

Patterns and uncertainties of species' range shifts under climate change

WILFRIED THUILLER

CEFE-CNRS, 1919 route de Mende, 34293 Montpellier cedex 5, France

Abstract

Although bioclimatic modelling is often used to estimate potential impacts of likely climate changes, little has been done to assess the reliability and variability of projections. Here, using four niche-based models, two methods to derive probability values from models into presence–absence data and five climate change scenarios, I project the future potential habitats of 1350 European plant species for 2050. All 40 different projections of species turnover across Europe suggested high potential species turnover (up to 70%) in response to climate change. However variability in the potential distributional changes of species across climate scenarios was obscured by a strong variability in projections arising from alternative, yet equally justifiable, niche-based models. Therefore, projections of future species distributions and derived community descriptors cannot be reliably discussed unless model uncertainty is quantified explicitly. I propose and test an alternative way to account for modelling variability when deriving estimates of species turnover (with and without dispersal) according to a range of climate change scenarios representing various socio-economic futures.

Keywords: BIOMOD, climate change scenarios, consensus analysis, niche-based models, plant diversity, species turnover, uncertainty

Received 9 March 2004; received in revised form 14 June 2004 and accepted 22 June 2004

Introduction

Significant effects of global climate change are already observed for a variety of ecosystems, especially with evidence for shifts in species ranges (Parmesan & Yohe, 2003; Root *et al.*, 2003). Because different species respond idiosyncratically to various ecological forces, ecological communities may disassemble as individual species shift their ranges in different directions. Recent insights into consequences of climate change on biodiversity have been gained from niche-based models that describe the potential responses of individual species to climate change. These models relate present-day distributions to current climate, and then project spatial shifts in species' climatic envelopes according to climate change scenarios (Bakkenes *et al.*, 2002; Peterson *et al.*, 2002; Thuiller, 2003; Thomas *et al.*, 2004). The percentage of species turnover (defined as an index of dissimilarity between the current and future

species composition within a given area) and the percentage of species that could persist in, disappear from, and colonize that area are often considered as a good measure of the degree of ecosystem perturbation, and have been used to assess the potential impact of climate change at regional to continental scales (Peterson *et al.*, 2002). Using climate change scenarios and niche-based models that project future suitable habitat from current distributions, several studies have suggested that species turnover may be very high in some regions, potentially resulting in modifications of community structure strong enough to lead to ecosystem disruption (Bakkenes *et al.*, 2002; Erasmus *et al.*, 2002; Peterson *et al.*, 2002).

However, to be able to highlight precisely where strong modifications of species diversity are expected to occur, projections of species turnover must be reliable and robust. Although niche-based models have been extensively assessed in their effectiveness to predict present-day distributions of various taxa from regional to continental scales (Moisen & Frescino, 2002; Thuiller *et al.*, 2003; Brotons *et al.*, 2004), their accuracy in predicting future conditions has not been evaluated yet. Because of the obvious fact that future realized

Correspondence: present address: W. Thuiller, Climate Change Research Group, Kirstenbosh Research Center, National Botanical Institute, P/Bag x7, Claremont 7735, Cape Town, South Africa, e-mail: wilfried.thuiller@cefe.cnrs-mop.fr or Thuiller@nbi.ac.za

distributions are unknown, reliability of niche-based models for projecting species distributions for future hypothetical conditions should not be taken for granted (Thuiller *et al.*, 2004). Recently, Thuiller (2003) showed that a range of niche-based models with similar accuracy for predicting species' present-day distributions could predict alarmingly different future distributions. Such results have critical implications when results for individual species are aggregated to derive turnover or extinction rates as used by Bakkenes *et al.* (2002), Peterson *et al.* (2002) or Thomas *et al.* (2004). Indeed, if projections from different models differ significantly for each individual species, then uncertainties will be accumulated when deriving species turnover by accumulating projections for large numbers of species. Hence, there is an urgent need to assess the impact of variability across modelling methods used to derive conclusions about the impacts of global change on biodiversity. This need increases when projections into the future are carried out according to a range of climate change scenarios as provided by the last report of the Intergovernmental Panel on Climate Change (IPCC) (Nakicenovic & Swart, 2000).

