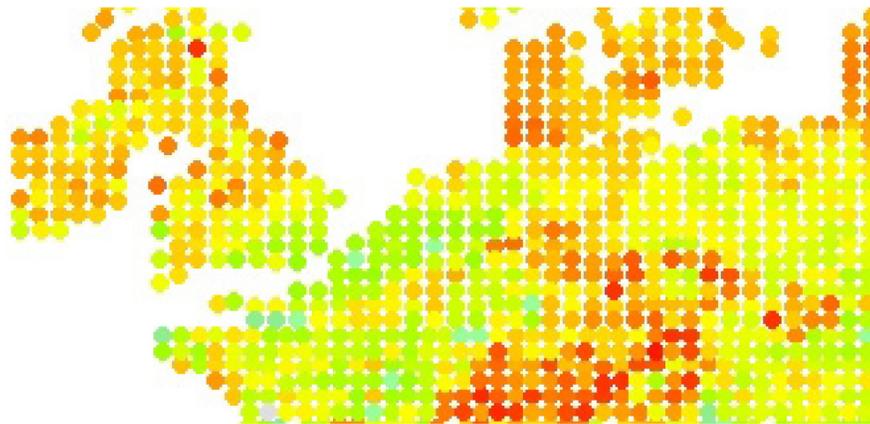




UPPSALA
UNIVERSITET

Choice of threshold alters projections of species range shifts under climate change



Hedvig Nenzén

Degree project in biology, Master of science (2 years), 2010

Examensarbete i biologi 30 hp till masterexamen, 2010

Biology Education Centre, Uppsala University, and Museo Nacional de Ciencias Naturales, Spain

Supervisors: Mats Björklund and Miguel Araújo

Table of Contents

Abstract	3
Introduction	4
Methods	
Climate data.....	7
Species data.....	7
Bioclimatic envelope modelling.....	8
Threshold optimization.....	8
Analysis of range changes and species temporal turnover.....	9
Results	
How are species range changes and turnover affected by the choice of thresholds?.....	11
How much variability in projections is attributed to the choice of thresholds?.....	13
How is variability in projections spatially and environmentally distributed?.....	14
Discussion	15
Acknowledgements	16
References	17
Supplementary Data	
List of species modelled.....	21

Abstract

Bioclimatic Envelope Models (BEM) have been extensively used to investigate climate change impacts on species potential distributions and make inferences about species extinction risk. In addition to the theoretical challenges of using BEM for inferring extinction risk there are a number of algorithmic uncertainties. One of the least explored sources of algorithmic uncertainty is the selection of thresholds to transform modelled probabilities of occurrence (or indices of suitability) into binary predictions of species presence and absence. I investigate the impacts of such thresholds in the specific context of species extinctions risk assessment under climate change. BEM for European tree species were fitted, using seven modelling techniques and 14 threshold-setting techniques. Estimated range shifts obtained by applying different threshold-setting methods were compared after grouping them by IUCN-based categories of threat. It was found that thresholds have a large impact on the inferred risks of extinction, producing 1.4- to 4.4-fold differences in the number of species projected to become threatened by climate change. I quantified sources of variability in the projections, and found that the selection of thresholds explained more variability in the results than the choice of the modelling technique. Results demonstrate that threshold selection has large – albeit often unappreciated - consequences for estimating species range shifts under climate change.

Introduction

Bioclimatic Envelope Models (BEM) characterise species climatic requirements by relating species occurrences with aspects of climate. These models have been used for a variety of theoretical as well as applied purposes (e.g., Guisan and Thuiller, 2005; Araújo and Peterson, 2010). Here, I focus on the use of the models for studying climate change impacts on species ranges and particularly for inferences about extinction risk (e.g., Thomas et al., 2004; Thuiller et al., 2005). The approach has been criticized on theoretical grounds (Akçakaya et al., 2006) and new tools are being devised to couple climatic and population processes (e.g. Keith et al., 2008; Anderson et al., 2009), thus providing more robust estimates of extinction risk.

However, in addition to the theoretical challenges for inferring extinction risk with BEM, there are a number of algorithmic uncertainties that contribute to uncertainty in projections (for review see Heikkinen et al., 2006; Araújo and New, 2007). One of the least explored sources of uncertainty is the rule to transform probabilities of occurrence (or indices of suitability) produced by models into binary predictions of species presence and absence. There are probably as many rules for setting thresholds (or cut-offs) as modelling methods and often they are chosen arbitrarily since no guidelines exist for helping the selection of the threshold-setting rules. Here, I investigate the impacts of different approaches of threshold optimization in the specific context of BEM used for making inferences of species extinction risk under climate change.

The impact of threshold-setting methods in BEM has previously been discussed in the literature. Fielding & Bell (1997) stated that a fixed threshold to transform values will perform badly by exaggerating prediction errors. A good threshold will minimize the presence of prediction errors: false negatives (modelled absences that really are presences) and false positives (modelled presences that really are absences). They highlighted that choice of modelling method may influence the values of probabilities of occurrence, and thus a single threshold for different models would be unsuitable. The authors also stated that prevalence; the number of grid cells that are occupied relative to the total number of grid cells in original species distribution data was important. Prevalence influences the absolute values of probabilities of occurrence produced by models, and thus the projected classified ranges if a fixed threshold is applied.

To avoid fixed thresholds many other threshold-optimization methods have been proposed (Table 1), but most have been rarely, if ever, used in climate change studies. Table 2 reviews published studies that have used at least two threshold-setting methods in BEM. Typically, in past studies, the number of threshold-setting methods or the number of modelling methods was low, preventing a thorough evaluation of threshold performance (Table 2).

Table 1. The fourteen threshold-setting methods used in this study with abbreviations in parenthesis. Table based on Liu et al. 2005, Pearson 2007 and Freeman and Moisen 2008. For common threshold setting methods I have cited studies where it was initially used, and studies that used the threshold in models of species distribution changes under climate change. For accuracy-based thresholds I calculated the values in the confusion matrix: TP = true positives, TN = true negatives, FP = false positives, FN = false negatives. Sensitivity = TP / TP + FN. Specificity = TN / FP + TN.

Threshold	Description	Citations in Bioclimatic Envelope Modelling
Subjective thresholds		
Fixed	0.5 in this study	Manel 1999, Hijmans and Graham 2006, Buckley et al. 2010.
Data-driven thresholds		
Observed prevalence (obsprev)	Using the original species prevalence as the threshold	Cramer 2003, Araújo and Luoto 2007, Baselga and Araújo 2009
Predicted prevalence same as observed (PredPrev=Obs)	Maintain the original prevalence	Hartley et al. 2006, Dormann et al. 2008
Average probability (avgprob)	Taking the mean (in this case) of the probabilities of occurrence of occupied locations for presence/absence data as the threshold	Cramer 2003
Mid-point probability (midptprob)	Taking the midpoint of the probabilities between the occupied and unoccupied sites.	Fielding and Haworth 1995
Accuracy-based thresholds		
Plot based Precision-recall (PRplotbased)	Minimize the distance to the 1,1 corner of the precision (TP / TP + FP) against recall (sensitivity) plot	
Minimize Precision-recall (PRmin)	Minimize the difference between the precision (TP / TP + FP) and recall (sensitivity)	Schapire et al. 1998
Overall Prediction Success (OPS)	Maximize OPS (TP + TN / number of data)	
F	Maximize $F = 1/\alpha P + (1 - \alpha) / R$ $\alpha = 0.5$ (no preference precision or recall)	Schapire et al. 1998
kappa	Maximize Cohen's kappa statistic	Huntley et al. 1995, Berry et al. 2002, Segurado and Araújo 2004, Araújo et al. 2005b, Elith et al. 2006
Maximize sum of Sensitivity and Specificity (SeSpMax)	Maximize the sum of sensitivity and specificity	Cantor et al. 1999, Manel et al. 2001, Svenning et al. 2008
Equalize sensitivity and specificity (SeSpeql)	Minimize the absolute difference between the sensitivity and specificity	Fielding and Bell 1997, Pearson et al. 2004, Pearson et al. 2006
TSS	True Skill Statistic, sensitivity + specificity -1	Allouche et al. 2006, Keenan et al. 2010
ROC	Minimize the distance to the 0,1 corner of the sensitivity against 1-specificity curve (Receiver Operating Curve).	Cantor et al. 1999, Pearce and Ferrier 2000, Araújo et al. 2005b

