Potential Distribution of Forest Species in Dehesas of Extremadura (Spain)

Ángel M. Felicísimo, Alicia Gómez and Jesús Muñoz

Abstract

Dehesas are artificial ecotypes derived from original forest clearing. Continuous forest cover disappeared centuries ago, and currently only dispersed patches remain. In some places deforestation was complete and not even the most open dehesas remain. Based on the current presence of forest species, the objectives of this research are:

- to generate potential distribution and suitability maps for each one of the three species of *Quercus* forming dehesas in Extremadura.
- to generate a potential distribution model of forests by combining suitability and distance models for the same three *Quercus* species that can be used as a guide for forest management.

Predictor variables used to generate the models include topographic, climatic and current vegetation data. Our work describes the construction of the statistical models and the evaluation of the results. A potential vegetation map was generated. Both suitability and synthesis maps could be used as a guide for forestry planning and management.

Keywords: Dehesas, distribution map, forest distribution model, predictive modeling.

Introduction

The aim of forestry planning is to resolve a set of problems that may include the conservation and restoration of the forests. In an area such as the Iberian Peninsula, where forests have been progressively eliminated over centuries, the main aims of forestry planning are the reduction of forest fragmentation, the conservation of diversity, and the restoration of degraded biotopes. To achieve these goals requires basic territorial information of high quality (Lund and Iremonger, 2000), including current vegetation distribution and a set of climate and substrate data integrated into a geographic information system (GIS). This information can help generate various models to facilitate objective decision-making.

In this paper we developed a set of likelihood or suitability models for the presence of forest types that are widely distributed in the Extremadura region of Spain. The suitability is expressed as a number, with zero being "incompatible" and 1 indicating "ideal". The number is derived from a set of physical and

ISBN 3-923381-50-6, US ISBN 1-59326-245-0

biological factors that favor or limit the growth of each type of forest. Knowing the suitability value of each point of the territory, decisions on land use can be made using objective criteria.

The set of suitability values for a territory constitutes a potential distribution model if presented as a map. This model is the reflection of the relationships between presence/absence of each forest type and the values of the potentially influential environmental variables for a given territory. Usually the current area of distribution is significantly smaller than the potential distribution area, since forests have been artificially eliminated from areas where they formerly grew. Potential distribution models allow recognizing and delimitating such former distribution areas in order to direct current management plans.

A previous hypothesis in the process of generating potential distribution models is that current forest distribution is a sufficient sample of its potential distribution. If a given forest type has only a residual presence, it would be impossible either to construct a representative sample or to generate an adequate model. Likewise, if the current area is strongly biased with respect to the environmental variables, the potential distribution model will only partially reflect the real potential area.

Objectives

The goal was to generate a set of suitability cartographic models for *Quercus* species in Extremadura. By combining the models, the potential distribution area of *Quercus* forests could be determined. These models constitute a tool for delimiting the most suitable zones for environmental restoration. The dominant forest species are of the genus *Quercus*, the most important in extension being *Quercus rotundifolia* Lam. (holm oak, 12 680 km^2, synonym: *Quercus ilex* L. ssp. *ballota* (Desf.) Samp.), *Quercus suber* L. (cork oak, 2130 km^2) and *Quercus pyrenaica* Wild. (Pyrenean oak, 950 km^2).

Study area

Extremadura is one of the 17 Autonomous Communities of Spain, covering 41,680 km^2 (Fig. 1). It has a Mediterranean climate, somewhat softened by the relative proximity to the sea and the entrance of oceanic fronts from the west.

Data sets

The dependent variable: *Quercus* distribution

Current *Quercus* species distribution maps derive from the "Mapa Forestal de España" (MFE, Forestry Map of Spain). MFE was elaborated by the Spanish General Directorate for Nature Conservation (Dirección General de Conservación de la Naturaleza, formerly ICONA) during the period 1986-96. The digital version includes a database with 72 information fields. This associated attribute table is complex and presents some problems of interpretation. For the present

work, only information necessary to identify the main vegetation classes was used and a hierarchic legend was created. This legend allows one to handle the vegetation classes on many levels of thematic resolution adapted to the information available in the source map. Figure 2 shows the current distribution of the *Quercus* species. Sampling was performed on this digital map layer using GIS.

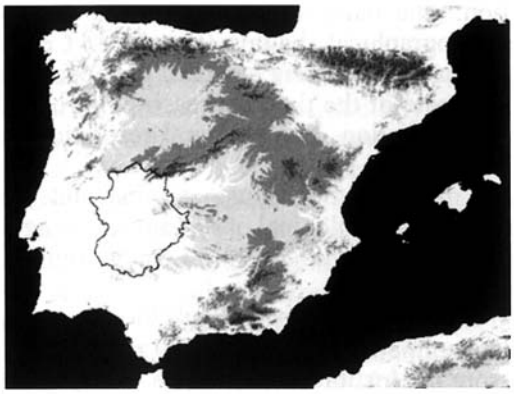

Fig. 1: Location of Extremadura in the Iberian Peninsula.

