Review TRENDS in Ecology and Evolution Vol.22 No.1

Ensemble forecasting of species distributions

Miguel B. Araújo[1] and Mark New[2]

[1] Department of Biodiversity and Evolutionary Biology, National Museum of Natural Sciences, CSIC, C/Gutiérrez Abascal, 2, 28006, Madrid, Spain
[2] Climate Research Laboratory, Oxford University Centre for the Environment, South Parks Road, Oxford, UK, OX1 3QY

Concern over implications of climate change for biodiversity has led to the use of bioclimatic models to forecast the range shifts of species under future climate-change scenarios. Recent studies have demonstrated that projections by alternative models can be so variable as to compromise their usefulness for guiding policy decisions. Here, we advocate the use of multiple models within an ensemble forecasting framework and describe alternative approaches to the analysis of bioclimatic ensembles, including bounding box, consensus and probabilistic techniques. We argue that, although improved accuracy can be delivered through the traditional tasks of trying to build better models with improved data, more robust forecasts can also be achieved if ensemble forecasts are produced and analysed appropriately.

Introduction

Attempts to predict climate change impacts on species distributions have often relied on the bioclimatic 'envelope' modelling approach, whereby empirical relationships between present-day distributions of species and climate variables are used to estimate distributions of species under future climate scenarios [1–4]. For several (usually) pragmatic reasons, modelling typically involves selecting a favoured technique from a range of alternatives, and then justifying the choice by making reference to one or more published studies. However, despite claims of superiority for any given technique [5–10], independent evaluations of models have often been unable to demonstrate the pre-eminence of any single one [11–13].

Furthermore, studies have shown that projections by alternative models can be so variable as to compromise even the simplest assessment of whether species distributions should be expected to contract or expand for any given climate scenario. For example, Pearson and colleagues [14] applied nine well documented bioclimatic modelling techniques to a standardised data set of four South African plant species and compared consistency in range predictions under current and future climates. Predicted distribution changes varied from a 92% loss to a 322% gain for one species and an equally wide variability in distribution change was predicted for the remaining species. Similarly divergent forecasts have been the rule in studies comparing alternative techniques to assess potential climate change-induced shifts in the distributions of European plants [15], amphibians and reptiles [16], and British breeding birds [17]. These results challenge the common practice of relying on one single method to make forecasts of the responses of species to climate change scenarios or, if one accepted a more sceptical view, the usefulness of bioclimatic modelling in general for climate change impact studies.

Such variability in forecasts is not surprising given that bioclimate 'envelope' models are correlative and therefore sensitive to the data and the mathematical functions utilized to describe the distributions of species in relation to climate parameters. Process-based models that simulate bioclimate interactions from theoretical and experimental knowledge provide an alternative that is less dependent on empirical relationships; however, their implementation at the species level is difficult because of the complex processes and interactions that have to be represented; and variability in forecasts is also common [18].

A solution to intermodel variations that has been used in other fields is to utilize several models (herein termed 'ensembles') and use appropriate techniques to explore the resulting range of projections. Here, we argue that significant improvements on the robustness of a forecast can be achieved if an ensemble approach is used and the results analysed appropriately. We provide an overview of ensemble forecasting, examine alternative techniques for combining ensembles, and discuss their uses and limitations for supporting policy decisions in biodiversity conservation.

Ensemble forecasting

An ensemble, as introduced into statistical mechanics by J. Willard Gibbs in 1878, is an idealization consisting of a large (possibly infinite) number of copies of a system, considered all at once, each of which represents a possible state that the real system might be in at some specified time. A forecast ensemble is more narrowly defined as multiple simulations (copies) across more than one set of initial conditions (IC), model classes (MC), parameters (MP) and boundary conditions (BC) (Box 1). Each combination of IC, MC, MP and BC is one possible state of the system being forecasted.

The idea of ensemble forecasting dates back to 1969, when J.M. Bates and Nobel Prize winner in Economics C.W.J. Granger published their influential article

Corresponding author: Araújo, M.B. (maraujo@mncn.csic.es).
Available online 29 September 2006.

