Non-productive; SC, Urban/Residential; HH, Water. The class Agriculture is divided into five sub-classes and the class Forest is divided into 13 sub-classes. For the purpose of the study and for the study area, a different set of land cover classes was thought to be more adequate. A new class was added: 'Burned areas' (FG). The class 'Forest' was divided into only four subclasses and the class IC used here also includes non-productive areas. A total of nine land cover classes were therefore considered. Figure 3 shows a sample photograph for each land cover class. A short description of these classes is presented in table 1. A void class was also used, to account for those pixels that were not observed in any of the two ASTER images (black areas in the RGB image of figure 1).

**2.5.1 Training areas.** Training areas were identified on the segmented image at level 1. In an initial stage, the existing land cover vector dataset, produced by air-photograph interpretation in 1995, was used. The polygons from this dataset that matched the segmented image polygons by at least 75% were selected as candidates for training. Ground surveys were then made to confirm the land cover of some of these objects. For practical reasons the training areas were chosen in two areas, one located in the southwest and the other on the northeast of the Vale do Sousa region. A total of 582 objects were identified for training. Table 1 shows the average size and number of objects used for training each class. Unfortunately only a small number



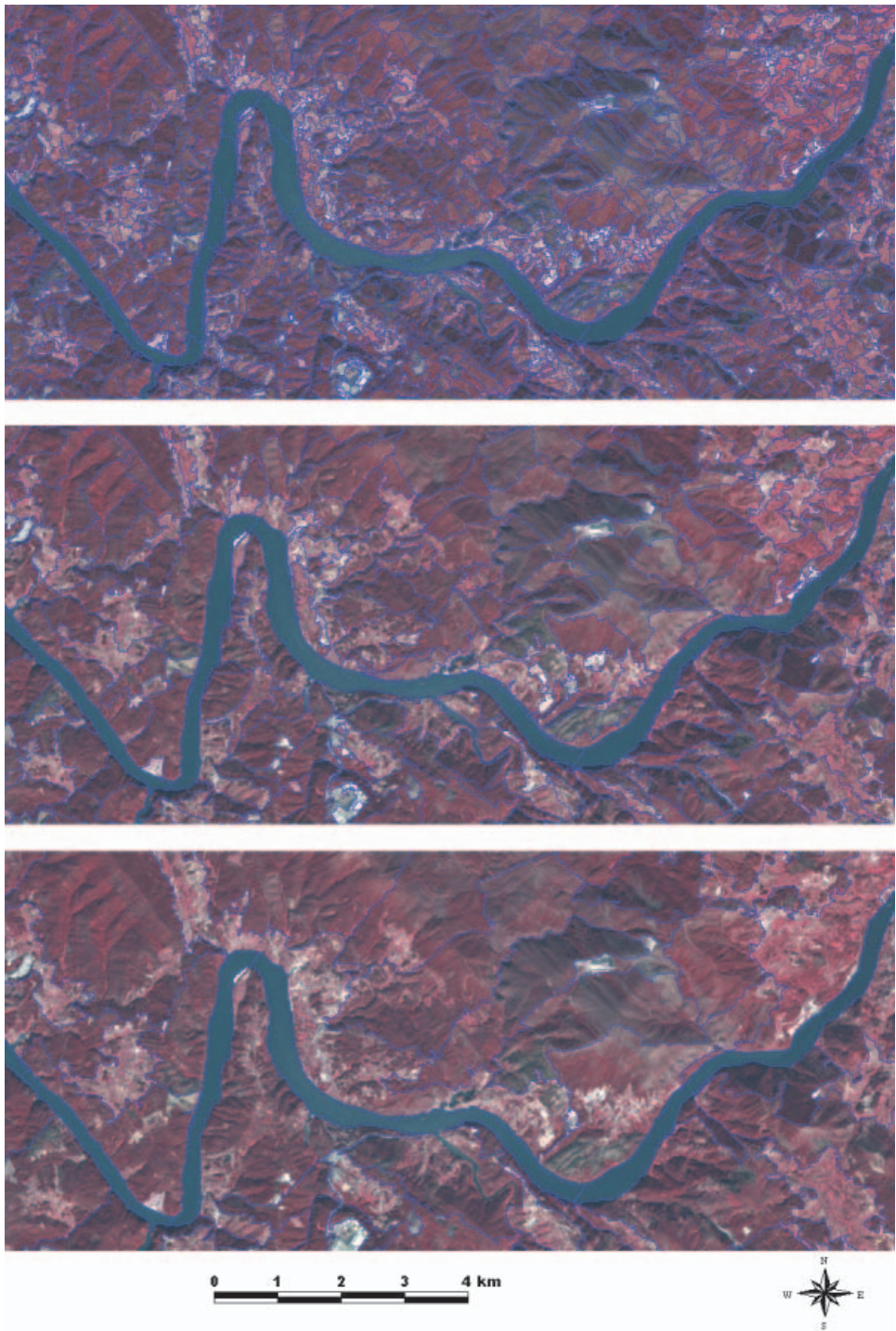


Figure 2. Section of the segmented image at levels 1 (a), 3 (b) and 5 (c), overlaid on a RGB colour composite of ASTER VNIR bands (R:3/2, G:1/2, B:1; 3/2 means the average of bands 3 and 2).



Figure 3. Photographs illustrating the nine land cover classes used: SC, Urban/residential; HH, Water; FG, Burned areas; FLEc, Forest/Eucalyptus; FLMix, Forest/mixed; FLFd, Forest/hardwoods; FLPb, Forest/Pinus; IC, Uncultivated/non-productive; AG, Agriculture.

of objects were available for some of the classes, because some types of land cover appear sparsely in the study area.

**2.5.2 Validation sites.** An inspection of the existing cartography showed that the class ‘Water’ identified on the classified image perfectly matches the existing rivers on the Vale do Sousa region. However, the existing cartography was inadequate for the evaluation of the classification results for all other land cover classes. A number of independent test sites (point locations) were selected on the image. These sites were chosen in two areas of the Vale do Sousa region by applying a rectangular grid with intervals of around 800 m. In a first stage 176 sites were selected, all assigned to one of the forest classes on the 1995 land cover survey. In a second stage 101 sites were randomly selected, following the relative distribution of each land cover class.

Table 1. Land cover classes, number of objects and average object size (ha) for the training areas of each class.

| Class identification | Class label | Class description           | No. of objects | Average size (ha) |
