Ecological Monographs, 77(4), 2007, pp. 615–630
© 2007 by the Ecological Society of America

WHAT MATTERS FOR PREDICTING THE OCCURRENCES OF TREES: TECHNIQUES, DATA, OR SPECIES' CHARACTERISTICS?

A. Guisan,[1,7] N. E. Zimmermann,[2] J. Elith,[3] C. H. Graham,[4] S. Phillips,[5] and A. T. Peterson[6]

[1]*Department of Ecology and Evolution, University of Lausanne, CH-1015 Lausanne, Switzerland*
[2]*Swiss Federal Research Institute WSL, Zürcherstrasse 111, CH-8903 Birmensdorf, Switzerland*
[3]*School of Botany, University of Melbourne, Parkville, Victoria, 3010 Australia*
[4]*Department of Ecology and Evolution, 650 Life Sciences Building, Stony Brook University, New York 11794 USA*
[5]*AT&T Labs-Research, Florham Park, New Jersey 07932 USA*
[6]*Natural History Museum and Biodiversity Research Center, University of Kansas, Lawrence, Kansas 66045 USA*

Abstract. Data characteristics and species traits are expected to influence the accuracy with which species' distributions can be modeled and predicted. We compare 10 modeling techniques in terms of predictive power and sensitivity to location error, change in map resolution, and sample size, and assess whether some species traits can explain variation in model performance. We focused on 30 native tree species in Switzerland and used presence-only data to model current distribution, which we evaluated against independent presence–absence data. While there are important differences between the predictive performance of modeling methods, the variance in model performance is greater among species than among techniques. Within the range of data perturbations in this study, some extrinsic parameters of data affect model performance more than others: location error and sample size reduced performance of many techniques, whereas grain had little effect on most techniques. No technique can rescue species that are difficult to predict. The predictive power of species-distribution models can partly be predicted from a series of species characteristics and traits based on growth rate, elevational distribution range, and maximum elevation. Slow-growing species or species with narrow and specialized niches tend to be better modeled. The Swiss presence-only tree data produce models that are reliable enough to be useful in planning and management applications.

Key words: data treatment; grain size; location error; model performance; niche-based modeling; sample size; species traits; Switzerland native tree species; tree occurrences.

Introduction

The various human-induced threats imposed on nature have recently reinvigorated the study of species' distributions, and more particularly the use of predictive models to describe and quantify them (Guisan and Thuiller 2005, Guisan et al. 2006). These empirical models relate known occurrences of species to environmental predictors using statistically derived response curves that aim to best reflect the species' environmental responses. The fitted model is then used to project the niche into geographic space, providing a spatial prediction of the most suitable/unsuitable areas for a given species.

Presence-only data currently represent the largest—and in many cases the only—source of information on species' distributions (Graham et al. 2004). These data, commonly compiled from natural-history collections in museums, herbaria, and national biological-data record centers, are often recorded without a sampling strategy, and thus contain sampling bias (Stockwell and Peterson 2002*a*, Kadmon et al. 2004). Depending on the species, date of collection, and geographic location, the amount of data and its spatial accuracy can also be highly variable (Hijmans et al. 2000, Loiselle et al. 2003, Graham et al. 2004). Bias and small sample sizes, in particular, have been shown to weaken model performance (Stockwell and Peterson 2002*a*, *b*, Kadmon et al. 2003, 2004). While these characteristics pose many technical problems when modeling species' distributions (Hirzel et al. 2002, Brotons et al. 2004, Engler et al. 2004, Pearce and Boyce 2006), the importance of presence-only records as a data source demands ongoing attention to determine the best methods for modeling them.

A large range of techniques now exists to predict species distributions (Guisan and Thuiller 2005, Elith et al. 2006). Some techniques were specifically designed for modeling presence-only data (Busby 1991, Walker and Cocks 1991, Carpenter et al. 1993, Hirzel et al. 2002). Alternatively, pseudo-absences can supplement presence-only records and then presence–absence techniques can be used (Manly et al. 1993, Stockwell and Peters 1999, Zaniewski et al. 2002, Engler et al. 2004).

Recently, a large international modeling experiment—the largest comparison of techniques conducted to date—assessed the reliability of 16 different modeling

Manuscript received 22 June 2006; revised 14 March 2007; accepted 17 April 2007; final version received 24 May 2007. Corresponding Editor: K. D. Woods.
[7] E-mail: antoine.guisan@unil.ch

techniques to predict species' occurrence with six data sets from study areas around the world (Elith et al. 2006). Models were trained with a set of presence-only records from natural-history collections (NHC; Graham et al. 2004) and tested with an independent, generally more accurate, set of observations that include both presence and absence data, thus providing a measure of the predictive power of each technique. Novel techniques, such as boosted regression trees or maximum entropy models, consistently outperformed more established techniques, such as simple rectilinear envelopes (e.g., BIOCLIM) or regression-based approaches (e.g., generalized additive models). However, several extrinsic properties of occurrence data, such as the choice of mapping resolution (grain size), the choice of predictors, location error associated with species' observations, or sample size, might also affect model performance, possibly unequally across different techniques.

Most studies show that it is difficult to create accurate models for all species regardless of the technique used. This suggests that model performance may also be influenced by species' intrinsic ecological characteristics, such as commonness (Karl et al. 2002, Kadmon et al. 2003), range size (Schwartz et al. 2006), niche width, and niche position (Boone and Krohn 2002, Kadmon et al. 2003, Thuiller et al. 2004, Hernandez et al. 2006). Other traits, such as successional status, light requirements, leaf longevity, or drought tolerance may also influence predictive performance.

Determining how species' characteristics influence model performance is particularly important because it could provide the means to use theoretical or expert knowledge to predict which species are suitable for modeling. For example, in plants one might expect models for species of late-succession stages to perform better and more consistently across techniques than those for pioneer species. Our rationale is that the realized niches of species are likely to be better quantified when geographic distributions are stable, as should be expected for dominant late-successional species. In contrast, distributions of early-successional species are likely to be conditioned by stochastic dynamics of gap openings and disturbances, which might make it difficult to determine their environmental requirements and to fit empirical models.

In this paper, we assess:

1) The ability of 10 modeling techniques to predict current Swiss tree distributions from presence-only data, and particularly their robustness to three data treatments: addition of location error, changes in resolution and changes in sample size;

2) Whether some species characteristics (traits) can explain variation in predictive success; and

3) The respective importance of data quality, species identity, or modeling technique to determine predictive success.

We used the Swiss tree data set used in Elith et al. (2006) because: (a) it contains thousands of evaluation sites and is therefore large enough to provide sufficient power to discriminate between modeling methods; (b) it includes environmental variables that are functionally relevant and likely to provide models with enough explanatory power to compare across ecological traits; (c) it provided a range of model performances, including some of the best-predicted species; and (d) detailed ecological information is available for all species.

Methods

Species data

Study species.—We used 30 tree species commonly found in Switzerland, representing various sample sizes for modeling (Table 1; exemplar maps in Fig. 1) and diverse ecological traits. Species selection included all native tree species that had at least 35 occurrences in both data sets. Compared with other data sets in Elith et al. (2006), the Swiss data set is the most comprehensive in terms of species records in both the training (presence-only) and test (presence–absence) sets. Species' traits (Table 2) were drawn from various sources: (a) species parameters from FORCLIM (Bugman 1996) and TREEMIG (Lischke et al. 2006) forest gap dynamic models; (b) forest plots database (Wohlgemuth 1992); and (iii) expert knowledge (A. Guisan and N. E. Zimmermann, *personal observations*).

Training data set.—The *Forest Plots* (FP) database contains irregularly and nonsystematically sampled forest vegetation relevés (inventories of all species within given plots) throughout Switzerland. The individual observers had their own particular sampling design or had no design. Hence, this data set has some similar properties to natural-history collection data (Graham et al. 2004) except that, here, exhaustive lists of species are available for each relevé, meaning that absences are fully recorded, and there is a substantial number of records for a relatively small area. The absences were ignored in our present study, and only presence observations were used for fitting the models. We reduced the data set so that only one record was allowed per 100-m grid cell, the highest resolution of the environmental maps. Absolute number of occurrences for modeling ranged from 36 (e.g., *Populus nigra*, *Salix alba*) to 5822 (*Picea abies*) occurrences, with many values in the range 100–1000 (first and third quartiles of 119 and 855 records, respectively).

Test data set.—The *Forest Inventory* (FI) is a regular 1-km lattice across Switzerland (Brassel and Braendli 1999). Approximately 25% of Switzerland is covered by forest, resulting in roughly 10 600 forest-inventory points. We used information on the presence and absence of tree species from the 1985 inventory. Subsets of this data set were also used to fit models in the "error" experiment (see *Modeling experiments*, below). The data were cleaned to allow only one sample per 100-m grid square and no sites were allowed in the same cells as the training data. In the rare cases with more than one record per cell, a presence for a species was recorded if

TABLE 1. Identification of and information about the 30 modeled tree species.

