

A study of drivers for agricultural land abandonment using GIS and Data Mining techniques

B. M. Zaragoza¹, J. T. Navarro¹, A. Ramón¹ & J. J. Rodríguez-Sala²

¹*Instituto Interuniversitario de Geografía, Universidad de Alicante, Spain*

²*Centro de Investigación Operativa, Universidad Miguel Hernández de Elche, Spain*

Abstract

This paper examines the problem of agricultural land abandonment through knowledge of driving factors. First, we reach out a review of the concept of abandonment in order to define the negative and positive aspects of this process in the landscape. Then we focus on a real case study in the Marina Baja region (SE Spain). Finally, we propose, from a theoretical perspective, the use of a methodology to define the most important variables for studying the abandonment processes in this region.

First of all, we performed various analyses using Geographical Information Systems (GIS) to create a geodatabase which stores the main factors generally associated with abandonment. The variables considered can be grouped into three groups: environmental (such as climate, soils or relief), socio-economic (such as demographic, accessibility or technology) and those related to cultural practices in the crop (such as irrigation or rain-fed stone walls).

Considering over a hundred variables we need to answer two questions: (1) which of the variables are really driving factors in the abandonment process of an agricultural plot? and (2) how important is each of them alone or in combination with others? The authors advocate the use of GIS for the feature creation or management and Data Mining techniques to perform the feature selection. Specifically, we propose the use of a concrete metric, aci (Attribute Correlation Index), to answer the questions raised above.

Keywords: *Geographical Information Systems, Data Mining, land abandonment.*



1 Introduction

The abandonment of agricultural lands is an important phenomenon in many regions in the European Union. We can appreciate these processes using some indicators like the continuous reduction of the population employed in agriculture or the decline in the number of exploitations. These processes are likely to continue over the coming years as a result of the lack of viability of many of these exploitations [1].

The concept of “abandoned lands” is not completely clear, there is no single definition of land abandonment in the EU countries and this process even does not exist in a few of these countries [2]. However, we can find different interpretations in the literature describing it from a static or a dynamic point of view [3]. Following a static point of view, we can find some visible effects of the abandonment like the spontaneous re-vegetation or the progressive deterioration of fences, roads, irrigation systems and other agricultural infrastructure. When all these changes make it very difficult or impossible for the continuation of the agricultural activities, the situation is described as total abandonment or effective abandonment [4].

However, many of the farms that in theory are not viable can continue to be active for a long time after the expected period. The reasons for this are many, including transfer of ownership, search for other incomes outside the agriculture (part-time agriculture), or continuation of activities beyond retirement age [1], among others. These processes can help continue the agricultural activity but they also indicate the state of the farm prior to an expected total end of activities. The processes that lead to the total abandonment can be prolonged for an indefinite period of time. In that case, the parcels are not strictly abandoned, and that is why the situation can be described as “semi-abandonment”. The semi-abandonment state is not easily perceived because the remaining activities can mask this situation. On the other hand, total abandonment is easier to locate geographically, but it includes in its definition many areas that could have been unexploited for a long time. For these reasons, in this paper we prefer an intermediate approach – we consider that an abandoned land is one that has an absence of any current land use but had a recent history of agricultural land use, normally, less in last 10 years.

A concept that helps to explain the semi-abandonment processes is “marginality”. It has been traditionally associated with those lands that have low potential productivity. In case of any adjustment marginal lands are usually the first to be abandoned, so it is an important concept in order to perform abandonment estimations in a short or mid-term period. Like some biophysical factors, other structural and socioeconomic factors can be related to the marginality [5]. In Table 1 we can find some of the most common factors related to marginality and abandonment.

Abandonment is related to many processes implied in land use changes. Depending on the type of change and the region, abandonment can be considered a positive or negative change. Some positive effects linked to abandonment are



Table 1: Some identified drivers of abandonment, based on [5].