Fig. 2: Current distribution of Quercus species in Extremadura (black, Q. pyrenaica, dark grey, Q. suber, and pale grey, Q. rotundifolia).

The independent variables: digital terrain models

Five digital terrain models representing independent variables were generated from the topographical map:

- Elevation: A digital elevation model (DEM) was constructed using Delaunay's triangulation algorithm followed by transformation to a regular 100 m cell size grid.
- Potential insolation: The models derivate from a simulation based on the DEM, analyzing topographical shading (Fernández Cepedal and Felicísimo, 1987) for the sun's path at different standard date periods (Heywood, 1964). The result is an estimate of the time that each point of the terrain is directly illuminated by solar radiation. The temporal resolution was 20 min and the spatial resolution 100 m.
- Mean temperature maps of the annual maxima and minima: These were interpolated from 140 meteorological observatories using the thin-plate spline method (Hutchinson and Dowling, 1991, Lennon and Turner, 1995) with a spatial resolution of 500 m.
- Quarterly rainfall maps: Interpolated from 276 meteorological observatories using the thin-plate spline method with a 500 m spatial resolution. Meteorological data are from the Instituto Nacional de Meteorología of Spain and the length of observation period is 40 years.

The above variables were chosen because they have potential influence on the distribution of the vegetation and enough data to generate GIS digital layers. Insufficient data eliminated other variables (e.g., soils) commonly used in ecological modeling.

Methods

Statistical methods

Methods used in ecological modeling can be considered as two main types: global parametric and local non-parametric. Global parametric models use a strategy of global variable selection. Each variable enters within the model as "a whole" to explain its contribution on the response. An advantage of global parametric models, such as linear and logistic regression, is that they are easy and quick to compute, and their implementation within GIS is straightforward. As an example of global parametric model we have used Logistic Multiple Regression (LMR).

Although widely used (Felicísimo et al., 2002), LMR has several important limitations. Ecologists frequently assume a response function, which is unimodal and symmetric and usually unsupported (Austin and Smith, 1989, Yee and Mitchell, 1991).

The alternative hypothesis is that in modeling organism/community distribution, the response is related to predictor variables in a non-linear and local mode. Local non-parametric models are suitable under such a hypothesis as they use a strategy of local variable selection and reduction, and are flexible enough to allow non-linear relationships. Multivariate Adaptive Regression Splines (MARS) and Classification and Regression Trees (CART) are examples.

In the present work all models were calculated from a stratified random sample of each zone of presence/absence for each *Quercus* species. Each random sample was about 10-20% of the total area for each forest type. The first sample was used to generate the model, and the second to test model reliability and independence.

Logistic Multiple Regression

Logistic multiple regression (LMR) has been used to generate likelihood models for forecasting in a variety of fields. It is particularly appropriate for our present purpose because the dependent variable is dichotomous (presence/absence), and admits independent variables with non-Gaussian distributions. LMR results take values between 0 and 1, and therefore the method is suitable to generate a likelihood model with probabilities as output (Jongman et al., 1995).

The introduction of the spatial component into LMR to generate cartographic models is recent. It is now usually integrated into the GIS as a development tool.Guisan, et al.(1998 used LMR in the GIS ArcInfo (ESRI Inc.) to generate a distribution model for the plant *Carex curvula* in the Swiss Alps. A similar study was applied to aquatic vegetation by (Van de Rijt et al., 1996) using the GIS Grass (US Army Construction Engineering Research Laboratory).

In the logistic model, the probability of presence, $P(i)$, takes the form: $P(i)=1/(1+\exp -(b(0)+b(1).x(1)+\ldots+b(n).x(n))$; where $P(i)$ is the probability of presence of the forest, $x(1) \ldots x(n)$ represents the values of the independent variables, and $b(1)\ldots b(n)$ the coefficients. The results for each point of the terrain vary between the extremes of 0 (incompatible) and 1 (ideal).

LMR was performed with SPSS 11.5 with the forward conditional stepwise method (P-to-enter=0.05 and P-to-remove=0.1).

CART, Classification and Regression Trees

CART (Breiman et al., 1984) is a rule-based method that generates a binary tree through 'binary recursive partitioning', a process that splits a node based on yes/no answers about the values of the predictors. Each split is based on a single variable, and while some variables are used one to many times in the model, others may not be used at all. The rule generated at each step minimizes the variability within each of the two resulting subsets. Each subset is split further based on different relationships. CART builds an overgrown tree based on a node purity criterion that later is pruned back to avoid over-fitting via cross-validation.