 0169-5347/$ – see front matter doi:10.1016/j.tree.2006.09.010

Box 1. Simulations for producing ensembles of forecasts

Ensembles of forecasts are produced by making multiple simulations across more than one set of IC, MC, MP and BC.

Initial conditions
The state of the real system (e.g. the distribution of species or factors that affect species) at the start of the simulation is often poorly known; it represents an incomplete realization of the real world. Small differences in IC will spawn different model trajectories (so-called 'chaos') around the system attractor(s). Models can be run with different ICs that are consistent with the available observations to explore the sensitivity of the predictions to IC uncertainty.

Model classes
Different MCs (e.g. polynomials and smoothing splines of different orders in general linear or additive models, nodes in classification and regression trees, hidden layers in neural nets, and various forms of process-based models) can all produce simulations that are consistent with available observations, and can be considered as competing and probably equally valid representations of the system of interest.

Model parameters
Statistical models typically have parameters (such as a and b in the linear regression model $y = a + bx$) that are estimated from the data. In classical statistics, the uncertainty in these parameters can be estimated. Multiple forecasts 'sampling' this parameter uncertainty are then possible. For process-based models, many important processes are parameterised, but the exact values of the parameters are unknown. Multiple simulations using different parameter values enable parameter uncertainty to be assessed.

Boundary conditions
Model forecasts are driven by an assumption about a change in BCs, defined broadly as predictors in a statistical model (e.g. climate variables). Typically, these BCs are uncertain, especially in the case of future anthropogenic pollution emissions. Alternative future BCs need to be explored, because the effect of differences between BCs in model predictions of species range shifts can be as large as differences between MCs.

'The combination of forecasts' [19]. Providing that individual forecasts contain some independent information, the authors observed that combined forecasts would yield lower mean error than any of the constituent individual forecasts. The idea had been formally developed by French mathematician P. Laplace in 1818 [20]: 'In combining the results of these two methods, one can obtain a result whose probability law of error will be more rapidly decreasing'. However it was not until the pioneering work of Bates and Granger that the idea of combining forecasts became established. Since then, hundreds of studies have been reviewed [21–23] and applied to a variety of fields of research, including economics [24], management [25], systematics [26], biomedicine [27], meteorology [28] and climatology [29]. Surprisingly, these ideas have been slow to penetrate the ecological literature and it was only recently that ensemble forecasting was explicitly attempted in bioclimatic modelling of species distributions [14–17,30]. However, these early attempts have assessed only a few of the possible combinations of MC, MP and BC in bioclimate models (Box 2; Figure Ie).

Combining ensembles

Given an ensemble of model forecasts, how should they be analysed? The traditional approach consists of identifying the 'best' model from an ensemble of forecasts [12,13,31], where the best model is often judged to be one in which outputs match observed data as closely as possible. However, the ability to describe a given situation by calibration of MPs does not always coincide with the ability to represent adequately new observations using the existing calibration. This problem is particularly severe when predicted observations have a degree of spatial or temporal independence from the calibration set [32–34], which is the case for models projecting distributions of species under future climate scenarios [32].

Instead of picking the 'best' model from an ensemble, a more promising approach is to explore the resulting range of projections (Figure 1). For small ensemble sizes, two contrasting approaches are to use the ensemble to define a 'bounding box', or to generate a 'consensus' forecast (Figure 1b). The appropriate approach is partly dependent on the question being asked, and the costs of being wrong. For medium to large ensemble sizes, raw data can be used generate probability distribution functions (PDFs) for the forecast variable, but these will always be conditional on the sampling strategy across IC, MC, MP and BC (Box 1).

The definition of a bounding box involves identification of the range in forecasts from the ensemble members. The approach is conservative in that it quantifies the range of forecasts, but makes no statement about the probability distribution or conditional probabilities of forecasts within the bounding box [14,15]. It is acknowledged that the ensemble members are a subset of all possible IC–MC–MP–BC combinations and, therefore, only represent some limited projection of reality. Any averaging of ensemble member forecasts is considered to be unlikely to match the truth, and modellers do not attempt to estimate ensemble average or confidence limits for the average.