Geo-bio-physical	Socio-economic	Management
<ul style="list-style-type: none"> • elevation • geology • slope • fertility • soil depth • soil erosion • climate • ... 	<ul style="list-style-type: none"> • market incentives • migration • technology • industrialization • farm characteristics • farmer age • accessibility • ... 	<ul style="list-style-type: none"> • unadapted agricultural systems • land mismanagement • soil degradation • frequent flooding • overexploitation • productivity loss • ...

connected to the re-vegetation of these areas with natural species adapted to the physical conditions. Due to this re-vegetation there are other positive effects caused by the abandonment process, like a better hydrological regulation and water quality, soil recovery and erosion mitigation, increased fertility, biomass increase and higher carbon sequestration. However, in Mediterranean areas this re-vegetation process can be quite slow and the pointed positive effects will arrive too late [6]. The negative effects associated to abandonment are also related to the re-vegetation process. These negative effects can be the loss of landscape heterogeneity and biodiversity, increased fire risk [7], soil erosion and desertification [8], reduction of water stocks, biodiversity loss and reduced population of adapted species or the loss of cultural and aesthetic values [9].

Once reviewed some important concepts, in this paper we consider that a better understanding of the driver factors that lead to the abandonment is a key subject, and we propose a new methodology that uses generally available geospatial data to find out the variables or drivers that lead to the abandonment of agricultural parcels (cadastral parcels).

2 A case study

2.1 Study area

In order to study the abandonment processes we have selected a Mediterranean region with a diverse geography where nowadays two socio-economic processes can be easily observed: firstly, agriculture and urbanisation intensification, and secondly, land abandonment. Now, we need a short geographical description of the selected study area in order to find out the most important factors that are structuring that region [10].

The study area is located in the southeast of the Iberian Peninsula, in the province of Alicante. Its extension covers about 578.8 squared kilometres and it is divided into a mountainous area and another coastal area. This region comprises 18 municipalities. In such a small area, the Marina Baja has many interesting contrasts. In Figure 1 you can appreciate that the altitudes range from sea level to 1558 m that is reached at the peaks of Sierra de Aitana. Between these mountains the most important river system is Guadalest-Algar (with



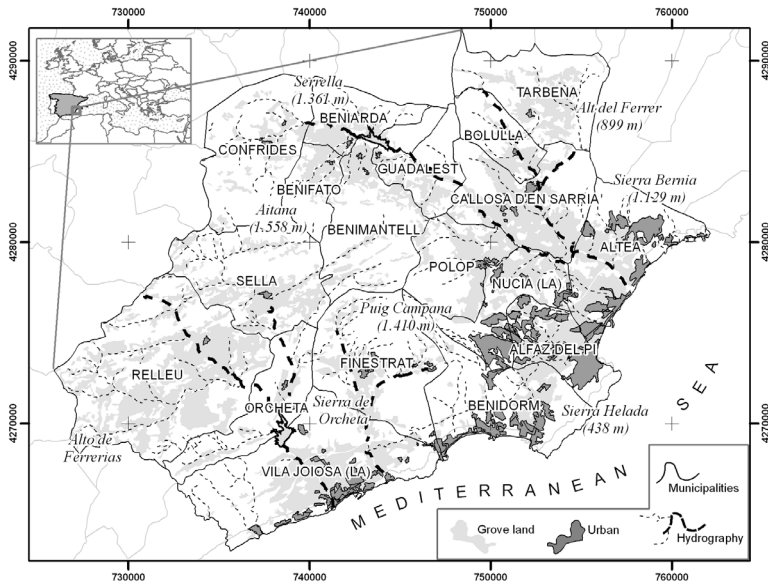


Figure 1: The Marina Baja.

modules from 0.90 to 0.78 m³/s, respectively), but it should be noted that the resources provided by this river allow most of the irrigated crops and supply to towns. Very related to this physical distribution we can appreciate a very diverse climate with sensible differences between sea and mountain and north to south. Just to give a few climatic data we can say that average temperatures range between 17 and 18°C and annual rainfall between 400 and 700 mm.