In this study I employ two methods that are appropriate to explore the consequences of particular threshold selection under climate change, and analyse performance on the basis of (1) species and (2) locations. First, to enable exploration of results for species, I converted changes in potential suitable climate, and the corresponding estimated range shifts, into categories of threat using a simplified interpretation of the IUCN criteria for Red Listing of species. Notice that the IUCN Red List criteria are used here as a strategy to explore the sensitivity of the results to different thresholds, and not to perform or recommend Red Listing of species based on the untransformed outputs of BEM models (Akçakaya et al., 2006; Brook et al., 2009). Second, to explore results for locations, I calculate the temporal turnover of species composition in each modelled grid cell. Then I partition and compare the variability brought to the measures of turnover by the choice of threshold-setting technique and modelling technique (Diniz-Filho et al., 2009).

Table 2. Results and number of models, thresholds and species from studies that explicitly investigated thresholds in BEM. Note that not all studies investigated performances of thresholds under climate change.

	Models	Thresholds	Species	Evaluation Data	Evaluation method	Results
Manel et al. 2001	1	2	34 invertebrate families	Cross-validation and data from other area	Predictive accuracy	No evidence for prevalence-dependency on kappa. ROC proposed as a threshold-setting technique.
Thuiller 2004 ^a	4	2	1350 plant species	Turnover results	Component loadings on PCA axes	Impact of threshold-setting method highlighted
Araújo et al. 2005b ^a	7	2	116 bird species	Cross-validation and independent data from different time	Predictive accuracy	kappa was more accurate than ROC
Liu et al. 2005	1	12	2 plant species ^b	Cross-validation	Predictive accuracy	obsprev, avgprob, SeSpmax, SeSpeql and ROC were most accurate thresholds
Allouche et al. 2006	1	2	128 plant species	Independent inventory data from same area	Predictive accuracy	TSS proposed as a threshold-setting technique
Jiménez-Valverde and Lobo 2007	1	4	1 virtual species ^b	Cross-validation	Predictive accuracy	SeSpmax and SeSpeql were most accurate thresholds
Freeman and Moisen 2008	1	11	13 tree species	Cross-validation	Predictive accuracy and change in predicted prevalence	PredPrev=Obs and kappa were most accurate thresholds
Present study ^a	7	14	116 tree species	Species range changes and turnover	Extinction risk, uncertainty	Threshold-setting method contributed much variation in results under climate change

^a Studies investigated threshold performance under climate change

^b Species was sampled at different prevalence

For this study, Bioclimatic Envelope Models are fitted for 116 European tree species, using seven modelling methods and 14 threshold-setting methods. By using many combinations of modelling algorithms and threshold-setting approaches I seek to distinguish trends and to find a number of generalities that might help to guide the selection of thresholds for studies assessing potential impacts of climate change on species distributions. Specifically, I ask: (I) How are projections of species extinction risk and temporal turnover affected by threshold selection? (II) How much variability in turnover values is attributed to the choice of threshold? (III) How is variability in projections spatially and environmentally distributed?

Methods

Climate data

A set of aggregated climate parameters were derived from an updated version of climate data provided by New et al., (2000). The updated data provides monthly values for the years 1901-2000 in a 10' grid resolution (Mitchell et al., 2004; Schröter et al., 2005). Average monthly temperature and precipitation in grid cells covering the mapped area of Europe were used to calculate mean values of eleven different climate parameters for the period of 1961-1991 (referred to as baseline data). To minimize model overfitting and ensure comparability across model projections I selected a smaller set of uncorrelated variables for inclusion in the models, after performing a Principal Components Analysis (PCA) of the climate data. The PCA identified two axes that together explain 94.7% of the variance in the data. I retained the two climatic variables with the highest component loadings in the first (Growing Degree Days) and second (Annual Precipitation) axes (see also Baselga and Araújo, 2009). These were expected to summarise important abiotic factors that directly limit the distributions of plant species (e.g., Woodward 1987; Sykes et al., 1996).

For the future climate data, I used the same bioclimatic variables modelled with the Hadley Global Climate Model version 3 (Schröter et al., 2005). As I was not interested in studying uncertainties arising from the choice of alternative emission scenarios, I used only one scenario: the fossil intensive A1FI scenario for 2050.

Species data

For this study, I considered native tree species distributed across Europe. Trees were chosen because: (i) their distribution and ecology is relatively well known and the data has been extensively used in other modelling studies (e.g., Araújo and Williams, 2000; Thuiller et al., 2003; Thuiller et al., 2005; Baselga and Araújo, 2009), (ii) their richness is correlated (Spearman correlation $\rho=0.80$, $P<0.001$) with the overall richness of the Atlas Flora Europaeae (AFE) data set (Araújo and Williams, 2000), and (iii) they are long-lived organisms and their distribution is relatively stable in comparison with some other groups. The species presence-absence data are a subset of AFE (Jalas and Suominen [Eds.]. 1972-1996), which was digitized by Lahti and Lampinen (1999). Data was originally located in 4419 UTM (Universal Transverse Mercator) 50×50 km grid cells, but I used only 2130 grid cells, excluding most of the eastern European countries (except for the Baltic States), because of low recording efforts in these areas (Williams et al., 2000). Species occurring in less than 25 grid cells were excluded from analyses to avoid problems of modelling species with small sample sizes (Stockwell and Peterson, 2002): the

reduced dataset comprised 116 tree species (Supplementary Data A). Original prevalence ranged from 1-80 % of all grid cells.

Bioclimatic envelope modelling

To characterise species potential distributions Bioclimatic Envelope Models were fitted using BIOMOD-R, which implements multiple model classes in a single platform (Thuiller et al., 2009). I employed the seven algorithms that generate continuous predictions of probabilities of occurrence as an output. Models included: generalized linear models (GLM, polynomial terms), generalized additive models (GAM, polynomial degree 3), multivariate adaptive regression splines (MARS), classification tree analysis (CTA), flexible discriminant analysis (FDA), random forests (RF, 750 trees), and generalized boosted models (GBM, 2000 trees). Since independent evaluation data were unavailable, a single-step cross-validation was performed by calibrating the models with a random sample of 70 % of the original presence records for each species, while setting thresholds with the remaining 30% (step 1 and 2 in Fig. 1).

