Combining classification techniques to define topo-climatic landscapes.

Memòria de recerca.
Doctorat en Ciències Ambientals.
Juny 2008.

Josep M Serra Díaz
Director: Miquel Ninyerola Casals

INDEX

PRESENTACIÓ

La present memòria de recerca exposa en format d'article científic un exemple de la recerca realitzada darrerament en diferents àmbits d'estudi sota un mateixa temàtica: les classificacions ambientals/ecològiques i la dinàmica de la vegetació. Així, la metodologia emprada ha estat utilitzada en àmbits tan diferents com el parc natural de l'Alt Pirineu, la vegueria de l'Alt Pirineu i Aran, Catalunya i la Península Ibèrica.

Cadascun d'aquests estudis ha representat variacions a nivell d'escala, variables seleccionades, mètodes d'agregació i matisos diferents en la classificació; essent la presentada en la següent memòria una de les més complertes.

En les classificacions realitzades al llarg d'aquest període s'ha tingut en compte variables físiques ambientals, tot i que també s'han realitzat proves afegint variables de cobertes del sòl o índexs d'estructura paisatgística. Els resultats d'algunes d'aquestes classificacions han servit de base per a la descripció i catalogació de paisatges, donant lloc categories (clústers) més aviat descriptives. En la memòria que es presenta a continuació s'ha donat més pes a relacionar les categories identificades amb el relleu i la vegetació en comptes de realitzar una anàlisi descriptiva exhaustiva de cada categoria.

Al llarg d'aquest període de recerca s'ha pogut analitzar i comprendre que l'objectiu i l'escala de treball influiran molt en la metodologia d'identificació de categories així com les variables d'anàlisi seleccionades.

Com a base de la recerca es parteix de l'aproximació global i sistèmica del paisatge dividida en tres subsistemes relacionats: Geosistema - Territori - Paisatge (GTP) descrita per G.Bertrand[1]. En la recerca que s'ha dut a terme, s'ha centrat en l'estudi del Geosistema o medi físic.

---

[1] BERTRAND, George (2001). "Le paysage et la géographie: un nouveau rendez-vous". *Treballs de la Societat Catalana de Geografia*, 50: 57-68

Combining classification techniques to map topo-climatic landscapes. A case study of Lleida-Pyrenees region (NE Iberian Peninsula).

ABSTRACT

Landscape classification tackles issues related to the representation and analysis of continuous and variable ecological data. In this study, a methodology is created in order to define topo-climatic landscapes (TCL) in the north-west of Catalonia (north-east of the Iberian Peninsula). TCLs relate the ecological behaviour of a landscape in terms of topography, physiognomy and climate, which compound the main drivers of an ecosystem. Selected variables are derived from different sources such as remote sensing and climatic atlas.

The proposed methodology combines unsupervised iterative cluster classification with a supervised fuzzy classification. As a result, 28 TCLs have been found for the study area which may be differentiated in terms of vegetation physiognomy and vegetation altitudinal range type. Furthermore a hierarchy among TCLs is set, enabling the merging of clusters and allowing for changes of scale. Through the topo-climatic landscape map, managers may identify patches with similar environmental conditions and asses at the same time the uncertainty involved.

Keywords: landscape classification, unsupervised classification, cluster analysis, topo-climatic analysis, remote sensing.

1. INTRODUCTION.

Defining landscape units and developing an ecological landscape classification has been the aim of many researchers for several years (Bailey 2004). Different approaches have been used to classify landscapes or, to a broader scale, define Eco-regions (Omernik 1987; Loveland et al 2004). Most of these classifications have been based on ecological variables related to landscape structure using landscape metrics (McGarigal 2002), hydro-ecological factors (Wolock et al 2004), topography and climatic variables (Host et al 1996, Blaschke and Strobl 2003).

According to Wolock (2004), these classifications should “identify patterns in biotic and abiotic factors thought to generally influence ecological processes at a relatively broad scale”, providing an ecological framework that encompasses specific uses and a broad ecological classification for management needs (Mackey et al 1988; Bailey 1996; Carter et al 1999).

Two general approaches are identified when classifying landscapes (Fairbanks et al 2000): human-landscape based classification (Blankson et al 1991) and biophysical approaches (Martin de Agar et al 1995; Bernert et al 1997). In the present study, a biophysical approach has been chosen in order to determine areas with similar topographic and climatic behaviour. Moreover, through remote sensing, land cover may be characterized and landscapes can be classified (Haines-Young 1992) adding information about a certain vegetation cover based on a set of widespread of remote sensing indexes used in ecology (Cohen and Goward 2004) such as the normalized difference vegetation index (NDVI), wetness index (WI) and land surface temperature (LST) (Quattrochi and Luvall 2000). This information, suitably combined, enables the determination of a regional ecological framework for management and other specific studies (Barnes et al 1982; Cleland et al 1992; Kirkpatrick et al 1994; Nolet et al 1995; Bunce et al 1996; Monjeau et al 1998; Fairbanks et al 2000).

Recently, a set of methods have been presented based on expert knowledge or quantitative methods. Both methods present a certain degree of subjectivity in classification but a quantitative method has been chosen for the present study since it is easier to apply an identical classification procedure for every map unit. However, the variables chosen for classification are also subjective since they will depend on the researcher and the aim of the output cartography.

Among quantitative methods, a wide range of new techniques have been applied to classify landscapes or environmental data such as fuzzy logics (Burrough et al 2000; MacMillan et al 2000; Burrough et al 2001; Nadeau et al 2004), neural networks (Chon et al 2000), information theory (Kraft et al 2004) or multivariate cluster analysis (Hutto et al 1999, Hargrove and Hoffman 2004).

One of the main issues when classifying ecological quantitative variables is dealing with heterogeneity, uncertainty and crispness of data and landscapes (Rocchini et al 2007) and the management need of such a classification; both will influence methodology for delineation of landscapes (Cleland et al 1997). A biophysical landscape classification must define areas with similar ecological behaviour based on continuous environmental factors. Therefore, environmental landscape classification must take into account uncertainty in classification since landscapes should not only be defined by crisp lines. However, landscape managers usually need to delimit landscape units in order to apply a certain policy or management tool.

Therefore, we deemed it necessary to combine different classification techniques (clustering and fuzzy) in order to tackle different uses of a map of landscape units.

Clustering classification techniques based on statistical distances create several classes or clusters. Hence, the membership value of an object to be classified is expressed as a binomial function ranging from 0 to 1 whether they belong to a class or not. In addition, a hierarchal cluster analysis may show similarity between clusters. As a result, this sort of analysis creates a clear classification of the environmental data that could be helpful for several uses or for managers, but too simplified to express complexity and heterogeneity of an environmental landscape classification.

On the other hand, fuzzy classification tools offer the possibility to assign a membership value of an object to the class, creating a fuzzy classification where objects may belong to several classes based on a membership function. Usually, this classification procedure requires training sites and the degree of membership of an object to a class in order to classify objects. However, it is difficult to identify these training sites when objects to be classified are not easily perceptible (i.e. topo-climatic classifications) and consequently, is highly recommendable to perform a first non-supervised clustering classification. Therefore, we deemed it necessary to create a GIS methodology that combine both techniques resulting in a classification based on a sharp cluster analysis overlaid with a fuzzy membership function, clarifying to what extent a pixel belongs to the assigned class.

Finally, testing a topo-climatic classification is often difficult since there are few independent variables that directly explain the whole variability of the amount of data involved (Host et al 1996). Instead, we used vegetation maps to relate topo-climatic classes to vegetation types assessing the correspondence between them and characterising topo-climatic landscapes.

The aim of the present study is to delimit ecological units at a regional scale with different biological and physical potentials (Cleland et al 1997) using different classification approaches. Specifically, this study develops a methodology to identify landscapes through topographic, climatic and physiognomic variables, combining crisp and fuzzy quantitative classification techniques and a biological characterization of the output maps through the Habitat Map of Catalonia (Housing and Environment Department of Generalitat de Catalunya and University of Barcelona, *2005*).

1.1 Study area.

The region under study has a total area of 11340 km^2 and it is located in Catalonia (33000 km^2), north-east of the Iberian Peninsula (Figure 1). This region is characterized by a high variability in habitats due to a wide range of environmental conditions such as seasonal rainfalls, temperature variations and altitude. For instance, alpine pastures over 2000 meters and semi-arid shrub communities could be found in relatively short distances.

The Pyrenees range (see Figure 1) is characterized by pastures, summits and mainly coniferous forests. Rainfalls are not distributed seasonally although there is a slight increase in autumn and spring. Temperature regimes vary largely during day and night and throughout the year.

