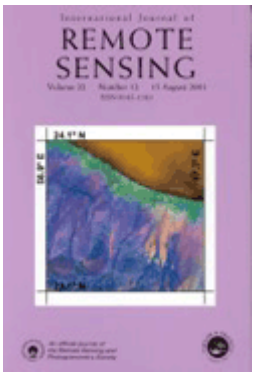

A method to incorporate uncertainty in the classification of remote sensing images

Journal:	*International Journal of Remote Sensing*
Manuscript ID:	TRES-SIP-2008-0049.R3
Manuscript Type:	Special Issue Paper
Date Submitted by the Author:	
Complete List of Authors:	Gonçalves, Luisa M S; Polytechnic Institute of Leiria, Civil Engineering Department Fonte, Cidália C; University of Coimbra, Department of Mathematics Júlio, Eduardo N B S; University of Coimbra, ISISE, Civil Engineering Department Caetano, Mário; Portuguese Geographic Institute (IGP), Remote Sensing Unit (RSU)
Keywords:	FUZZY CLASSIFICATION, IKONOS, LANDSCAPE
Keywords (user defined):	Hybrid classification, Uncertainty

Catchline (head of first page only) *International Journal of Remote Sensing*, Vol. X, No. X, Month 2008, xxx–xxx

Running heads (verso) *Gonçalves, Fonte, Júlio and Caetano*
(recto) *A method to incorporate uncertainty in the classification*

Article Type (Research Paper)

A method to incorporate uncertainty in the classification of remote sensing images

LUISA M S GONÇALVES*†, CIDÁLIA C FONTE§, EDUARDO N B S JÚLIO&, MARIO CAETANO‡

†Polytechnic Institute of Leiria, School of Technology and Management, Portugal
† Institute for Systems and Computers Engineering at Coimbra, Portugal
§Institute for Systems and Computers Engineering at Coimbra, Portugal
§Department of Mathematics, University of Coimbra, Portugal
&ISISE, Civil Engineering Department, University of Coimbra, Portugal
‡Portuguese Geographic Institute (IGP), Remote Sensing Unit (RSU), Lisboa, Portugal
‡ CEGI, Instituto Superior de Estatística e Gestão de Informação, ISEGI, Universidade Nova de Lisboa, 1070-312 Lisboa, Portugal

*†Email: luisa@estg.ipleiria.pt

Abstract

The aim of this paper is to investigate if the uncertainty associated with the classification of surface elements into the classification of landscape units increases the results accuracy. To this end, a hybrid classification method is developed, incorporating uncertainty information in the classification of very high spatial resolution multispectral satellite images, to obtain a map of landscape units. The developed classification methodology includes the following steps: 1) a pixel-based hard classification with a probabilistic Bayesian classifier; 2) computation of the posterior probabilities and quantification of the classification uncertainty using an uncertainty measure; 3) image segmentation; and 4) object classification based on decision rules. The classification of the resulting objects into landscape units is performed considering a set of decision rules that incorporate the pixel based classification uncertainty. The proposed methodology was tested on the classification of an IKONOS satellite image. The accuracy of the classification was computed using an error matrix. The comparison between the results obtained with the proposed approach and those obtained without considering the classification uncertainty revealed a 12% increase in the overall accuracy. This shows that the information about uncertainty can be valuable when making decisions and can actually increase the accuracy of the classification results.

AMS Subject Classification:

1. Introduction

Very High Spatial Resolution (VHSR) images opened some new challenges to the remote sensing community and led to the development of new studies to improve the identification of land cover features. These images allow the identification of smaller objects and landscape units, and therefore finer maps can be produced. Although, despite their great potential, they also present drawbacks and limitations, such as the increase of the spectral variability and the amount of shadows, as well as the enormous amount of data (Blaschke *et al.* 2004, Goetz *et al*. 2003). In addition, since these images have lower spectral resolution than the images obtained with sensors with smaller spatial resolution, such as the Landsat-TM, some limitations occur, for example, in the characterization of forest cover, increasing the difficulty in separating different forest species (Goetz *et al*. 2003).

Several studies show that traditional pixel-based classification methods are not suitable to identify many types of land cover classes in VHSR images (e.g., Scheiewe *et al*. 2001, Carleer and Wolff 2004). The complex relationships between pixels and objects, as well as the increased number of Landscape Units (LU) that are mosaics of single entities or spatial arrangements of land cover classes, like Agro-forest areas, impose new demands and the need to develop new methods that incorporate shape and context, which are some of the main clues used by a human interpreter (e.g., Wang *et al.* 2004, Plantier and Caetano 2007), as well as a closer integration of remote sensing and Geographic Information Systems (GIS) (Donay *et al*. 2001, Blaschke *et al.* 2004). Object-based classification methods and hybrid methods composed of pixel-based and object-based classifications have been proposed to incorporate such spatial information into the classification procedure (e.g., Wang *et al*. 2004, Carleer and Wolff 2004, Gonçalves and Caetano 2004; Guerrero *et al*. 2006; Plantier and Caetano 2007). However, as highlighted by Guerrero *et al*. (2006), the difficulty in discriminating forest species and Mediterranean landscapes with an object-based classification using only the multispectral bands of VHSR images still remains.

During the last decade considerable research has also been done in the development of soft classifiers to extract information from remote sensing images (e.g., Maselli *et al*. 1995, Foody 2000, Brown *et al*. 2000, Zhang and Foody 2001, Ibrahim *et al.* 2005, Doan and Foody 2007). These methods assign, to each pixel or object, different degrees of probability, possibility or membership associated with the several classes. These extra data provide additional land cover information at the pixel or object level and allow the assessment of the classification uncertainty. Within the context of remote sensing, several semantics may be associated with the degrees of probability, possibility or membership obtained with these classifiers. These may represent partial membership of the classes to the spatial units, as in the case of mixed pixels or objects; a degree of similarity between what exists in the ground and the pure classes; or the uncertainty associated with the correct allocation of a class to a pixel. However, in practice, it is not generally possible to know which interpretation is the correct one, since the spectral responses are used to determine what exists in reality and therefore the real conditions are not known in advance. For this reason, in most cases, this additional information actually reveals the degree of uncertainty associated with the correct allocation of a class to a pixel, even though this information is frequently used to estimate the mixture between classes (e.g., Zhang and Foody 2001, Maselli *et al.* 1996, Bastin 1997).

