Species distribution models applied to plant species selection in forest restoration: are model predictions comparable to expert opinion?

Aitor Gastón*, Juan I. García-Viñas, Alfredo J. Bravo-Fernández, César López-Leiva, Juan A. Oliet, Sonia Roig & Rafael Serrada

EGOGESFOR Research Group, Universidad Politécnica de Madrid, Escuela Técnica Superior de Ingeniería de Montes, Forestal y del Medio Natural, Ciudad Universitaria s/n, 28040, Madrid, Spain.

* Email, aitor.gaston@upm.es

Abstract

An expert on local flora usually is the best option for plant species selection in most ecological restoration projects; although species selection often needs to be dealt with swiftly as well as on a limited budget, and obtaining the opinion of a local expert may not always be an economically viable alternative. In such cases, species distribution models (SDM) may offer a faster and more cost effective alternative. We asked six experts to rank native tree species according to their suitability at 24 forest sites. The predictive performance of the suitability rankings was evaluated by assessing their ability to discriminate present from absent species in the observed tree assemblages at each evaluation site. We used the area under the receiver operating characteristic curve to calculate the probability that the estimated suitability for a species present at a particular evaluation site is greater than the estimate for an absent species (both picked at random). Suitability rankings were also obtained from the predictions of SDM and the same procedure was used to estimate the predictive performance of the set of models at each site. The experts offered concordant suitability rankings at almost every evaluation site. There were no significant differences in the predictive performance of the SDM and four of the experts, although the SDM performed slightly better than the other two experts. Our results point to the suitability of the proposed species distribution modeling approach to obtain fast and cost effective recommendations for species selection in forest restoration projects.

Keywords: ecological niche modeling; forest restoration; tree species selection guidelines

1. **Introduction**

Ecological restoration is an intentional activity that initiates or accelerates the recovery of a degraded ecosystem with respect to its health, integrity and sustainability (SER 2004). Although passive restoration (i.e., secondary succession after the disappearance or removal of the stresses that caused degradation) may be appropriate in some cases, more active approaches have been found to be more successful in highly degraded ecosystems (McIver & Starr 2001). Active restoration is especially important in degraded Mediterranean ecosystems due to extreme climatic conditions and poor soil fertility (Cuesta et al. 2012). Active restoration of forests frequently involves planting or seeding (Clewell et al. 2005), therefore, appropriate species selection is a central issue in such projects. Although many different criteria may be used in species selection (e.g., late-successional vs. mid-successional species, Padilla et al. 2009; adaptation to wildfires, Peman Garcia et al. 2008), the selected species should always match local environmental conditions (Peman Garcia et al. 2008). Using species that fulfill other important selection criteria for the purpose of the forest restoration project, but which are not adapted to local environmental conditions, would lead to failure of the restoration project.

An expert on local flora usually constitutes the best option for plant species selection in most ecological restoration projects (Ruiz de la Torre et al. 1990); although species selection often needs to be dealt with swiftly as well as on a limited budget, and obtaining the opinion of a local expert may not always be an economically viable alternative. In such cases, decision support tools may be very useful to forestry practitioners. A simple approach is to build a database of known distribution limits of species along environmental gradients and then to identify which species are suitable for the site to be restored by comparing local environmental conditions to the limiting values (e.g., Webb et al. 1980; Booth & Jones 1998). This approach offers binary predictions (i.e. the species are either suitable or not), therefore, it is not possible to rank species according to their suitability for a given site. Ranking species is desirable for species selection in restoration projects and requires continuous habitat suitability estimations. Such estimations make up part of the output of predictive species distribution models (SDM), which relate field observations to environmental predictor variables (Guisan & Thuiller 2005). Early applications of SDM to support species selection were developed as part of forest restoration initiatives, using environmental envelopes to estimate site specific suitability (Gandullo & Sánchez Palomares 1994, chap. XIII). These early applications used presence-only data from previous research on the autoecology of pines and fitted models using an approach similar to BIOCLIM, the first widely used SDM (Busby 1991; Booth et al. 2013). Later, more complex modeling techniques have been used to fit SDM which support species selection in forest restoration, e.g., a convex hull approach (García López & Allúe Camacho 2004) or generalized linear models (Felicísimo 2003).

As with any predictive empirical model, SDM should be validated through an evaluation of the predictive performance prior to use by practitioners (e.g., Fielding & Bell 1997). The comparison of model predictions to the observed distribution of either individual species (e.g., Pearce & Ferrier 2000) or species assemblages (Gastón & García-Viñas 2013) are valid procedures to evaluate predictive performance. When models are used to predict areas for restoration, translocation, or reintroductions, performance evaluations should focus on the reduction of errors of omission (Araújo & Peterson 2012). However, the aim of a restoration project is not always to predict areas for restoration of populations of given species, but to predict the most suitable species for a given site (e.g., restoring a degraded plant community). The number of suitable plant species can easily exceed the maximum number of species that can be introduced for practical reasons. Therefore, when restoring site *i*, practitioners must identify the most suitable n_i species from among a limited number of commercially available species (*N*). This kind of decision requires one SDM for each one of the *N* species in order to predict the occurrence probability for each species at site *i* to be restored. The concern of the practitioner is the ability of the *N* models to rank the *N* species according to habitat suitability at site *i*. In the present study we focus on the latter approach and we use a novel method to evaluate SDM applied to plant species selection in ecological restoration (Gastón & García-Viñas 2013). The evaluation method does not assess the accuracy of individual species models, but focuses on the ability of a set of models to discriminate between present and absent species at a given evaluation site.

