Assessing the spatial variability of classification accuracy using uncertainty information

Cidália C Fonte
Department of Mathematics
University of Coimbra /
INESC Coimbra
Coimbra, Portugal
cfonte@mat.uc.pt

João Apolinário
Department of Mathematics
University of Coimbra
Coimbra, Portugal

Luísa M S Gonçalves
Polytechnic Institute of Leiria
Civil Engineering Department /
INESC Coimbra
Coimbra, Portugal
luisa.goncalves@ipleiria.pt

Abstract

In this paper an approach is presented that enables the assessment of the spatial variability of the classification accuracy of a land cover map generated with soft classifiers, using the uncertainty information associated with the classification. Soft classifiers enable the computation of uncertainty measures for each pixel of the classified image. This information is used to identify regions with different levels of uncertainty. The accuracy of these regions is then evaluated separately, building spatially constrained confusion matrices, instead of confusion matrices for the whole image. This enables the identification of regions of the map with different levels of accuracy, allowing the computation of accuracy indices for different regions. The methodology is applied to two sets of multispectral images with 4m and 20m spatial resolutions, using two soft classifiers in each case. The results show that lower levels of accuracy are in general obtained for the regions with higher levels of uncertainty, and therefore the uncertainty information is a valuable indicator to assess the spatial variability of the classification accuracy.

Keywords: accuracy, uncertainty, soft classifiers, multispectral images.

1 Introduction

The accuracy assessment of land cover maps is usually performed building confusion matrixes with the classification results and reference data for a sample of points. From these matrices accuracy indices may be computed, such as the overall accuracy and the user's and producer's accuracy for each class [7]. However, since one confusion matrix is built to the whole map, these indices are applied to the whole area and, even though there may be different levels of accuracy in different regions of the map, this variability cannot be estimated with this methodology. Some developments addressing this limitation of accuracy assessment can be found in [3, 2, 1].

The aim of the study herein presented is to determine if the identification of regions with different degrees of uncertainty may be an indicator to find regions where different levels of accuracy are more likely to occur. To this aim, the classification uncertainty is computed, and regions with different levels of uncertainty are identified in the uncertainty image. The classification accuracy within each region is then assessed separately, enabling the construction of spatial constrained confusion matrixes and the computation of independent accuracy indices for the regions with different levels of uncertainty. The proposed methodology is applied to two sets of multispectral images with different spatial resolutions, using two soft classifiers, to determine if the results are consistent in all cases.

2 Methodology

The methodology proposed in this paper consists of the following steps: 1) Multispectral image classification with a soft classifier; 2) Computation of the classification uncertainty using an uncertainty index; 3) Generation of a hard version of the classification produce in step 1; 4) Identification of regions with different levels of uncertainty; 5) Generation of confusion matrixes for each region obtained in the previous step using reference data.

2.1 Classification

The proposed methodology requires the classification of the multispectral image with a soft classifier. Two distinct classifiers were used for this study: 1) A Bayesian Classifier (BC), corresponding to the soft version of the maximum likelihood classifier, which uses the posterior probabilities as indicators of the probability of each pixel belonging to each class [3]; 2) A classifier based on the Dempster-Shafer theory (DSC) [6]. This theory considers the existence of ignorance and enables the computation of degrees of belief and plausibility of the membership of each pixel to each class, where beliefs reflect the degree to which the evidence supports the assignment of each pixel to each class and plausibilities the degree to which the evidence does not support the opposite hypothesis (assignment of the pixel to the other classes). Degrees of belief were used in this study.

A traditional hard classification can be obtained when the class corresponding to the higher probability or belief value for each pixel is assigned to the pixel.