The main drawback of CART models when used to predict organism distributions is that the generated models can be extremely complex and difficult to interpret. For example, work on Australian forests by Moore et al. (1991) produced a tree with 510 nodes for just 10 predictors. In the present study, the optimal tree we obtained with the *Quercus rotundifolia* data set has 4840 terminal nodes. Although the complexity of such a tree does not diminish its predictive power, it makes the tree impossible to interpret, whereas in many

studies interpretability is a key issue. Moreover, implementation of such a tree within GIS is very difficult. Nevertheless, a contribution of our work has been to elaborate a method to translate the large CART reports (text files) to AML (Arc Macro Language) files, legible with ArcInfo GIS. As an example, the file with the CART decision rules for constructing the *Q. rotundifolia* suitability map is a 1.8 Mb AML. Our method has no limitations although the execution time can be long (about 55 h for the *Q. rotundifolia* model).

MARS, Multivariate Adaptive Regression Splines

MARS is a relatively novel technique that combines classical linear regression, mathematical construction of splines, and binary recursive partitioning to produce a local model where relationships between response and predictors are either linear or non-linear (Friedman, 1991). To do this, MARS approximates the underlying function through a set of adaptive piecewise linear regressions termed 'basis functions'. For example, the first four basis functions from the *Q. pyrenaica* model are:

BF1 = MAX (0, PT4 - 3431)
BF2 = MAX (0, 3431 - PT4)
BF3 = MAX (0, MDE50 - 1181)
BF4 = MAX (0, 1181 - MDE50)
PT4: mean rainfall of the period October-December
MDE50: elevation

Changes in slope of those *basis functions* occur at points called *knots* (the values 3431 or 1181, in the above examples). The regression line is thus allowed to bend at the knots, which mark the end of one region of data and the beginning of another with different behavior of the function. Like splits in CART, knots are established in a forward/backward stepwise way. A model which clearly overfits is produced first, and then those knots that contribute least to the efficiency of the model are discarded in the backwards pruning step to avoid overfitting. The best model is selected via cross-validation, a process that applies a penalty to each term – namely, a knot – added to the model to keep low complexity values.

As in the CART case, we transformed the text report files into AML files and generated the suitability models using ArcInfo GIS.

Evaluation of the models

To check model performance we used the area under the Receiver Operating Characteristic (ROC) curve, commonly termed AUC (Hanley and McNeil, 1982) and the specific statistics "cross validated relative cost" for CART and GCV "generalized cross validation" for MARS. The Receiver Operating Characteristic curve is recommended for comparing two-class classifiers, as it does not merely summarize performance at a single arbitrarily selected decision threshold, but across all possible decision thresholds (Fielding and Bell, 1997; Ivers et al., 2001). AUC is a measure of model accuracy and is equivalent to the normalized

Mann-Whitney 2-sample statistic, which makes it equivalent to the Wilcoxon statistic.

The specific statistics used by CART and MARS are based on cross-validation methods along with a complexity penalty that looks for the most harmonious solution. AUC was also calculated for the best option in the MARS case, but it has not been possible with CART.

Results

LMR

We generated a logistic model for each of the three *Quercus* formations present in the study zone. Coefficients and the values of the area under the ROC curve are listed in Table 1.

Table 1: Regression coefficients and area under the ROC curve (AUC) for the three Quercus formations.

	MODEL		
Var	QPYR	QROT	QSUB
constant	-8.223	-1.031	1.614
elevation	0.003	-0.002	-
slope	-0.005	0.021	0.050
insol (-23)	-0.053	0.037	-0.048
insol (-12)	0.076	-0.098	-
insol (0)	-0.064	0.107	-
rainfall (1)	-0.003	0.002	0.004
rainfall (2)	-0.002	-0.002	-0.009
rainfall (3)	0.004	-0.002	-0.006
rainfall (4)	0.005	-0.001	0.002
min temp	-0.061	0.032	0.100
max temp	0.080	-0.023	-0.052
AUC	0.924	0.627	0.790

Only QPYR (*Q. pyrenaica*) and QSUB (*Q. suber*) models show significant fit with the data.

MARS

MARS consistently achieves better results than logistic regression. The AUC values are 0.969 (*Q. pyrenaica*), 0.910 (*Q. suber*), and 0.783 (*Q. rotundifolia*). The improvement is more evident in the most difficult models, indicating better adaptability of this method.