|----------------------|-------------|-----------------------------|----------------|-------------------|
| 1                    | SC          | Urban/residential areas     | 222            | 1.9               |
| 2                    | HH          | Water                       | 13             | 15.5              |
| 3                    | FG          | Burned areas                | 65             | 3.4               |
| 4                    | FIEc        | Forest—Eucalyptus           | 35             | 6.4               |
| 5                    | Flmix       | Forest—mixed                | 70             | 5.8               |
| 6                    | FIFd        | Forest—hardwoods            | 12             | 2.6               |
| 7                    | FIPb        | Forest—Pinus                | 16             | 4.3               |
| 8                    | IC          | Uncultivated/non-productive | 58             | 5.6               |
| 9                    | AG          | Agricultural areas          | 91             | 3.6               |

A total of 277 test sites were used for validation. This set of sites provides a uniform distribution for the land cover classes, except for the forest areas that were over sampled. The land cover of each of these sites was obtained by field surveys carried out in 2002. The time gap between the satellite image acquisition and the field surveys carried for validation were between a few months and one year. This time gap might be responsible for a few wrong site attributions, mainly due to forest fires occurring in 2002.

### 3. Methods

A total of four supervised classification methods were used: an algorithm based on a fuzzy classifier, Support Vector Machines (SVM), K Nearest Neighbours (K-NN) and a Logistic Discrimination classifier (LD).

#### 3.1 Fuzzy classifier

The eCognition software uses a supervised classification algorithm based on fuzzy logic (Baatz *et al.* 2001). A given object is assigned a probability value ( $p_i$ ) to belong to class  $i$ . The probability value is computed by membership functions, which can be defined manually or automatically on eCognition software (Baatz *et al.* 2001). The membership functions used were defined automatically, based exclusively on the training data. The object will have information about the three classes most likely to belong to ( $i^{1st}$ ,  $i^{2nd}$ ,  $i^{3rd}$ ) and the corresponding probability values ( $p^{1st}$ ,  $p^{2nd}$ ,  $p^{3rd}$ ). The class assigned will be that with highest probability. The value of  $p^{1st}$  is one indication of the reliability of the classification. Another and perhaps more consistent indicator is the stability measure,  $\Delta p = p^{1st} - p^{2nd}$ . In the best-case scenario we would have  $p^{1st} = 1$  and  $p^i = 0$  for all the other classes ( $i \neq i^{1st}$ ). That would result in a categorical choice of the object class and the value for  $\Delta p$  would be 1. On the opposite case we would have  $p^{1st} = p^{2nd}$  and that would result in  $\Delta p = 0$ . Although this is a good concept, the end user is usually interested in a single class assignment for each object, and often only the highest scoring class is in fact used.