2.2 Classification uncertainty

The classification uncertainty was computed considering the Relative Maximum Deviation Measure (RMD) given by equation (1), where $m_i(x)$ are the degrees of assignment of

pixels x to class i (posterior probabilities in the case of the BC and beliefs for the DSC) and n is the number of classes.

$$RMD(x) = 1 - \frac{\max_{i=1,..n}\left[m_i(x)\right] - \frac{\sum_{i=1}^{n} m_i(x)}{n}}{1 - \frac{1}{n}} \qquad (1)$$

This uncertainty measure assumes values in the interval [0,1] and evaluates the degree of compatibility of the chosen class with a perfect match, corresponding to $m_i(x) = 1$. Since this value is computed to every pixel of the image, an image reflecting the classification uncertainty is obtained.

2.3 Identification of regions with diferent levels of uncertainty

To identify regions with different levels of uncertainty, a three steps procedure is used: firstly a segmentation algorithm is applied to the uncertainty image to generate objects corresponding to the patterns visually identified in the uncertainty image, grouping pixels with similar values of uncertainty; then, the mean uncertainty of the pixels included in each object is computed; and thirdly the objected are grouped into three regions corresponding to low, medium and high levels of mean uncertainty.

The segmentation algorithm used in this study, available in software IDRISI, groups adjacent pixels into objects according to their similarity. The variance of the input image is computed for each pixel using a moving window, whose size may be defined by the user. The values of the variance image are treated as elevation values in a digital elevation model and a watershed delineation process is used to group pixels into objects. An iterative process to merge adjacent objects is then used. Objects are merged if they satisfy simultaneously the following conditions: 1) the objects are adjacent; 2) they must be mutually most similar and 3) the similarity must be less than a user-specified threshold, referred to as similarity tolerance. The similarity is evaluated computing a difference between objects using the mean and standard deviation of the pixel values belonging to each object as well as user defined weights for both the mean and the standard deviation. For the objects to be merged this difference must be less than the similarity tolerance, which may take values equal or larger than zero. For a value of zero no merging is done and the larger the value the larger the degree of generalization, resulting in larger objects [5].

Once the objects are identified, the mean uncertainty of the pixels located inside the objects is computed. Since a large number of objects is likely to be obtained, to identify a limited number of regions to assess the accuracy, the objects were aggregated into three regions, considering the objects mean uncertainty.

2.4 Accuracy assessment

To assess the classification accuracy within each of the obtained regions, a stratified sample with 50 points per class was used for each uncertainty region. A confusion matrix was then built for each region, enabling the construction of confusion matrixes constrained to these regions, and the computation of independent accuracy indices, namely the overall accuracy and the user's and producer's accuracy per class. The accuracy of the hardened classification of the complete image was evaluated building a confusion matrix with all the points collected for the three levels of uncertainty considered, corresponding to a total of 750 points per classification and per multispectral image.

3 Case study

3.1 Data

Two sets of multispectral images were used in this study (Figure 1). A CARTERRA-Geo image of 2000 obtained by the IKONOS-2 sensor, with a spatial resolution of 4m in the multispectral mode (XS), and a SPOT-4 image of 2006, with a spatial resolution of 20m. The study was performed using 4 multispectral bands in both cases.

Both multispectral images cover an area of 81.5 km^2 located near the Portuguese coast, and includes regions with different characteristics, such as built up areas, agricultural fields and forest. The nomenclature used in this study considers the classes Urban Areas (UA), Herbaceous Vegetation (HV), Shrub Lands (SL), Forest Areas (FA) and Barren Areas (BA).

