Journal of Biogeography (J. Biogeogr.) (2006) 33, 1677–1688

SPECIAL ISSUE

Five (or so) challenges for species distribution modelling

Miguel B. Araújo[1,2] and Antoine Guisan[3]

[1]Department of Biodiversity and Evolutionary Biology, National Museum of Natural Sciences, CSIC, 28006, Madrid, Spain, [2]Centre for Macroecology, Institute of Biology, Universitetsparken 15, DK-2100 Copenhagen, Denmark, [3]Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland

*Correspondence: Miguel Araújo, Department of Biodiversity and Evolutionary Biology, National Museum of Natural Sciences, CSIC, C/Gutiérrez Abascal, 2, 28006, Madrid, Spain. E-mail: maraujo@mncn.csic.es

ABSTRACT

Species distribution modelling is central to both fundamental and applied research in biogeography. Despite widespread use of models, there are still important conceptual ambiguities as well as biotic and algorithmic uncertainties that need to be investigated in order to increase confidence in model results. We identify and discuss five areas of enquiry that are of high importance for species distribution modelling: (1) clarification of the niche concept; (2) improved designs for sampling data for building models; (3) improved parameterization; (4) improved model selection and predictor contribution; and (5) improved model evaluation. The challenges discussed in this essay do not preclude the need for developments of other areas of research in this field. However, they are critical for allowing the science of species distribution modelling to move forward.

Keywords

Fundamental niche, model evaluation, niche models, parameterization, realized niche, sampling, variable selection.

INTRODUCTION

Species distribution modelling is central to both fundamental and applied research in biogeography. Over the past 10 years species distribution models have become commonplace in studies of biogeography, conservation biology, ecology, palaeoecology and wildlife management. Model fitting is usually based on pattern-recognition approaches, whereby associations between geographic occurrence of a species and a set of predictor variables are explored to allow or support statements of the mechanisms governing species' distributions. These models allow estimation of species' ecological requirements, although the degree to which causal relationships between species distributions and the predictor variables are unveiled depends on the adequacy of the predictors used for model building. In spite of the widespread use of niche-based species distribution models (e.g. Guisan & Thuiller, 2005), important conceptual, biotic and algorithmic uncertainties still need to be investigated if models are to make important contributions for conservation and biogeographical research. As with climate change research (e.g. Smith, 2002), the species' distribution modelling community needs to deepen the ongoing debate, where the strengths and limitations of available approaches are investigated, fuelled by more rigorous assessments of the sensitivity of model outcomes to initial assumptions and parameters (Whittaker *et al.*, 2005). In this essay, we identify five high-priority areas of inquiry for niche-based species distribution modelling: (1) clarification of the niche concept; (2) improved designs for sampling data for building models; (3) improved parameterization; (4) improved model selection and predictor contribution; and (5) improved model evaluation. The challenges discussed in this essay do not preclude the need for development of other areas of species distribution modelling research, but are critical for allowing this science to move forward.

CHALLENGE 1: CLARIFICATION OF THE NICHE CONCEPT

The foundations of niche modelling are deeply rooted in Hutchinson's (1957) fundamental and realized niche concepts. Even though a majority of modellers would subscribe to Hutchinson's framework, there are conflicting views about what the models truly represent. For example, recently Soberón & Peterson (2005) concluded that niche models provide an approximation to the species' fundamental niche. Others have regarded models as providing a spatial representation of the realized niche (e.g. Austin *et al.*, 1990; Guisan & Zimmermann, 2000; Pearson & Dawson, 2003) on the grounds that the observed species' spatial distributions that are utilized

doi:10.1111/j.1365-2699.2006.01584.x

to estimate species–climate relationships are constrained by non-climatic factors (e.g. Araújo & Pearson, 2005). Conflicting interpretations of the niche component being represented by models arise from ambiguities in the original formulation of the fundamental and realized niche concepts, and from difficulties in translating Hutchinson's conceptual framework into niche modelling, particularly at large biogeographical scales.

Hutchinson defined the fundamental niche (N) as the 'n-dimensional hypervolume' where a species S_1, in the absence of competition with species S_2, is able to persist indefinitely. The realized niche of species S_1 (N_1') is the part of the fundamental niche (N_1) where the species is not absent due to competition with species S_2. Following Hutchinson's notation, N_1' would be the subset of N_1 that does not contain N_2:

$$N_1' = N_1 - N_2 \cap N_1 \times N_2, \quad \text{where } S_1 \text{ survives.} \qquad (1)$$

And, reciprocally,

$$N_2' = N_2 - N_1 \cap N_1 \times N_2, \quad \text{where } S_2 \text{ survives.} \qquad (2)$$

There are a number of difficulties in adopting such a strict Hutchinsonian framework in the context of niche modelling. First, the word niche became firmly entangled with the notion of interspecific competition (Chase & Leibold, 2003). However, there is increasing evidence that positive biotic interactions (e.g. mutualism and facilitation) may be as important as negative interactions for species survival (e.g. Callaway *et al.*, 2002; Yamamura *et al.*, 2004; Travis *et al.*, 2005). The entanglement of the niche concept with competition partly reflects the disproportionate weight given to competition in Hutchinson's original writings, but also to ambiguity as to how other types of interactions would fit in the fundamental and realized niche framework. Two statements from his 'concluding remarks' (1957) may, however, shed some light on Hutchinson's thinking regarding how to integrate biotic interactions into the niche framework:

(1) it will be apparent that if this procedure [defining the hypervolume] could be carried out, all…variables, both physical and biological, being considered, the fundamental niche of any species will completely define its ecological properties. (p. 416)

(2) Interaction of any of the considered species [defining the realized niche] is regarded as competitive…All species other than those under consideration are regarded as part of the coordinate system. (p. 417)

It is reasonable to interpret these statements as indicating that Hutchinson saw biotic interactions other than competition as comprising the n-dimensional hypervolume defining the fundamental niche.

It follows from our reading of Hutchinson that limiting factors (e.g. temperature and presence of mutualist species) and resource factors (e.g. energy and presence of prey) should be part of the coordinate system characterizing the fundamental niche. Some authors have already implemented this idea by incorporating distributions of interacting species as predictor variables within niche models (e.g. Leathwick & Austin, 2001; Anderson *et al.*, 2002; Gutiérrez *et al.*, 2005). One consequence of including both positive and negative interactions within the niche framework is that the clear-cut dichotomy between fundamental and realized niches becomes artificial and its usefulness debatable. If positive interactions were considered part of the fundamental niche (as implied by Hutchinson's statements), then why should negative interactions alone be associated to the realized niche? If the rationale – as we interpret it – is that the fundamental niche is defined by the resources and limiting factors required for species' persistence, and that the realized niche is defined by the constraints preventing the exploitation of resources, should the absence of mutualists or facilitators (thus preventing the use of resources) be included as part of the factors defining the realized niche? Ambiguities concerning the role of biotic interactions within the niche framework need to be resolved in order to allow appropriate integration of these neglected issues into niche models.