In consensus forecasting, no assumption is made over the expected frequency distribution of the combined forecasts, but a measure of the central tendency (e.g. the mean or median) is calculated for the ensemble of forecasts [17]. The rationale behind consensus forecasts is that, in averaging several models, the 'signal' that one is interested in emerges from the 'noise' associated with individual model errors and uncertainties. When combining forecasts for consensus, one can produce weighted and unweighted averages. Committee averaging takes a simple unweighted average of the predictions, essentially giving equal probability to each model. Examples include implementations of artificial neural networks, where models are run several times and the mean prediction used [14,15], or when results of different modelling techniques are averaged [16,17,35] (Box 2). With Bayesian approaches, weights are proportional with the posterior probability of each model, which depend on how well the model fits the data and how many parameters are used [36]. Stacking is one analogous procedure for estimating weights with least square regressions [36], but the idea is more general and can be used to obtain weights from measures of accuracy with any modelling technique. The discussion of whether to weight is a long one, but there is evidence from other disciplines that unweighted methods can yield cost-effective solutions [21], although this is only correct if model predictions are equally robust [16,17]. In the simple

Box 2. The production of ensembles of forecasts in practice

The simplest and most widely used approach for modelling species distributions involves a single combination of IC, MC, MP and BC to produce P [1] (Figure Ia). An approach using the concept of ensemble forecasting utilizes *n* realisations of IC (either by boot-strapping or *k* fold cross-validation) to select the model that either best fits the data (e.g. bumping) or that produces a single forecast by averaging all individual P (e.g. bagging) [35] (Figure Ib). A range of alternative MC can also be fitted enabling the model that best fits the data to be selected [31]; alternatively, a single forecast can be produced by averaging all individual P [17] (Figure Ic). Technical developments have recently enabled the production of alternative MP for the same MC, enabling the production of single forecasts by committee averaging (e.g. boosting and random forests) [13,35] (Figure Id), although the same methods can be used to produce several P with different parameterisations [14]. Recent attempts to model species responses to climate change have used ensembles that combine different MC and BC [14–16,30]. In Figure Ie, three MC and three BC are combined to produce nine P, which can then be synthesized using bounding box [14,15] or model averaging [16].

The production of forecast ensembles requires software that automates simulations across a range of IC, MC, MP and BC. The implementation of ensemble forecasting for modelling species distributions is still in its infancy, but there are several modelling techniques that incorporate the notion of ensemble forecasting. Typically, they sparsely sample all possible combinations of IC, MC, MP and BC, yielding an incomplete representation of the potential model uncertainties. In most cases, available techniques simulate across IC and MP or IC and MC (Table I). In the future, software platforms that automate simulations across different techniques will enable comprehensive combinations of IC $\times$ MC $\times$ MP $\times$ BC, yielding potentially large forecast ensembles [38–40].

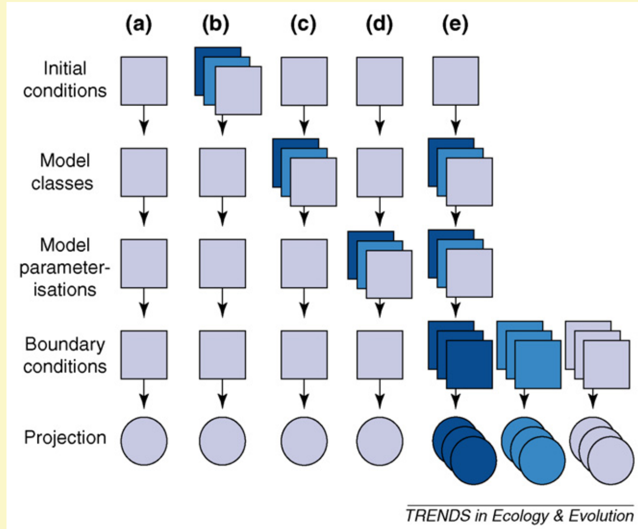

Figure I. Fitting models of species distributions and the production ensembles of forecasts. The squares represent different steps in the production of ensembles and circles represent the predictions from models.