The classification of VHSR images with a hybrid classification approach that combines pixels and objects has shown to be suitable for the identification of LU classes that contain a variety of land cover objects (e.g., Wang *et al.* 2004, Plantier and Caetano 2007), but the integration of the classification uncertainty in this process has never been investigated. Since it is difficult to discriminate between different species of forest

with VHSR images (e.g., Goetz *et al*. 2003, Carleer and Wolff 2004), this paper analyses whether the introduction of uncertainty information in the hybrid classification process increases the accuracy of LU classification and enables the identification of the main forest species in Portugal.

The new hybrid classification method presented in this study involves two steps: (1) a pixel based classification to obtain a Surface Elements Map (SEM), containing the elementary entities, like crown trees and parts of buildings, that are the basic units of landscape, called Surface Elements (SE); (2) an object-based classification to obtain a Landscape Units Map (LUM), on a 1:10 000 scale, presenting the spatial patterns of the LU classes. The pixel based classification was performed with a probabilistic Bayesian classifier similar to the maximum likelihood classifier. The posterior probabilities associated to the classes were then computed for all pixels, and were used to evaluate the classification uncertainty, for each pixel, with an uncertainty measure. This uncertainty information was used in the creation of a LUM from the SEM, considering a set of decision rules that incorporate the arrangement of the SE classification within each object and the degree of uncertainty given by the uncertainty measure. The main objective of integrating uncertainty in the classification process is to avoid the use of misclassified SE in the classification of the LU classes. The results are then compared with the ones obtained with a similar method where uncertainty is not considered. With this approach the final LUM is in the vector format, which is well suited to GIS users.

2. Study area and data

The study was conducted in a rural area with a smooth topographic relief, located in a transition zone between the centre and south of Portugal, which includes diverse landscapes representing Mediterranean environments. The area is mainly occupied by agriculture, pastures, forest and agro-forestry areas, where the dominant forest species are eucalyptus, coniferous and cork trees. An image obtained by the IKONOS sensor was used, with a spatial resolution of respectively 1 m in the panchromatic mode and 4 m in the multi-spectral mode (XS). The product acquired was the Geo Ortho Kit and the study was performed using the four multi-spectral bands. The image acquisition details are presented in table 1. The geometric correction of the multi-spectral image consisted of its orthorectification. The average quadratic error obtained for the geometric correction was 1.39 m, inferior to half the pixel size, which guarantees an accurate geo-referencing. Pixels in the image are recorded in 16 bits to keep the 11 bits original image information.

Since in unitemporal studies carried out in regions with no significant topographic relief, and presenting uniform atmospheric conditions in the image data, the radiometric corrections do not improve the results (Caetano 1995), no radiometric corrections were applied to the image.

3. Methodology

The main goal of the classification is to obtain a LUM on a 1:10 000 scale that includes the main forest species of the Portuguese mainland, using a new hybrid classification approach that integrates uncertainty information. To determine if the inclusion of information about the SE classification uncertainty increases the accuracy of the LU classification, a similar method, where the classification uncertainty is not considered, was performed. For this reason, two hybrid classification methods are presented in this study. The first method introduces the uncertainty in the classification process and includes the following steps: 1) pixel-based classification; 2) evaluation of the previous classification uncertainty; 3) image segmentation and 4) object classification based on decision rules. The second classification method does not take uncertainty into consideration and includes three steps: 1) pixel-based classification of the image; 2) image segmentation and 3) object classification based on decision rules. This pixel/object combined approach was initially presented in Plantier and Caetano (2007). Since the goal is to evaluate if the introduction of the uncertainty information in

the classification of LU can increase the results accuracy, the paper focuses mainly on the first methodology (see figure 1).

3.1 Classification with uncertainty

3.1.1 Surface elements map

To identify and map the SE a pixel-based classification of the image was performed using a Bayesian classifier similar to the maximum likelihood classifier. The traditional use of this classification method assigns each pixel to the class corresponding to the highest probability density function value. However, posterior probabilities can be computed with the probability density functions, which may be interpreted as representing the proportional cover of the classes in each pixel or indicators of the uncertainty associated with the pixel allocation to the classes (e.g., Ibrahim *et al.* 2005, Shi *et al.* 1999). The second interpretation is used in this paper, where the posterior probabilities are used to compute classification uncertainty measures.

Before the classification itself, several preliminary processing steps were carried out. First, an analysis of the image by a human interpreter was made, to define the most representative classes and their SE. Since the main goal was to discriminate the dominant forest species of the region, the nomenclature used contains mainly classes corresponding to the dominant forest of the Portuguese mainland. The SE classes used in this study are: Eucalyptus Trees (ET); Cork Trees (CKT), Coniferous Trees (CFT); Shadows (S); Shallow Water (SW), Deep Water (DW), Herbaceous Vegetation (HV), Sparse Herbaceous Vegetation (SHV) and Non-Vegetated Area (NVA).

The second step was the establishment of the protocol to select the training and testing sample elements. The training and testing dataset consisted in a semi-random selection of sites. A human interpreter delimited twenty five polygons for each class and a stratified random selection of 300 samples per class was performed: one half was used to train the classifier and the other half to test it. The sample unit was the pixel. The total sample size included 2700 pixels, corresponding to 5% of the pixels inside the chosen polygons. Only pixels representative of pure SE were considered (Plantier and Caetano 2007). This testing set presented a spectral variability in class response very similar to the training set and was used to evaluate the classifier's behaviour.

To evaluate the classification accuracy a second testing set was used. A stratified random sampling of 100 pixels per class was selected considering the entire image scene, which also included mixed pixels. The number of pixels was chosen to obtain a standard error of 0.05 for the estimation of the accuracy indexes of each class (Stehman 2001). Each land cover class was sampled independently. Hereafter, the first testing set, which was used to assess the classifier performance, is referred to as testing set 1, while the one used to evaluate the map accuracy is referred to as testing set 2.

The accuracy assessment was made with an error matrix, where the p_{ij} entry in the matrix is the estimation, obtained from the sample data, of the proportion of pixels that are class i in the map and class j in the reference. For a stratified random sample considering the mapped land-cover classes as strata, $p_{ij} = (n_{ij} / n_{i+})(N_{i+} / N)$, where n_{ij} is the number of sample pixels classified as map class i and reference class j, n_{i+} is the sample size in class i, N_{i+} is the population size in class i and N is the total number of pixels in the map (Stehman and Czaplewski 1998). Accuracy parameters can be estimated using the p_{ij} entries in the error matrix. When considering proportions, the "user's accuracy" corresponds to the conditional probability of correctly classifying a location given that it has been mapped as class i, referred here as the Conditional Probability of the Map (CPM) and the "producer's accuracy ",which is the conditional probability of having correctly mapped a location given that it is truly class j, is the Conditional Probability of the Reference (CPR) (Stehman and Czaplewski 2003).