The best model for *Q. pyrenaica* consists of 45 basis functions. It has a global correct prediction of 0.933 for a threshold value of 0.5 with a specificity of 0.990 and a sensitivity of 0.875. Not all the variables enter in the final equation: the most relevant is the elevation, followed by quarterly mean rainfall and annual

maxima and minima temperatures. Slope and potential insolation were not significant. Values for remainder models are:

Quercus suber best model: 45 basis functions; correct prediction: 0.837 (threshold value = 0.5), specificity: 0.910, sensitivity: 0.765. Most relevant variables are, in decreasing order, quarterly mean rainfall, elevation, annual maxima and minima temperatures, winter potential insolation and slope.

Quercus rotundifolia best model: 40 basis functions; correct prediction: 0.712 (threshold value = 0.5), specificity: 0.778, sensitivity: 0.645. Most relevant variables are, in decreasing order, quarterly mean rainfall, elevation, annual maxima and minima temperatures and slope. The length of basis functions does not allow presentation of the complete models. A sample is shown in Table 2.

Table 2: A sample of the MARS basis functions and final equation (Quercus suber best model).

No.	BASIS FUNCTION
1	max (0, PT4 - 2238.0)
2	max (0, 2238.0 - PT4)
3	max (0, PT3 - 381.0)
4	max (0, 381.000 - PT3)
5	max (0, PT4 - 3020.000) * BF3
...	...
45	max (0, MDE - 115.0) * BF30

EQUATION

Y = 2.075 - .663801E-03 * BF1 - 0.002 * BF2 + .638964E-03 * BF3 - 0.013 * BF4 + .279240E-05 * BF5 - ... - .102701E-08 * BF42 + .489292E-06 * BF43 + .172231E-07 * BF44 - .236239E-07 * BF45

CART

CART results are difficult to compare with LMR and MARS because there are no data to calculate the AUC statistic. Nevertheless, the cross-validation tests show improvements in all the cases for a threshold value of 0.5. The results were:

Quercus pyrenaica: global correct prediction: 0.976 (specificity: 0.987, sensitivity: 0.966), 817 terminal nodes tree (Gini construction rule). Variable importance in decreasing order: quarterly mean rainfall, elevation, slope, annual maxima and minima temperatures and potential insolation.

Quercus suber: global correct prediction: 0.956 (specificity: 0.966, sensitivity: 0.946), 2353 terminal nodes tree (Gini symmetric construction rule). Variables: quarterly mean rainfall, annual maxima and minima temperatures, elevation, slope, and potential insolation.

Quercus rotundifolia: global correct prediction: 0.858 (specificity: 0.867, sensitivity: 0.850), 4889 terminal nodes tree (Gini construction rule). Variables: quarterly mean rainfall, elevation, annual maxima and minima temperatures and potential insolation.

CART classification rules are much more complex than the MARS functions or LMR equations. Each terminal node is determined by a logical expression where every variable may be present. Table 3 shows an example.

Table 3: An example of Quercus rotundifolia CART classification rules. The tree has 4889 terminal nodes that must be calculated testing the logical expressions within the GIS.

NODE	CLASSIFICATION RULE
1	IF (PT4 <= 1966.5 & MDI-12 <= 30.5 & T_JULY <= 34.5 & PT1 <= 1634.5 & T_JANUARY <= 18.5 & MDP <= 9.5) THEN P = 0.000
2	IF (PT4 <= 1966.5 & MDI-12 <= 30.5 & T_JULY <= 34.5 & PT1 <= 1634.5 & T_JANUARY <= 18.5 & MDP > 9.5) THEN P = 1.000
...	...
4889	IF (PT4 > 1966.5 & PT1 > 3260.5 & T_JULIO > 32.5 & PT2 > 1969 & PT3 > 677.5 & MDE50 > 503) THEN P = 0.003

Suitability models

All LMR equations, MARS basis functions and CART classification rules have been translated to AML syntax. ArcInfo GIS was used to create the suitability models as maps. The statistics confirm the improved results obtained with local non-parametric methods. However, CART shows a clear overfitting and a conspicuous discretization effect due to binary rules. Likewise, due to weight of the climate variables, the suitability models frequently replicate the isopleths shape, making such models spatially less convincing. For the above reasons, we accepted MARS models to perform the statistics shown in the Table 4.

Table 4: Area below suitability classes according to MARS models.

	Quercus		
	pyrenaica	*suber*	*rotundifolia*
current area (km^2)	945	2127	12679
suitability areas			
incompatible	10180	1755	3317
unfavorable	16550	6117	5511
indifferent	6435	10323	8161
favorable	3008	9089	9521
very favorable	2625	8770	10929
suitable	2722	5436	4031

Compared with the current distribution area (945 km^2), we can observe that *Q. pyrenaica* has a great potential: 5347 km^2 of very favorable or suitable areas. *Quercus suber* shows a similar pattern: 14,206 km^2 very favorable or suitable area against the current 2127 km^2. On the other hand, *Q. rotundifolia* shows a potential close to the present distribution: 14,960 km^2 of very favorable or suitable areas against the current 12,679 km^2. Figures 3, 4 and 5 show the suitability models classified into six classes for greater legibility from white to black (1: incompatible, 2: unfavorable, 3: indifferent, 4: favorable, 5: very favorable, 6: suitable). All models were constructed from MARS basis functions.