Trees			Species used in data experiment	No. presence records		Mean AUC value across techniques‡	pAUC§
Code	Genus	Species		Training data set	Test data set†		
abialb	*Abies*	*alba*		3357	3326	0.730	useful
acecam	*Acer*	*campestre*		710	119	0.841	useful
acepla	*Acer*	*platanoides*	x	482	107	0.795	useful
acepse	*Acer*	*pseudoplatanus*	x	2800	1520	0.686	poor
alnglu	*Alnus*	*glutinosa*		297	134	0.771	useful
alninc	*Alnus*	*incana*		293	278	0.592	poor
betpen	*Betula*	*pendula*	x	468	391	0.699	poor
carbet	*Carpinus*	*betulus*		857	222	0.897	useful
cassat	*Castanea*	*sativa*		458	308	0.957	good
fagsyl	*Fagus*	*sylvatica*	x	5528	4246	0.781	useful
fraexc	*Fraxinus*	*excelsior*		2830	1366	0.730	useful
lardec	*Larix*	*decidua*	x	986	1493	0.725	useful
ostcar	*Ostrya*	*carpinifolia*		91	35	0.969	good
picabi	*Picea*	*abies*		5822	6953	0.702	useful
pincem	*Pinus*	*cembra*	x	279	238	0.955	good
pinmug	*Pinus*	*mugo*		89	22	0.795	useful
pinsyl	*Pinus*	*sylvestris*		2142	978	0.771	useful
pinunc	*Pinus*	*uncinata*	x	291	142	0.654	poor
popnig	*Populus*	*nigra*		36	19	0.892	useful
poptre	*Populus*	*tremula*	x	154	100	0.648	poor
pruavi	*Prunus*	*avium*		613	271	0.740	useful
quepet	*Quercus*	*petraea*		1452	477	0.834	useful
quepub	*Quercus*	*pubescens*		382	26	0.889	useful
querob	*Quercus*	*robur*		734	395	0.829	useful
salalb	*Salix*	*alba*		37	20	0.6340	poor
sorari	*Sorbus*	*aria*	x	1245	298	0.7182	useful
sorauc	*Sorbus*	*aucuparia*		426	224	0.7032	useful
tilcor	*Tilia*	*cordata*		560	182	0.8235	useful
tilpla	*Tilia*	*platiphyllos*	x	749	104	0.8226	useful
ulmgla	*Ulmus*	*glabra*		937	306	0.7525	useful

† Out of a total of 10 013 sites.
‡ AUC = area under the curve.
§ "pAUC" represents the classes of performance according to Swets (1988).

any of the records was a presence. This resulted in a final set of 10 013 test sites. Absences are unlikely to be a result of forest management because Swiss forest management relies on fostering natural regeneration from local seed sources, so species are unlikely to be eliminated from stands within their natural range.

Pseudo-absences.—As several modeling techniques required records of absences, 10 000 pseudo-absences were generated at random across the whole country. For many techniques, these were weighted when fitting the models to give them a similar weight as for the sum of presences (mimicking prevalence = 0.5, but not implying any a priori knowledge).

Environmental Predictors

We used 13 spatially explicit environmental variables reflecting ecologically meaningful information to predict tree distributions (Table 3). Among these, seven climatic variables and two topographic variables were derived from spatial climate mapping and GIS modeling, following Zimmermann and Kienast (1999). Climatic predictors were chosen to reflect physiological needs and included: growing degree-days (DDEG), average temperature of the coldest month (TAV), the sum of frost events during the frost-sensitive time (SFRO), the yearly sum of all monthly precipitation (PREC), the number of summer (June–August) precipitation days with rainfall >1 mm (PDSUM), the site water balance (SWB; also integrating soil parameters) and the potential yearly global radiation (SRAD; daily clear-sky average). All basic climatic variables (temperature, precipitation) were based on 1961–1990 monthly normals. One hundred eighty stations were available for temperature and 360 for precipitation. The two topographic predictors, were: slope (SLOPE), summarizing the intensity of gravitational processes, and a relative index of topographic position (TOPO; a multi-scale measure of convexity/concavity), ranging from positive values expressing ridges, peaks, and exposed sites to negative values, describing sinks, gullies, valleys or toe slopes. The two variables reflecting the type of substratum were: nutrient availability (NUTRI), derived from the Swiss soil suitability map (Anonymous 1980), and the presence of strictly calcareous bedrock type (CALC), reclassified from the Swiss Geotechnical map. Finally, two variables related to leaf type were derived from a Landsat TM (thematic mapper)-based forest composition map (GEOSTAT; BFS 2001). BCC is a continuous measure of the broadleaved tree cover, while CCC represents the cover fraction of needle-leaf trees. All predictors,

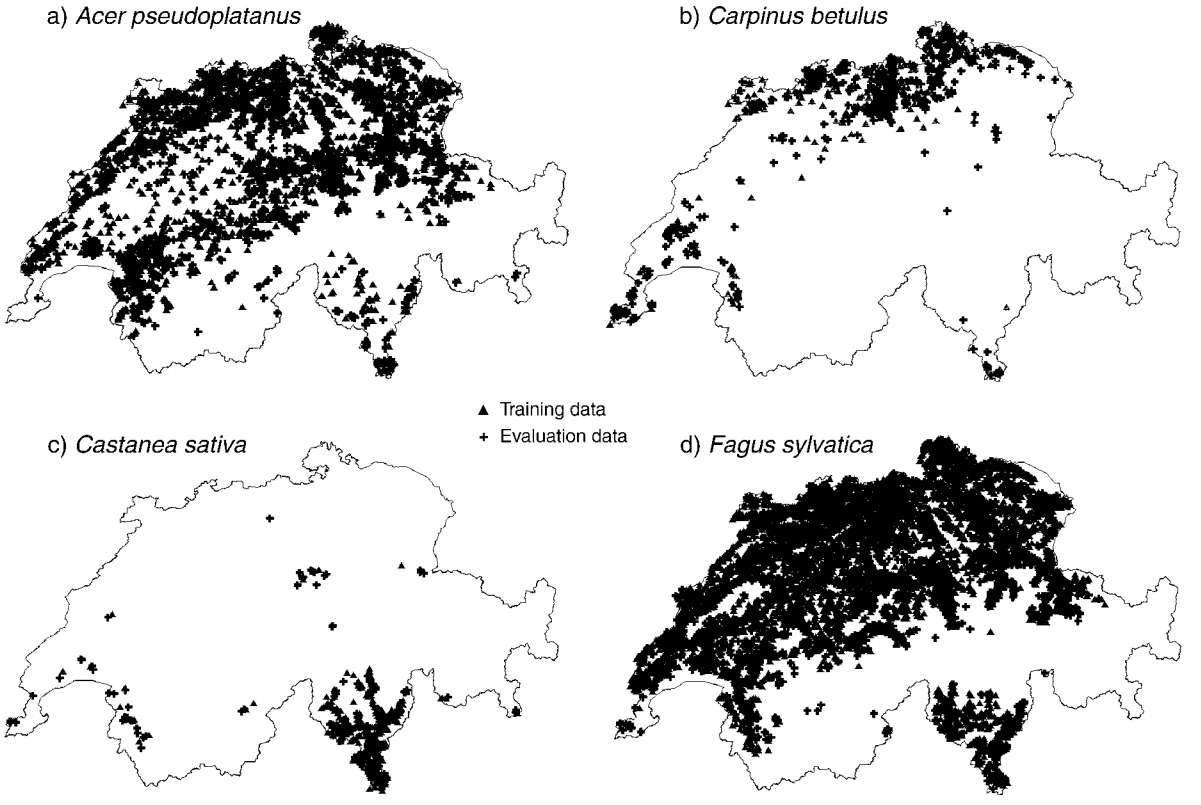

FIG. 1. Distribution maps for four of the 30 modeled Swiss tree species: (a) *Acer pseudoplatanus*, (b) *Carpinus betulus*, (c) *Castanea sativa*, and (d) *Fagus sylvatica*.

originally available at 25 m-resolution, were resampled to a 100 m resolution (1000 m for the grain experiment) using the "aggregate" function in ArcGIS (ESRI 2005). Additional details for each predictor are in Appendix A.

Modeling techniques

We compared 10 single-species modeling techniques (Appendix B): BIOCLIM, BRUTO, BRT, DOMAIN, GDMSS GAM, GLM, MAXENT, MARS, and OM-

TABLE 2. Biological and ecological traits for the 30 tree species modeled.

Traits		Variable type	Details	Classes in Fig. 6	
Code	Description [source]			Lower	Upper
LLON	leaf longevity [1]	2 classes		evergreen	deciduous
MGR	maximum growth rate [1]	continuous	range 82–310; 10× (mm/year)	<167	≥167
DDGM	minimum degree-days for growth [1]	continuous	range 323–1339	<898	≥898
DRTOL	drought tolerance [1]	5 classes	range 1–5; 1 = weak, 5 = high	<3	≥3
LREQ	light requirement of adult E7smp;trees [1]	9 classes	range 1–9; 1 = low, 9 = high	<5	≥5
PREV	prevalence (relative sample size) [2]	continuous	absolute number; max is 10 610	<270	≥270
ELVMIN	minimum elevation [2]	continuous	range 200–1400; observed from the data (empirical)	<250	≥250
ELVMAX	maximum elevation [2]	continuous	range 650–2300; observed from the data (empirical)	<1550	≥1550
ELVRNG	elevation range [2]	continuous	range 450–2000; difference between max and min	<1275	≥1275
SUCC	successional status [3]	2 classes		early (pioneer)	late
HUMAN	human influence [3]	2 classes		planted frequently	never planted
AVABU50	relative mean basal area (BA) [2]	2 classes		larger than median BA	smaller than median BA

Notes: The "lower" and "upper" classes of each trait are represented in Fig. 6 by "–" and "+," respectively. For continuous variables the boundary between the two classes is the median. Sources: [1], species parameters of FORCLIM/TreeMig (Bugmann 1996, Lischke et al. 2006); [2], derived from the Swiss forest plots database (Wohlgemuth 1992); [3], expert knowledge.

TABLE 3. Predictor variables used for the modeling of tree species using the 10 predictive techniques.

Code	Variable type†	Description (units)
Climatic‡		
DDEG	C	growing degree-days above the threshold of 0°C (°C × days)
TAV	C	average temperature of the coldest month (°C)
SFRO	C	summer frost frequency (no. days)
PREC	C	average yearly precipitation sum (mm)
PDSUM	C	number of days with rainfall >1 mm (no. days)
SWB	C	site water balance (mm)
SRAD	C	potential yearly global radiation (daily average) ($kJ \cdot m^{-2} \cdot d^{-1}$)
Topographic§		
SLOPE	C	slope (degrees)
TOPO	C	topographic position (range)
Substratum		
NUTRI	C	soil nutrients (index between 0 and 45) (milival/cm^2)||
CALC	B bedrock strictly calcareous vs. other type (present or absent)	
Leaf type		
BCC	C	broadleaved continuous cover (TM-based classification) (%)
CCC	C	coniferous continuous cover (TM-based classification) (%)

Notes: The 13 spatially explicit environmental variables reflect ecologically meaningful information to predict tree distributions. See Appendix A for details on their calculation. Summer = June–August.

† C = continuous, B = binary.

‡ Chosen to reflect physiological needs. Site water balance integrates soil parameters. Temperature and precipitation were based on 30-year average (1961–1990 monthly normals).

§ Slope summarizes the intensity of gravitational processes. TOPO is a relative index of convexity–concavity.

|| Explanation of units: 1 milival/cm^2 = 1 mmol/cm^2 if the valence is ±1.