Additionally, expert opinion can be used to assess the reliability of model predictions (Thuiller 2003). Although several studies have used expert opinion to fit SDM (e.g., Pearce et al. 2001; Seoane et al. 2005), few references report the use of experts for model evaluation (e.g., Anderson et al. 2003; van Zonneveld et al. 2013). Moreover, SDM evaluation by experts focuses on the performance of individual species models and are usually qualitative (Anderson et al. 2003; but see van Zonneveld et al. 2013 for a quantitative approach to formalize expert knowledge for SDM outcome validation). We are not aware of any existing quantitative expert validation that evaluates the predictions of SDM against observed species assemblages.

In the case of SDM applied to plant species selection for ecological restoration, comparing model predictions to expert opinion may be especially useful. If the modeler wants to recommend a set of models as an alternative when expert opinion is not available, it should be checked whether the models offer similar recommendations to those expected from the experts. The objective of this study is to validate the application of SDM to species selection in forest restoration by comparing model predictions to expert opinion. Our work only deals with a part of the species selection procedure, the identification of the species that can survive and reproduce successfully given the environmental conditions of the restoration site and we do not consider other criteria. We hypothesized that; given that models and experts use the same environmental data, the ability of SDM to discriminate between present and absent species in an independent sample of evaluation sites is at least as good as the ability of a group of experts.

2. Materials and methods

We asked the experts to rank tree species according to their suitability for the environment of a sample of evaluation sites. SDM predictions were used to rank the same species at the evaluation sites. First, we assessed the concordance between the rankings of the experts. If there were discrepancies in the rankings produced by the different experts, then their use as a reference for the model predictions would be compromised. Secondly, we estimated the ability of the experts to discriminate between present and absent species in observed tree assemblages at the sample of evaluation sites and finally we compared the results to the discriminative ability of the models.

2.1. Evaluation sites

The evaluation sites were sampled from the Third Spanish National Forest Inventory (NFI), a systematic grid with over 90,000 plots, each of which was 0.2 ha in size. Only those NFI plots with relatively low degrees of human disturbance were considered. Plots were visible human activity have been recorded (e.g. plantations of exotic species) were first excluded. As intermediate disturbance levels are not always easily detectable in the field, we added two extra criteria to characterize the degree of disturbance: tree cover of more than 75% and six or more native tree species. These highly demanding criteria resulted in a subset of plots with relatively low probability of human disturbance, as it only contains the upper 3% of the plots in a gradient related to disturbance (tree cover and species richness). Anyway, the subset was large enough (*ca.* 3,000 plots) to allow the subsequent sampling procedure. We used a stratified sampling approach to guarantee that the evaluation sites covered most of the environmental conditions of forests in the study area. Based on the CLATERES biogeoclimatic classification of Spain (Elena Rosselló 1997), 12 strata were used for sampling. CLATERES considers both climate and soil gradients to build strata and is specially designed for stratified ecological sampling as it follows the guidelines of Bunce et al. (1983). We randomly chose two plots with low levels of disturbance for each stratum, giving a final sample of 24 evaluation sites (see fig. 1).

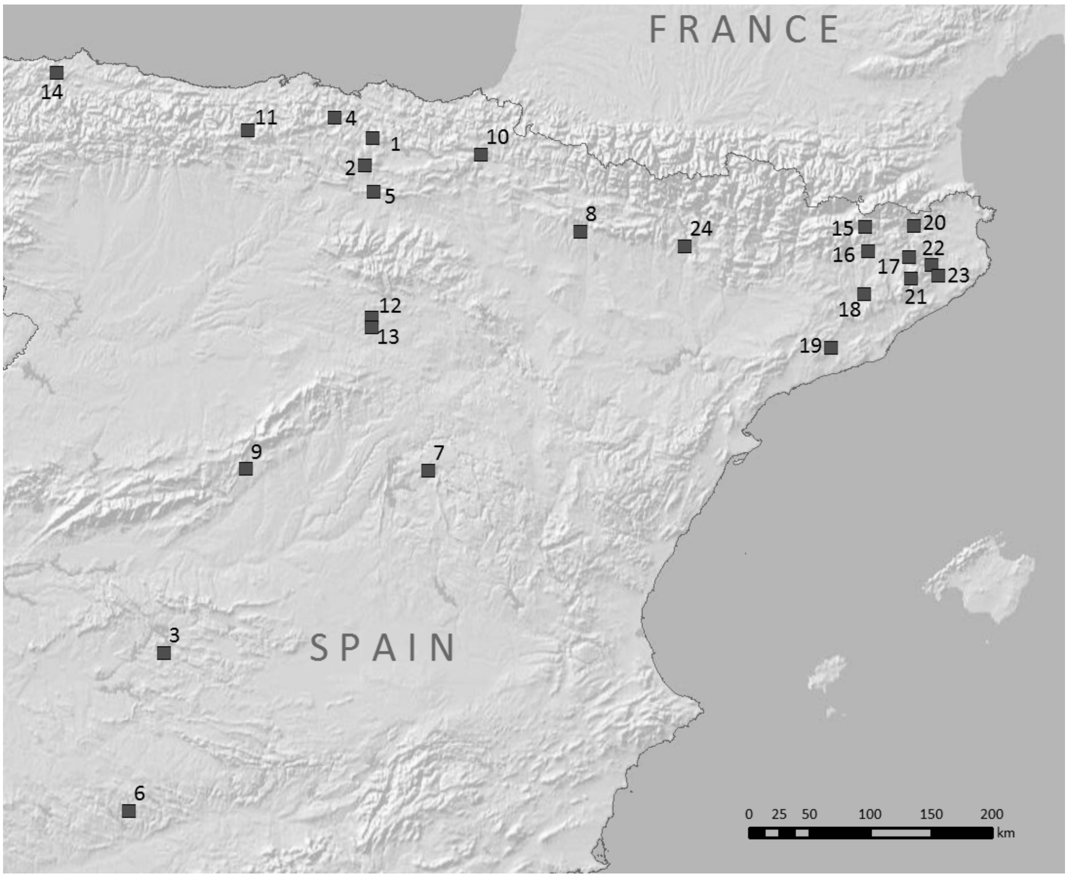

Fig. 1 Location of the 24 evaluation sites, covering most of the environmental conditions of forests in continental Spain (see section 2.1 for more details on sampling procedure). Site codes in the map correspond to those in table 2