#### 3.2 Support Vector Machines

The Support Vector Machine (SVM) classifier is an extension to the optimum margin algorithm. This extension is made by mapping the input space,  $\mathbf{X}$ , into



another one of higher dimension (eventually infinite): the feature space  $\mathbf{H}$  (Vapnik 1998).

Consider the training set as  $\{(x_i, y_i); x_i \in \mathbb{R}^P; y_i \in \{-1, 1\}, i=1, \dots, l\}$ . The optimum margin algorithm searches for a linear boundary given by a hyper plane  $\mathbf{w} \cdot \mathbf{x} + b = 0$  that separates the two classes. In this way, the decision function is given by  $f(\mathbf{x}) = \mathbf{w} \cdot \mathbf{x} + b$ . The separating hyperplane is the one that maximizes the margin between the two classes. The solution can be achieved by solving the convex optimization problem:

$$\begin{aligned} \min \|\mathbf{w}\|^2 \\ \text{subject to } \mathbf{y}_i(\mathbf{w} \cdot x_i + b) \geq 1, i = 1, \dots, l \end{aligned} \tag{1}$$

Using Lagrange Multipliers  $\alpha_i$  to solve this problem, the decision function is given by

$$f(\mathbf{x}) = \text{sign} \left( \sum_{i=1, \dots, l} \alpha_i y_i (x_i \cdot \mathbf{x}) + b \right) \tag{2}$$

To build a SVM, we need to define a feature space  $\mathbf{H}$ , with higher dimension than  $\mathbf{X}$ . To do so, we need a function  $\phi: \mathbf{X} \rightarrow \mathbf{H}$  that maps the initial space  $\mathbf{X}$  into  $\mathbf{H}$ . This mapping will lead us to a higher space  $\mathbf{H}$  where it will be possible to apply the optimum margin algorithm to get a linear classifier. The separating hyperplane in  $\mathbf{H}$  will thus be given by:

$$\sum_{i=1, \dots, l} \alpha_i y_i (\phi(x_i) \cdot \phi(x)) + b = 0 \tag{3}$$

By using a kernel function  $K(\mathbf{x}, \mathbf{y})$ , such that  $K(\mathbf{x}, \mathbf{y}) = \phi(\mathbf{x}) \cdot \phi(\mathbf{y})$ , the linear boundary on  $\mathbf{H}$  will correspond to a nonlinear one in input space  $\mathbf{X}$ . Doing this, we do not need to specify the space  $\mathbf{H}$  nor the mapping  $\phi(\mathbf{x})$ . The decision function in the input space is given by:

$$f(\mathbf{x}) = \text{sign} \left( \sum_{i=1, \dots, l} \alpha_i y_i K(x_i, \mathbf{x}) + b \right) \tag{4}$$

The most commonly used kernels to build SVM for classification are the Gaussian,  $K(\mathbf{x}, \mathbf{y}) = e^{-\|\mathbf{x}-\mathbf{y}\|^2/2\sigma^2}$  and polynomial  $K(\mathbf{x}, \mathbf{y}) = (\gamma \mathbf{x} \cdot \mathbf{y} + \beta)^p$  kernels.

For each classification problem, the user must choose a kernel function and the associated parameters. This choice is crucial for the accuracy of the final classification. In principal, SVMs can only solve binary classification problems. One commonly used technique that allows for multi-class classification is the one-against-one method. This method fits a total of  $k(k - 1)/2$  binary sub-classifiers and finds the correct class by a voting mechanism. The one-against-all method is an alternative approach, in which  $k$  SVM models are constructed. The  $i$ th SVM is trained with all members of  $i$  having positive labels, and all remaining members having negative labels. The one-against-one method was used here, as it has been shown that it performs better than the one-against-all method (Platt *et al.* 2000).