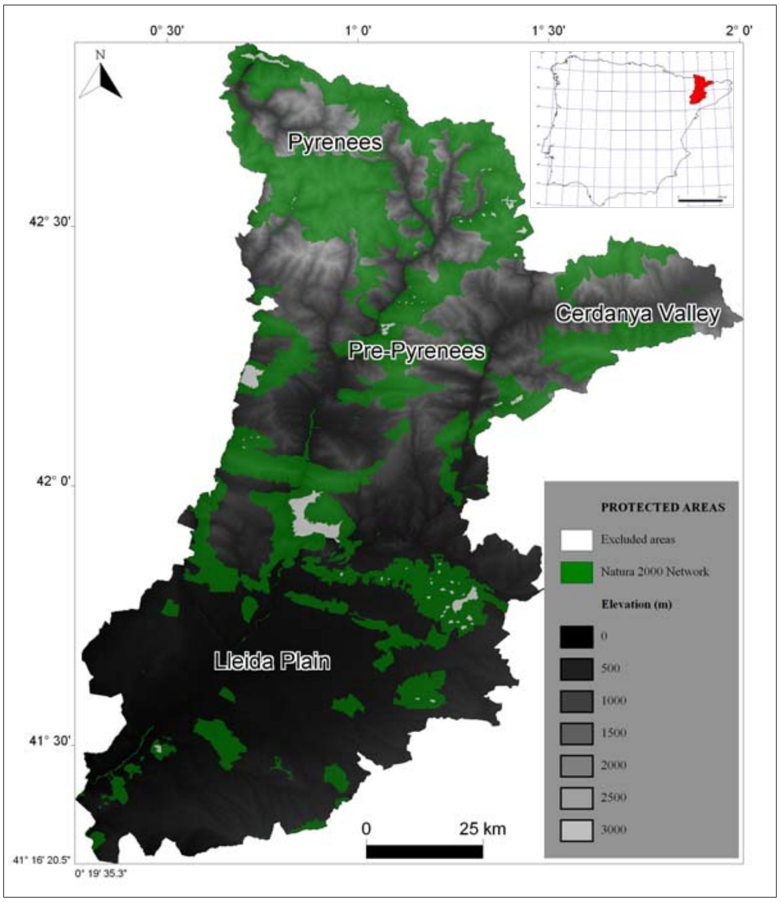

Figure 1. Digital Elevation Model of the study area. Natura 2000 European protected areas are indicated.

South of the Pyrenees, the Pre-Pyrenees range is characterized by intermediate climate conditions between alpine and purely Mediterranean climatic conditions. Precipitations are distributed seasonally throughout the year, peaking during autumn and spring. It could be assumed that regions of Pre-Pyrenees range could present similar habitat compositions as found in other high mountains of the Pyrenees. However, precipitation in this area is much lower than in the Pyrenees. That pluviometric shadow is due to the Pyrenees range, which stops cloud fronts coming from the north. This leads to higher rainfalls in the Northern side of the Pyrenees and a dryer climate in the southern face (Garcia de Pedraza and Reija 1994; Clavero et al 1996).

The southern zone of the study area is characterized by big plains and hills. Dry land crops and irrigated cultivation areas dominate this region, although there are relevant patches of semi-arid shrub communities.

The area under study includes over 3614 km^2 of land within the declared Natura 2000 network of European protected areas, which covers over the 30 % of the total area included in this study, underlining the ecological importance of this region and the relevance of its classification in order to determine conservation goals and strategic planning alternatives.

2. METHODS.

2.1 Methodology overview.

In order to identify and map topo-climatic landscapes (TCL), a set of relevant variables: climatic is selected: topographic and physiognomic variables (remote sensing). Principal Component Analysis (PCA) was performed among these variables to reduce redundancy (reefer to section 2.3) and a clustering classification technique was used to group cells with similar climatic behavior (reefer to section 2.4) to obtain TCLs. (see Figure 2)

Finally, a hierarchical clustering technique was applied to investigate similarities between TCLs and an uncertainty index was calculated for each cell using a fuzzy classifier (sections 2.4, 2.5 and 2.6). These three outputs (clustering classification, hierarchical classification and the fuzzy uncertainty map) compound the Topo-Climatic Landscape Map. Furthermore, a habitat characterization of each TCL was undertaken through the Habitat Map of Catalonia (HMC).

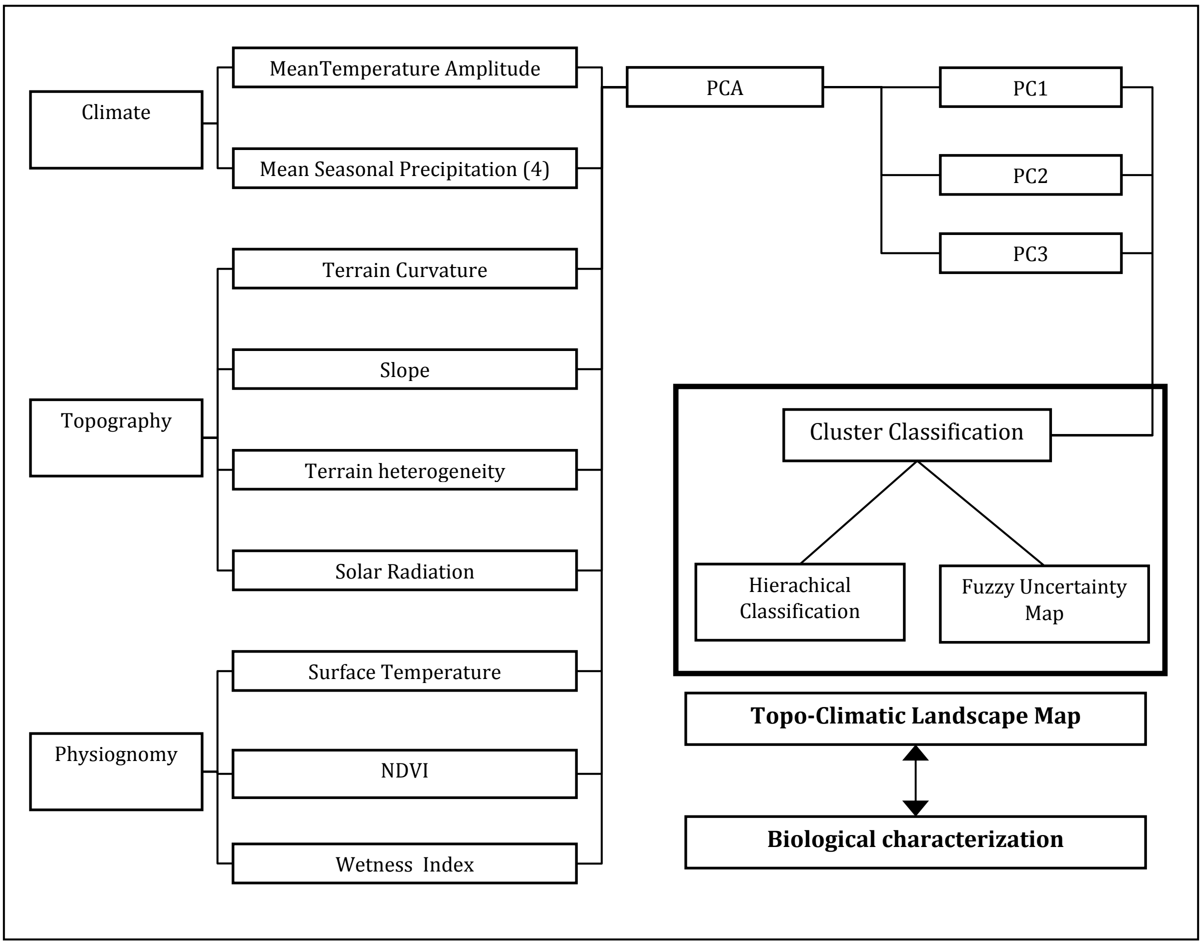

Figure 2. Methodology scheme for delineation of Topo-climatic landscapes. Variables are reduced through a principal component analysis (PCA) to three variables or principals components. These principal components (PC) will be used for subsequent cluster analysis.

2.2 Topographic, climatic and physiognomic variables.

2.2.1 Climatic variables.

Variables selected for classification were: mean air temperature amplitude and mean seasonal precipitation for each season (autumn, winter, spring and summer), hence five variables in total. Mean air temperature amplitude results from the difference between mean annual maximum temperatures and mean annual minimum temperatures. This variable measures the degree of continentality as well as the mean range of temperatures within a region. In the case of precipitation, four mean seasonal precipitation values were used due to seasonal differences in the Mediterranean climate.

Climatic variables were derived from the Digital Climatic Atlas of Catalonia (Ninyerola et al 2000; available at: http://opengis.uab.es/WMS/acdc/index.htm). This cartography consists of 65 monthly maps of mean air temperature (minimum, mean and maximum), precipitation and solar radiation derived from 160 to 257 meteorological stations, depending on the variable. It is based on statistical methods, for example multiple regression, and spatial interpolation techniques (inverse distance weighted and kriging), implemented in a GIS environment. Data from meteorological stations has been combined with altitude, latitude, distances to sea, solar radiation and terrain curvature to obtain 180 m spatial resolution grids of every climatic variable. Cross-validation results show a root mean square error (RMSE) between 6-20 mm, for precipitation, and between 0.8-1.5 ºC for mean temperatures.

2.2.2 Topographic variables.

Topographic and geomorphological characteristics of a landscape have been shown to be related to biodiversity and vegetation spatial distribution (del Barrio et al 1997; Burnett et al 1998; Pfeffer et al 2003; Bailey 2004). Here, four topographic variables were used for classification: slope, terrain curvature, terrain heterogeneity and solar radiation.

Slope and terrain curvature were derived from a 30 m DEM of the Catalan Cartographic Institute through basic GIS tools. Terrain heterogeneity was calculated for each cell as the standard deviation of elevation in a 3 by 3 cell window, thus detailing the degree of hilliness of a terrain. These variables have been resampled to a 180 m spatial resolution to match them with the resolution of climatic variables.

Solar Radiation was also obtained from the Digital Climatic Atlas of Catalonia (Ninyerola et al 2000). Solar Radiation was first obtained through a computational model, using astronomic equations and relief (DEM) (Pons and Ninyerola, in press). The subsequent correction of these maps of solar radiation through meteorological stations enabled the inclusion of cloudiness effects. Note that this variable integrates other variables such as slope, aspect and self and cast shadows.