Table I. Modelling techniques that incorporate the notion of ensemble forecasting

Approach	Procedure	Shown in	Refs
Artificial neural networks	Models are run several times and the mean prediction used Alternatively, the best fitting model can be selected	Figure Id	[12,14–17]
Bagging trees	Multiple boot-strapped regression trees are fitted without pruning and the mean prediction used	Figure Ib,d	[35]
Boosted additive trees	The boosting algorithm iteratively calls the regression-tree algorithm to construct an ensemble of trees The regression trees are fitted sequentially on weighted versions of the data, where the weights continuously adjust to take account of observations that are poorly fitted by the preceding models Predictions are finally combined using a majority vote criterion	Figure Ib,d	[13]
GARP	A genetic algorithm evolves a set of rules that best predicts the distribution of species based on bootstrapped samples of available information Rules developed are ranked by predictive performance, and applied to the environmental conditions	Figure Ib,c	[2,13,14]
Maximum entropy (MAXENT)	Algorithm estimates the distribution of a species by finding the probability distribution of maximum entropy (i.e. closest to uniform) subject to the constraint that the expected value of each of a set of features (environmental variables or functions thereof) under this estimated distribution closely matches its empirical average The modelled probability is a 'Gibbs' distribution (i.e. exponential in a weighted sum of the features) The algorithm used to find the MAXENT distribution is similar to boosting	Figure Ib,d	[13,50]
Random forests	Similar to bagging trees but each tree is grown with a randomized subset of predictors Several trees are grown and the predictions aggregated by averaging	Figure Ib,d	[35]

case of committee methods where the median forecast is taken, the consensus will always be more accurate than at least half of the individual forecasts [37]. This is true independently of the distribution of individual forecasts, or their position in relation to the phenomenon of interest (the 'truth'). If the truth falls within the range encompassed by all forecasts, less than half of the individual forecasts will be superior to the median forecast; at worst, if the truth lies outside the forecast range, consensus will be better than 50% of the forecasts.

Probabilistic forecasting can be considered the 'end game' of ensemble forecasting (Figure 1d). We accept the fact that models are different from reality and that, in most cases, we have many possible candidate models that pass some criteria about their ability to represent key aspects of the real system that we are interested in. For a large ensemble across multiple ICs, MPs, MCs and BCs, the frequency distribution of forecasts approaches a probability distribution. New developments in climate modelling where many tens of thousands of simulations [38–40] have produced initial results where frequency histograms are starting to resemble PDFs [41]. However, even these large ensembles are only sparsely sampling the possible IC–MC–MP–BC combinations. Estimating a PDF from