GARP. Some techniques, like BIOCLIM and DOMAIN, only need presence data and can be called "profile techniques," whereas others, like GAM or BRT, require absences and can be called "group discrimination techniques" (Robertson et al. 2001).

BIOCLIM fits simple rectilinear envelopes and was applied within DIVA-GIS (Hijmans et al. 2001), an alternative implementation to the original climate-based method by Nix (1986). DOMAIN (Carpenter et al. 1993) differs from BIOCLIM by using Gower metric to define point-wise similarities rather than outlining envelopes. OM-GARP is an improved, recently released version of the GARP package (Stockwell and Peters 1999) that uses genetic algorithms to select a set of rules that best describes the species' distribution. Rule types include range rules akin to BIOCLIM models, negated range rules (i.e., *not* within a given range), and "atomic" rules (i.e., specifying single values as opposed to ranges), as well as a simple logistic regression analogue.

All remaining techniques generally require presence and absence data to model probabilities of occurrence, but here were implemented with pseudo-absences instead of true absences. GAM (generalized additive models), GLM (generalized linear models), MARS (multivariate adaptive regression splines), and BRUTO are all forms of generalized multiple regressions, with differences in the fitting procedures and the methods used to model complex responses. GAMs are similar to GLMs, with nonparametric smoothers replacing the parametric functions in a GLM. MARS is similar to GAM but relies on fitting piecewise linear basis functions rather than smoothed functions (Leathwick et al. 2005) , and includes a recursive simplification procedure. This makes it faster than a GAM and simpler to transfer to a GIS. BRUTO is a generalized additive model with a different and faster fitting procedure (Leathwick et al. 2006*a*).

BRT (boosted regression trees) comprises two algorithms (boosting and regression trees) and computes a sequence of simple regression trees, where each successive tree is fitted to the residuals of the existing set of trees (Leathwick et al. 2006*b*). By allowing trees with more than one node, interactions can be modeled automatically. BRT can be prone to overfitting, but this was controlled here by using cross-validation to select the final model.

MAXENT is a maximum-entropy approach recently applied to predict species distributions (Phillips et al. 2006). The general idea behind MAXENT is to find a probability distribution, defined over the study area, that satisfies a set of constraints derived from the occurrence data. Each constraint requires that the expected value of an environmental variable (or function thereof) must be within a confidence interval of its empirical mean (the mean over the presences). Among distributions that satisfy the constraints, MAXENT chooses the one that maximizes entropy, i.e., is the closest to uniform, as any other choice would represent constraints on the distribution that are not justified by the data.

GDMSS is the single-species implementation of the generalized dissimilarity modeling approach usually applied in a community context (Ferrier et al., *in press*). It differs from the community version by only using data from one species to generate the dissimilarities and to transform the environmental space More detail and specifics of settings used are published elsewhere (Elith et al. 2006).

Modeling experiments

As part of a prior modeling exercise (Elith et al. 2006), we conducted experiments on a subset of 10 species per region to compare the robustness of the various techniques to changes in parameters. Here we bring together the Swiss results for manipulations of three parameters—location error, sample size, and grain size—to compare model responses to data perturbation. Nine of the 10 techniques reported here were applied (excluding GDMSS), and some experiments had more than one treatment.

The aim of the first experiment, the error experiment (E), was to mimic locational error in presence-only data, and to assess how error affected predictive performance (see also C. H. Graham, J. Elith, R. Hijmans, A. Guisan, A. T. Peterson, B. A. Loiselle, and the NCEAS species distribution group, *unpublished manuscript*). To control for data quality we used only the higher quality test data set (*forest inventory*, FI) for this experiment. We used 21% of presence records for modeling, and the remaining 79% for evaluation, along with all the absence records. Error was added to the data used for modeling by shifting the coordinates by an amount sampled from a normal distribution with a mean of zero and a standard deviation of 5 km (this resulted in shifted values ranging between 6 m and 16 km for easting, and between 9 m and 17 km for northing). The shifts were constrained so that the sites did not move outside of the region. Model performance based on these shifted data was compared to performance on the same data, unshifted, and both sets of models were evaluated using the independent evaluation data set (i.e., the 79% of presence records that were set aside).

In the remaining two experiments, grain size and sample size, models were fit with the training data (forest plots, FP) and tested with the full evaluation set (FI).

The *grain-size* experiment (G) investigated the effect of grain size of environmental data on model performance (see also Guisan et al. 2007). Grain size was enlarged by a factor of 10, changing pixel size for environmental predictors from 100 m to 1 km. All environmental and species data were aggregated accordingly in a GIS to build models at the coarser resolution. At the finer grain size, there was a maximum of one species record per cell. At the coarse grain size there were sometimes multiple records per cell, but overall sample size remained the same.

The *sample-size* experiment (S) investigated the effect of sample size on predictive performance. From all available records, two smaller sample sizes of 10 and 30 occurrences were created, using random sampling.

Model evaluation and comparison

A strength of this study is that models were fitted using presence-only data (FP data set) but evaluated using presences and absences (FI data set), thus ensuring proper independence when testing their predictive capacity. We used the area under the curve (AUC) of a receiver-operating characteristic plot (ROC-plot; Swets 1988, Fielding and Bell 1997) as the evaluation measure, because it is threshold independent. AUC values were interpreted on the scale proposed by Swets (1988): AUC $>$ 0.90: good; 0.90 $>$ AUC $>$ 0.70: useful; AUC $<$ 0.70: poor. As the choice of the evaluation metrics may influence the results, we also tested model predictions with a second measure, the maximized kappa (Manel et al. 2001). Because max-kappa provided similar results we present only the AUC results here.

Analyses of model performance

Variances of techniques across species were compared to variances of species across techniques. To assess whether some techniques are consistently better than others, we found the best-performing technique for each species. For each technique, we then compared its performance across all species to the vector of best performance (Wilcoxon test), and plotted the corresponding P values, which provide an index of the deviation from the best performance. Variation in AUC across the 10 techniques was also tested using a generalized linear mixed model (GLMM) with the AUC as the response (Elith et al. 2006). Technique was fitted as a fixed effect, and Species as a random effect. Analyses were performed using WinBUGS (Spiegelhalter et al. 2003) and the 50 000 iterations were summarized in terms of the percentage of runs where the AUC for method A was greater than that for method B. More details on the GLMM procedure are given in Appendix C.

To visualize results from the experimental data manipulations, boxplots by techniques were drawn and pairwise comparisons (Wilcoxon tests) were conducted to assess the respective robustness of each technique against error, change in resolution, and sample size.

To explore whether individual species' traits could explain variation in model performance, we fitted a generalized linear model (GLM), with Gaussian family and identity link, relating model performance (AUC) to multiple species traits. Predictors were selected by AIC-based stagewise selection procedure. No quadratic or higher polynomial terms were included, as no unimodal relationship was expected a priori.

Results

Performance of the tree-distribution models varied significantly more within techniques across species (0.10297 $\pm$ 0.00907 [mean $\pm$ SD]) than within species across techniques (0.04191 $\pm$ 0.02028; with both t or Wilcoxon signed-ranks test, $P \sim 0$). Even so, the effect of modeling technique was statistically important, because the deviance information criterion (DIC) in the generalized linear mixed model (GLMM) changed by 192 (full model compared to a model without an effect for modeling technique), where changes in DIC of $>$10 are considered important (Burnham and Anderson 2002).

Differences can be seen in the resulting prediction maps prepared for four selected species (*Acer pseudoplatanus*, *Carpinus betulus*, *Castanea sativa* and *Fagus sylvatica*) and four techniques (GARP, MARS, GDMSS, and MAXENT; Fig. 2) representing a range of performances. Based on mean performance of species across techniques, 3 species fell into the category of "good" models (*Ostrya carpinifolia*, *Castanea sativa*, *Pinus cembra*), 21 species fell into the category "useful," and 6 species were rated "poor."

Ranking of overall model performance (area under the curve, AUC) was (from best to worst): (1) BRT (boosted regression tree), (2) MAXENT (maximum entropy), (3) GAM (generalized additive model; regression), (4) MARS (multivariate adaptive regression splines), (5) GDMSS (generalized dissimilarity model, single species), (6) BRUTO (a fast GAM), (7) GLM (generalized linear model), (8) OM-GARP (new version of GARP), (9) DOMAIN (multivariate distance model), and (10) BIOCLIM (envelope model) (Figs. 3 and 4). The GLMM analyses (Table 4) consistently showed that BRT and MAXENT ranked first, all regression-based techniques (GAM, MARS, BRUTO, GLM; little evidence for true differences between them) ranked next, and the techniques that include profile-based information (i.e., relying only on presence data) ranked last (BIOCLIM, DOMAIN). The single-species dissimilarity approach (GDMSS) showed similar performance to the regression approaches. OM-GARP ranked just below these regression and dissimilarity techniques, and above the two envelope approaches (DOMAIN, BIOCLIM), but there was no detectable difference between these latter two.

Results from the experiments testing the response of model performance to experimental data manipulation showed significant and consistent effects, with all treatments causing a decline in the predictive performance (Fig. 5). The patterns in mean AUC per technique across the 10 species selected for these experiments were generally consistent with trends from the full suite of 30 species, though for these 10 species OM-GARP was closer to the regression methods than to the profile methods. Most modeling techniques responded significantly to error introduced in the geographic coordinates of the species data ($P = 0.0039$; Fig. 5), with profile-based algorithms (DOMAIN, BIOCLIM) being more affected. Coarsening the grain (resolution) had a consistent but almost negligible effect (Fig. 5).

Models performed more poorly with smaller sample size ($P = 0.0020$ for $n = 10$, 0.0039 for $n = 30$; Fig. 5c, d), but techniques varied in robustness. DOMAIN, OM-GARP, and MAXENT maintained average performance close to their original levels even at $n = 10$. At $n = 30$ all methods were recovering towards their original effectiveness (Fig. 5). Even though the regression methods were relatively more degraded at $N = 10$, they were still close to DOMAIN in terms of average performance.

The ability to model species' distributions can be successfully predicted from a series of additional species traits (58% variance explained). Only maximum growth rate (MGR), maximum elevation (ELVMAX), and elevation range (ELVRNG) showed significant effects on model performance ($P < 0.05$; Fig. 6). The multiple GLM of AUC evaluation as a function of quantitative species traits confirmed the strong influence of MGR and identified ELVRNG as another important variable (Table 5).