2.2.3 Physiognomy variables

Variables related to the physiognomy of different land covers of the study area have been calculated from remote sensing imagery. Land surface temperatures (LST), normalized difference vegetation index (NDVI) and wetness index computed through tasselled cap (WI) were obtained from Landsat-5 TM and Landsat-7 ETM+ within the 2002-2006 period. The

resulting images were corrected radiometrically (Pons and Solé-Sugrañes 1994; Valor et al 2000) and geometrically, obtaining a RMSE of less than 30m (Palà and Pons 1995).

Physiognomy variables were resampled from a 30 meter spatial resolution to 180 meters, in order to match spatial resolution of climatic and topographic data.

2.3 Reducing redundancy among variables (PCA).

Principal Component Analysis (PCA) was applied to the 12 variables selected using ENVI 4.3 software (ITT industries, Inc. 2006). Through PCA, we have reduced redundancy among the variables and three components have been extracted for subsequent classification, using eigenvalues criteria (Kaiser's rule) coupled with a variability threshold criteria: only eigenvalues over 1 and explaining more than 80 % of total variability are selected. Principal component over 1 explain more than a single variable itself (Mora and Iverson, 2002), thus they are selected for classifications. In addition to this criterion, all principal components selected must explain at least the 80% of the total variability; otherwise additional principal components should be included in order to achieve this threshold.

Table 1 shows that first and second components explain much of the variability: both explain more than the 72 % of the total variance, however a third component had to be added to the classification procedure in order to achieve the 80 % variability.

Principal component	1	2	3	4	5
Eigen Values	7.36	1.30	0.99	0.90	0.68
% explained variability	61.31	10.84	8.31	7.51	5.74
Main correlated variables	CURV,THETER, SLP,SR,TA,SmP, WiP,NDVI,LST	WI,NDVI AuP	SpP	-	-

Table 1. Principal Component Analysis summary table. Eigen values and percentage of explained variability are given for the first five principal components. Through interpretation of coefficients of Eigen vectors, the most influencing variables for every principal component are indicated. CURV: curvature; THETER: terrain

heterogeneity; SLP: slope; SR: solar radiation; TA: temperature amplitude; SmP: summer precipitation; WiP: winter precipitation; NDVI: normalized difference vegetation index; LST: land surface temperature; WI: wetness index; AuP: autumn precipitation; SpP: spring precipitation.

The first principal component is linked to the topographic characteristics of the terrain. Mean air temperature amplitude, summer precipitation and winter precipitation are also significant, as well as remote sensing variables of LST and NDVI. Coefficients of eigenvectors for this first principal component do not show large differences between these significant variables, although topographic variables coefficients are slightly larger.

For the second principal component, wetness index becomes the most relevant although autumn precipitation, terrain curvature and NDVI play an important secondary role in this component. Finally, the third component reflects the maximum precipitation falling in a Mediterranean region, with spring precipitation having the highest influence.

2.4 Clustering classification: identifying TCLs.

A non-supervised cluster classification technique was performed for the three principal components specified above, using ISOMM module of MiraMon GIS (Pons 1998) based on the Duda and Hart algorithm (Duda and Hart 1973).

This module groups cells with similar values of the three principal components mentioned above, resulting in a raster where cells are classified into categories.

The module randomly assigns a certain set number of seeds through three spaces: the whole multidimensional space of variables, the multivariate diagonal and the physical space (seeds sampled over the image). Then, it calculates the euclidean distance of every cell to these cluster centres (seeds). Subsequently, ISOMM aggregates cells in clusters, hence classifying input rasters (the 3 principal components) in different groups. ISOMM determines whether some of these groups have to be merged or considered separately, depending on the parameters defined. Once cells are assigned to a specific class, it re-calculates a new statistical centre of the class and performs a distance calculation for every cell to this new center of the cluster.

These operations run iteratively until 98% of cells in the raster are classified or a maximum of 200 iterations are performed, which are parameters defined by the researcher.

Two parameters significantly influence the final number of clusters thus, the whole classification: the minimum similarity between classes and the minimum cluster area.

A minimum similarity threshold between two classes must be set, ensuring that output classes (clusters) are distinct and therefore, have meaning as a class. This depends basically on the amount of clusters and how different are classes expected to be, as well as the number of input variables (or bands) considered.

Concerning to the initial number of class centres parameter, it must be set to be sufficiently large to cover a big percentage of the multivariate diagonal. We rejected introducing any initial seed throughout the whole multidimensional space since it has been demonstrated that there is little or no change by adding seeds randomly in the whole multidimensional space: there is little chance for a seed to fall in a real value or nearby, hence the majority of seeds introduced in the whole multidimensional space are erased after the first iteration (Pons et al 2006).

An initial idea of the approximate number of classes to identify is necessary when performing this analysis. Many classifications with different number of categories can be created by changing some parameters; therefore many test-beds have to be performed. For managing purposes we accepted a maximum of 30 TCLs as the most detailed working scale (see also section 4.Discussion and conclusions). Subsequent aggregation of centroids will enable working at broader scales with fewer categories (see section 2.5).

2.4.1 IsoMM behaviour.

Seven series of test-beds have been performed, taking into account influencing parameters of IsoMM module in the resulting number of classes: minimum similarity between classes and minimum cluster area.

As expected, number of categories decreases when raising cluster minimum area parameter thus hampering the creation of TCLs of small area. On the contrary, it is difficult to predict the trend for minimum similarity (understood as distance) between clusters. Theoretically, number of categories would decrease when increasing this parameter: statistical distances between cluster centroids become larger as fewer categories are created. Nevertheless, this distance is implemented as a Manhattan distance applied to initial seeds over the image, which makes difficult to predict its behaviour and becomes largely dependable on number of initial seeds and structure of data.

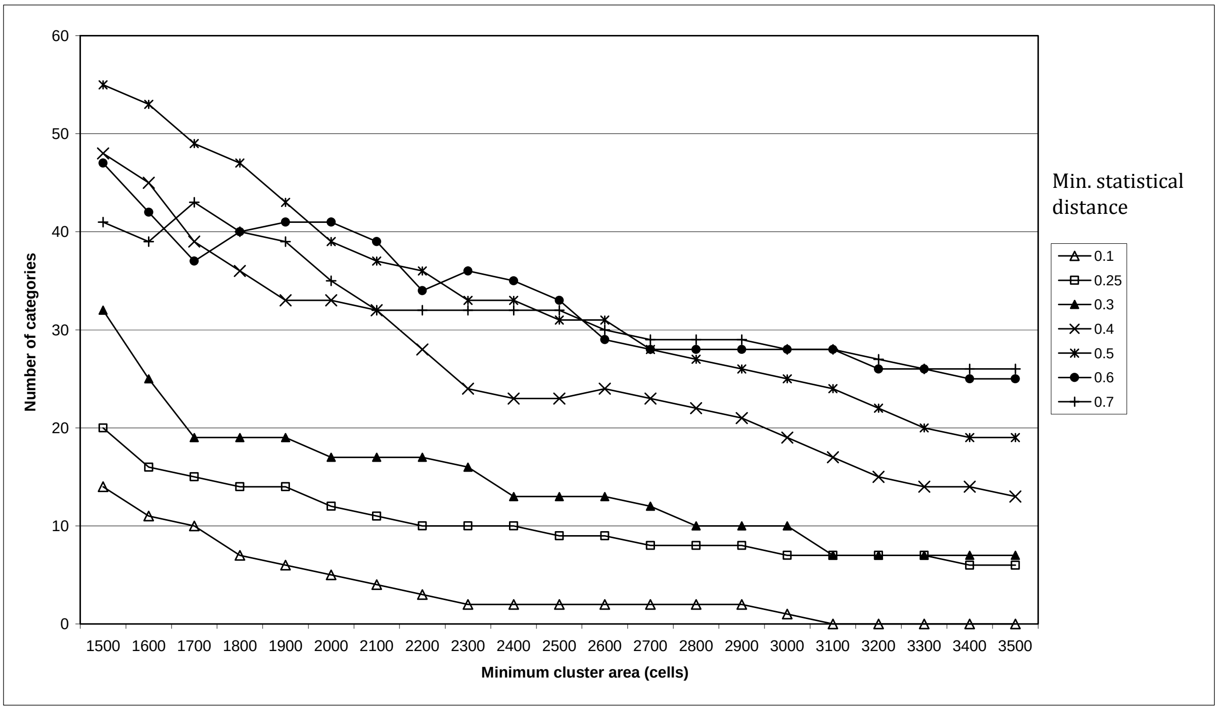

Figure 3. IsoMM behaviour has been evaluated through several test-beds. Line-marks identify series with different minimum statistical distance between clusters and the resulting number of classes when changing minimum cluster area parameter.

After performing these test-beds, parameters chosen for classification were a minimum cluster area of 2700 cells and the parameter of minimum distance between clusters was set to 0.7. These parameters ensure a number of categories (28) that match managing purposes, which accepted a maximum of 30 classes. Other tests with similar parameters result in the same number of clusters (see Figure 3) and barely exact boundaries between TCLs, which is considered to be a stable behaviour. However a Kappa analysis among these results will be considered to statistically prove little changes between similar cartographies.

2.5 Hierarchical relationships between TCLs.

The next step is realized after classifying the whole image: a classification for each cluster centroid (mean values of each cluster in multidimensional space) through Ward's method

(Ward, 1963) using STATISTICA 6.0 software (StatSoft Inc. 2001). A statistical distance between the clusters' centroids was calculated thus, a dissimilarity matrix could be built. The method uses analysis of variance to evaluate distances between clusters. Results show similarity relationship between clusters (Figure 4), which enables grouping TCLs and work at broader scales (see Figure 6) .