Additional difficulties arise from the utilization of the Volterra–Gause principle (Hutchinson, 1957) to justify distinction between fundamental and realized niches. The principle postulates that two species utilizing, and limited by, a common resource cannot coexist in the same physical space. This observation is equivalent to stating that the realized niches of two co-occurring species do not intersect and that the species distribution can be restricted by a superior competitor. However, the relevance of this principle is likely to be contingent on the spatial grain of the analysis and the type of organism being considered. Hutchinson was mainly concerned with small species' ranges and niches at the scale of the community, but species distribution models are often fitted at regional to continental scales. The characterization of niches is made from gridded species' occurrence records on maps; the size of the grids can be large (1–50 km) and species competing for the same resource may well co-occur within the same grid since they can shift positions in order to avoid competition. In practice this means that species with intersecting modelled realized niches can co-occur in geographical space. Even if geographical units were small enough to detect patterns of competitive exclusion, species competing for the same resources could still reach local equilibrium. Examples include the case of species with lower degrees of competitiveness that establish in randomly vacant portions of geographical space (Hutchinson's 'fugitive' species), or species that use the same resources, at the same place, but in different periods of time (e.g. diurnal vs. nocturnal species). Therefore, co-occurring species *can* have intersecting realized niches and coexist in space and time (see also Amarasekare, 2003).

Given the discussion above, it is worth asking whether the distinction between fundamental and realized niches is useful for niche modelling. A possibility is to discard the fundamental and realized niche concepts altogether, accepting that any characterization of the niche is an incomplete description of

the abiotic and biotic factors allowing species to satisfy their minimum ecological requirements. In a major revision of the niche theory, Chase & Leibold (2003) dropped the fundamental and realized niche concepts and provided an updated definition of niche: the environmental conditions that allow a species to satisfy its minimum requirements so that birth rate of a local population is equal to or greater than its death rate. (p. 19)

This definition is essentially a modern translation of the original formulation of the niche concept proposed by Grinnell (1917, 1924). However, Chase and Leibold extended this concept to include 'the per capita impacts of that species on these environmental conditions'. Even if conceptually useful, this extension is of limited utility for niche-based modelling because niche models explore snapshot correlations and it is difficult to assess feedback mechanisms of species with their environment. Furthermore, models are usually implemented for areas with broad extents and/or using coarse resolutions where feed-back mechanisms are difficult to detect.

Niche theory is also hampered by inadequate consideration of species dispersal. Some authors have noted that dispersal limitation, not competition, causes species to be absent from significant portions of the fundamental niche at large ecological scales (e.g. Pulliam, 2000; Svenning & Skov, 2004; Araújo & Pearson, 2005; Guisan & Thuiller, 2005; Soberón & Peterson, 2005; Peterson, 2006). However dispersal is a spatial (and temporal) process. Hence, dispersal and dispersal limitation are more adequately treated as extrinsic dimensions of the niche concept (e.g. Araújo & Williams, 2000). We suggest that modellers should make a clearer distinction between niche models and the modelling of spatially explicit features. Niche models based on environmental predictors only yield projections of *potential habitats* for species (Guisan & Zimmermann, 2000). When niche and spatially explicit factors are combined, models can be said to yield projections of *potential geographical distributions* of species. This distinction is important, although researchers have often treated the two concepts interchangeably. For example, assessments of species extinction-risk that make assumptions on the relationship between shifts in niche space for species and species range losses (Thomas *et al.*, 2004) assume a linear relationship between species range size and availability of suitable habitats for species, which may not always apply (Araújo *et al.*, 2005c).

We have highlighted some difficulties in applying Hutchinson's (1957) niche concept to niche modelling. The definition of niche recently proposed by Chase & Leibold (2003) offers much promise, but it needs to be considered in the context of larger scale species' distribution modelling, where detailed population parameters are not measurable.

CHALLENGE 2: IMPROVED DESIGNS FOR SAMPLING DATA FOR BUILDING MODELS

A major assumption of traditional parametric procedures is that data used for calibration represent a random sample of the population being studied. Of course, it is rarely the case that such unbiased data exist for large numbers of species across wide geographical areas. Even if random samples of species locational records were obtainable, they would unlikely be the best for niche modelling (Hirzel & Guisan, 2002), particularly for species with a high degree of niche specialization and/or restricted range (e.g. Edwards *et al.*, 2005; Guisan *et al.*, 2006a). In practice, modellers would be reasonably happy if they could stratify sampling in order to obtain data that were representative of the species' frequencies of occurrence in a given region and the kinds of environments in which species can live. These attributes are important because niche models are sensitive to both sample size and biases in the distribution of data (e.g. Peterson & Cohoon, 1999; Stockwell & Peterson, 2002; Kadmon *et al.*, 2003; Araújo *et al.*, 2005c; Visscher, 2006). The purpose of improved statistical sampling design is to limit these biases while increasing model performance.

Currently, the vast majority of data available for species distribution modelling come from museum and other natural history collections (e.g. Graham, 2001; Ponder *et al.*, 2001; Reutter *et al.*, 2003; Stockman *et al.*, 2006). These data are often incomplete and biased in relation to the true spatial or environmental distributions of species. The option of sampling entirely new data by means of a statistically-sound recording scheme (e.g. random-stratified or other random-based procedure) is appealing but unfeasible in many circumstances (Balmford & Gaston, 1999). Here, we argue that intermediate solutions for handling and improving poor quality data may exist, but they need to be thoroughly tested in the context of species distribution modelling. In particular, we suggest the sub-sampling of existing data or, as a better alternative, the use of design-based or model-based environmental stratifications to help targeting additional field sampling.

Sub-sampling consists of selecting observations from a larger set with the aim to remove or reduce an identified bias. For instance, a non-random sample might bias observations to the neighbourhood of urban centres or roads because these areas are more easily accessed by recorders. Such biases in geographical space might translate into biases in environmental space, which can produce spurious assessments of the relationship between the response and predictor variables. However, when sampling occurs across broad geographical areas or sharp environmental gradients, there is a possibility that the environmental conditions limiting species distributions are well sampled, thus enabling useful models to be fitted (Peterson, 2006). If these conditions are not met, sub-sampling may help reduce existing biases and produce more balanced samples for model calibration, but with the corollary that the model can then only be applied within the range of resampled environmental conditions. Furthermore, this approach will not remove any undetected bias. The sub-sampling of existing data should be performed using a robust design (e.g. providing the best performing model for a target species; Hirzel & Guisan, 2002) and this operation repeated several times to assess the sensitivity of model outcomes to the sample data. Alternatively, sub-sampling can be replaced by a process of down-weighting observations that are over-represented or linked

with an identified sampling bias. However, these two *post hoc* strategies are not optimal and are likely to be useful only if: (1) the initial data set is large-enough to allow sub-sampling or weightings, leaving sufficient data for calibrating models; and (2) environmental space has been comprehensively sampled but in a very unbalanced way, such that some environmental combinations are much more represented than others.

A more appropriate choice is to use design- or model-based environmental stratifications (see Buckland *et al.*, 2000) to target additional sampling-effort needed to complement existing species' occurrence records. Use of this approach assumes that if an environmental pattern is sampled across the range of abiotic and biotic factors governing species distributions, an unbiased sample is gained of the species' occurrence along these gradients. This approach provides a framework for sampling species when only incomplete distributional data are available. The approach also allows the assessment of the expected degree of environmental representativeness of existing records for modelling. The idea of using pattern variation to sample biodiversity was first proposed for reserve selection (Faith & Walker, 1996), but it has also been discussed in the context of sampling biodiversity or single species for modelling applications (e.g. Ferrier *et al.*, 2002; Araújo *et al.*, 2003; Ferrier *et al.*, 2004; Edwards *et al.*, 2005; Hortal & Lobo, 2005; Guisan *et al.*, 2006a). Model-based stratifications are not very different from previous design-based environmental-stratification schemes used, for example, in vegetation science. The difference is that stratifications using models can produce species-specific quantitative assessments, thus increasing the likelihood that outputs are realistic for large numbers of species. An early example of model-based sampling (Mohler, 1983) showed that sampling more intensively the tails of predefined unimodal species' response curves along environmental gradients reduced the standard error and improved the adjustment of the curves to the data, and thus improved the fit of the model.