### 3.3 K Nearest Neighbours classifier

The K Nearest Neighbours (K-NN) classifier is a simple, well known method frequently used in various types of classification problems. Nonparametric K-NN

algorithms have been used to produce estimates of various forest parameters over large areas in Sweden, using optical satellite data (Reese *et al.* 2002, Tomppo *et al.* 2002). The K-NN classifier does not require the fitting of a model, it is a memory-based method (Fukunaga 1990). Given a new observation to be classified,  $x_0$ , this algorithm finds the K closest observations in distance to  $x_0$ . The set of K closest observations in distance are the K nearest neighbours. The class most frequent in the set of the K nearest neighbours is the one to be assigned to the observation  $x_0$ . The Euclidean distance is used to measure the proximity of the observations. Each time a new object is classified, the K-NN classifier needs to compute the distance of that object to all other objects in the training set. This requires that all objects are stored in memory, while other methods only need to have the training set available in memory to build the classifier. The training objects are usually no longer required once the classifier is established. This is therefore a disadvantage of the K-NN classifier.

### 3.4 Logistic discrimination classifier

The classification performed by logistic discrimination methodology is based on Bayes rule, this means that posterior probabilities,  $P(\text{Class}=k|X=x)$ , are needed (Hastie *et al.* 2001). The model estimates posterior probabilities of the  $C$  classes via linear functions in  $x$  (vector in the feature space), ensuring that they sum up to 1, and are kept within the range 0 to 1. The model is then specified in terms of  $K-1$  logit transformations. This leads to the following estimators:

$$P(\text{Class}=k|X=x) = \frac{\exp(\beta_0 + \beta_k \cdot x)}{1 + \sum_{i=1, \dots, C-1} \exp(\beta_{i0} + \beta_{ik} \cdot x)}, \quad k=1, \dots, C-1 \quad (5)$$

$$P(\text{Class}=C|X=x) = \frac{1}{1 + \sum_{i=1, \dots, C-1} \exp(\beta_{i0} + \beta_{ik} \cdot x)} \quad (6)$$

The rule is assign to  $x_0$  (new observation) the class  $k$  that maximizes the posterior probability  $P(\text{Class}=k|X=x_0)$ . To achieve the maximizing probability, the coefficients  $\beta_k$  are determined by maximum likelihood method. As a consequence, the user does not need to define any values for the parameters.

### 3.5 Jeffries-Matusita distance

The Jeffries-Matusita (J-M) distance is a good indicator of the separability between classes often used. The J-M distance between a pair of probability distributions (spectral classes) converges asymptotically to 2.0 as a function of the distance between class means (Richards and Jia 1999). In general, the larger the separability values between classes, the better the classification results should be. As a commonly used rule, the separability between two classes is considered good when the J-M distance is above 1.90. Two classes are considered to be very poorly separated when the J-M distance is below 1.0 (Richards and Jia 1999).

## 4. Results and discussion

The segmented ASTER image was classified using the 582 objects as training for the nine land cover classes (table 1). A total of 19 features were used—the mean and

standard deviation values from ASTER bands 1–9, and an additional feature with the object density (related with its shape). These features were exported from the eCognition software for all 51 186 objects. The K-NN, LD and SVM classifiers were implemented using R project software (<http://www.r-project.org/>). The K value was set to 4 for the K-NN classifier, as this was the value in the range 2 to 200 that minimized the error estimate. The fuzzy classifier was implemented on eCognition. A section of the classified image by the fuzzy classifier is presented in figure 4, together with a number of validation sites (numbered black dots).

#### 4.1 Image classification results

An evaluation of the classification methods was performed for the SVM, K-NN and LD methods using a cross validation technique. These classifiers were applied using 90% of the training objects, and the remaining 10% were used for validation purposes. This process was repeated 10 times, changing the subset of the training areas selected for validation. The overall accuracy of the cross validation was 81.8% for SVM and 79.7% for K-NN and 82.8% for LD.

An efficient projection method to reduce the dimensionality of the data is the canonical variates transformation. The canonical variates are the variables resulting from the linear discriminant analysis developed by Fisher (Hastie *et al.* 2001). A canonical variates transformation was applied to the original dataset in order to improve the classification results achieved by the SVM, K-NN and LD methods.