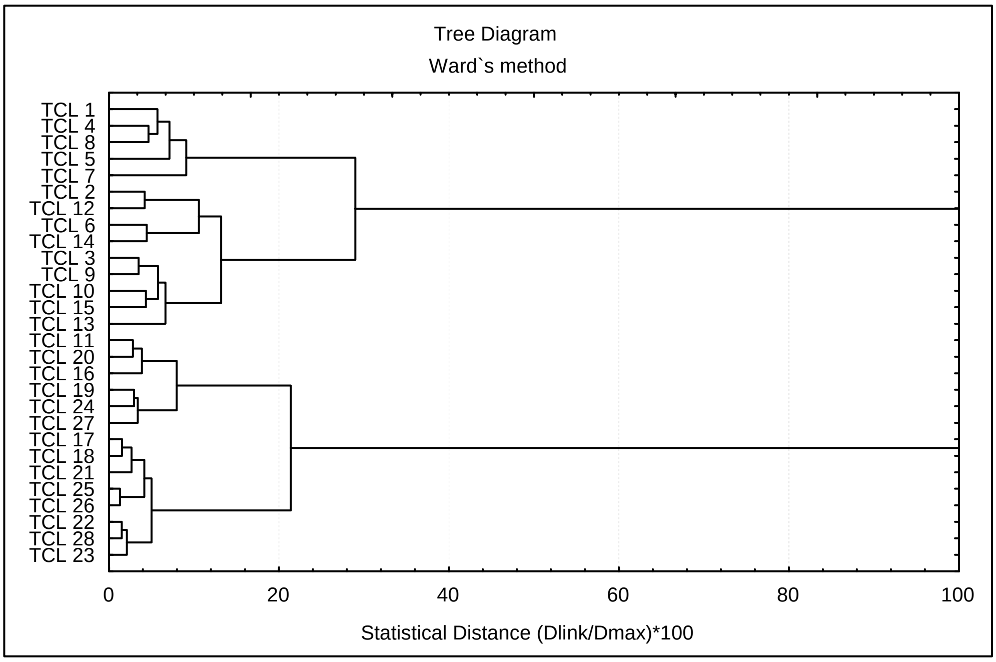

Figure 4. Diagram of aggregation of the 28 TCLs through a standardized distance (distance of linkage / maximum distance) using Ward's clustering method.

2.6 Adding uncertainty to the crisp classification.

In order to assess the uncertainty of the non-supervised clustering classification realized, an uncertainty map was created through a membership function derived from a fuzzy classification tool (supervised classification).

TCL patches over 20 ha and with a mean Euclidean distance to the TCL class centroid of under 0.25 (see section 3.1 TCLs biological characterization) have been used as training sites, as they are considered to be the statistical core of a TCL, hence the most representative patches have been selected.

GIS Fuzzy classification tool used for the present analysis was Idrisi 32 FUZCLAS module (Eastman 1999). FUZCLAS produces a classification of uncertainty which gives information about the maximum degree of membership of a cell to a certain class, ranging from 0 to 1.

The aim is to create a multi-layer raster map where the upper layer corresponds to a TCL map and the bottom layer to an uncertainty value for every cell. Therefore relevant information is added to the TCLs map and it can be easily displayed through GIS transparency functions (see Figure 5).

Twenty different combinations of TCL training sites have been chosen, hence 20 fuzzy uncertainty grids produced. In every grid, combinations of 25 percent of the total selected patches of every TCL (area over 20 ha and less than 0.25 to class centroid) were used as training sites. The uncertainty linked to a cell is calculated as the mean of these 20 fuzzy uncertainty grids. Moreover, we have assessed the reliability of the resulting fuzzy uncertainty map by calculating the standard deviation of the membership value for every pixel using all 20 uncertainty grids. Results show a mean standard deviation of 0.1 for the whole map and a maximum of 0.49 standard deviation. This value enhances reliability on the classification performed and indicates an acceptable stability of the fuzzy uncertainty map.

3. RESULTS: The topo-climatic landscape map.

Following the methodology presented in section 2, a multilayer display compounded by a clustering classification map and fuzzy uncertainty grid is presented as the Topo-climatic Landscape Map. In the present study 28 TCLs were obtained from cluster classification for the study area, following management consideration and taking into account ISOMM stability behaviour (see section 2.4.1).

The main relief features such as mountains, plains and valleys can be identified although we observe different categories within them. High mountains of the Pyrenees are grouped in TCL no.5 because of their similar climatic behaviour and land cover (rocky areas). Main valleys show higher complexity of TCLs. North and south slope are considered separately in these valleys due to different climatic characteristics (i.e. radiation) which will lead to different vegetation patterns. Similarly, valleys show different TCLs depending on their geomorphology (narrow valleys or wide valleys).

Main plain relieves are characterised by several TCLs (e.g. Lledia Plain and Cerdanya plain, see Figure 1 and Figure 5). Regardless of the exact meaning of every TCL, it is important to stress out the different topographic and climatic nuance detected through the present methodology. Hence, different plains geographically apart may be considered together with regard to their climate and topography.

In Figure 5 a zoom of the TCL map is shown. It can be appreciated in detail the delimitation of TCLs as well as the information provided by the uncertainty map. Within a TCL, light colours are linked to high uncertainty in classification whereas dark colours indicate low uncertainty. As expected, observed results show that low uncertainty values correspond to high mountains regions due to their specific climatic conditions, very different to low steep areas or hilly areas, where higher uncertainty values are present.

After analyzing similarity relationships between clusters (see section 2.5). TCLs can be grouped following a hierarchy derived from applying Ward's method clustering (Figure 4) to TCLs' centroids. These levels of generalization (Figure 6) enable adapting scale to ecological processes under study.

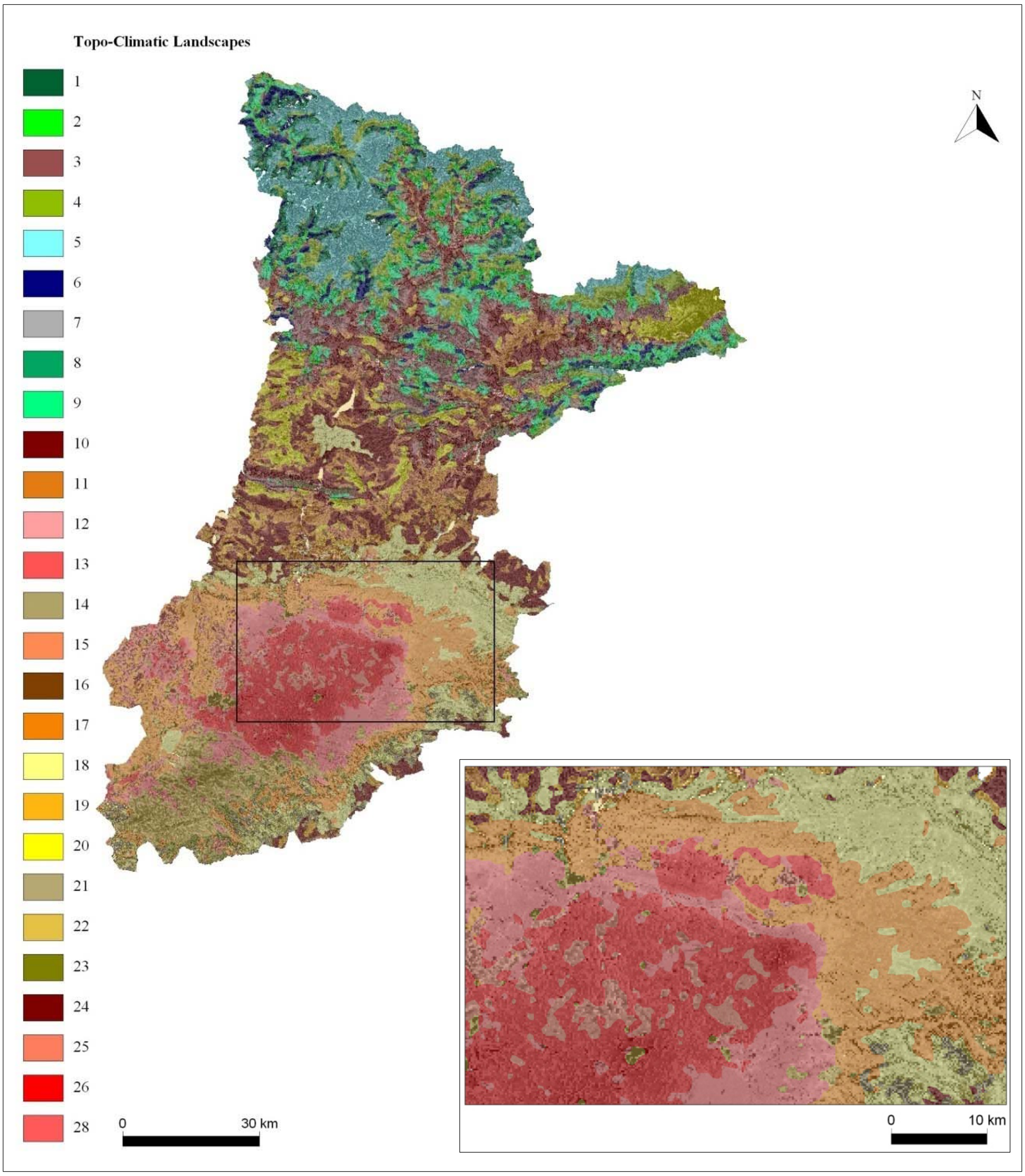

Figure 5. Topo-Climatic Landscape Map. Different degrees of intensity in colors can be appreciated in the Topo-Climatic Landscape Map. In the area zoomed, two intensity of colors can be appreciated: deep colors show a strong degree of membership to the class whereas light colors represent less degree of belonging of a pixel to a class. This is due to the fuzzy membership function which produced a raster image with uncertainty values for every pixel.