Fig. 3: Quercus pyrenaica suitability model (MARS)

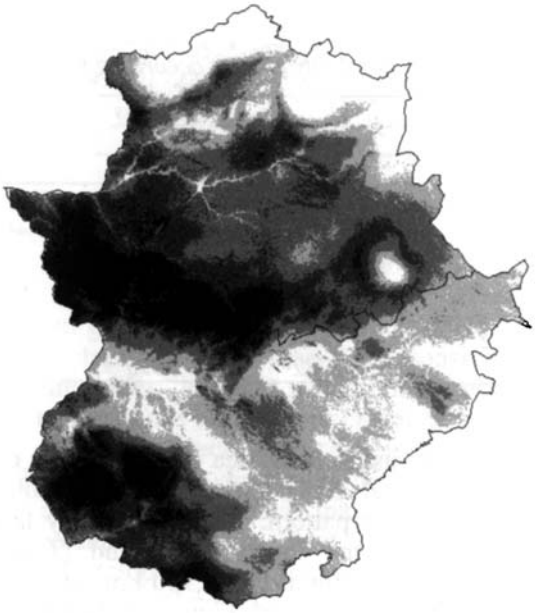

Fig. 4: Quercus suber suitability model (MARS)

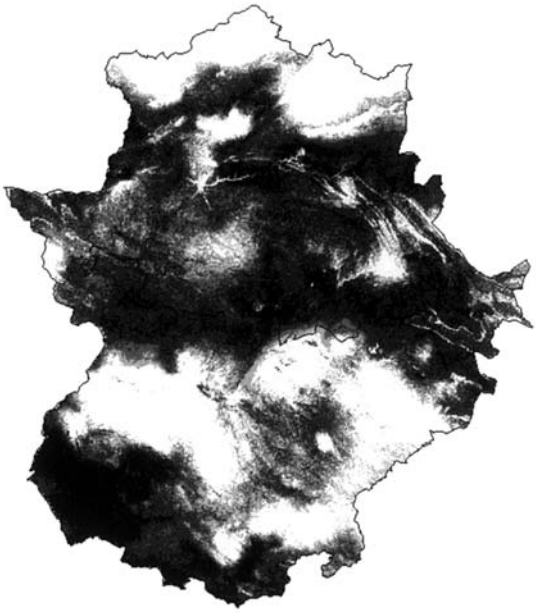

Fig. 5: Quercus rotundifolia suitability model (MARS)

Potential areas and current vegetation

We superimposed the suitability models over the MFE to determine current vegetation on each suitability value of *Quercus* species. Such analysis is similar to the work by Higgins et al. (1999), who estimated diversity loss resulting from invasive plants expansion. This comparison between suitability and present vegetation maps allows one to understand some aspects of vegetation dynamics and facilitates further planning proposals.

As an example, we will show some results obtained with *Quercus pyrenaica*. Some 450 km^2 of very favorable area for this species are presently covered with *Q. suber* or *Q. rotundifolia*, species typical of other vegetation series. Figure 6 shows frequency histograms for both in the *Q. pyrenaica* distribution area. *Quercus pyrenaica* suitability values have been grouped into 10 classes. Both histograms clearly show that the environmental requirements of both *Quercus suber* and *Q. rotundifolia* differ from those of *Q. pyrenaica*, since their presence on the areas of highest suitability is small (110 km^2 for *Q. suber* and 340 km^2 for *Q. rotundifolia*). These areas could be considered as transition areas with mixed formations.

In contrast, Fig. 7 shows the results corresponding to two plantations species: *Castanea sativa* (sweet chestnut) and *Pinus pinaster* (maritime pine). From the histograms it is clear that both grow on areas of high suitability for *Q. pyrenaica*. Other species not shown, *Pinus pinea* L. and *Eucaliptus camaldulensis* Dehnh, have different preferences and grow in areas with low suitability for *Q. pyrenaica*.