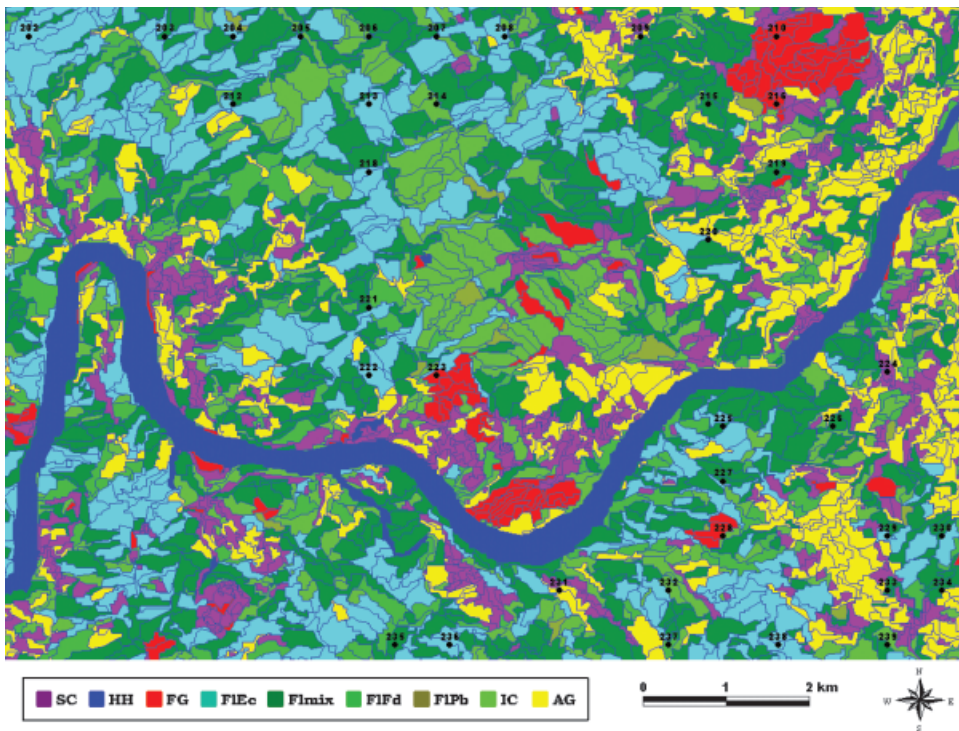


Figure 4. Section of the classified image (fuzzy method) with several validation sites (black dots with a numeric identifier).

After the canonical variates transformation, the cross validation results were improved to 85.4% for SVM, 84.0% for K-NN and 84.0% for LD.

The classification results were also evaluated using 277 validation sites, obtained independently of the training areas. The percentage of correctly classified sites for each land cover class is presented in table 2 for each of the four methods used. The last line on this table is the average accuracy for all the validation sites. A perhaps more statistically meaningful average accuracy could be obtained using the relative incidence of each land cover class used. However, as the main concern of this work was to classify different types of forest, a decision was made to over sample these classes in the validation process. The average accuracy would nevertheless be improved if this over sample was not used, as the classification results are generally worst for all of the forest classes. Overall, the fuzzy classifier is the worst performer with only 44.6% of the sites correctly classified. The K-NN method classified 60.9% of the sites accurately. The best methods were the SVM (70.9% sites correct) and LD classifiers (72.2% sites correct). The results obtained were somehow disappointing. In a related work, Zhu and Blumberg (2002) reported a classification accuracy of nearly 90% using a SVM classifier on an ASTER image. However, only six classes were considered for the urban area of that study and the classes were likely to be spectrally more diverse than those considered here.

#### 4.2 Evaluation of the spectral signature of land cover classes

The poor results from the classification methods are probably linked with a difficulty in separating the spectral signal of the various classes. The J-M distance was computed for the nine classes using the training areas, with the results presented in table 3. As expected, the separability between 'Water' (HH) and all the other classes is excellent. The class 'Burned' (FG) is also well separated from most classes. The classes IC and SC are reasonably separated from each other, as is SC and the four forest classes. The main problems occur between the various forest classes and between these and IC.

A less common investigation on the spectral similarities between classes was carried out. An unsupervised classification algorithm was applied to the ASTER image on a pixel-by-pixel basis. The algorithm used is based, with minor modifications, on the ISODATA method (Tou and Gonzalez 1974). The ISODATA method is similar in principle to the K-means procedure in the sense

Table 2. Summary of the evaluation of the classification results. Percentage of correctly classified validation sites for each method—Fuzzy classifier, Support Vector Machine (SVM), K Nearest Neighbours classifier (K-NN) and Logistic Discrimination classifier (LD).