3.1.2 Uncertainty

Even though the posterior probabilities already provide some information about the pixel-based classification uncertainty, another indicator of the uncertainty was used, which can be applied to possibility distributions or probabilities distributions. This uncertainty measure is available in the commercial software IDRISI and is given by

$$U = 1 - \frac{\max_{i=1,\ldots,n}(p_i) - \frac{\sum_{i=1}^{n} p_i}{n}}{1 - \frac{1}{n}} \quad (1)$$

where p_i (i=1,…,n) are the probabilities associated with the several classes and n is the number of classes under consideration. Since in this study a pixel-based classification with a probabilistic Bayesian classifier was used, $\sum_{i=1}^{n} p_i = 1$. This uncertainty measure assumes values in the interval [0,1] and only depends on the maximum probability and the total number of classes. The numerator of the second term expresses the difference between the maximum probability assigned to a class and the probability that would be associated with the classes if a total dispersion for all classes occurred, that is, a probability of $\frac{1}{n}$ was assigned to all classes. The denominator expresses the extreme case of the numerator, where the maximum probability is one (and thus a total commitment to a single class occurs). The ratio of these two quantities expresses the degree of commitment to a specific class relative to the largest possible commitment (Eastman 2006). The classification uncertainty is thus the complement of this ratio and evaluates the degree of compatibility with the most probable class and until at which point the classification is dispersed over more than one class, providing information regarding the classifier's difficulty in assigning only one class to each pixel.

To illustrate the behaviour of this measure, some examples are shown in table 2. If, for one pixel, the maximum probability assigned to a particular class is one, the probabilities assigned to all other classes are zero, and a minimum value of zero is obtained for U (e.g., pixel p_1 on table 2). On the other hand, if the maximum probability is relatively low (e.g., pixel p_4 on table 2), since the probabilities have to add up to one, the probabilities associated to the other classes are still relatively high, and therefore the uncertainty increases. The maximum value of U is reached when equal probabilities are assigned to all classes (e.g., pixel p_{10} on table 2).

Notice that higher uncertainty values are obtained when lower values for the maximum probability occur and when the probabilities are dispersed over almost all classes. For example, for pixel p_3, with the probability distribution 0.8, 0.1, 0.1, the U measure gives an uncertainty value of 0.30 and for pixel p_4, with the same number of classes and probability distribution 0.4, 0.4, 0.2, the U measure uncertainty value increases to 0.90. This means that this uncertainty measure is sensitive to the compatibility between the classes and the pixel characteristics, expressed by the maximum probability. However, if the number of classes increases, for example to eight, even though the non-zero values of the probabilities distribution are the same, as for pixel p_7, the U measure decreases to 0.686. This means that this uncertainty measure is also sensitive to the dispersion of the classification over the total number of classes, expressed by the probabilities that are associated to the classes which do not correspond to the maximum probability. This is illustrated in table 2, for example by pixels p_4, p_6 and p_7, which have considerable different U values. This characteristic is due to the term $\frac{1}{n}$ that plays an important role in this measure, since it allows the analysis of the dispersion of the probability distribution. For example, pixels p_4, p_6, p_7, p_9 and p_{11} in table 2 have an equal maximum

probability and have different U values, because the term $\frac{1}{n}$ is different for each one of them. This uncertainty measure provides important information since: (1) the importance of dispersion increases when the maximum probability decreases (e.g., p_1, p_2, p_3, p_4) and the number of classes decreases (e.g., p_3, p_5, p_8, p_{12} and p_{13}; or p_4, p_6, p_7, p_9 and p_{11}); and (2) in the classification of remote sensing images the number of classes is, in general, not high and relatively low maximum probabilities may be obtained.

3.1.3 Landscape units map

The Landscape Units Map (LUM) was built combining the SEM, its uncertainty information and the objects obtained with the segmentation algorithm. In the segmentation stage the whole image was partitioned into a series of closed objects, corresponding to the spatial patterns. The objects extraction was driven using the "Fractal Net Evolution Approach" (FNEA) segmentation method, implemented in eCognition software, which can be described as a region merging technique (Baatz and Schape 2000).

This method starts with the assumption that each pixel is an object, and proceeds with the aggregation of neighbouring objects. The decision to fuse adjoining objects depends on the criteria of local homogeneity. The adjoining objects are fused into one if the spectral heterogeneity of the object resulting from the fusion does not exceed a certain maximum value, which determines the maximum heterogeneity. As a consequence, the size of the objects resulting from the fusion depends upon the value given to that parameter, called, for this reason, scale parameter (Baatz and Schape 2000). The drawback of this approach is that the final decision about the scale parameter is made by visual inspection of the image rather than by quantitative criteria. In this study only one segmentation level was considered, chosen from a series of experiments done with different parameters, whose results were visually analyzed. The criteria that led to their choice was the identification of meaningful image-objects i.e., groups of pixels that represented the LU existing in the study area, with a mean area of 0.5 ha. The parameters used in the chosen segmentation are shown in table 3.

The next step was the development of rules that incorporate the information provided by the previous pixel-based classification within each object and the results given by the uncertainty measure U. The rules construction requires a preliminary analysis of the uncertainty assigned to the SE classes in order to choose the appropriate thresholds.

The transformation of a SEM into a LUM is similar to a decision tree which, for geographical objects, is a hierarchical structure consisting of several levels. At each level a test is applied to one or more attribute values. The application of a rule results either in a leaf, allocating an object to a class, or a new decision node, specifying a further decision rule. The eight LU classes used in this study are presented in table 4.

Figure 2 shows the LU classes classification workflow and table 5 shows the classification rules. The aim of rule 1 is to make a distinction between 'Forest Areas' and 'Non-Forest Areas'. Rule 2 assigns the objects considered 'Non-Forest Areas", to one of the three LU classes: Water Bodies, Agriculture, and Non-Vegetated Areas. Rule 3 classifies the Forest regions into 'Dense Forest' and 'Non-Dense Forest'. Rule 4 assigns the objects classified as 'Dense Forest' to one of four possible LU classes, namely Broad-Leaved Forest, Coniferous Forest, Cork Forest and Mixed Forest. Finally, rule 5 assigns the objects considered 'Non-Dense Forest' to one of two possible LU classes: Agro-Forestry Areas and Mixed Forest.