Here, using five climate change scenarios and four different niche-based models, I project the future climatic envelopes of 1350 European plant species for the middle of this century. Combinations of individual species envelopes for a given method and scenario were used to project the associated species turnover and percentage of stable/gained habitats. I then analysed the relative contributions of modelling methods and climate change scenarios to variability across these projections. Based on these results, I propose a method to incorporate methodological uncertainty into the modelling process and derive robust estimates of species turnover across a range of climate scenarios.

Methods

The general approach to modelling climate change effects on species distributions used herein is developed in detail elsewhere (Thuiller, 2003).

Plant species data sets

I used a subset (20%) of the Atlas Florae Europaeae (Lahti & Lampinen, 1999), which contains digitized maps for more than 2600 species across Europe. Species with less than 20 records in the data set were excluded from the analysis, so as to reduce errors associated with excessively small sample sizes (Stockwell & Peterson, 2002). Retained species included all European pteridophytes and a sample of spermatophytes comprising of all gymnospermae families (Coniferales, Taxales and

Gnetales), a fraction of angiospermae dycotyledones (Salicales, Myricales, Juglandales, Fagales, Urticales, Proteales, Santales, Aristolochiales, Balanophorales, Polygonales, Centrospermae and Ranales), but no monocotyledones. Species data were located in 4419 UTM (Universal Transverse Mercator) $50 \times 50 \text{ km}^2$ grid cells. I used only 2089 grid cells excluding most of the eastern European countries (except for the Baltic States) because of low recording efforts in these areas (Williams & Araújo, 2000).

Bioclimatic data sets

I used seven climatically derived variables considered critical to plant physiological function and survival (Bartlein *et al.*, 1986; Woodward, 1987). Data were obtained from the Climatic Research Unit (<http://www.cru.uea.ac.uk/>) for: mean annual, winter, and summer precipitation, mean annual temperature and minimum temperature of the coldest month, growing degree days ($\text{GDD} > 5^\circ$) and an index of humidity (AET/PET: mean ratio of annual actual over annual potential evapotranspiration). Mean values were averages for the period of 1961–1990. These data sets are supplied on a $10'$ grid, covering Europe. Then data were aggregated by averaging to $50 \times 50 \text{ km}^2$ UTM in order to match the resolution of species data.

Climate change scenarios

Future projections for the 2050 time slice (averages for 2020–2050) were derived using two general circulation models (GCMs) experiment (Carson, 1999) made available in the context of the EC-funded ATEAM project and conducted, respectively, at the UK Hadley Centre for Climate Prediction and Research (HadCM3) and Australia's Commonwealth Scientific and Industrial Research Organization (CSIRO2). These GCMs included different storylines according to the IPCC's (Nakicenovic & Swart, 2000). The A1 storyline describes a future world of very rapid economic growth, global population that peaks by mid-century and then declines, and the rapid introduction of new and more efficient technologies. The A2 storyline describes a very heterogeneous world. The underlying theme is self-reliance and preservation of local identities. Economic development is primarily regionally oriented and *per capita* economic growth and technological changes are more fragmented and slower than in the other storylines. The B1 storyline describes a convergent world with the same global population as the A1 storyline that peaks by mid-century and declines thereafter, but with rapid change in economic structures towards a service and information economy, with the introduction of

Table 1 Variation of mean annual temperature and precipitation over Europe across the different GCMs and storylines used in the study between 2050 and 1990

	HadCM3				A2 CSIRO2
	A1	A2	B1	B2	
Annual temperature (°C)	2.16	1.90	1.89	1.95	1.76
Annual precipitation (mm)	-10.1	-7.99	-10.6	-5.5	10.5

GCM, general circulation model; HadCM3, Hadley Centre for Climate Prediction and Research; CSIRO2, Commonwealth Scientific and Industrial Research Organization.

clean and resource-efficient technologies. The B2 storyline describes a world in which the emphasis is on local solutions to economic, social, and environment sustainability. It is a world with continuously increasing global population at a rate lower than A2, intermediate levels of economic development, and less rapid and more diverse technological change than in the B1 and A1 storylines (Nakicenovic & Swart, 2000). I used the four storylines under HadCM3 GCM experiments and only A2 for CSIRO2 because of data availability (see Table 1 for more details on expected mean temperature increase and variation of precipitation between 1990 and 2050).