Even though model-based stratifications offer promise for prioritizing the distribution of sampling effort for niche modelling, further testing of the assumptions underlying these models is needed. Indeed, tests developed for reserve selection have provided only limited evidence for the value of particular implementations of stratification-based models (Araújo *et al.*, 2001, 2004; Bonn & Gaston, 2005), particularly for restricted range species (but see Sarkar *et al.*, 2005; Trakhtenbrot & Kadmon, 2005). These results contrast with those of recent studies investigating the usefulness of stratification-based models in the context sampling recording effort for niche modelling (Edwards *et al.*, 2005; Guisan *et al.*, 2006a).

CHALLENGE 3: IMPROVED PARAMETERIZATION STRATEGIES

Many modelling techniques are now available for species distribution modelling and there is increased recognition that different techniques yield different results, even when models are calibrated with the same response and predictor variables (e.g. Elith, 2000; Olden & Jackson, 2002; Thuiller *et al.*, 2003; Brotons *et al.*, 2004; Segurado & Araújo, 2004; Araújo *et al.*, 2006; Elith *et al.*, 2006; Pearson *et al.*, 2006). In addition to variability of model outputs due to using different modelling techniques, variability arises from using different implementations of the same technique (e.g. Elith *et al.*, 2006; Pearson *et al.*, 2006). Factors that affect parameterizations include the type of variable selection strategy used, the way absences are estimated, and the way spatial structures are considered (e.g. Maggini *et al.*, 2006; Segurado *et al.*, 2006).

It is important to consider several implementations of the same technique in order to improve our understanding of the sensitivity of models to initial assumptions and parameters, and to allow more robust model comparisons (e.g. Elith *et al.*, 2006; Pearson *et al.*, 2006). In a modern regression framework, tests and diagnostic tools are available and they should be used more systematically to check for: (1) the appropriateness of the distribution of residuals, and particularly of their variance and dispersion parameters; (2) the appropriateness of the link function; (3) the influence of outlying individual observations on model fit; (4) the presence of autocorrelation in the response and predictor variables; and (5) the appropriate level of complexity allowed in response curves. Flexible solutions to some of these problems are currently available in most GLM and GAM utilities, but they are also available and still being developed for other predictive modelling frameworks (for review see Hastie *et al.*, 2001).

When comparing different techniques using a single implementation of each technique, results may lead to concluding that a technique is inferior to another simply because the latter is more powerfully implemented (for discussion see Thuiller *et al.*, 2003; Segurado & Araújo, 2004). As a result, when making comparisons of different techniques, the variation in parameterization within each technique should be also assessed. A useful question is whether variation arising from alternative parameterizations within a single technique (i.e. within-model comparison) differs significantly from variation between modelling techniques (i.e. between-model comparisons).

CHALLENGE 4: IMPROVED MODEL SELECTION AND PREDICTOR CONTRIBUTION

Model outputs are primarily driven by the choice of predictor variables entering the models and by the type and level of adjustment between the response and predictor variables. The selection of predictor variables (model selection) is, therefore, a central step in most modelling efforts (e.g. Guisan & Zimmermann, 2000; Heikkinen *et al.*, 2006). We suggest that greater focus be given to the relative weight (explanatory power) and causality (ecological basis for choosing variables) of each predictor entering species distribution models.

It is unsurprizing that the choice of predictors affects the modelled spatial distribution of species. More interestingly, it was shown recently that differences among predictions can be very great when models are used to project distributions of

species into independent situations (e.g. Araújo *et al.*, 2005c; Pearson *et al.*, 2006). There is no easy fix for this problem, but increasing the stability of model selection is worth exploring. For example, regression techniques use different stepwise variable selection algorithms, but these are high-variance operations with small perturbations in the response variable leading to potentially very different subsets of the predictor variables being selected (e.g. Guisan *et al.*, 2002; Visscher, 2006). These problems have led statisticians to develop alternative approaches for the selection of predictors in regression-based frameworks, such as coupling a stepwise approach with cross-validation (Maggini *et al.*, 2006), using shrinkage rules, such as ridge regression or lasso (Harrell, 2001; Hastie *et al.*, 2001), or averaging competing models (Johnson & Omland, 2004). Similarly, traditional model selection procedures for classification and regression tree (CART) algorithms have now been out-competed by boosted versions of these techniques that make use of a multitude of small sub-trees branched to each other to form a final tree (Hastie *et al.*, 2001). These alternatives still need more thorough testing in the context of species distribution modelling. Nonetheless, recent analyses indicate that they may hold great promise (Elith *et al.*, 2006).

Another important issue is the calculation of the contribution of each predictor in the model. In regressions, ANOVA tables summarize the amount of deviance explained by each predictor in a sequential way and this is a function of the order in which each predictor enters the model. Hence, each predictor 'only explains what is left to explain' once the previous predictors have accounted for a part of the initial deviance. Therefore, except the cases where predictors are orthogonal, the deviance is not and cannot be interpreted as an absolute measure of the contribution of predictors in the model. This sensitivity weakens the utility of the model as it confounds two key factors: which are the most important predictors, and how much variance each predictor can potentially explain. Solutions to the problem have been proposed, but again they have not been extensively used in species distribution modelling and sufficient statistical validation is lacking.

One approach for generalized regressions is hierarchical partitioning (MacNally & Walsh, 2004; Heikkinen *et al.*, 2005). This procedure tests, for a single model, all entry order of predictors (i.e. all permutations of the model formula), and provides an assessment of the mean contribution of each predictor. As these mean contributions are not additive, hierarchical partitioning does not produce adequate models for prediction – indeed that is not its purpose (MacNally, 2002). Nonetheless, hierarchical partitioning might provide a useful framework for generating a detailed basis for inferring causality in multivariate regression settings (Heikkinen *et al.*, 2005). Still, the rational for calculating these average contributions needs further statistical testing. When the contribution of predictors needs to be calculated from many competing models, for the same species, model averaging (Burnham & Anderson, 2002) can be used instead by summing the weights of the models in which each predictor is present. This procedure allows the calculation of the contribution of each predictor as a weighted sum of all AIC-based scores attributed to the models in which the predictor is present.

Variance partitioning is another approach, embedded within a regression framework, which allows comparison of groups of predictors (rather than individual predictors). The idea was developed in the 1950s (Yoccoz & Chessel, 1988) to compare two groups of predictors in ordinary least squares regressions (OLS) and associated ordination techniques (Borcard *et al.*, 1992) and it is based on partial analyses of regression residuals. First a model is fitted with the first group of predictors, then a second model is fitted by regressing the residuals of the first model with the second group of predictors. The same procedure is then reversed, by fitting the first model with the second group and the second model with the first group. A full model is also fitted with both groups of predictors merged to calculate the total amount of unexplained deviance. By comparing the different models, one can derive the proportion of pure deviance explained by group 1, of deviance explained by group 2, of shared deviance between the two groups, and of total unexplained deviance. An extension of the approach was recently proposed to take more than two groups into account (Heikkinen *et al.*, 2005). While OLS is a powerful approach, complications arise when trying to implement it into other modelling frameworks. For instance, when a binary response is used, such as the presence or absence of a species, and modelled with a binomial generalized linear model, the residuals take values along a probability scale, and thus have a different distribution than the parent binary response variable. This difference can make estimates of explained deviance in the second model incomparable with those from the first model. A clear statistical rationale needs to be developed before the procedure can be safely generalized to non-OLS situations.