For the accuracy assessment of remote sensing images classifications the three primary areal sampling units are pixels, polygons, and fixed-area plots and no consensus exists on which sampling unit is the best. According to Stehman and Czaplewski (1998), the choice of the sampling unit should be guided by the characteristics of the landscape, features of the mapping process, project objectives and practical constraints. Since in this case the objective of the mapping was to produce a polygon map where the LU had a mean area

of 0.5 ha, the choice of the sampling unit to assess the accuracy of the LUM was a fixed-area square plot sampling unit with an area of 0.5 ha.

A stratified random sampling of 50 samples per class was chosen, which guarantees a standard error of 0.07 for the CPM and CPR estimates for each class, assuming that the classification accuracy is superior to 50% (Wickham *et al*. 2004), which is acceptable because the construction of the LUM already involved a prior pixel-based classification and an analysis of the terrain. The accuracy assessment was made with an error matrix, where the p_{ij} entry is the proportion of area that is class i in the map and class j in the reference within the square areas with 0.5 ha. The CPM and CPR accuracy parameters were then derived from the error matrix (Stehman and Czaplewski 1998, Stehman and Czaplewski 2003).

3.2 Classification without uncertainty

To evaluate if the use of uncertainty in the classification improves the results, a classification method very similar to the previously described, but where the uncertainty is not considered, was applied to the same image. The identification and mapping of the SE were made with the Bayesian classifier. The SE classes and the sampling design were the ones used for the previous method and described in section 3.1.1. The rules developed to transform the SEM into a LUM are similar to the ones explained in section 3.1.3. but without considering the pixel-based classification uncertainty. The accuracy assessment was made with the same protocol used in the classification method explained above.

4. Results and discussion

The procedure to evaluate the pixel-based classifier and the accuracy of the SEM classification was the same for both classification methodologies and was performed using respectively testing sets 1 and 2. A global accuracy of 95% was obtained with testing set 1, and of 66% with testing set 2. The CPR and the CPM obtained for all classes with both testing sets are presented in figure 3 and figure 4.

The values obtained with testing set 1 indicate that the classifier was able to identify the pure classes correctly. However, the CPR and CPM obtained with testing set 2, where mixed pixels were considered, show worse results. The water classes (DW and SW) and Herbaceous Vegetation (HV) were well identified and the non forestry classes presented better results than forestry classes. Forestry species were often confused with several other SE, but mainly with Herbaceous Vegetation. Significant confusion was also observed between Cork Trees and Sparse Herbaceous and between Eucalyptus and Coniferous. This confusion was due to the proximity of their spectral signatures.

Figure 5 shows the average of the uncertainty measure U per class. The comparison of these results with figure 4, shows that the results are consistent, since forestry species present higher values of uncertainty than non forest species and lower values of accuracy. The forest species that show higher values of U and lower accuracy is the Eucalyptus Trees. The correlation between U and the classification accuracy was also evaluated. The correlation coefficient between U and CPM was 0.71 and between U and CPR was 0.39. These results reveal that there is a good agreement between the U measure and CPM, and therefore uncertainty measure U may be used to estimate the CPM.

The global classification accuracy of the LUM obtained with the classification method that includes uncertainty was 66%, and the one obtained with the classification method without uncertainty was 54%. Even though the accuracy is not high in both cases, due mainly to the difficulty in discriminating between the SE representing the forest species, a 12% increase was observed for the LUM obtained using the uncertainty information. Furthermore, the LUM obtained with this new approach also achieved better accuracy results than the SEM. This reveals that when SE with high values of U are excluded from the process of transforming

a SEM into a LUM, some of the LU classes were better identified, such as Broad-Leaved Forest (BF) and Coniferous Forest (CFF).

Figure 6 and figure 7 allow the comparison between the results of the CPR and CPM for the LUM obtained with both classification methods. The classification results obtained with the new method using uncertainty are considerably better for almost all LU classes and this improvement is more evident in the forest classes.

A comparison between figure 5, figure 6 and figure 7, shows that the LU classes showing better results when considering the uncertainty were the ones formed by the SE that presented higher values of uncertainty such as Broad-Leaved Forest (BF) and Coniferous Forest (CFF).

This new classification method proved that the uncertainty information allowed the identification of the misclassified SE and avoided their use in the transformation from a SEM into a LUM. Figure 8 shows an example of a LU which was only correctly classified as agriculture when the uncertainty of the SE classification was considered.

Figure 9 shows the final results of the classification with the proposed method. A relevant aspect of this hybrid classification methodology is that the final map is in the vector format, without the salt – and - pepper effect, which is caused mainly by a small percentage of isolated pixels in the results of the SEM often located at the limits between different SE.

5. Conclusions

The goal of this study was to highlight the influence and usefulness of the uncertainty information associated with the classification of Surface Elements into the classification of Landscape Units. The obtained results show that the hybrid pixel-object classification integrating the SE classification uncertainty increases by 12% the global accuracy of the classification of Landscape Unit classes in a Mediterranean environment, when compared to a similar classification method that does not take classification uncertainty into consideration. Therefore, the use of uncertainty proved to be valuable in the classification process. In fact, this strategy allowed the identification of the misclassified surface elements, avoiding their use in the construction of the Landscape Unit Map, and therefore this approach seems promising and worthy of further studies. From a methodological viewpoint, the hybrid approach also proved to be adequate for the transformation of a Surface Elements Map, with detailed land cover features, into a Landscape Unit Map with a vector format well suited for integration in a GIS.

Acknowledgments

The authors are grateful to the reviewers for their helpful comments, that contributed to improve the paper.

References

BAATZ, M., and SCHAPE, A., 2000, Multiresolution segmentation – an optimization approach for high quality multi-scale image segmentation. In *Angewandte Geographische Informationsverarbeitung XII*. Beitrage Zum AGIT – Symposium Salzburg 2000, Strobl, J. *et al.* (Ed.) pp. 12-23(Karlsruhe: Herbert Wichmann Verlag).

BASTIN, L., 1997, Comparison of fuzzy c-means classification, linear mixture modelling and MLC probabilities as tools for unmixing coarse pixels. *International Journal of Remote Sensing,* **18,** 3629-3648.