Tree species presence or absence data were extracted from the NFI database for each evaluation site. Only native trees recorded at species level were considered. Species that usually occur as shrubs in continental Spain were excluded (e.g., *Salix atrocinerea, Erica scoparia*). Nine species with insufficient occurrence data in the model training dataset (i.e., the Forest Map of Spain, see section 2.3) were dropped from the initial species list, leaving a set of 54 native tree species (see table 1). The habitats of the considered species cover the wide range of environmental conditions that can be found in forest restoration projects across the study area.

Table 1. List of species considered in the study. Only trees native to continental Spain recorded in the National Forest Inventory at species level were considered. Species that usually occur as shrubs in continental Spain and those with insufficient occurrence data in the model training dataset were excluded.

Abies alba	*Corylus avellana*	*Olea europaea*	*Salix alba*	*Ulmus minor*
Abies pinsapo	*Fagus sylvatica*	*Pinus halepensis*	*Quercus canariensis*	*Sorbus aria*
Acer campestre	*Fraxinus angustifolia*	*Pinus nigra*	*Quercus faginea*	*Sorbus aucuparia*
Acer monspessulanum	*Fraxinus excelsior*	*Pinus pinaster*	*Quercus humilis*	*Sorbus domestica*
Acer pseudoplatanus	*Fraxinus ornus*	*Pinus pinea*	*Quercus ilex*	*Sorbus torminalis*
Alnus glutinosa	*Ilex aquifolium*	*Pinus sylvestris*	*Quercus petraea*	*Tamarix canariensis*
Arbutus unedo	*Juniperus communis*	*Pinus uncinata*	*Quercus pyrenaica*	*Taxus baccata*
Betula alba	*Juniperus oxycedrus*	*Populus alba*	*Quercus robur*	*Tilia cordata*
Betula pendula	*Juniperus phoenicea*	*Populus nigra*	*Quercus suber*	*Tilia platyphyllos*
Castanea sativa	*Juniperus thurifera*	*Populus tremula*	*Salix fragilis*	*Ulmus glabra*
Celtis australis	*Malus sylvestris*	*Prunus avium*	*Sambucus nigra*	

Climatic and soil variables were extracted from existing datasets (see section 2.3 and table 2) for each evaluation site and were used by the experts to rank species according to their suitability. The same climatic and soil variables were used to fit the SDM and estimate the probability of occurrence of each species at each evaluation site. The elevation and approximate location (municipality) of the evaluation sites were also available to the experts but were not used to train the models.

Table 2. Location (municipality and province, see fig. 1 for map) and values for the environmental variables at the evaluation sites. Z, elevation (m.a.s.l.); T, mean annual temperature (ºC); Tw, mean maximum temperature of the warmest month (ºC); Tc, mean minimum temperature of the coldest month (ºC); R, mean annual rainfall (mm); SuR, mean summer rainfall (mm); DSL, mean dry season length (months).

Site code	Municipality	Z	T	Tw	Tc	R	SuR	DSL	Lithology	FAO soil group
1	Zuia (Álava)	441	12.3	25.8	1.8	1382	177	0	Siliceous	Cambisol
2	Ribera Alta (Álava)	785	10.1	24.6	-0.3	953	137	0.15	Calcareous	Rendzina
3	Fuenlabrada de los Montes (Badajoz)	690	14.6	33.3	1.3	616	48	3.5	Siliceous	Regosol
4	Valle de Mena (Burgos)	670	11	24.3	0.8	1575	228	0	Siliceous	Cambisol
5	Miranda de Ebro (Burgos)	517	11.8	27	0.9	574	115	2.3	Calcareous	Lithosol
6	Villaviciosa de Córdoba (Córdoba)	316	16.7	35.3	3.5	611	29	4.1	Siliceous	Regosol
7	Alcantud (Cuenca)	987	11.6	30.7	-1.8	658	87	2.3	Calcareous	Cambisol
8	Agüero (Huesca)	801	11.3	28.7	-1.2	711	145	0.7	Calcareous	Cambisol
9	El Escorial (Madrid)	911	12.2	30.1	-0.2	734	69	2.7	Siliceous	Regosol
10	Arakil (Navarra)	473	11.5	26.1	0.4	1157	157	0	Calcareous	Rendzina
11	Hermandad de Campoo de Suso (Cantabria)	1121	8.9	22.5	-1.8	1803	250	0	Siliceous	Cambisol
12	Torreblacos (Soria)	1057	9.8	27.6	-2.4	631	105	1.8	Siliceous	Cambisol
13	Valdenebro (Soria)	1036	10	27.8	-2.2	601	101	1.9	Siliceous	Cambisol
14	Candamo (Asturias)	58	13.5	23.5	4.5	970	144	0.4	Siliceous	Luvisol
15	Castellar de n'Hug (Barcelona)	1310	8.8	23.3	-2.8	1003	323	0	Calcareous	Rendzina
16	Lluçà (Barcelona)	827	11.3	27.3	-1.8	868	256	0	Calcareous	Rendzina
17	Tavertet (Barcelona)	1118	9.9	23.3	-1.1	971	249	0	Calcareous	Rendzina
18	Calders (Barcelona)	346	13.8	30.4	-0.1	669	159	0.7	Calcareous	Rendzina
19	Sant Martín Sarroca (Barcelona)	315	14.1	28.2	2.1	638	132	1.3	Calcareous	Cambisol
20	La Vall de Bianya (Girona)	481	13.2	27.7	0.6	915	281	0	Siliceous	Cambisol
21	Espinelves (Girona)	868	11.2	24.9	-0.1	898	208	0	Siliceous	Regosol
22	Anglès (Girona)	233	14.5	29.2	1.6	743	165	0.5	Siliceous	Regosol
23	Vilobí d'Onyar (Girona)	145	15	29.4	2.3	697	134	1.3	Siliceous	Fluvisol
24	El Grado (Huesca)	551	12.7	30.4	-0.8	696	165	0.5	Calcareous	Cambisol