Discussion

Elith et al. (2006) compared the predictive power of 16 modeling techniques to predict species distribution using six data sets. In this study we look more extensively within one region, comparing the robustness of 10 techniques to various data issues—errors, sample size, and grain—and assessing species' predictability as a function of biological and ecological species' characteristics.

Species vary more than techniques

The variance in model effectiveness across species proved to be higher than the variance across techniques. This result confirms results from previous studies (e.g., Austin and Meyers 1995, Ferrier and Watson 1997, Elith 2000, Moisen and Frescino 2002, Thuiller 2003), but across a larger sample of techniques and with a focus on presence-only data. For the Swiss data set, species poorly modeled by one technique tend also to be poorly modeled by other techniques. Clearly, the variance results depend on the selection of species and techniques, but it is notable that the result is a repeatable one.

We found less evidence in the literature showing consistency in the ranking of techniques across species, as demonstrated by our study. In some data sets there are trends for more variation across techniques for species that are harder to model (Elith et al. 2006), so our result might be most relevant to data sets with functionally relevant predictors, strong environmental gradients, reasonable numbers of presence records for modeling, and enough evaluation data to distinguish between methods. In such cases, these results might be used to select a single best technique, for instance to test a hypothesis in biogeography or macro-ecology, or to develop a ready-to-use model-based management tool.

Ranking species

Model accuracy varied by species. The models for three species—*Castanea sativa*, *Ostrya carpinifolia*, and *Pinus cembra*—ranked as "good." These three species all have narrow ecological niches (specialized on specific substrates or for high elevations) and narrow geographical distribution in Switzerland. *Castanea sativa* and *Ostrya carpinifolia* are mainly found in the Southern Alps in the state of Ticino, and *C. sativa* in a small number of sites in the Jura mountains, whereas *Pinus*

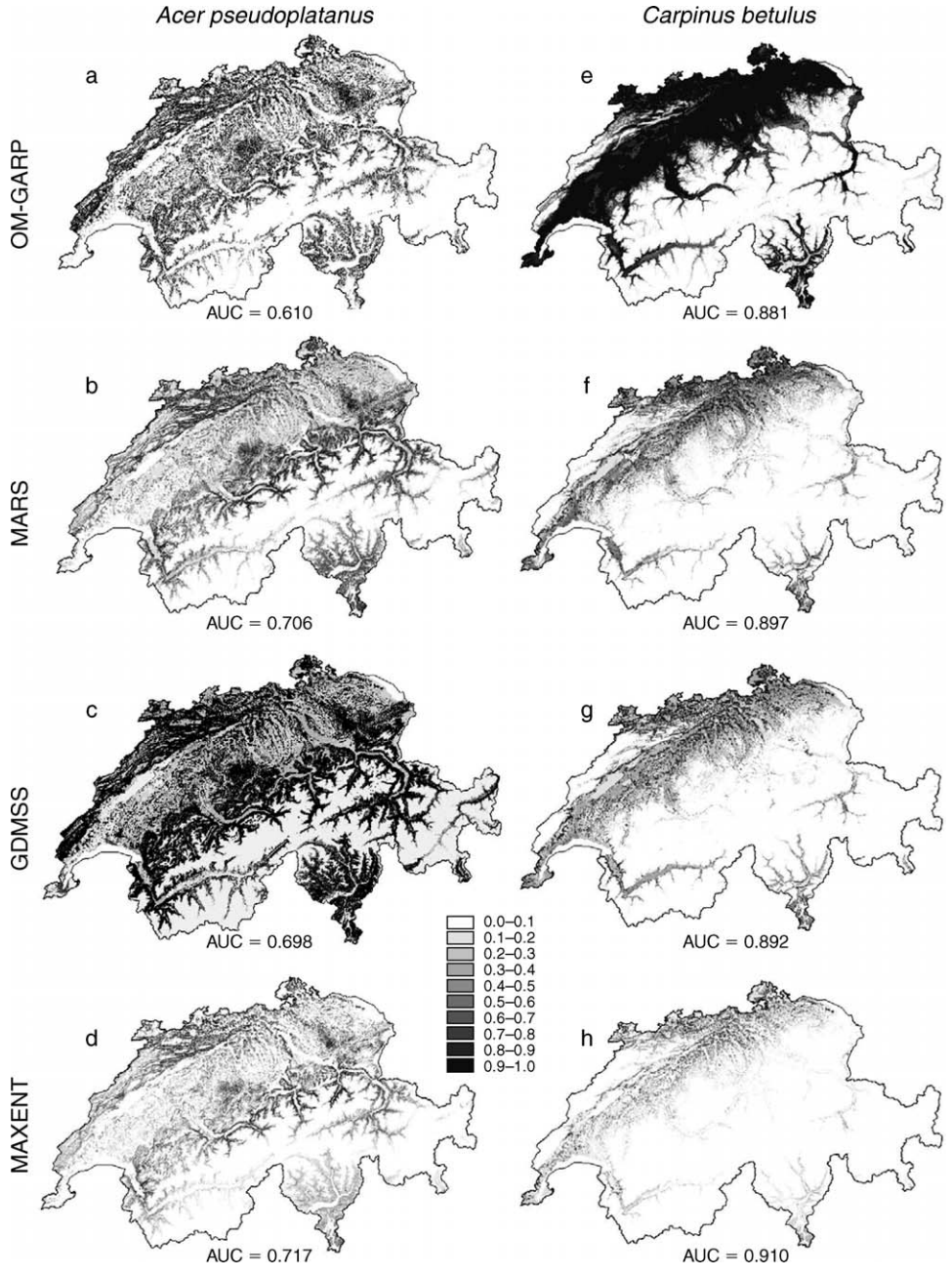

FIG. 2. Prediction maps for the four species in Fig. 1 and four modeling techniques. Species: (a–d) *Acer pseudoplatanus,* (e–h) *Carpinas betulus*, (i–l) *Castanea sativa*, (m–p) *Fagus sylvatica*. Techniques (see Appendix B): row 1 (a, e, i, m) is OM-GARP, row 2 (b, f, j, n) is MARS, row 3 (c, g, k, o) is GDMSS, and row 4 (d, h, l, p) is MAXENT. See *Methods: Species data* and *Modeling techniques* and Tables 1 and 4 for details on species and techniques. See Fig. 1 for actual species-distribution maps. The area-under-the-curve (AUC) values below each map were calculated on the independent test data set.

cembra is mainly found in the dry and continental internal valleys of the Alps.

Next, 21 species had models that were judged "useful." Several of these species—e.g., *Abies alba*, *Fagus sylvatica*, *Fraxinus excelsior*, *Larix decidua*, *Picea abies*, *Pinus sylvestris*, *Quercus petraea*—are dominant species in at least one type of forest in Switzerland, possibly reflecting some degree of environmental determinism. Their suboptimal ranking might partly result from the fact that some have been managed by humans in a number of places (mostly *F. sylvatica*, *L. decidua*, *P. abies*). *Fraxinus excelsior* can dominate forests in moist gullies and on sites close to springs (preferring subsurface water), but the extent of this type of forest was

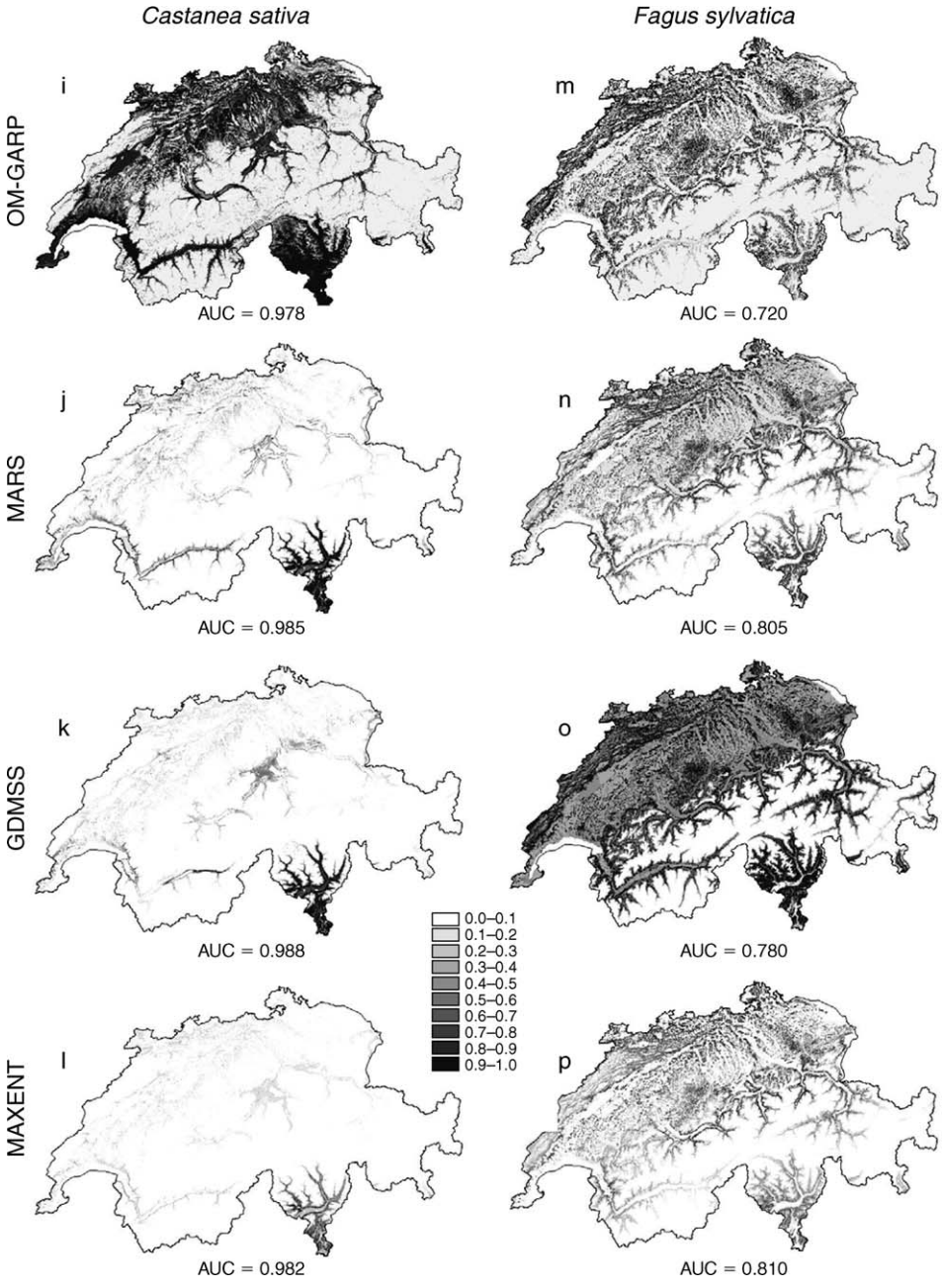

FIG. 2. Continued.