Niche-based models

Current habitats at $50 \times 50 \text{ km}^2$ and $10'$ resolution, and future potential habitats at $10'$ resolution were projected for each plant species using the BIOMOD framework (Thuiller, 2003). For each species, generalized linear models (GLMs), generalized additive models (GAMs), classification tree analysis (CTA), and artificial neural networks (ANNs) were calibrated on a random sample of the initial data (70%) and tested on the remaining data sets with both the receiver-operating characteristic (ROC) curve and the κ statistic (Pearce & Ferrier, 2000; Thuiller *et al.*, 2003).

Projected distributions of presence-absence were derived from probability values generated by niche-based models using two methodological approaches (cut-off methods hereafter): first, by maximizing the κ statistic, and second by maximizing the percentage of presence and absence correctly predicted for present conditions (Manel *et al.*, 2001). The last one is done by estimating the number of presence and absence correctly predicted for a range of cut-off values and by selecting the cut off that maximizes both presence and absence jointly (Pearce & Ferrier, 2000, Thuiller *et al.*, 2003).

Each species was predicted to lose or gain suitable habitat according to each combination of modelling and cut-off methods on one hand, and climate change

scenarios on the other. To evaluate geographic patterns of numbers of species potentially lost or gained across Europe, I summed the number of species gained (G) or lost (L) by pixel for each model. I summed simulated present-day distributions across all species to determine the simulated species richness (SR) and estimated the percentage of species turnover as: $100 \times (G + L) / (SR + G)$. As this process implied unlikely universal dispersal for species to track suitable bioclimatic conditions, I also derived species turnover without dispersal as: $100 \times (L / SR)$.

Quantifying uncertainties

The three factors tested here (4 niche-based models \times 2 cut-off methods \times 5 climate change scenarios) resulted in 40 projections of species turnover. To assess and account for the differences in sources of variability among projections, I used a consensus analysis (Westerhuis *et al.*, 1998). This is based on a multivariate approach in order to search among a set of variables for the one that summarizes the highest amount of information. I used a principal component analysis (PCA consensus) on the different combinations of modelling methods and scenarios to run a consensus analysis on projections of species turnover. The first axis (called consensus axis hereafter) of PCA captured consistent spatial patterns in turnover rate across the different projections. If the 40 projections were exactly similar (i.e. only one axis would be retained explaining 100% of the variation), there would be no variability across projections. If the 40 projections were randomly simulated and completely unrelated to each other, 40 axes would be retained, each of them explaining $1/40 = 2.5\%$ of the variation. Based on the basic expectation that at least climate change scenarios would differ, the variability explained by the first axis should be comprised between 2.5% and 100%. If the consensus axis explains less than 100% of the variability, then the subsequent axes (termed variability axes hereafter) would explain the orthogonal variation because of variability across models, cut-off methods or/and

scenarios of climate change. I analysed these variability axes using CTA (Breiman *et al.*, 1984) to measure the proportion of variability explained by each of factor. I used CTA because this method is not perturbed by interrelated components and make it possible to separate variation within each factor (for instance between scenario A1 and scenario A2).

Results

Quantifying uncertainties

The first five axes of the consensus PCA explained 90% of the total variability across projections (Table 2). The consensus axis accounted for 56.1% of variability (i.e. the 40 projections were 56.1% similar). Hence, the variability across the three experimental factors was 43.9% and described by the variability axes of the PCA.