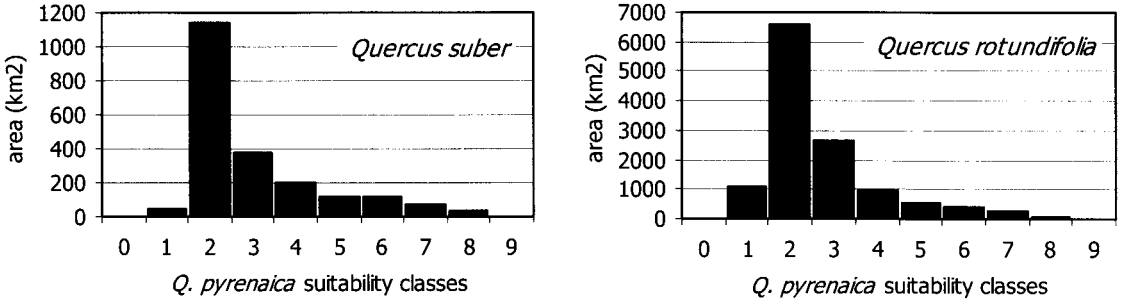

Fig. 6: Presence of Q. suber and Q. rotundifolia in Q. pyrenaica suitability classes. Both species almost disappear from areas with high suitability for Q. pyrenaica. Suitability classes have equal intervals: 0 (0.00-0.10), 1 (0.11-0.20)... and 9 (0.91-1.00).

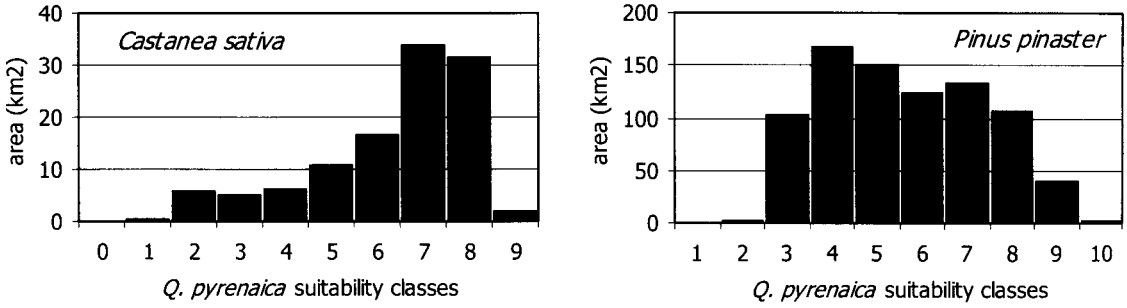

Fig. 7: Presence of C. sativa and P. pinaster in Q. pyrenaica suitability classes. Both species have a significant presence on areas of high suitability for Q. pyrenaica.

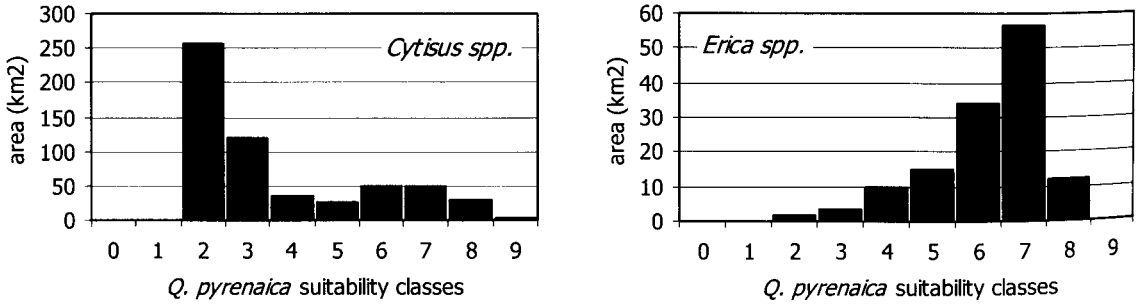

Fig. 8: Presence of Cytisus and Erica in Q. pyrenaica suitability classes. Both shrubs must be considered as substitution stages of the primary forests.

Finally, Fig. 8 shows the presence of shrublands dominated by *Cytisus* spp. (brooms) and *Erica* spp. (heather) in classes of high suitability for *Q. pyrenaica*.

This shrublands should be considered as substitution stages in which primary forests were cleared. When soil degradation is moderate, *Cytisus* formations dominate, while *Erica* formations commonly cover more degraded soils.

Table 5 shows the more widespread MFE formations in areas with *Q. pyrenaica* suitability values higher than 0.70 (rock excluded). Two new formations appear: gum rock-rose (*Cistus ladanifer* L.) formations and croplands. The first must be interpreted just as shrubland and heath: a substitution stage on partially or totally deforested lands. Croplands are mainly pastures, meadows and cherry trees.

Table 5: Classes covering larger extensions of Q. pyrenaica high suitability areas.

MFE formations	area (km^2)
Q. pyrenaica	860
Cropland	774
Q. rotundifolia	340
Mixed shrubland	259
Pinus pinaster	146
Cistus ladanifer	120
Quercus suber	115
Castanea sativa	68
Heath	53

These results permit preliminary guidelines to be developed for land management. As an example, we can define three main action guidelines:

Type 1 Zones, currently *Q. pyrenaica* and *Castanea sativa* woodlands, where conservation is a priority.