BLASCHKE, T., BURNETT, C. and PEKKARINEN, A., 2004, Image segmentation methods for object-based analysis and classification-including the spatial domain. In *Remote Sensing Image Analysis*, Jong, S. and van der Meer, F. (Ed.), pp. 211-234 (Dordrecht: Springer).

BROWN, M., LEWIS, H. and GUNN, S., 2000, Linear spectral mixing models and support vector machines for remote sensing. *IEEE Transactions on Geoscience and Remote Sensing*, **38**, 2346–2360.

CAETANO, M., 1995, Burned vegetation mapping in mountainous areas with satellite remote sensing. M. A. thesis. University of California, Santa Barbara.

CARLEER, A., and WOLFF, E., 2004, Exploitation of very high resolution satellite data for tree species identification. *Photogrammetric Engineering & Remote Sensing*, **70**, 135–140.

DOAN, H. T. X. and FOODY, G. M., 2007, Increasing soft classification accuracy through the use of an ensemble of classifiers. *International Journal of Remote Sensing*, **28**, 4609-4623.

DONNAY, J.P., BARNSLEY, M., and LONGLEY, P., 2001, Remote sensing and urban analysis. In *Remote Sensing and Urban Analysis*. J.P. Donnay, M. Barnsley P. Longley (Ed.), pp. 3–18 (New York: Taylor & Francis).

EASTMAN, J.R., 2006, Idrisis Andes Guide to GIS and Image Processing. Clark Labs., Clark University.

FOODY, G.M., 2000, Estimation of sub-pixel land cover composition in the presence of untrained classes. *Computers & Geosciences*, **26**, 469–470.

GOETZ, S. J., WRIGHT, R.K., SMITH, A.J., ZINECKER, E., and SCHAUB, E., 2003, IKONOS imagery for resource management: tree cover, impervious surfaces and riparian buffer analyses in the mid-atlantic region. *Remote Sensing of Environment*, **88**, 195–208.

GONÇALVES, L. and CAETANO, M., 2004, Classificação das imagens do satélite IKONOS utilizando uma abordagem orientada por objectos. In *Actas da Conferência de Cartografia e Geodesia*, L. Bastos e J. Matos (Ed.), pp.287–298 (Lisboa: Lidel).

GUERRERO I., TANASE, M., MANAKOS I., and GITAS, I., 2006, A semi-operational approach for land cover mapping in the Mediterranean. In *Proceedings of the 26th EARSeL Symposium: New developments and challenges in remote sensing*, May 29–2 June 2006, Warsaw, Poland.

IBRAHIM, M. A., ARORA, M. K., and GHOSH, S. K., 2005, Estimating and accommodating uncertainty through the soft classification of remote sensing data. *International Journal of Remote Sensing*, **26**, 2995–3007.

MASELLI, F., CONESE, C., FILIPPIS, T.D., and NORCINI, S., 1995, Estimation of forest parameters through fuzzy classification of TM data. *IEEE Transactions on Geoscience and Remote Sensing*, **33**, 77–84.

MASELLI, F., RODOLFI, A. and CONESE, C., 1996, Fuzzy classification of spatially degraded thematic mapper data for the estimation of sub-pixel components. *International Journal of Remote Sensing*, **17**, 537–551.

PLANTIER, T. and CAETANO, M., 2007. Mapas do Coberto Florestal: Abordagem Combinada Pixel/Objecto. In *Acta da V Conferência Nacional de Cartografia e Geodesia*, J. Casaca and J. Matos (Ed.), pp. 157–166 (Lisboa: Lidel).

SCHEIEWE, J., TUFTE, L., and EHLERS, M., 2001, Potential and problems of multi-scale segmentation methods in remote sensing. *GeoBIT/GIS*, **6**, 34–39.

STEHMAN, S.V., 2001, Statistical rigor and practical utility in thematic map accuracy assessment. *Photogrammetric Engineering & Remote Sensing*, **67**, 727–734.

STEHMAN, S.V. and CZAPLEWSKI, R.L., 2003, Introduction to special issue on map accuracy. *Environmental and Ecological Statistics*, **10**, 301–308.

STEHMAN, S.V. and CZAPLEWSKI, R.L., 1998, Design and analysis for thematic map accuracy assessment: fundamental principles. *Remote Sensing of Environment,* **64**, 332–334.

WANG, L., SOUSA, W. P., and GONG, P., 2004, Integration of object-based and pixel-based classification for mapping mangroves with IKONOS imagery. *International Journal of Remote Sensing*, **20**, 5655–5668.

WICKHAM, J.D., STEHMAN, S.V., SMITH, J.H., WADE,T.G. and YANG,L., 2004, A priori evaluation of two-stage cluster sampling for accuracy assessment of large-area land cover maps. *International Journal of Remote Sensing***, 25,** 1235–1252.

ZHANG, J. and FOODY, G. M., 2001, Fully-fuzzy supervised classification of sub-urban land cover from remotely sensed imagery: statistical and artificial neural network approaches. *International Journal of Remote Sensing*, **22**, 615–628.

SHI, W.Z., EHLERS, M. and TEMPFLI, K., 1999, Analytical modelling of positional and thematic uncertainties in the integration of remote sensing and geographical information systems. *Transactions in GIS*, **3**, 119–136.

Table 1. Acquisition details of IKONOS image.

Date	06/04/2004
Sun elevation (deg)	74.8
Sensor elevation (deg)	55.5
Dimension (m x m)	11884 x 14432
Bits/pixel	11

Table 2. Classification uncertainty (CU) values obtained for some probability distributions

i	Ordered probability distributions for pixel p_i	$\max_{i=1..n}(p_i)$	n	U
1	1;0;0	1	3	0
2	0.9;0.1;0	0.9	3	0.150
3	0.8;0.1;0.1	0.8	3	0.300
4	0.4;0.4;0.2	0.4	3	0.900
5	0.8;0.1;0.1;0;0	0.8	5	0.250
6	0.4;0.4;0.2;0;0	0.4	5	0.750
7	0.4;0.4;0.2;0;0;0;0;0	0.4	8	0.686
8	0.8;0.1;0.1;0;0;0;0;0	0.8	8	0.229
9	0.4;0.3;0.2;0.1;0;0;0;0;0;0	0.4	10	0.667
10	0.1;0.1;0.1;0.1;0.1;0.1;0.1; 0.1; 0.1;0.1	0.1	10	1
11	0.4;0.3;0.2;0.1;0;0;0;0;0;0;0;0;0;0;0	0.4	15	0.643
12	0.8;0.1;0.1;0;0;0;0;0;0;0;0;0;0;0;0	0.8	15	0.214
13	0.8;0.1;0.1;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0	0.8	30	0.206

Table 3. Segmentation parameters.