Before the 1960s, the Marina Baja was isolated because its abrupt topography and there were no highways or main roads as known as today but due to many socio-economic changes the Marina Baja structure has experimented a radical change. Today, over 60% of the Valencian tourist activity is concentrated here, mostly located in Benidorm, main municipality in the region [10]. These trends are still going on and producing constant land-use changes associated with two main processes: tourism/urbanisation and agricultural intensification/abandonment.

2.2 Data sources and GIS analysis

In this section we explain three steps of the GIS analysis performed for each data source: analysis of source, derivation of features and extraction to the parcel level (Object Oriented Analysis). Of course, all the data sources referred here are available for other Spanish regions and due to the INSPIRE initiative they should be available for all EU countries (Directive 2007/2/EC). This way the methodology posed here can be extrapolated to many other sites.

We acquired many different datasets that model as much as possible the abandonment drivers identified during the literature review and derived from the knowledge of the study area. These datasets include remote sensing images and altimetry datasets from the Valencian Institute of Cartography (ICV), the parcel boundaries defined by the Spanish Land Parcel Identification System (SIGPAC), and a geospatial database used for agriculture management. At last, the climatic data were obtained at the meteorological station level from the National Institute of Meteorology (INM). Of course, all these datasets needed to be pre-processed to get some useful and continuous variables. This way many of the attributes defined were derived from the previous datasets by application of different methods.

The cartographic boundaries to define the agricultural plots were obtained from the SIGPAC. In this database, each polygon represents a continuous area within a parcel, with normally a single agricultural use [11]. The total number of plots of the study area was 108,000. Directly from the parcels we calculated some shape-metrics including some very basic metrics like area or perimeter, but these metrics can be very useful because for example the shape of the area highly related to the agriculture viability [6].

The orthophotos employed were facilitated by the ICV. The images have a spatial resolution of 0.25 m/pixel and three spectral bands, including the near infrared, necessary to calculate the vegetation index. These images were acquired in 2007. To perform the feature extraction of the images we used the software FETEX 2.0 (Feature Extraction). This software was created by the CGAT (Geoenvironmental Mapping and Remote Sensing) research group, from the Department of Cartographic Engineering, Geodesy and Photogrammetry of the Universidad Politécnica de Valencia. It is programmed in IDL 6.2 as an ENVI extension [12]. This software is designed to work with land cover databases assisting in the classification process of existing plots or even to determine land cover changes by comparing the current and the older database. Using this software it is possible to obtain over 50 characteristics describing an agricultural plot (spectral, textural and structural properties of the parcel image).

Using the same orthophotos we performed a first classification of the abandoned parcels following some criteria:

- No evidence of certain local agricultural practices
- Stonewalls are broken
- There are crop species mixed with natural vegetation
- Agricultural terraces are totally recolonized by natural species
- Presence of abandonment effects, such as erosion or wildfire
- The size of species is highly diverse
- There is bare soil but the databases indicate any agricultural use

Using the SAGA GIS software (System for Automated Geoscientific Analyses) and the ICV elevation data, we created a DEM by the Triangulation method. After that we converted the DEM to a raster format with a fine grid cell size (2x2 meters) to facilitate the feature extraction. SAGA GIS is useful for extraction of many features related to the DEM, but there are other proximity and interpolation methods that we used for preparing the other layers.



Table 2: Summary of possible features to include in this methodology.

Group / Number	Variables	Source
Shape/Plot features (6)	area, perimeter, shape index, fractal index...	SIGPAC geometry
Image features (56)	spectral features, texture features, structural features...	ICV CIR Images
DEM features (34)	Mean elevation, slope, aspect... Topographic Humidity Index...	Isohypes 1/10,000
Database features (26)	landuse03, landuse09, municipality, irrigation associations...	SIGPAC, Cadastre, Agricole cooperatives...
Climate features (60)	Precipitation, temperatures, water balance...	INM, IVIA... (Interpolation)
Soil features (12)	pH+, Organic matter...	Lab análisis (Interpolation)
Proximity features (10)	Accessibility, proximity to urban areas, to water sources...	ICV10
Aesthetic landscape (6)	Visibility, fragility...	DEM, Landuse...

