COMPARING SATELLITE IMAGERY CLASSIFICATION METHODS IN RIVER SEGURA BASIN (MURCIA REGION)

Nicolás Del Toro Espín
Francisco Gomariz-Castillo[1]
Fulgencio Cánovas-García[2]
Francisco Alonso-Sarría[2]
1 Departamento de Ciencias del Mar y Biología aplicada. Universidad de Alicante.
Instituto Euromediterráneo del Agua (Campus de la Merced Universidad de Murcia)
2 Departamento de Geografía de la Universidad de Murcia.
Instituto del Agua y del Medio Ambiente de la Universidad de Murcia.
francisco.gomariz@ua.es, fulgencio.canovas@um.es, alonsarp@um.es

I. INTRODUCTION

Supervised classification of remote sensing imagery consists of 4 steps:
1. To establish the number of classes and to characterise them
2. To calibrate a classification model using training areas
3. To validate the classification model using validation areas
4. To use the model to obtain a final land use map

Class characterisation is generally achieved by using reflectivity values, although it is also possible to include textural and contextual features. Zhou and Robson (2001) claim that using textural features is necessary to obtain an accurate classification. In addition, Berberoglu et al. (2007) highlight the importance of such information, especially in Mediterranean landscapes that usually show a high variety and fragmentation in their land use patterns. The objective of contextual classification is usually to obtain larger and more cohesive land use polygons, minimising the speckle pattern (Tso and Mather, 2009).

Maximum likelihood is one of the most used classification models (Chuvieco, 2006; Tso and Mather, 2009). However, because of its limitations, a new set of machine learning techniques, such as Random Forest, have been been adopted.

The validation of a classification model is usually achieved using a confusion matrix and the kappa index (Congalton and Green, 1999). One of the most interesting criticisms of this approach is that the results of a classification model should not be compared with a random

classification (as kappa index does) but with what Pontius and Millones (2011) call a naive classification method that would be used as base classification method.

II. OBJECTIVES

In this study we tried to classify land use in the River Argos Basin from a Landsat TM image using three different classification methods: Maximum likelihood, Sequential Maximum a Posteriori, and Random Forest. Another objectives were to test the hypothesis that including textural features would increase the classification accuracy; and to use minimum distance classification as a base classifier to compare the results of the three methods being tested.

III. STUDY AREA

For the results to be representative and the evaluation of the different methods significative, the study area had to be large and complex enough. River Argos basin in southeast Spain has a surface of 51.786 ha. Three, main land uses appear in this basin: Natural coverages (forests and scrubs) representeing the 51.4 % of the total surface; about 2,2% of urban areas and the rest is almost equally divided between irrigated and rainfed.

Taking into account this land use variability, river Argos basin is adequate to compare the results of the four classification methods analysed in this study. However, due to the existence of rainfed and irrigated crops, the different methods can be confounded. Moreover, due to the landsat satellite spatial resolution, irrigated crops (herbaceous and trees) may also be confounded.

IV. MATERIAL AND METHODS

IV.1. Used data

The Landsat TM image used in this study was taken on the 8th of August, 2000. The image was georreferenced using control points, the atmospheric and illumination correction were carried out using the Chávez (1988) and Teillet (1982) approaches respectively. A more detailed explanation of the process can be found in Alonso Sarría et al (2010).

As the image had been taken more than 10 years before the study was carried out, training and validation areas could not be obtained by field work. Instead, a set several land use maps an orthophotographies was used to extract them (Alonso Sarría et al, 2010). A total of 111 training areas and 56 validation areas were used in this study.

In order to obtain textural features, the semivariogram function for one pixel step was applied to 2 layers: the first component of a principal components analysis carried out on the reflectivity layers, and the normalised difference vegetation index.

IV.2. Classification methods

Minimum distance is one of the oldest classification methods, and today it is not commonly used in research. However, it has been included as a base classifier to compare the

three main methods with. It consists in assigning each pixel to the nearest class in the space of variables. The position of a class in such space is defined by its average vector.

Maximum likelihood (Chuvieco, 2006; Tso and Mather, 2009) is a method based on the assumption that reflectivity values are multivariate normally distributed. The average vector and the covariance/correlation matrix for each class are then used to estimate the probability of a pixel belonging to each of the classes. Finally, the pixel is assigned to the class that maximises the probability.

Sequential Maximum a posteriori (SMAP) is a contextual classification method that classify groups or regions of pixels instead of individual pixels. The basic assumption is that pixels close in the image are more likely to belong to the same class, in this sense it can be considered a segmentation method. It works classifying the image in several resolutions and using the coarser classification to obtain an a priori density function to the finer. Then the spectral information is used to obtain the posterior density function using a bayesian framework (Bouman and Shapiro, 1994; Cheng and Bouman, 2001). The final result is a map with larger and more cohesive polygons than would have been obtained using non-contextual methods.

Random Forest (Breiman, 2001) is an ensemble classifier that uses a large set of decision trees labelling each pixel to the class that was mostly decide by the individual trees. This method has proven to produce very accurate classification comparing with other methods based or not on decision trees (Breiman, 2001; Liam and Wiener, 2002) even where there are more features than cases or when most of the features are very noisy. It is especially interesting that Random Forest do not overfit the model (Ghimire rt al. 2010) providing a great generalisation capacity (Breiman, 2001; Pal, 2005; Prasad et al., 2006). Random Forest also provide a feature importance ranking to identify which are the most relevant to classify the image.