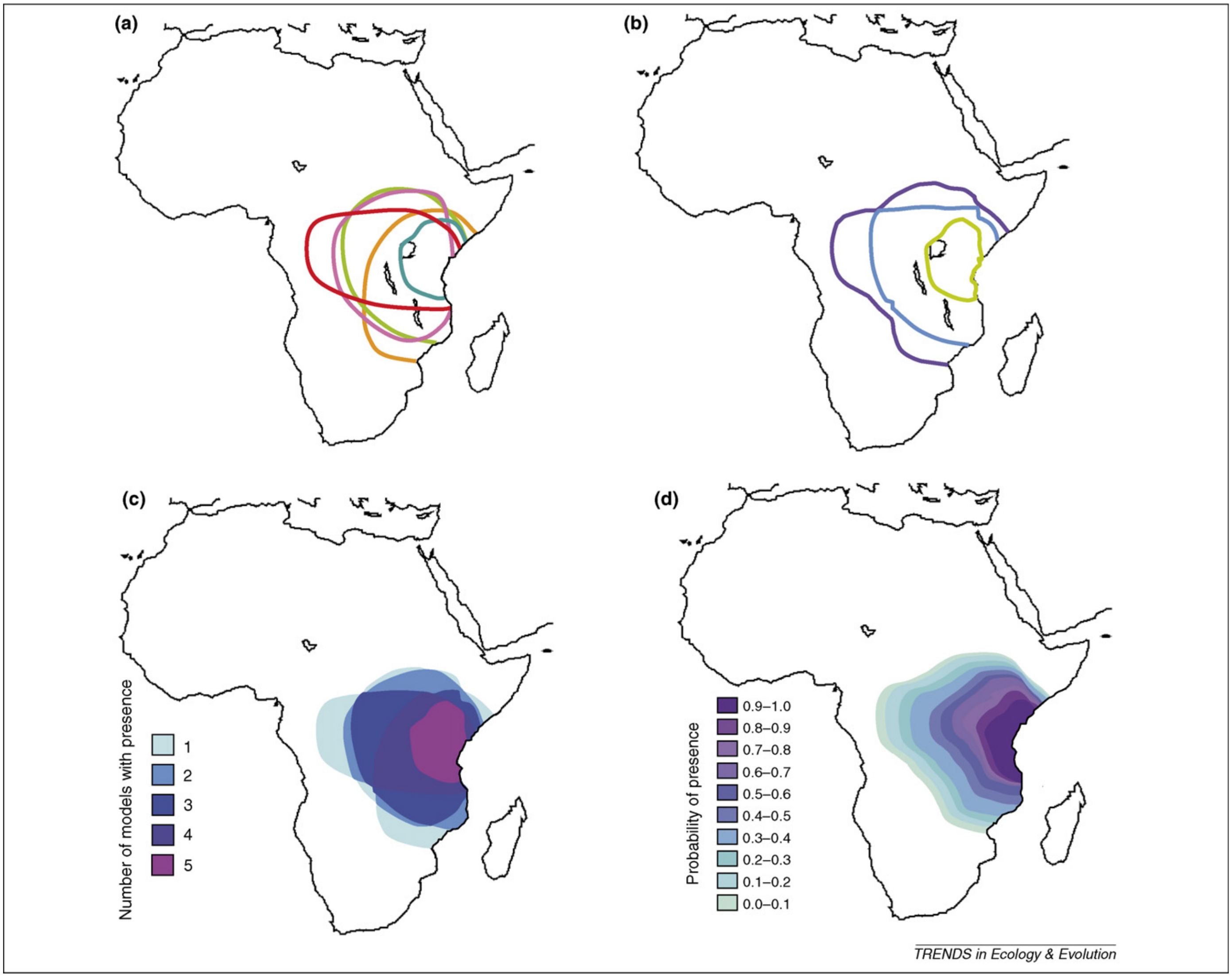

Figure 1. Examples of alternative approaches to analysing ensemble forecasts using artificial data projected onto the map of Africa: **(a)** Individual results from five hypothetical bioclimatic models (shown by coloured lines) predicting the area occupied by a key species under a climate change scenario (no combination of the ensemble forecast is performed); **(b)** a bounding box showing the area where at least one (purple) or all models (green) predict species presence in the future, and a consensus forecast (blue) showing the area where at least half the models (the median) forecast species presence; **(c)** a frequency histogram, showing the number of models (1–5) forecasting the presence of the species at any point; and **(d)** a probability density function showing the likelihood of species presence estimated from a large ensemble.

the frequency distributions requires some form of emulation of the forecast over unsampled combinations, and the resulting PDFs remain conditional [42]. For simpler systems, exhaustive sampling across uncertainties is more feasible, enabling PDFs to be constructed. Yet these will remain conditional on the model or family of models being sampled, and a frequency distribution can only be considered robust when inclusion of additional MPs or MCs does not make much difference to the forecast distribution of a particular event [43]. Whether we can ever know that model stability has been achieved is a moot point, and how to use probabilistic information from large ensemble simulations remains an area of debate [42,44,45].

Ensembles in practice

The idea of combining forecasts is particularly appealing for those who are not convinced that a single model is closest to the truth in all circumstances [11,12] and who sympathise with the view that all models are flawed, but provide useful information [46,47]. Yet ensemble forecasting should not be viewed as an alternative to the more traditional approach of trying to build better models with improved data. Combined forecasts, although emphasizing the 'signal' emerging from the noise associated with different model outputs, remain dependent on individual predictions; better individual forecasts will yield a better combined forecast [17].