The position of each projection and associated sources of variation (models \times cut-off method \times scenarios) along each of the four variability axes was extracted and used to run CTA (Fig. 1). The second, third, and fourth orthogonal axes were essentially explained by the model and cut-off method used, while the fifth axis, which accounted only for 4% of the variability, was mainly driven by variability across scenarios. The second axis (Fig. 1a), which explained the greatest part of variability across turnover projections (18.3%), was mainly driven by differences between rule-based and generalized regression methods. The projection scores on the second PCA axis, plotted on the European map, showed that this variability across methods was spatially explicit. High scores (darker colours) showed high species turnover predicted by generalized models (GAM and GLM) and low scores (lighter colours) showed high species turnover predicted by rule-based models (ANN and CTA) but not by the generalized models. The third and fourth axes, which explained noticeably less variability, were also mainly determined by variability across models and then by cut-off methods.

Variability across climate change scenarios was discernable only for the fifth axis separating broadly the GCM (CSIRO2 vs. HadCM3) and then the storylines

(global vs. regional economy). The PCA fifth axis scores, plotted on the European map, displayed higher species turnover predicted under HadCM3 than under CSIRO2.

Integrating modelling uncertainties

To decrease the uncertainty among projection methods, an alternative approach is followed here. For each species and for each selected climate change scenario, a PCA consensus was run on the coupled model \times cut-off method (eight projections by species). The projections that were the most correlated with the PCA consensus axis were selected. This represents the projections that best captures shared information among the whole set of projections for a given climate change scenario. This approach was carried out for the 1350 plants species to derive one set of projections for the four climate change scenarios (Table 3). Species turnover under universal and no dispersal hypothesis were then derived for each climate change scenario and compared (Figs 2 and 3)

Percentage of species turnover in each pixel under universal and no dispersal hypothesis showed contrasts across the range of climate change scenarios. The A1 storyline that represents economic and global world appeared the most critical for species distributions in many regions, while the other storylines showed similar patterns. For instance, the Pannonian region was the biogeographical region most affected by climate change in terms of species turnover under all storylines, whether no or universal species dispersal ability was assumed. Across all GCMs and storylines, rates of species turnover were noticeably lower for western and northern biogeographic regions than for southern and eastern regions.

While the regions where high rate of species turnover were predicted consistently across the five climate change scenarios, turnover values under universal and no dispersal hypothesis differed between biogeographical regions. Under the more likely 'no dispersal' hypothesis, Scandinavia should be the least affected by climate change with a very low rate of species turnover, except for extreme northern part. Conversely, under universal dispersal, this region could be prone to high

Table 2 Summary of the PCA consensus carried out on the 40 projections

	Axis 1	Axis 2	Axis 3	Axis 4	Axis 5
Variance	4.747	2.704	1.637	1.424	1.241
Proportion of variance	0.561	0.183	0.067	0.057	0.039
Cumulative proportion \times 100	56.1	74.4	81.1	86.2	90.0

PCA, principal component analysis.