Figure 1: False color representation of the multispectral images: a) IKONOS-2 sensor, b) SPOT-4.
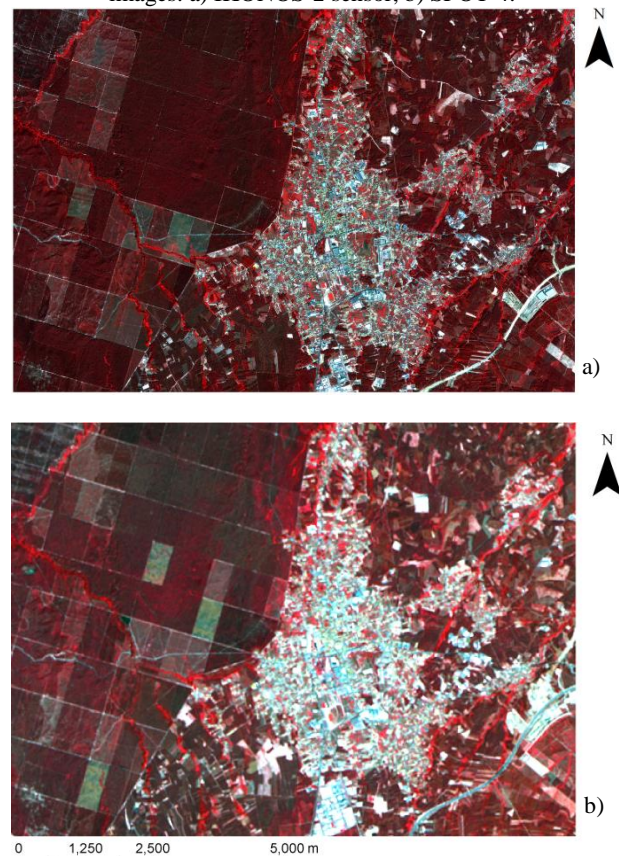
3.2   Results

Figure 2 shows the classification results obtained with both classifiers for the two sets of multispectral images. Considerable differences in the classification results with both classifiers can be observed, mainly for the SPOT image, where some regions are clearly wrongly classified. Figure 3 shows the uncertainty values obtained with each classification. The uncertainty values are much higher for the DSC than for the BC, since the maximum value of beliefs is in most cases considerably low.