2.2. The experts

Six experts participated in the study. All of them are forest engineers with extensive field work experience (20 years or more) in botany or silviculture, they are university professors who frequently provide advice on tree species selection for forest restoration. Each expert ranked the native tree species according to their suitability for the environmental conditions of the 24 evaluation sites (see table 2). As differences among very unsuitable species are not easy to estimate, ties were allowed in the suitability rankings.

2.3. Modeling strategy

We fitted one SDM for each of the considered native tree species (see table 1). The known distribution of each species in continental Spain was used as the model-training sample. We extracted the species occurrence data from 120,938 vegetation relevés included in the Forest Map of Spain (Ruiz de la Torre 1990). Variables related to climate and soils were used as predictors.

1. Climatic data grids were obtained by applying a multiple regression model based on meteorological station data (Sánchez Palomares et al. 1999) to the STRM 3-arc-second (≈ 90 m) elevation dataset (Farr et al. 2007). Six climatic predictors were used: mean annual temperature (ºC), mean maximum temperature of the warmest month (ºC), mean minimum temperature of the coldest month (ºC), mean annual rainfall (mm), mean summer rainfall (mm) and mean dry season length (number of months). The European Soil Database (ESDB) was used as a source for soil (Van Liedekerke 2006). The ESDB comprises a coarse-scale soil map (1 km resolution grid) and an associated database with the values of several soil-related variables for each cell of the map. Two soil-related predictors were extracted from the ESDB: calcareous nature of the parent material (a binary value) and major FAO soil group (14 classes).

We used penalized logistic regression to fit the SDM (Harrell 2001). The penalized regression outperformed an alternative regularization technique called *lasso* (Tibshirani 1994) with small sample sizes in a comparison of regularization methods applied to species distribution models (Reineking & Schröder 2006) and performed at least as well as Maxent (Gastón & García-Viñas 2011).

We fitted penalized models using the standard deviation of each predictor as a scaling factor and used a modified version of Akaike's Information Criterion to select the optimal penalty factor (Harrell 2001). We used the *lrm* and *pentrace* functions from the *rms* package (Harrell 2013) in the R environment for statistical computing (R Core Team 2013).

We expected nonlinear relationships between species occurrence and environmental predictors. Moreover, a significant proportion of responses may be skewed (Oksanen & Minchin 2002). Therefore, we prespecified continuous predictor complexity to four-knot restricted cubic splines (Harrell 2001). This way of adding nonlinear terms allows modelling responses from linear to skewed unimodal and requires the estimation of three parameters per continuous predictor.

2.4. Concordance among experts

The concordance among the suitability rankings provided by the experts was estimated using Kendall's coefficient of concordance (*W*). First, an overall test of independence of all the experts was carried out (Legendre 2005). If the null hypothesis (i.e., the six experts are not concordant with one another) was rejected, *a posteriori* tests were computed using permutation, to determine which of the individual experts are concordant with the others (Legendre 2005). This procedure was repeated for each evaluation site. The calculations were made using the *kendall.global* function from the *vegan* package (Oksanen et al. 2013) in the R environment for statistical computing (R Core Team 2013).

2.5. Evaluation of the predictive performance

We used a novel approach which is particularly appropriate for evaluating the predictive performance of species suitability estimates applied to plant species selection in ecological restoration (Gastón & García-Viñas 2013). The evaluation involves calculating the ability of the suitability estimates to discriminate present from absent species in observed tree assemblages at each evaluation site. We used the area under the receiver operating characteristic curve (AUC) to estimate the probability that the set of SDM offer a higher suitability estimate for a species present at a particular evaluation site than for an absent species (both picked at random). The calculation procedure of Harrell's Concordance Index was used, as it is identical to the AUC in the case of binary outcomes (Harrell et al. 1982). For each evaluation site, the occurrence probability of the 54 tree species was predicted using the set of models. Each species present at an evaluation site was paired with every species absent and the proportion of pairs in which the present species had a higher occurrence probability was calculated (see Gastón & García-Viñas 2013, for more details).

The predictive performance of the suitability estimates of each expert was evaluated following the same procedure used for model predictions, replacing model predicted probabilities with the suitability rankings of the experts.

As the normality and sphericity assumptions were not verified, we used the Friedman test (the non-parametric alternative to the repeated measures ANOVA) to compare the predictive performance of the models against the performance of each expert.