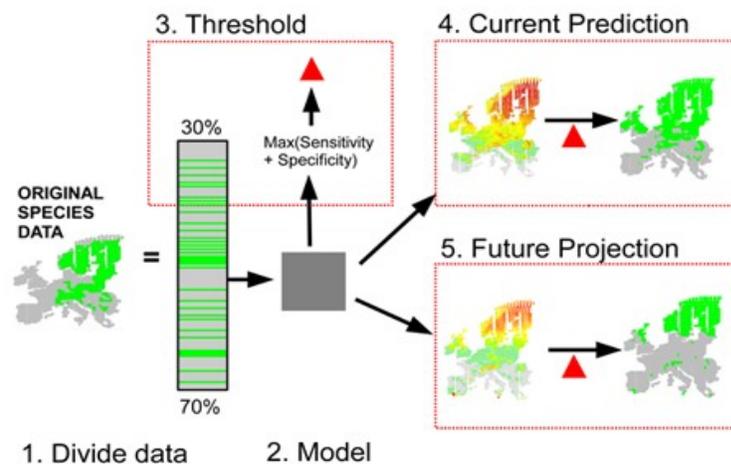


Fig. 1. Flowchart illustrating the steps to optimize one accuracy-based threshold for *Picea abies* (spruce). (1) First original species data was randomly split into calibration and evaluation data. (2) The calibration data was used to train models. (3) The models were projected on the 30% evaluation data. A sequence of thresholds were applied to predictions to find the value that best satisfies the threshold criteria, in this case the threshold that maximized the sum of sensitivity and specificity. (4) The model was applied to current climate data and the threshold from step 3 was applied to transform data to binary species distributions. (5) The model was applied to future climate data. The process was repeated for 116 species, seven model techniques and nine accuracy-based threshold-setting methods. Figure based on Pearson (2007) and redrawn with permission.

Threshold optimization

A threshold is required for transforming continuous probabilities of occurrence (or indices of suitability) from the models into binary presence-absence values. The procedure is shown in Fig. 1, and the R script used for threshold optimization in this study is documented in the Supplementary Data B. First, all thresholds were calculated (step 3), then the same thresholds

were used as cut-offs for transforming probabilities of occurrence into presence and absence of species in the baseline and future periods (steps 4 and 5). A large number of methodologies exist for optimizing the thresholds and here I used 14 different methods (Table 2). Broadly, I identified three families of approaches for selecting thresholds: (I) subjective; (II) data-driven – using the data for model building and predicted probability values; (III) accuracy-based – finding the threshold that produces the best agreement between the evaluation data and the original data. Subjective thresholds are decided by the modeller, whereas data-driven thresholds are set by using the data for calibration of the models and predicted probability values.

For accuracy-based thresholds, agreement between the predicted range and the withheld evaluation data is assessed (30% of original distribution data). The threshold that produces the fewest prediction errors is selected. I applied a progression of 2000 'testing' thresholds to the continuous probability data. Each 'testing' threshold was used as the cut-off to classify the continuous predicted probability values to present or absent. The classified prediction was then compared to the original species data. Four outcomes were possible: true positives, true negatives, false positives and false negatives. Outcomes from comparisons in all grid cells were tallied in a confusion matrix (a contingency table of presences and absences in the predicted data and original data, Fielding and Bell, 1997) for each 'testing' threshold. From the confusion matrix I were able to calculate the various accuracy-based evaluation indices in Table 2. The 'testing' threshold that generated the prediction that, when compared to the original species distribution, best satisfied each criterion was selected as the final threshold. If several thresholds produced identical accuracy values, I chose the lowest threshold.

Finally, I calibrated models for each species with 100% of the species data and projected probabilities of occurrence into both the baseline and future climates. They were transformed into binary ranges using previously calculated thresholds. The resulting 98 final projections for each species arise from combinations of 7 modelling techniques and 14 threshold-setting methods. For the 116 species modelled, 11368 model outputs were generated.

Analysis of range changes and temporal turnover of species

To compare performance of models and thresholds between baseline and future scenarios I calculated two measures: range changes per species; and temporal turnover of species composition per grid cell. Range changes were defined as the difference between the numbers of sites gained and lost relative to the sum of the currently occupied grid cells. To assess the impact of thresholds on species range changes, I calculated how many species would be candidates to the IUCN Red List if the magnitude of range changes were implemented as a criterion for predicting extinction risk. I applied IUCN criterion A3a population size reduction projected or suspected to be met within the next 10 years or three generations, whichever is the longer (IUCN 2001, page 16). More specifically, I applied criterion A3c that refers to the decline in the area of occupancy. Hence, to identify a critically endangered species I used a range-size reduction of 80%; endangered 50%; and vulnerable 30%. Species that were projected to lose their entire range were considered critically endangered and not extinct, as European trees modelled typically survive outside the study area. Since I were explicitly interested in modelling potential ranges rather than actual ones I considered unlimited migration in all calculations. Also, because trees were modelled, I assumed long generation time and that range changes from 2000 to 2050 correspond to three generations. However, for some species the 50 years modelled might correspond to just two generations, so

there is a chance that overestimation of risk occurs for those species, and underestimation for others.

Temporal turnover of species between the two time periods was defined as $\text{species gained} + \text{species lost} / \text{species richness} + \text{species gained}$ in each grid cell (e.g. Peterson et al., 2002). This formula measures changes in species composition, rather than changes in species richness. To distinguish trends in turnover values among the 98 results for each location, I employed various statistical methods. All analyses were carried out in R statistical environment (R Development Core Team 2010) using packages BIOMOD (Thuiller et al., 2009) and vegan (Oksanen et al., 2010).

First, I performed a Principal Components Analysis (PCA) to determine if certain thresholds and models were more similar than others (Araújo et al., 2005b). If the first PCA axis explains a large proportion of the variation in results, then it is assumed that the axis is close to all values (e.g. Thuiller 2004). That was not the case here, as the first axis explained 52% of the variation in the entire data set. Many combinations of methods had a small component loading and there was no clear separation of different threshold-setting methods (maximum loading on the first axis was 0.14).

To further identify the differences in turnover between threshold-setting methods, a clustering technique was used to group the most similar methods. First I calculated the Euclidean-distance matrix that measures distances between all pairs of measurements. From the distance matrix I implemented an agglomerative hierarchical classification (single linkage clustering) to determine which groups of threshold-setting methods and models were alike (Araújo et al., 2006). From this clustering I identified groups of thresholds that produced the most similar values of turnover.

Nonparametric analysis of similarity (ANOSIM, Clarke and Green, 1988) was calculated to test if the resulting clusters were statistically different. A high value of the R global statistic implies that a high degree of separation exists between groups. As grouping factor I used the groups observed by the clustering procedure. Permutations of the grouping factor were done 999 times, allowing for the value of the R statistic to be calculated and compared against a null model.

To examine sources of variability in all projections, I partitioned the sources of variability in turnover values. The two factors (7 modelling methods x 14 threshold-setting methods) resulted in 98 values of turnover in each grid cell, which was considered as separate experiments. In each grid cell variability in turnover was assessed by two-way ANOVAs with threshold-setting method and model class as the two factors (Diniz-Filho et al., 2009). From the ANOVAs I obtained the total variance of turnover values in each grid cell, expressed in total sums of squares. I also obtained the relative contribution to the sums of squares by each factor and by the interaction of factors. I divided the sums of squares for each factor by the total sums of squares in each grid cell. Thus I found the percentage of variability contributed by each factor in each grid cell.