A further complication arises in calculations of predictor contribution if predictor interactions are allowed in the models. The two approaches previously discussed – i.e. hierarchical partitioning and variance partitioning – assume additive effects of predictors. When interaction terms are allowed, their contribution cannot be easily attributed to any of the predictors involved in the interaction and new solutions need to be developed (Guisan *et al.*, 2006b).

The use of automated solutions to predictor selection and contribution should not be seen as a substitution for pre-selecting sound ecophysiological predictors based on deep knowledge of the biogeographical and ecological theory (Austin, 2002), No meaningful model can be built without knowledge of the species' ecology, population dynamics and sensitivity to human and other disturbances and the approaches described previously can hardly be successful if the initial number of predictors is too large. A crucial question is thus 'How much [variables or explained variance] is enough?' (Huston, 2002). For instance, when predicting the likely impact of climate change on species distributions, across large regions, one can reasonably assume that using climate

predictors alone should prove sufficient to assess the main changes in distributions (e.g. Pearson & Dawson, 2003). Nonetheless, it is reasonable to ask what else is left, when all the climate-related variance has been explained. Answering this question requires quantifying how much climate can explain species distributions compared to other predictors, such as soils, site history, human influences, or other factors (e.g. Pearson *et al.*, 2004; Thuiller *et al.*, 2004a; del Barrio *et al.*, 2006; Coudun *et al.*, 2006; Luoto *et al.*, 2006). Clearly, this amount will depend upon the type of organism being modelled and the spatial scale considered (for discussion see Pearson & Dawson, 2003; Guisan & Thuiller, 2005). Indeed, in most cases, it would be suspect to explain 100% of the variance through climate predictors alone. One can reasonably expect climate to explain more in regions with extreme climates, e.g. alpine or arctic environments, than in tropical regions where biotic interactions are likely to play a greater role. We know very little about these issues. A further complication is that the maximum possible variance explained by all climatic predictors in a model does not necessarily correspond to the absolute maximum climatic variance that can be potentially explained, because some key climatic predictors are likely to be missing. Further research is needed to assess the varying influence of climate on the distribution of different groups of species (particularly animals) in different regions (e.g. Thuiller *et al.*, 2004a; Heikkinen *et al.*, 2005; Luoto *et al.*, 2006). Once the maximum climatic variance has been assessed for many species in many distinct groups, one should be able to relate these amounts to various species traits and the climatic characteristics of the study areas, in order to seek ways to predict *a priori* the maximum explainable climatic variance for a given organism.

CHALLENGE 5: IMPROVED MODEL EVALUATION STRATEGIES

Model evaluation describes the testing process required to justify the acceptance of a model for its intended purpose. The semantics, logic and philosophy underpinning model evaluation are thorny and have been widely debated in the environmental modelling literature (e.g. Konikow & Bredehoeft, 1992; Oreskes *et al.*, 1994; Rykiel, 1996). However, in the context of species' distribution modelling a rather pragmatic question still remains to be addressed convincingly: *What evaluation procedure is required to justify acceptance of models fitted for a variety of different purposes*? Although most modellers would accept that validation (a model's ability to predict events with independent test data) is preferable to verification (a model's ability to fit the training data), there are cases where simple forms of verification are sufficient, and others where model evaluation may be unnecessary or even impossible. Because model evaluation is inextricably related to its intended purpose, the choice of the evaluation strategy needs to be explicitly related to the subject and goals of modelling.

We suggest that evaluation strategies be discussed in the context of three possible uses: description, understanding and prediction. A purely descriptive model is one in which the strength of the relationship between a response and predictor variables is measured without additional considerations. A model aimed at understanding goes a step further in that the hypotheses concerning the relationships being measured are tested. Finally a predictive model is one where confidence in the hypothesized relationships allows projections of observed patterns into independent situations. Complexity of model evaluation increases from explanation to prediction to the point where models that simply seek to describe a given pattern may not need to be evaluated, whereas the evaluation of models aiming at prediction is desirable but not always conceptually possible.

Take the simplest case of 'curve-fitting' models, commonly applied to the product of unsupervized data mining (reviewed by Hastie *et al.*, 2001). The goal of such models is to describe a given pattern; the greater the flexibility given to 'curve-fitting' (e.g. number of polynomials in GLM, smoothing splines in GAM, nodes in CART, or hidden layers in artificial neural networks), the greater the likelihood that models will overfit the data. Overfitting is not a problem if the goal is to describe a pattern and simultaneously reduce false negatives, i.e. true observations that are not predicted by models. This is the case of models seeking to convert all presence records of a species into a suitability score for an application, for example, in conservation planning methodologies (e.g., Araújo & Williams, 2000; Araújo *et al.*, 2002; Cabeza *et al.*, 2004). Here, simple forms of verification, such as measuring the number of false negatives, can be implemented to check whether models are performing as intended. We expect that, if appropriately implemented, models fitting more flexible and complex response shapes will reduce the number of false negatives. In this context, verification is a tool to evaluate model implementation rather than model performance.

When models are implemented to understand the likely mechanisms governing species' distributions, model evaluation [verification] is appropriately used to assess the 'robustness' of inferred mechanisms. In a correlative setting, variable selection is sensitive to the training data (e.g. Thuiller *et al.*, 2004b; Coudun *et al.*, 2006; Luoto *et al.*, 2006; Maggini *et al.*, 2006). Hence, a major purpose of model evaluation is to measure the stability of selected variables using, e.g. grouped cross-validation (also known as k fold partitioning), bootstrap and jackknife approaches (also known as leave-one-out; see extended discussion in the 'Challenge 4' section). It is important to note that the purpose of model evaluation in a 'seeking to understand' context is not to measure the accuracy of model projections, but the probability that any given predictor variable is selected as the potential driving force of existing species distributions.

Finally, when models are implemented for prediction, it is generality or transferability (the ability to predict independent events) that needs to be evaluated [validation]. Examples include models that use sample distribution data to allow predictions of species' potential distributions within the same region and resolution (e.g. Guisan *et al.*, 1998; Olden & Jackson, 2002;

Segurado & Araújo, 2004; Elith *et al.*, 2006), same region but different resolution (e.g. Araújo *et al.*, 2005b; McPherson *et al.*, 2006), different regions (e.g. Fielding & Haworth, 1995; Thuiller *et al.*, 2005b; Randin *et al.*, 2006; Segurado *et al.*, 2006), or different time periods (e.g. Austin, 1992; Huntley *et al.*, 1995; Sykes *et al.*, 1996; Berry *et al.*, 2002; Peterson *et al.*, 2002; Thuiller *et al.*, 2005a; Araújo *et al.*, 2006; Harrison *et al.*, 2006). It is precisely in the context of the two latter examples that over-fitted models are likely to be less robust than more parsimonious solutions (Randin *et al.*, 2006).