Band Weights				Scale	Color	Shape	Smoothness	Compactness
Red	Blue	Green	Nir					
1	0.5	1	1	250	0.9	0.1	0.9	0.1

Table 4. Landscape Unit classes description and label used.

Class	Label	Description
Non-Vegetated Areas	NVA	Built –up areas, transport network and bare soil areas.
Agriculture and pastures areas	A	Annual crops and grasslands
Water Bodies	WB	Natural or artificial water bodies
Broad-Leaved Forest	BF	Vegetation formation composed mainly of trees, and its understory, where broad-leaved species predominate.
Coniferous Forest	CFF	Vegetation formation composed mainly of trees, and its understory, where coniferous species predominate.
Cork Forest	CKF	Vegetation formation composed mainly of trees, and its understory, where cork species predominate.
Agro-Forestry Areas	AFA	Annual crops or grazing land and fallow land covering less than 50 % of the surface. Agricultural land shaded by palm trees in Mediterranean context.
Mixed Forest	MF	Vegetation formation composed mainly of trees, including shrub and bush understory, where neither broad-leaved nor coniferous species predominate.

Table 5. Classification Rules.

Rules	Test	Class if true
Rule 1	Objects have more than 10% of SE classified as tree crowns, regardless of species, with uncertainty less than 0.25	Forest
	Objects do not satisfy the previous test	Non-Forest
Rule 2	The mode of the SE, inside the object, with uncertainty less than 0.25 is Deep Water or Shallow Water	Water Bodies
	The mode of the SE, inside the objects, with uncertainty less than 0.25 is Herbaceous Vegetation or Sparse Herbaceous Vegetation	Agriculture
	The mode of the SE, inside the objects, with uncertainty less than 0.25 is Non-Vegetated Area or Shadow	Non-Vegetated Areas
Rule 3	Objects have more than 75% of trees with uncertainty less than 0.25, regardless of the species	Dense Forest
	Objects do not satisfy the previous test	Non-Dense Forest
Rule 4	Eucalyptus Trees represent more than 75% of the trees	Broad-Leaved Forest
	Coniferous Trees represent more than 75% of the trees	Coniferous Forest
	Cork Trees represent more than 75% of the trees	Cork Tree Forest
	Neither Eucalyptus nor Coniferous species predominates	Mixed Forest
Rule 5	The percentage of trees is less than 50%; the percentage of Herbaceous or Sparse Herbaceous is superior than Cork Trees and 80% of trees is Cork Trees with uncertainty less than 25%	Agro-Forestry Areas
	Objects do not satisfy the previous test	Mixed Forest

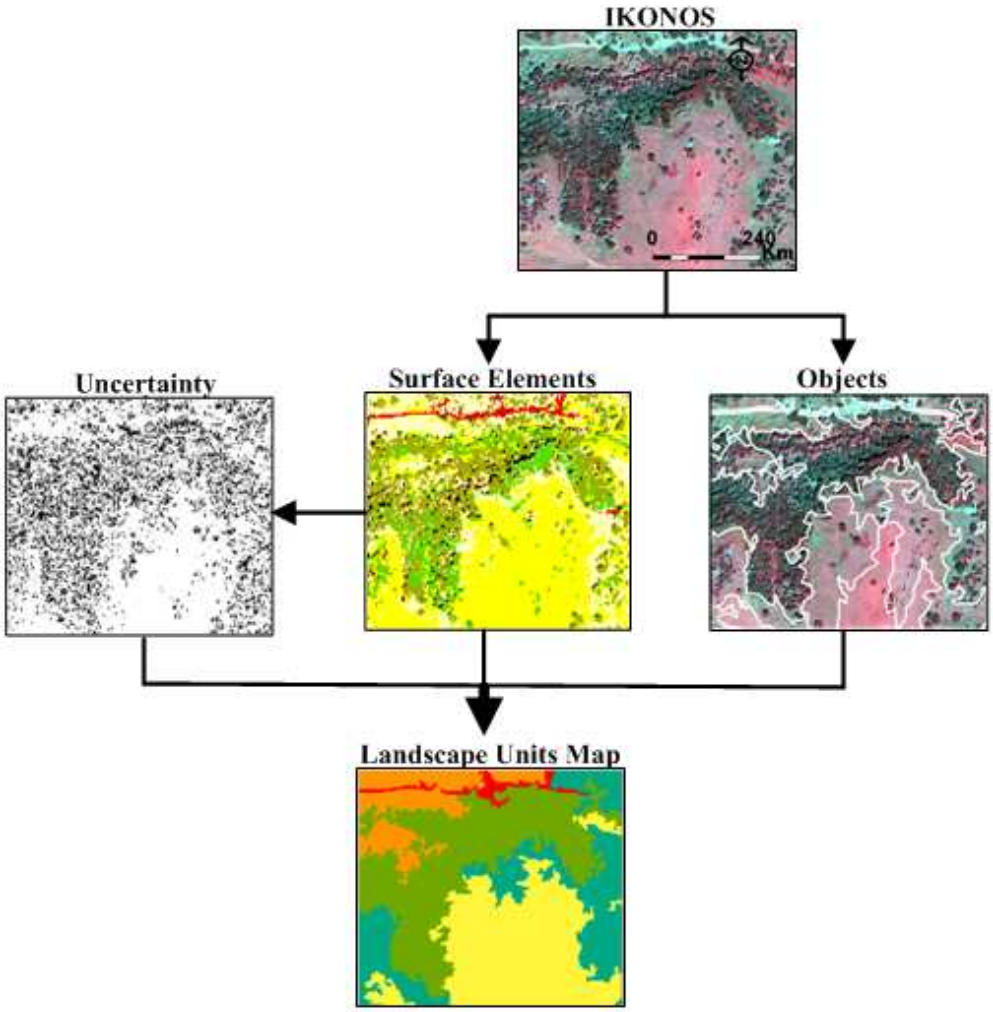

Figure 1. Flowchart of the methodology to integrate uncertainty into the classification of an IKONOS image.
21x22mm (600 x 600 DPI)