Figure 2: Hardened classification of the multispectral images of Figure 1: IKONOS image using the BC (a) and the DSC (b); SPOT image with the BC (c) and the DSC (d).
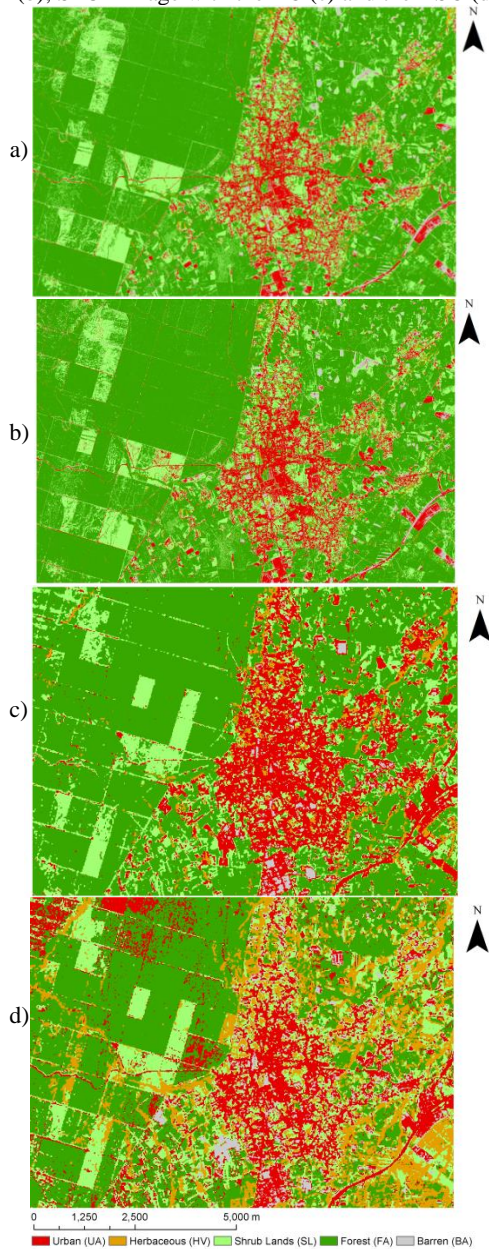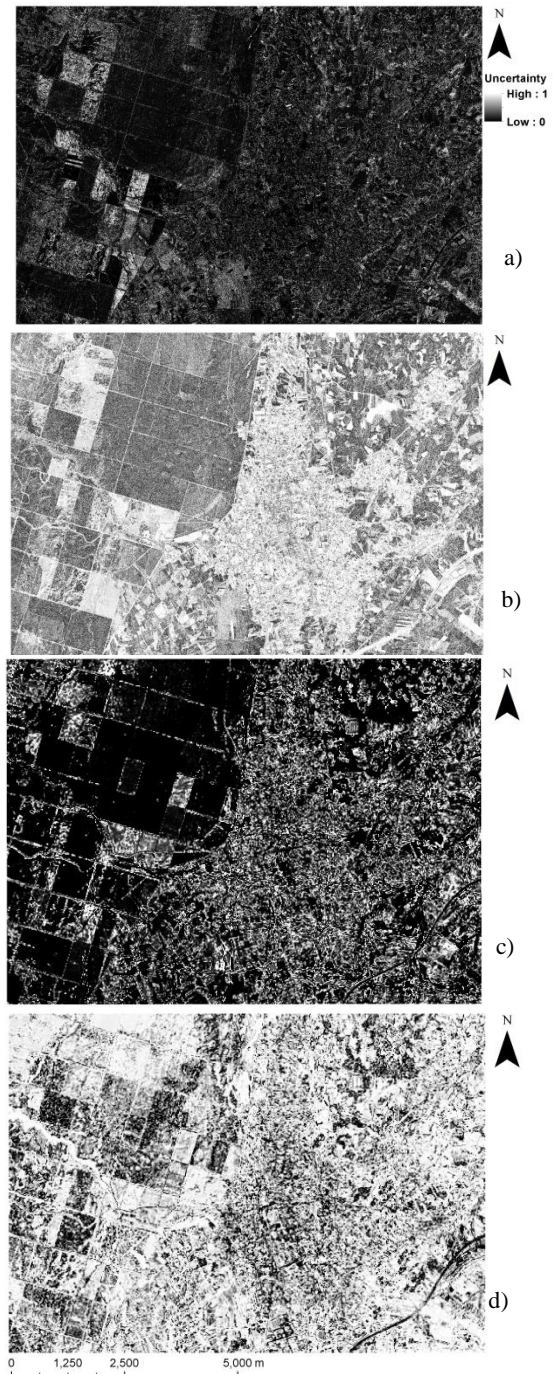

Figure 3: Uncertainty of the classification of: a) IKONOS using the BC, b) IKONOS with the DSC, c) SPOT with the BC and d) SPOT with the DSC.

The segmentation was performed considering a three by three pixels moving window, a weight of 0.5 for the variance and mean and a similarity tolerance of 100 for the IKONOS image and of 50 for the SPOT image. The segmentation parameters were chosen so that the patterns visually identified in the uncertainty image were captured by the segmentation. The mean uncertainty of the pixels located inside each object was

computed. Figure 4 shows the segmentation of the uncertainty obtained with the BC for the IKONOS image and the SPOT image. The shades of grey represent the objects mean uncertainty.

Figure 4: Images a) and b) show respectively the results of the segmentation of the images represented in Figure 3a) and 3c). The shades of grey represent the objects mean uncertainty.

a)