At least two difficulties arise when validating niche models for predictive purposes. First, models predict events (e.g. species distributions in unsampled locations, different regions or times) that have a degree of independence from the events used to make the predictions. Therefore, the data used to train the models should be evaluated against independent data. However, a recent review showed that the majority of studies using niche models to project distributions into the future use a simple form of verification (the resubstituition approach) in which the same data used for model calibration are used to test the models (Araújo *et al.*, 2005a); a similar observation was documented for studies modelling current species distributions (e.g. Olden & Jackson, 2000). The ability to describe a given situation by training of given model parameters does not imply that models are able to predict independent situations with accuracy, as demonstrated for models projecting distributions of species into different regions (Randin *et al.*, 2006) or times (Araújo *et al.*, 2005a). Attempts to circumvent this problem have included one-time data splitting, bootstrap and jack-knife approaches. These approaches share the assumption that randomly selected samples from the original data constitute suitable independent observations for testing, but independence is not guaranteed if the training and test sets are spatially autocorrelated (Araújo *et al.*, 2005a). The problem of non-independence is not overridden by carrying out additional field sampling for testing models within the modelled region (e.g. Feria & Peterson, 2002; Raxworthy *et al.*, 2003; Elith *et al.*, 2006; Stockman *et al.*, 2006) because test data may be spatially autocorrelated with the data used to train the models, thus providing unrealistic estimates of model performance outside the training set. The solution for this problem implies that models are tested with data recorded in different regions or times. In the specific case of models used to project species distributions under future climate change scenarios, it is usually not possible to evaluate predictive performance of models because events being predicted have yet not occurred. An alternative is to make backward projections (hind-casting) of species distributions and use the temporal record for model validation (e.g. Martinez-Meyer *et al.*, 2004; Martínez-Meyer & Peterson, 2006).

Secondly, even when there is suitable independent distribution data for the testing of models (e.g. based on records compiled from fossil and pollen records) it is often the case that models cannot be validated because of inaccurate formulation of the modelling problem (see discussion in the 'Clarification of the niche concept' section). For example, the chief aim of niche modelling – to characterize a species' suitable environmental space (or *potential* habitat range) – is usually tested against the realized spatial distribution of a species. However, unless spatial explicit features of species distributions are modelled, models remain untested by simple comparison of observed and modelled data since only *realized*, rather than *potential*, ranges against which to validate models are used (Araújo *et al.*, 2005c).

CONCLUDING REMARKS

We have discussed five topics that are critical for the development of the science of species distribution modelling. They included the following.

Clarification of the niche concept

A majority of modellers subscribes to Hutchinson's realized vs. fundamental niche framework. Nonetheless, there are conflicting views about what the models truly represent. Conflicting interpretations arise from ambiguities in the original formulation of the niche concept. We argue that a simpler definition, such as the modern translation of Grinnell's original formulation (Grinnell, 1917, 1924) by Chase & Leibold (2003), suffices for the purposes of niche modelling: the environmental conditions that allow a species to satisfy its minimum requirements so that birth rate of a local population is equal to or greater than its death rate. It is also argued that a clearer distinction between the modelling of niche and spatial-explicit features would be useful. Niche models yield projections of *potential habitats for species*, but when niche and spatial factors (e.g. dispersal) are combined, models yield projections of *potential geographical distributions of species*. This distinction should be clarified further.

Improved designs for sampling data for building models

Model outputs are sensitive to sampling biases in the input data. Even though well-designed recording schemes are more likely to produce useful data for modelling, it is the poor quality of data that justifies the use of species' distribution models in many applications. Sub-sampling of existing data to remove or reduce existing biases in the data may help improve the predictive power of the models. However, this procedure requires that there are enough observations in the data to allow removal of sub-samples without decreasing the ability of models to fit the data. Hence, a better alternative, albeit more expensive procedure, is to target strategically the location of additional samples in the field. Model-based environmental stratifications can help target the minimum number of samples that is required to obtain a representative coverage of niche space. Yet, these ideas are in their infancy and more effort is required to assess their strengths and limitations in the context of niche-based modelling of species distributions.

Improved parameterization

Different parameterizations of the same model may yield considerably different projections of species potential habitats or distributions. The existence of variability in model outputs due to differences in model parameterization constitutes a form of uncertainty that has been previously underestimated. A better understanding is necessary of *why* and *when* different modelling techniques, or parameterizations of the same technique, provide different results.

Improved model selection and predictor contribution

There are a number of relatively novel model selection strategies available that should be more widely used by ecological modellers. Calculating the individual contribution of predictors in multiple models is a distinct, and more difficult, problem. Some of the proposed solutions, like hierarchical or variance partitioning, still require further testing by ecologists and statisticians.

Improved model evaluation

Evaluation of models is inextricably related to their intended purpose. Even though this seems a trivial statement, modellers often use model evaluation strategies without considering the object and goals of the modelling exercise. We argue that evaluation strategies can be usefully discussed in the context of three possible implementations: explanation, understanding and prediction. In some cases validation (a model's ability to predict events with independent test data) is preferable to verification (a model's ability to fit the training data), but there are circumstances where simple forms of verification are sufficient, and others where model evaluation may be unnecessary or even impossible.

By selecting these five challenges to species distribution modelling, we hope to have highlighted important areas for future work but have doubtless thereby neglected others. Some issues have been recently debated in the literature (e.g. Guisan & Thuiller; 2005, Ferrier & Guisan, 2006; Guisan *et al.*, 2006b; Heikkinen *et al.*, 2006; Pearce & Boyce, 2006; Peterson, 2006), but there is still considerable ground for discussion and clarification. Some of the most prominent challenges missing from our discussion include (1) the effects of spatial and temporal autocorrelation on models – when does spatial and temporal autocorrelation matter and what can be done about it? (e.g. Segurado *et al.*, 2006); (2) the effects of geographical extent and resolution – how does variation in the extent and resolution of the studied area affect variable selection and the performance of models? (e.g. Nogués-Bravo & Araújo, 2006); (3) the strategies for selecting pseudo-absences for model fitting – do approaches that stratify the selection of pseudo-absences improve the predictive ability of models on independent data when compared with pseudo-absences chosen randomly? (e.g. Zaniewski *et al.*, 2002; Engler *et al.*, 2004); and (4) the rules used for transforming modelled probabilities of occurrence into presence absence – how does choice of the rule (e.g. maximizing the kappa statistic or using AUC index) affect the characterization of the niche or spatial distributions of species? (e.g. Liu *et al.*, 2005).

The perceived usefulness of niche-based species distribution models for applied biogeography and ecology is partly dependent on the successful resolution of some of these problems. We hope to have convinced the reader that there is a need for deepening the debate on the strengths and limitations of existing species distribution models and undertake more rigorous assessments of the sensitivity of model outcomes to starting assumptions and parameters. Our goal with this essay was to spark discussion and stimulate researchers to invest their energies into this vibrant and relatively new field of enquiry.

ACKNOWLEDGEMENTS

We thank for discussion and useful comments on the manuscript: M. Bush, J. Hortal, G. Le Lay, J.M. Lobo, D. Nogués-Bravo, P. Pearman, W. Thuiller, A. Townsend Peterson, C.F. Randin, J. Soberón, N.G. Yoccoz and one anonymous referee. M.B.A. is funded by EC FP6 ALARM project (GOCE-CT-2003-506675). A.G. thanks the Swiss National Science Foundation for support (SNF Grant Project No 3100A0-110000 and NCCR Plant Survival TG3 Subproject).