Whether to use a synthetically combined forecast (consensus or probabilistic) or bounded forecasts depends in part upon the way in which the forecast will be used. In financial futures, for example, analysts take a long-term view and are prepared for forecasts to be wrong [37]. In seasonal climate forecasting, where users of information might include small-scale farmers, agribusinesses and commodity traders, the utility of consensus or probability forecasts will vary according to the decision maker. Agribusiness and traders can cope with the financial costs of incorrect forecasts in a similar way to financial futures traders. For small farmers, the costs of acting on an incorrect forecast can be so severe that they might prefer a more conservative forecast, or might even ignore forecasts, adopting a cropping strategy that minimises vulnerability to a poor season.

Unlike small-scale farmers, conservation planners need to take a long-term view and rely on forecasts to support conservation decisions [48]. The costs of being wrong can be high. For example, when planning for climate change, the consequences of acting upon a forecast yielding false positives (e.g. species ranges predicted to expand in fact contract) are that resources are not spent on the species most in need of conservation action. Alternatively, acting on the basis of false-negative information (e.g. species ranges predicted to contract in fact expand) might lead to an investment of resources in species that are not threatened by climate change. In both cases, investments might be directed away from the most vulnerable species, leading to potential increases in their risks of extinctions.

The crucial issue is whether the benefits of using a set of combined forecasts in decision-making outweigh, on average, the costs. In reserve selection, it has been shown that acting upon ignorance and opportunism can be more expensive than acting with the support of data and models [49]. Recent analyses of species range shifts under climate change have shown that the costs of relying on a single forecast severely compromise their usefulness [14–16], even when excluding the uncertainties arising from different global climate models (GCM) projections (the BCs). However, a recent study demonstrated the success of a simple implementation of consensus forecasting (Box 2; Figure Ie) in reducing both false negative and positive errors in predictions of observed distribution shifts among British breeding birds [17]. False negatives were reduced from an average of 50% error to 0% (lower quartile = 31% versus 0%; upper quartile = 69% versus 0%, respectively) and comparable reductions in false positives were obtained. Surely, running ensembles of models and combining the results using consensus or probabilistic approaches will not always remove the uncertainties, but the likelihood of making conservation decisions based on forecasts that are far from the truth is reduced.

Some writers have criticized the use of any single forecast (combined from several models or single-model), as it can lead to a decision that, although appropriate for the forecast, imposes a rigidity that might have serious negative consequences if the forecast deviates significantly from truth. The use of bounded forecasts offers an approach that, although honest, might be challenging for decision makers; it enables us to say, with some confidence, what will not happen. The onus then lies with the decision maker to develop policy that is as robust as possible to the predictive uncertainty. A potentially useful pathway is the development of hybrid approaches that combine bounding box with consensus or probabilistic forecasting [16].

Conclusions and recommendations

Using ensemble forecasting has clear advantages over single-model forecasts. Different approaches to the analysis of the ensemble data have their own advantages and disadvantages, and their suitability will depend on the questions being asked. But if used appropriately, either individually, in combination, or in hybrid form, these approaches can enable more robust decision making in the face of uncertainty, and have much to offer to conservation planning. There are additional reasons to adopt ensemble forecasting as part of the mainstream practice of species distribution modelling. Climatologists are now producing tens of thousands of simulations of future climates [38–40]. Exploring these data will be necessary to provide comprehensive assessments of the possible impacts of climate change on biodiversity and the ensembles framework will be required to enable such an exploration. The most comprehensive attempts to run ensembles of models of species distributions have spawn a limited number of combinations of MCs and BCs (Box 2; Figure Ie), yielding no more than 40 projections per species. Developments in bioclimate and climate modelling will rapidly force this number to increase to several thousands projections per species. Considering that models need to be fitted for several hundreds or thousands of species, it is clear that we currently have no ability to cope with such large problems.

Interactions with other disciplines, including statistics, will help us to decide upon the most appropriate analytical tools, but a serious limitation still includes the lack of appropriate software to run and combine large ensembles of models. This is essentially the same conclusion reached by Robert T. Clemen in his 1989 review of ensemble forecasting for management and business [21]. If progress is to be made in the field of ensemble forecasting of species distributions, ecologists need to move fast, join efforts and abandon parochialism in software production. Open source platforms, such as that provided by the R project for statistical computing (http://www.r-project.org/), might provide an adequate source of inspiration.