b)
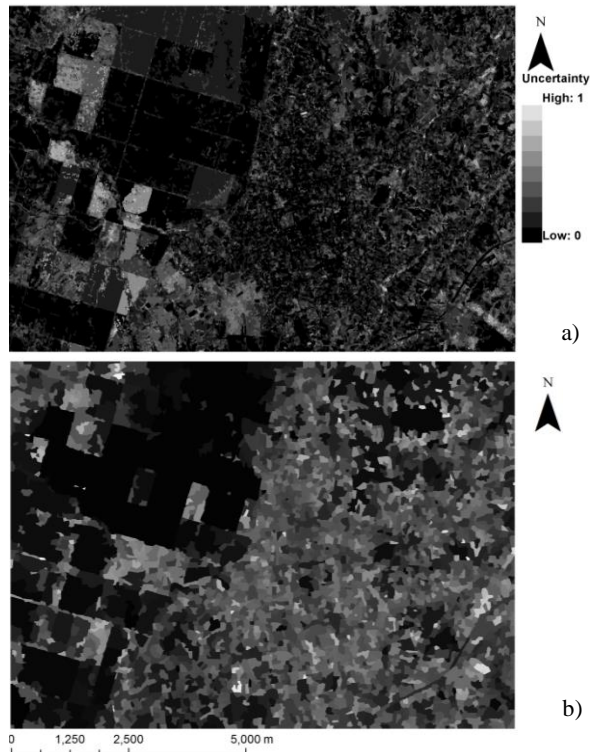
The segmentation process produced in the four cases considered several thousands of objects. To separate the study area in three regions with different levels of uncertainty, the objects were aggregated into low, medium and high levels of uncertainty. Several mean uncertainty threshold values may be used to identify these regions. Within this paper, the objects grouping was made considering mean uncertainty thresholds that enabled the inclusion of the same number of objects into the three regions. Figure 5 show the regions with low and high uncertainty obtained for the classification of the IKONOS image with the BC (respectively 5a) and 5b)) and for the SPOT image with the DSC (respectively 5c) and 5d)). A discrepancy between the size of the areas contained in each level of uncertainty occurs because the objects obtained with the segmentation have in some cases very different areas.

The results show that the classes are not evenly distributed by all levels of uncertainty. For example, it can be easily seen in Figure 5b) that most pixels located in the region with high uncertainty were classified as belonging to the classes FA and SL and most of the UA were located in the regions with low uncertainty. In Figure 5c) it can be seen that most of the UA of the study area (see Figure 1b) were classified with the lower levels of uncertainty, even though some UA are found in the regions with high uncertainty, which are clearly misclassified.

Figure 5: Regions with low (a) and high (b) uncertainty obtained for the IKONOS image with the BC and regions with low (c) and high (d) uncertainty obtained for the SPOT image with the DSC. The class assigned to each pixel by the classifier is represented.

a)

b)

c)

d)
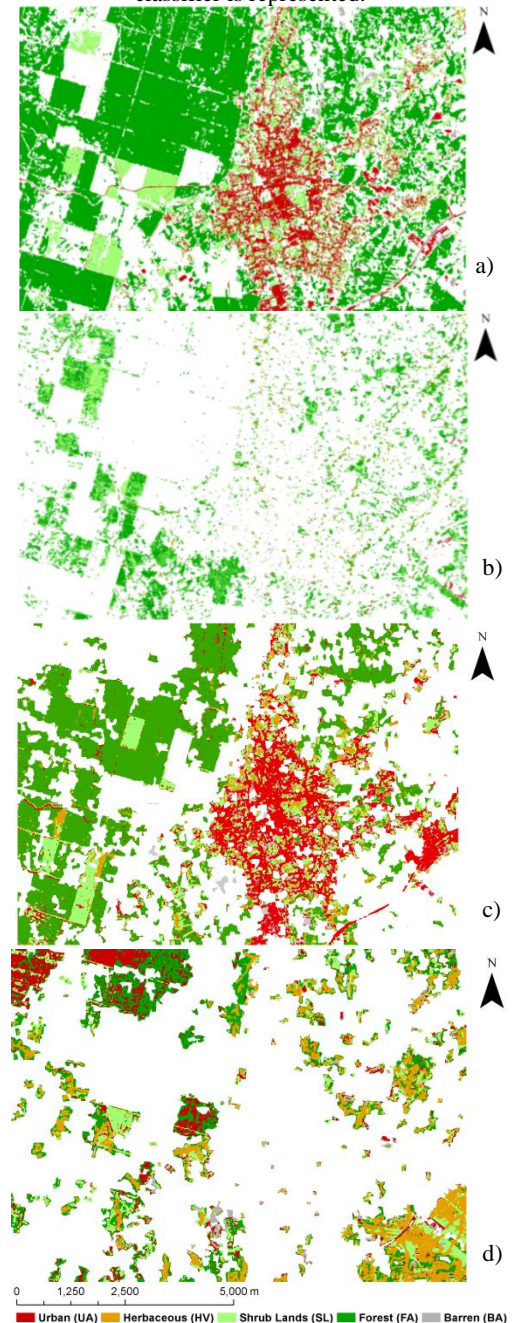
The classifications accuracy was then computed building one confusion matrix for each level of uncertainty of the classification with both classifiers and both multispectral images, corresponding to a total of eight confusion matrixes. Table 1 shows the overall accuracy as well as the user's and producer's accuracy per class obtained for the complete image and the regions with low, medium and high levels of uncertainty, for both multispectral images and both classifiers.