REFERENCES

Amarasekare, P. (2003) Competitive coexistence in spatially structured environments: a synthesis. *Ecology Letters*, **6,** 1109–1122.

Anderson, R.A., Peterson, A.T. & Gomez-Laverde, M. (2002) Using niche-based GIS modelling to test geographic predictions of competitive exclusion and competitive release in South American pocket mice. *Oikos*, **98,** 3–16.

Araújo, M.B. & Pearson, R.G. (2005) Equilibrium of species' distributions with climate. *Ecography*, **28,** 693–695.

Araújo, M.B. & Williams, P.H. (2000) Selecting areas for species persistence using occurrence data. *Biological Conservation*, **96,** 331–345.

Araújo, M.B., Humphries, C.J., Densham, P.J., Lampinen, R., Hagemeijer, W.J.M., Mitchell-Jones, A.J. & Gasc, J.P. (2001) Would environmental diversity be a good surrogate for species diversity? *Ecography*, **24,** 103–110.

Araújo, M.B., Williams, P.H. & Fuller, R.J. (2002) Dynamics of extinction and the selection of nature reserves. *Proceedings of the Royal Society London Series B, Biological Sciences*, **269,** 1971–1980.

Araújo, M.B., Densham, P.J. & Humphries, C.J. (2003) Predicting species diversity with ED: the quest for evidence. *Ecography*, **26,** 380–383.

Araújo, M.B., Densham, P.J. & Williams, P.H. (2004) Representing species in reserves from patterns of assemblage diversity. *Journal of Biogeography*, **31,** 1037–1050.

Araújo, M.B., Pearson, R.G., Thuiller, W. & Erhard, M. (2005a) Validation of species-climate impact models under climate change. *Global Change Biology*, **11,** 1504–1513.

Araújo, M.B., Thuiller, W., Williams, P.H. & Reginster, I. (2005b) Downscaling European species atlas distributions to a finer resolution: implications for conservation planning. *Global Ecology and Biogeography*, **14,** 17–30.

Araújo, M.B., Whittaker, R.J., Ladle, R. & Erhard, M. (2005c) Reducing uncertainty in projections of extinction risk from climate change. *Global Ecology and Biogeography*, **14,** 529–538.

Araújo, M.B., Thuiller, W. & Pearson, R.G. (2006) Climate warming and the decline of amphibians and reptiles in Europe. *Journal of Biogeography*. doi: 10.1111/j.1365-2699.2006.01482.x.

Austin, M.P. (1992) Modelling the environmental niche of plants – implications for plant community response to elevated CO_2 levels. *Australian Journal of Botany*, **40,** 615–630.

Austin, M.P. (2002) Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological Modelling*, **157,** 101–118.

Austin, M.P., Nicholls, A.O. & Margules, C.R. (1990) Measurement of the realized qualitative niche: environmental niches of five *Eucalyptus* species. *Ecological Monographs*, **60,** 161–177.

Balmford, A. & Gaston, K.J. (1999) Why biodiversity surveys are good value? *Nature*, **18,** 204–205.

del Barrio, G., Harrison, P.A., Berry, P.M., Butt, N., Sanjuan, M.E., Pearson, R.G. & Dawson, T. (2006) Integrating multiple modelling approaches to predict the potential impacts of climate change on species' distributions in contrasting regions: comparison and implications for policy. *Environmental Science and Policy*, **9,** 129–147.

Berry, P.M., Dawson, T.E., Harrison, P.A. & Pearson, R.G. (2002) Modelling potential impacts of climate change on the bioclimatic envelope of species in Britain and Ireland. *Global Ecology and Biogeography*, **11,** 453–462.

Bonn, A. & Gaston, K.J. (2005) Capturing biodiversity: selecting priority areas for conservation using different criteria. *Biodiversity and Conservation*, **14,** 1083–1100.

Borcard, D., Legendre, P. & Drapeau, P. (1992) Partialling out the spatial component of ecological variation. *Ecology*, **73,** 1045–1055.

Brotons, L., Thuiller, W., Araújo, M.B. & Hirzel, A.H. (2004) Presence-absence versus presence-only based habitat suitability models for bird atlas data: the role of species ecology and prevalence. *Ecography*, **27,** 437–448.

Buckland, S.T., Goudie, I.B.J. & Borchers, D.L. (2000) Wildlife population assessment: past developments and future directions. *Biometrics*, **56,** 1–12.

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

Cabeza, M., Araújo, M.B., Wilson, R.J., Thomas, C.D., Cowley, M.J.R. & Moilanen, A. (2004) Combining probabilities of occurrence with spatial reserve design. *Journal of Applied Ecology*, **41,** 252–262.

Callaway, R.M., Brooker, R.W., Choler, P., Kikvidze, Z., Lortie, C.J., Michalet, R., Paolini, L., Pugnaire, F.I., Newingham, B., Aschehoug, E.T., Armas, C., Kikodze, D. & Cook, B.J. (2002) Positive interactions among alpine plants increase with stress. *Nature*, **417,** 844–848.

Chase, J.M. & Leibold, M.A. (2003) *Ecological niches – linking classical and contemporary approaches*. The University of Chicago Press, Chicago, IL.

Coudun, C., Gégout, J.C., Piedallu, C. & Rameau, J.C. (2006) Soil nutritional factors improve plant species distribution an illustration with *Acer campestre* (L.) in France. *Journal of Biogeography* (doi: 10.1111/j.1365.2699.2005.01443.x).

Edwards, T.C.J., Cutler, R., Zimmermann, N.E., Geiser, L. & Alegria, J. (2005) Model-based stratifications for enhancing the detection of rare ecological events: lichens as a case study. *Ecology*, **86,** 1081–1090.

Elith, J. (2000) Quantitative methods for modeling species habitat: comparative performance and an application to Australian plants. *Quantitative Methods in Conservation Biology* (ed. by S. Ferson & M.A. Burgman), pp. 39–58. Springer, New York.

Elith, J., Graham, C.H., Anderson, R.P., Dudik, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettman, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L., Loiselle, B.A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J.M., Peterson, A.T., Phillips, S., Richardson, K., Schachetti Pereira, R., Schapire, R.E., Soberón, J., Williams, S.E., Wisz, M. & Zimmermann, N.E. (2006) Novel methods improve predictions of species' distributions from occurrence data. *Ecography*, **29,** 129–151.

Engler, R., Guisan, A. & Rechsteiner, L. (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.

Faith, D.P. & Walker, P.A. (1996) Environmental diversity: on the best-possible use of surrogate data for assessing the relative biodiversity of sets of areas. *Biodiversity and Conservation*, **5,** 399–415.

Feria, T.P. & Peterson, A.T. (2002) Prediction of bird community composition based on point-occurrence data and inferential algorithms: a valuable tool in biodiversity assessments. *Diversity and Distributions*, **8,** 49–56.

Ferrier, S. & Guisan, A. (2006) Spatial modelling of biodiversity at the community level. *Journal of Applied Ecology*, **43,** 393–404.

Ferrier, S., Drielsma, M., Manion, G. & Watson, G. (2002) Extended statistical approaches to modelling spatial pattern in biodiversity in north-east New South Wales: II. Community-level modelling. *Biodiversity and Conservation*, **11,** 2309–2338.