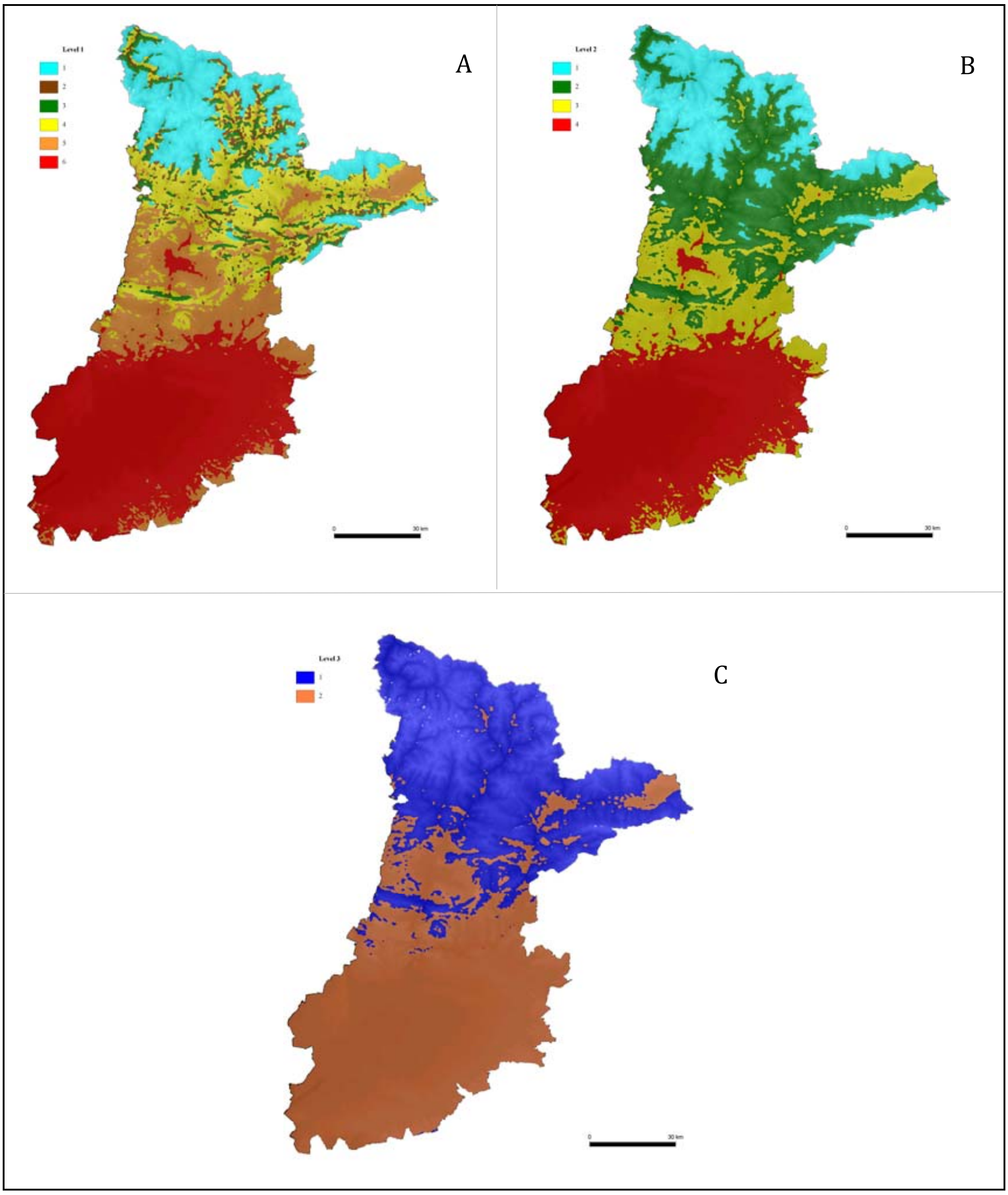

Figure 6. Different levels of generalization. Using the hierarchy derived from Ward's clustering method (see Figure 4), TCLs can be grouped according to a statistical distance criteria. A: 6 groups (10 Dlink/Dmax*100); B: 4 groups (20 Dlink/Dmax*100); C: 2 groups (40 Dlink/Dmax*100).

3.1 TCLs biological characterization.

We considered a biological characterization of TCLs very interesting; since it is necessary to asses to what extent TCLs explain vegetation composition.

It is expected that different TCLs present different composition of habitats. Through GIS layer combination tools; we evaluated habitat homogeneity within each TCL and differences in terms of habitat composition between TCLs.

Habitat information has been extracted from the HMC**,** which identifies up to 276 habitats based on “CORINE bitotopes manual”(see http://biodiversity-chm.eea.europa.eu/information/document/F1088156525/F1125582140). To avoid excessively complex analysis, we classified these categories into two different groups: physiognomic types (conifer forests, schlerophyllous forests, Mediterranean shrubs and grasslands) and altitudinal distribution types (high mountains, mid mountains and lowlands vegetation). Rocky habitats have been reclassified as nodata.

Two multiple correspondence analysis (MCA) have been performed using the habitat composition of each TCL, thus characterization and similarity among TCLs can be evaluated. MCA has been firstly applied to physiognomic and secondly applied to altitude distribution of habitats. Such an approach is commonly used to relate abiotic factors and vegetation responses (e.g., Jongman et al 1995).

Due to scale constraints and accuracy of analysis, only the most representative patches of TCLs were selected: those over 20 ha and with a mean Euclidean distance to the TCL class centroid under 0.25.

In order to perform MCA, vegetation types have been reduced to several dimensions that explain much of the total variability. In the case of physiognomy, 4 dimensions have been considered whereas 3 dimensions for the altitudinal distribution type.

Each TCL is characterized by a mean value of every dimension of physiognomic type and another mean value of every dimension of altitudinal distribution type. These mean values correspond to each centroid of the TCL for the two vegetation types considered. Consequently, every patch of a TCL is located at a certain euclidean distance to its altitudinal type and physiognomic type centroid. Patches are considered homogeneous for a vegetation type when euclidean distances are less than 0.5 for every dimension to the TCL vegetation type centroid.

The total statistical distance accepted for a patch in order to be considered homogeneous was: $1[(0.5)^2 * 4\text{dimensions}]$ in the case of physiognomy and $0.75\ [(0.5)^2 * 3\ \text{dimensions}]$ for altitude range. Moreover, a homogeneity index has been created in order to assess how different are patches of the same TCL (see Table 2). This index indicates the percentage of patches considered homogeneous for a single TCL in terms of vegetation types and it is based on the distance to the centroid patch of every TCL.

	Physiognomy type homogeneous patches (%)	**Altitude range vegetation type homogeneous patches (%)**		**Physiognomy type homogeneous patches (%)**	**Altitude range vegetation type homogeneous patches (%)**
TCL 1	56	29	**TCL 15**	37	28
TCL 2	54	62	**TCL 16**	55	55
TCL 3	52	87	**TCL 17**	88	93
TCL 4	33	34	**TCL 18**	79	87
TCL 5	74	79	**TCL 19**	22	17
TCL 6	78	39	**TCL 20**	37	73
TCL 7	73	00	**TCL 21**	71	69
TCL 8	49	79	**TCL 22**	70	89
TCL 9	70	70	**TCL 23**	82	93
TCL 10	36	37	**TCL 24**	50	61
TCL 11	29	27	**TCL 25**	92	96
TCL 12	53	66	**TCL 26**	93	-
TCL 13	33	61	**TCL 27**	59	69
TCL 14	70	64	**TCL 28**	84	92

Table 2. Homogeneity indexes for every TCL. Percentage of patches considered homogeneous for every TCL. Homogeneity has been analyzed for physiognomy and for altitudinal range of the vegetation since it does not necessarily have to follow the same trend.

Figures 7 and 8 show results of the MCA performed. In every TCL (circles) and each variable of composition (crosses) are placed in two axes (dimensions). For physiognomy, 74.06 percent of the total variance is explained and 87.11 percent of the total variance is explained in the case of altitude range. It is important to state that as a TCL gets closer to a variable (crosses) it is expected to find a higher degree of that particular sort of vegetation compared to other TCLs. However, it does not imply a higher representation of particular vegetation over other types of vegetation considered; instead it shows the degree in which a certain vegetation physiognomy is overrepresented in a TCL. The same interpretation must also be applied to altitudinal range classification.

MCA. Physiognomy type.