| Class    | No. of sites | Fuzzy | SVM   | K-NN  | LD    |
|----------|--------------|-------|-------|-------|-------|
| 1. SC    | 36           | 56%   | 81%   | 75%   | 75%   |
| 2. HH    | 1            | 100%  | 100%  | 100%  | 100%  |
| 3. FG    | 3            | 100%  | 66%   | 66%   | 66%   |
| 4. FlEc  | 79           | 34%   | 77%   | 44%   | 68%   |
| 5. Flmix | 88           | 48%   | 55%   | 71%   | 74%   |
| 6. FlFd  | 4            | 50%   | 50%   | 0%    | 25%   |
| 7. FlPb  | 15           | 7%    | 62%   | 18%   | 68%   |
| 8. IC    | 20           | 55%   | 88%   | 73%   | 73%   |
| 9. AG    | 31           | 52%   | 81%   | 74%   | 81%   |
| Total    | 277          | 44.6% | 70.5% | 60.9% | 72.2% |



Table 3. Jeffries-Matusita distance ( $d_{J-M}$ ) between the land cover classes.

| $d_{J-M}$ | SC   | HH   | FG   | FIEc | Flmix | FIFd | FIPb | IC   |
|-----------|------|------|------|------|-------|------|------|------|
| HH        | 2.00 |      |      |      |       |      |      |      |
| FG        | 1.68 | 2.00 |      |      |       |      |      |      |
| FIEc      | 1.95 | 2.00 | 1.95 |      |       |      |      |      |
| Flmix     | 1.89 | 2.00 | 1.88 | 0.42 |       |      |      |      |
| FIFd      | 1.85 | 2.00 | 1.91 | 1.25 | 0.84  |      |      |      |
| FIPb      | 1.90 | 2.00 | 1.82 | 1.49 | 1.02  | 1.63 |      |      |
| IC        | 1.82 | 2.00 | 1.75 | 1.40 | 0.96  | 1.06 | 1.36 |      |
| AG        | 1.57 | 2.00 | 1.91 | 1.92 | 1.74  | 1.58 | 1.88 | 1.71 |

that cluster centres are iteratively determined sampled means. Unlike the latter algorithm, however, ISODATA represents a fairly comprehensive set of additional heuristic procedures that have been incorporated into an interactive scheme (PCI Geomatics 2001), allowing, for example, for the number of clusters to vary between a predefined range. The ISODATA classifier was run using PCI software (PCI Geomatics 2001). The process converged after 147 iterations, resulting in 33 classes. These classes can be thought of as probability density clusters in the 9-dimensional multi-spectral space. Ideally each one of the classes characterized in the training areas would correspond to one of these clusters, or perhaps more than one cluster close together in the multi-spectral space.

A comparison between the nine land cover classes and the clusters from the unsupervised classification was carried out to evaluate the spectral similarity between classes. The 33 clusters were structured hierarchically according to their relative distance in the multi-spectral space (Duda *et al.* 2001). Figure 5 shows the hierarchical tree for 25 clusters, labelled A to Y. The cluster X is an aggregation of four original clusters, all related to non-observed areas of the image. The reason that non-observed pixels were assigned four different clusters has to do with slightly different coverage by the VNIR and SWIR instruments. Another seven clusters

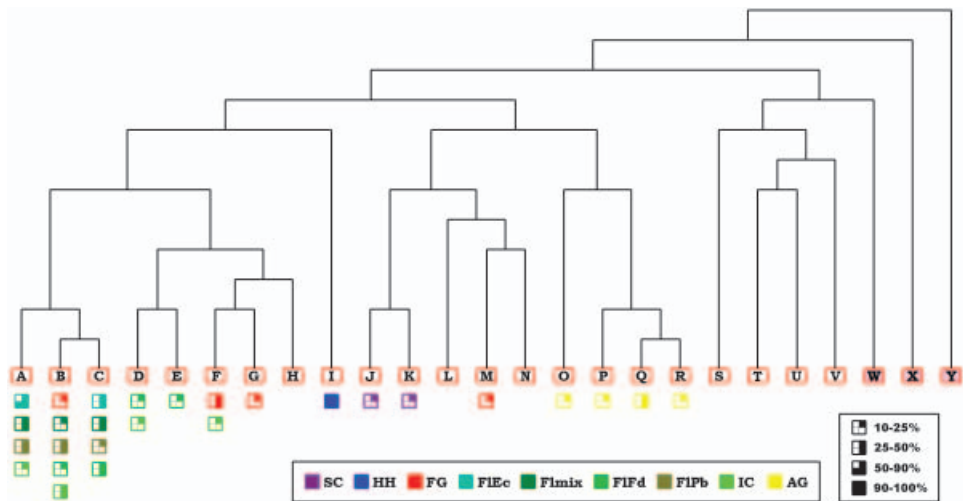


Figure 5. Hierarchical structure of clusters identified by unsupervised classification and correspondence between these clusters and the land cover classes searched for, according to the training areas.