Ferrier, S., Powell, G.V.N., Richardson, K.S., Manion, G., Overton, J.M., Allnutt, T.F., Cameron, S.E., Mantle, K., Burgess, N.D., Faith, D.P., Lamoreux, J.F., Kier, G., Hijmans, R.J., Funk, V.A., Cassis, G.A., Fisher, B.L., Flemons,

P., Lees, D., Lovett, J.C. & Van Rompaey, R.S.A.R. (2004) Mapping more of terrestrial biodiversity for global conservation assessment: a new approach to integrating disparate sources of biological and environmental data. *BioScience*, **54,** 1101–1109.

Fielding, A.H. & Haworth, P.F. (1995) Testing the generality of bird-habitat models. *Conservation Biology*, **9,** 1466–1481.

Graham, C.H. (2001) Factors influencing movement patterns of Keel-billed Toucans in a fragmented tropical landscape in Southern Mexico. *Conservation Biology*, **15,** 1789–1798.

Grinnell, J. (1917) Field tests of theories concerning distributional control. *The American Naturalist*, **51,** 115–128.

Grinnell, J. (1924) Geography of evolution. *Ecology*, **5,** 225–229.

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

Guisan, A. & Zimmermann, N.E. (2000) Predictive habitat distribution models in ecology. *Ecological Modelling*, **135,** 147–186.

Guisan, A., Theurillat, J.-P. & Kienast, F. (1998) Predicting the potential distribution of plant species in an alpine environment. *Journal of Vegetation Science*, **9,** 65–74.

Guisan, A., Edwards, J., Thomas, C. & Hastie, T. (2002) Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecological Modelling*, **157,** 89–100.

Guisan, A., Broennimann, O., Engler, R., Vust, M., Yoccoz, N.G., Lehmann, A. & Zimmermann, N.E. (2006a) Using niche-based distribution models to improve the sampling of rare species. *Conservation Biology*, **20,** 501–511.

Guisan, A., Lehmann, A., Ferrier, S., Aspinall, R., Overton, R., Austin, M.P. & Hastie, T. (2006b) Making better biogeographic predictions of species distribution. *Journal of Applied Ecology*, **43,** 386–392.

Gutiérrez, D., Fernández, P., Seymour, A.S. & Jordano, D. (2005) Habitat distribution models: are mutualist distributions good predictors of their associates? *Ecological applications*, **15,** 3–18.

Harrell, F.E. (2001) *Regression modeling strategies – with applications to linear models, logistic regression and survival analysis*. Springer-Verlag, New York.

Harrison, P.A., Berry, P.M., Butt, N. & New, M. (2006) Modelling climate change impacts on species' distributions at the European scale: implications for conservation policy. *Environmental Science and Policy*, **9,** 116–128.

Hastie, T., Tibshirani, R. & Friedman, J. (2001) *The elements of statistical learning: data mining, inference, and prediction*. Springer, New York.

Heikkinen, R.K., Luoto, M., Kuussaari, M. & Pöyry, J. (2005) New insights to butterfly–environment relationships with partitioning methods. *Proceedings of the Royal Society London Series B, Biological Sciences*, **272,** 2203–2210.

Heikkinen, R.K., Luoto, M., Araújo, M.B., Virkkala, R., Thuiller, W. & Sykes, M.T. (2006) Methods and uncertainties in bioclimatic modelling under climate change. *Progress in Physical Geography* (in press).

Hirzel, A. & Guisan, A. (2002) Which is the optimal sampling strategy for habitat suitability modelling. *Ecological Modelling*, **157,** 331–341.

Hortal, J. & Lobo, J.M. (2005) An ED-based protocol for optimal sampling of biodiversity. *Biodiversity and Conservation*, **14,** 2913–2947.

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. *Journal of Biogeography*, **22,** 967–1001.

Huston, M.A. (2002) Introductory essay: critical issues for improving predictions. *Predicting species occurrences: issues of accuracy and scale* (ed. by J.M. Scott, P.J. Heglund, M.L. Morrison, J.B. Haufler, M.G. Raphael, W.A. Wall and F.B. Samson), pp. 7–21. Island Press, Covelo, CA.

Hutchinson, G.E. (1957) Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, **22,** 145–159.

Johnson, J.B. & Omland, K.S. (2004) Model selection in ecology and evolution. *Trends in Ecology and Evolution*, **19,** 101–108.

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

Konikow, L.F. & Bredehoeft, J.D. (1992) Ground-water models cannot be validated. *Advances in Water Resources*, **15,** 75–83.

Leathwick, J.R. & Austin, M.P. (2001) Competitive interactions between tree species in New Zealand's old-growth indigenous forests. *Ecology*, **82,** 2560–2573.

Liu, C., Berry, P.M., Dawson, T.P. & Pearson, R.G. (2005) Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, **28,** 385–393.

Luoto, M., Heikkinen, R.K., Pöyry, J. & Saarinen, K. (2006) Determinants of biogeographical distribution of butterflies in boreal regions. *Journal of Biogeography*. doi: 10.1111/j.1365-2699.2005.01395.x.

MacNally, R. (2002) Multiple regression and inference in ecology and conservation biology: further comments on identifying important predictor variables. *Biodiversity and Conservation*, **11,** 1397–1401.

MacNally, R. & Walsh, C.J. (2004) Hierarchical partitioning of public domain software. *Biodiversity and Conservation*, **13,** 659–660.

Maggini, R., Lehmann, A., Zimmermann, N.E. & Guisan, A. (2006) Improving generalized regression analysis for spatial predictions of forest communities. *Journal of Biogeography*. doi: 10.1111/j.1365-2699.2006.01465.x.

Martínez-Meyer, E. & Peterson, A.T. (2006) Conservatism of ecological niche characteristics in North American plant species over the Pleistocene-to-Recent transition. *Journal of Biogeography* (doi: 10.1111/j.1365.2699.2006.01612.x).

Martinez-Meyer, E., Townsend Peterson, A. & Hargrove, W.W. (2004) Ecological niches as stable distributional constraints on mammal species, with implications for

Pleistocene extinctions and climate change projections for biodiversity. *Global Ecology and Biogeography*, **13,** 305–314.

McPherson, J.M., Jetz, W. & Rogers, D.J. (2006) Using coarse-grained occurrence data to predict species distributions at finer spatial resolutions – possibilities and limitations. *Ecological Modelling*, **192,** 499–522.

Mohler, C.L. (1983) Effect of sampling pattern on estimation of species distributions along gradients. *Vegetatio*, **54,** 97–102.

Nogués-Bravo, D. & Araújo, M.B. (2006) Species richness, area and climate correlates. *Global Ecology and Biogeography*, doi: 10.1111/j.1566-822x.2006.00240.x.

Olden, J.D. & Jackson, D.A. (2000) Torturing data for the sake of generality: how valid are our regression models? *Ecoscience*, **7,** 501–510.

Olden, J.D. & Jackson, D.A. (2002) A comparison of statistical approaches for modelling fish species distributions. *Freshwater Biology*, **47,** 1976–1995.

Oreskes, N., Shrader-Frechette, K.S. & Belitz, K. (1994) Verification, validation, and confirmation of numerical models in the earth sciences. *Science*, **263,** 641–646.

Pearce, J.L. & Boyce, M.S. (2006) Modelling distribution and abundance with presence-only data. *Journal of Applied Ecology*, **43,** 405–412.