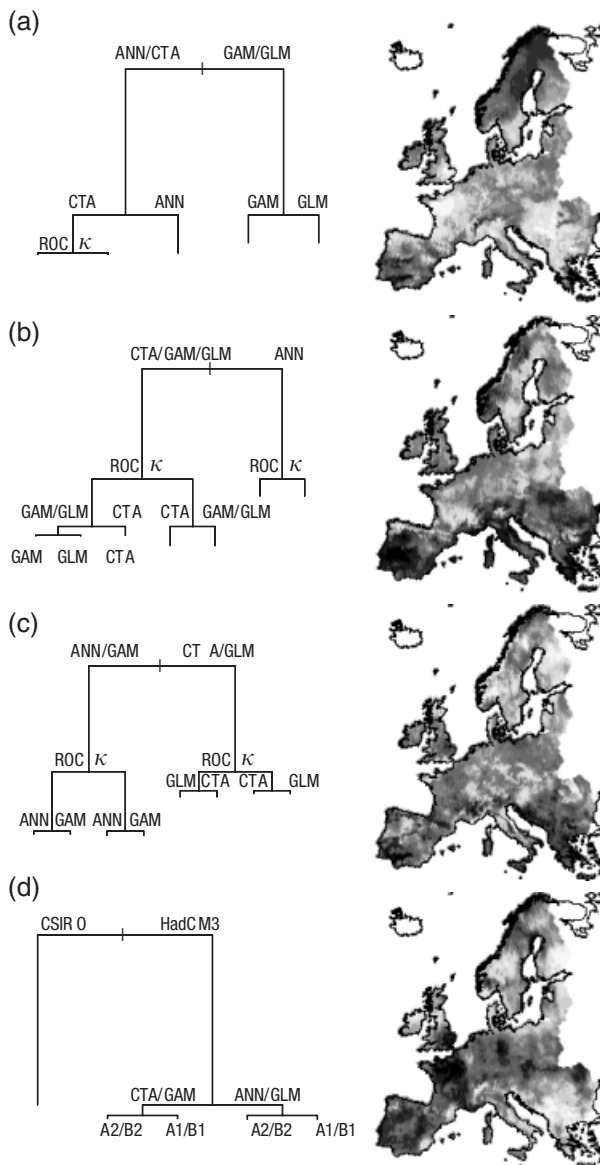


Fig. 1 Details of the classification trees carried out on the position of each triplet (niche-based model \times cut-off method \times scenarios) on principal component analysis axes 2(a), 3(b), 4(c) and 5(d) and the mapped scores of each axes. The trees showed the factors that explained the proportion of explained deviance.

species turnover with high rates of colonization per pixel. At the same time, the Atlantic region and Central Europe (e.g. Germany) were the least affected by climate change under almost all scenarios.

Discussion

Modelling uncertainty

Limitations of niche-based modelling methods are not discussed here, as recent papers have provided

insightful comments, descriptions and critiques (Peterson *et al.*, 2002; Pearson & Dawson, 2003; Thuiller, 2003; Thuiller *et al.*, 2004). In this present study, I analysed the impact of variability across niche-based models and the impact of cut-off method used to derive probability values from niche-based models into presence–absence form. Although cut-off methods are largely understudied, their impact on the quality of projections appeared crucial as both methods (maximizing κ statistic and maximizing % of presence and absence correctly predicted) could lead to very different projected distributions, a result that could be problematic for biodiversity risk assessment and conservation purpose (Manel *et al.*, 2001).

Here, I demonstrated that even when using common niche-based modelling and cut-off methods, the variability across projections was large and may even hide the variability of using a range of climate change scenarios as recommended by the IPCC (Nakicenovic & Swart, 2000). My aim here is not to deny the suitability of niche-based models to derive broad estimates of sensitivity of biodiversity to climate change, but to argue that care must be taken before deriving any conclusion about future impacts on species diversity. Recently, several authors have published interesting results on species sensitivity to climate change and derived geographic values of species extirpation, colonization or turnover (Bakkenes *et al.*, 2002; Erasmus *et al.*, 2002; Peterson *et al.*, 2002; Thomas *et al.*, 2004). Using GLM, and the same models as Bakkenes *et al.* (2002), for the same flora except Russia, I estimated that on average up to 41% of the plant species present in a pixel in 1990 would disappear by 2050, which was slightly higher than the 32% they obtained, but much higher than the 29% I estimated using artificial neural networks. This example shows the strong variability of species turnover estimates from different niche-based models applied on the same data. In addition, differences in estimated rates of species extinction or colonization across niche-based models appeared to be spatially structured, showing that different models have different abilities to deal with particular species or geographic locations with specific environmental conditions.

Since there is no way currently to assess which universal niche-based model is most appropriate (Franklin, 1998; Thuiller *et al.*, 2003), one could argue the use for each species of the model with the highest predictive accuracy for present-day conditions, and project future distributions according to a range of climate change scenarios. However, there are at least two problems: First, there are several methods to assess predictive accuracy (e.g. ROC curve or κ statistic) and they do often not provide the same estimates as they are

Table 3 Percentage of best combination model \times cut-off selected by the consensus analysis to project future species distribution across the 1350 species

	HadCM3				A2 CSIRO2
	A1	A2	B1	B2	
<i>GLM</i>					
ROC	15.3	15.1	15.4	15.3	15.9
κ	9.30	8.10	8.70	8.90	10.5
<i>GAM</i>					
ROC	38.6	40.0	40.0	38.9	40.1
κ	22.0	23.0	21.5	21.7	21.7
<i>CTA</i>					
ROC	5.30	5.10	5.00	4.80	4.40
κ	1.50	1.50	1.20	1.60	1.10
<i>ANN</i>					
ROC	4.80	4.10	5.00	5.00	3.40
κ	3.20	3.10	3.20	3.80	2.90

HadCM3, Hadley Centre for Climate Prediction and Research; CSIRO2, Commonwealth Scientific and Industrial Research Organization; GLM, generalized linear model; GAM, generalized additive model; CTA, classification tree analysis; ANN, artificial neural network; ROC, receiver–operating characteristic.