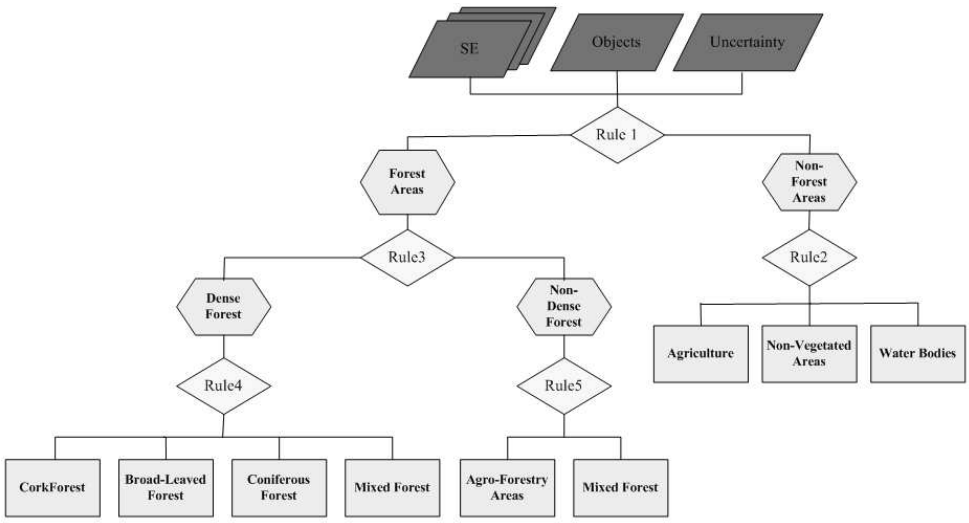

Figure 2. Landscape Unit Classes classification workflow.
40x21mm (600 x 600 DPI)

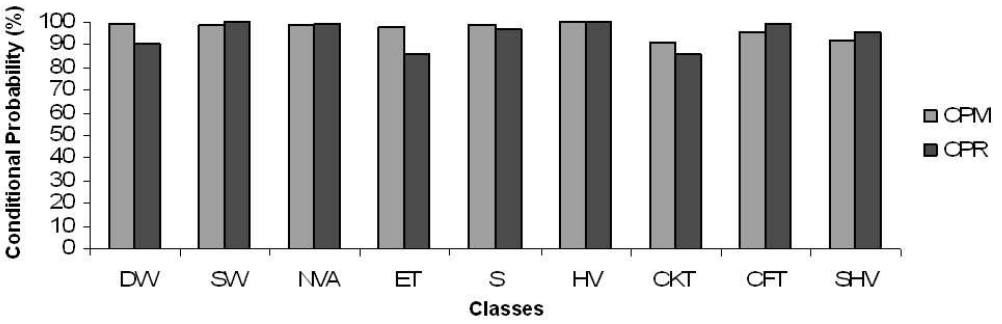

Figure 3. Conditional Probability of the Reference (CPR) and Conditional Probability of the Map (CPM) of the Surface Elements Map (SEM) obtained with testing set 1 (used to evaluate the classifier accuracy).
40x14mm (600 x 600 DPI)

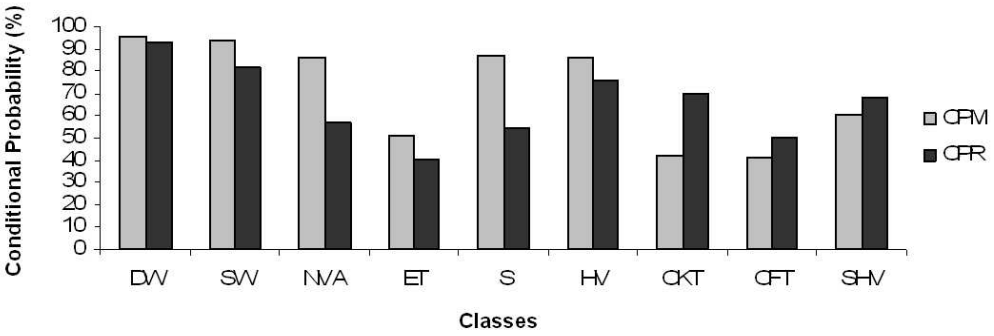

Figure 4. Conditional Probability of the Reference (CPR) and Conditional Probability of the Map (CPM) of the Surface Elements Map (SEM) obtained with testing set 2 (used to evaluate the classification accuracy).
44x16mm (600 x 600 DPI)

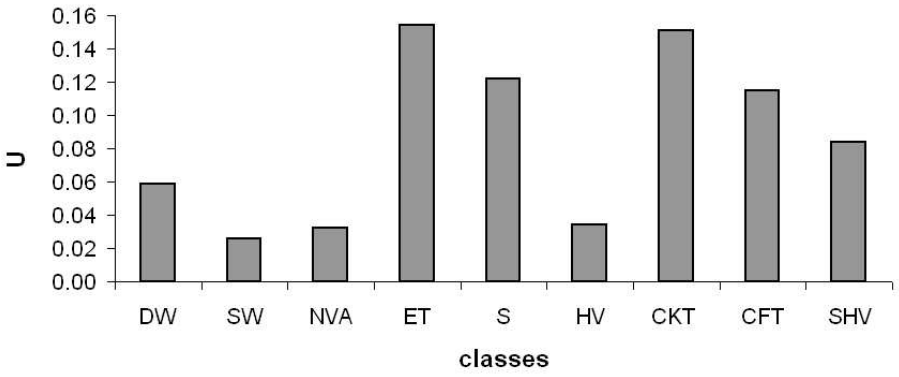

Figure 5. Average values of the uncertainty measure (U) per class of the Surface Elements Map (SEM).
39x19mm (600 x 600 DPI)