locally reduced by drainage practices over the last century. Yet this species still has a relatively narrow ecological niche, which might explain its position in this usefully modeled group. Some nondominant species are also partly influenced by human activities, including *Quercus robur*, *Populus nigra*, and *P. tremulus*. A bias in occurrence data may sometime result from human activities, but recorded absences should be more reliable since human management in Switzerland never attempted to eradicate a species from its natural range. Other species may have intermediate performance because their ecological requirements are difficult to map. *Quercus pubescens* needs highly thermal environments (e.g., shallow and stony soils), frequent in the southern slopes of the internal Alps but scattered elsewhere. *Pinus mugo* is only found in subalpine heaths of the eastern Swiss Alps. *Acer pseudoplatanus* can dominate on screes. These features are not frequent in the plateau and can be difficult to capture in these models. For instance, detailed soil characteristics are likely to be important additional predictors for many species, as recently illustrated for *Acer campestre* in France (Coudun et al.

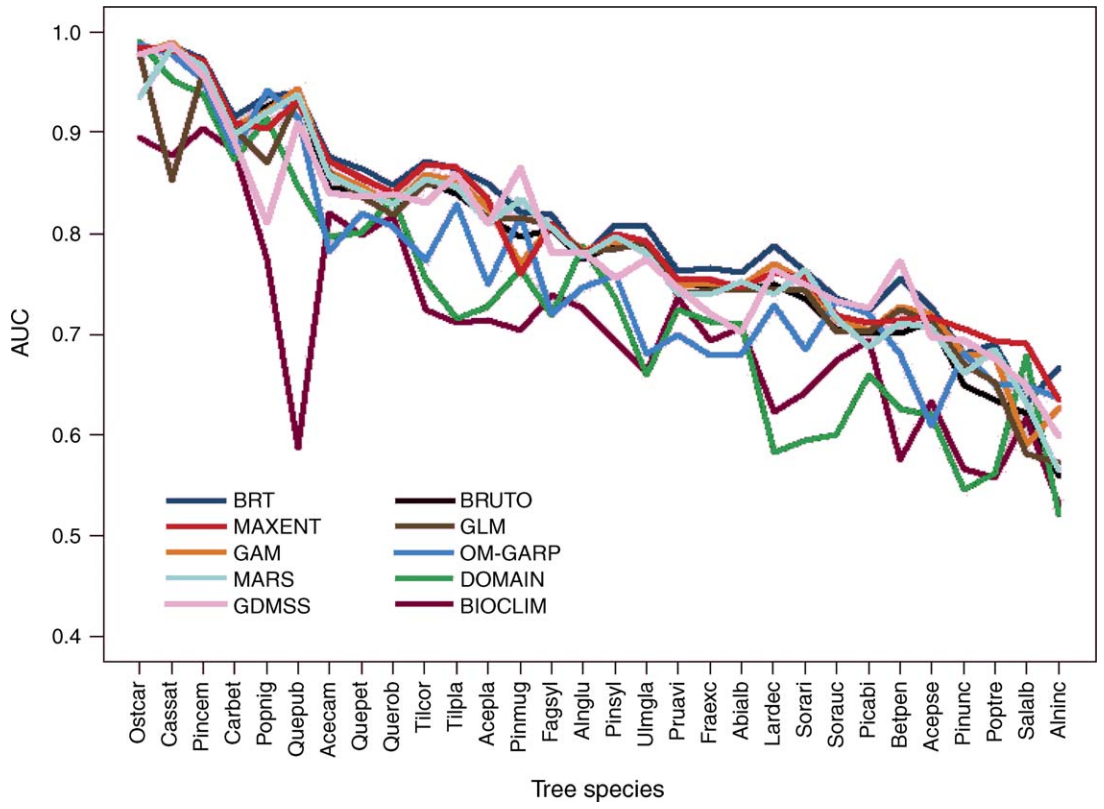

FIG. 3. Performance of species for each technique, in order of decreasing mean performance (area under the curve, AUC) of species across techniques. Techniques in the key are ordered by their mean performance across species (from best to worst); for species codes, see Table 1.

2006), but are not available as spatially explicit predictors. A few other species in this category were affected by historical events, as was the case for *Ulmus glabra* whose populations in western Europe were decimated by a fungal disease (*Ophiostoma ulmi*) during the 1970s.

Finally, six species had "poor" models. Three of them—*Acer pseudoplatanus*, *Betula pendula*, *Populus tremula*—rarely dominate in the landscape and tend to occur scattered throughout forests dominated by other species. For these species, the bulk of unexplained variance may result from the models' inability to either account for competitive performance (*Acer*) and/or the unstable occurrence of early successionals (*Betula*, *Populus*). Other species in this category have very specialized ecological requirements that could not have been distinguished using the available predictors. *Alnus incana* and *Salix alba* are only found in wet, potentially flooded environment (e.g., lake sides, remnant alluvial zones), but no predictor of aquatic habitats was included in the modeling. Finally, one species—*Pinus uncinata*—is mostly found on peat bogs and in mountains on subalpine rocky outcrops, and may have been affected by taxonomic confusion with its sister species *P. mugo* with which it is known to hybridize.

Based on expert knowledge, the following trend can be suggested: (1) best modeled species have distinct (narrow) distribution and are good competitors; (2) dominant and widespread species still support useful models; (3) species with low competitive strength, early successional status, or with narrowly specialized topographic requirements that are hard to model in a GIS are the most difficult to model. Although certainly oversimplified, these results may be helpful in predicting which species can be modeled well.

Ranking techniques

The fairly consistent ranking of techniques across species is consistent with results of Elith et al. (2006). The top two techniques were BRT and MAXENT, followed by GAM, MARS, GDMSS, BRUTO, and GLM (all approximately equal), all outperforming BIOCLIM, OM-GARP, and DOMAIN (see Appendix B for model information).

Not all available techniques were included in this study. For instance, ENFA (ecological niche factor analysis; Hirzel et al. 2002) is a profile technique that was not tested because it requires a dynamic link with a GIS and could not be run effectively for many species and many experiments. RF (random forest; Breiman 2001) is a machine-learning method (as neural networks)

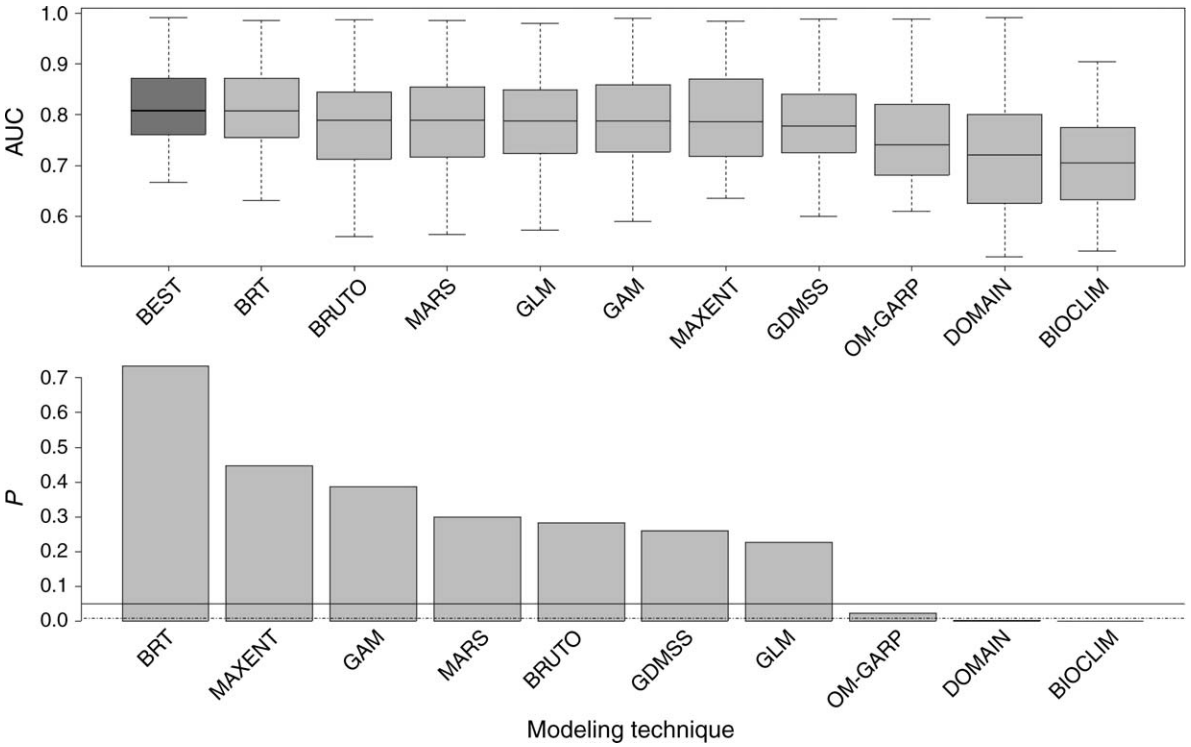

FIG. 4. Performance of modeling techniques compared to the best prediction (throughout all techniques) for each species. (Top panel) Boxplots give the distribution of performance measures (area under the curve, AUC) for each technique (second to last boxplots) compared to the distribution of the best performance for each species (first boxplot, dark gray). The line across the box indicates the median; the box boundaries show the interquartile range; and whiskers identify extreme data points that are no more than 1.5 times the interquartile range on both sides. (Bottom panel) For each technique, the histograms give the *P* value of differences from the best technique.

used occasionally in ecology with promising results (Lawler et al. 2006, Schwartz et al. 2006, Broennimann et al. 2007). It would also be interesting to test regression methods with other model-selection techniques such as the lasso (Reineking and Schröder 2006), because these are likely to be more robust to poor data than the stepwise methods used here.