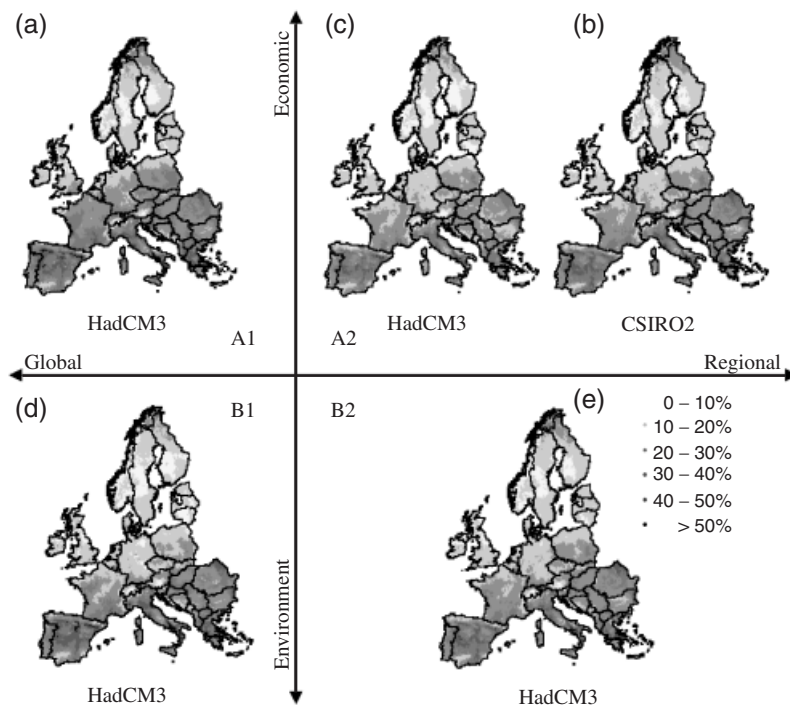


Fig. 2 Rate of plant species turnover in Europe estimated using the consensus approach and assuming no species dispersal across a range of climate change scenarios. (a) A1 HadCM3; (b) A2 CSIRO2; (c) A2 HadCM3; (d) B1 HadCM3; (e) B2 HadCM3.

based on different algorithms and assumptions (Manel *et al.*, 2001). Second, it is not certain that the model with the highest predictive accuracy for present-day conditions will be the best to estimate future distributions (Thuiller, 2003). Therefore, I suggest here that it is most appropriate to choose a projection (one niche-based

model \times one cut-off combination) into the future that conservatively summarizes agreements among projections generated by different modelling techniques. The projection most correlated with the first PCA axis represents such a consensus by extracting shared variations among future projections. It does not mean