Using the above-mentioned software and the methods available in there it is possible to extract hundreds of features related to the drivers of abandonment at the parcel level (see Table 2). In the methodology proposed here we must analyze this huge amount of information to get an explanation about the abandonment process in the study area, but this is high time and effort demanding, so we have selected a varied list of 10 features to test a Data Mining (DM) method that will help us to select the most important features or variables.

The first feature is the area (AREA) that can be easily calculated from the SIGPAC polygons. All the GIS programs should be able to calculate such a simple metric. We also calculated a shape index (SH_INDEX).

The easiest DEM feature we can extract is the mean elevation of each parcel (MEAN_ELEV). The crops in the study area are highly influenced by the elevation because this variable modifies the climatic characteristics decreasing the temperature range that is a limiting factor for many subtropical crops. Also based on the DEM, we extracted the Topographic Wetness Index (TWI). The TWI has been used to study different hydrological processes and biological processes based on topography [13].

Using the high resolution images available, we calculated the Normalized Difference Vegetation Index (NDVI). The NDVI is useful to distinguish artificial from natural plots, or dry crops from irrigated crops, for example. At the parcel level we could calculate many statistics relating the index to each parcel but we have selected some of the most simple: mean and standard deviation of the NDVI (MEANNDVI and DEVSTNDVI). We also have selected one of the many possible structural features we can extract from the orthoimages. The First Derivative near the Origin (FDO) is a feature based on the empirical semivariogram of the pixels that cover one parcel and it is also calculated using FETEX.

There are many climatic features we could extract or derive from the climatic data. Using the climatic data from INM (temperature and precipitation data from



1950 to 2007) we performed an interpolation as explained by [14]. Finally, we obtained monthly GIS layers for mean precipitations and temperatures (mean, max and min). However, in this study we had to limit the number of variables so we decided to calculate a climatic index used in vegetation studies. The Patterson index or Forest Productivity Index (FPI) has been widely used and adapted to different regions. Nowadays, it is still integrated in more complex methodologies to consider the climatic biomass productivity (e.g.: in [15]). This index is based in the assumption that high temperatures and strong precipitations are necessary for the biomass production and it gives an estimation of productivity in m³ of wood/ha·year. The FPI formula is:

$$FPI = \frac{tm_{12} \cdot G \cdot P \cdot f}{12 \cdot [T_{12} - t_1]} \quad (1)$$

where P is the annual precipitation in mm, tm_{12} is the mean temperature in the warmest month, t_1 is the mean temperature in the coldest month, T_{12} is the mean of the maximum temperatures of the warmest month, G is the duration of the vegetative period (in months) and f is a sunstroke factor.

We used the SIGPAC land uses for two different time periods 2003 and 2009 (LU03 and LU09). Since they were already at the parcel level we did not need any additional process to add these two variables to the DM analysis. We also used information indicating if a parcel is irrigated or not.

Finally, we wanted to use any socio-economic variable and we selected the population density in 2009 (DENS_POP). This kind of sources is normally aggregated at the municipality level but it could be still explanatory. In these cases we apply for each parcel the value of its overlapping municipality.

The abandonment process has been studied using photo-interpretation as described before, and we have conducted some field work to validate a 10% of the total parcels in order to be sure that the abandonment is real.

3 Feature selection using DM techniques

Usually, statistical techniques provide useful tools to manage huge databases with millions of records. This is a length problem. But sometimes, as the case study proposed in this paper, the dimension problem is not in length but in width. Thus, considering 10 variables each of them acquiring only 5 possible values could imply a maximum of 100,000 rules (10^5 combinations). Although there are several decision tree algorithms with contrasted accuracy [16] that implement heuristics and pruning techniques, as ID3 for nominal attributes or C4.5 for nominal or numerical ones, sometimes a Feature Selection procedure [17] becomes necessary.

In this paper, the *aci* measure, provided by the RBS algorithm [18], is used to carry out the Feature Selection procedure. The RBS (Reduction Based on Significance) is an algorithm that calculates the critical significance and confidence boundaries to classify every rule in their specific region of significance: $A1$ contains significant and discriminating rules; $A2$ contains

supportable and reliable rules and $A3$ contains rules with insufficient support to be classified as significant or not. All the rules outside those regions are considered as non-significant ones.