3. Results

The null hypothesis that the six experts are not concordant with one another in the ranking of species suitability was rejected at every evaluation site (see third column in table 3). *A posteriori* testing showed that in most tests the hypothesis that each expert is not concordant with the other five can be rejected, only in 8 tests can we accept that the suitability rankings of one expert are discordant with those of the other experts (see shaded cells in table 3). Even the most discordant expert (E2) was concordant with the rest of the experts at 80% of the evaluation sites. As significant discordance among experts was found only in 5.6 % of the evaluated cases, we can accept that the experts ranked species harmoniously. The observed low heterogeneity among experts means that the use of their suitability estimations as a reference for model predictions is not compromised.

Table 3. Results of the overall and the *a posteriori* tests of Kendall's coefficient of concordance (W) among the experts (E1 to E6). Permutational probabilities based upon 999 random permutations. H_0 (overall test): the six experts are not concordant with one another in the ranking of species suitability. H_0 (*a posteriori* test) each expert is not concordant with the other five in the ranking of species suitability. The cell is shaded in the cases where H_0 could not be rejected.

Evaluation site	**Overall W**	**Overall permuted probability**	**A posteriori tests Permutional probabilities for each expert**					
			E 1	**E 2**	**E 3**	**E 4**	**E 5**	**E 6**
1	0.514	0.001	0.006	0.006	0.006	0.006	0.006	0.006
2	0.584	0.001	0.006	0.006	0.006	0.006	0.006	0.006
3	0.528	0.001	0.006	0.006	0.006	0.006	0.006	0.006
4	0.570	0.001	0.006	0.006	0.006	0.006	0.006	0.006
5	0.526	0.001	0.006	0.010	0.006	0.006	0.006	0.006
6	0.736	0.001	0.006	0.006	0.006	0.006	0.006	0.006
7	0.606	0.001	0.006	0.006	0.006	0.006	0.006	0.006
8	0.619	0.001	0.006	0.006	0.006	0.006	0.006	0.006
9	0.414	0.001	0.042	0.088	0.006	0.006	0.006	0.006
10	0.606	0.001	0.006	0.006	0.006	0.006	0.006	0.006
11	0.435	0.001	0.193	0.016	0.006	0.006	0.006	0.006
12	0.525	0.001	0.006	0.006	0.006	0.006	0.006	0.006
13	0.422	0.001	0.006	0.056	0.006	0.006	0.006	0.006
14	0.445	0.001	0.501	0.024	0.006	0.006	0.006	0.006
15	0.359	0.001	0.022	0.205	0.006	0.006	0.006	0.006
16	0.571	0.001	0.006	0.006	0.006	0.006	0.006	0.006
17	0.561	0.001	0.006	0.080	0.006	0.006	0.006	0.006
18	0.590	0.001	0.006	0.006	0.006	0.006	0.006	0.006
19	0.578	0.001	0.006	0.006	0.006	0.006	0.006	0.006
20	0.516	0.001	0.006	0.006	0.006	0.006	0.006	0.006
21	0.426	0.001	0.006	0.172	0.006	0.136	0.006	0.006
22	0.440	0.001	0.010	0.011	0.006	0.006	0.006	0.006
23	0.616	0.001	0.006	0.006	0.006	0.006	0.006	0.006
24	0.621	0.001	0.006	0.006	0.006	0.006	0.006	0.006

Expert suitability estimations were converted to rankings with ties that ranged from 1 (the lowest suitability) to 54 (the highest suitability) and paired with model predictions (probability of occurrence). Table 4 shows the data used for the calculation of the predictive performance using the evaluation site 13 and expert 4 just as an example. In this case, predictive performance measured using the AUC is better (0.969) for the set of SDM than for the expert (0.778). The expert made more discrimination mistakes, i.e., species that actually occur in the evaluation site with low rankings from the expert (e.g., *Quercus pyrenaica, Pinus pinaster*) and absent species with high rankings (e.g., *Juniperus oxycedrus, Acer campestre*). The discrimination mistakes of the set SDM were only two (high probability of occurrence for two absent species, *Quercus ilex* and *Pinus sylvestris*), resulting in a better ability to discriminate between present and absent species.

Table 4. An example of data corresponding to expert 4 and evaluation site 13 (see table 2 for environmental data). Species are ordered according to habitat suitability as estimated by the set of SDM. AUC is 0.969 for the set of SDM and 0.778 for the expert. Shaded rows correspond to present species (data not shown to experts).

Species	**Expert rank**	**Model prediction**	**Species**	**Expert rank**	**Model prediction**
Quercus ilex	53	0.35	*Quercus suber*	1	0.00
Juniperus thurifera	54	0.24	*Sorbus aucuparia*	1	0.00
Quercus pyrenaica	1	0.23	*Sorbus domestica*	1	0.00
Juniperus communis	50	0.20	*Taxus baccata*	42	0.00
Pinus sylvestris	40	0.16	*Corylus avellana*	1	0.00
Quercus faginea	50	0.13	*Celtis australis*	1	0.00
Pinus pinaster	1	0.13	*Arbutus unedo*	1	0.00
Pinus nigra	49	0.12	*Malus sylvestris*	1	0.00
Populus nigra	44	0.09	*Quercus humilis*	1	0.00
Fraxinus angustifolia	42	0.06	*Fraxinus excelsior*	1	0.00
Salix alba	44	0.03	*Pinus uncinata*	1	0.00
Juniperus oxycedrus	50	0.03	*Betula alba*	1	0.00
Ulmus minor	41	0.02	*Quercus petraea*	1	0.00
Juniperus phoenicea	1	0.02	*Tilia cordata*	1	0.00
Salix fragilis	44	0.01	*Olea europaea*	1	0.00
Populus alba	1	0.01	*Tilia platyphyllos*	1	0.00
Acer monspessulanum	1	0.01	*Sorbus torminalis*	1	0.00
Populus tremula	1	0.01	*Fagus sylvatica*	1	0.00
Sambucus nigra	44	0.01	*Acer pseudoplatanus*	1	0.00
Pinus pinea	1	0.01	*Quercus robur*	1	0.00
Ilex aquifolium	1	0.00	*Ulmus glabra*	1	0.00
Prunus avium	1	0.00	*Betula pendula*	1	0.00
Acer campestre	48	0.00	*Abies alba*	1	0.00
Sorbus aria	1	0.00	*Abies pinsapo*	1	0.00
Pinus halepensis	1	0.00	*Tamarix canariensis*	1	0.00
Alnus glutinosa	1	0.00	*Quercus canariensis*	1	0.00
Castanea sativa	1	0.00	*Fraxinus ornus*	1	0.00