Table 1: Accuracy results.

IKONOS (Bayesian Classifier)

	All image	Uncertainty		
		Low	Medium	High
Overall accuracy (%)	66	78	65	54
User's Accuracy (%) UA	59	76	58	42
HV	64	74	68	50
SL	69	78	68	60
FA	78	92	74	68
BA	59	68	56	52
Producer's Accuracy (%) UA	67	69	71	57
HV	92	95	87	96
SL	50	63	47	41
FA	65	87	66	47
BA	71	83	68	63

IKONOS (Dempster-Shafer Classifier)

	All image	Uncertainty		
		Low	Medium	High
Overall accuracy (%)	60	69	61	53
User's Accuracy (%) UA	68	76	72	58
HV	53	52	56	56
SL	51	50	48	56
FA	68	90	70	44
BA	62	76	60	50
Producer's Accuracy (%) UA	68	81	61	66
HV	79	100	85	67
SL	47	81	49	33
FA	49	47	50	54
BA	73	76	77	64

SPOT (Bayesian Classifier)

	All image	Uncertainty		
		Low	Medium	High
Overall accuracy (%)	59	68	59	49
User's Accuracy (%) UA	67	74	70	58
HV	34	42	38	22
SL	64	68	62	62
FA	76	92	82	58
BA	51	66	44	44
Producer's Accuracy (%) UA	66	82	57	63
HV	91	100	100	69
SL	43	53	44	36
FA	53	58	57	44
BA	75	83	81	63

SPOT (Dempster-Shafer Classifier)

	All image	Uncertainty		
		Low	Medium	High
Overall accuracy (%)	51	61	51	40
User's Accuracy (%) UA	42	68	44	14
HV	32	34	38	24
SL	57	62	58	52
FA	79	86	80	70
BA	45	56	36	42
Producer's Accuracy (%) UA	64	69	65	44
HV	79	85	73	80
SL	39	50	38	31
FA	41	49	45	32
BA	82	88	75	81

The results show that the overall accuracy of the hardened version of the classifications with both classifiers for both images decreases with the increase of uncertainty. Moreover, the overall accuracy of the total image is in all cases between the overall accuracy of the regions with low and high uncertainty, and equal or very close to the overall accuracy of the regions with mean uncertainty.

In general, the user´s accuracy of all classes also decreases for the regions with higher uncertainty, with only a few exceptions. The same trend is also observed for the producer's accuracy, although with a few more exception.

Figure 6 shows a detail of the regions with low, medium and high levels of uncertainty obtained with the BC for the IKONOS image, where the classes assigned to the pixels by the classifier are shown. It can be seen that the regions classified with lower uncertainty are more homogeneous. The asphalt road is very well identified, as well as some parcels of shrubs and forest. In the larger forest zone present in the region with medium uncertainty (on the right of Figure 6c), smaller vegetation exists, corresponding to young trees. In the regions of higher uncertainty, the largest regions of forest and shrubs correspond in reality to vegetated regions with small trees and shrubs, where considerable confusion exists between both.