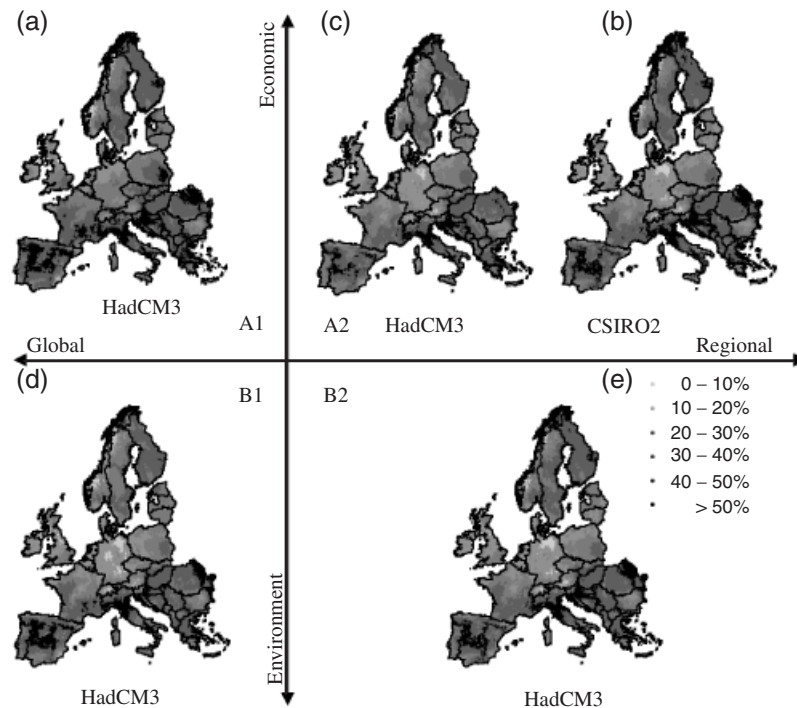


Fig. 3 Rate of plant species turnover in Europe estimated using the consensus approach and assuming universal dispersal across a range of climate change scenarios. (a) A1 HadCM3; (b) A2 CSIRO2; (c) A2 HadCM3; (d) B1 HadCM3; (e) B2 HadCM3.

that it will necessarily be the most accurate projection into the future but it is at least the most consensual and therefore conservative.

Integrating modelling uncertainties

Across the emission scenarios describing different societal responses to climate change (Nakicenovic & Swart, 2000), the A1 storyline appeared as the most threatening plant species diversity and enhanced species turnover across some specific regions such as southeast France or northern Italia. The A2 appeared less harmful for plant diversity than B1 and B2 representing a more environmentally minded world. Indeed, climate is changing continuously and incidence of present pollution will be effective at least one or two decades later. That means that in average from 2020 to 2050, there are lower CO₂ and NO₂ emissions under A2 than B1 and B2 storylines but that the trend is reversed from 2050 to 2080 (Nakicenovic & Swart, 2000). That trend explains why A2 could be considered as safer for species diversity than B1 and B2 scenarios, although it is considered as one of the most negative scenarios for global equilibrium (Nakicenovic & Swart, 2000).

It is often stressed that under climate change, northward and upslope migration of species is expected (Parmesan & Yohe, 2003) but little attention has been given to which regions would be the most affected by

high species turnover and community changes. Under the presented simulations it appears that Pannonian and Continental biogeographical regions could undergo dramatic rate of species turnover (under universal and no dispersal hypotheses) that could strongly modify existing ecological communities. However, results highlight that species dispersal could strongly moderate future species turnover. The simulations presented here show the possible range between no dispersal and universal dispersal ability. For instance, the Boreo-Alpine region could be the least affected by climate change assuming no dispersal, but one of the most affected assuming universal dispersal. Such a phenomenon is expected in a region that could undergo massive immigrations from neighbouring regions as Boreo-Alpine or Pannonian regions.

To conclude, the methodological uncertainty associated to niche-based models is problematic for forecasting impact of global climate change on biodiversity. This uncertainty hides the effects of using different GCMs and different emissions scenarios as recommended by IPCC (2001). There is an inherent uncertainty for forecasting anthropogenic climate change and so predictions of current and future species distributions should be exempt of uncertainty in order to produce relevant patterns of extinctions, colonizations and species turnover. In this paper, I propose an alternative method based on consensus among

modelling methods to incorporate species modelling uncertainty in order to produce more reliable estimates of biodiversity risk under global climate change.

Acknowledgements

This research was funded by the European Commission's ATEAM (Advanced Terrestrial Ecosystem Analysis and Modelling) project (EVK2-CT-2000-00075). Species data were kindly supplied by M. B. Araújo. The author thanks particularly T. D. Mitchell of the Tyndall Centre for Climate Change Research for providing climate and scenario data and S. Zaehle and M. Erhard of Potsdam-Institute for Climate Impact Research for aggregating climate data. The author also thanks Lluís Brotons, Sandra Lavorel, Miguel B. Araújo and Guy Midgley for their valuable comments on the manuscript.