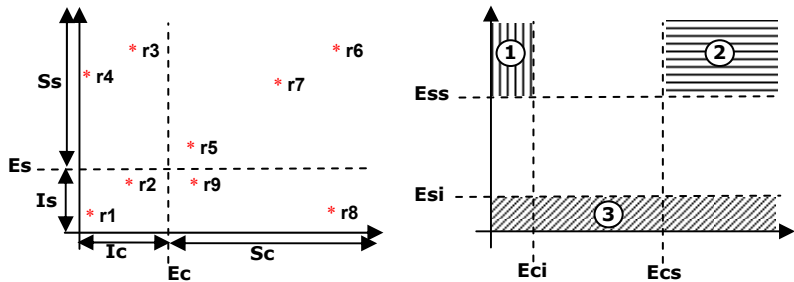


Figure 2: Example of rule distribution and significance region division obtained with RBS.

Besides, this algorithm calculates an empirical measure of variables correlation: *aci* (Attribute Correlation Index), the incidence of the attributes over the target variable.

The *aci* value associated to this Rule Set is defined as follows:

$$RS_{A \rightarrow C}.aci = \frac{RS_{A \rightarrow C}.srp}{A1 + A2 + A3} \quad (2)$$

where $RS_{A \rightarrow C}.aci$ is a set of rules typed as: antecedent (A) implying consequent (C). While antecedent is a nominal attribute or group of them, the consequent is a unique attribute, frequently called class attribute or target variable. For example $RS_{A \rightarrow C}$ can be instantiated to $A(\text{area, crop_type}) \rightarrow C(\text{abandonment})$, and a rule in this rule set could be (area=very low, crop_type=arable lands) \rightarrow (abandonment=yes) with confidence=93.5%. $RS_{A \rightarrow C}.srp$ is the proportion of significant rules in this set, and Ai the area of every significant region ($i=1, 2, 3$). High values of *aci* mean A is highly correlated with C , and vice versa.

Using this algorithm we can select the variables that are related to abandonment. However, this algorithm requires nominal variables so before applying the algorithm we performed a discretization process.

The landuse fields LU03 and LU09 were already discretized having 20 and 28 values respectively. Nevertheless, some parcels did not have any value so we redefined them as “NO_DEF”.

Population Density (PD) has been classified in three groups (Low-Medium-High) to separate the crowded municipalities on the coast from the interior agricultural towns and finally the mountainous and depopulated towns.

In the most of the parcels the irrigation values (irrigated percent area of the parcel) are 0 or 100, but there are a few plots where the values remain between these two values. In those cases using the 50% limit, we have decided to use only two ranges (YES/NO) to classify the irrigation field (IRRIG).



Table 3: Discretization applied to get the nominal values.

LU03 (20 values) and LU09 (28 values)	MEANNDVI (5 values): (-inf, -0.3] => VL (-0.3, -0.1] => L (-0.1, 0.1) => M [0.1, 0.3) => H [0.3, +inf) => VH	DEVSTNDVI (3 values): (0.00, 0.15] => L (0.15, 0.35) => M [0.35, 0.50) => H
FDO (4 values): (-200, 0] => L (0, 200) => M [200, 400) => H [400, +inf) => VH	SH_INDEX (4 values) (0, 1] => L (1, 2) => M [2, 5) => H [5, +inf) => VH	AREA (4 values): (0, 2000] => VL (2000, 10000] => L (10000, 50000) => M [50000, 100000) => H [100000, +Inf) => VH
MEAN_ELEV (4 values): (0, 200] => L (200, 500] => M (500, 800) => H [800, +Inf) => VH	MEAN_FPI (5 values): (0, 3] => VL (3, 4] => L (4, 5) => M [5, 6) => H [6, +7) => VH	MEAN_TWI (5 values): (0, 5] => VL (5, 10] => L (10, 15) => M [15, 20) => H [20, +) => VH
IRRIG =>50 => Yes <50 => No	DP <100 => L [100, 400] => M > 400 = H	