Figure 6: – Detail showing the IKONOS false color image (a), the regions with b) low, c) medium and d) high uncertainty of the image classification with the BC.

a) b)

c) d)

🟥 Urban Areas (UA) 🟧 Herbaceous Vegetation (HV)
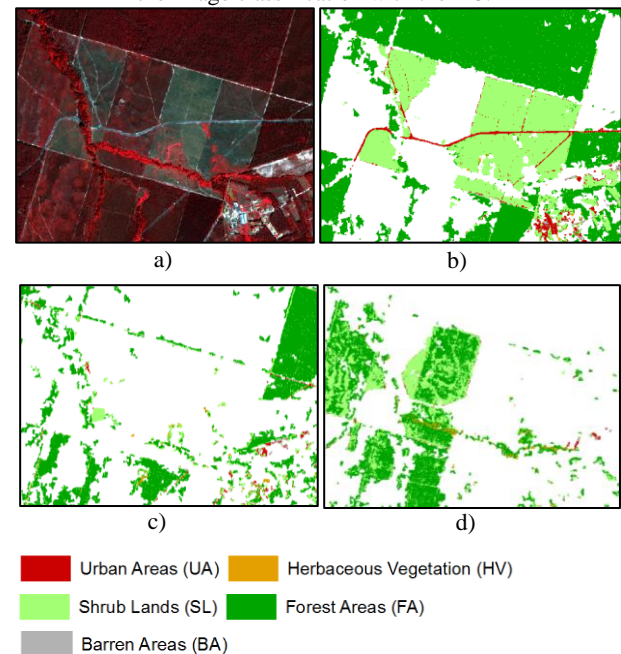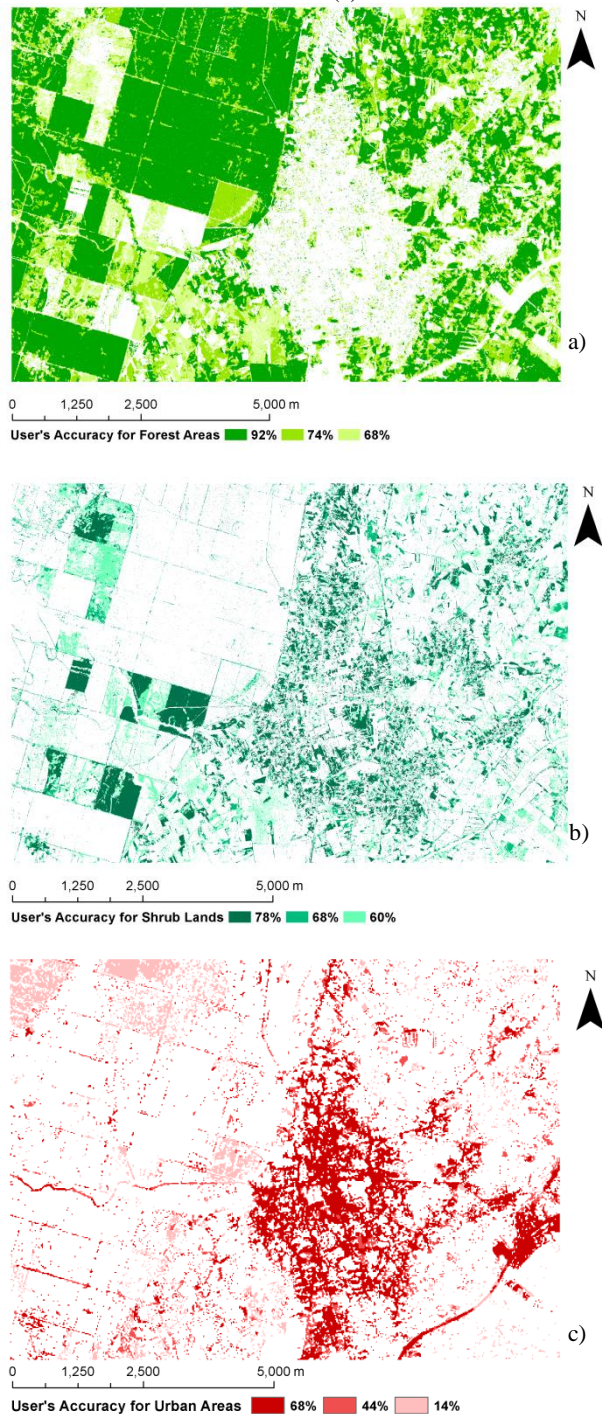🟩 Shrub Lands (SL) 🟩 Forest Areas (FA)
⬜ Barren Areas (BA)