Acknowledgements

M.B.A is funded by the BBVA Foundation BIO-IMPACT project. The authors thank Tim Carter and David Stainforth for discussion; Townsend Peterson and Steve Philips for clarification of the GARP and MAXENT methods; and four anonymous referees for insightful comments on an early version of the article.

References

1 Huntley, B. *et al.* (1995) Modelling present and potential future ranges of some European higher plants using climate response surfaces. *J. Biogeogr.* 22, 967–1001
2 Peterson, A.T. *et al.* (2002) Future projections for Mexican faunas under global climate change scenarios. *Nature* 416, 626–629
3 Thomas, C.D. *et al.* (2004) Extinction risk from climate change. *Nature* 427, 145–148
4 Thuiller, W. *et al.* (2005) Climate change threats plant diversity in Europe. *Proc. Natl. Acad. Sci. U. S. A.* 102, 8245–8250
5 Lehmann, A. *et al.* (2003) GRASP: generalized regression analysis and spatial prediction. *Ecol. Model.* 160, 165–183
6 Walker, P.A. and Cocks, K.D. (1991) HABITAT: a procedure for modelling a disjoint environmental envelope for a plant or animal species. *Global Ecol. Biogeogr. Lett.* 1, 108–118
7 Busby, J.R. (1986) A biogeographical analysis of *Nothofagus cunninghamii* (Hook.) Oerst. in southeastern Australia. *Aust. J. Ecol.* 11, 1–7
8 Carpenter, G. *et al.* (1993) DOMAIN: a flexible modelling procedure for mapping potential distributions of plants and animals. *Biodiv. Conserv.* 2, 667–680