Type 2 Zones, currently covered by *Cytisus* and *Cistus* shrublands where soil degradation is moderate, and *Erica* spp. heaths where soil degradation is more severe. Actions should reduce forest fragmentation through reforestation of these secondary formations (in the sense of ecological succession) with *Q. pyrenaica*.

Type 3 Zones, where exotic species, mainly *Pinus* plantations, will be progressively replaced with *Q. pyrenaica*.

Priority actions should be taken, both for Type Zones 2 and 3, in areas with high suitability values and proximity to Type 1 Zones, which facilitate seed arrival and, consequently, rapid regeneration.

Potential vegetation model

In the previous process we generated three suitability models, one for each type of potential forest present in the zone. To generate the final map (Fig. 9), we defined a function whereby, for each cell, the species maximum suitability values were corrected by the inverse of the distance to the closest cell where such species grow. This correction must be considered as an indicator of colonization likelihood. The result of this evaluation process is a model showing, for each cell, the type of forest with the highest potential value after considering colonization processes, and may be interpreted as a more sound potential vegetation model (Figure 9).

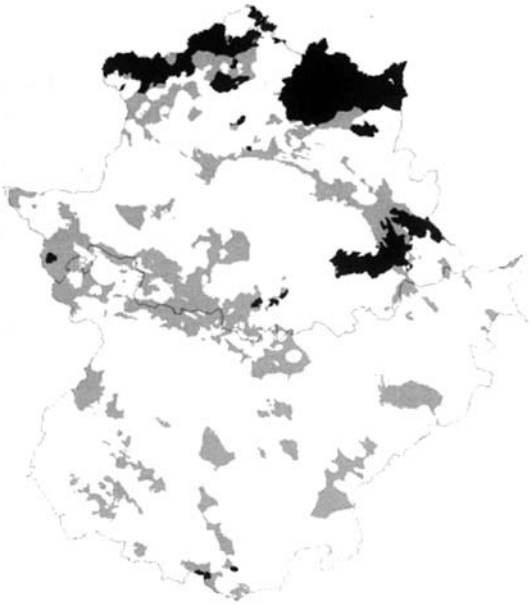

Fig: 9: Potential distribution model of Quercus in Extremadura; Q. pyrenaica (black), Q. suber (dark grey), Q. rotundifolia (pale grey).

Discussion

Traditionally, environmental factors have been considered as the main forces driving vegetation distribution. First attempts to clarify the relationships between environment and vegetation were analytical, exploring the association between species and one variable at a time (Gauthier et al., 1977) Later, computers allowed generalization of multivariate analyses (factorial analysis, classification, etc). Suitability models synthesize the response of the species to a set of environmental variables affecting its spatial distribution. Such models are based on real data (*data driven*) and statistical techniques, which make them objective methods. These methods have advantages and drawbacks. A limitation is that suitability models can only be applied in the same climatic and spatial scenario, although several authors have made predictions under climatic change scenarios (Peterson et al., 2002). Another problem is related to the sampling process: suitability models are credible only if the source samples represent vegetation classes with a good conservation state. Marginal or relictic formations may not be included in the analysis.

In our experience those methods give good results in mountain areas because limiting factors are mainly physical: elevation, potential insolation, slope, etc. These factors are generally known (elevation), or they can be derived with enough accuracy (potential insolation).

Quercus pyrenaica, *Q. suber,* and *Q. rotundifolia* are widespread species in the Iberian Peninsula. The generated suitability models reflect rather precisely their current spatial distribution, and they predict a potential area distribution of

reasonable extension and characteristics from a biological point of view. The models are static and do not integrate historical or biotic factors. The limitations recommend a cautions interpretation: the potential vegetation model is a synthesis of suitability models, but it does not have to be interpreted like a map of climax formations in a biological sense.

Transferring models into actions is an issue for which complementary information is needed, especially on soil property and its current management. To deal with this aspect is beyond the aim of this paper, and the proposed actions should be considered only as examples. It must be pointed out, however, that the generated maps and statistics reflect objective data of obvious utility. Combinations of model maps with current land use and landscape management could successfully complete an objective decision-making system extremely useful in territorial planning.

Potential vegetation maps are theoretical constructions that may be interpreted as biologically reasonable hypotheses, based on sets of evidence and underpinned by some of the prevailing theoretical frameworks in ecology. Understood as hypotheses, however, they can not be readily subjected to experimental tests, and in this sense lack one of the properties that are usually required in the scientific method: that of being refutable or, vice versa, verifiable.