Figure 7 shows the regions presenting different levels of user's accuracy of FA and SL for the IKONOS image classification with the BC, and the user's accuracy of UA for the SPOT image classification with the DSC. It can be clearly seen that there are regions in the image which were assigned by the classifiers to the same class, which in reality have different characteristics, resulting in a classification with different levels of accuracy.

Figure 7: Regions with different levels of user's accuracy for the classification of the: IKONOS image with the BC for the classes FA (a) and SL (b); SPOT image with the DSC for the class UA (c).

a)

0   1,250   2,500   5,000 m

User's Accuracy for Forest Areas 92% 74% 68%

b)

0   1,250   2,500   5,000 m

User's Accuracy for Shrub Lands 78% 68% 60%

c)

0   1,250   2,500   5,000 m

User's Accuracy for Urban Areas 68% 44% 14%

4   Conclusions

The proposed methodology uses the classification uncertainty computed with the information provided by soft classifiers to identify regions with different levels of accuracy.

The results showed that in the four cases analysed, using images with different spatial resolutions and different soft classifiers, the overall accuracy of the regions with low, medium and high levels of uncertainty decreases with the increase of uncertainty. A similar trend is also observed for the user's and producer's accuracy per class. However, for these indices some exceptions were found, mainly for the producer's accuracy. A more close analysis of the obtained regions and the confusion matrixes is in these cases necessary, to identify the causes of the discrepancies.

The results suggest that the levels of uncertainty computed with the RMD uncertainty index may be used as indicators of the classification accuracy, as it is more likely to have misclassifications in the regions with larger values of uncertainty. Therefore, with the prior knowledge of the regions where different levels of accuracy are expected, it is possible to establish geographically constrained confusion matrixes, which provide information on the spatial distribution of the classification accuracy. The proposed methodology showed therefore to be a promising approach to identify regions with different levels of accuracy of a hardened version of a classification performed with soft classifiers. This information may be quite useful, for example, to report the limitations of a land cover map or to identify regions with different characteristics within the same class.

References

[1] A. Comber, P.F. Fisher, C. Brunsdon and A. Khmag. Spatial analysis of remote sensing image classification accuracy. *Remote Sensing of Environment*, 127: 237–246, 2012.

[2] C.C. Fonte and L.M.S. Gonçalves. Assessing the spatial variability of the accuracy of multispectral images classification using the uncertainty information provided by soft classifiers. In *Proceeddings of the 7th International Symposium on Spatial Data Quality*, Coimbra, Portugal, 2011.

[3] G.M. Foody, N.A. Campbell, N.M. Trodd and T.F. Wood. Derivation and applications of probabilistic measures of class membership from maximum likelihood classification. *Photogrammetric Engineering and Remote Sensing*, 58: 1335-1341, 1992.

[4] G.M. Foody. Local characterization of thematic classification accuracy through spatially constrained confusion matrices. *International Journal of Remote Sensing*, 26: 1217-1228, 2005.

[5] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 3rd ed., Prentice Hall, New Jersey, 2008.

[6] L. I. Kuncheva. *Fuzzy Classifier Design*. Physica-Verlag, Springer-Verlag, 2000.

[7] S. V. Stehman and L. Czaplewski. Design and Analysis for Thematic Map Accuracy Assessment: Fundamental Principles. *Remote Sensing of Environment*, 64: 331-344, 1998.