The predictive performance of the experts varied from site to site, but the average performance was well above chance. According to (Swets 1988), the average ability of the experts to discriminate between present and absent species at the evaluation sites was acceptable (AUC 0.7 - 0.9, see fig. 2). The good predictive performance of the experts strengthens the case for their suitability estimations to be used as a reference for model predictions.

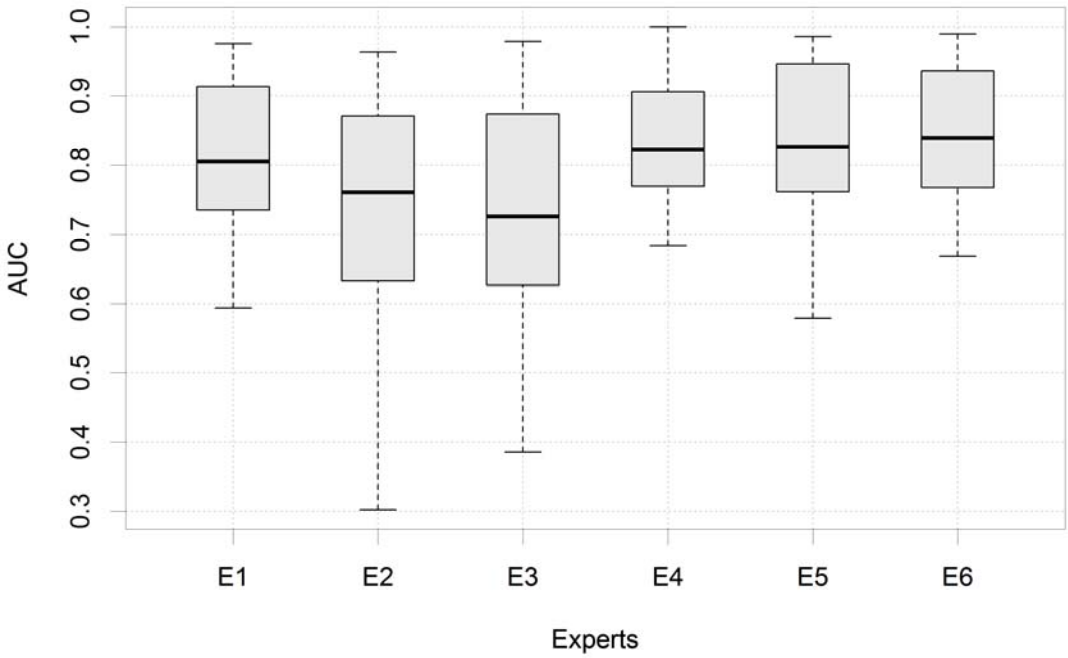

Fig. 2 Boxplot of the predictive performance of each expert measured by the site AUC (i.e., the probability that the expert offers a higher suitability estimate for a species present at a particular evaluation site than for an absent species, both picked at random)

The overall difference in predictive performance between models and experts was small, but significant better performance was observed for models, average AUC difference between models and experts was 0.079 (95% confidence interval: 0.060 - 0.099). Comparing each expert's rankings against the set of SDM yielded similar results. The mean discrimination ability of four of the experts (E1, E4, E5 and E6) showed no significant differences with the AUC of the set of models (see fig. 3). The suitability estimates of the other two experts (E2 and E3) showed a lower discriminative ability than the models ($p<0.05$, see fig. 3). These results support the hypothesis that, provided the models and experts use the same environmental data, the ability of SDM to discriminate between present and absent species in an independent sample of evaluation sites is at least as good as the ability of the experts.