References

- Bakkenes M, Alkemade RM, Ihle F *et al.* (2002) Assessing effects of forecasted climate change on the diversity and distribution of European higher plants for 2050. *Global Change Biology*, **8**, 390–407.
- Bartlein PJ, Prentice IC, Webb T (1986) Climatic response surfaces from pollen data for some eastern North American taxa. *Journal of Biogeography*, **13**, 35–57.
- Breiman L, Friedman JH, Olshen RA *et al.* (1984) *Classification and Regression Trees*. Chapman & Hall, New York.
- Brotons L, Thuiller W, Araújo MB *et al.* (2004) Presence-absence vs presence-only modelling methods for predicting bird habitat suitability. *Ecography*, **27**, 165–172.
- Carson DJ (1999) Climate Modelling: achievements and prospects. *Quarterly Journal of the Royal Meteorological Society*, **125** 1, 27.
- Erasmus BFN, Van Jaarsveld AS, Chown SL *et al.* (2002) Vulnerability of South African animal taxa to climate change. *Global Change Biology*, **8**, 679–693.
- Franklin J (1998) Predicting the distribution of shrub species in southern California from climate and terrain-derived variables. *Journal of Vegetation Science*, **9**, 733–748.
- IPCC (2001) In: *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change* (eds Houghton JT, Ding Y, Griggs DJ *et al.*), Cambridge University Press, Cambridge.
- Lahti T, Lampinen R (1999) From dot maps to bitmaps – Atlas Florae Europaeae goes digital. *Acta Botanica Fennica*, **162**, 5–9.
- Manel S, Williams HC, Ormerod SJ (2001) Evaluating presence-absence models in ecology: the need to account for prevalence. *Journal of Applied Ecology*, **38**, 921–931.
- Moisen GG, Frescino TS (2002) Comparing five modelling techniques for predicting forest characteristics. *Ecological Modelling*, **157**, 209–225.
- Nakicenovic N, Swart R (eds). (2000) *Emissions Scenarios: A Special Report of Working Group III of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge.
- Parmesan C, Yohe G (2003) A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, **421**, 37–42.
- Pearce J, Ferrier S (2000) Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling*, **133**, 225–245.
- Pearson RG, Dawson TE (2003) Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology and Biogeography*, **12**, 361–372.
- Peterson AT, Ortega-Huerta MA, Bartley J *et al.* (2002) Future projections for Mexican faunas under global climate change scenarios. *Nature*, **416**, 626–629.
- Root TL, Price JT, Hall KR *et al.* (2003) Fingerprints of global warming on wild animals and plants. *Nature*, **421**, 57–60.
- Stockwell DRB, Peterson AT (2002) Effects of sample size on accuracy of species distribution models. *Ecological Modelling*, **148**, 1–13.
- Thomas CD, Cameron A, Green RE *et al.* (2004) Extinction risk from climate change. *Nature*, **427**, 145–148.
- Thuiller W (2003) BIOMOD: optimising predictions of species distributions and projecting potential future shifts under global change. *Global Change Biology*, **9**, 1353–1362.
- Thuiller W, Araújo MB, Lavorel S (2003) Generalized models versus classification tree analysis: a comparative study for predicting spatial distributions of plant species at different scales. *Journal of Vegetation Science*, **14**, 669–680.
- Thuiller W, Brotons L, Araújo MB *et al.* (2004) Effects of restricting environmental range of data to project current and future species distributions. *Ecography*, **27**, 165–172.
- Westerhuis JA, Kourti T, MaxGregor JF (1998) Analysis of multiblock and hierarchical pca and pls methods. *Journal of Chemometrics*, **12**, 463–482.
- Williams PH, Araújo MB (2000) Using probability of persistence to identify important areas for biodiversity conservation. *Proceedings of the Royal Society of London B*, **267**, 1959–1966.
- Woodward FI (1987) *Climate and Plant Distribution*. Cambridge University Press, Cambridge.