BRT and MAXENT are new in ecological modeling, with few ecological applications available (e.g., Kawakita et al. 2005, Leathwick et al. 2006*a*, Moisen et al. 2006, Phillips et al. 2006). All regression techniques—GLM, GAM, BRUTO, and MARS—showed similar overall performance. The single-species dissimilarity approach (GDMSS) also showed comparable performance to the regression approaches. In the other regions tested, GDMSS performed relatively better (Elith et al. 2006). OM-GARP is based on a mix of logistic regression (GLM) and BIOCLIM-like approaches, so its intermediate performance (between profile techniques and regression approaches) is unsurprising. Yet, statistically, its

TABLE 4. Pairwise comparisons of predictive performance (area under the curve, AUC) of the 10 modeling techniques.

	BIOCLIM	BRT	BRUTO	DOMAIN	GAM	GDMSS	GLM	OMGARP	MARS
BRT	**0.000**								
BRUTO	**0.000**	**0.996**							
DOMAIN	**0.005**	**1.000**	**1.000**						
GAM	**0.000**	0.960	0.184	**0.000**					
GDMSS	**0.000**	**0.994**	0.440	**0.000**	0.773				
GLM	**0.000**	**1.000**	0.732	**0.000**	0.933	0.781			
OMGARP	**0.000**	**1.000**	**0.998**	**0.000**	**1.000**	**0.999**	**0.984**		
MARS	**0.000**	**0.991**	0.385	**0.000**	0.730	0.447	0.184	**0.001**	
MAXENT	**0.000**	0.853	0.052	**0.000**	0.237	0.071	**0.013**	**0.000**	0.093

Notes: For details of modeling techniques, see Appendix B; for details of generalized linear mixed model (GLMM) see Appendix C. Values were obtained from a GLMM of performance as a function of techniques and species (random effect) and were summarized over 50 000 Monte Carlo iterations. Values indicate the probability that the true difference between the techniques is greater than zero. Low values indicate that the technique in the row tends to give *higher* AUC than the technique in the column, and vice versa. Values outside the arbitrary limits $P = (0.025, 0.975)$ are highlighted in bold for this two-tailed test.

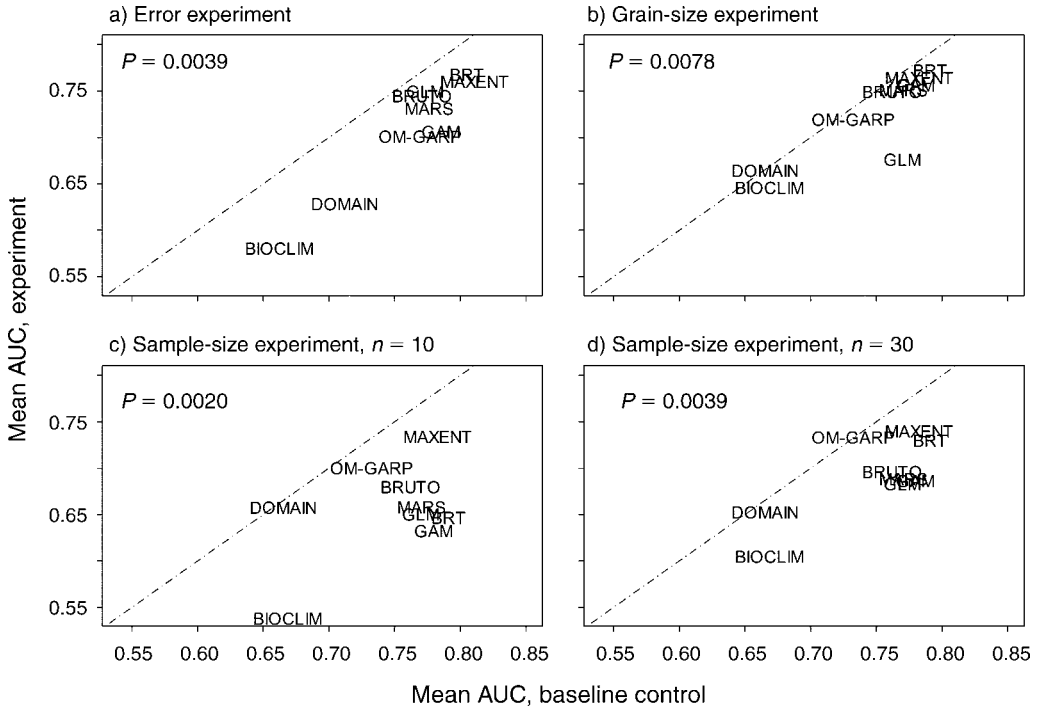

FIG. 5. Results of the modeling experiments. Response of the different techniques to various treatments of the data, with model performance measured by the area under the curve (AUC) (see *Methods: Model evaluation and comparison* for explanations). The *x*-axis shows baseline modeling (control); the *y*-axis is the experiment. The dashed lines represent the 1:1 relationship (perfect correspondence). For each method, the position of the label represents the mean value across 10 species. On average, all experiments either degrade the performance of techniques or provide similar results. The *P* value of a two-sided Wilcoxon signed-ranks test between the baseline modeling and experiment is provided in the upper-left corner of each panel (all significant with one-sided tests).

performance is not distinguishable from those of profile models. We note that the profile methods, particularly BIOCLIM, were originally introduced to enable broad-scale biogeographic analyses of species (e.g., Nix 1986). While not originally intended for the finer-grained scales here, we included them because their use has not been confined to broad scales. Methods that define unsuitable environments within broadly suitable envelopes are more successful at finer grains, as in this study.

Influence of extrinsic parameters of the data

Error.—Most modeling techniques were weakened by artificially introduced errors in the data set, with profile-based algorithms being more affected. Consistent degradation of model performance (area under the curve [AUC] always lower for the faulty data, except for some cases for DOMAIN) suggests that this is a significant issue. Overall, this result was expected. Introducing errors in the geographic coordinates should result in incorrect values for environmental predictors. The particularly consistent and significant effect on model performance here may be due to the high geographic accuracy of the Swiss data set. BIOCLIM may have been particularly affected because predictor variables are not weighted and erroneous extreme values (e.g., max/min or some lower percentiles) have a large potential effect on the fitting of the species' realized niche compared to regression-based techniques.

The same errors likely had less impact on techniques that fit more complex response curves (e.g., smoothing splines in GAM), where the influence of observations on the fitting process strongly depends on their position and their density along each environmental gradient.

Grain.—Coarsening the grain (resolution) had only a weak effect on the performance of techniques. GLM exhibited a somewhat greater sensitivity, but this likely resulted from three outlying species for which no satisfactory models could be obtained at the coarser scale. In general, coarsening the grain can both increase and decrease predictive capacity. Increased performance may originate from averaging or smoothing errors in both environmental and species data, for instance in the case of mobile species accessing resources available in a large neighborhood (Guisan and Thuiller 2005). Reduced performance may result from a lack in spatial matching between species observations—specifically of sessile organisms like trees—and their associated environments (Guisan and Thuiller 2005) and/or from limited/reduced precision of environmental predictors at the coarser grain (Guisan et al. 2007).

The effect of changing grain size might be confounded with a change in sample sizes of presence data between models built at coarse and fine grain. Coarsening the

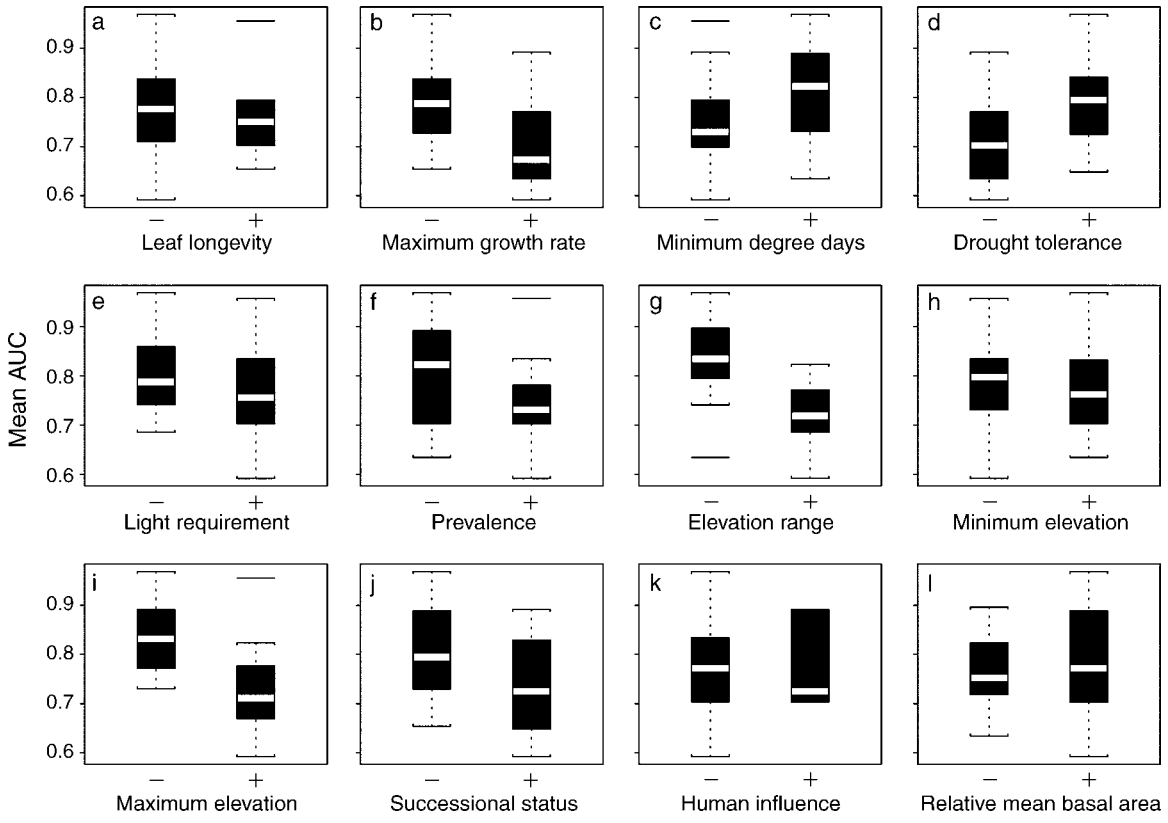

FIG. 6. Boxplots of mean AUC (area under the curve) values per species across techniques for the binary classes of species traits (30 species in total). The "–" on the *x*-axis represents the lower class whereas the "+" represents the upper class (e.g., low and high growth rate). For each continuous trait, separation between the two classes is obtained by taking the median value as threshold (see Table 1). The difference is only significant for (b), (g), and (i). The line across the box indicates the median; the box boundaries show the interquartile range. Whiskers identify extreme data points that are no more than 1.5 times the interquartile range on both sides. Horizontal lines outside the square brackets denote outliers.

grain requires either (a) aggregating the environmental and species data, or (b) allowing multiple species observations per coarser cell (to keep the sample size constant). In the second case—the one chosen in our study—the likelihood of finding both presence and (pseudo)absence data co-occurring in coarser cells increases. This should therefore result in a reduction of model fit, because such cells reduce the model's ability to distinguish between suitable and unsuitable locations. With only a few such cases, the model fit should not be seriously affected, whereas with many such cases, the fit could be decreased considerably, cancelling out positive effects of coarsening grain. The similar model performance observed across species at the two grain sizes may also indicate that the important environmental gradients in Switzerland can still be captured at the coarser grain, but this result is likely to be regionally variable, and will depend on the complexity of the environmental space.