First I plotted values for all proportions of variability (in total 2130) contributed by each factor (Fig. 3). To determine if results are significantly different, I calculated two-sided Kolmogorov-Smirnov tests between pairs of results. Kolmogorov-Smirnov tests are non-parametric and compare two distributions of samples without making assumptions about the shape of their

distribution (Legendre and Legendre, 1998). The null hypothesis is that the two values have the same distribution, and the test gives confidence intervals to reject the null hypothesis. The relative proportion of variability contributed by each factor in each grid cell was plotted in both geographical and environmental space (Fig. 4).

Results

How are projections of species extinction risk and temporal turnover affected by threshold selection?

The projections of range changes were greatly affected by threshold-setting method, as expressed by estimates of extinction risk. The percentage of species forecasted to become critically endangered in 2050 ranged from 0 to 28% depending on the modelling and threshold technique (Table 3). The threshold/model combination that yielded the highest percentage of threatened species was the precision-recall threshold method with random forests modelling method, which projected that 74% of the European tree species would be candidates for Red Listing under climate change. Within the same modelling methods, estimates of extinction risks differed from one another from 1.4 up to 4.4-fold, due to differences in threshold performance. Conversely, comparing modelling methods within the same threshold-setting method revealed that extinction estimates only differ by as much as 2.8-fold. Therefore estimates of probabilities of extinction were more sensitive to performance of threshold-setting method than performance of modelling method.

Table 3. Percentage of European Tree species that became candidates to Red Listing among the 116 species modelled. The first value in each cell is the percentage of species that could become “Critically Endangered” under the corresponding modelling class (columns) and threshold-setting method (rows). The second number is the percentage of species that is potentially “Endangered”. The third number is the percentage of potentially “Vulnerable” species. See text and Table 2 for meanings of abbreviations.

		CTA	GAM	GBM	GLM	MARS	FDA	RF
Subjective	fixed	2, 3, 22	3, 3, 16	7, 6, 16	2, 6, 15	6, 13, 21	0, 14, 24	5, 31, 22
	obsprev	1, 5, 29	0, 7, 31	0, 5, 32	0, 6, 30	2, 0, 15	0, 5, 30	0, 0, 22
Data- driven	predPrev=Obs	2, 8, 21	3, 9, 27	2, 11, 29	2, 11, 25	9, 16, 22	3, 13, 26	7, 30, 20
	avgprob	1, 5, 29	0, 7, 31	0, 5, 32	0, 6, 30	6, 3, 26	0, 8, 30	0, 0, 22
	midtprob	1, 10, 30	1, 8, 30	1, 9, 33	0, 9, 30	7, 8, 27	0, 12, 26	2, 16, 33
Accuracy- based	PRplotbased	3, 9, 22	9, 16, 7	9, 16, 28	7, 18, 7	28, 17, 9	9, 16, 20	23, 40, 11
	PRmin	0, 5, 26	4, 7, 28	2, 13, 30	1, 13, 23	10, 14, 22	3, 14, 22	3, 8, 23
	OPS	0, 8, 20	3, 6, 16	6, 10, 26	4, 8, 16	15, 12, 18	25, 7, 16	25, 29, 15
	F	2, 12, 23	3, 9, 25	3, 13, 28	3, 10, 27	9, 17, 22	4, 14, 24	3, 23, 16
	kappa	1, 8, 22	1, 8, 31	2, 9, 28	0, 10, 27	9, 12, 25	1, 9, 26	0, 9, 13
	SeSpmax	0, 4, 20	0, 4, 32	0, 5, 31	0, 7, 30	6, 5, 23	0, 7, 32	0, 1, 16
	SeSpeql	0, 3, 22	0, 7, 34	0, 6, 32	0, 8, 34	6, 3, 28	0, 10, 30	0, 0, 20
	TSS	1, 4, 22	0, 6, 31	0, 4, 32	0, 7, 30	6, 5, 22	0, 8, 30	0, 1, 17
	ROC	0, 3, 20	0, 7, 34	0, 6, 30	0, 9, 32	6, 5, 26	0, 9, 30	0, 0, 21

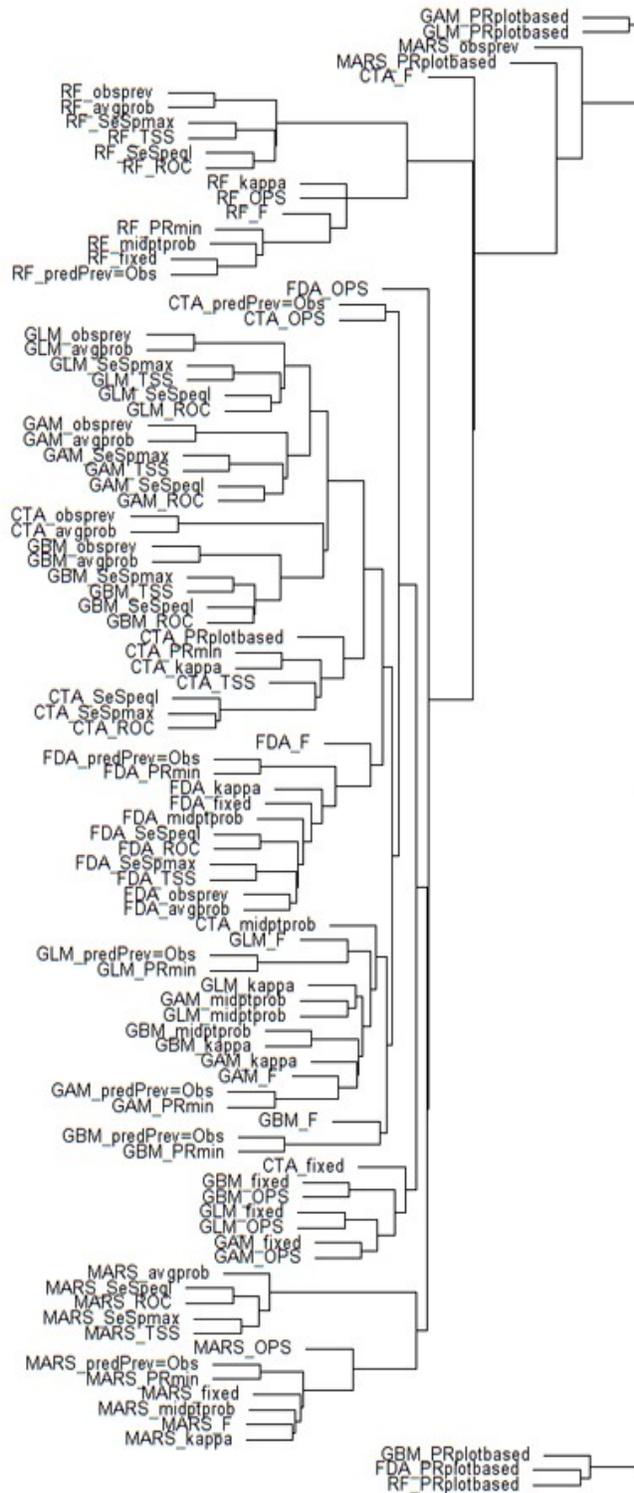


Fig. 2. Cluster dendrogram from hierarchical single linkage clustering analysis of turnover values, based on the Euclidean distance matrix. I separated the following clusters: (1) PR plot based, (2) RF, (3) MARS, (4) FDA, (5) Fixed threshold and OPS (6) GLM, GBM, GAM modelling methods with thresholds kappa, TSS, ROC, SeSpeql and predPrev=Obs. The global ANOSIM was 0.57 ($P = 0.01$).