9 Pearson, R.G. *et al.* (2002) SPECIES: a spatial evaluation of climate impact on the envelope of species. *Ecol. Model.* 154, 289–300
10 Stockwell, D.R.B. and Peters, D.P. (1999) The GARP modelling system: problems and solutions to automated spatial prediction. *Int. J. Geogr. Inf. Syst.* 13, 143–158
11 Olden, J.D. and Jackson, D.A. (2002) A comparison of statistical approaches for modelling fish species distributions. *Freshw. Biol.* 47, 1976–1995
12 Segurado, P. and Araújo, M.B. (2004) An evaluation of methods for modelling species distributions. *J. Biogeogr.* 31, 1555–1568
13 Elith, J. *et al.* (2006) Novel methods improve predictions of species' distributions from occurrence data. *Ecography* 29, 129–151
14 Pearson, R.G. *et al.* (2006) Model-based uncertainty in species' range prediction. *J. Biogeogr.* 33, 1704–1711
15 Thuiller, W. *et al.* (2004) Uncertainty in predictions of extinction risk. *Nature* 430, 10.1038/nature02716
16 Araújo, M.B. *et al.* (2006) Climate warming and the decline of amphibians and reptiles in Europe. *J. Biogeogr.* 33, 1712–1728
17 Araújo, M.B. *et al.* (2005) Reducing uncertainty in projections of extinction risk from climate change. *Glob. Ecol. Biogeogr.* 14, 529–538
18 Cramer, W. *et al.* (2001) Global response of terrestrial ecosystem structure and function to CO2 and climate change: results from six dynamic global vegetation models. *Glob. Change Biol.* 7, 357–373
19 Bates, J.M. and Granger, C.W. (1969) The combination of forecasts. *Operat. Res. Quart.* 20, 451–468
20 Laplace, P.S. (1818) *Deuxième supplément à la Théorie Analytique des Probabilités,* Courcier
21 Clemen, R.T. (1989) Combining forecasts: a review and annotated bibliography. *Int. J. Forecast.* 5, 559–583
22 Palm, F.C. and Zellner, A. (1992) To combine or not combine? Issues of combining forecasts. *J. Forecast.* 11, 687–701
23 Cheung, K.K.W. (2001) A review of ensemble forecasting techniques with a focus on tropical cyclone forecasting. *Meteorol. Appl.* 8, 315–332
24 Gregory, A.W. *et al.* (2001) Testing for forecast consensus. *J. Bus. Econ. Stat.* 19, 34–43
25 Makridakis, S. and Winkler, R.L. (1983) Averages of forecasts: some empirical results. *Manage. Sci.* 29, 987–996
26 Miyamoto, M.M. (1985) Consensus cladograms and general classifications. *Cladistics* 1, 186–189
27 Nilsson, J. *et al.* (2000) Consensus predictions of membrane protein topology. *FEBS Lett.* 486, 267–269
28 Sanders, F. (1963) On subjective probability forecasting. *J. Appl. Meteorol.* 2, 191–201
29 Benestad, R.E. (2004) Tentative probabilistic temperature scenarios for northern Europe. *Tellus* 56A, 89–101
30 Thuiller, W. (2004) Patterns and uncertainties of species' range shifts under climate change. *Glob. Change Biol.* 10, 2220–2227
31 Thuiller, W. (2003) BIOMOD: optimising predictions of species distributions and projecting potential future shifts under global change. *Glob. Change Biol.* 9, 1353–1362
32 Araújo, M.B. *et al.* (2005) Validation of species-climate impact models under climate change. *Glob. Change Biol.* 11, 1504–1513
33 Randin, C.F. *et al.* (2006) Are niche-based species distribution models transferable in space? *J. Biogeogr.* 33, 1689–1703
34 Araújo, M.B. *et al.* (2005) Downscaling European species atlas distributions to a finer resolution: implications for conservation planning. *Glob. Ecol. Biogeogr.* 14, 17–30
35 Prasad, A.M. *et al.* (2006) Newer classification and regression tree techniques: bagging and random forests for ecological prediction. *Ecosystems* 9, 181–199
36 Hastie, T. *et al.* (2001) *The Elements of Statistical Learning: Data Mining, Inference and Prediction,* Springer
37 McNees, S.K. (1992) The uses and abuses of "consensus" forecasts. *J. Forecast.* 11, 703–710
38 Stainforth, D.A. *et al.* (2004) climateprediction.net: a global community for research in climate physics. In *Environmental Online Communication* (Scharl, A., ed.), pp. 101–112, Springer
39 Allen, M. (1999) Do-it-yourself climate prediction. *Nature* 401, 642
40 Allen, M.R. and Stainforth, D.A. (2002) Towards objective probabilistic climate forecasting. *Nature* 419, 228
41 Stainforth, D.A. *et al.* (2005) Uncertainty in predictions of the climate response to rising levels of greenhouse gases. *Nature* 433, 403–406
42 Smith, L.A. (2002) What might we learn from climate forecasts? *Proc. Natl. Acad. Sci. U. S. A.* 99, 2487–2492
43 Allen, M. *et al.* (2002) Model error in weather and climate forecasting. In *Predictability of Weather and Climate* (Palmer, T. and Hagedorn, R., eds), pp. 391–427, Cambridge University Press
44 Frame, D.J. *et al.* (2005) Constraining climate forecasts: the role of prior assumptions. *Geophys. Res. Lett.* 32, art. no.-L09702
45 Dessai, S. and Hulme, M. (2004) Does climate adaptation policy need probabilities? *Climate Policy* 4, 107–128
46 Winkler, R.L. (1989) Combining forecasts: a philosophical basis and some current issues. *Int. J. Forecast.* 5, 605–609
47 Box, G.E.P. (1979) Some problems of statistics and everyday life. *JAMA* 74, 1–4
48 Araújo, M.B. *et al.* (2004) Would climate change drive species out of reserves? An assessment of existing reserve selection methods. *Glob. Change Biol.* 10, 1618–1626
49 Pressey, R.L. and Cowling, R.M. (2001) Reserve selection algorithms and the real world. *Conserv. Biol.* 15, 275–277
50 Phillips, S.J. *et al.* (2006) Maximum entropy modeling of species geographic distributions. *Ecol. Model.* 190, 231–259