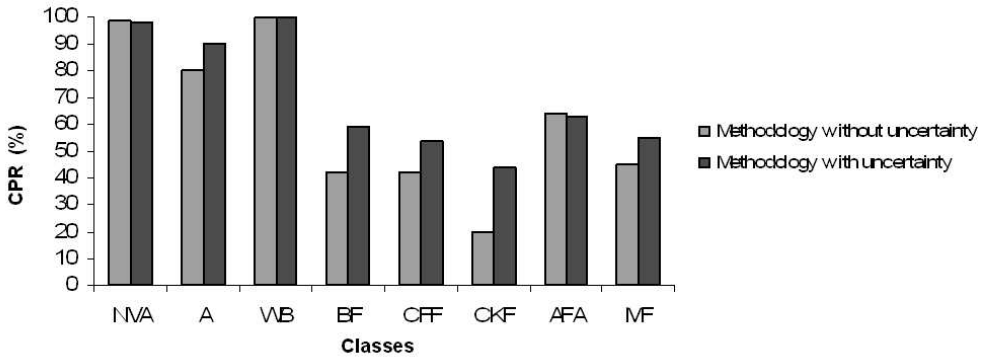

Figure 6. Conditional Probability of Reference (CPR) obtained with the hybrid approach (LUM) with and without uncertainty.
41x16mm (600 x 600 DPI)

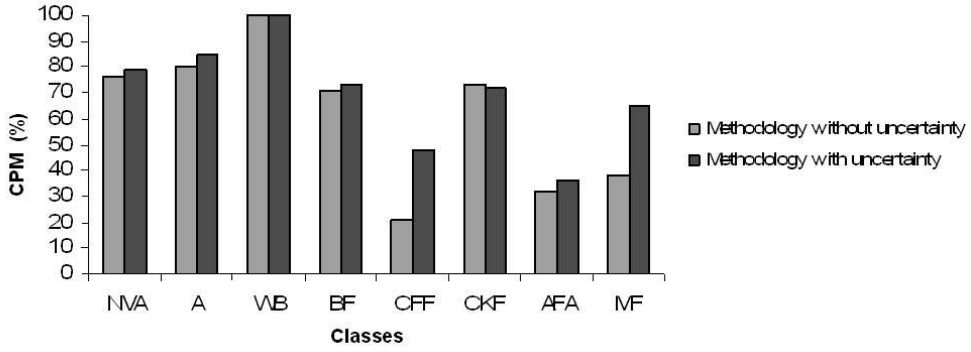

Figure 7. Conditional Probability of Map (CPM) obtained with the hybrid approach (LUM) with and without uncertainty.
41x16mm (600 x 600 DPI)

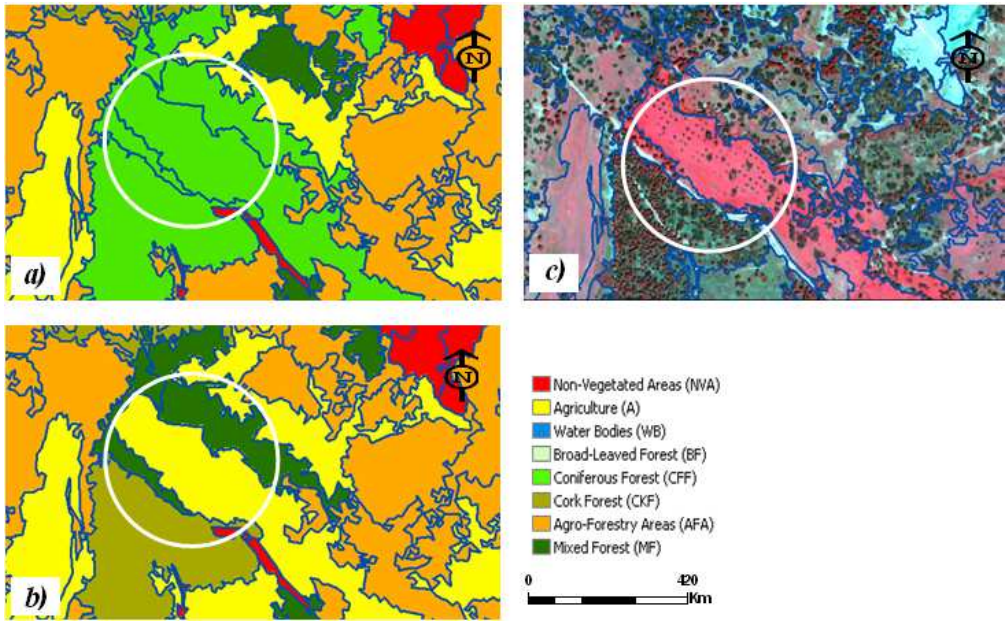

Figure 8. LUM obtained with both methodologies and an extract of the IKONOS image (RGB432). a) classification obtained without uncertainty; b) classification obtained with uncertainty; c) extract of the IKONOS image RGB432. The circles mark a land unit occupied by agriculture that was correctly classified when the uncertainty was considered in the classification rules.
30x19mm (600 x 600 DPI)

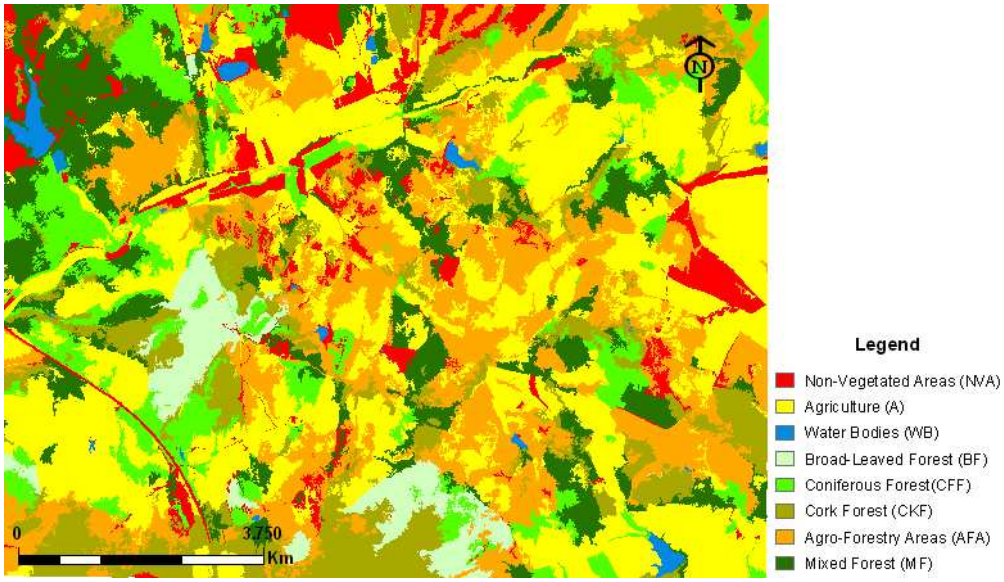

Figure 9. Final Land Units Map
33x18mm (600 x 600 DPI)