Despite the above considerations, potential vegetation maps represent a useful tool for environmental management since they synthesize different types of knowledge about the reality of a given territory that are difficult to integrate in any other way. Until recently, potential vegetation maps were prepared mostly by subjective methods, usually by a "specialist". The problems with this approach are obvious: maps will vary in quality according to the "ability" of the expert, and especially they will not be repeatable since there is no an explicit method of explaining them (an algorithm). In contrast, we have proposed here a method based on robust statistical operations and objective cartographical information. There exists an explicit procedure to reach the final result, and the entire flow of information is accessible.

Acknowledgements

This paper is part of the Project 2PR01C023 co-funded by the Junta de Extremadura (Consejería de Educación, Ciencia y Tecnología - II Plan Regional de Investigación, Desarrollo Tecnológico e Innovación de Extremadura) and FEDER (Fondo Europeo de Desarrollo Regional).

References

Austin, M.P. and Smith, T.M. (1989): A new model for the continuum concept. Vegetatio **83**: 35-47.

Breiman, L., Friedman, F., Olshen, R. and Stone, C. (1984): Classification and Regression Trees. Wadsworth, Belmont, CA, 358 pp.

Felicísimo, A.M., Francés, E., Fernández, J.M., González-Díez, A. and Varas, J. (2002):

Modeling the potential distribution of forests with a GIS. Photogrammetric Engineering & Remote Sensing **68**: 455-461.
Fernández Cepedal, G. and Felicísimo, A.M. (1987): Método de cálculo de la radiación solar incidente en áreas con apantallamiento topográfico. Revista de Biología de la Universidad de Oviedo **5**: 109-119.
Fielding, A.H. and Bell, J.F. (1997): A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental Conservation **24**: 38-49.
Friedman, J.H. (1991): Multivariate adaptive regression splines. Annals of Statistics **19**: 1-141.
Gauthier, B., Godron, M., Hiernaux, P. and Lepart, J. (1977): Un type complémentaire de profil écologique : le profil écologique indicé. canadian Journal of Botany **55**: 2859-2865.
Guisan, A., Theurillat, J.-P. and Kienast, F. (1998): Predicting the potential distribution of plant species in an alpine environment. Journal of Vegetation Science **9**: 65-74.
Hanley, J.A. and McNeil, B.J. (1982): The meaning and use of the area under a Receiver Operating Characteristic (ROC) curve. Radiology **143**: 29-36.
Heywood, H. (1964): Standard date periods with declination limits. Nature **204**: 678.
Hutchinson, M.F. and Dowling, T.I. (1991): A continental hydrological assessment of a new grid-based digital elevation model of Australia. Hydrological Processes **5**: 45-58.
Ivers, R.Q., Macaskill, P., Cumming, R.G. and Mitchell, P. (2001): Sensitivity and specificity of tests to detect eye disease in an older population. Ophthalmology **108**: 968-975.
Jongman, R.H.G., Ter Braak, C.J.F. and van Tongeren, O.F.R. (1995): Data analysis in community and landscape ecology. Cambridge University Press, Cambridge, 229 pp.
Lennon, J.J. and Turner, J.R.G. (1995): Predicting the spatial distribution of climate temperature in Great Britain. Journal of Animal Ecology **64**: 370-392.
Lund, H.G. and Iremonger, S. (2000): Omissions, commissions, and decisions: the need for integrated resource assessments. Forest Ecology and Management **128**: 3-10.
Moore, D.M., Lee, B.G. and Davey, S.M. (1991): A new method for predicting vegetaion distributions using decision tree analysis in a geographic information system. Environmental Management **15**: 59-71.
Peterson, A.T., Ortega-Huerta, M.A., Bartley, J., Sánchez-Cordero, V., Soberón, J., Buddemeier, R.H. and Stockwell, D.R.B. (2002): Future projections for Mexican faunas under global climate change scenarios. Nature **416**: 626-629.
Van de Rijt, C.W.C.J., Hazelhoff, L. and Blom, C.W.P.M. (1996): Vegetation zonation in a former tidal area: A vegetation-type response model based on DCA and logistic regression using GIS. Journal of Vegetation Science **7**: 5005-518.
Yee, T.W. and Mitchell, N.D. (1991): Generalized additive models in plant ecology. Journal of Vegetation Science **2**: 587-602.

Addresses of authors:
Ángel M. Felicísimo
Alicia Gómez
Ingeniería Cartográfica, Geodesia y Fotogrametría, Universidad de Extremadura
10071 Cáceres, Spain
amfeli@unex.es
Jesús Muñoz
Real Jardín Botánico, CSIC
Plaza de Murillo 2
28014 Madrid, Spain
jmunoz@ma-rjb.csic.es