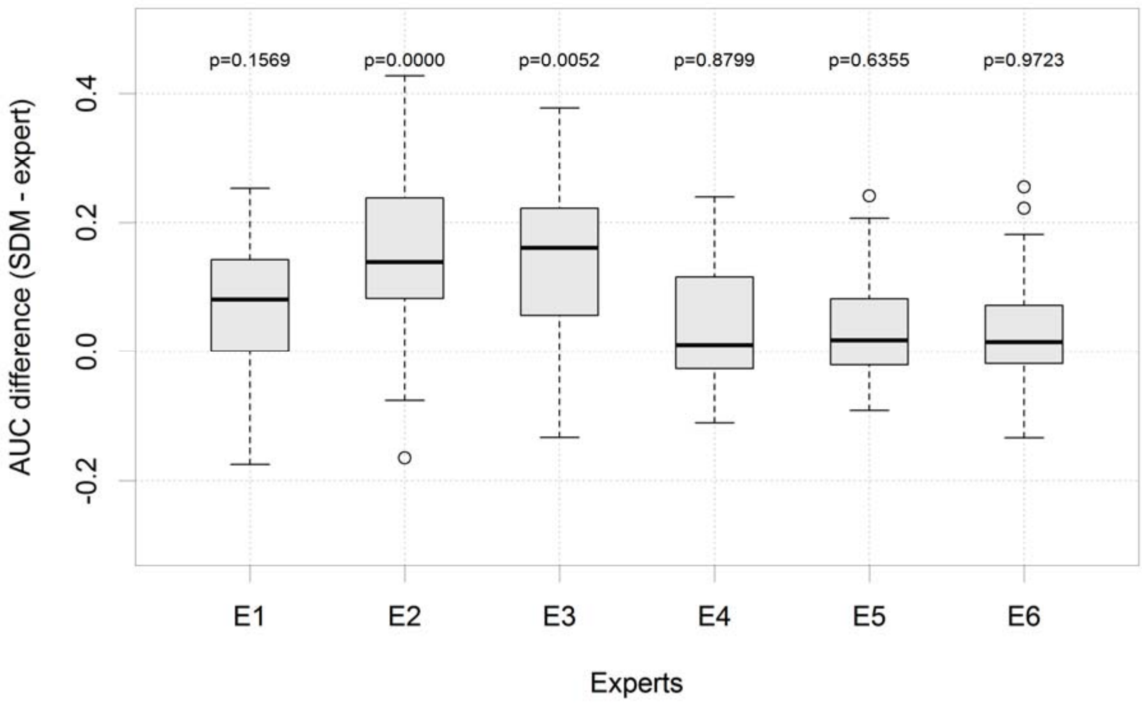

Fig. 3 Boxplot of the differences in predictive performance between the set of models and each expert. The performance is measured by the site AUC (i.e., the probability that the estimated suitability for a species present at a particular evaluation site is greater than the estimate for an absent species, both picked at random). The null hypothesis that the mean difference between the expert and the SDM is zero can be rejected for the experts with p-values under 0.05

4. Discussion

The present study constitutes an additional step in the validation of SDM applied to plant species selection for forest restoration. Most of the previous research work on species distribution modeling applied to tree species selection used presence-only data (e.g., Gandullo & Sánchez Palomares 1994; García López & Allúe Camacho 2004), hence, common evaluation techniques based on the discrimination ability (e.g. AUC) or the reliability of occurrence probabilities (e.g. calibration slope) were not applied (see Pearce & Ferrier 2000 for a review of such techniques). Recent work on SDM applied to species selection for ecological restoration using presence-absence data and independent validation datasets allows a more exhaustive evaluation of model performance. The previous research work comparing model predictions to the observed distribution of individual species (considering both discrimination and calibration, Gastón & García-Viñas 2011) and observed species assemblages (Gastón & García-Viñas 2013) is now completed with a comparison between the performance of the models and that of experts.

Our work adds to the relatively few studies that use expert opinion to validate SDM (van Zonneveld et al. 2013). Previous studies have focused on the validation of individual species models using qualitative (Anderson et al. 2003) or quantitative approaches (van Zonneveld et al. 2013). Our proposal, which involves an expert based quantitative evaluation of a set of SDM against observed species assemblages, is therefore novel. Moreover, to our knowledge this is the first time that SDM used for species selection in ecological restoration have been evaluated against expert opinion.

The models used in this study have known limitations (see Guisan & Thuiller 2005 for a review). Many factors affecting species distribution such as dispersion and biotic interactions are not considered due to data scarcity. For other environmental predictors like soil related variables the available data is not as detailed as desirable. Additionally, some of the assumptions of the models may not be true, e.g., we assume that the species are in equilibrium with the environment. Our results support the idea that, despite their limitations, SDM can perform as well as the experienced experts that participated in the study.

Other limitation of our study is that the information given to the experts is less detailed than the data that the experts could gather in the field. If data on microsite conditions are available, experts are more likely to outperform regional or national models like those used in this study. Our models are cost-effective as they are fitted using free, publicly available data, although gathering new data for models based on microsite conditions may be more expensive and time consuming than direct estimation of species suitability by local experts. Nevertheless, the use of national or regional cost-effective models may be useful even when species selection is performed by a local expert, as an ordered list offered by the set of SDM could be used by the expert as a preliminary check-list.

In conclusion, as the ability of the set of SDM to correctly rank tree species according to habitat suitability was at least as good as the ability of the experts, our results point to the suitability of the proposed species distribution modeling approach to obtain fast and cost effective recommendations for matching species to sites in forest restoration projects.

References

Anderson RP, Lew D, Peterson AT (2003) Evaluating predictive models of species' distributions: criteria for selecting optimal models. Ecol Model 162: 211-232

Araújo MB, Peterson AT (2012) Uses and misuses of bioclimatic envelope modeling. Ecology 93: 1527-1539

Booth TH, Jones PG (1998) Identifying climatically suitable areas for growing particular trees in Latin America. For Ecol Manage 108: 167-173

Booth TH, Nix HA, Busby JR, Hutchinson MF (2013) bioclim: the first species distribution modelling package, its early applications and relevance to most current MaxEnt studies. Divers Distrib : 1-9

Bunce RGH, Barr CJ, Whittaker HA (1983) A stratification system for ecological sampling. In: Fuller RM (ed) Ecological mapping from ground, air and space. Institute of Terrestrial Ecology, Cambridge, pp 39–46

Busby JR (1991) BIOCLIM - a bioclimatic analysis and prediction system. In: Margules CR, Austin M (ed.) Nature Conservation: Cost Effective Biological Surveys and Data Analysis, CSIRO, Melbourne ed., pp. 64-68

Clewell A, Rieger J, Munro J (2005) Guidelines for developing and managing ecological restoration projects. Society for Ecological Restoration, Tucson (USA)

Cuesta B, Benayas JMR, Gallardo A, Villar-Salvador P, González-Espinosa M (2012) Soil chemical properties in abandoned Mediterranean cropland after succession and oak reforestation. Acta Oecol 38: 58-65

Elena Rosselló R (1997) Clasificación biogeoclimática de España peninsular y balear. Ministerio de Agricultura, Pesca y Alimentación, Madrid (España).