It was observed that both models and thresholds contribute to similarities in turnover values. The clustering dendrogram (Fig. 2) revealed that reasons for variation in turnover were the threshold-setting method (groups 1 and 5), the modelling method (groups 2-4), or a combination of both. In the analysis of clustering of turnover values, it was possible to separate the following groups: (1) PR plot-based thresholds; (2) Random Forest modelling method; (3) MARS modelling method; (4) FDA modelling method; (5) Fixed and OPS thresholds; and (6) remaining models and thresholds that were not easily separated. Group 6 corresponded to the most utilized accuracy-based thresholds such as TSS, ROC and maximising and equalising sensitivity and specificity. The ANOSIM analysis performed, that tested the degree of separation among all groups, had an R statistic of 0.57 ($P= 0.01$), which indicates weak grouping.

How much variability in projections is attributed to the choice of threshold?

The ANOVA applied to each grid cell indicated that the selection of threshold optimization method contributed to greater variability in projections of species temporal turnover than the selection of modelling method (Fig. 3). Across all grid cells, the proportion of variability explained by the threshold-setting method was generally the greatest, and contributed to a median of 43 % of deviance in all grid cells. The modelling method and interaction between thresholds and modelling methods contributed similar amount of variability, i.e, 27%. All distributions of values were found to be significantly different ($p<0.05$) when compared using pair-wise Kolmogorov-Smirnov tests.

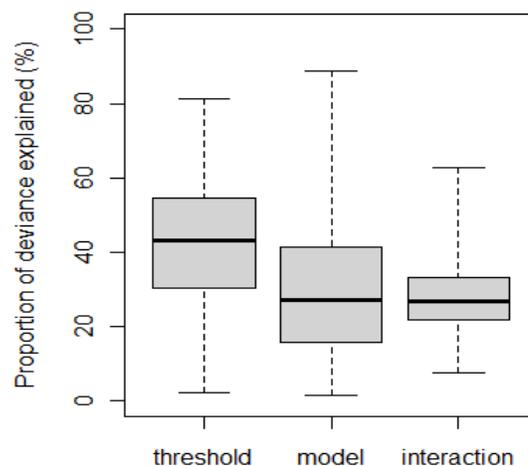


Fig. 3. Importance of sources of variation in species turnover from two-way ANOVAs in each grid cells. $n= 2130$. Boxes show median, quantiles and extremes in proportions of total sums of squares for each source of variation (factor) and their interaction. All pair-wise comparisons between values were found to be significantly different ($p<0.05$) in Kolmogorov-Smirnov tests.

How is variability in projections spatially and environmentally distributed?

Variability across projections of species temporal turnover was spatially and climatically structured (Fig. 4). In the centre of Europe, such as Northern France, Germany and Southern Scandinavia, the threshold-setting method contributed to a greater proportion of the variability in turnover values. At the edges of Europe, such as Southern Spain, Italy, Northern Scandinavia, and England, the choice of the modelling technique contributed more to such variability (Fig. 4, upper panel). When the values in environmental space were plotted, I found the same trend (Fig. 4, lower panel). At the climatic extremes of the study area, models explained a greater proportion of the variability in modelled turnover values, representing a higher proportion of the sum of squares. In centre of climatic space where all models produced similar projections, the choice of thresholds in turn contributed to more variability.

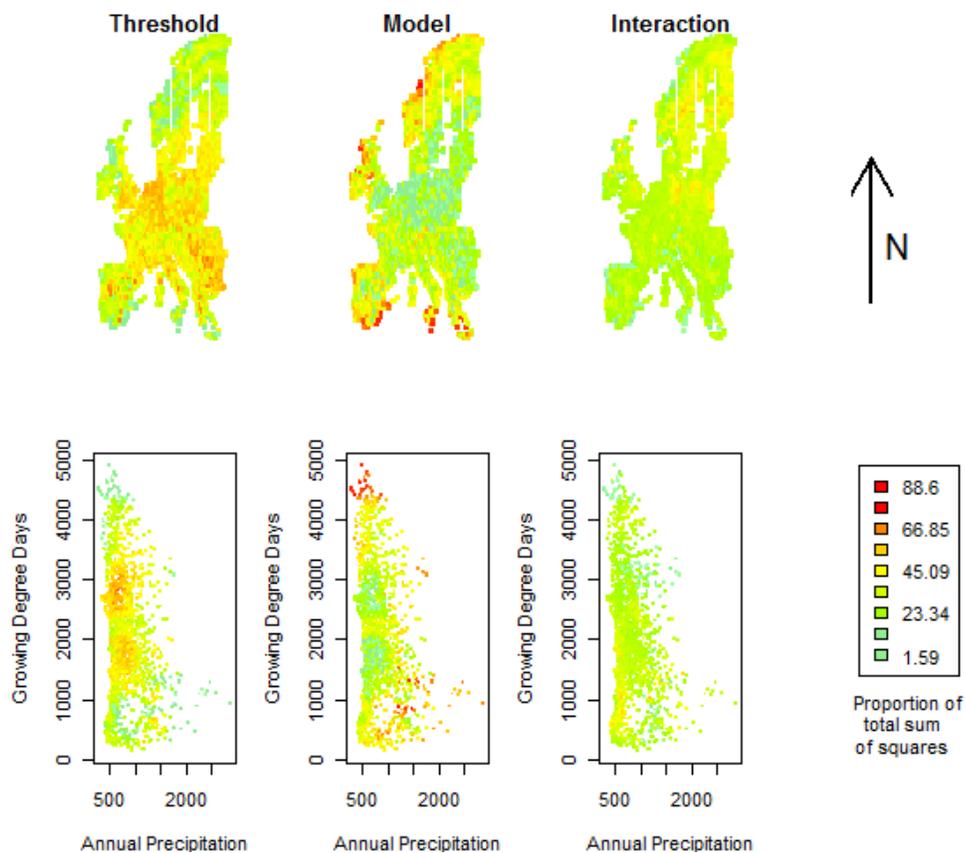


Fig. 4. The distribution of the proportion of sum of squares (in percentages) for ANOVAs of turnover values among 116 species of European trees in geographical space (upper panel) and environmental space (lower panel). Axes represent the two environmental variables used for modelling in this study. Threshold-setting methods are represented on the left, model classes in the middle and interaction between factors on the right.

Discussion

In this study I evaluated the impact of threshold-setting methods on climate-change induced projections of species range shift by bioclimatic envelope models. I found that the choice of the threshold method markedly altered inferences of extinction risks and temporal turnover of species, more so than choice of the modelling method. Threshold selection generated up to 4.4-fold differences in percentages of species projected to become threatened under climate change (Table 3). Within the same threshold method but with different modelling methods the largest magnitude of differences in results is only 2.8.

In addition to illustrating threshold performances by projected differences in species extinction risks, I also quantified the amount of variability added to projections by threshold selection. A number of studies have explored sources of algorithmic uncertainty in BEM (for review see Heikkinen et al., 2006, Araújo and New, 2007), but to date no study has systematically investigated the amount of uncertainty accrued by using different thresholds. When partitioning the variability in results contributed by various sources of variability, earlier studies have consistently reported that modelling method explained most of the variation in results (Thuiller 2004; Dormann et al., 2008; Diniz-Filho et al., 2009; Buisson et al., 2010). Here, I show that including an additional source of variability – choice of threshold optimization method – can contribute to a larger variability in estimates of both species range changes and turnover than the choice of the modelling method.