The other numerical values have been discretized in 3, 4 or 5 values based on the following criteria (with these criteria, the discretization is as explained in Table 3):

1. If the distribution was uniform in all its values range the variable has been divided into 5 values VL, L, M, H, VH (Very Low, Low, Medium, High, Very High) symmetrical, except the attribute “DEVSTNDVI” which has been divided into three values (L, M, H) because it has a very narrow range (between 0 and 0.5)
2. When the average of the values of a particular attribute lies far to one end, particularly to the minimum value, we have decided to discretize it in 4 ranges (L, M, H, VH), so that the labels L, M and H are for values around the mean and VH label is reserved for the more distant values above the average.

After this reclassification process we ran our own program and after 21 minutes of process time we obtained 8191 rules/group of rules ordered by the *aci* values indicating the variables or combinations that explain better the abandonment process.

Analysing all the rules with a higher *aci* value, we find that a key field in all the examples is LU03, and the second most recurring is M_FPI so we can say that both are particularly important in determining the abandonment of an agricultural parcel. The LU09 field seems that is somewhat less decisive than

LU03 as it appears in other sets of rules with lower *aci*, but in both variables we can see that the abandonment is highly associated with the “arable lands” category (those lands with a recent agricultural past but with no vegetation now). It is obvious that this definition of arable land matches in many cases with our abandonment definition, but this can not help to estimate marginality; this is why LU needs always to be combined with other variables.

The abandonment is related to a low or very low area value (less than one Hectare). The mean elevation seems also to be important. Normally, a low value in elevation is highly associated with the abandoned parcels, but this is not a direct relationship, in this study area the zones with low elevation (0-200 m) are those where we can find a higher urbanisation process. This means that these preliminary results are going to change once we include more variables in the analysis.

4 Conclusions

In this paper we have shown that nowadays we are able to generate and use a huge geodatabase for a relatively small study area. The analysis and interpretation of such a huge database could take a long time and in many cases we only would be able to get the obvious conclusions. For these reasons we have tested and confirmed that the *aci* metric/algorithm can help us to explore a very particular problem as the agricultural land abandonment.

We have conducted a test with a small number of variables and many of them were already well known and related to abandonment by the bibliography. However, we have confirmed those relationships for our particular study area. Variables obtained from the existing landuse databases (SIGPAC), the parcel area, the NDVI or the mean elevations are related to abandonment in this area. On the other hand, the climatic index (Patterson; MEAN_FPI) has not been used in previous studies, but we found that it is (in second position) one of the most important explanatory variables.

Performing the same analysis for over a hundred of new variables we will be able to explore new relationships that can help us in the study of a complex problem such as abandonment.

5 Future research

The GIS feature extraction process can be performed using different software but due to the amount of GIS layers we need to process, it is necessary to automate the whole process once the total variables are determined. In our next research we will automate the feature extraction from GIS layers.

The proposed software implementing the RBS algorithm is nowadays a prototype. Therefore in this paper a simplified pilot study is approached, considering a dataset with only 10 attributes.

For future research, a complete computational experiment (with over a hundred variables) will be developed and analyzed in detail. As the proposed algorithm needs nominal variables as input data, the discretization procedure



accuracy is a critical constraint. Therefore, dynamic clustering techniques will be developed in order to perform an accurate and efficient discretization task.