Farr TG, Rosen PA, Caro E, Crippen R, Duren R, Hensley S, Kobrick M, Paller M, Rodriguez E, Roth L, Seal D, Shaffer S, Shimada J, Umland J, Werner M, Oskin M, Burbank D, Alsdorf D (2007) The Shuttle Radar Topography Mission. Rev Geophys 45: RG2004

Felicísimo A M (2003) Uses of spatial predictive models in forested areas territorial planning. CIOT-IV International Conference on Spatial Planning: 1-15

Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ Conserv 24: 38-49

Gandullo JM, Sánchez Palomares O (1994) Estaciones ecológicas de los pinares españoles. ICONA, Madrid (España)

García López J M, Allúe Camacho C (2004) Ensayo de un sistema fitoclimático de carácter autoecológico para especies arbóreas forestales en la Península Ibérica y su aplicación en labores de repoblación forestal. Actas IV Congreso Forestal Español: 1-8

Gastón A, García-Viñas JI (2011) Modelling species distributions with penalised logistic regressions: A comparison with maximum entropy models. Ecol Model 222: 2037-2041

Gastón A, García-Viñas JI (2013) Evaluating the predictive performance of stacked species distribution models applied to plant species selection in ecological restoration. Ecol Model 263: 103-108

Guisan A, Thuiller W (2005) Predicting species distribution: offering more than simple habitat models. Ecol Lett 8: 993-1009

Harrell FE (2001) Regression modeling strategies: with applications to linear models, logistic regression and survival analysis. Springer, New York

Harrell FE (2013) rms: Regression Modeling Strategies. R package version 3.6-3

Harrell FE, Califf RM, Pryor DB, Lee KL, Rosati RA (1982) Evaluating the Yield of Medical Tests. JAMA-J Am Med Assoc 247: 2543-2546

Legendre P (2005) Species associations: the Kendall coefficient of concordance revisited. J Agric Biol Envir S 10: 226-245

McIver J, Starr L (2001) Restoration of degraded lands in the interior Columbia River basin: passive vs. active approaches. For Ecol Manage 153: 15-28

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Wagner H (2013) vegan: Community Ecology R package version 2.0-8

Oksanen J, Minchin PR (2002) Continuum theory revisited: what shape are species responses along ecological gradients? Ecol Model 157: 119-129

Padilla FM, Ortega R, Sánchez J, Pugnaire FI (2009) Rethinking species selection for restoration of arid shrublands. Basic Appl Ecol 10: 640-647

Pearce JL, Cherry K, M. D, S. F, Whish G (2001) Incorporating expert opinion and fine-scale vegetation mapping into statistical models of faunal distribution. J Appl Ecol 38: 412-424.

Pearce JL, Ferrier S (2000) Evaluating the predictive performance of habitat models developed using logistic regression. Ecol Model 133: 225-245

Peman Garcia J, Navarro Cerrillo RM, Serrada Hierro R (2008) Species selection guidelines in reforestation. Ruiz de la Torre's contributions. Inv Agrar-Sist Rec F 15: 87-102

R Core Team (2013) R: A Language and Environment for Statistical Computing

Reineking B, Schröder B (2006) Constrain to perform: Regularization of habitat models. Ecol Model 193: 675-690

Ruiz de la Torre J (1990) Mapa forestal de España. Memoria general. ICONA, Madrid (España)

Ruiz de la Torre J, Gil P, García-Viñas JI, González-Adrados JR, Gil F (1990) Catálogo de especies vegetales a utilizar en plantaciones de carreteras. Ministerio de Obras Públicas y Urbanismo, Madrid (España)

Sánchez Palomares O, Sánchez Serrano F, Carretero MP (1999) Modelos y cartografía de estimaciones climáticas termopluviométricas para la España peninsular. Instituto Nacional de Investigaciones Agrarias, Madrid, España

Seoane J, Bustamante J, Díaz-Delgado R (2005) Effect of Expert Opinion on the Predictive Ability of Environmental Models of Bird Distribution. Conserv Biol 19: 512-522

SER (2004) The SER International Primer on Ecological Restoration. Society for Ecological Restoration International, Tucson (USA)

Swets J (1988) Measuring the accuracy of diagnostic systems. Science 240: 1285-1293.

Thuiller W (2003) BIOMOD - optimizing predictions of species distributions and projecting potential future shifts under global change. Global Change Biol 9: 1353-1362

Tibshirani R (1994) Regression shrinkage and selection via the Lasso. J Roy Stat Soc B 58: 267-288

Van Liedekerke M, Jones A, Panagos P (2006) ESDBv2 Raster Library, a set of rasters derived from the European Soil Database distribution v2.0. European Commission and the European Soil Bureau Network, CDROM, EUR 19945 EN

van Zonneveld M, Castañeda N, Scheldeman X, van Etten J, Van Damme P (2013) Application of consensus theory to formalize expert evaluations of plant species distribution models. Appl Veg Sci. doi: 10.1111/avsc.12081

Webb DB, Wood PJ, Smith J (1980) A guide to species selection for tropical and sub-tropical plantations. Tropical Forestry Papers 15, University of Oxford, Oxford.