One reason why such high variability in performances of threshold-setting methods was recorded in our study is that I took into account a greater variety of thresholds, including some that are rarely used, such as PRplotbased thresholds. However, I also employed many different modelling types, ranging from statistical models such as generalized linear and additive model, to machine learning models such as boosted regression trees. I also found a high interaction between effects of modelling type and thresholds, which highlights the difficulty of separating the effects of modelling methods and thresholds. This was reflected both when examining the low separation of clusters of range change estimates (Fig. 2), and the high ANOVA interaction terms when examining turnover values over the study area (Fig. 3). Model methods and threshold methods interact, making it difficult to recommend a single threshold for all modelling methods.

The close relationship between models and thresholds could also be inferred from examining spatial distribution of uncertainty in turnover results (Fig. 4). At the centre of the study area, models contributed little to uncertainty and were able to produce similar projections. Since models performed equally, instead choice of thresholds produced most uncertainty. At the climatic extremes of the study area models were forced to extrapolate beyond the species ranges and contributed greater variation. This is probably because modelling algorithms have difficulties in projecting distributions to novel climatic conditions (Thuiller et al., 2004; Fitzpatrick and Hargrove, 2009) and these are more prone to emerge in climatic edges of study areas, which in this particular case, also coincide with the geographical edges. Since climate change may result in novel climates, there is much discussion on the ability of models to predict in a climate in which it has not been calibrated (Randin et al., 2006; Araújo and Rahbek, 2006).

Some patterns emerge from examining the results from different threshold optimization methods.

Certain commonly used accuracy-based thresholds such as SeSpmax, SeSpeql, TSS and ROC produce more conservative estimates of extinction risk (Table 3). Their performance was also less sensitive to different modelling methods, except to the MARS modelling method. This was not the case when maximising kappa as a threshold, which showed higher values under certain modelling methods. Kappa has received criticism for performance as an index of prediction accuracy, especially at low prevalences (McPherson et al., 2004; Allouche et al., 2006; and Jimenez-Valverde and Lobo, 2007), and should also be treated with caution as a threshold-setting method. Finally, using stability of results as a performance criterion, observed prevalence and average probability were also found to be acceptable threshold optimization methods, thus supporting the conclusions by Liu et al. (2005). Observed prevalence and average probability as thresholds were also grouped close to the accuracy-based thresholds in the clustering dendrogram, which further supports their application (Fig. 2).

To be able to use outputs of BEM for decision-making purposes, modellers have to be confident that the results are as reliable as possible. I have to understand and minimize all sources of error and uncertainty in order to maximize the usefulness of models (Barry and Elith 2006; Heikkinen et al., 2006). Threshold selection complicates comparisons of projections from multiple studies. Even if most BEM studies choose one threshold to transform the modelled probability values, this choice still can generate biases in the resulting projections. Thresholds also complicate studies that compare projections from several modelling methods, since differences between models can actually be due to differences in threshold performance.

In this study, I have documented that threshold selection adds a large component of uncertainty when modelling changes in species distributions under climate change, and should be chosen cautiously. Several ways to adapt threshold methods to the modelling aims have been proposed, such as when modelling for conservation purposes (Freeman and Moisen, 2008) or studying invasion risk (Hartley et al., 2006). For climate change modelling it has been suggested to avoid applying a threshold and instead to directly analyse projected probability values (Araújo et al., 2002). Hopefully advances to more robust predictions that minimize all types of uncertainty can be made by further developing threshold methods and mechanistic modelling methods.

Acknowledgements

Species distributions data was kindly supplied by Raino Lampinen. Thanks to Mats Björklund for supporting this project. I would also like to thank all the staff at IBG – you know who you are, and you know who I am. Much grateful for Ann Gurell and Stina Weststrand for proofreading the report. I would like to thank everyone in the Biodiversity and Evolutionary Biology laboratory at the Natural History Museum in Madrid, Spain and the Rui Nabeiro Biodiversity Chair in Évora, Portugal for all the help and simulating discussions. Thanks to Miguel Araújo for help with writing and for teaching me about the scientific process. Finally I would like my parents for always supporting me.

References

- Akçakaya, H. et al., 2006. Use and misuse of the IUCN Red List Criteria in projecting climate change impacts on biodiversity. *Global Change Biol.* 12, 2037-2043.
- Allouche, O., Tsoar, A., Kadmon R., 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *J. Appl. Ecol.* 43, 1223-1232.
- Anderson, B.J, Akçakaya, H.R., Araújo, M.B., Fordham, D.A., Martinez-Meyer, E., Thuiller, W., Brook, B.W., 2009. Dynamics of range margins for metapopulations under climate change. *Proc. Roy. Soc. Lond. B Biol.* 276, 1415-1420.
- Araújo, M.B., Williams, P.H., Fuller, R.J., 2002. Dynamics of extinction and the selection of nature reserves. *Proc. Roy. Soc. Lond. B Biol.* 269, 1971-1980.
- Araújo, M.B. et al., 2005a. Validation of species-climate impact models under climate change. *Global Change Biol.* 11, 1504-1513.
- Araújo, M.B. et al., 2005b. Reducing uncertainty in projections of extinction risk from climate change. *Glob. Ecol. Biogeogr.* 14, 529-538.
- Araújo, M.B, Rahbek, C., 2006. Ecology: How Does Climate Change Affect Biodiversity? *Science* 313, 1396-1397.
- Araújo, M.B., Thuiller, W. Pearson, R., 2006. Climate warming and the decline of amphibians and reptiles in Europe. *J. Biogeogr.* 33, 1712-1728.
- Araújo, M. B., Luoto, M., 2007. The importance of biotic interactions for modelling species distributions under climate change. *Glob. Ecol. Biogeogr.* 16, 743-753.
- Araújo, M.B., Williams, P.H., 2000. Selecting areas for species persistence using occurrence data. *Biol. Cons.* 96, 331-345.
- Araújo, M.B., New, M., 2007. Ensemble forecasting of species distributions. *Trends Ecol. Evol.* 22, 42-47.
- Araújo, M.B. Peterson, A.T. Uses and misuses of bioclimatic envelope modelling. *Trends Ecol. Evol.* In review.
- Barry, S., Elith, J., 2006. Error and uncertainty in habitat models. *J. Appl. Ecol.* 43, 413-423.
- Baselga, A., Araújo, M.B., 2009. Individualistic vs community modelling of species distributions under climate change. *Ecography* 32, 55-65.
- Berry, P.M., Dawson, T.P., Harrison, P.A., Pearson, R.G., 2002. Modelling potential impacts of climate change on the bioclimatic envelope of species in Britain and Ireland. *Glob. Ecol. Biogeogr.* 11, 453-462.
- Brook, B. et al., 2009. Integrating bioclimate with population models to improve forecasts of species extinctions under climate change. *Biol. Lett.* 5, 723-725.
- Buckley, L.B., Urban, M.C., Angilletta, M.J., Crozier, L.G., Rissler, L.J. Sears, M.W., 2010. Can mechanism inform species' distribution models? *Ecol. Lett.* 13, 1041-1054.
- Buisson, L. et al., 2010. Uncertainty in ensemble forecasting of species distribution. *Global Change Biol.* 16, 1145-1157.