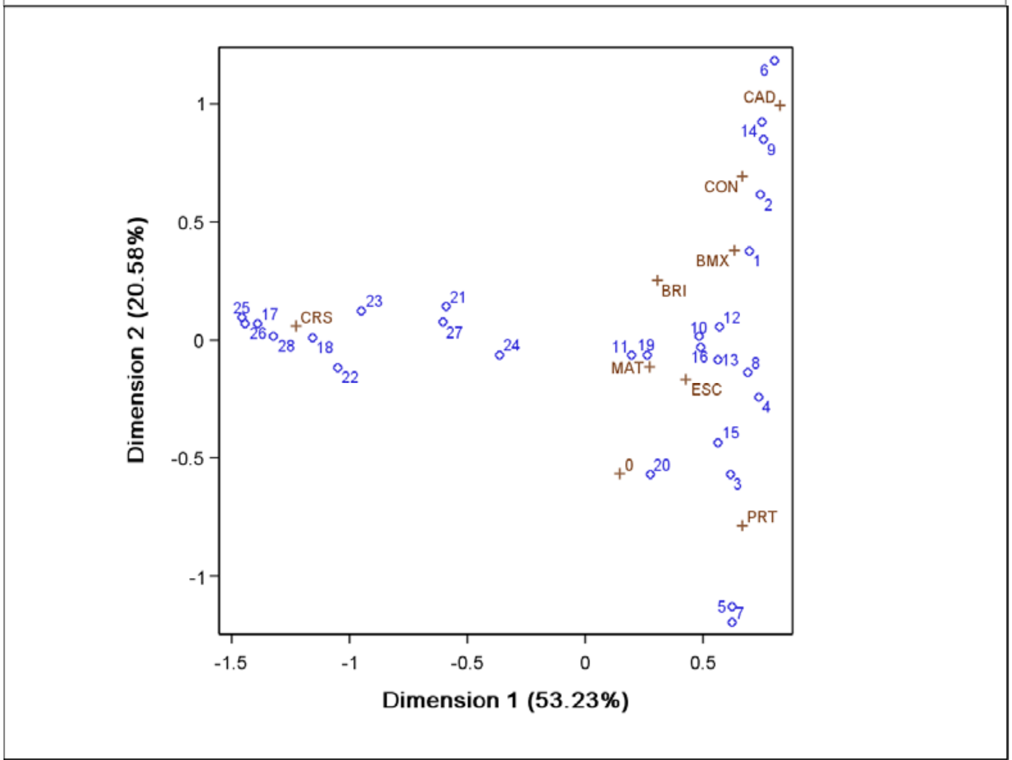

Figure 7. Multiple correspondence analysis using physiognomy of vegetation in patches of every TCL. Every variable case is marked with a cross and every TCL is marked with a circle. Two dimensions explain 73.81 % of the total variability. CRS=crops; MAT=shrubs; BRI= river forests; BMX=mixed forests; CON=conifer forests; CAD=deciduous forests; ESC=sclerophyllous forests; PRT=meadow; 0= no data.

MCA. Altitudinal range type.

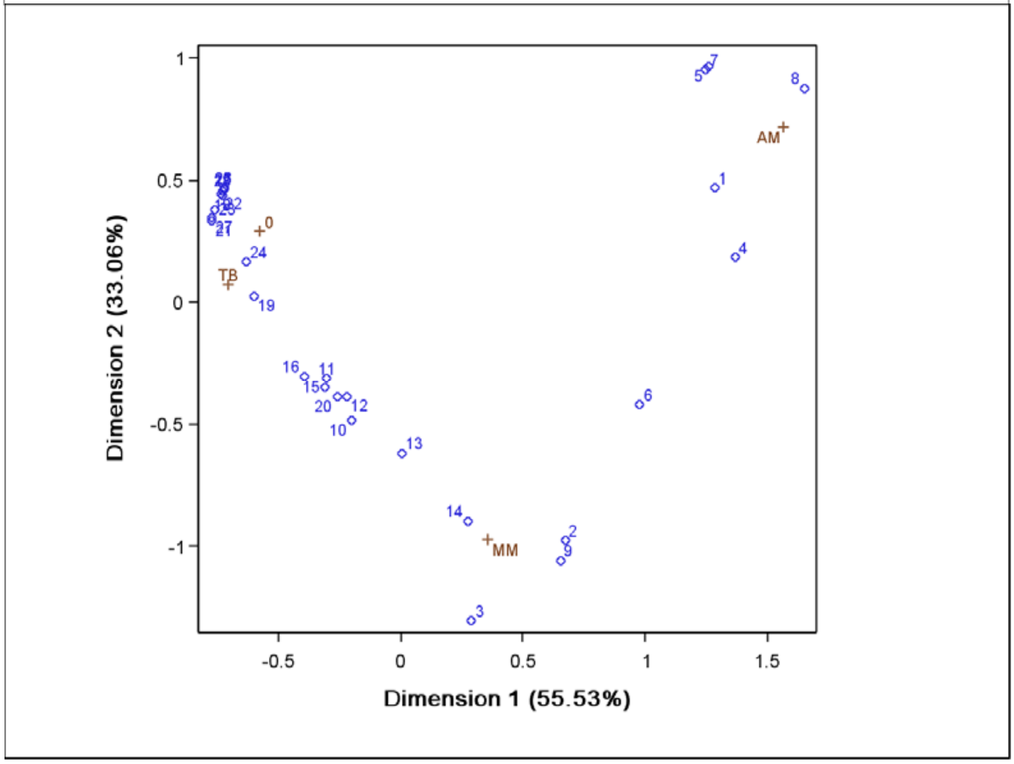

Figure 8. Multiple correspondence analysis using altitudinal range of vegetation in patches of every TCL. Every variable case is marked with a cross and every TCL is marked with a circle. Two dimensions explain 88.59 % of the total variability. AM=High mountains; MM=mid-mountain belt; TB=lowlands; 0=no data.

We did not perform a deeper analysis of TCLs no's: 17, 18, 22, 23, 25, 26 and 28 as they are mainly composed by crops and that was not the aim of our study, which is focused on natural vegetation landscapes.

A transition region between crops (CRS) and shrubby areas (MAT) is observed and identified by TCL no's 21, 27, 24. MCA shows that crops are quite dominant in these TCLs but shrubs and sclerophyllous are significantly present in TCLs to a different degree as seen in Figure 7. Therefore, MCA suggest a link between these TCLs and agricultural abandonment, but further analysis is required in order to statistically prove such a hypothesis.

A third group of TCLs could be identified by an over-representation of conifers (CON) and deciduous forests (CAD) including TCLs 2,14,9 and 6. These TCLs are located around the category of mid-mountains (MM) but there are different trends between them (Figure 8). Similarity between TCLs 2, 9 and 14 for the altitudinal range is obvious however, it is important to detect a different trend in TCL 6, in which the presence of high mountain communities is relevant. Concerning homogeneity, patches of TCLs 9 and 14 are considered homogeneous whereas patches of TCL 2 are considered homogeneous only for altitudinal range and TCL 6 only for physiognomy.

TCL 1 has been considered as a group itself, since mixture woodlands are overrepresented and it is related to high mountain vegetation. This TCL is linked to the transition between conifers and deciduous.

TCL 5 and TCL 7 are characterized by high mountain grasslands (see also Figure 8) and are located along all summits of the Pyrenees (Figure 1). Note that both TCLs present homogeneous patches in physiognomy (grasslands) but TCL 7 is not homogeneous for altitudinal range. This is due to reclassification as nodata of rocky areas category. That decreased the number of patches that could be included in the MCA and strongly affected TCLs not accounting for a large area, for instance TCL 7.

TCLs 3 and 15 become a heterogeneous group, since only TCL 3 scores homogeneous for altitudinal ranges (mid-mountain vegetation). However, both clusters are characterized by high percentages of grasslands and to a lesser extent by shrubs. They both belong to a mid-mountain altitudinal range. TCL 15 is strongly linked to lowlands and its patches spread through the study area when decreasing latitude. TCL 20 could be added to this group (TCL 15 and 3) as it is also a heterogeneous TCL showing similar representation of sclerophyllous, shrubs and grasslands as well as belonging to the same altitudinal range. But we considered it separately since it has significant number of patches with nodata (rocky habitats) and crops.

TCL no’s 4, 8, 10, 11, 12, 13, 16 and 19 represent a heterogeneous group where sclerophyllous and shrubs are highly dominant. Other categories such as crops or conifers, occupy significant percentages of these TCLs. Differences in Figure 7 are due to different compositions of sclerophyllous, shrubs, conifers crops and grasslands and subgroups could be identified.

3.2 Statistical differences for similar TCLs.

Many TCLs can be characterized and differenced for altitudinal range or physiognomy through correspondence analysis (see Figures 7 and 8). However, some TCLs show similarity in both physiognomy and altitudinal range and therefore differences are not obvious, which is the case for TCL 14 and TCL 9.

A second statistical analysis was performed for TCLs showing little statistical distance in the multiple correspondence analysis. A Kruskal-Wallis test was used for each physiognomy variable as well as for each variable of altitudinal range. This test showed a *p* significance level for the Kruskal-Wallis test, determining that there exist significant differences ($p<0.05$) between TCL 9 and TCL 14 for mid-mountains and lowlands habitats and for conifers and sclerophyllous distribution in patches. These results suggest that these differences are due to the drier and lower character of TCL 14, where sclerophyllous seem to occur more frequently than in TCL 9.

3.3 Spatial meaning of TCLs: the case of Cerdanya Valley

We have assessed the spatial meaning of TCLs, and hence the usefulness of the elaborated map through Cerdanya Valley case (see Figure 1).

A transect of 24 km has been drawn in a NE-SW direction. As seen in Figure 9, there exist different TCLs' composition between north and south face. A wide plain of ca. 8000 m is located in between. In the South face relief exerts a relevant role, showing small abrupt changes in altitude, which lead to a wide range of environmental conditions, mainly because of differences in radiation interception angle, shade distribution, etc.

Analyzing the elevation profile from left to right (north-east to south-west) (see Figure 9), it can be appreciated that TCL 5 is characterized by the highest slopes of the Pyrenean Mountains, with a southern orientation. Habitats found within this TCL are mostly grasslands with species such as *Festuca eskia* and *Nardus stricta*.

TCL 4 is geographically located between TCL5 and TCL3. This TCL is not as steep as TCL5 and is mainly south oriented. It is found in both sides of the valley as the aspect changes, between 1600 m and 2200 m. Concerning to habitat composition, the HMC show that grasslands are still prevalent in this TCL although these grassland species indicate mesophile and sub alpine conditions of the mid-mountain belt.