References

- [1] Sineiro, F., López, E., Lorenzana, R. & Valdés, B., La tipología de las explotaciones en función de su viabilidad económica y demográfica; aplicación a las explotaciones de bovino de Galicia. *Economía Agraria y Re- cursos Naturales*, **4(8)**, pp. 63-85, 2004.
- [2] Moravec, J & Zemeckis, R., Cross Compliance and Land Abandonment, Deliverable D17 of the CC Network Project, SSPE-CT-2005-022727, 2007.
- [3] Corbelle, E. & Crecente, R., El abandono de tierras: concepto teórico y consecuencias. *Revista Galega de Economía*. **17(2)**, 2008.
- [4] Food and Agriculture Organization (FAO). The role of agriculture and rural development in revitalizing abandoned/depopulated areas. 34th Session of the European Commission on Agriculture. Riga. 2006.
- [5] Rey, J., Martins, A., Nicolau, J.M. & Schulz, J.J., Abandonment of agricultural land: an overview of drivers and consequences. *CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources*, **2 (57)**, pp. 14, 2010.
- [6] Bielsa, I., Pons, X. & Bunce, B., Agricultural Abandonment in the North Eastern Iberian Peninsula: The Use of Basic Landscape Metrics to Support Planning. *Journal of Environmental Planning and Management*. **48(1)**, pp. 85-102, 2005.
- [7] Lloret, F., Calvo, E., Pons, X. & Díaz-Delgado, R. Wildfires and landscape patterns in the Eastern Iberian Peninsula. *Landscape Ecology*, **17**, pp. 745 – 759, 2002.
- [8] Koulouri, M. & Giourga, C., Land Abandonment And Slope Gradient As Key Factors Of Soil Erosion In Mediterranean Terraced Lands. *Catena*, **69(3)**, pp. 274-281, 2007.
- [9] Otero, I., Boada, M., Badia, A., Pla, E., Vayreda, J., Sabaté, S., Gracia, C.A. & Peñuelas, J. Loss of water availability and stream biodiversity under land abandonment and climate change in a Mediterranean catchment (Olzinelles, NE Spain). *Land Use Policy* **In Press**, 2010.
- [10] Peña, J., Bonet, A., Bellot, J., Sánchez, J.R., Eisenhuth, D., Hallet, S. & Aledo, A., Driving forces of land-use change in a cultural landscape of Spain. A preliminary assessment of the human-mediated influences (Chapter 6). *Modelling Land-Use Change: Progress and Applications*, eds E. Koomen, J. Stillwell, A. Bakema, H.J. Scholten. Springer: Netherlands, pp. 398, 2007.
- [11] Mirón, J., Cadastre and the reform of European Union's common agricultural policy. Implementation of the SIGPAC (1). *Catastro*, **54**, pp. 161-172, 2005.
- [12] Ruíz, L.A., Recio, J.A., Fernández-Sarría, A. & Hermosilla, T., A Tool For Object Descriptive Feature Extraction: Application To Image Classification And Map Updating. *The International Archives of the Photogrammetry*,



Remote Sensing and Spatial Information Sciences, Vol. XXXVIII-4/C7, 2010.

- [13] Sørensen, R., Zinko, U., & Seibert, J., On the calculation of the topographic wetness index: evaluation of different methods based on field observations. *Hydrol. Earth Syst. Sci. Discuss.*, **2**, pp. 1807-1834, 2005.
- [14] Ninyerola, M., Pons, X. & Roure, J.M. A methodological approach of climatological modelling of air temperature and precipitation through GIS techniques. *International Journal of Climatology*, 20, pp. 1823-1841. 2000.
- [15] Benavides, R., Roig, S. & Osoro, K., Potential productivity of forested areas based on a biophysical model. A case study of a mountainous region in northern Spain. *Ann. For. Sci.* **66 (1)** 108, 2009.
- [16] Zhang, H. & Su, J. Learning Probabilistic Decision Trees For AUC. *Pattern Recognition Letters*, 27, pp. 892–899, 2006.
- [17] Chen, Y-L., Hsu, W-H. & Lee, Y-H., TASC: Two-Attribute-Set Clustering Through Decision Tree Construction. *European Journal of Operational Research*, 174, pp. 930–944, 2006.
- [18] Abadía, R., Almiñana, M., Escudero, L.F., Pérez, A., Rabasa, A. & Santamaría, L. A, Rules Reduction Algorithm Based On Significance Measure. Proc. IX Data Mining and Information Engineering, eds. Zanasi, Alzamora, Ebecken & Brebbia, WIT Press. Cádiz, pp. 63-72, 2008.