Sample size.—Sample size had an important effect on model performance. At the smallest sample size ($n = 10$), the performance of all regression approaches, together with BRT and BIOCLIM, was strongly degraded compared to the full sample. By contrast, techniques that were able to maintain similar performance at very small sample size were, by decreasing order of performance: MAXENT, OM-GARP and DOMAIN. Maintenance of performance does not imply a better performance overall compared with other methods; regression methods still performed as well as DOMAIN at small sample sizes. Sample sizes of 30 produced the same trend but loss of performance was smaller. Among the two best techniques, BRT only remains superior to MAXENT until the sample size gets very small, showing the superiority of some techniques with small data sets.

Techniques fitting simpler response curves (like DOMAIN or GARP) tend to be less affected by decreased sample sizes than more complex (and data-

TABLE 5. Results of ANOVA of the GLM (generalized linear model) of mean AUC (area under the curve) performance as a function of species' traits.

Predictor†	Deviance	df	Residual deviance	*F*	*P*
MGR	0.025141	1, 27	0.254315	5.49336	0.027008
ELVRNG	0.135322	1, 26	0.118992	29.56820	0.000011

Note: Overall, the model explained 51.6% of the deviance.
† Species traits (predictors): MGR, maximum growth rate; ELVRNG, elevation range.

hungry) techniques (like GAM, MARS and BRT) that fit more complex response curves of species along environmental predictors. More observations are usually required to fit group-discrimination models (e.g., GAM) compared to profile techniques (e.g., DOMAIN); the techniques generally performing most poorly cannot make use of additional data in the same way as the best-performing techniques. MAXENT probably remains fairly robust at small sample sizes because the settings used vary with sample size, and simpler models are fitted with smaller samples (Elith et al. 2006, Phillips et al. 2006). Equivalent approaches (adjusting model complexity with sample size) could be applied to other methods (e.g., BRT and regression), but were not tested here. Results from all these experiments need to be considered with caution because, with only 10 species, one or two species can have substantial impacts on the results. It would be instructive to build more comprehensive evaluations with comparative analyses of fitted functions and mapped predictions (Hernandez et al. 2006).

Influence of species traits

Predicting how well species' distributions can be modeled as a function of a given set of environmental predictors and biological traits is important in optimizing investment in, for example, conservation of particular target species. Among species traits tested here, growth rate and elevation range had significant effects on model performance, and maximum elevation had a marginal effect.

Maximum growth rate is closely related to species' successional status, since fast growth rates are linked to low shade tolerance (Shugart 1984, Bugmann 2001). Fast-growing pioneer species are more difficult to model, since they require early successional stages that are neither easily mapped nor evaluated, and may not be frequent over the landscape. Hence, from a habitat perspective, absence of early successional species does not necessarily represent low habitat suitability (Guisan and Thuiller 2005), assuming that light is the primary factor for competitive exclusion. In contrast, species with slow growth rates are often late-successional, dominant, restricted-habitat specialists, which are more persistent in the landscape. Such persistence should allow better distinction between suitable and unsuitable habitats.

We found that species with narrower elevation ranges and with slow growth rate are also more likely to be modeled successfully. Stenoecious species with narrow elevation ranges have already been shown to be modeled more efficiently than ubiquitous species that occur widely along environmental gradients (e.g., Guisan and Hofer 2003). The importance of the maximum growth rate and the elevation range was confirmed in a multiple GLM model relating AUC performance to these traits. This result supports the idea that some species might be easier to model than others, and that this ability can be partly predicted from a limited number of species' traits. One way to overcome difficulties in modeling the actual distribution of pioneer species may be to expand models to include more dynamic components (Guisan and Thuiller 2005). Other species traits, such as dispersal or population-dynamic properties (e.g., dispersal-limited species; Pulliam 2000) or types of geographic distribution (core, satellite, urban, or rural species; see Collins et al. 1993), may affect model performance and should be explored.

Conclusion

We point out that the modeling reported in this study focuses on one particular problem: that of modeling species' distributions and maximizing discrimination between presences and absences. This challenge may not always coincide with the uses to which such models are to be put (Peterson 2006); as such, conclusions regarding which modeling approach is "best" may depend on the goals of the modeler. For our data and study area, our study suggests the following conclusions:

1) Presence-only data, here from the Swiss FP inventory, can be usefully modeled to predict species' distribution using random pseudo-absences and flexible modeling methods.

2) Variation in performance is greater among species than among modeling techniques, so a species modeled well (or poorly) by one technique is likely to be also well (or poorly) modeled by other techniques (but see point 4, below).

3) Ranking species by their average model performance, three categories can be identified, related to geographic distribution of the species (widespread vs. localized) and level of habitat specialization (narrow vs. generalist): species with both narrow distributions and specialized habitats are easier to model, but this effect may result in part from the fixed study extent and the smaller sample sizes for these species.

4) Different modeling techniques yield consistently distinctive model performance, so it is worth using the best techniques to improve overall performance; in the case of our Swiss forest trees data, the best two techniques are BRT and MAXENT. Profile-based techniques performed least well, but may still be useful in cases of small sample sizes, such as characterizing rare or endangered species. We have not tested comparative performance of techniques at the broader scales for which BIOCLIM was originally intended.

5) Among data-set properties experimentally manipulated, location error and sample size affected model performance. Changing grain had little impact on model performance of any techniques. In this study profile-type techniques were less sensitive to small sample size, but more sensitive to location errors, while the opposite was observed for regression-based approaches. This result depends on the relative perturbation of the data in each experiment.

6) Modeling success for particular species can be predicted, for the current data set, from a series of additional species traits (~52% variance explained, 30 species), based principally on growth rate, elevation

range and maximum elevation. No technique can "save" difficult-to-predict species.

Acknowledgments

We thank Pascal Vittoz for additional insights into the species' ecology; Robert Anderson, Louis Iverson, and an anonymous referee for useful comments on the manuscript; and other members of our NCEAS Species Distribution Modeling Group for useful discussions during the three workshops: Miro Dudik, Simon Ferrier, John Leathwick, Anthony Lehmann, Jin Li, Bette Loiselle, Glenn Manion, Craig Moritz, Miguel Nakamura, Yoshinori Nakazawa, Jake Overton, Karen Richardson, Ricardo Scachetti Pereira, Rob Schapire, Jorge Soberón, Stephen Williams, and Mary Wisz. A. Guisan benefited from support by the European Commission and by the Swiss National Science Foundation.

Literature Cited

Anonymous. 1980. Soil suitability map of Switzerland [in German]. In Eidgenossiche Drucksachen und Materialzentrale (EDMZ), Bern, Switzerland.

Austin, M. P., and J. A. Meyers. 1995. Modelling of landscape patterns and processes using biological data, subproject 4: real data case study. Consultant report to CSIRO Division of Wildlife and Ecology, Canberra, Austrailian Capital Territory, Australia.

BFS [Bimdesamt fuer Statistik]. 2001. Benutzerhandbuch GEOSTAT. Ausgabe 11/2001. Bundesamt fuer Statistik, Bern, Switzerland.

Boone, R. B., and W. B. Krohn. 2002. Modeling tools and accuracy assessment. Pages 265–270 *in* J. M. Scott, P. J. Heglund, F. Samson, J. Haufler, M. Morrison, M. Raphael, and B. Wall, editors. Predicting species occurrences: issues of accuracy and scale. Island Press, Covelo, California, USA.

Brassel, P., and U.-B. Braendli. 1999. Schweizerisches Landesforstinventar. Ergebnisse der Zweitaufnahme 1993–1995. Haupt Verlag, Bern, Switzerland.

Breiman, L. 2001. Random forests. Machine Learning 45:5–32.

Broennimann, O., U. Treier, H. Müller-Schärer, A. T. P. Peterson, W. Thuiller, and A. Guisan. 2007. Evidence of climatic niche shift during biological invasion. Ecology Letters 10:701–709.

Brotons, L., W. Thuiller, M. B. Araujo, and A. H. Hirzel. 2004. Presence–absence versus presence-only modelling methods for predicting bird habitat suitability. Ecography 27:437–448.

Bugmann, H. 1996. A simplified forest model to study species composition along climate gradients. Ecology 77:2055–2074.

Bugmann, H. 2001. A review of forest gap models. Climatic Change 51:259–305.

Burnham, K. P., and D. R. Anderson. 2002. Model selection and multi-model inference: a practical information-theoretic approach. Springer-Verlag, New York, New York, USA.

Busby, J. R. 1991. BIOCLIM—A bioclimate analysis and prediction system. Pages 64–68 *in* C. R. Margules and M. P. Austin, editors. Nature conservation: cost effective biological surveys and data analysis. CSIRO, Melbourne, Victoria, Australia.

Carpenter, G., A. N. Gillison, and J. Winter. 1993. DOMAIN: a flexible modelling procedure for mapping potential distributions of plants and animals. Biodiversity and Conservation 2:667–680.