IV.3. Validation

To validate the different classification, confusion matrices were built and the kappa index was calculated with their 95 % confidence intervals. Omission and commission errors and their confidence intervals for each class were calculated as well. Besides these usual accuracy measurements, the Corine Land Cover land use map was also used to measure discrepancies in occupied surface.

IV.4. Computer resources

All the study was developed on a Linux computer using GRASS (Neteler and Mitasova, 2008) to store and manage satellite imagery. GRASS is open source software under GPL licence. It was used to carry out the previous geometric, atmospheric and illumination corrections. Maximum likelihood and SMAP classification used also the modules of the program. Minum distance was carried out in GRASS using a module specially written for this study by the authors.

Random Forest classification was accomplished using the package randomForest (Liaw and Wiener, 2002) in R software (Ihaka and Gentleman (1996) using the pacakge spgrass6 (Bivand, 2007) to read GRASS raster layers into R.

V. RESULTS AND DISCUSSION

The two classification methods that, globally, obtain better results are Random Forest (accuracy=79.1% ± 0.48) and SMAP (accuracy=80.2% ± 0.49). The accuracy of both methods increase when adding textural variables to the dataset, especially SMAP (83.4 % ± 0.43). Maximum likelihood accuracy is slightly lower (72.9 % ± 0.52).

The three methods obtain an accuracy clearly higher than minimum distance's (42.2 % ± 0.59) with a significant separation between confidence intervals.

These global results, however, mask some important differences when the results are analysed for each different class. Natural coverages are quite properly classified. The best results in forest appear when using SMAP without textural features (user's kappa=0.9653 ± 0.0047 and producer's kappa=0.9748 ± 0.004). Using textural features increases producer's kappa but reduces user's kappa. The best results for scrubs are obtained with the SMAP method using textural features (user's kappa 0.9076 ± 0.0073 and producer's kappa=0.8754 ± 0.0081), user's kappa increase slightly when including textural features.

Crop classes present more problems. Errors for tree crops are around 30 %, slightly better using SMAP with textural features. User's kappa equals 0.784±0.0191 and producer's kappa equals 0.6132 ± 0.0199 for rainfed tree crops. In irrigated tree crops user's kappa equals 0.6649 ± 0.0285 and producer's kappa equals 0.8471 ± 0.0246. Rainfed herbaceous crops follow a similar pattern. SMAP with textural features reaches the highest accuracy (user's kappa equals 0.8256 ± 0.0143 and producer's kappa equals 0.6212 ± 0.0156). However, irrigated herbaceous trees show very high errors, near 100 % of error. It is a case of systematic missclassification.

Urban areas have large errors, being Random Forest using textural features the methods that classifies them with best results (user's kappa = 0.4895 ± 0.0432 and producer's kappa = 0.3598 ± 0.0354), although the kappa values for the other methods are not significantly different. Water bodies are very well classified; Random Forest with textural features obtain the best results (user's kappa = 0.9087 ± 0.0492 and producer's kappa = 0.9018 ± 0.05), although the differences with other methods, are once more, not significantly different. In vineyars, the best method, Random Forest with textural features, reach a user's kappa equal to 0.6420 ± 0.0642 and a producer's kappa of 03741 ± 0.0493.

The final land use maps obtained with SMAP include larger and more cohesive homogeneous polygons than the Random Forest or Maximum Likelihood maps, especially in irrigated herbaceous crops. Although larger and more cohesive polygons can be a desirable property from a cartographic point of view, the results of Random Forest appear more realistic. It is noteworthy that including textural features increases the size of such polygons with Random Forest and reduces it with SMAP. A certain convergence of both methods and an increase in their accuracy occurs.

When comparing the classification results with the Corine-Land Cover map, the forest areas show little discrepancy, scrubs are overestimated while bare soil is underestimated. We think that the reason is a criterion difference between Corine Land Cover and our training and validation areas when separating bare soil and disperse scrub. Irrigated tree crops are infraestimated by all methods while rainfed tree crops are overestimated by all methods. Water areas appear overestimated with all methods; the reason is two-fold:

firstly, Corine Land-Cover do not represent most of the irrigation pools; second, the water layer of the reservoir in the study area was larger in the image than in the map.

Classes in which small patches are expected (urban areas, water layers) are more overestimated with SMAP than with Random Forest. It seems that the cohesion effect of contextual classification generated polygons too large for these classes.

VI. CONCLUSIONS

The main results obtained allow us to extract these main conclusions:

1. The three analysed classification methods improve substantially the accuracy of the base classifier (minimum distance).
2. The accuracy of both Random Forest an SMAP are similar, being the latter the one with higher kappa index. This kapa index has a significant increase when using textural features to classify.
3. When analysing missclassifications class by class, the results are very similar for all the analysed methods. Crops are more frequently missclassified than natural covers, being irrigated herbaceous crop the most difficult class to correctly classify
4. The percentage areas obtained for the whole River Argos basin reproduce quite accurately the obtained using Corine Land Cover.