- Cantor, S. B. et al., 1999. A comparison of C/B ratios from studies using receiver operating characteristic curve analysis. *J. Clin. Epidemiol.* 52: 885-892.
- Clarke, K.R., Green, R.H., 1988. Statistical design and analysis for a 'biological effects' study. *Mar. Ecol. Prog. Ser.* 46, 213–226.
- Cramer, J. S., 2003. *Logit models: from economics and other fields.* Cambridge University Press.
- Diniz-Filho, J. et al., 2009. Partitioning and mapping uncertainties in ensembles of forecasts of species turnover under climate change. *Ecography* 32, 897-906.
- Dormann, C. F., Purschke, O., García Márquez, J.R., Lautenbach, S., Schröder, B., 2008. Components of uncertainty in species distribution analysis: a case study of the great grey shrike. *Ecology* 89, 3371-3386.
- Elith, J. et al., 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29, 129-151.
- Fielding, A., Haworth P.F., 1995. Testing the generality of bird-habitat models. *Conserv. Biol.* 9, 1466-1481
- Fielding, A., Bell, J., 1997. A review of methods for the assessment of prediction errors in conservation presence/ absence models. *Environ. Conserv.* 24, 38-49.
- Fitzpatrick, M., Hargrove, W., 2009. The projection of species distribution models and the problem of non-analog climate. *Biodivers. Conserv.* 18, 2255-2261.
- Freeman, E.A., Moisen, G.G., 2008. A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa. *Ecol. Model.* 217, 48-58.
- Guisan, A., Thuiller, W., 2005. Predicting species distribution: offering more than simple habitat models. *Ecol. Lett.* 8, 993-1009.
- Hartley, S., Harris, R., Lester, P.J., 2006. Quantifying uncertainty in the potential distribution of an invasive species: climate and the Argentine ant. *Ecol. Lett.*, 9, 1068-1079.
- Heikkinen, R.K., Luoto, M., Araújo, M.B., Virkkala, R., Thuiller, W., Sykes, M.T., 2006. Methods and uncertainties in bioclimatic envelope modelling under climate change. *Prog. Phys. Geog.* 30, 751-777.
- Hijmans, R.J., Graham, C.H., 2006. The ability of climate envelope models to predict the effect of climate change on species distributions. *Global Change Biol.* 12, 2272-2281.
- Huntley, B., Berry, P.M., Cramer, W., McDonald, A.P., 1995. Modelling present and potential future ranges of some European higher plants using climate response surfaces. *J. Biogeogr.* 22, 967-1001.
- IUCN. 2001. *IUCN Red List Categories and Criteria: Version 3.1.* IUCN Species Survival Commission. IUCN, Gland, Switzerland and Cambridge, UK.
- Jalas, J., Suominen, J., 1972. *Atlas Flora Europaeae: distribution of vascular plants in Europe, 1.* The Committee for Mapping the Flora of Europe and Societas Biologica Fennica Vanamo, Helsinki.
- Jalas, J. and Suominen, J. 1973. *Atlas Flora Europaeae: distribution of vascular plants in Europe, 2.* The Committee for Mapping the Flora of Europe and Societas Biologica Fennica Vanamo,

Helsinki.

- Jiménez-Valverde, A., Lobo, J. M., 2007. Threshold criteria for conversion of probability of species presence to either-or presence-absence. *Acta Oecol.* 31, 361–369.
- Keenan, T., Serra, J.M., Bastiansen, F., Ninyerola, M., Sabate, S., 2010. Predicting the future of forests in the Mediterranean under climate change, with niche- and process-based models: CO2 matters! *Global Change Biol.* no. doi: 10.1111/j.1365-2486.2010.02254.x
- Keith, D. A., Akçakaya, H. R., Thuiller, W., Midgley, G.F., Pearson, R.G., Phillips, S.J., Regan, H.M., Araújo, M.B., Rebelo, T.G., 2008. Predicting extinction risks under climate change: coupling stochastic population models with dynamic bioclimatic habitat models. *Biol. Lett.* 4, 560-563.
- Lahti, T., Lampinen, R. 1999. From dot maps to bitmaps Atlas Florae Europaeae goes digital. *Acta Bot. Fenn.* 162, 5-9.
- Legendre, P., Legendre, L., 1998. Numerical ecology. 2nd English edition. Elsevier Science BV, Amsterdam.
- Liu, C. et al., 2005. Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, 28, 385-393.
- Manel, S., Dias, J.M., Buckton, S.T., Ormerod, S.J., 1999. Comparing discriminant analysis, neural networks and logistic regression for predicting species distributions: a case study with a Himalayan river bird. *Ecol. Model.* 120, 337–347.
- Manel, S., Williams, H.C., Ormerod, S.J., 2001. Evaluating presence-absence models in ecology: the need to account for prevalence. *J. Appl. Ecol.* 38, 921-931.
- McPherson, J.M., Jetz, W., Rogers, D.J., 2004. The effects of species' range sizes on the accuracy of distribution models: ecological phenomenon or statistical artefact? *J. Appl. Ecol.* 41, 811-823.
- Mitchell, T.D., Carter, T.R., Jones, P.D., Hulme, M. New, M., 2004. A comprehensive set of high-resolution grids of monthly climate for Europe and the globe: the observed record (1990–2000) and 16 scenarios (2001–2100). Tyndall Centre WorkingPaper. Tyndal Centre for Climate Change Research, Norwich.
- New, M., Hulme, M., Jones, P.D., 2000. Representing twentieth century space–time climate variability. Part 2: Development of 1901–96 monthly grids of terrestrial surface climate. *J. Climate* 13, 2217–2238.
- Oksanen, J, Blanchet, F., Kindt, R., Legendre, P., O'Hara, R. B., Simpson, G.L., Solymos, P., Stevens, M.H. , Wagner, H., 2010. vegan: Community Ecology Package. R package version 1.17-2. <http://CRAN.R-project.org/package=vegan>
- Pearce, J., Ferrier, S., 2000. Evaluating the predictive performance of habitat models developed using logistic regression. *Ecol. Model.* 133, 225–245.
- Pearson, R.G., Dawson, T.P., Liu, C. 2004. Modelling species distributions in Britain: a hierarchical integration of climate and land-cover data. *Ecography* 27, 285-298.
- Pearson, R. et al., 2006. Model-based uncertainty in species range prediction. *J. Biogeogr.* 33, 1704-1711.