Collins, S. L., S. M. Glenn, and D. W. Roberts. 1993. The hierarchical continuum concept. Journal of Vegetation Science 4:149–156.

Coudun, C., J.-C. Gégout, C. Piedallu, and J.-C. Rameau. 2006. Soil nutritional factors improve plant species distribution models: an illustration with *Acer campestre* (L.) in France. Journal of Biogeography 33:1750–1763.

Elith, J. 2000. Quantitative methods for modeling species habitat: comparative performance and an application to Australian plants. Pages 39–58 *in* S. Ferson and M. A. Burgman, editors. Quantitative methods in conservation biology. Springer-Verlag, New York, New York, USA.

Elith, J., et al. 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29:129–151.

Engler, R., A. Guisan, and L. Rechsteiner. 2004. An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. Journal of Applied Ecology 41:263–274.

ESRI [Environmental Science Research institute]. 2005. ArcGIS 9.2. Environmental Science Research Institute, Redlands, California, USA.

Ferrier, S., G. Manion, J. Elith, and K. Richardson. 2007. Using generalised dissimilarity modelling to analyse and predict patterns of beta-diversity in regional biodiversity assessment. Diversity and Distributions 13.

Ferrier, S., and G. Watson. 1997. An evaluation of the effectiveness of environmental surrogates and modelling techniques in predicting the distribution of biological diversity. Environment Australia, Canberra, Australian Capital Territory, Australia.

Fielding, A. H., and J. F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence-absence models. Environmental Conservation 24:38–49.

Graham, C. H., S. Ferrier, F. Huettman, C. Moritz, and A. T. Peterson. 2004. New developments in museum-based informatics and applications in biodiversity analysis. Trends in Ecology and Evolution 19:497–503.

Guisan, A., C. Graham, J. Elith, and F. Huettman, the NCEAS Species Distribution Group. 2007. Sensitivity of predictive species distribution models to change in grain size. Diversity and Distributions 13:332–340.

Guisan, A., and U. Hofer. 2003. Predicting reptile distributions at the mesoscale: relation to climate and topography. Journal of Biogeography 30:1233–1243.

Guisan, A., A. Lehmann, S. Ferrier, R. Aspinall, R. Overton, M. Austin, and T. Hastie. 2006. Making better biogeographic predictions of species distribution. Journal of Applied Ecology 43:386–392.

Guisan, A., and W. Thuiller. 2005. Predicting species distribution: offering more than simple habitat models. Ecology Letters 8:993–1009.

Hernandez, P. A., C. H. Graham, L. L. Master, and D. L. Albert. 2006. The effect of sample size and species characteristics on performance of different species distribution modeling methods. Ecography 29:773–785.

Hijmans, R. J., K. A. Garrett, Z. Huamán, D. P. Zhang, M. Schreuder, and M. Bonierbale. 2000. Assessing the geographic representativeness of genebank collections: the case of Bolivian wild potatoes. Conservation Biology 14:1755–1765.

Hijmans, R. J., L. Guarino, M. Cruz, and E. Rojas. 2001. Computer tools for spatial analysis of plant genetic resources data: 1. DIVA-GIS. Plant Genetic Resources Newsletter 127: 15–19.

Hirzel, A. H., J. Hausser, D. Chessel, and N. Perrin. 2002. Ecological-niche factor analysis: How to compute habitat-suitability maps without absence data? Ecology 83:2027–2036.

Kadmon, R., O. Farber, and A. Danin. 2003. A systematic analysis of factors affecting the performance of climatic envelope models. Ecological Applications 13:853–867.

Kadmon, R., O. Farber, and A. Danin. 2004. Effect of roadside bias on the accuracy of predictive maps produced by bioclimatic models. Ecological Applications 14:401–413.

Karl, J. W., L. K. Svancara, P. J. Heglund, N. M. Wright, and J. M. Scott. 2002. Species commonness and the accuracy of habitat-relationship models. Pages 573–580 *in* J. M. Scott, P. J. Heglund, M. L. Morrison, J. B. Haufler, M. G. Raphael, W. A. Wall, and F. B. Samson, editors. Predicting species occurrences: issues of accuracy and scale. Island Press, Covelo, California, USA.

Kawakita, M., M. Minamia, S. Eguchi, and C. E. Lennert-Cody. 2005. An introduction to the predictive technique AdaBoost with a comparison to generalized additive models. Fisheries Research 76:328–343.

Lawler, J. J., D. White, R. P. Neilson, and A. R. Blaustein. 2006. Predicting climate-induced range shifts: model differences and model reliability. Global Change Biology 12:1568–1584.

Leathwick, J. R., J. Elith, M. P. Francis, T. Hastie, and P. Taylor. 2006*a*. Variation in demersal fish species richness in the oceans surrounding New Zealand: an analysis using boosted regression trees. Marine Ecology Progress Series 321: 267–281.

Leathwick, J. R., J. Elith, and T. Hastie. 2006*b*. Comparative performance of generalized additive models and multivariate adaptive regression splines for statistical modelling of species distribution. Ecological Modelling 199:188–196.

Leathwick, J. R., D. Rowe, J. Richardson, J. Elith, and T. Hastie. 2005. Using multivariate adaptive regression splines to predict the distribution of New Zealand's freshwater diadromous fish. Freshwater Biology 50:2034–2052.

Lischke, H., N. E. Zimmermann, J. Bolliger, S. Rickebusch, and T. J. Löffler. 2006. TreeMig: a forest-landscape model for simulating spatio-temporal patterns from stand to landscape scales. Ecological Modelling 199:409–420.

Loiselle, B. A., C. A. Howell, C. H. Graham, J. M. Goerck, T. Brooks, K. G. Smith, and P. H. Williams. 2003. Avoiding pitfalls of using species distribution models in conservation planning. Conservation Biology 17:1591–1600.

Manel, S., H. C. Williams, and S. J. Ormerod. 2001. Evaluating presence–absence models in ecology: the need to account for prevalence. Journal of Applied Ecology 38:921–931.

Manly, B. F., L. L. McDonald, and D. L. Thomas. 1993. Resource selection by animals: statistical design and analysis for field studies. First edition. Chapman and Hall, New York, New York, USA.

Moisen, G. G., E. A. Freeman, J. A. Blackard, T. S. Frescino, N. E. Zimmermann, and T. C. J. Edwards. 2006. Predicting tree species presence and basal area in Utah: a comparison of stochastic gradient boosting, generalized additive models and tree-based methods. Ecological Modelling 199:176–187.

Moisen, G. G., and T. S. Frescino. 2002. Comparing five modelling techniques for predicting forest characteristics. Ecological Modelling 157:209–225.

Nix, H. 1986. A biogeographic analysis of Australian elapid snakes. Pages 4–5 *in* R. Longmore, editor. Atlas of elapid snakes of Australia. Australian Government Publishing Service, Canberra, Australian Capital Territory, Australia.

Pearce, J., and M. Boyce. 2006. Modelling distribution and abundance with presence-only data. Journal of Applied Ecology 43:405–412.

Peterson, A. T. 2006. Uses and requirements of ecological niche models and related distributional models. Biodiversity Informatics 3:59–72.

Phillips, S. J., R. P. Anderson, and R. E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. Ecological Modeling 190:231–259.

Pulliam, H. R. 2000. On the relationship between niche and distribution. Ecology Letters 3:349–361.

Reineking, B., and B. Schröder. 2006. Constrain to perform: regularization of habitat models. Ecological Modelling 193: 675–690.

Robertson, M. P., N. Caithness, and M. H. Villet. 2001. A PCA-based modelling technique for predicting environmental suitability for organisms from presence records. Diversity and Distributions 7:15–27.

Schwartz, M. W., L. R. Iverson, A. Prasad, A. Matthews, and R. J. O'Connor. 2006. Predicting extinctions as a result of climate change. Ecology 87:1611–1615.

Shugart, H. H. 1984. A theory of forest dynamics: the ecological implications of forest succession models. Springer-Verlag, New York, New York, USA.

Spiegelhalter, D., A. Thomas, N. Best, and D. Lunn. 2003. WinBUGS user manual. Version 1·4. MRC Biostatistics Unit, Cambridge, UK.

Stockwell, D., and D. Peters. 1999. The GARP modelling system: problems and solutions to automated spatial prediction. International Journal of Geographical Information Science 13:143–158.

Stockwell, D. R. B., and A. T. Peterson. 2002*a*. Controlling bias in biodiversity data. Pages 537–546 *in* J. M. Scott, P. J. Heglund, M. L. Morrison, J. B. Haufler, M. G. Raphael, W. A. Wall, and F. B. Samson, editors. Predicting species occurrences: issues of accuracy and scale. Island Press, Covelo, California, USA.

Stockwell, D. R. B., and A. T. Peterson. 2002*b*. Effects of sample size on accuracy of species distribution models. Ecological Modelling 148:1–13.

Swets, J. A. 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 Biology 9:1353–1362.

Thuiller, W., S. Lavorel, G. F. Midgley, S. Lavergne, and A. G. Rebelo. 2004. Relating plant traits and species distributions along bioclimatic gradients for 88 *Leucadendron* taxa. Ecology 85:1688–1699.

Walker, P. A., and D. Cocks. 1991. HABITAT: a procedure for modelling a disjoint environmental envelope for a plant or animal species. Global Ecology and Biogeography Letters 1: 108–118.

Wohlgemuth, T. 1992. Die vegetationskundliche Datenbank. Schweizerische Zeitschrift für das Forstwesen 143:22–36.

Zaniewski, A. E., A. Lehmann, and J. M. C. Overton. 2002. Predicting species spatial distributions using presence-only data: a case study of native New Zealand ferns. Ecological Modelling 157:261–280.

Zimmermann, N. E., and F. Kienast. 1999. Predictive mapping of Alpine grasslands in Switzerland: species versus community approach. Journal of Vegetation Science 10:469–482.

APPENDIX A

Supplementary details on the derivation of environmental predictors used to predict Swiss tree distribution (*Ecological Archives* M077-018-A1).

APPENDIX B

Summary table on the ten modeling techniques used to predict Swiss tree distribution (*Ecological Archives* M077-018-A2).

APPENDIX C

Supplementary details on the generalized linear mixed models (*Ecological Archives* M077-018-A3).