TCL 3 appears three times in the profile and it seems not to be strongly dependent on relief characteristics or aspect. However, a broader analysis of this TCL indicates that it often

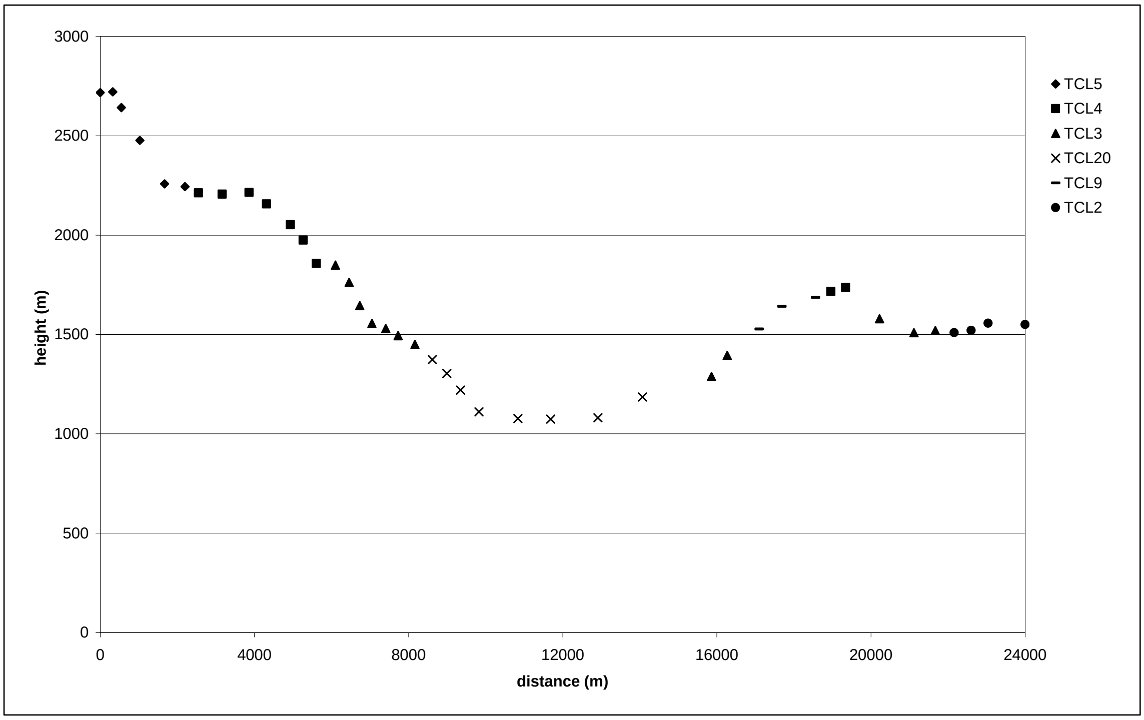

Figure.9 Elevation profile of Cerdanya valley (see Fig.1). The profile suggests a strong relationship between altitude, aspect and TCLs.

appears in south slopes of high mountains, at their base. A habitat analysis found that grasslands with montane species composition are still relevant in this TCL, but managed grasslands are significantly present. We concluded that managed grasslands present at both the south and north oriented slopes of the mountain allow for this TCLs occurrence at both sides. However, the sunny character of TCL3 meant it only appeared as a small patch in the north face.

TCL 20 relates to valleys located around a height of 1000 m and wide enough to be soundly influenced by solar radiation. These characteristics differentiate TCL 20 from plains located at lower latitudes or elevations (see Figure 5). Habitats in this TCL are mainly composed of rushes and managed grasslands of mid-mountains or even lowlands with relatively arid environmental conditions.

TCL9 and TCL2 are found at the northern slope of the valley. The first is related to steep and north-face conditions in mountains at an altitude of between 1500 m and 1800 m. Its vegetation is characterized by conifer forests of high mountains and mid-mountains.

According to the HMC, forests of *Pinus sylvestris* occupy around 30% of this TCL and forests of *Pinus uncinata* around 20 %. In both cases the HMC specify that these forests are located in shady north faces of mountains, which confers with results obtained through cluster analysis.

TCL2 is related to relatively plain areas occupied by small streams forming secondary valleys between Pyrenean Mountains. A lot of small habitats occur in such areas, according to HMC, which gives very detailed classifications. However grouping them as realized in section 2.5, it is found that high mountain conifers occupy a high percentage of this TCL, but there is a high percentage of deciduous river species when compared to other TCLs.

4. DISCUSSION AND CONCLUSIONS.

In the present article a methodology is proposed in order to define homogeneous topo-climatic areas, useful for management and planning needs. Other landscape classifications have shown that defining topographic and climatic behaviour may be crucial for several human activities such as reforestation, strategic agricultural settlement and forest conservation policies (see 1.Introduction).

It is important to state that high quality and large quantity of data used for classification is needed in order to face a broad ecological classification. We have combined information provided from the Digital Climatic Atlas of Catalonia, remote sensing imagery from the period 2002-2006 and topography derived from a 30 m DEM. The indices computed through remote sensing have added relevant information about the physiognomy of the land cover such as water content, structure and vegetation physiology. For instance, TCL 22 identifies water reservoirs thanks to the introduction of these indices.

The number of topo-climatic landscapes to be obtained is a key issue in such methodologies, since we need to consider variance inside clusters and take into account the working scale of users of the presented topo-climatic landscapes cartography. Therefore, an initial idea of the maximum categories that should be defined is needed, but without being restricted to that exact number. In the present article a maximum of 30 categories defined *a priori* were considered as the most detailed scale that could be identified for our study area.

The unsupervised classifier uses certain statistical parameters that must be set. Hence, expert knowledge in clustering techniques is highly recommended to successfully perform such a classification. These parameters as well as variables of interest are defined by the researcher; therefore a subjective component is present in classification although this quantitative approach is in principle less judgemental than other visual methods (Fairbanks et al 2000).

The Ward's method of tree-joining has reduced the number of TCLs and it has made possible to define different scales, depending on a statistical distance threshold. Furthermore, the output graph of this second classification (Figure 4) has enabled analysing similarity relationships between different TCLs. Ward's method takes into account not just the mean conditions (the centroid of every cluster) but also minimizes variances when merging topo-climatic landscapes. That is quite a relevant issue since it takes into account the inherent variability of a topo-climatic landscape.

Results from the correspondence analysis showed that Topo-climatic Landscapes obtained can be characterized and differentiated in terms of vegetation physiognomy types and the altitudinal range of vegetation. However, some TCLs cannot be defined by a single type of physiognomy or altitudinal range of vegetation. That is in line with expectations, as vegetation is subjected to human and natural perturbations and other factors, although climate and relief exert large influences. Moreover, some TCLs (e.g. TCL14 and TCL9) do show both similarities in vegetation physiognomy types and in altitudinal ranges of vegetation. However, the Kruskal-Wallis test showed that there exist significant differences in occurrences of both physiognomy and altitudinal ranges between these similar TCLs.

It might be argued that some TCLs are quite similar in terms of vegetation types, although significantly different in their composition but it does not represent a clear differentiation in terms of vegetation. It is important to stress out that differences in relief and climate might lead to eco-physiological differences between these similar TCLs or differences in terms of potential vegetation or species suitability, but not necessarily lead to differences in the presence or absence of a certain habitat or specie or composition. Unfortunately, assessing eco-physiological differences for the same habitats in different TCLs has not been tested yet.

In fact, it becomes difficult to interpret topo-climatic landscapes because there is not a perceptible direct response of vegetation. That could hamper managing uses since it is difficult to understand the whole set of conditions within a TCL, but such classifications represent a very useful tool for strategic planning of vegetation.

Further information is added in order to tackle with the issue of crisp classifications: managers need to *draw lines* to apply policies but at the same time understand environmental gradients and complexity of ecological data. In the present case, a crisp classification of environmental topo-climatic gradients is performed through the first cluster analysis that defines Topo-climatic Landscapes, where each TCL patch is assigned to a single class exclusively as in many thematic classifications. However, a fuzzy analysis which produced the final classification of uncertainty, improves the crisp classification by adding a pixel membership value. Other studies, for instance Mitchell et al (2008), explore pixel membership values of a cluster classification through the calculation of the euclidean distances to second nearest cluster centroid in order to tackle with uncertainty, thus the confidence of class membership is mapped.

In our approach, fuzzy classification tool is used not as a classifier in itself, but rather as a measure of pixel membership to a certain TCL. Although it could be argued that fuzzy classification would improve traditional crisp classifiers (Rocchini et al 2007), expert knowledge cannot identify training sites of topo-climatic conditions. Therefore, training sites in this case have been identified statistically as patches near the mean conditions of a TCL.

As showed in the present study, combining unsupervised hierarchical cluster classifications and fuzzy analysis could help in the development of environmental classifications that would be useful for managers and researchers. Accordingly, some fuzzy clustering algorithms have recently been created (Salski 2007) but still need to be implemented in a GIS environment.

ACKNOWLEDGEMENTS.

We are grateful to the research group GRUMETS of the Autonomous University of Barcelona, and especially to Gerard Moré for his advice on GIS cluster analysis. We also want to thank to the research group the Applied Geography Group (2005SGR00942) of the Autonomous University of Barcelona for the development of this methodology in the High Pyrenees Natural Park (2005-2006PNATAPI).

We would like to express our gratitude to the Catalan Water Agency and to the Department of the Environment and Housing of the Generalitat (Autonomous Government) of Catalonia

for their investment policy and the availability of Remote Sensing data, which has made it possible to conduct this study under optimal conditions.

REFERENCES

Bailey R (1996) Ecosystem Geography. Springer. Berlin.

Bailey R (2004) Identifying ecoregion boundaries. Environ Manage 34: S14-S26.