- Pearson, R.G., 2007. Species' Distribution Modeling for Conservation Educators and Practitioners. Synthesis. American Museum of Natural History. Available at <http://ncep.amnh.org>.
- Peterson, A.T. et al., 2002. Future projections for Mexican faunas under global climate change scenarios. *Nature* 416, 626-629.
- Randin, C.F., Dirnbock, T., Dullinger, S., Zimmermann, N.E., Zappa, M., Guisan, A., 2006. Are niche-based species distribution models transferable in space? *J. Biogeogr.* 33, 1689-1703.
- R Development Core Team 2010. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- Schapire, R. E., Singer, Y., Singhal, A. 1998. Boosting and Rocchio applied to text filtering. *Proc. ACM SIGIR*, 215-223.
- Schröter, D., Cramer, W., Leemans, R., Prentice, I.C. Araújo, M.B., Arnell, N.W., Bondeau, A., et al., 2005. Ecosystem service supply and vulnerability to global change in Europe. *Science* 310, 1333-7.
- Segurado, P., Araújo, M.B., 2004. An evaluation of methods for modelling species distributions. *J. Biogeogr.* 31, 1555-1568.
- Stockwell, D. R. B., Peterson, A.T., 2002. Effects of sample size on accuracy of species distribution models. *Ecol. Model.* 148, 1-13.
- Svenning, J., Normand, S., Kageyama, M., 2008. Glacial refugia of temperate trees in Europe: insights from species distribution modelling. *J. Ecol.* 96, 1117-1127.
- Sykes, M.T., Prentice, I.C., Cramer, W., 1996. A bioclimatic model for the potential distributions of north European tree species under present and future climates. *J. Biogeogr.* 23, 203-233.
- Thuiller, W., 2004. Patterns and uncertainties of species' range shifts under climate change. *Global Change Biol.* 10, 2020-2027.
- Thuiller, W., Araújo, M.B., Lavorel, S., 2003. Generalized models vs. classification tree analysis: Predicting spatial distributions of plant species at different scales. *J. Veg. Sci.* 14, 699-680.
- Thuiller W., Brotons L., Araújo M.B., 2004. Effects of restricting environmental range of data to project current and future species distributions. *Ecography* 27, 165–172.
- Thuiller, W., Lavorel, S., Araújo, M.B., Sykes, M.T. et al., 2005. Climate change threats to plant diversity in Europe. *Proc. Natl. Acad. Sci. U.S.A.* 102, 8245-8250.
- Thuiller, W., Lafourcade, B., Engler, R., Araújo, M.B., 2009. BIOMOD - a platform for ensemble forecasting of species distributions. *Ecography* 32, 369-373.
- Thomas, C.D., A. Cameron, R.E. Green, M. Bakkenes, L.J. Beaumont, Y.C. Collingham, B.F.N. Erasmus, et al., 2004. Extinction risk from climate change. *Nature* 427, 145-8.
- Williams, P. H. et al., 2000. Endemism and important areas for conserving European biodiversity: a preliminary exploration of Atlas data for plants and terrestrial vertebrates. *Belgian J. Entomol.* 2, 21-46.
- Woodward, F.I., 1987. *Climate and plant distribution*. Cambridge University Press, Cambridge.

Supplementary Data

List of species modelled.

1. *Abies alba* Mill.
2. *Abies borisii-regis* Mattf.
3. *Alnus cordata* (Loisel.) Loisel.
4. *Alnus incana* (L.) Moench subsp. *incana*
5. *Alnus incana* (L.) Moench subsp. *kolaensis* (N.I.Orlova) A.Löve & D.Löve
6. *Alnus viridis* (Chaix) DC.
7. *Betula humilis* Schrank
8. *Betula nana* L.
9. *Betula pendula* Roth
10. *Betula pubescens* Ehrh.
11. *Carpinus betulus* L.
12. *Carpinus orientalis* Mill.
13. *Castanea sativa* Mill.
14. *Celtis australis* L.
15. *Corylus avellana* L.
16. *Corylus colurna* L.
17. *Fagus orientalis* Lipsky
18. *Fagus sylvatica* L. subsp. *orientalis* (Lipsky) Greuter & Burdet
19. *Fagus sylvatica* L. subsp. *Sylvatica*
20. *Ficus carica* L.
21. *Juglans regia* L.
22. *Juniperus communis* L.
23. *Juniperus foetidissima* Willd.
24. *Juniperus oxycedrus* L. subsp. *macrocarpa* (Sm.) Ball
25. *Juniperus oxycedrus* L. subsp. *oxycedrus*
26. *Juniperus phoenicea* L.
27. *Juniperus sabina* L.
28. *Larix decidua* Mill.
29. *Myrica gale* L.
30. *Ostrya carpinifolia* Scop.
31. *Picea abies* (L.) H.Karst. subsp. *abies*
32. *Picea abies* (L.) H.Karst. subsp. *alpestris* (Brügger) Domin
33. *Picea abies* (L.) H.Karst. subsp. *obovata* (Ledeb.) Hultén
34. *Pinus cembra* L.
35. *Pinus halepensis* Mill.
36. *Pinus heldreichii* H.Christ
37. *Pinus mugo* Turra
38. *Pinus nigra* J.F.Arnold subsp. *nigra*
39. *Pinus nigra* J.F.Arnold subsp. *pallasiana* (Lamb.) Holmboe
40. *Pinus nigra* J.F.Arnold subsp. *salzmannii* (Dunal) Franco
41. *Pinus pinaster* Aiton
42. *Pinus pinea* L.

43. *Pinus rotundata* Link
44. *Pinus sylvestris* L.
45. *Pinus uliginosa* Neumann
46. *Pinus uncinata* Mill. ex Mirb.
47. *Populus alba* L.
48. *Populus canescens* (Aiton) Sm.
49. *Populus nigra* L.
50. *Populus tremula* L.
51. *Quercus cerris* L.
52. *Quercus coccifera* L.
53. *Quercus crenata* Lam.
54. *Quercus dalechampii* Ten.
55. *Quercus faginea* Lam.
56. *Quercus frainetto* Ten.
57. *Quercus ilex* L.
58. *Quercus macrolepis* Kotschy
59. *Quercus pedunculiflora* K.Koch
60. *Quercus petraea* (Matt.) Liebl.
61. *Quercus pubescens* Willd. subsp. *anatolica* O.Schwarz
62. *Quercus pubescens* Willd. subsp. *pubescens*
63. *Quercus pyrenaica* Willd.
64. *Quercus robur* L.
65. *Quercus rotundifolia* Lam.
66. *Quercus suber* L.
67. *Quercus trojana* Webb
68. *Salix alba* L.
69. *Salix alpina* Scop.
70. *Salix amplexicaulis* Bory
71. *Salix appendiculata* Vill.
72. *Salix arbuscula* L.
73. *Salix atrocinerea* Brot.
74. *Salix aurita* L.
75. *Salix breviserrata* Flod.
76. *Salix burjatica* Nasarov
77. *Salix caesia* Vill.
78. *Salix caprea* L.
79. *Salix cinerea* L.
80. *Salix daphnoides* Vill.
81. *Salix eleagnos* Scop.
82. *Salix foetida* Schleich. ex DC. in Lam. & DC.
83. *Salix fragilis* L.
84. *Salix hastata* L.
85. *Salix glabra* Scop.
86. *Salix glauca* L.
87. *Salix glaucosericea* Flod.
88. *Salix hegetschweileri* Heer

89. *Salix helvetica* Vill
90. *Salix herbacea* L.
91. *Salix lanata* L.
92. *Salix lapponum* L.
93. *Salix myrsinifolia* Salisb.
94. *Salix myrsinites* L.
95. *Salix myrtilloides* L.
96. *Salix pedicellata* Desf.
97. *Salix pentandra* L.
98. *Salix phylicifolia* L.
99. *Salix polaris* Wahlenb.
100. *Salix purpurea* L.
101. *Salix repens* L. subsp. *arenaria* (L.) Hiitonen
102. *Salix repens* L. subsp. *repens*
103. *Salix reticulata* L.
104. *Salix retusa* L.
105. *Salix rosmarinifolia* L.
106. *Salix salviifolia* Brot.
107. *Salix serpillifolia* Scop.
108. *Salix silesiaca* Willd.
109. *Salix starkeana* Willd.
110. *Salix triandra* L.
111. *Salix viminalis* L.
112. *Salix waldsteiniana* Willd.
113. *Taxus baccata* L.
114. *Ulmus glabra* Huds.
115. *Ulmus laevis* Pall.
116. *Ulmus minor* Mill.