Pearson, R.G. & Dawson, T.E. (2003) Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology and Biogeography*, **12,** 361–371.

Pearson, R.G., Dawson, T.E. & Liu, C. (2004) Modelling species distributions in Britain: a hierarchical integration of climate and land-cover data. *Ecography*, **27,** 285–298.

Pearson, R.G., Thuiller, W., Araújo, M.B., Martinez-Meyer, E., Brotons, L., McClean, C., Miles, L.J., Segurado, P., Dawson, T.E. & Lees, D.C. (2006) Model-based uncertainty in species' range prediction. *Journal of Biogeography*. doi: 10.1111/j.1365-2699.2006.01460.x.

Peterson, A.T. (2006) Uses and requirements of ecological niche models and related distributional models. *Biodiversity Informatics* (in press).

Peterson, A.T. & Cohoon, K.C. (1999) Sensitivity of distributional prediction algorithms to geographic data completeness. *Ecological Modeling*, **117,** 159–164.

Peterson, A.T., Ortega-Huerta, M.A., Bartley, J., Sanchez-Cordero, V., Soberon, J., Buddemeier, R.H. & Stockwell, D.R.B. (2002) Future projections for Mexican faunas under global climate change scenarios. *Nature*, **416,** 626–629.

Ponder, W.F., Carter, G.A., Flemons, P. & Chapman, R.R. (2001) Evaluation of museum collection data for use in biodiversity assessment. *Conservation Biology*, **15,** 648–657.

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

Randin, C.F., Dirnböck, T., Dullinger, S., Zimmermann, N.E., Zappa, M. & Guisan, A. (2006) Are niche-based species distribution models transferable in space? *Journal of Biogeography*. doi: 10.1111/j.1365-2699.2006.01466.x.

Raxworthy, C.J., Martinez-Meyer, E., Horning, N., Nussbaum, R.A., Schneider, G.E., Ortega-Huerta, M.A. & Townsend Peterson, A. (2003) Predicting distributions of known and unknown reptile species in Madagascar. *Nature*, **426,** 837–841.

Reutter, B.A., Helfer, V., Hirzel, A.H. & Vogel, P. (2003) Modelling habitat-suitability on the base of museum collections: an example with three sympatric *Apodemus* species from the Alps. *Journal of Biogeography*, **30,** 581–590.

Rykiel, E.J., Jr (1996) Testing ecological models: the meaning of validation. *Ecological Modelling*, **90,** 229–244.

Sarkar, S., Justus, J., Fuller, T., Kelley, C., Garson, J. & Mayfield, M. (2005) Effectiveness of environmental surrogates for the selection of conservation area networks. *Conservation Biology*, **19,** 815–825.

Segurado, P. & Araújo, M.B. (2004) An evaluation of methods for modelling species distributions. *Journal of Biogeography*, **31,** 1555–1568.

Segurado, P., Araújo, M.B. & Kunin, W.E. (2006) Consequences of spatial autocorrelation for niche-based models. *Journal of Applied Ecology*, **43,** 433–444.

Smith, L.A. (2002) What might we learn from climate forecasts? *Proceedings of the National Academy of Sciences of the United States of America*, **99,** 2487–2492.

Soberón, J. & Peterson, A.T. (2005) Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics*, **2,** 1–10.

Stockman, A.K., Beamer, D.A. & Bond, J.E. (2006) An evaluation of a GARP model as an approach to predicting the spatial distribution of non-vagile invertebrate species. *Diversity and Distributions*, **12,** 81–89.

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

Svenning, J.-C. & Skov, F. (2004) Limited filling of the potential range in European tree species. *Ecology Letters*, **7,** 565–573.

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 climate. *Journal of Biogeography*, **23,** 203–233.

Thomas, C.D., Cameron, A., Green, R.E., Bakkenes, M., Beaumont, L.J., Collingham, Y., Erasmus, B.F.N., de Siqueira, M.F., Grainger, A., Hannah, L., Hughes, L., Huntley, B., van Jaarsveld, A.S., Midgley, G.F., Miles, L.J., Ortega-Huerta, M.A., Townsend Peterson, A., Phillips, O. & Williams, S.E. (2004) Extinction risk from climate change. *Nature*, **427,** 145–148.

Thuiller, W., Araújo, M.B. & Lavorel, S. (2003) Generalized models vs. classification tree analysis: predicting spatial distributions of plant species at different scales. *Journal of Vegetation Science*, **14,** 669–680.

Thuiller, W., Araújo, M.B. & Lavorel, S. (2004a) Do we need land-cover data to model species distributions in Europe? *Journal of Biogeography*, **31,** 353–361.

Thuiller, W., Brotons, L., Araújo, M.B. & Lavorel, S. (2004b) 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. & Prentice, I.C. (2005a) Climate change threats plant diversity in Europe. *Proceedings of the National Academy of Sciences of the United States of America*, **102,** 8245–8250.

Thuiller, W., Richardson, D.M., Pysek, P., Midgley, G.F., Hughes, G.O. & Rouget, M. (2005b) Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. *Global Change Biology*, **11,** 2234–2250.

Trakhtenbrot, A. & Kadmon, R. (2005) Environmental cluster analysis as a tool for selecting complementary networks of conservation sites. *Ecological Applications*, **15,** 335–345.

Travis, J.M.J., Brooker, R.W. & Dytham, C. (2005) The interplay of positive and negative species interactions across an environmental gradient: insights from an individual-based simulation model. *Biology Letters*, **1,** 5–8.

Visscher, D.R. (2006) GPS measurement error and resource selection functions in a fragmented landscape. *Ecography*, **29,** 458–464.

Whittaker, R.J., Araújo, M.B., Jepson, P., Ladle, R., Watson, J.E. & Willis, K.J. (2005) Conservation biogeography: assessment and prospect. *Diversity and Distributions*, **11,** 3–23.

Yamamura, N., Higashi, M., Behera, N. & Wakaon, J.Y. (2004) Evolution of mutualism through spatial effects. *Journal of Theoretical Biology*, **226,** 421–428.

Yoccoz, N.G. & Chessel, D. (1988) Ordination sous contraintes de relevés d'avifaune: elimination d'effets dans un plan d'observation à deux facteurs. *Compte-Rendus de l'Académie des Sciences, Paris, Série III*, **307,** 189–194.

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

BIOSKETCHES

Miguel B. Araújo has a PhD from the University of London. He is a Research Fellow at the National Museum of Natural Sciences in Madrid (http://mncn.csic.es/maraujo.htm), a Senior Research Associate to the Oxford University Centre for the Environment, and visiting Associate Professor at the University of Copenhagen. He has a range of interests in macroecology and conservation biogeography having focused in reserve selection, species distribution modelling, climate change and species diversity.

Antoine Guisan has a PhD from the University of Geneva. He is the leader of the Spatial Ecology Group at the University of Lausanne (http://ecospat.unil.ch). He specialized in predictive modelling of species distributions with a special interest in alpine landscapes, climate change research and invasive species.

Editors: Miguel B. Araújo and Mark B. Bush

This paper is part of the Special Issue, Species distribution modelling: methods, challenges and applications, which owes its origins to the workshop on Generalized Regression Analyses and Spatial Predictions: Grasping Ecological Patterns from Species to Landscape held at the Centre Pro Natura Aletsch in Riederalp, Switzerland, in August 2004.