from the original 33 were aggregated into clusters W and Y, as the number of pixels from the training areas in these clusters was negligible. Figure 5 also shows the correspondence between the land cover classes, based on the training areas, and the clusters identified by the unsupervised classification. For example, 8790 of the 8936 pixels (98.4%) identified as 'Water' in training belong to cluster I. This is represented in figure 5 in blue with a solid square (90–100%). The class 'Water' is very well defined in the multi-spectral space, not only because it is mostly associated with a single cluster, but also because this cluster only merges with others at a high level in the hierarchy tree. Another example of a class spectrally well defined is 'Agricultural' (AG). Most of the pixels identified in this class in training (74.4%) belong not to one but to four different clusters (O, P, Q and R). However, these clusters merge together before they mix with other classes, as can be easily seen in the hierarchical tree.

An example of a class poorly characterized spectrally is 'Burned areas' (FG). Over 91% of the pixels from the training of this class belong to four clusters (B, F, G and M), but, unlike 'Agricultural', these clusters are far apart in the hierarchical tree. In order to have them all merged together an aggregation of 18 clusters would have to be done. Furthermore, cluster B is shared with several other classes.

As shown in figure 5, the forest classes are reasonably well defined in the multi-spectral space, but the problem with these classes is that they share areas (and clusters) in the multi-spectral space. This analysis shows that it is almost impossible to properly discriminate objects in the various forest classes. This is consistent both with the ground validation results and the analysis of the J-M distance between classes.

## 5. Conclusions

The wider availability of high quality multi-spectral satellite images, such as ASTER, should allow for more frequent updates of land cover maps, at reduced costs. The accuracy of these maps is nevertheless dependent on ground truth data, which is a time consuming task. One of the problems in importing satellite derived land cover maps into a GIS is the fixed pixel resolution of the images. The segmentation of an image into a set of objects should provide a more realistic result, from the end user perspective, than treating pixels as individual observations independently from their neighbourhood. Furthermore, some of the most sophisticated classifiers (such as SVM) are very intensive computationally and are not applicable on a pixel-by-pixel basis on very large multi-spectral images. The data reduction achieved by the initial image segmentation is therefore a crucial stage in order for those classifiers to be used.

A practical application of object based supervised classification for land cover update was undertaken on an ASTER satellite image. The land cover classes searched for and the characteristics of the Vale do Sousa region were a major test for the classification methods. The classification accuracy was checked by cross validation analysis and by a total of 277 validation sites obtained from independent field surveys.

The Support Vector Machines (SVM) and Logistic Discrimination (LD) classifiers provided the best results, both with over 70% of the validation sites accurately classified. SVMs are a very powerful technique for classification. The main problem with this method is the choice of the right parameters. We carried out extensive experimentation in order to estimate the parameters for the SVM,

which resulted in a low classification error. Logistic Discrimination (LD) is a linear method of classification, which makes weaker assumptions than Linear Discriminant Analysis. This means that for a large class of problems, this method provides either the best classifier or close to best. This is probably the case for our recognition problem, as the results obtained with LD were good. The K Nearest Neighbours (K-NN) only classified 60.9% of sites accurately and the fuzzy classifiers performed even worse (44.6%). The results would have been slightly better if the relative incidence of each land cover class in the region was used for the selection of validation sites.

The land cover classes with different forest types were found to be difficult to discriminate between. The main reason for this was the spectral similarity between these classes, according to the analysis performed on the training data. The unsupervised classification with hierarchical clustering showed that most classes were composed of several clusters in the multi-spectral space, and that some of the classes shared the same clusters. This analysis proved quite effective in identifying the possible limitations in the land cover classification with this specific dataset.

### Acknowledgments

The authors wish to thank the ASTER Earth Remote Sensing Data Analysis Center (ERSDAC, Tokyo, Japan) for providing the satellite data through an Announcement of Research Opportunity (ARO-070) and Centro de Investigação em Ciências Geo-Espaciais (Porto, Portugal) for financial support through the POCTI FEDER program.