Barnes BV, Pregitzer K S, Spies TA, Spooner VH (1992) Ecological forest site classification. J. For. 80: 493-498.

Barrio G del, Alvera B, Puigdefabregas J, Diez C (1997) Response of high mountain landscape to topographic variables: Central Pyrenees. Landscape Ecol 12: 95-115.

Bernert JA, Eilers JM, Freemark KE, Ribic C (1997) A quantitative method for delineating regions: An example for the western corn belt plains ecoregion of the USA. Environ Manage 21: 405-420.

Blankson EJ, Green BH (1991) Use of landscape classification as an essential prerequisite to landscape evaluation. Landscape and Urban planning 21: 149-162.

Blaschke J, Strobl G (2003) Defining landscape units through integrated morphometric characteristics. In: Buhmann E. and Ervin S. (eds), Landscape Modeling: Digital Techniques for Landscape Architecture, Wichmann-Verlag, Heidelberg , pp. 104–113.

Bunce RGH, Barr CJ, Clarke RT, Howard D C, Lane AMJ (1996) Land classification for strategic ecological survey. Journal of Environmental Management 47: 37-60.

Burnett MR, August PV, Brown JH, Killingbeck KT (1998) The influence of geomorphological heterogeneity on biodiversity I. A patch-scale perspective. Conserv Biol 12: 363-370.

Burrough PA, van Gaans PFM, MacMillan RA (2000) High-resolution landform classification using fuzzy k-means. Fuzzy Sets Syst 113: 37-52.

Burrough PA, Wilson JP, van Gaans PFM, Hansen AJ (2001) Fuzzy k-means classification of topo-climatic data as an aid to forest mapping in the greater yellowstone area, USA. Landscape Ecol 16: 523-546.

Carter RE, MacKenzie MD, Gjerstad DH (1999) Ecological land classification in the southern loam hills of south Alabama. Forest Ecology and Management 114: 395-404.

Chon T, Park Y, Park JH (2000) Determining temporal pattern of community dynamics by using unsupervised learning algorithms. Ecological Modelling 132: 151-166.

Cohen WB and Goward SN (2004) Landsat's role in ecological applications of remote sensing. Bioscience 54 (6): 535-545.

Clavero P, Martin Vide J, Raso Nadal JM (1996) Atles climàtic de Catalunya. Termopluviometria, Generalitat de Catalunya (Departament de Politica Territorrial i Obres Públiques), Institut Cartogràfic de Catalunya i Departament de Medi Ambient, Barcelona.

Cleland DT, Crow TR, Avers PE, Probst JP (1992) Principles of land stratification for delineating ecosystems. In: Taking an ecological approach to management. US Forest Service Watershed and Air Management. Pp.40-50.

Cleland DT, Avers PE, McNab WH, Jensen ME, Bailey RG, King T, Russell WE (1997) National Hierarchical Framework of Ecological Units. In: Boyce M, Haney A (eds), Ecosystem Management Applications for Sustainable Forest and Wildlife Resources. Yale University Press, New Haven, CT, pp 181-200.

Duda RO, Hart PE (1973) Pattern classification and scene analysis. Wiley-Interscience Publication. New York.

Eastman JR (1999) IDRISI32.Guide to GIS and Image Processing. User's Guide, Version I32.01, Clark University, Worcester, MA, USA.

Fairbanks DHK, Benn GA (2000) Identifying regional landscapes for conservation planning: A case study from KwaZulu-natal, South Africa. Landscape and Urban planning 50: 237-257.

García de Pedraza L, Rejia G (1994) Tiempo y clima en España. Meteorología de las Autonomías. DOSSAT-2000, Madrid, Spain.

Habitats Map of Catalonia. (2005. Housing and Environment Department of Genarlitat de Catalunya and University of Barcelona, Barcelona. Available form http://mediambient.gencat.net/cat/el_medi/habitats/habitats_cartografia.htm#cd. (accessed February 2008)

Haines-Young RH (1992) The use of remotely-sensed satellite imagery for landscape classification in Wales (U.K.). Landscape Ecol 7: 253-274.

Hargrove W, Hoffman F (2004) Potential of multivariate quantitative methods for delineation and visualization of ecoregions. Environ Manage. 34: S39-S60.

Host GE, Polzer PL, Mladenoff DJ, White MA, Crow TR (1996) A quantitative approach to developing regional ecosystem classifications. Ecological Applications 6: 608-618.

Hutto JC, Shelburne BV, Jones SM (1999) Preliminary ecological land classification of the Chauga Ridges region of South Carolina.114: 385-393.

ITT Industries Inc. (2006) ENVI version 4.3. http://www.RSInc.com/envi.

Jongman RHG, Ter Braak CJF, van Tongeren OFR (1995). Data analysis in community and landscape ecology. Cambridge University Press, UK.

Kirkpatrick JB, Brown MJ (1994) A comparison of direct and environmental domain approaches to planning reservation of forest higher plant communities and species in Tasmania. Conserv Biol 8: 217-224.

Kraft J, Einax JW, Kowalik C (2004) Information theory for evaluating environmental classification systems. Analytical and Bioanalytical Chemistry. 380: 475-483.

Loveland TR, Merchant JM (2004) Ecoregions and ecoregionalization: Geographical and ecological perspectives. Environ Manage 34: S1-S13.

McGarigal K, Cushman SA, Neel MC, Ene E (2002) FRAGSTATS: Spatial Pattern Analysis Program for Categorical Maps. University of Massachusetts, Amherst. Available at the following web site: http://www.umass.edu/landeco/research/fragstats/fragstats.html.

Mackey BG, Nix HA, Hutchinson MF, Macmahon JP, Fleming PM (1988) Assessing representativeness of places for conservation reservation and heritage listing. Environ Manage 12: 501-514.

MacMillan RA, Pettapiece WW, Nolan SC, Goddard TW (2000) A generic procedure for automatically segmenting landforms into landform elements using DEMs, heuristic rules and fuzzy logic. Fuzzy Sets Syst 113: 81-109.

Martin de Agar P, de Pablo C, Pineda F (1995) Mapping the ecological structure of a territory: A case study in Madrid (central Spain). Environ Manage 19: 345-357.

Mitchell SW, Remmel TK, Csillag F, Wulder MA (2008) Distance to second cluster as a measure of classification confidence. Remote Sensing of Environment, 112: 2615-2626.

Monjeau JA, Birney EC, Ghermandi L, Sikes RS, Margutti L, Phillips CJ (1998) Plants, small mammals, and the hierarchical landscape classifications of Patagonia. Landscape Ecol 13: 285-306.

Mora F, Iverson L (2002) A spatially constrained ecological classification: rationale, methodology and implementation. Plant Ecology 158: 153-169.

Nadeau LB, Li C, Hans H (2004) Ecosystem mapping in the lower foothills subregion of Alberta: application of fuzzy logic. Forestry Chronicle 80: 359-365.

Ninyerola M, Pons X, Roure JM (2000) A methodological approach of climatological modelling of air temperature and precipitation through GIS techniques. International Journal of Climatology 20: 1823-1841.

Nolet P, Domon G, Bergeron Y (1995) Potentials and limitations of ecological classifications as a tool for forest management: a case study of disturbed deciduous forests, Québec. Forest Ecology and Management 78: 85-98.

Omernik JM (1987) Ecoregions of the conterminous United States. Annals of the Association of American Geographers 77: 118-125.

Palà V, Pons X (1995) Incorporation of relief into geometric corrections based on polynomials, Photogramm. Eng. Rem. S. 61: 935-944.

Pfeffer K, Pebesma E, Burrough P (2003) Mapping alpine vegetation using vegetation observations and topographic attributes. Landscape Ecol 18: 759-776.

Pons X, Solé-Sugrañes L (1994) A Simple Radiometric Correction Model to Improve Automatic Mapping of Vegetation from Multispectral Satellite Data. Remote Sens. Environ. 47: 1-14.

Pons X (1998) Manual of MiraMon. Geographic Information System and Remote Sensing Software. (http://www.creaf.uab.es/miramon). Centre de Recerca Ecològica i Aplicacions Forestals (CREAF): Bellaterra; 150.

Pons X, Moré G, Serra P (2006) Improvements on Classification by Tolerating NoData Values. Application to a Hybrid Classifier to Discriminate Mediterranean Vegetation with a Detailed Legend Using Multitemporal Series of Images. In: 2006 IEEE International Geoscience and Remote Sensing Symposium And 27th Canadian Symposium on Remote Sensing. Denver. pp: 192-195.

Pons X, Ninyerola M (in press). Mapping a topographic global solar radiation model implemented in a GIS and calibrated with ground data. International Journal of Climate.

Quattrochi DA and Luvall JC (Eds.) (2000) Thermal Remote Sensing in Land Surface Processes. CRC Press. New York.

Rocchini D, Ricotta C (2007) Are landscapes as crisp as we may think?. Ecol. Modelling. 204: 535-539.

Salski A (2007) Fuzzy clustering of fuzzy ecological data. Ecological Informatics 2: 262-269.

StatSoft Inc. (2001) STATISTICA (data analysis software system), version 6. http://www.statsoft.com.

Valor E, Caselles V, Coll C, Sánchez F, Rubio E, Sospedra F (2000) Simulation of a medium-scale-surface-temperature instrument from Thematic Mapper data. Int. J. Remote Sens. 21: 3153-3159.

Ward JH (1963) Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association 58: 236.

Wolock D. M., Winter T. C. and McMahon G. 2004. Delineation and evaluation of hydrologic-landscape regions in the United States using geographic information system tools and multivariate statistical analyses. Environ Manage 34: S71-S88.