### References

- BAATZ, M., BENZ, U., DEGHANI, S., HEYENEN, M., HÖLTJE, A., HOFMANN, P., LINGENFELDER, I., MIMLER, M., SOHLBACH, M., WEBER, M. and WILLHAUCK, G., 2001, eCognition Object Oriented Image Analysis, User Guide, Definiens Imaging, München, Germany.
- CNIG, 1990, Carta de Ocupação do Solo (COS' 90). Centro Nacional de Informação Geográfica, Lisboa, Portugal.
- COLBY, J.D. and KEATING, P.L., 1998, Land cover classification using Landsat TM imagery in the Tropical Highlands: the influence of anisotropic reflectance. *International Journal of Remote Sensing*, **19**, pp. 1479–1500.
- DUDA, R.O., HART, P.E. and STORK, D.G., 2001, *Pattern Classification*, 2nd edn (New York: John Wiley and Sons).
- FUKUNAGA, K., 1990, *Introduction to Statistical Pattern Recognition*, 2nd edn (New York: Academic Press).
- GAO, J. and SKILLCORN, D., 1998, Capability of SPOT XS data in producing detailed land cover maps at the urban-rural periphery. *International Journal of Remote Sensing*, **19**, pp. 2877–2891.
- GENELETTI, D. and GORTE, B.G.H., 2003, A method for object-oriented land cover classification combining Landsat TM data and aerial photographs. *International Journal of Remote Sensing*, **24**, pp. 1273–1286.
- HASTIE, T., TIBSHIRANI, R. and FRIEDMAN, J., 2001, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (Berlin: Springer Verlag).
- IGEOE, 1998, Série Cartográfica M888 1/25000 de Portugal Continental. Instituto Geográfico do Exército, Lisboa, Portugal.
- LABA, M., SMITH, S.D. and DEGLORIA, S.D., 1997, Landsat-based land cover mapping in the Lower Yuna River watershed in the Dominican Republic. *International Journal of Remote Sensing*, **18**, pp. 3011–3025.

- PCI GEOMATICS, 2000, Ortho Engine Reference Manual, Version 7.0, PCI Geomatics, Ontario, Canada.
- PCI GEOMATICS, 2001, X-Pace Reference Manual, Version 8.2, PCI Geomatics, Ontario, Canada.
- PLATT, J.C., CRISTIANINI, N. and SHAW-TAYLOR, J., 2000, Large Margin DAGs for multiclass classification. *Advances in Neural Network Information Processing Systems*, **12**, pp. 547–553.
- R PROJECT, The R Project for Statistical Computing, <http://www.r-project.org/>
- REESE, H., NILSSON, M., SANDSTRÖM, P. and OLSSON, H., 2002, Applications using estimates of forest parameters derived from satellite and forest inventory data. *Computers and Electronics in Agriculture*, **37**, pp. 37–55.
- RICHARDS, J.A. and JIA, X., 1999, *Remote Sensing Digital Image Analysis—An Introduction*, 3rd edn (Berlin: Springer Verlag).
- TOMÉ, M., FIDALGO, B. and GASPAR, J., 2002, Normas de fotointerpretação para as áreas de demonstração da Associação Florestal do Vale do Sousa. Relatórios Técnicos do GIMREF, No.1/2002, Universidade Técnica de Lisboa, Instituto Superior de Agronomia, Lisboa, Portugal.
- TOMPO, E., NILSSON, M., ROSENGREN, M., AALTO, P. and KENNEDY, P., 2002, Simultaneous use of Landsat-TM and IRS-1C WiFS data in estimating large area tree stem volume and aboveground biomass. *Remote Sensing of Environment*, **82**, pp. 156–171.
- TOU, J.T. and GONZALEZ, R.C., 1974, *Pattern Recognition Principles* (Reading, MA: Addison-Wesley).
- VAPNIK, V., 1998, *Statistical Learning Theory* (New York: John Wiley & Sons).
- YAMAGUCHI, Y., KAHLE, A.B., TSU, H., KAWAKAMI, T. and PNIEL, M., 1998, Overview of Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER). *IEEE Transactions on Geoscience and Remote Sensing*, **36**, pp. 1062–1071.
- ZHU, G. and BLUMBERG, D.G., 2002, Classification using ASTER data and SVM algorithms; the case study of Beer Sheva, Israel. *Remote Sensing of Environment*, **80**, pp. 233–240.